Du côté des mathématiciens

12

Il n'est pas nécessaire d'être grand mathématicien pour utiliser les fonctions rangées dans la catégorie Maths et trigonométrie*. D'un abord un peu austère, cette liste recèle pourtant LA fonction magique d'Excel. Nous avons bien entendu nommé la fonction SOMME qui, à elle seule, justifie l'utilisation d'Excel pour de nombreux utilisateurs.*

SOMMAIRE

- Arrondis
- Décompositions en facteurs premiers
- Sommes, produits
- Exponentielles, logarithmes
- Matrices
- Probabilités
- Fonctions circulaires
- Fonctions hyperboliques

MOTS-CLÉS

- Aléa
- Angle
- Arrangement
- Arrondi
- Combinaison
- Cosinus
- Déterminant
- Exponentielle
- Factorielle
- Logarithme
- Matrice
- Nombre premier
- Produit
- Quotient
- Racine
- Signe
- Sinus
- Somme
- Tangente

La catégorie *Maths et trigonométrie* regroupe à la fois des outils d'usage courant, comme les fonctions d'arrondi, et d'autres plus sophistiqués, comme les fonctions de calcul matriciel ou logarithmique. Toutefois, quelle que soit la complexité des mécanismes qui les sous-tendent, ne négligez pas ces fonctions, car bien souvent, elles permettent de répondre simplement à des problèmes épineux. De plus, il n'est pas nécessaire de comprendre les tenants et les aboutissants des formules qu'elles traduisent pour bien les utiliser.

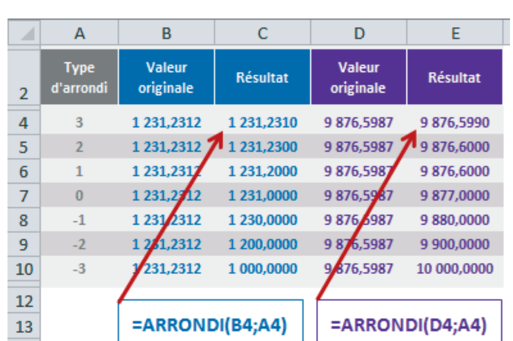
Soixante-quatorze fonctions Maths et trigonométrie

Cette catégorie rassemble plusieurs familles. Vous avez d'abord les fonctions d'arrondi et quelques fonctions d'arithmétique élémentaire. Ensuite, vous trouvez les fonctions hyperboliques, logarithmiques, matricielles ainsi que quelques fonctions de calcul de probabilités. Enfin, vous disposez de toute une série de fonctions circulaires (trigonométrie).

Fonctions d'arrondi

Une façon d'arrondir une valeur consiste à appliquer à la cellule un format de nombre ; mais dans ce cas, vous ne jouez que sur l'apparence du nombre (la valeur elle-même est conservée avec toute sa précision). En revanche, les fonctions d'arrondi transforment profondément les valeurs auxquelles elles s'appliquent et leur font réellement perdre leur précision. Les douze fonctions présentées ici appliquent toutes un arrondi, mais en respectant à chaque fois des règles différentes.

Le second argument des quatre fonctions *ARRONDI*, *ARRONDI.INF*, *ARRONDI.SUP* et *TRONQUE* donne des niveaux d'arrondis différents. Il faut utiliser une valeur entière positive pour arrondir au-delà de la virgule, ou négative pour arrondir en-deçà. Les deux séries d'exemples entrés dans les plages *C4:C10* et *E4:E10* de la figure 12-1 exposent les différentes réactions de la fonction *ARRONDI* suivant que la valeur à arrondir est supérieure ou strictement inférieure à 5.



	A	B	C	D	E
2	Type d'arrondi	Valeur originale	Résultat	Valeur originale	Résultat
4	3	1 231,2312	1 231,2310	9 876,5987	9 876,5990
5	2	1 231,2312	1 231,2300	9 876,5987	9 876,6000
6	1	1 231,2312	1 231,2000	9 876,5987	9 876,6000
7	0	1 231,2312	1 231,0000	9 876,5987	9 877,0000
8	-1	1 231,2312	1 230,0000	9 876,5987	9 880,0000
9	-2	1 231,2312	1 200,0000	9 876,5987	9 900,0000
10	-3	1 231,2312	1 000,0000	9 876,5987	10 000,0000
12					
13		=ARRONDI(B4;A4)		=ARRONDI(D4;A4)	
14					

Figure 12–1
Des fonctions ARRONDI ont été entrées dans les plages C4:C10 et E4:E10. Elles utilisent pour deuxième argument les valeurs entières de la plage A4:A10.

Tableau 12–1 Fonctions d'arrondi

Fonction	Description
ARRONDI	Cette fonction utilise deux arguments. Elle arrondit un nombre décimal (premier argument) au niveau de précision indiqué dans le deuxième argument (un entier). Elle applique les règles d'arrondi standard, c'est-à-dire le choix de la valeur inférieure jusqu'à 5 exclu et de la valeur supérieure à partir de 5. Si le deuxième argument n'est pas un entier, Excel le tronque à sa valeur entière.
ARRONDI.SUP	Cette fonction utilise deux arguments. Elle arrondit un nombre décimal (premier argument) au niveau de précision indiqué dans le deuxième argument (un entier). Elle n'applique pas les règles d'arrondi standard, mais choisit systématiquement la valeur supérieure. Si le deuxième argument n'est pas un entier, Excel le tronque à sa valeur entière.

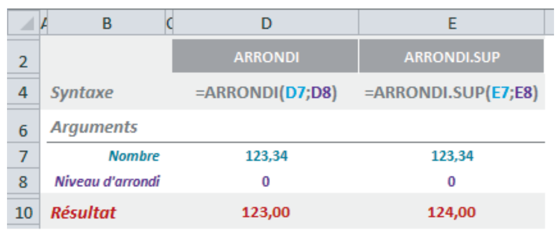


Figure 12–2
Mise en œuvre des fonctions ARRONDI et ARRONDI.SUP.

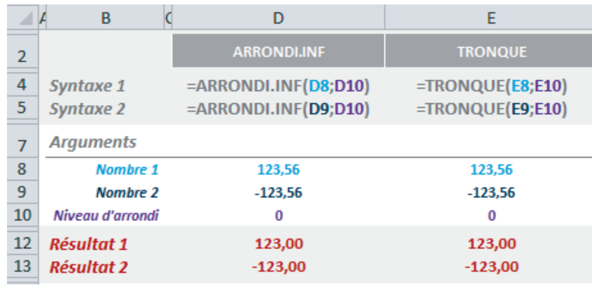
ATTENTION **Comportement inversé sur les valeurs négatives**

Les fonctions *ARRONDI.INF* et *ARRONDI.SUP* raisonnent en valeur absolue quand on aborde les valeurs négatives. `=ARRONDI.INF(-25;-1)` renvoie `-20` alors que `=ARRONDI.SUP(-25;-1)` renvoie `-30`.

Tableau 12–2 Fonctions d'arrondi

Fonction	Description
ARRONDI.INF	Cette fonction utilise deux arguments. Elle arrondit un nombre décimal (premier argument) au niveau de précision indiqué dans le deuxième argument (un entier). Elle n'applique pas les règles d'arrondi standard, mais choisit systématiquement la valeur inférieure. Si le deuxième argument n'est pas un entier, Excel le tronque à sa valeur entière.
TRONQUE	Cette fonction utilise deux arguments. Elle supprime tous les chiffres composant un nombre décimal (premier argument) jusqu'au niveau de précision indiqué dans le deuxième argument (un entier). De ce fait, elle renvoie des résultats similaires à la fonction *ARRONDI.INF*. Si le deuxième argument n'est pas un entier, Excel le tronque à sa valeur entière.

Figure 12–3
Mise en œuvre des fonctions TRONQUE et ARRONDI.INF.



	B	D	E
2		ARRONDI.INF	TRONQUE
4	*Syntaxe 1*	=ARRONDI.INF(D8;D10)	=TRONQUE(E8;E10)
5	*Syntaxe 2*	=ARRONDI.INF(D9;D10)	=TRONQUE(E9;E10)
7	*Arguments*		
8	*Nombre 1*	123,56	123,56
9	*Nombre 2*	-123,56	-123,56
10	*Niveau d'arrondi*	0	0
12	*Résultat 1*	123,00	123,00
13	*Résultat 2*	-123,00	-123,00

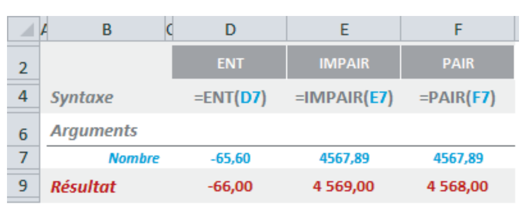
HISTOIRE Pourquoi deux fonctions pour un même objectif ?

La fonction *TRONQUE* existe depuis les premières versions d'Excel, alors que la fonction *ARRONDI.INF* est apparue plus tard. La fonction *TRONQUE* est donc conservée dans Excel 2010 et Excel 2013 pour assurer une continuité avec les versions précédentes.

Tableau 12–3 Fonctions d'arrondi

Fonction	Description
ENT	Cette fonction utilise un seul argument (un nombre décimal), dont elle renvoie la partie entière.
IMPAIR	Cette fonction utilise un seul argument (un nombre décimal), dont elle renvoie la valeur entière impaire la plus proche en s'éloignant de zéro.
PAIR	Cette fonction utilise un seul argument (un nombre décimal), dont elle renvoie la valeur entière paire la plus proche en s'éloignant de zéro.

Figure 12–4
Mise en œuvre des fonctions ENT, IMPAIR et PAIR.



	B	D	E	F
2		ENT	IMPAIR	PAIR
4	*Syntaxe*	=ENT(D7)	=IMPAIR(E7)	=PAIR(F7)
6	*Arguments*			
7	*Nombre*	-65,60	4567,89	4567,89
9	*Résultat*	-66,00	4 569,00	4 568,00

À SAVOIR Valeurs négatives traitées « normalement »

La fonction *ENT* aborde les valeurs négatives en prenant en compte leur valeur réelle et non leur valeur absolue. Ainsi, à partir de `-65,4`, la fonction *ENT* renverra systématiquement `-66`.

Tableau 12–4 Fonctions d'arrondi

Fonction	Description
ARRONDI.AU.MULTIPLE	Cette fonction utilise deux arguments. Elle arrondit un nombre décimal (premier argument) au multiple le plus proche de la valeur spécifiée dans le deuxième argument (nombre décimal). Dans l'exemple présenté figure 12-5, on a bien `25 * 60 = 1 500` et `26 * 60 = 1 560`. `1 534` étant plus proche de `1 560`, c'est cette dernière qui est renvoyée par la formule. Si le nombre et le multiple sont de signes différents, la fonction renvoie une valeur d'erreur.

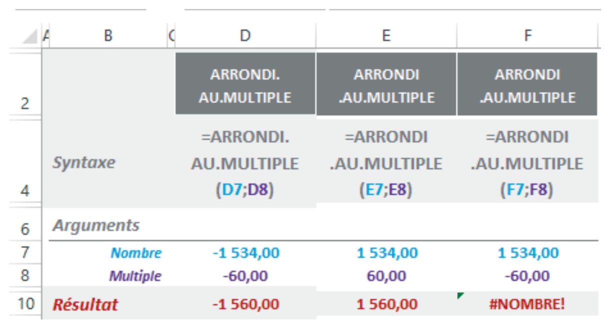


Figure 12–5
Mise en œuvre de la fonction ARRONDI.AU.MULTIPLE.

Tableau 12–5 Fonctions d'arrondi

Fonction	Description
PLANCHER	Cette fonction utilise deux arguments. Elle arrondit un nombre décimal (premier argument) au multiple le plus proche de la valeur spécifiée dans le deuxième argument (nombre décimal) en s'approchant de zéro. Elle applique les mêmes règles que la fonction *ARRONDI.AU.MULTIPLE* mais, au lieu d'arrondir à la valeur la plus proche, elle choisit systématiquement celle qui est plus près de zéro. Quand les deux arguments sont positifs (`1 534` et `60` dans l'exemple proposé à la figure 12-6), elle renvoie donc la plus petite valeur (`1 500`), mais quand ils sont tous les deux négatifs (`-1 534` et `-60`), elle renvoie la plus grande (`-1 500`, qui est en effet plus proche de zéro que la plus petite, `-1 560`). Lorsque les deux arguments sont de signes différents, la fonction réagit autrement. Si c'est le multiple qui est négatif, la fonction renvoie une valeur d'erreur. Si c'est le nombre (`-1 534`), la fonction renvoie la valeur arrondie la plus proche en s'éloignant de zéro (`-1 560`). Cette fonction a disparu de la version 2013.
PLANCHER.MATH ou *PLANCHER.PRECIS*	Sous Excel 2010, cette fonction s'intitule *PLANCHER.PRECIS* et utilise deux arguments. Sous Excel 2013, elle s'intitule *PLANCHER.MATH* et utilise un argument supplémentaire. Lorsque les deux arguments sont positifs (`1 534` et `60`), elle a un comportement similaire à la fonction *PLANCHER* et renvoie `1 500`. Elle réagit également de la même façon lorsque le nombre est négatif (`-1 534`) et le multiple positif (`60`) ; elle renvoie `-1 560`.

Tableau 12–5 Fonctions d'arrondi

	Elle diffère de la fonction *PLANCHER* dans les deux derniers cas de figure. Lorsque les deux arguments sont négatifs, elle renvoie `-1 560` et non pas `-1 500`. Et enfin, lorsque le nombre est positif (`1 534`) et le multiple négatif (`-60`), elle ne renvoie plus de valeur d'erreur mais `1 500`. Si vous travaillez sous Excel 2013 et souhaitez arrondir un nombre négatif, vous pouvez préciser un troisième argument, *mode*. En entrant une valeur différente de `0` (`1`, par exemple), vous arrondissez en vous approchant de zéro. **Nouveauté Excel 2013**.

Figure 12–6
Mise en œuvre de la fonction PLANCHER.

	PLANCHER	PLANCHER	PLANCHER	PLANCHER
Syntaxe	=PLANCHER (D7;D8)	=PLANCHER (E7;E8)	=PLANCHER (F7;F8)	=PLANCHER (G7;G8)
Arguments				
Nombre	1 534,00	-1 534,00	-1 534,00	1 534,00
Multiple	60,00	-60,00	60,00	-60,00
Résultat	1 500,00	-1 500,00	-1 560,00	#NOMBRE!

Figure 12–7
Mise en œuvre de la fonction PLANCHER.MATH (sous Excel 2013) ou PLANCHER.PRECIS (sous Excel 2010).

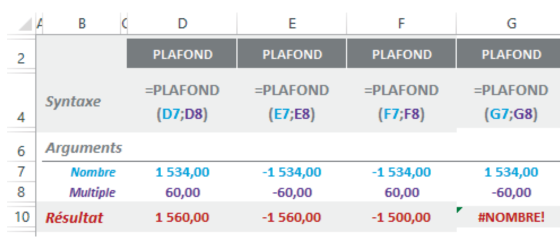
	PLANCHER. MATH	PLANCHER. MATH	PLANCHER. MATH	PLANCHER. MATH
Syntaxe sous Excel 2010	=PLANCHER. PRECIS(D8;D9)	=PLANCHER. PRECIS(E8;E9)	=PLANCHER. PRECIS(F8;F9)	=PLANCHER. PRECIS(G8;G9)
Syntaxe sous Excel 2013	=PLANCHER. MATH(D8;D9)	=PLANCHER. MATH(E8;E9;E10)	=PLANCHER. MATH(F8;F9;F10)	=PLANCHER. MATH(G8;G9)
Arguments				
Nombre	1 534,00	-1 534,00	-1 534,00	1 534,00
Multiple	60,00	-60,00	60,00	-60,00
Mode		0,00	1,00	
Résultat sous Excel 2010	1 500,00	-1 560,00	-1 560,00	1 500,00
Résultat sous Excel 2013	1 500,00	-1 560,00	-1 500,00	1 500,00

Tableau 12–6 Fonctions d'arrondi

Fonction	Description
PLAFOND	Cette fonction utilise deux arguments. Elle arrondit un nombre décimal (premier argument) au multiple le plus proche de la valeur spécifiée dans le deuxième argument (nombre décimal) en s'éloignant de zéro. Elle applique les mêmes règles que la fonction *ARRONDI.AU.MULTIPLE* mais, au lieu d'arrondir à la valeur la plus proche, elle choisit systématiquement celle qui est plus loin de zéro.

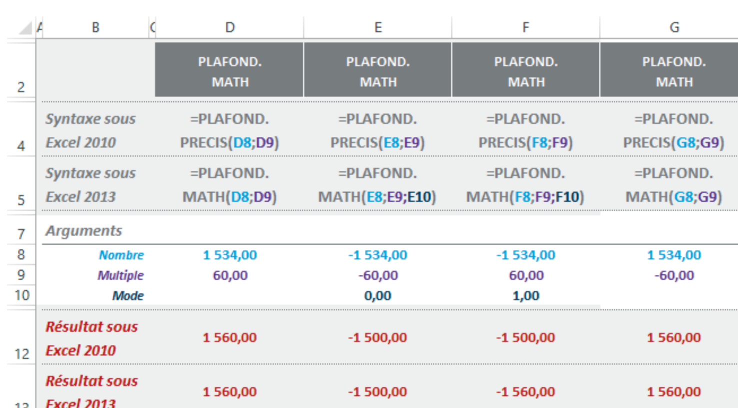
Tableau 12–6 Fonctions d'arrondi (suite)

	Quand les deux arguments sont positifs (1 534 et 60 dans l'exemple proposé à la figure 12-8), elle renvoie donc la plus grande valeur (1 560), mais quand ils sont tous les deux négatifs (-1 534 et -60), elle renvoie la plus petite (-1 560, qui est en effet plus loin de zéro que la plus grande, -1 500). Lorsque les deux arguments sont de signes différents, la fonction réagit autrement. Si c'est le multiple qui est négatif, la fonction renvoie une valeur d'erreur. Si c'est le nombre (-1 534), la fonction renvoie la valeur arrondie la plus proche en s'approchant de zéro (-1 500). Cette fonction a disparu de la version 2013.
PLAFOND.MATH ou *PLAFOND.PRECIS*	Sous Excel 2010, cette fonction s'intitule *PLAFOND.PRECIS* et utilise deux arguments. Sous Excel 2013, elle s'intitule *PLAFOND.MATH* et utilise un argument supplémentaire. Lorsque les deux arguments sont positifs (1 534 et 60), elle a un comportement similaire à la fonction *PLAFOND* et renvoie 1 560. Elle réagit également de la même façon lorsque le nombre est négatif (-1 534) et le multiple positif (60) ; elle renvoie -1 500. Elle diffère de la fonction *PLAFOND* dans les deux derniers cas de figure. Lorsque les deux arguments sont négatifs, elle renvoie -1 500 et non pas -1 560. Et enfin, lorsque le nombre est positif (1 534) et le multiple négatif (-60), elle ne renvoie plus de valeur d'erreur mais 1 560. Si vous travaillez sous Excel 2013 et souhaitez arrondir un nombre négatif, vous pouvez préciser un troisième argument, *mode*. En entrant une valeur différente de 0 (1, par exemple), vous arrondissez en vous éloignant de zéro. **Nouveauté Excel 2013**.



	B	D	E	F	G
2		PLAFOND	PLAFOND	PLAFOND	PLAFOND
4	*Syntaxe*	=PLAFOND (D7;D8)	=PLAFOND (E7;E8)	=PLAFOND (F7;F8)	=PLAFOND (G7;G8)
6	*Arguments*				
7	*Nombre*	1 534,00	-1 534,00	-1 534,00	1 534,00
8	*Multiple*	60,00	-60,00	60,00	-60,00
10	*Résultat*	1 560,00	-1 560,00	-1 500,00	#NOMBRE!

Figure 12–8
Mise en œuvre de la fonction PLAFOND.



	B	D	E	F	G
2		PLAFOND. MATH	PLAFOND. MATH	PLAFOND. MATH	PLAFOND. MATH
4	*Syntaxe sous Excel 2010*	=PLAFOND. PRECIS(D8;D9)	=PLAFOND. PRECIS(E8;E9)	=PLAFOND. PRECIS(F8;F9)	=PLAFOND. PRECIS(G8;G9)
5	*Syntaxe sous Excel 2013*	=PLAFOND. MATH(D8;D9)	=PLAFOND. MATH(E8;E9;E10)	=PLAFOND. MATH(F8;F9;F10)	=PLAFOND. MATH(G8;G9)
7	*Arguments*				
8	*Nombre*	1 534,00	-1 534,00	-1 534,00	1 534,00
9	*Multiple*	60,00	-60,00	60,00	-60,00
10	*Mode*		0,00	1,00	
12	*Résultat sous Excel 2010*	1 560,00	-1 500,00	-1 500,00	1 560,00
13	*Résultat sous Excel 2013*	1 560,00	-1 500,00	-1 560,00	1 560,00

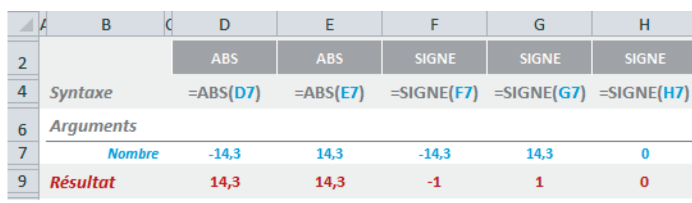
Figure 12–9
Mise en œuvre de la fonction PLAFOND.MATH (sous Excel 2013) ou PLAFOND.PRECIS (sous Excel 2010).

Fonctions afférentes au signe des nombres

Deux fonctions s'intéressent au signe des nombres : *ABS* et *SIGNE*. La première en fait abstraction et l'autre renvoie une information à son propos.

Tableau 12–7 Fonctions afférentes au signe des nombres

Fonction	Description
ABS	Cette fonction utilise un seul argument, un nombre décimal, dont elle renvoie la valeur absolue (le nombre privé de son signe).
SIGNE	Cette fonction utilise un seul argument, un nombre décimal, qu'elle analyse pour renvoyer 1, 0 ou −1 suivant qu'il est positif, nul ou négatif.



	A	B	C	D	E	F	G	H
2				ABS	ABS	SIGNE	SIGNE	SIGNE
4		*Syntaxe*		=ABS(D7)	=ABS(E7)	=SIGNE(F7)	=SIGNE(G7)	=SIGNE(H7)
6		*Arguments*						
7		*Nombre*		-14,3	14,3	-14,3	14,3	0
9		*Résultat*		14,3	14,3	-1	1	0

Figure 12–10
Mise en œuvre des fonctions ABS et SIGNE.

Fonctions afférentes aux nombres entiers

Certaines fonctions présentées dans cette section utilisent les nombres premiers, qui interviennent largement dans certains domaines des mathématiques appliquées, comme les algorithmes de cryptographie.

D'autres fonctions concernent le champ de l'arithmétique modulaire, dont le principe consiste à ne pas travailler sur les nombres eux-mêmes, mais sur le reste de leur division par une valeur quelconque.

> **RAPPEL Décomposition en facteurs premiers**
>
> Un nombre premier n'est divisible que par 1 et par lui-même. Un théorème fondamental de l'arithmétique nous apprend qu'un nombre entier se décompose de manière unique en un produit de facteurs premiers. Par exemple, 24 se décompose en $2^3 \times 3$ et 180 en $2^2 \times 3^2 \times 5$. Une telle décomposition permet de trouver le PGCD (plus grand commun diviseur) et le PPCM (plus petit commun multiple) de ces deux nombres.
> Pour trouver le PGCD, on ne prend que les facteurs premiers communs avec leur plus petit exposant (soit $2^2 \times 3 = 12$). Pour trouver le PPCM, on prend tous les facteurs premiers avec leur plus grand exposant (soit $2^3 \times 3^2 \times 5 = 360$).

Tableau 12–8 Fonctions afférentes aux décompositions en facteurs premiers

Fonction	Description
PGCD	Cette fonction utilise un nombre variable d'arguments. Elle renvoie le plus grand diviseur commun de tous les arguments (des entiers).
PPCM	Cette fonction utilise un nombre variable d'arguments. Elle renvoie le plus petit multiple commun de tous les arguments (des entiers).

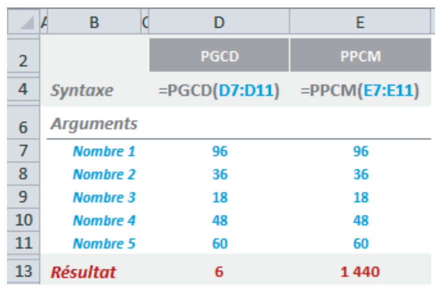


Figure 12–11
Mise en œuvre des fonctions PGCD et PPCM.

Les arguments utilisés par les fonctions *PGCD* et *PPCM* doivent être des entiers positifs. Les lois mathématiques impliquent que lorsqu'une valeur négative se glisse parmi eux, la fonction renvoie une valeur d'erreur. Une limite d'Excel fait que si l'un d'eux est supérieur à `9 007 199 254 740 990` (`2^53`), la fonction renvoie également une valeur d'erreur. Si la fonction rencontre une valeur décimale, elle la tronque à l'unité.

Tableau 12–9 Fonctions afférentes à l'arithmétique modulaire

Fonction	Description
QUOTIENT	Cette fonction utilise deux arguments (nombres décimaux). Elle considère le premier comme un numérateur, le second comme un dénominateur et réalise la division correspondante dont elle renvoie la partie entière du résultat.
MOD	Cette fonction utilise deux arguments (nombres décimaux). Elle divise le premier par le second et renvoie le reste (valeur entière) de cette opération pour un résultat entier.

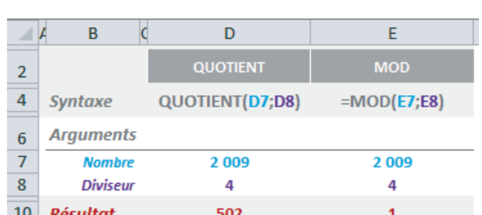


Figure 12–12
Mise en œuvre des fonctions QUOTIENT et MOD.

Sommes

L'addition est l'une des quatre opérations de l'arithmétique élémentaire. Comme ses petites sœurs, elle peut être directement mise en œuvre dans une formule grâce à l'opérateur +. Néanmoins, dans la majorité des cas, elle doit être appliquée à un très grand nombre de valeurs et il vaut mieux utiliser la fonction correspondante, qui évite de détailler les termes de l'opération. Dans la même famille, Excel propose des fonctions qui cumulent les valeurs d'une plage en excluant celles qui ne répondent pas à certains critères.

Figure 12–13
Pour illustrer les fonctions de la catégorie Sommes, nous utiliserons ce petit tableau. La feuille sur laquelle il se trouve s'appelle Données.

	A	B	C	D	E
1	Prénom	Âge	Sport	Nb enfants	Budget vacances annuel
2	Jean	47	Tennis	2	10 000
3	Paul	24	Judo	1	1 000
4	Roger	36	Tennis		3 000
5	Albert	57	Natation	3	12 000
6	Felix	62	Natation	4	8 000
7	Michel	27	Natation		1 200
8	Georges	32	Judo		600
9	Simon	29	Tennis	2	700
10	Ernest	43	Tennis	1	3 000
11	Amelin	26	Judo		1 300

ASTUCE Des noms pour clarifier les formules

Dans le tableau présenté figure 12-13, la plage *B2:B11* a été nommée `Age`, la plage *C2:C11* `Sport`, la plage *D2:D11* `Enfants` et la plage *E2:E11* `Budget` (pour savoir comment baptiser une plage, consultez le chapitre 2).

Tableau 12–10 Sommes

Fonction	Description
SOMME	Cette fonction renvoie la somme des valeurs stockées dans ses arguments (nombres décimaux). Il peut s'agir d'une plage unique, de plusieurs plages ou de valeurs exprimées « en dur » dans la formule. Elle utilise donc un nombre variable d'arguments. Quelle que soit la forme de sa syntaxe, elle ne prend en compte que les valeurs numériques (les valeurs de texte sont considérées comme nulles).

Figure 12–14
Mise en œuvre de la fonction SOMME.

	B	D	E	F
2		SOMME	SOMME	SOMME
4	*Syntaxe*	=SOMME (E2:E11)	=SOMME (Données!E2:E11)	=SOMME (Budget)
6	*Arguments*			
7	*Plage*	E2:E11	Données!E2:E11	Budget
9	*Résultat*	40 800	40 800	40 800

La figure 12-14 propose trois syntaxes différentes pour mener à bien le même calcul à partir de la même plage. Si la formule est entrée dans une cellule de la feuille `Données`, l'exemple de la colonne *D* convient tout à fait. Si elle se trouve dans une autre feuille située dans le même classeur, il faut utiliser la formule proposée dans la colonne *E*.

Dans notre exemple, la plage *E2:E11* a été nommée `Budget`. Aussi, lors de la saisie de la fonction *SOMME*, dès que vous cliquez-glissez sur elle, c'est le nom `Budget` qui s'inscrit dans la formule. Vous pouvez alors utiliser la troisième syntaxe, présentée colonne *F*.

Réaliser une somme respectant une condition

Pour que la fonction *SOMME.SI* ait un sens, les deux plages (premier et troisième argument) doivent avoir la même taille, car les données de l'une sont liées à celles de l'autre par leur position. La première cellule de la plage de somme correspond à la première cellule de la plage de filtre et ainsi de suite. Seules les cellules remplissant la condition exprimée dans le deuxième argument sont prises en compte dans la somme.

Tableau 12–11 Sommes

Fonction	Description
SOMME.SI	Cette fonction fait la somme des valeurs d'une plage respectant un certain critère. Dans la figure 12-15, on fait la somme des valeurs de la plage `Budget` uniquement lorsque la valeur correspondante de la plage `Enfants` est strictement supérieure à zéro. Le résultat obtenu est le cumul des budgets de vacances pour les personnes ayant au moins un enfant. Cette fonction utilise trois arguments. Le troisième est la plage de cellules contenant les données à additionner (`Budget`). Le premier est une plage de cellules (`Enfants`) sur lesquelles vous appliquez le filtre indiqué dans le deuxième argument (`>0`).

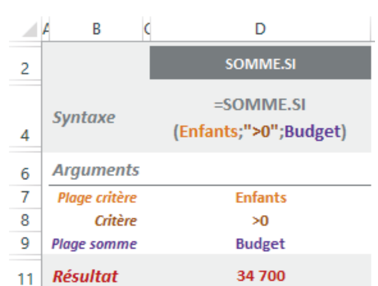


Figure 12–15
Mise en œuvre de la fonction SOMME.SI.

Un critère s'exprime à l'aide de l'un des six opérateurs de comparaison (pour les connaître, consultez le début du chapitre 4). Si vous ne précisez aucun opérateur, Excel comprend par défaut qu'il s'agit du signe égal. Pour cumuler les budgets des joueurs de tennis, vous pouvez utiliser les critères suivants : `"=Tennis"`, `"Tennis"`, `A1` (si la cellule *A1* contient le texte `Tennis`), `"T*"` ou `"T?????"` (Excel comprend les caractères génériques). S'il ne s'agit pas du signe égal, vous devez préciser l'opérateur. Pour cumuler les budgets des individus de moins de 30 ans, vous pouvez utiliser les critères suivants : `"<30"`, `"<"&A1` (si la cellule *A1* contient la valeur `30`).

Figure 12–16
Les lignes sélectionnées par le critère (arguments 1 et 2) apparaissent sur un fond mauve.

Prénom	Âge	Sport	Nb enfants	Budget vacances annuel
Jean	47	Tennis	2	10 000
Paul	24	Judo	1	1 000
Roger	36	Tennis		3 000
Albert	57	Natation	3	12 000
Felix	62	Natation	4	8 000
Michel	27	Natation		1 200
Georges	32	Judo		600
Simon	29	Tennis	2	700
Ernest	43	Tennis	1	3 000
Amelin	26	Judo		1 300

= 10 000 + 1 000 + 12 000 + 8 000 + 700 + 3 000
= 34 700

> **MISE EN GARDE Limite de la fonction SOMME.SI**
>
> Attention, le deuxième argument de la fonction *SOMME.SI* ne peut prendre en compte qu'une condition unique. Si vous souhaitez soumettre votre cumul à des conditions multiples, il faut utiliser la fonction *SOMME.SI.ENS*.

Réaliser une somme respectant plusieurs conditions

La fonction *SOMME.SI.ENS* (voir la figure 12-17) utilise un nombre variable d'arguments, qui dépend des conditions auxquelles vous souhaitez soumettre votre somme. Le premier argument est la plage de cellules contenant les données à additionner (`Budget`). Ensuite, les arguments vont par paires. Le deuxième argument correspond à la plage (`Enfants`) sur laquelle s'applique la première condition (`>0`), elle-même exprimée dans le troisième argument. Si vous souhaitez filtrer encore davantage votre somme, utilisez un quatrième argument pour indiquer une nouvelle plage (`Sport`) sur laquelle s'appliquera la deuxième condition (`Tennis`, dans le cinquième argument) et ainsi de suite (jusqu'à la 127e paire !).

Tableau 12–12 Sommes

Fonction	Description
SOMME.SI.ENS	Cette fonction suit la même logique que *SOMME.SI*, mais en autorisant l'emploi simultané de plusieurs critères. Elle additionne les valeurs d'une plage en omettant celles qui ne remplissent pas les critères exprimés à partir du deuxième argument. Dans la figure 12-17, on fait la somme des valeurs de la plage `Budget` uniquement lorsque la valeur correspondante de la plage `Enfant` est strictement supérieure à zéro et lorsque la valeur correspondante de la plage `Sport` est égale à `tennis`. Le résultat obtenu est le cumul des budgets de vacances pour les personnes jouant au tennis et ayant au moins un enfant.

Figure 12–17
Mise en œuvre de la fonction SOMME.SI.ENS.

	A	B	C	D
2				SOMME.SI.ENS
4		*Syntaxe*		=SOMME.SI.ENS(Budget;Enfants;">0";Sport;"Tennis")
6		*Arguments*		
7		*Plage somme*		Budget
8		*Plage critère 1*		Enfants
9		*Critère 1*		>0
10		*Plage critère 2*		Sport
11		*Critère 2*		Tennis
13		***Résultat***		**13 700**

Figure 12–18
Les lignes sélectionnées par les critères (arguments 2, 3, 4 et 5) apparaissent sur un fond mauve.

Prénom	Âge	Sport	Nb enfants	Budget vacances annuel
Jean	47	Tennis	2	10 000
Paul	24	Judo	1	1 000
Roger	36	Tennis		3 000
Albert	57	Natation	3	12 000
Felix	62	Natation	4	8 000
Michel	27	Natation		1 200
Georges	32	Judo		600
Simon	29	Tennis	2	700
Ernest	43	Tennis	1	3 000
Amelin	26	Judo		1 300

= 10 000 + 700 + 3 000 = 13 700

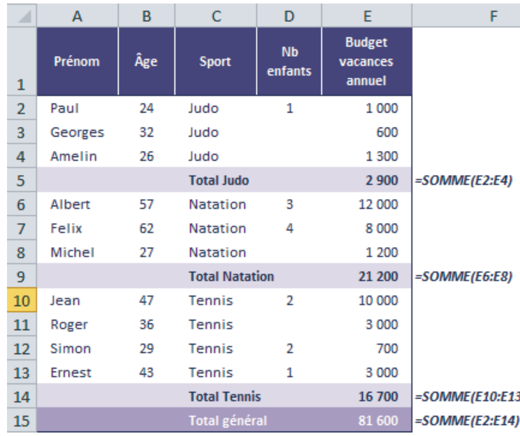
Calculer des sous-totaux

Lorsque, dans un tableau, vous gérez plusieurs niveaux, la présence de totaux intermédiaires vous fait toujours courir le risque de compter les valeurs en double, voire en triple ! C'est ce que la fonction *SOUS.TOTAL* vous évite. Même si les plages de calcul se recouvrent, à aucun moment, vous ne risquerez ces cumuls intempestifs. La façon la plus simple d'intégrer des sous-totaux à un tableau est d'utiliser la commande *Données>Plan>Sous-total* (consultez le chapitre 3). Néanmoins, il peut être utile de savoir manipuler cette fonction indépendamment de la commande *Sous-total*.

ASTUCE Sous-totaux et plan

Si vous utilisez la commande *Données>Plan>Sous-total*, un plan est automatiquement installé avec les sous-totaux. Si ce plan vous dérange, cliquez sur n'importe quelle cellule du tableau, déroulez *Données>Plan>Dissocier* et sélectionnez *Effacer le plan*.

Figure 12–19
Ce tableau illustre une utilisation maladroite de la fonction SOMME engendrant des dysfonctionnements facilement résolus par la mise en œuvre de la fonction SOUS.TOTAL.

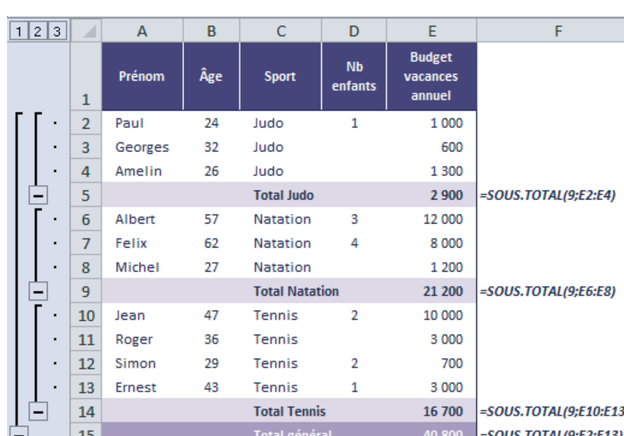


	A	B	C	D	E	F
1	Prénom	Âge	Sport	Nb enfants	Budget vacances annuel	
2	Paul	24	Judo	1	1 000	
3	Georges	32	Judo		600	
4	Amelin	26	Judo		1 300	
5			Total Judo		2 900	=SOMME(E2:E4)
6	Albert	57	Natation	3	12 000	
7	Felix	62	Natation	4	8 000	
8	Michel	27	Natation		1 200	
9			Total Natation		21 200	=SOMME(E6:E8)
10	Jean	47	Tennis	2	10 000	
11	Roger	36	Tennis		3 000	
12	Simon	29	Tennis	2	700	
13	Ernest	43	Tennis	1	3 000	
14			Total Tennis		16 700	=SOMME(E10:E13)
15			Total général		81 600	=SOMME(E2:E14)

Tableau 12–13 Sommes

Fonction	Description
SOUS.TOTAL	Cette fonction fait la somme, la moyenne, le dénombrement (ou huit autres opérations statistiques) d'une ou plusieurs plage(s) de cellules. Si on la compare aux fonctions *SOMME* ou *MOYENNE*, elle présente l'avantage d'introduire dans un tableau des calculs intermédiaires dont les résultats ne sont pas pris en compte dans les calculs généraux. Le tableau présenté figure 12-19 utilise des fonctions *SOMME* à mauvais escient. La formule entrée en *E15* cumule non seulement les données initiales, mais aussi les résultats intermédiaires, renvoyant une valeur qui est le double de ce qu'elle devrait être. Dans la figure 12-20, la fonction *SOUS.TOTAL* a été entrée en *E5*, *E9*, *E14* et *E15*. La syntaxe des quatre fonctions est donnée en *F5*, *F9*, *F14* et *F15*. La formule entrée en *E15* réalise une somme (code 9 indiqué dans le premier argument) et cumule les valeurs de la plage *E2:E14*. Dans cet exemple, la fonction *SOUS.TOTAL* utilise deux arguments, mais elle pourrait en comporter davantage. Si le calcul doit, par exemple, porter sur trois plages de cellules, la fonction comprendra quatre arguments (le code de l'opération et les trois plages de cellules).

Figure 12–20
Mise en œuvre de la fonction SOUS.TOTAL.



	A	B	C	D	E	F
1	Prénom	Âge	Sport	Nb enfants	Budget vacances annuel	
2	Paul	24	Judo	1	1 000	
3	Georges	32	Judo		600	
4	Amelin	26	Judo		1 300	
5			Total Judo		2 900	=SOUS.TOTAL(9;E2:E4)
6	Albert	57	Natation	3	12 000	
7	Felix	62	Natation	4	8 000	
8	Michel	27	Natation		1 200	
9			Total Natation		21 200	=SOUS.TOTAL(9;E6:E8)
10	Jean	47	Tennis	2	10 000	
11	Roger	36	Tennis		3 000	
12	Simon	29	Tennis	2	700	
13	Ernest	43	Tennis	1	3 000	
14			Total Tennis		16 700	=SOUS.TOTAL(9;E10:E13)
15			Total général		40 800	=SOUS.TOTAL(9;E2:E13)

ATTENTION Compter ou non les cellules masquées

Le premier argument de la fonction indique le code de l'opération. Vous disposez de deux séries de codes. La première série (de 1 à 11) met en œuvre une fonction statistique (somme, moyenne, dénombrement, etc.) qui prend en compte toutes les valeurs référencées dans les arguments 2 à n de la fonction (cellules masquées incluses). La deuxième série (de 101 à 111) réalise les mêmes calculs, mais sans tenir compte des valeurs stockées dans les cellules masquées.

Sélection de la fonction de calcul		
Intègre les valeurs masquées	Exclut les valeurs masquées	Fonction
1	101	MOYENNE
2	102	NB
3	103	NBVAL
4	104	MAX
5	105	MIN
6	106	PRODUIT
7	107	ECARTYPE
8	108	ECARTYPEP
9	109	SOMME
10	110	VAR
11	111	VAR.P

Figure 12–21 Deux séries de codes sont disponibles pour préciser le premier argument de la fonction SOUS.TOTAL.

Par cellules masquées, on comprend les lignes que vous avez masquées à l'aide de la commande éponyme. Si des lignes se retrouvent masquées par l'application d'un filtre (voir le chapitre 3), les données correspondantes ne seront pas prises en compte dans le calcul, quelle que soit la série de codes utilisée.
Enfin, il faut savoir que la différence de comportement en fonction du code ne s'applique qu'aux lignes. Pour les colonnes, quelle que soit la nature du code, les données seront prises en compte, qu'elles soient masquées ou non.

Fonctions particulières

Voici quatre fonctions inclassables, qui permettent de régler des problèmes propres à Excel ou de répondre à des besoins spécifiques.

Fonctions statistiques et valeurs d'erreur

Les fonctions *MOYENNE*, *NB*, *MIN*, *MAX*, etc. réalisent des calculs statistiques sur des plages de cellules (voir le chapitre 13). Si des valeurs d'erreur se sont glissées dans ces plages, ces fonctions sont dans l'incapacité de renvoyer un résultat.

	A	B	C	D
2				AGREGAT
3		Syntaxe		=AGREGAT(D5;D6;Budget)
4		Arguments		
5		N° fonction		1
6		N° option		1
7		Référence		Budget
8		Résultat		4 080

Figure 12–22
Mise en œuvre de la fonction AGREGAT. L'argument Budget désigne la plage E2:E11 du tableau présenté figure 12-13. Le premier argument (1) indique qu'Excel calcule une moyenne et le deuxième argument (1) indique que ce calcul ignore les lignes masquées et les fonctions SOUS.TOTAL et AGREGAT imbriquées (voir la figure 12-23).

La fonction *AGREGAT* utilise un code (nombre entier compris entre 1 et 19) qui lui permet de jouer le rôle de dix-neuf fonctions statistiques différentes (voir la figure 12-23). En ignorant les erreurs, la fonction *AGREGAT* est capable de renvoyer un résultat, même à partir d'une plage « polluée ». L'utilisation de cette fonction résout donc certains problèmes surgissant à l'occasion d'une mise en forme conditionnelle, lorsque les barres de données, les jeux d'icônes et les nuances de couleurs sont incapables d'afficher une mise en forme pour cause d'erreurs dans les plages.

Tableau 12–14 Fonctions particulières

Fonction	Description
AGREGAT	La fonction *AGREGAT* permet de faire dix-neuf calculs statistiques différents (moyenne, somme, dénombrement, etc.). Elle utilise un nombre variable d'arguments. Le premier, un entier compris entre 1 et 19, représente le code de la fonction à utiliser (voir la figure 12-23). Le deuxième, un entier compris entre 1 et 7, précise la manière dont doivent être traités calculs intermédiaires, valeurs d'erreur et lignes masquées (voir également la figure 12-23). Les arguments suivants (en nombre variable) indiquent les références des plages contenant les valeurs faisant l'objet du calcul. La fonction réalise l'opération statistique précisée dans le premier argument sur les données indiquées dans le troisième argument et les suivants.

Figure 12–23
Le tableau de gauche présente les dix-neuf codes à utiliser comme premier argument de la fonction AGREGAT. Le tableau de droite liste les huit options disponibles pour mener à bien son calcul (à indiquer dans le deuxième argument de la fonction).

N° fonctions			
1	MOYENNE	11	VAR.P.N
2	NB	12	MEDIANE
3	NBVAL	13	MODE.SIMPLE
4	MAX	14	GRANDE.VALEUR
5	MIN	15	PETITE.VALEUR
6	PRODUIT	16	CENTILE.INCLURE
7	ECARTYPE.STANDARD	17	QUARTILE.INCLURE
8	ECARTYPE.PEARSON	18	CENTILE.EXCLURE
9	SOMME	19	QUARTILE.EXCLURE
10	VAR.S		

N° options	
Ignore :	
0	fonctions SOUS.TOTAL et AGREGAT imbriquées
1	lignes masquées et option 0
2	valeurs d'erreur et option 0
3	lignes masquées, valeurs d'erreur et option 0
5	lignes masquées
6	valeurs d'erreur
7	lignes masquées et valeurs d'erreur
N'ignore rien :	
4	N'ignore rien

Développements limités

En physique et en mathématiques, le développement limité d'une fonction F au voisinage d'un point est une approximation polynomiale de cette fonction en ce point. En physique, il est fréquent de confondre la fonction avec son développement limité, à condition que le reste soit inférieur à l'erreur autorisée. *SOMME.SERIES* sert à calculer des fonctions polynomiales.

$$a_1x^n + a_2x^{(n+m)} + a_3x^{(n+2m)} + \ldots + a_ix^{(n+(i-1)m)}$$

Figure 12–24
Syntaxe de la fonction polynomiale construite par SOMME.SERIES.

$$A = \boxed{a_1 \mid a_2 \mid a_3 \mid \ldots \mid a_i}$$

Figure 12–25
Vecteur des coefficients. C'est sa taille qui détermine le nombre de termes de la fonction polynomiale.

Tableau 12–15 Fonctions particulières

Fonction	Description
SOMME.SERIES	*SOMME.SERIES* calcule la fonction polynomiale présentée figure 12-24. Sa syntaxe est `SOMME.SERIES(x;n;m;A)`. `x`, `n` et `m` sont les trois premiers arguments de la fonction. Le quatrième est un vecteur (consultez le chapitre 5). Il correspond à la liste des coefficients a_1, a_2 … a_i (voir la figure 12-25).

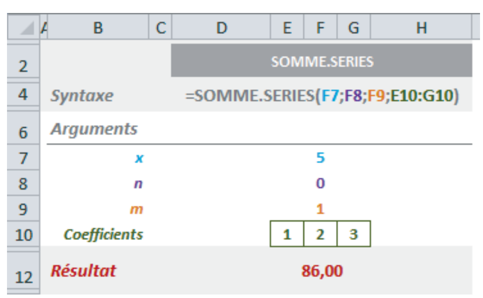


Figure 12–26
Mise en œuvre de la fonction SOMME.SERIES.

$$F(x) = 1x^{(0)} + 2x^{(0+1)} + 3x^{(0+2*1)}$$
$$= 1 + 2x + 3x^2$$

$$F(5) = 1 + 2*5 + 3*5^2$$
$$= 1 + 10 + 75 = 86$$

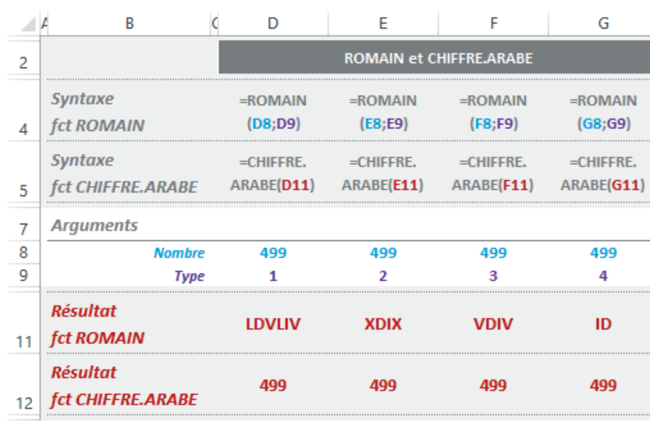
Figure 12–27
Fonction polynomiale utilisée par l'exemple proposé figure 12-26 et son calcul pour x=5.

L'un des exemples développés dans la deuxième section de ce chapitre utilise la fonction *SOMME.SERIE*.

Convertir un nombre en chiffres romains ou en chiffres arabes

Tableau 12–16 Fonctions particulières

Fonction	Description
ROMAIN	Cette fonction convertit en chiffres romains un entier compris entre 0 et 3 999 indiqué dans le premier argument de la fonction. Elle effectue cette conversion selon le type d'écriture précisé dans le deuxième argument (ce dernier va de 1 à 4 impliquant une concision croissante).
CHIFFRE.ARABE	Cette fonction convertit en chiffres arabes une valeur numérique exprimée en chiffres romains et donnée en argument. Elle n'a pas les mêmes limitations que la fonction *ROMAIN* et peut convertir des valeurs qui sont bien supérieures à 3 999. **Nouveauté Excel 2013**.



		D	E	F	G
2		ROMAIN et CHIFFRE.ARABE			
4	*Syntaxe fct ROMAIN*	=ROMAIN (D8;D9)	=ROMAIN (E8;E9)	=ROMAIN (F8;F9)	=ROMAIN (G8;G9)
5	*Syntaxe fct CHIFFRE.ARABE*	=CHIFFRE. ARABE(D11)	=CHIFFRE. ARABE(E11)	=CHIFFRE. ARABE(F11)	=CHIFFRE. ARABE(G11)
7	*Arguments*				
8	*Nombre*	499	499	499	499
9	*Type*	1	2	3	4
11	*Résultat fct ROMAIN*	LDVLIV	XDIX	VDIV	ID
12	*Résultat fct CHIFFRE.ARABE*	499	499	499	499

Figure 12–28
Mise en œuvre des fonctions ROMAIN et CHIFFRE.ARABE.

Convertir un nombre décimal en diverses bases et inversement

> **PLUS DE FONCTIONS Fonctions de l'ingénieur**
>
> Dans le chapitre 14, vous trouverez toute une section réservée à une catégorie de fonctions de l'ingénieur dédiées aux conversions en base 2, 8, 10 et 16.

Tableau 12–17 Fonctions de conversion

Fonction	Description
BASE	Cette fonction utilise trois arguments. Elle convertit un nombre entier exprimé en base 10 (premier argument compris entre 0 et 9 007 199 254 740 990, autrement dit 2^{53}) en un nombre (au format texte) exprimé dans une base quelconque (deuxième argument qui doit être une valeur entière comprise entre 2 et 36). Le troisième argument est facultatif et sert à préciser la longueur minimale de la chaîne à renvoyer (entier positif). Pour comprendre le rôle de ce troisième argument, voir la figure 12-29. La fonction *DECIMAL* permet de revenir à la valeur initiale. **Nouveauté Excel 2013**.

Tableau 12–17 Fonctions de conversion (suite)

Fonction	Description
DECIMAL	Cette fonction utilise deux arguments. Elle convertit un nombre au format texte (premier argument) exprimé dans une base quelconque (deuxième argument qui doit être une valeur entière comprise entre 2 et 36) en un nombre entier exprimé en base 10. La fonction *BASE* permet de revenir à la valeur initiale. **Nouveauté Excel 2013**.

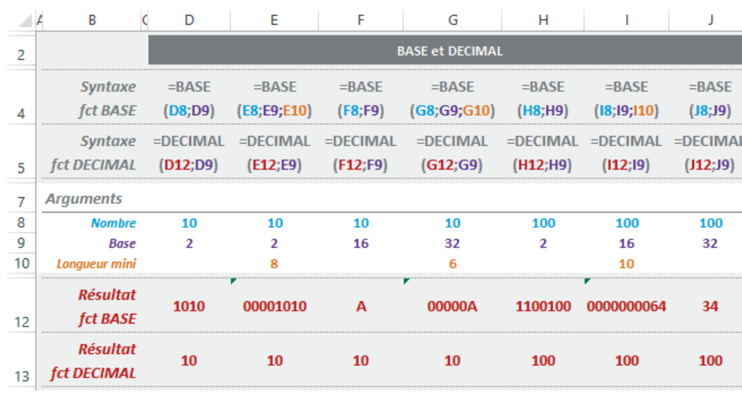


Figure 12–29
Mise en œuvre des fonctions BASE et DECIMAL.

Produits

La multiplication est l'une des quatre opérations de l'arithmétique élémentaire. Comme ses petites sœurs, elle peut être utilisée directement dans une formule grâce à l'opérateur *. Excel propose trois fonctions pour faciliter sa mise en œuvre. L'une d'elles permet de multiplier tous les termes d'une plage sans les détailler, l'autre facilite les élévations à la puissance et la dernière calcule les racines carrées.

Tableau 12–18 Produits

Fonction	Description
PRODUIT	Cette fonction calcule le produit des nombres décimaux contenus dans les plages spécifiées en argument. Sa syntaxe est =PRODUIT(X_i:X_j;X_k:X_l;…).

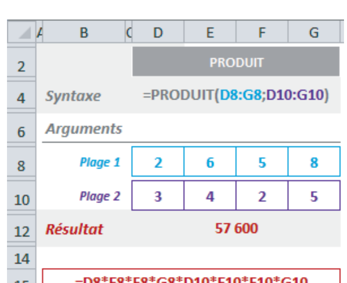
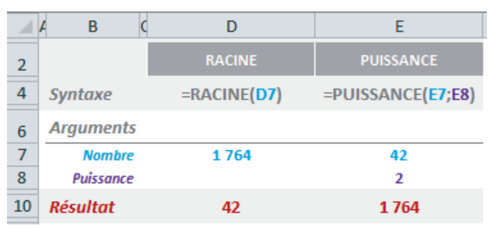


Figure 12–30
Mise en œuvre de la fonction PRODUIT. La formule donnée dans le cadre rouge explicite l'opération effectuée par cette fonction. Dans cet exemple, elle est calculée à partir de deux plages, mais elle pourrait en utiliser davantage.

> **RAPPEL Puissance, racine**
>
> Si a est multiplié n fois par lui-même (a × a × … × a), il est plus synthétique d'écrire cette opération sous la forme a^n (on lit cette expression `a puissance n`).
> La racine est l'opération réciproque de la puissance. La racine carrée de a^2 est a. La fonction *RACINE* renvoie une racine carrée, mais on peut calculer tous les types de racine (cubiques, etc.). Dans ce cas, il faut utiliser l'opérateur ^ avec une puissance fractionnaire ; par exemple, pour calculer la racine cubique de 8, il faut utiliser la formule `=8^(1/3)`, dont le résultat est 2.

Tableau 12–19 Produits

Fonction	Description
PUISSANCE	Cette fonction élève un nombre à une certaine puissance. Elle utilise deux arguments (nombres décimaux). Le premier est le nombre et le deuxième est la puissance. Pour faire ce calcul, Excel fournit également l'opérateur ^. Ainsi, `=PUISSANCE(2;3)` peut aussi s'écrire `=2^3`, les deux formules renvoyant 8 (2 × 2 × 2).
RACINE	Cette fonction utilise un seul argument, un nombre décimal, dont elle renvoie la racine carrée.



	B	D	E
2		RACINE	PUISSANCE
4	*Syntaxe*	=RACINE(D7)	=PUISSANCE(E7;E8)
6	*Arguments*		
7	*Nombre*	1 764	42
8	*Puissance*		2
10	*Résultat*	42	1 764

Figure 12–31
Mise en œuvre des fonctions PUISSANCE et RACINE.

Exponentielles et logarithmes

L'objet d'une fonction logarithme est de transformer un produit en une somme : `Log(a × b) = Log(a) + Log(b)`. C'est la réciproque d'une fonction exponentielle. Les plus connus sont le logarithme naturel ou népérien (de base e), le logarithme décimal (de base 10), très utilisé en physique, et le logarithme binaire (de base 2), très utilisé en informatique.

> **HISTOIRE D'où viennent les logarithmes ?**
>
> Vers la fin du XVI^e siècle, le développement de l'astronomie, de la navigation et des calculs bancaires d'intérêts composés poussent les mathématiciens à mettre au point des méthodes de simplification des calculs et, en particulier, à chercher des relations entre des suites arithmétiques et des suites géométriques.

Dans les calculs numériques, le logarithme le plus pratique est le logarithme décimal. Ainsi :

- `Log(10) = 1`
- `Log(100) = Log(10 × 10) = Log(10) + Log(10) = 1 + 1 = 2`

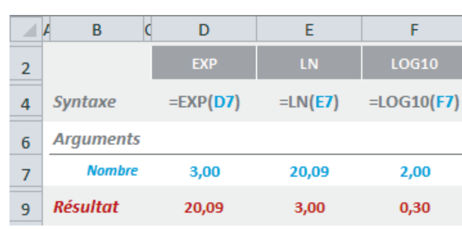
- `Log(1000) = Log(10`^{3}`) = 3 × Log(10) = 3 × 1 = 3`
- `Log(0,01) = Log(10`$^{-2}$`) = -2`
- etc.

La valeur du logarithme d'autres nombres que les puissances de `10` demande un calcul approché. `Log(2)`, par exemple, peut s'approcher en remarquant que `2`^{10}` = 1024`, soit environ `1000`. Donc, `Log(2`^{10}`)` est à peu près égal à `Log(1000)`. Ainsi, `10 × Log(2)` est à peu près égal à `3`, d'où `Log(2)` vaut environ `3/10`, c'est-à-dire `0,3`.

> **COMPRENDRE La constante e**
>
> Elle est probablement la constante réelle la plus importante des mathématiques après π. Elle est égale à environ `2,71828182845904` et est la base du logarithme népérien (ln). La fonction *EXP* est la réciproque de la fonction *LN*.

Tableau 12–20 Exponentielles et logarithmes

Fonction	Description
EXP	Cette fonction n'utilise qu'un argument, un nombre décimal. Elle s'en sert pour élever la constante `e` à la puissance et renvoyer le résultat de ce calcul. Pour élever à la puissance d'autres bases, utilisez l'opérateur `^`.
LN	Cette fonction utilise un seul argument, un nombre décimal, dont elle renvoie le logarithme népérien.
LOG10	Cette fonction utilise un seul argument, un nombre décimal, dont elle renvoie le logarithme en base `10`.



	B	D	E	F
2		EXP	LN	LOG10
4	Syntaxe	=EXP(D7)	=LN(E7)	=LOG10(F7)
6	Arguments			
7	Nombre	3,00	20,09	2,00
9	Résultat	20,09	3,00	0,30

Figure 12–32
Mise en œuvre des fonctions EXP, LN et LOG10. Ayant utilisé les mêmes valeurs pour les fonctions EXP et LN, on constate bien que l'une est la réciproque de l'autre. Le calcul de LOG10(2) renvoie une valeur qui correspond à l'approximation proposée un peu plus haut.

Tableau 12–21 Exponentielles et logarithmes

Fonction	Description
LOG	Cette fonction utilise deux arguments (nombres décimaux). Le second argument lui sert à savoir dans quelle base elle doit calculer le logarithme du nombre indiqué dans le premier argument. Si vous ne précisez pas le deuxième argument, la fonction renvoie le logarithme en base `10`.

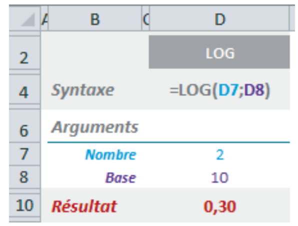


Figure 12–33
Mise en œuvre de la fonction LOG. Ayant indiqué 10 en deuxième argument et 2 en premier argument, le calcul effectué ici est équivalent à LOG10(2).

Calculs matriciels

En mathématiques, « linéaire » signifie « du premier degré ». La résolution d'une équation du premier degré à une inconnue ou d'un système de n équations à n inconnues correspond à un calcul d'algèbre linéaire. Dès que ces calculs deviennent trop compliqués pour être effectués séparément, ils peuvent être traités « en bloc » grâce à un outil mathématique introduit vers 1850 par James Joseph Sylvester : les matrices, dont la théorie a été établie par Hamilton et Cayley.

> **RAPPEL Bien valider les formules matricielles**
>
> N'oubliez pas que vous devez valider ces fonctions en pressant simultanément les touches *Ctrl+Maj+Entrée*. Consultez la fin du chapitre 4 pour consolider vos connaissances en matière de calcul matriciel.

Tableau 12–22 Calculs matriciels

Fonction	Description
PRODUITMAT	Cette fonction utilise deux arguments, deux matrices, dont elle renvoie le produit matriciel. Le résultat est une matrice comportant le même nombre de lignes que le premier argument et le même nombre de colonnes que le second.

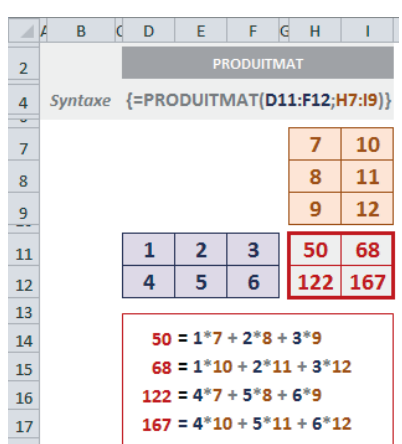


Figure 12–34
Mise en œuvre de la fonction PRODUITMAT. Elle a été entrée dans la plage H11:I12 et validée en pressant simultanément les touches Ctrl+Maj+Entrée. Pour information, on a indiqué, tout en bas, les calculs menés par Excel pour renvoyer la matrice résultat.

Tableau 12–23 Calculs matriciels

Fonction	Description
DETERMAT	Cette fonction utilise un argument, une matrice carrée, dont elle renvoie le déterminant. Il n'est pas nécessaire de la valider avec les touches *Ctrl+Maj+Entrée*. Si l'utilisateur distrait spécifie une matrice non carrée, *DETERMAT* renvoie la valeur d'erreur #VALEUR!.

Matrice carrée

Une matrice carrée a le même nombre de lignes et de colonnes.

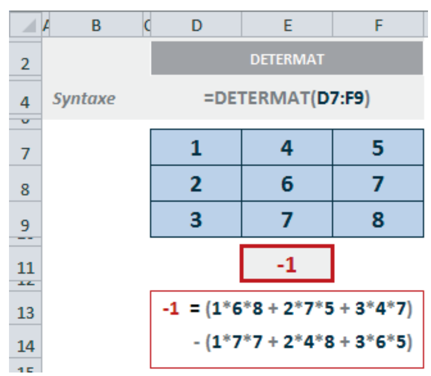


Figure 12–35
Mise en œuvre de la fonction DETERMAT. Elle a été entrée dans la cellule E11. Pour information, on a indiqué, tout en bas, les calculs menés par Excel pour renvoyer le déterminant.

Tableau 12–24 Calculs matriciels

Fonction	Description
INVERSEMAT	Cette fonction utilise un seul argument, une matrice carrée, dont elle renvoie la matrice inverse. Pour mener à bien le calcul d'une matrice inverse, il faut diviser chaque élément de calcul par le déterminant de la matrice : les matrices carrées dont le déterminant est égal à 0 ne peuvent donc pas être inversées. Si l'utilisateur distrait spécifie une matrice non carrée, *INVERSEMAT* renvoie une matrice remplie des valeurs d'erreur #VALEUR!.

Les fonctions *SOMME.X2MY2*, *SOMME.X2PY2* et *SOMME.XMY2* correspondent à des calculs courants dans le domaine statistique. Par ailleurs, elles renvoient une valeur unique et il n'est pas nécessaire de les valider avec les touches *Ctrl+Maj+Entrée*.

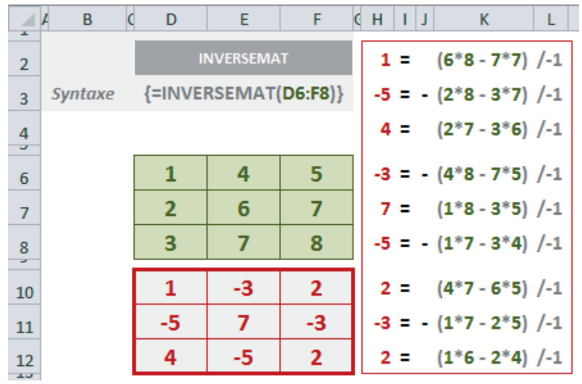


Figure 12–36
Mise en œuvre de la fonction INVERSEMAT. Elle a été entrée dans la plage D10:F12 et validée en pressant simultanément les touches Ctrl+Maj+Entrée. Pour information, on a indiqué à droite les calculs menés par Excel pour renvoyer la matrice inverse.

NOUVEAUTÉ EXCEL 2013 Matrice identité

Le produit d'une matrice par son inverse est égal à la matrice identité, c'est-à-dire une matrice carrée qui a des 1 sur sa diagonale et des 0 partout ailleurs. Excel 2013 offre une nouvelle fonction, *MATRICE.UNITAIRE*. Elle utilise un seul argument, la dimension de la matrice, et renvoie la matrice identité correspondante. Pour créer une matrice identité :

1. Sélectionnez une plage de cellules avec le même nombre de lignes que de colonnes (par exemple, trois lignes sur trois colonnes).
2. Saisissez `=MATRICE.UNITAIRE(3)`.
3. Validez la fonction en pressant les touches *Ctrl+Maj+Entrée*.

Vous obtenez une matrice de trois lignes sur trois colonnes, avec des 1 sur sa diagonale et des 0 partout ailleurs.
En fait, cette façon d'utiliser la fonction n'est pas la plus utile. En revanche, imbriquée dans d'autres fonctions, elle prendra tout son intérêt pour mener à bien des calculs matriciels élaborés.

Tableau 12–25 Calculs matriciels

Fonction	Description
SOMME.X2MY2	Cette fonction utilise deux arguments, deux matrices, dont elle renvoie la somme de la différence des carrés. Il faut que les deux matrices aient le même nombre de valeurs.
SOMME.X2PY2	Cette fonction utilise deux arguments, deux matrices, dont elle renvoie la somme de la somme des carrés. Il faut que les deux matrices aient le même nombre de valeurs.
SOMME.XMY2	Cette fonction utilise deux arguments, deux matrices, dont elle renvoie la somme des carrés des différences. Il faut que les deux matrices aient le même nombre de valeurs.

Figure 12–37
Mise en œuvre des fonctions SOMME.X2MY2, SOMME.X2PY2 et SOMME.XMY2. Pour information, on a indiqué, tout en bas, les calculs menés par Excel pour renvoyer les résultats.

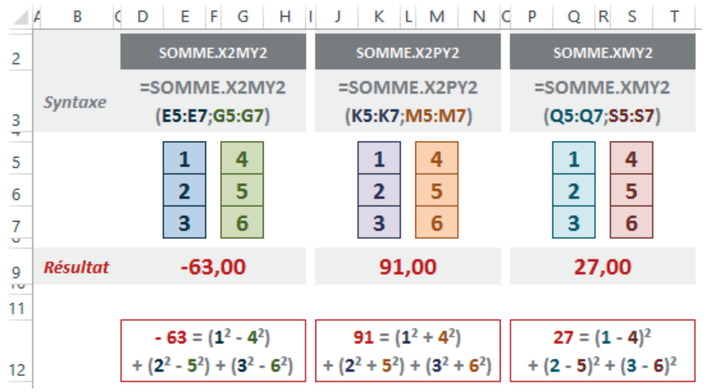
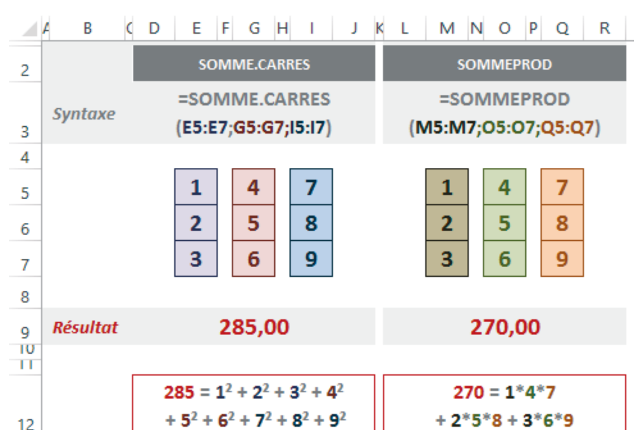


Tableau 12–26 Calculs matriciels

Fonction	Description
SOMME.CARRES	Cette fonction utilise un nombre variable d'arguments, suivant les données à intégrer au calcul. Elle élève au carré toutes les valeurs rencontrées dans les diverses plages et en renvoie la somme.
SOMMEPROD	Cette fonction utilise un nombre variable d'arguments, selon les vecteurs à intégrer au calcul. En fonction du nombre d'arguments, elle fait les produits deux à deux, trois à trois, quatre à quatre… et renvoie la somme de ces produits. Les matrices spécifiées dans les divers arguments doivent donc avoir la même dimension.

Figure 12–38
Mise en œuvre des fonctions SOMME.CARRES et SOMMEPROD. Pour information, on a indiqué, tout en bas, les calculs menés par Excel pour renvoyer les résultats.



Probabilités

L'étude des probabilités a connu de nombreux développements au cours des trois derniers siècles. En travaillant sur le caractère aléatoire et en partie imprévisible de certains phénomènes, les mathématiciens ont développé une théorie qui a eu des implications dans des domaines aussi variés que la météorologie, la finance ou la chimie.

La probabilité (du latin *probabilitas*) est une évaluation du caractère probable d'un événement. La probabilité d'un événement est un nombre réel compris entre `0` et `1` ; plus ce nombre est grand, plus l'événement a de chance de se produire. On dit que deux événements sont indépendants lorsque le fait de connaître le résultat du premier événement ne nous aide pas pour prévoir le second et inversement.

Factorielles

COMPRENDRE Factorielle et factorielle double d'un nombre n

La factorielle d'un entier positif `n`, notée `n!`, est le produit $n \times (n-1) \times (n-2) \times \ldots \times 3 \times 2 \times 1$. Les factorielles sont fréquemment utilisées dans les calculs de probabilités (voir un peu plus loin, les formules de calcul des combinaisons et des arrangements).
La factorielle double d'un nombre `n` (notée `n!!`) égale $n \times (n-2) \times \ldots \times 4 \times 2$ si le nombre est pair et $n \times (n-2) \times \ldots \times 3 \times 1$ si le nombre est impair.

Tableau 12–27 Probabilités

Fonction	Description
FACT	Cette fonction utilise un seul argument, un entier, dont elle renvoie la factorielle. Si un utilisateur distrait indique un nombre décimal, Excel le tronque à sa valeur entière.
FACTDOUBLE	Cette fonction utilise un seul argument, un entier, dont elle renvoie la factorielle double. Si un utilisateur distrait indique un nombre décimal, Excel le tronque à sa valeur entière.

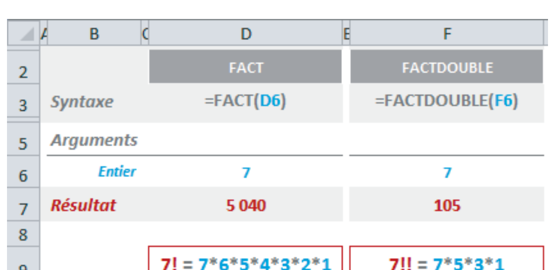


Figure 12–39
Mise en œuvre des fonctions FACT et FACTDOUBLE.
Pour information, on a indiqué, tout en bas, les calculs menés par Excel pour renvoyer les résultats.

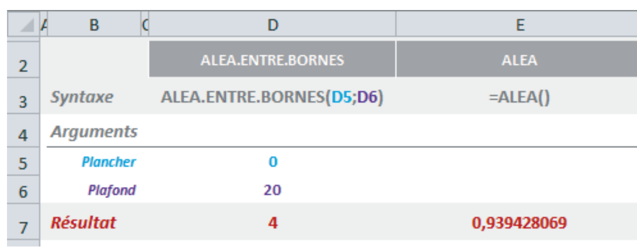
> **ANECDOTE Retrouver la constante e**
>
> La somme de 0 à l'infini des inverses des factorielles donne la constante e :
> `1/0! + 1/1! + 1/2! + 1/3! + … + 1/n! = 2,7182818…`

Valeurs aléatoires

Tableau 12–28 Probabilités

Fonction	Description
ALEA.ENTRE.BORNES	Cette fonction renvoie une valeur entière aléatoire comprise entre les deux nombres entiers (*plancher* et *plafond*), précisés en arguments.
ALEA	Cette fonction renvoie un nombre réel aléatoire compris entre 0 et 1. Elle n'utilise pas d'argument.

Figure 12–40
Mise en œuvre des fonctions ALEA.ENTRE.BORNES et ALEA.



	A	B	C	D	E
2				ALEA.ENTRE.BORNES	ALEA
3		*Syntaxe*		ALEA.ENTRE.BORNES(D5;D6)	=ALEA()
4		*Arguments*			
5		*Plancher*		0	
6		*Plafond*		20	
7		*Résultat*		4	0,939428069

Combinaisons et arrangements

Au même titre que les factorielles, les combinaisons et les arrangements sont des notions de base en probabilités. Supposons que nous choisissions k éléments parmi n, et que nous souhaitions connaître le nombre de possibilités dont nous disposons pour faire ce choix. On démontre aisément que si l'on tient compte de l'ordre dans lequel on a choisi les k éléments, le résultat correspond au nombre d'arrangements (avec ou sans répétitions) de k dans n, dont les deux formules sont données figure 12-41, Si l'on ne tient pas compte de cet ordre, le résultat correspond au nombre de combinaisons (avec ou sans répétitions) dont les deux formules sont également données figure 12-41.

En appliquant les formules de calcul données figure 12-41, on a bien :

- `5! / (2! × (5 - 2)!) = 120 / (2 × 6) = 120 / 12 = 10` combinaisons.
- `(5 + 2 - 1)! / (2! × (5 - 1)!) = 720 / 48 = 15` combinaisons avec répétition.
- `5! / (5 - 2)! = 120 / 6 = 20` arrangements.
- $5^2 = 25$ arrangements avec répétition.

Figure 12–41
Choix de deux éléments parmi cinq (symbolisés par les carrés de couleur). Si l'on tient compte de l'ordre dans lequel on a choisi les deux éléments, on a 20 choix possibles, et même 25 si l'on accepte les répétitions. Si l'on ne tient pas compte de l'ordre, le nombre de possibilités est réduit à 10, ou 15 si l'on accepte les répétitions.

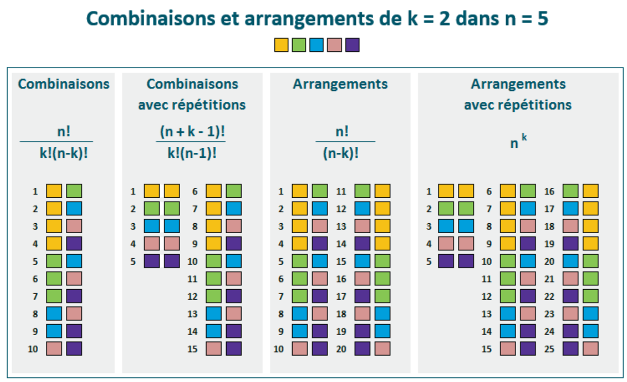


Tableau 12–29 Probabilités

Fonction	Description
COMBIN	Cette fonction utilise deux arguments, deux nombres entiers. Le premier argument représente le nombre total d'éléments et le second est le nombre d'éléments choisis. À partir de ces deux valeurs, la fonction renvoie le nombre de combinaisons.
COMBINA	Cette fonction utilise deux arguments, deux nombres entiers. Le premier argument représente le nombre total d'éléments et le second est le nombre d'éléments choisis. À partir de ces deux valeurs, la fonction renvoie le nombre de combinaisons avec répétition. **Nouveauté Excel 2013**.

Figure 12–42
Calcul du nombre de combinaisons avec les fonctions COMBIN et COMBINA. Les formules entrées dans les deux dernières colonnes permettent de déduire le nombre d'arrangements et le nombre d'arrangements avec répétition.

	A	B	C	D	E	F	G
2				COMBIN	COMBINA	*Arrangements*	*Arrangements avec répétitions*
3		*Syntaxe*		=COMBIN (D5;D6)	=COMBINA (E5;E6)	=COMBIN (F5;F6)*FACT(F6)	=G5^G6
4		***Arguments***					
5		***Nb éléments***		5	5	5	5
6		***Eléments choisis***		2	2	2	2
7		***Résultat***		**10**	**15**	**20**	**25**

Oups **Des fonctions dispersées dans plusieurs catégories**

Excel 2010 et 2013 proposent la fonction *PERMUTATION* pour calculer automatiquement le nombre d'arrangements, mais cette dernière étant stockée dans la catégorie *Statistiques*, elle est traitée dans le chapitre 13. Excel 2013 propose la fonction *PERMUTATIONA* (également stockée dans la catégorie *Statistiques*) qui renvoie le nombre d'arrangements avec répétition (**nouveauté Excel 2013**).

Fonction multinomiale

Une distribution multinomiale est une généralisation de la distribution binomiale à plus de deux catégories (voir le chapitre 13). Elle modélise l'expérience consistant à répéter n fois, indépendamment, une épreuve admettant r issues différentes, de probabilités respectives p_1, p_2 … p_r, telles que :

$P(X_1=n_1,\ X_2=n_2\ \ldots\ X_r=n_r) = [n!\ /\ (n_1!\times n_2!\times \ldots \times n_r!)] \times (p_1^{n_1}\times p_2^{n_2}\times \ldots \times p_r^{n_r})$.

Tableau 12–30 Probabilités

Fonction	Description
MULTINOMIALE	Cette fonction renvoie le rapport de la factorielle de la somme sur le produit des factorielles. Le nombre de valeurs sur lesquelles le calcul est fait n'est pas fixe. Les arguments peuvent désigner des valeurs isolées comme des plages de cellules.

Figure 12–43
Formule de la fonction MULTINOMIALE.

$$\frac{\left(\sum_{i=1}^{n} a_i\right)!}{\prod_{i=1}^{n} a_i!}$$

Sur la figure 12-44, nous avons pris l'exemple de dix jets de dés successifs. Nous cherchons la probabilité pour que le 1 apparaisse trois fois, le 2 et le 3, deux fois, le 4, le 5 et le 6, une fois, sachant que le dé est parfaitement équilibré et, donc, que la probabilité d'obtenir une face donnée lors d'un jet est égale à `1/6`. Si l'on applique la formule donnée au début de cette section, la probabilité cherchée doit être égale à :

$[10!\ /\ (3!\times 2!\times 2!\times 1!\times 1!\times 1!)] \times ((1/6)^3\times(1/6)^2\times(1/6)^2\times(1/6)^1\times(1/6)^1\times(1/6)^1)$

La formule de la fonction *MULTINOMIALE* d'Excel, donnée figure 12-43, donne directement la première partie de notre calcul, c'est-à-dire le résultat du quotient :

`(3+2+2+1+1+1)! / (3!x2!x2!x1!x1!x1!) = 3 628 800 / 24 = 151 200`.

Pour obtenir notre probabilité, il ne reste plus qu'à multiplier cette valeur par $(1/6)^{(3+2+2+1+1+1)}$, c'est-à-dire $(1/6)^{10}$. Cela donne une probabilité de `0,0025`.

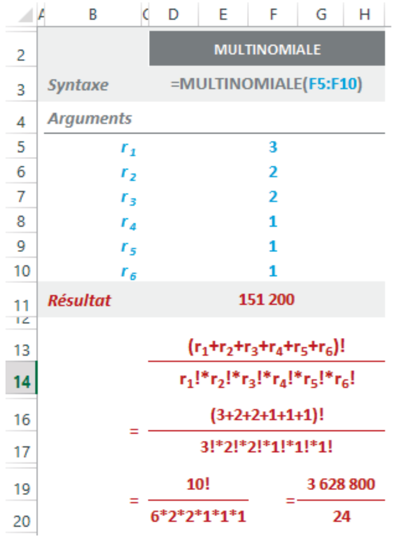


Figure 12–44
Mise en œuvre de la fonction MULTINOMIALE.

Fonctions circulaires

Les fonctions circulaires sont massivement utilisées en mathématiques (trigonométrie, étude des triangles, des cercles, etc.) et en physique pour modéliser des phénomènes périodiques (ondes électromagnétiques, lumière, traitement du signal, etc.).

HISTOIRE Origines de la trigonométrie

La trigonométrie (étymologiquement « mesure des triangles ») a été inventée par les astronomes grecs pour calculer les éléments d'un triangle (ses angles et ses côtés). Elle a conduit à associer à chaque angle des grandeurs appelées rapports trigonométriques, ou fonctions circulaires.

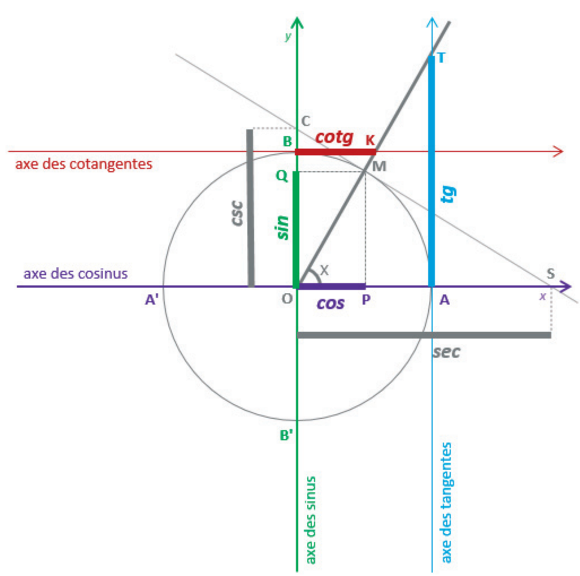
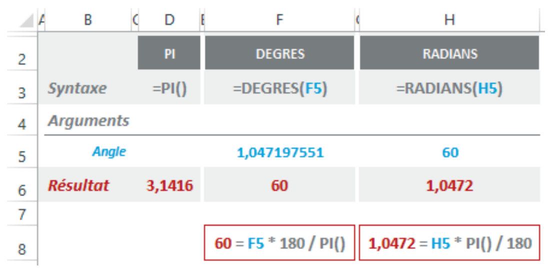


Figure 12–45
On appelle fonctions circulaires de l'arc X les nombres réels qui constituent les mesures algébriques des segments OP, OQ, AT, BK, OS et OC. Ces nombres sont appelés respectivement le cosinus, le sinus, la tangente, la cotangente, la sécante et la cosécante de l'arc X.

Tableau 12–31 Fonctions circulaires

Fonction	Description
PI	Cette fonction ne prend aucun argument. Elle renvoie la valeur `3,14159265358979`, approximation de la constante mathématique π, avec une précision de 15 décimales.
DEGRES	Cette fonction utilise un seul argument : un angle exprimé en radians. Elle renvoie une valeur correspondant à sa conversion en degrés.
RADIANS	Cette fonction utilise un seul argument : un angle exprimé en degrés. Elle renvoie une valeur correspondant à sa conversion en radians.



	B	D	F	H
2		PI	DEGRES	RADIANS
3	*Syntaxe*	=PI()	=DEGRES(F5)	=RADIANS(H5)
4	*Arguments*			
5	*Angle*		1,047197551	60
6	*Résultat*	3,1416	60	1,0472
7				
8			60 = F5 * 180 / PI()	1,0472 = H5 * PI() / 180

Figure 12–46
Mise en œuvre des fonctions PI, DEGRES et RADIANS. Pour information, on a donné en clair, en bas du tableau, les formules qui permettent de passer des radians aux degrés et inversement.

Comprendre Mesure des angles

Un angle, dans le plan, est une partie du plan limitée par deux demi-droites qui ont une origine commune, appelée sommet de l'angle. Les demi-droites constituent les côtés de l'angle.
Il est commode d'associer la mesure d'un angle dont le sommet coïncide avec le centre d'un cercle à celle d'un arc de cercle intercepté par les côtés de cet angle.
L'angle au centre interceptant un arc égal au quart de la circonférence est un angle droit. Celui qui intercepte la moitié de la circonférence est un angle plat.
Le degré est la 90e partie d'un angle droit. Donc, un angle droit égale 90° et un angle plat égale 180°. La division en degrés est commode car le cercle se prête facilement à une division sexagésimale. En construisant six arcs consécutifs dont la corde (en bleu sur la figure 12-47) est égale au rayon, on partage le cercle en six parties égales qui sont six arcs de 60° chacun.

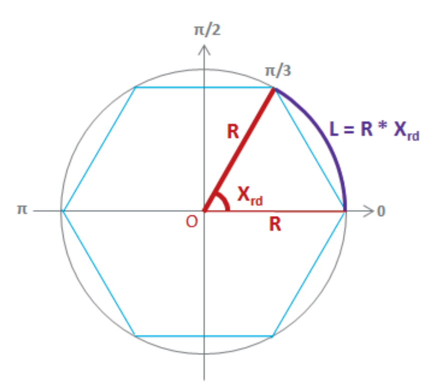


Figure 12–47
Angles remarquables et mesure de la longueur d'un arc à partir d'un angle exprimé en radians.

Dans les problèmes théoriques, on mesure les angles en radians. Étant donné un cercle de rayon R, un arc de longueur L a pour mesure $L = R \times X_{rd}$ (le rayon que multiplie l'angle exprimé en radians). Si l'on choisit de rapporter les angles à un cercle de rayon $R = 1$ mètre, un arc de cercle de longueur 1 mètre sur ce cercle aura donc pour mesure 1 radian. Ainsi, la circonférence (360°) dont la longueur est $2\pi \times 1 = 2\pi$ mètres, mesure 2π radians. On a donc $180^\circ = \pi$ radians, $90^\circ = \pi/2$ radians, $60^\circ = \pi/3$ radians et $30^\circ = \pi/6$ radians.

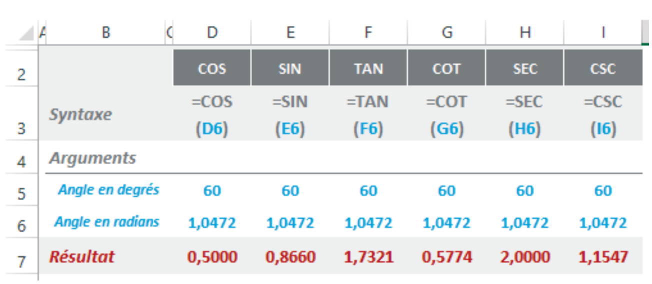
Toutes les fonctions présentées ici nécessitent un angle exprimé en radians. Si vous partez d'un angle mesuré en degrés, multipliez-le par `PI()/180` ou utilisez la fonction *RADIANS* pour le convertir.

Tableau 12–32 Fonctions circulaires

Fonction	Description
COS	Cette fonction utilise un seul argument : un angle exprimé en radians. Elle renvoie son cosinus.
SIN	Cette fonction utilise un seul argument : un angle exprimé en radians. Elle renvoie son sinus.
TAN	Cette fonction utilise un seul argument : un angle exprimé en radians. Elle renvoie sa tangente.
COT	Cette fonction utilise un seul argument : un angle exprimé en radians. Elle renvoie sa cotangente. **Nouveauté Excel 2013**.
SEC	Cette fonction utilise un seul argument : un angle exprimé en radians. Elle renvoie sa sécante. **Nouveauté Excel 2013**.
CSC	Cette fonction utilise un seul argument : un angle exprimé en radians. Elle renvoie sa cosécante. **Nouveauté Excel 2013**.

RAPPEL Relations entre les fonctions

- `tg x = sin x / cos x`.
- `cotg x = cos x / sin x = 1 / tg x`.
- `sec x = 1 / cos x`.
- `csc x = 1 / sin x`.



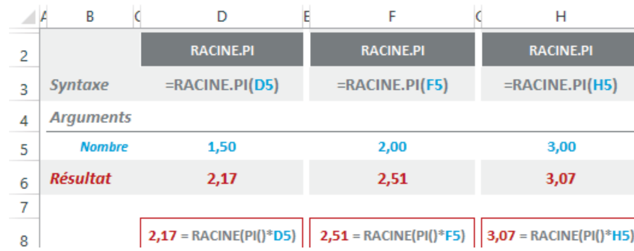
	A	B	C	D	E	F	G	H	I
2				COS	SIN	TAN	COT	SEC	CSC
3		*Syntaxe*		=COS (D6)	=SIN (E6)	=TAN (F6)	=COT (G6)	=SEC (H6)	=CSC (I6)
4		*Arguments*							
5		*Angle en degrés*		60	60	60	60	60	60
6		*Angle en radians*		1,0472	1,0472	1,0472	1,0472	1,0472	1,0472
7		*Résultat*		0,5000	0,8660	1,7321	0,5774	2,0000	1,1547

Figure 12–48
Mise en œuvre des fonctions COS, SIN, TAN, COT, SEC et CSC.

Tableau 12–33 Fonctions circulaires

Fonction	Description
RACINE.PI	Cette fonction renvoie la racine carrée du produit de π par le nombre décimal précisé dans l'argument.

Figure 12–49
Mise en œuvre de la fonction RACINE.PI. Pour information, on a indiqué sous le tableau le calcul effectué par la fonction pour chaque valeur.



	A	B	C	D	E	F	G	H
2				RACINE.PI		RACINE.PI		RACINE.PI
3		*Syntaxe*		=RACINE.PI(D5)		=RACINE.PI(F5)		=RACINE.PI(H5)
4		*Arguments*						
5		*Nombre*		1,50		2,00		3,00
6		*Résultat*		2,17		2,51		3,07
7								
8				2,17 = RACINE(PI()*D5)		2,51 = RACINE(PI()*F5)		3,07 = RACINE(PI()*H5)

Les fonctions trigonométriques ne sont pas bijectives. Par exemple, `0,5` est le cosinus d'un angle de `60°` (`π/3`), mais aussi celui d'un angle de `-60°` (`2π/3`), ou encore d'un angle de `420°` (`7π/3`), et ainsi de suite. De même, les angles `π/3` et `2π/3` ont un sinus identique. En restreignant la fonction `cos` à l'intervalle `[0, π]`, la fonction `sin` à l'intervalle `[-π/2, π/2]`, la fonction `tg` à l'intervalle `]-π/2, π/2[` et la fonction `cotg` à l'intervalle `]0, π[`, elles réalisent des bijections. Les fonctions trigonométriques réciproques partent donc d'un nombre décimal et renvoient l'angle en radians correspondant, à l'intérieur des intervalles précisés ci-dessus.

Figure 12–50
Pour que les fonctions trigonométriques soient bijectives et que l'on puisse considérer leurs réciproques sans ambiguïté, on les définit sur des intervalles restreints ([0, π] pour la fonction cos et [-π/2, π/2] pour la fonction sin).

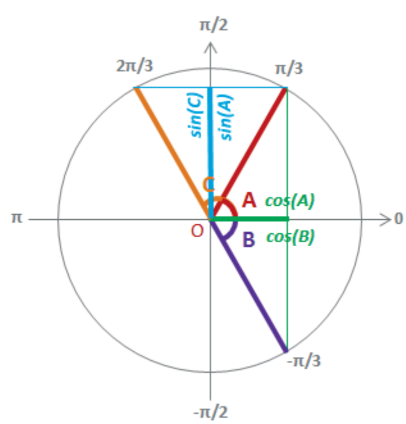
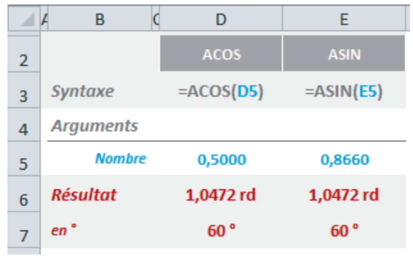


Tableau 12–34 Fonctions circulaires

Fonction	Description
ACOS	Cette fonction renvoie l'arc cosinus du nombre décimal précisé en argument. Le résultat est un angle exprimé en radians, compris entre `0` et `π`. Le cosinus de cet angle correspond à la valeur indiquée dans l'argument.
ASIN	Cette fonction renvoie l'arc sinus du nombre décimal précisé en argument. Le résultat est un angle exprimé en radians, compris entre `-π/2` et `π/2`. Le sinus de cet angle correspond à la valeur indiquée dans l'argument.

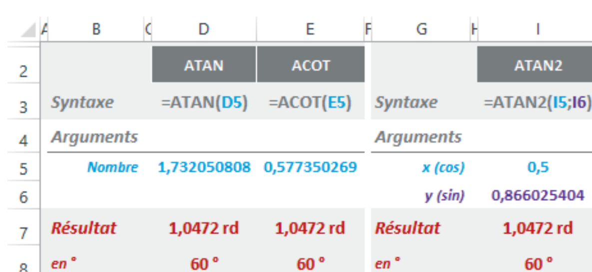


	B	D	E
2		ACOS	ASIN
3	*Syntaxe*	=ACOS(D5)	=ASIN(E5)
4	*Arguments*		
5	*Nombre*	0,5000	0,8660
6	*Résultat*	1,0472 rd	1,0472 rd
7	*en °*	60 °	60 °

Figure 12–51
Mise en œuvre des fonctions ACOS et ASIN. Le résultat est un angle en radians. Dans les cellules D7 et E7, on a entré les formules =DEGRES(D6) et =DEGRES(E6) pour obtenir leur équivalent en degrés.

Tableau 12–35 Fonctions circulaires

Fonction	Description
ATAN	Cette fonction renvoie l'arc tangente du nombre décimal précisé en argument. Le résultat est un angle exprimé en radians, compris entre $-\pi/2$ et $\pi/2$. La tangente de cet angle correspond à la valeur indiquée dans l'argument.
ATAN2	Cette fonction utilise deux arguments : un couple de coordonnées `x` et `y`, `x` représentant le cosinus d'un angle et `y` son sinus. Cette fonction renvoie l'angle en radians (compris entre $-\pi$ et π) passant par ce couple de coordonnées. Dans l'exemple proposé, on a bien `0,5` qui est le cosinus de l'angle `60°` et `0,866` qui est son sinus. La fonction `ATAN2(0,5;0,866)` donne bien comme résultat `1,0472`, qui, converti en degrés, donne `60`.
ACOT	Cette fonction renvoie l'arc cotangente du nombre décimal précisé en argument. Le résultat est un angle exprimé en radians, compris entre `0` et π. La cotangente de cet angle correspond à la valeur indiquée dans l'argument. **Nouveauté Excel 2013**.



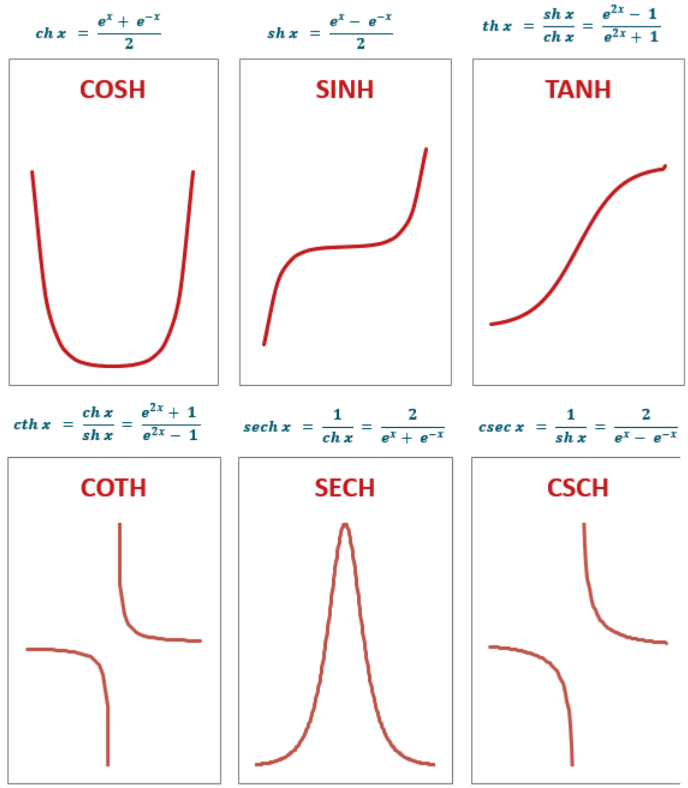
	B	D	E	G	I
2		ATAN	ACOT		ATAN2
3	*Syntaxe*	=ATAN(D5)	=ACOT(E5)	*Syntaxe*	=ATAN2(I5;I6)
4	*Arguments*			*Arguments*	
5	*Nombre*	1,732050808	0,577350269	*x (cos)*	0,5
6				*y (sin)*	0,866025404
7	*Résultat*	1,0472 rd	1,0472 rd	*Résultat*	1,0472 rd
8	*en °*	60 °	60 °	*en °*	60 °

Figure 12–52
Mise en œuvre des fonctions ATAN, ACOT et ATAN2.
Le résultat est un angle en radians. Dans les cellules D8, E8 et I8, on a entré les formules =DEGRES(D7), =DEGRES(E7) et =DEGRES(I7) pour obtenir leur équivalent en degrés.

Fonctions hyperboliques

Ces fonctions sont fréquemment utilisées en mathématiques et en physique. La fonction cosinus hyperbolique, par exemple, intervient dans la définition de la chaînette (forme que prend un câble suspendu à ses extrémités et soumis à son propre poids).

Figure 12–53
Définition et représentation graphique des six fonctions hyperboliques proposées par Excel.



HISTOIRE Les fonctions hyperboliques

Les fonctions hyperboliques ont été inventées par le jésuite Vincenzo Riccati dans les années 1760 alors qu'il cherchait à calculer l'aire sous l'hyperbole d'équation $x^2 - y^2 = 1$. La méthode géométrique qu'il employa alors était très similaire à celle que l'on peut utiliser pour calculer l'aire d'un cercle d'équation $x^2 + y^2 = 1$. Le calcul de l'aire du cercle fait intervenir les fonctions trigonométriques classiques que Riccati nommait cosinus et sinus circulaires. Par analogie, il appela alors les fonctions qu'il venait de créer cosinus et sinus hyperboliques.

Tableau 12–36 Fonctions hyperboliques

Fonction	Description
COSH	Cette fonction utilise un seul argument, un nombre décimal compris entre −709 et 709 (limite d'Excel) ; elle renvoie son cosinus hyperbolique.
SINH	Cette fonction utilise un seul argument, un nombre décimal compris entre −709 et 709 (limite d'Excel) ; elle renvoie son sinus hyperbolique.
TANH	Cette fonction utilise un seul argument, un nombre décimal quelconque ; elle renvoie sa tangente hyperbolique.
COTH	Cette fonction utilise un seul argument, un nombre décimal quelconque différent de 0 ; elle renvoie sa cotangente hyperbolique. **Nouveauté Excel 2013**.
SECH	Cette fonction utilise un seul argument, un nombre décimal quelconque ; elle renvoie sa sécante hyperbolique. **Nouveauté Excel 2013**.
CSCH	Cette fonction utilise un seul argument, un nombre décimal quelconque différent de 0 ; elle renvoie sa cosécante hyperbolique. **Nouveauté Excel 2013**.

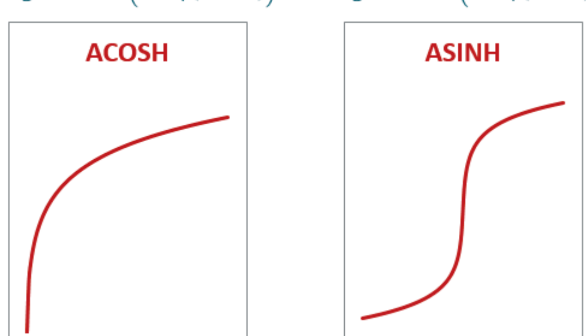
Figure 12–54
Mise en œuvre des fonctions COSH, SINH, TANH, COTH, SECH et CSCH.

	B	D	E	F	G	H	I
2		COSH	SINH	TANH	COTH	SECH	CSCH
3	*Syntaxe*	=COSH (D5)	=SINH (E5)	=TANH (F5)	=COTH (G5)	=SECH (H5)	=CSCH (I5)
4	*Arguments*						
5	*Nombre*	2,00	2,00	3,00	1,86	3,63	3,76
6	*Résultat*	3,76	3,63	1,00	1,05	0,05	0,05

Excel offre les quatre fonctions hyperboliques inverses *ACOSH*, *ASINH*, *ATANH* et *ACOTH*.

Figure 12–55
Définition et représentation graphique des quatre fonctions hyperboliques inverses proposées par Excel.

$$\mathrm{argch}\, x = \ln\left(x + \sqrt{(x^2 - 1)}\right) \qquad \mathrm{argsh}\, x = \ln\left(x + \sqrt{(x^2 + 1)}\right)$$



$$\mathit{argth}\, x = \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) \qquad \mathit{argcoth}\, x = \frac{1}{2} \ln\left(\frac{x+1}{x-1}\right)$$

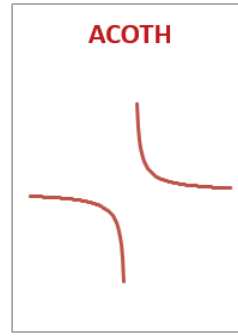


Tableau 12–37 Fonctions hyperboliques inverses

Fonction	Description
ACOSH	Cette fonction utilise un seul argument, un nombre décimal supérieur ou égal à 1 ; elle renvoie son cosinus hyperbolique inverse.
ASINH	Cette fonction utilise un seul argument, un nombre décimal quelconque ; elle renvoie son sinus hyperbolique inverse.
ATANH	Cette fonction utilise un seul argument, un nombre décimal compris strictement entre -1 et 1 ; elle renvoie sa tangente hyperbolique inverse.
ACOTH	Cette fonction utilise un seul argument, un nombre décimal strictement inférieur à -1 ou strictement supérieur à 1 ; elle renvoie sa cotangente hyperbolique inverse. **Nouveauté Excel 2013**.

Figure 12–56
Mise en œuvre des fonctions ACOSH, ASINH, ATANH et ACOTH.

	A	B	C	D	E	F	G
2				ACOSH	ASINH	ATANH	ACOTH
3		*Syntaxe*		=ACOSH (D5)	=ASINH (E5)	=ATANH (F5)	=ACOTH (G5)
4		*Arguments*					
5		*Nombre*		3,76	3,63	1,00	1,05
6		*Résultat*		2,00	2,00	3,00	1,86

Deux exemples d'utilisation des fonctions mathématiques

Pour mieux comprendre comment mettre en œuvre les fonctions mathématiques, voici deux exemples d'application. Le premier met à contribution le calcul matriciel pour résoudre facilement un système de quatre équations à quatre inconnues et le second crée un développement limité de la fonction e^x au voisinage de zéro.

Résolution d'un système de 4 équations à 4 inconnues

À l'aide des matrices, on peut résoudre facilement les systèmes de n équations à n inconnues. Par exemple, le système de quatre équations à quatre inconnues présenté dans le coin supérieur gauche de la figure 12-57 peut être considéré comme le produit des deux matrices A et X dont le résultat donne la matrice B (les parties droites des équations). On a donc l'égalité $A \times X = B$.

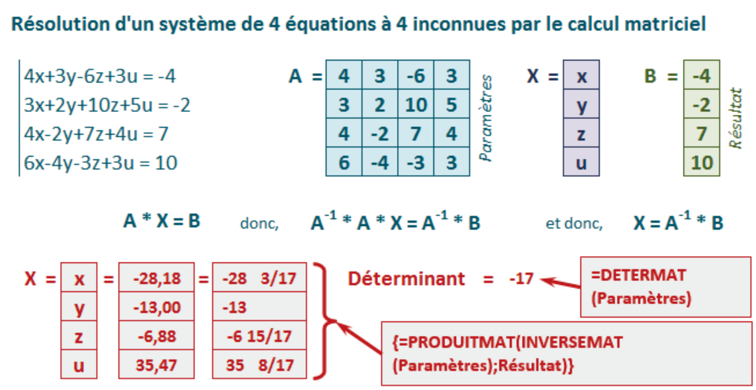


Figure 12–57
Le résultat de ce système de quatre équations à quatre inconnues est présenté sous sa forme décimale et sous sa forme fractionnaire.

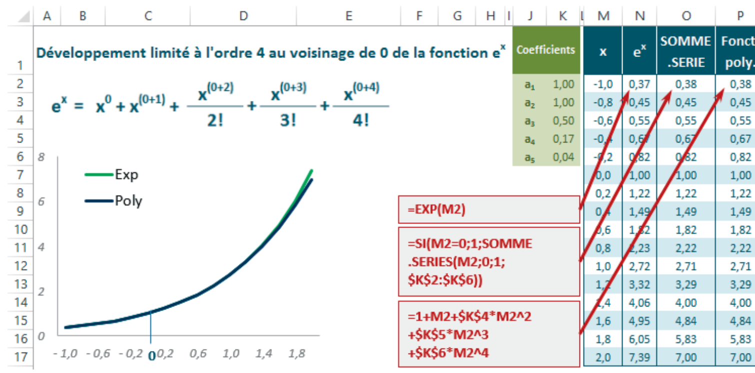
Si l'on note A^{-1} la matrice inverse de A, on peut déduire que $A^{-1} \times A \times X = A^{-1} \times B$ et, $A^{-1} \times A$ étant égal à la matrice identité, on a finalement $X = A^{-1} \times B$.

Le calcul de l'inverse d'une matrice n'étant possible que lorsque son déterminant est non nul, on a utilisé *DETERMAT* pour trouver sa valeur, 17. Le calcul est donc possible… et immédiat, puisqu'Excel fournit tout ce dont on a besoin. La formule utilisée pour résoudre ce problème est présentée dans un cartouche rouge sur fond gris, en bas de la figure. Pour entrer cette formule, il faut sélectionner quatre cellules contiguës dans une même colonne, entrer la formule et, surtout, ne pas oublier de la valider en pressant simultanément les touches *Ctrl+Maj+Entrée*.

Développement limité

La syntaxe du développement limité de la fonction e^x au voisinage de zéro est donnée dans le coin supérieur gauche de la figure 12-58.

Figure 12–58
La représentation graphique de la fonction e^x et de son approximation polynomiale au voisinage de zéro montre une très grande similarité entre les deux courbes.



Dans le tableau bleu situé à droite de la figure, on trouve tout d'abord la colonne x contenant les valeurs pour lesquelles le calcul doit être fait. La colonne suivante utilise *EXP* pour calculer les valeurs de la fonction e^x. Dans celle d'après, on trouve la fonction *SOMME.SERIE* avec les paramètres 0 et 1, ainsi que les coefficients de la plage *K2:K6* pour calculer l'approximation polynomiale. Enfin, la dernière colonne sert simplement à vérifier qu'en construisant nous-mêmes la fonction polynomiale, on obtient bien les mêmes résultats qu'avec *SOMME.SERIE*.

Du côté des statisticiens 13

Même si vous n'êtes pas statisticien, vous manipulez en permanence des concepts statistiques : salaire moyen, taux de rendement, espérance de vie et autres litanies récurrentes qu'égrènent vos flashs d'information quotidiens.

SOMMAIRE

- Dénombrement
- Tendance centrale
- Dispersion
- Corrélation
- Régression
- Distributions théoriques
- Tests statistiques
- Intervalles de confiance

MOTS-CLÉS

- Asymétrie
- Centile
- Corrélation
- Dénombrement
- Écart-type
- Fréquence
- Intervalle de confiance
- Khi-deux
- Kurtosis
- Loi binomiale
- Loi de Fisher
- Loi de Student
- Loi Gamma
- Loi normale
- Médiane
- Mode
- Moyenne
- Population
- Probabilité
- Quartile
- Régression
- Variable
- Variance

Parmi les 460 fonctions de calcul, la catégorie *Statistiques* est la plus volumineuse (à elle seule, elle en réunit plus d'une centaine !)… sans compter les 40 fonctions statistiques supplémentaires (en doublon des premières) qui ne sont là que pour assurer la compatibilité avec les versions antérieures d'Excel.

Dans les deux premières sections de ce chapitre, nous en présenterons la moitié. Leur rôle principal est la mise à disposition de toute une série d'indices (moyenne, écart-type, etc.) dont l'objet est de mieux appréhender la structure des données étudiées.

Dans la troisième section, nous exposerons l'autre moitié. À cette occasion, nous aborderons une douzaine de distributions théoriques. Basées sur des modèles mathématiques sophistiqués, ces outils statistiques s'adressent clairement à des utilisateurs avertis.

Tendance centrale et dispersion

Pour aborder les premières fonctions de ce chapitre, nous allons utiliser un tableau réunissant les notes obtenues par 25 élèves dans cinq matières sur trois trimestres (figure 13-1). En termes statistiques, on parle d'une population de 25 individus (n) pour lesquels on dispose de 15 variables quantitatives.

	A	B	D	E	F	H	I	J	L	M	N	P	Q	R	T	U	V
1			Maths			Physique			Français			Histoire			Langues		
2	Nom	Prénom	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
3	ATIER	Modeste	7	5	6	6	4	7	14	13	15	15	14	14	14	12	13
4	BAYON	Basile	19	17	18	17	15	16	12	13	12	11	10	11	13	12	10
5	BEDAS	Genevieve	20	18	17	19	17	19	16	13	14	15	15	14	14	16	13
6	BERTHIER	Edouard	11	10	11	10	9	10	11	11	11	10	11	12	11	9	11
7	CAFU	Melaine	19	18	18	17	17	17	Abs	13	14	15	15	12	16	Abs	13
8	CANTREL	Raymond	12	11	11	11	9	10	13	12	12	11	9	9	10	10	10
9	DUPONT	Nina	14	11	15	12	10	14	14	11	14	15	12	15	14	13	16
10	DURAND	Rémi	11	12	14	8	8	12	6	4	9	7	6	8	9	7	7
11	FATEU	Marcel	7	8	9	6	9	10	5	4	11	8	7	9	9	9	7
12	FROUIN	Roseline	2	3	6	Abs	4	7	13	15	14	14	11	12	10	11	9
13	GRENADA	Marius	11	11	12	13	8	14	12	7	13	10	9	12	9	9	11
14	LANGOT	Angèle	11	12	14	6	8	13	5	7	9	7	3	9	9	4	11
15	LAPEYRE	Martine	7	6	9	8	5	7	15	12	15	16	12	16	14	12	14
16	LEGRAND	Véronique	3	2	8	5	4	9	9	6	11	8	7	9	7	4	10
17	LIERT	Gaston	14	*Abs*	17	Abs	*19*	Abs	18	*16*	14	17	*11*	15	10	*11*	16
18	MEILLAC	Bernadette		4	5	3	5	8	6	5	7	7	4	9	6	6	9
19	MILLET	Aimée	13	12	14	13	11	16	9	6	9	8	7	8	7	4	11
20	OPERT	Casimir	6	4	9	7	3	8	8	7	9	6	6	8	7	5	7
21	PELLISSON	Justine	*Abs*	3	6	*2*	2	7	*11*	10	11	*12*	9	13	*13*	10	14
22	PILON	Patrice	5	17	16	17	16	17	17	15	16	16	14	15	15	13	14
23	RUNIN	Diane	20	10	19	11	Abs	17	Abs	10	11	12	11	Abs	14	11	10
24	TALAMON	Sylvain	20	18	18	15	17	19	3	2	4	4	2	12	5	2	13
25	TEROIN	Pascal	18	14	19	12	12	18	17	11	14	15	12	14	16	13	12
26	TERTI	Didier	14	11	18	13	12	17	17	14	18	16	13	13	15	12	12
27	VALADE	Florent	18	14	18	18	15	19	4	3	8	6	2	7	7	4	8

Figure 13–1
Ce tableau réunit les notes trimestrielles de 25 élèves dans cinq matières. La chaîne de caractères « Abs » indique l'absence de note dans la matière pour un trimestre donné.

> **VOCABULAIRE Variables, individus et population**
>
> Les 25 élèves représentent les « individus » de la population étudiée. Il se trouve qu'employé dans ce contexte, le terme est bien adapté. Pourtant, en matière statistique, un individu ne désigne pas nécessairement un être humain. Il peut tout aussi bien désigner un livre, une entreprise ou un objet quelconque. Dans ce cas, il est bien évident que la « population » qui désigne l'ensemble des individus étudiés ne représente absolument pas un rassemblement d'êtres humains.
> Un « caractère » est une propriété mesurée selon le même procédé pour tous les individus d'une population (dans notre exemple, on étudie 15 caractères). On appelle « modalité » une valeur prise par un caractère (ici, comme il s'agit d'une notation sur 20, chaque caractère offre 21 modalités possibles, 21 nombres entiers compris entre 0 et 20). Un caractère peut être qualitatif ou quantitatif. Par convention, on appelle variable statistique quantitative tout caractère quantitatif.
> Une variable quantitative est discrète si les valeurs qu'elle peut prendre sont distinctes les unes des autres. Les 15 variables étudiées ici sont des variables discrètes, puisque leurs valeurs sont l'un des 21 entiers compris dans l'intervalle [0-20].
> Une variable quantitative est continue lorsqu'elle peut prendre n'importe quelle valeur contenue dans un intervalle. Nous travaillerons un peu plus loin avec la moyenne annuelle par matière qui est une variable continue, puisque ces moyennes peuvent être n'importe quel nombre décimal compris dans l'intervalle [0-20].

Compter les individus

La première phase de l'étude d'une population est d'en dénombrer les individus. Les trois fonctions *NBVAL*, *NB* et *NB.VIDE* renvoient respectivement le nombre total des individus (toutes les cellules non vides), le nombre de valeurs numériques et le nombre de cellules vides.

	B	D	E	F	H	I	J	L	M	N	P	Q	R	T	U	V
1		Maths			Physique			Français			Histoire			Langues		
2		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	=NBVAL(Notes!D3:D27)	24	25	25	25	25	25	25	25	25	25	25	25	25	25	25
5	=NB(Notes!D3:D27)	23	24	25	23	24	24	23	25	25	25	25	24	25	24	25
6	=NB.VIDE(Notes!D3:D27)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 13–2
Le tableau présenté à la figure 13-1 se trouve sur une feuille nommée « Notes ». Ses données sont utilisées par les trois fonctions de dénombrement NBVAL, NB et NB.VIDE. La colonne B affiche la syntaxe des formules entrées dans la plage D4:D6.

Entrées d'abord en colonne *D*, les formules ont été ensuite recopiées dans le reste du tableau. On remarque que la fonction *NB* renvoie régulièrement des valeurs inférieures à celles de la fonction *NBVAL*, puisqu'elle ne prend pas en compte les valeurs de texte `Abs`. La seule colonne pour laquelle la fonction *NB.VIDE* renvoie une valeur différente de zéro est la colonne *D* puisqu'elle correspond à la seule plage contenant une cellule vide.

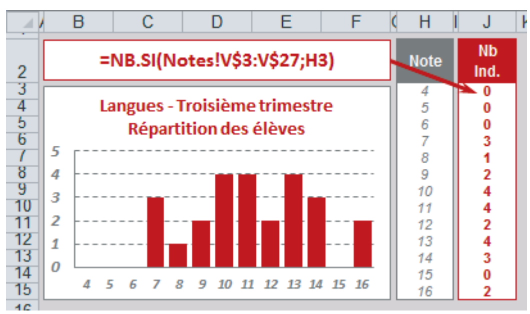
Un graphique pour représenter la répartition des élèves

À l'issue de ce dénombrement global, on cherche à savoir combien d'élèves ont obtenu `0`, `1`, `2`, … , `20` dans une matière donnée. En termes statistiques, cela revient à connaître le nombre d'individus pour chaque modalité x_i. Pour y parvenir, utilisez la fonction *NB.SI*. La statistique descriptive utilise ensuite le résultat de ces calculs pour en faire une représentation graphique qui donne une première idée de la structure de la population étudiée.

LISIBILITÉ Choisir l'intervalle

Afin de ne pas surcharger la figure, le calcul a été fait sur l'intervalle `[4,16]` (au lieu de `[0,20]`). En effet, avant `4` et après `16`, tous les cumuls sont nuls. Dans un souci de symétrie de la représentation, on a conservé les notes 4 à 7 (pour lesquelles les cumuls sont pourtant nuls).

Figure 13–3
La fonction NB.SI, entrée en J3 et dont la syntaxe est présentée dans le cadre rouge, permet de calculer le nombre d'élèves ayant obtenu 4.



Les résultats affichés en colonne *J* correspondent à la répartition des notes du troisième trimestre pour les langues. À nouveau, on a utilisé les données de la feuille `Notes` dont le contenu est présenté figure 13-1. La formule entrée en *J3* a ensuite été recopiée dans la plage *J4:J15*. La somme des valeurs de la plage *J3:J15* redonne bien `25`, nombre total d'individus dans la population étudiée.

COMPRENDRE Fréquences

On peut calculer la fréquence de chaque modalité : $f_i = n_i / n$. On dit alors que la distribution de la variable X est l'ensemble des couples $\{(x_1,f_1), (x_2,f_2), \ldots ,(x_I,f_I)\}$. On a donc $n_1 + n_2 + \ldots + n_I = n$ et $f_1 + f_2 + \ldots + f_I = 1$.

ALLER PLUS LOIN La fonction NB.SI.ENS

La fonction *NB.SI* permet d'exprimer un critère unique. Pour faire des dénombrements impliquant davantage de critères, utilisez la fonction *NB.SI.ENS*. Ses mécanismes sont les mêmes que ceux de la fonction *SOMME.SI.ENS* (consultez le chapitre 12).

Figure 13–4
Pour exprimer des critères à la fois sur les notes du premier et du troisième trimestre, utilisez la fonction NB.SI.ENS.

	B	D	F	H
2	Critères		Nb.	Langues - Trimestres 1 et 3
3	1	3	Individus	
4	<10	<11	6	=NB.SI.ENS(Notes!T$3:T$27;B4;Notes!$V$3:$V$27;D4)
5	>=10	<11	4	=NB.SI.ENS(Notes!T$3:T$27;B5;Notes!$V$3:$V$27;D5)
6	<10	>=11	4	=NB.SI.ENS(Notes!T$3:T$27;B6;Notes!$V$3:$V$27;D6)
7	>=10	>=11	11	=NB.SI.ENS(Notes!T$3:T$27;B7;Notes!$V$3:$V$27;D7)

Les arguments de *NB.SI.ENS* fonctionnent par paires. À titre d'exemple, les deux premiers arguments de la formule entrée en *F4* (`Notes!T$3:T$27` et `B4`) indiquent qu'il faut appliquer le critère `B4` (`<10`) à la plage `Notes!T$3:T$27` (langues au premier trimestre). Les deux derniers arguments (`Notes!$V$3:$V$27` et `D4`) indiquent qu'il faut appliquer le critère `D4` (`<11`) à la plage `Notes!$V$3:$V$27` (langues au troisième trimestre). En d'autres termes, on dénombre (pour les langues) les élèves ayant obtenu à la fois une note strictement inférieure à `10` au premier trimestre et strictement inférieure à `11` au troisième trimestre. Excel trouve 6 individus remplissant à la fois ces deux critères. La formule entrée en *F4*, dont la syntaxe apparaît en *H4*, a ensuite été recopiée dans la plage *F5:F7* (chacune utilisant les critères entrés dans sa propre ligne).

Regrouper en classes

Si la distribution étudiée concerne peu d'individus, il est possible qu'une représentation graphique par modalité n'affiche que des valeurs de type `1/n` (un individu par modalité). Or, en se plaçant toujours dans la perspective d'affiner peu à peu notre compréhension de la structure de la population, une telle représentation n'est pas très utile.

Pour obtenir un graphique plus intéressant, il est souvent préférable de regrouper les modalités et de définir des classes. La fonction *FREQUENCE* d'Excel répond à ce besoin. À partir d'un vecteur de seuils (`8` et `12` dans notre exemple), la fonction *FREQUENCE* renvoie un dénombrement des individus pour les trois classes `[0,8]`, `]8,12]` et `]12,20]`.

Figure 13–5
Pour chaque matière, la fonction FREQUENCE montre la répartition des individus en fonction des deux seuils 8 et 12.

	B	D	E	F	H	I	J	L	M	N	P	Q	R	T	U	V	X
1		Maths			Physique			Français			Histoire			Langues			
2		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
4	{=FREQUENCE (Notes!D3:D27; $X4:$X5)}	7	8	5	9	10	6	7	10	3	9	9	4	6	8	4	8
5		5	9	6	5	7	5	6	7	11	6	11	11	8	12	12	12
6		11	7	14	9	7	13	10	8	11	10	5	9	11	4	9	

À la figure 13-5, les deux seuils 8 et 12 ont été entrés en *X4* et *X5*. Ces deux seuils définissent trois classes : `[0-8]`, `]8-12]` et `]12-20]`, ou encore (les notes sont ici des valeurs entières), les trois classes : `[0-8]`, `[9-12]` et `[13-20]`.

> **ATTENTION FREQUENCE est une fonction matricielle**
>
> La fonction *FREQUENCE* renvoie plusieurs valeurs simultanément. Il s'agit donc d'une fonction matricielle qu'il convient de traiter comme telle (voir le chapitre 4). Pour obtenir la répartition des notes de mathématiques au premier trimestre en fonction des deux seuils 8 et 12, il faut sélectionner la plage *D4:D6*, entrer la fonction *FREQUENCE* (sa syntaxe est indiquée en *B4*) et la valider en pressant simultanément les touches *Ctrl+Maj+Entrée*. La plage *D4:D6* est ensuite recopiée dans les colonnes suivantes.
> Attention, les accolades qui apparaissent dans la syntaxe de la formule entrée en *D4:D6* ne doivent pas être saisies. Excel les ajoutera automatiquement quand vous validerez la formule à l'aide des touches *Ctrl+Maj+Entrée*.

Dans notre exemple, les trois classes ne sont pas égales (les deux classes extrêmes sont deux fois plus grandes que la classe centrale). Pour respecter le principe de proportionnalité (figure 13-6), il faut soit vous ramener à des classes égales (graphique de droite), soit faire en sorte que votre graphique ressemble à la représentation de gauche.

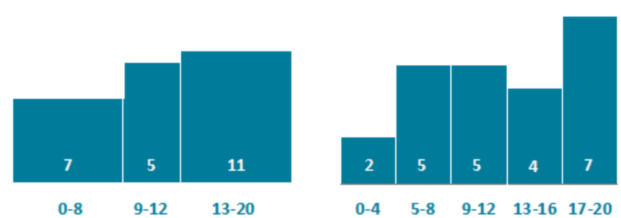


Figure 13–6
Deux manières de représenter la répartition par classes des individus pour la variable « Mathématiques au premier trimestre ».

Indicateurs de tendance centrale

Toujours dans la perspective de mieux comprendre les quinze variables quantitatives de notre exemple, nous nous intéressons maintenant à leur moyenne. Cet indice donne une sorte de résumé de chacune des variables et constitue l'une des mesures de tendance centrale couramment utilisées en statistiques. Une moyenne offre un résumé d'informations nombreuses, mais entraîne du même coup une inévitable perte d'information.

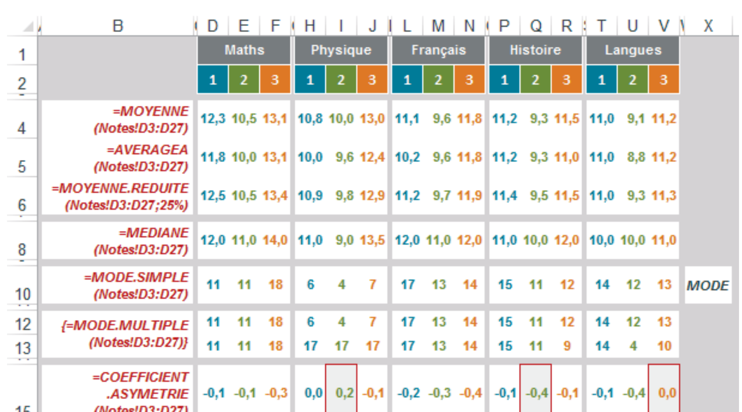
Par la suite, nous en aborderons d'autres comme la médiane ou les modes, qui compléteront la description ébauchée avec le calcul de la moyenne.

Excel fournit trois fonctions relatives au calcul de la moyenne arithmétique : *MOYENNE*, *AVERAGEA* et *MOYENNE.REDUITE*.

> **ATTENTION Coefficient d'asymétrie**
>
> Le coefficient d'asymétrie présenté en ligne *15* de la figure 13-7 a été placé ici car il témoigne de la position respective de la moyenne et de la médiane, mais la fonction qui lui correspond sera traitée un peu plus loin dans ce chapitre. D'ailleurs, considérant que l'on travaille avec la population totale, il vaudrait mieux utiliser la fonction *COEFFICIENT.ASYMETRIE.P*, disponible uniquement sous Excel 2013.

Tableau 13–1 Moyennes

Fonction	Description
MOYENNE	Cette fonction renvoie la moyenne arithmétique de tous ses arguments. En d'autres termes, elle fait la somme de toutes les valeurs numériques qu'elle divise par leur dénombrement. Elle ne prend en compte ni les valeurs logiques, ni les valeurs de texte, ni les cellules vides (voir la ligne *4* de la figure 13-7). Les moyennes des quinze variables quantitatives s'étendent de `9,1` (langues au deuxième trimestre) à `13,1` (mathématiques au troisième trimestre).
AVERAGEA	Cette fonction fait le même calcul que *MOYENNE*, mais en prenant en compte les valeurs logiques (`1` pour `VRAI` et `0` pour `FAUX`), ainsi que les valeurs de texte, qu'elle remplace par `0` (voir la ligne *5* de la figure 13-7). La présence des chaînes de caractères `Abs` dans certaines variables rend quelques résultats légèrement plus faibles qu'avec la fonction *MOYENNE*.
MOYENNE .REDUITE	Cette fonction calcule la moyenne d'un ensemble de données en excluant un certain pourcentage des valeurs extrêmes (inférieures et supérieures). À la figure 13-7, l'exemple donné ligne *6* exclut 25 % des notes aux deux extrémités, c'est-à-dire `0,25 × 25 = 6,25` qui, arrondi au multiple de *2* inférieur, donne *6*. Il exclut donc les trois notes les plus faibles, ainsi que les trois plus fortes, sans tenir compte ni des valeurs logiques, ni des valeurs de texte.



	B	D	E	F	H	I	J	L	M	N	P	Q	R	T	U	V	X
1		Maths			Physique			Français			Histoire			Langues			
2		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
4	=MOYENNE (Notes!D3:D27)	12,3	10,5	13,1	10,8	10,0	13,0	11,1	9,6	11,8	11,2	9,3	11,5	11,0	9,1	11,2	
5	=AVERAGEA (Notes!D3:D27)	11,8	10,0	13,1	10,0	9,6	12,4	10,2	9,6	11,8	11,2	9,3	11,0	11,0	8,8	11,2	
6	=MOYENNE.REDUITE (Notes!D3:D27;25%)	12,5	10,5	13,4	10,9	9,8	12,9	11,2	9,7	11,9	11,4	9,5	11,5	11,0	9,3	11,3	
8	=MEDIANE (Notes!D3:D27)	12,0	11,0	14,0	11,0	9,0	13,5	12,0	11,0	12,0	11,0	10,0	12,0	10,0	10,0	11,0	
10	=MODE.SIMPLE (Notes!D3:D27)	11	11	18	6	4	7	17	13	14	15	11	12	14	12	13	MODE
12	{=MODE.MULTIPLE (Notes!D3:D27)}	11	11	18	6	4	7	17	13	14	15	11	12	14	12	13	
13		11	11	18	17	17	17	17	13	14	15	11	9	14	4	10	
15	=COEFFICIENT .ASYMETRIE (Notes!D3:D27)	-0,1	-0,1	-0,3	0,0	0,2	-0,1	-0,2	-0,3	-0,4	-0,1	-0,4	-0,1	-0,1	-0,4	0,0	

Figure 13–7
Ce tableau réunit les principaux indicateurs de tendance centrale. La colonne X donne l'ancienne forme de certaines fonctions, conservées dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

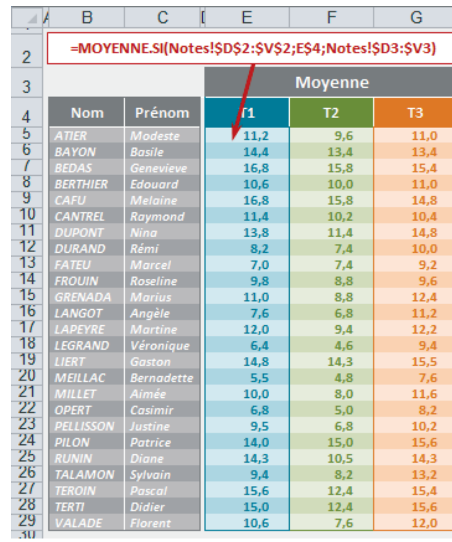
La colonne *B* expose la syntaxe des formules entrées dans la plage *D4:D15*. Ces dernières ont ensuite été recopiées dans la plage *E4:V15*.

Bon à savoir **Moyennes conditionnelles**

Tout comme les sommes conditionnelles étudiées au chapitre 12 et les dénombrements conditionnels abordés au tout début de ce chapitre, *MOYENNE.SI* et *MOYENNE.SI.ENS* calculent des moyennes sur une partie de la population filtrée par un critère. Le tableau présenté figure 13-8 affiche la moyenne trimestrielle de chaque élève, toutes matières confondues.

La formule dont la syntaxe apparaît dans le cadre rouge a été entrée en *E5*, puis recopiée en *E5:G29*. Elle utilise la plage *D2:V2* de la feuille `Notes` (présentée figure 13-1) comme plage de critères et lui applique le critère entré cellule *E4* (deuxième argument). Comme cette dernière contient la chaîne de caractères `T1`, elle fait la moyenne de la plage *D3:V3* (troisième argument, notes de l'élève `ATIER`), mais en ne prenant en compte que celles du premier trimestre (celles dont la position en *D3:V3* correspond à la position des `T1` dans la plage *D2:V2*). Reportez-vous à la description de la fonction *SOMME.SI* faite dans le chapitre 12 pour bien comprendre le mécanisme d'expression des critères.

Si le filtre de calcul doit porter simultanément sur plusieurs critères, il faut utiliser *MOYENNE.SI.ENS*. L'articulation de ses arguments est exactement la même qu'avec *SOMME.SI.ENS*.



=MOYENNE.SI(Notes!$D$2:$V$2;E$4;Notes!$D3:$V3)

			Moyenne		
	Nom	Prénom	T1	T2	T3
5	ATIER	Modeste	11,2	9,6	11,0
6	BAYON	Basile	14,4	13,4	13,4
7	BEDAS	Genevieve	16,8	15,8	15,4
8	BERTHIER	Edouard	10,6	10,0	11,0
9	CAFU	Melaine	16,8	15,8	14,8
10	CANTREL	Raymond	11,4	10,2	10,4
11	DUPONT	Nina	13,8	11,4	14,8
12	DURAND	Rémi	8,2	7,4	10,0
13	FATEU	Marcel	7,0	7,4	9,2
14	FROUIN	Roseline	9,8	8,8	9,6
15	GRENADA	Marius	11,0	8,8	12,4
16	LANGOT	Angèle	7,6	6,8	11,2
17	LAPEYRE	Martine	12,0	9,4	12,2
18	LEGRAND	Véronique	6,4	4,6	9,4
19	LIERT	Gaston	14,8	14,3	15,5
20	MEILLAC	Bernadette	5,5	4,8	7,6
21	MILLET	Aimée	10,0	8,0	11,6
22	OPERT	Casimir	6,8	5,0	8,2
23	PELLISSON	Justine	9,5	6,8	10,2
24	PILON	Patrice	14,0	15,0	15,6
25	RUNIN	Diane	14,3	10,5	14,3
26	TALAMON	Sylvain	9,4	8,2	13,2
27	TEROIN	Pascal	15,6	12,4	15,4
28	TERTI	Didier	15,0	12,4	15,6
29	VALADE	Florent	10,6	7,6	12,0

Figure 13–8 Bien que les données du calcul soient stockées dans des cellules non contiguës, on obtient très facilement la moyenne trimestrielle de chaque élève grâce à la fonction MOYENNE.SI.

Lisser **Moyennes mobiles**

Les moyennes mobiles permettent de lisser des valeurs et, donc, de mettre en évidence une tendance en gommant les accidents de parcours. Vous pouvez utiliser la fonction *MOYENNE* sur quelques valeurs pour calculer une moyenne arithmétique, mais Excel met également à disposition les fonctions *MOYENNE.GEOMETRIQUE* et *MOYENNE HARMONIQUE* qui fondent leur calcul sur d'autres algorithmes.

À-la figure 13-9, la syntaxe de la fonction *MOYENNE.HARMONIQUE* entrée en *D7* apparaît dans le premier cadre orange. Elle réalise un lissage sur les cinq valeurs précédentes. La cellule *E7* contient une formule qui renvoie le même résultat. Cette dernière met en évidence l'algorithme sur lequel repose la fonction *MOYENNE.HARMONIQUE* (rapport du nombre de valeurs dont on fait la moyenne sur la somme des inverses de ces valeurs). Ces deux formules ont été respectivement recopiées dans les plages *D8:D32* et *E8:E32*.

La syntaxe de la fonction *MOYENNE.GEOMETRIQUE* entrée en *F12* apparaît dans le premier cadre vert. Elle réalise un lissage sur les dix valeurs précédentes. La cellule *G12* contient une formule qui renvoie le même résultat. Cette dernière met en évidence l'algorithme sur lequel repose la fonction *MOYENNE.GEOMETRIQUE* (racine énième du produit des `n` valeurs dont on fait la moyenne). Ces deux formules ont été respectivement recopiées dans les plages *F13:F32* et *G13:G32*. Calculée sur davantage de valeurs, cette fonction effectue un lissage plus important que la formule entrée en colonne *D*.

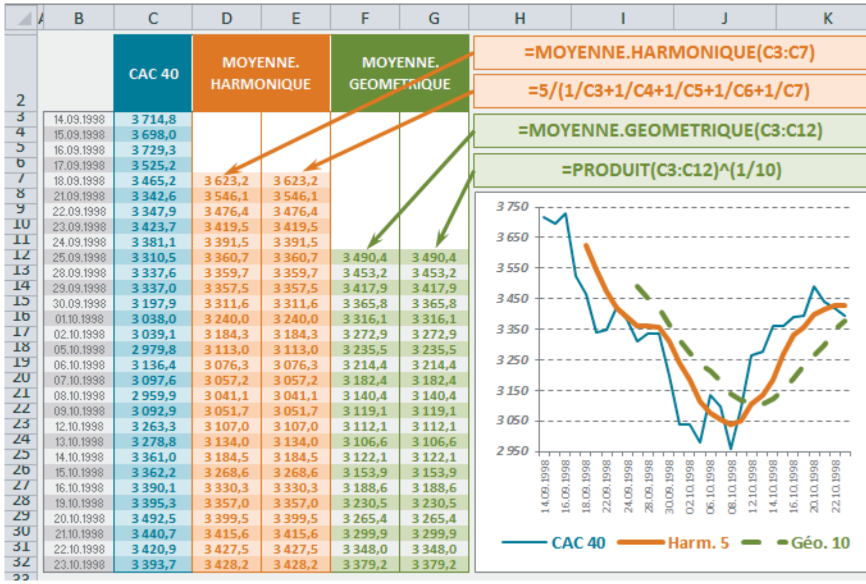


Figure 13–9 Les cours du CAC 40 sont lissés par l'utilisation des fonctions MOYENNE.GEOMETRIQUE et MOYENNE.HARMONIQUE en tant que moyennes mobiles.

Tableau 13–2 Valeurs centrales et répétitives

Fonction	Description
MEDIANE	Cette fonction renvoie la valeur centrale d'un ensemble de nombres. En d'autres termes, elle ordonne toutes les valeurs de la variable et choisit celle qui se trouve exactement au milieu. Si la variable contient un nombre pair de valeurs, la médiane correspond à la moyenne des deux valeurs centrales (voir la ligne 8 de la figure 13-7).
MODE.SIMPLE	À partir d'un ensemble de valeurs, *MODE.SIMPLE* renvoie celle qui apparaît le plus fréquemment. Si chaque valeur est unique, elle renvoie `#N/A`. Si plusieurs arrivent ex-æquo dans ce palmarès, la fonction renvoie la première valeur rencontrée dans la plage (voir la ligne *10* de la figure 13-7). Elle ne tient pas compte des valeurs logiques ou des valeurs de texte.
MODE.MULTIPLE	Cette fonction remplit le même office que la fonction *MODE.SIMPLE*, mais peut renvoyer une matrice de valeurs. Dans les lignes *12* et *13* de la figure 13-7, la variable quantitative `Physique du premier trimestre` (colonne *H*) possède trois ex-æquo (`6`, `13` et `17`), répétées chacune en trois exemplaires. La fonction ayant été entrée dans deux cellules seulement, elle ne peut renvoyer que deux d'entre elles. Elle les choisit dans leur ordre d'apparition dans la plage. Si, au contraire, la plage ne contient qu'un lauréat, la valeur correspondante est répétée plusieurs fois dans la matrice résultat (cas des mathématiques au premier trimestre en colonne *D*). Si chaque valeur est unique, elle renvoie une matrice de `#N/A`. *MODE.MULTIPLE* étant une fonction matricielle, elle doit être validée comme telle (voir le chapitre 4).

Indicateurs de dispersion

Plus une population est regroupée autour de sa moyenne, plus cette dernière est représentative de sa distribution. Obtenir une moyenne de 10 à partir des valeurs 9 et 11 n'a pas la même représentativité qu'à partir de 0 et 20.

Toujours dans la perspective de mieux appréhender la structure d'une population, nous allons maintenant nous intéresser à une vingtaine de fonctions renvoyant une série d'indicateurs de dispersion. Il y a d'abord celles qui renvoient les valeurs extrêmes, puis celles qui donnent une idée de la distance moyenne séparant chaque valeur de la moyenne et, enfin, celles qui découpent la population en tranches régulières plus ou moins larges et plus ou moins éloignées de la moyenne.

Valeurs extrêmes

Les fonctions *MAX* et *MIN* (lignes 4 et 5 de la figure 13-10) renvoient la plus grande et la plus petite valeur numérique d'une population. *MAXA* et *MINA* (lignes 6 et 7 de la figure 13-10) remplissent le même office en tenant compte des valeurs logiques (1 pour `VRAI` et 0 pour `FAUX`) et des valeurs de texte (prises en compte pour 0).

GRANDE.VALEUR et *PETITE.VALEUR* (lignes 9 et 10 de la figure 13-10) renvoient le même genre d'information en excluant un certain nombre de valeurs extrêmes (deuxième argument). En *D9*, le résultat 19 correspond bien à la plus haute note de mathématiques du premier trimestre, abstraction faite des quatre valeurs de tête (le 5 saisi en deuxième argument indique que l'on veut renvoyer la cinquième plus grande valeur). Ces fonctions ne tiennent pas compte des valeurs logiques et des valeurs de texte.

	B	D	E	F	H	I	J	L	M	N	P	Q	R	T	U	V	X
1		Maths			Physique			Français			Histoire			Langues			
2		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
4	=MAX (Notes!D3:D27)	20,0	18,0	19,0	19,0	19,0	19,0	18,0	16,0	18,0	17,0	15,0	16,0	16,0	16,0	16,0	
5	=MIN (Notes!D3:D27)	2,0	2,0	5,0	2,0	2,0	7,0	3,0	2,0	4,0	4,0	2,0	7,0	5,0	2,0	7,0	
6	=MAXA (Notes!D3:D27)	20,0	18,0	19,0	19,0	19,0	19,0	18,0	16,0	18,0	17,0	15,0	16,0	16,0	16,0	16,0	
7	=MINA (Notes!D3:D27)	0,0	0,0	5,0	0,0	0,0	0,0	0,0	2,0	4,0	4,0	2,0	0,0	5,0	0,0	7,0	
9	=GRANDE.VALEUR (Notes!D3:D27;5)	19,0	17,0	18,0	17,0	16,0	17,0	16,0	13,0	14,0	15,0	13,0	14,0	14,0	12,0	14,0	
10	=PETITE.VALEUR (Notes!D3:D27;5)	7,0	4,0	8,0	6,0	4,0	8,0	6,0	5,0	9,0	7,0	6,0	9,0	7,0	4,0	9,0	
12	=QUARTILE .INCLURE (Notes!D3:D27;1)	7,0	5,8	9,0	6,5	5,0	8,8	7,0	6,0	9,0	8,0	7,0	9,0	9,0	5,8	10,0	QUARTILE
13	=QUARTILE .EXCLURE (Notes!D3:D27;3)	18,0	14,0	18,0	15,0	15,0	17,0	15,0	13,0	14,0	15,0	12,0	14,0	14,0	12,0	13,0	
15	=CENTILE .INCLURE (Notes!D3:D27;20%)	7,0	4,6	8,8	6,0	4,6	8,0	6,0	5,8	9,0	7,0	6,0	9,0	7,0	4,6	9,0	CENTILE
16	=CENTILE .EXCLURE (Notes!D3:D27;20%)	6,8	4,0	8,2	6,0	4,0	8,0	5,8	5,2	9,0	7,0	6,0	9,0	7,0	4,0	9,0	
18	=ECART.MOYEN (Notes!D3:D27)	4,7	4,1	4,2	4,1	4,3	4,0	3,9	3,6	2,5	3,4	3,2	2,3	3,0	3,1	2,1	

Figure 13–10
Ce tableau réunit les principaux indicateurs de dispersion. La colonne X donne la forme ancienne de certaines fonctions, conservées dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

20	=VAR.P.N (Notes!D3:D27)	31,2	25,5	21,5	23,5	25,4	19,0	20,8	16,6	9,6	15,1	14,8	7,1	11,2	13,5	6,5	VAR.P
21	=VAR.S (Notes!D3:D27)	32,7	26,6	22,4	24,6	26,5	19,8	21,7	17,3	10,0	15,7	15,4	7,4	11,6	14,1	6,8	VAR
22	=VARPA (Notes!D3:D27)	35,9	28,7	21,5	30,3	28,2	24,6	28,2	16,6	9,6	15,1	14,8	11,9	11,2	16,2	6,5	
23	=VARA (Notes!D3:D27)	37,5	29,9	22,4	31,5	29,3	25,7	29,3	17,3	10,0	15,7	15,4	12,4	11,6	16,9	6,8	
25	=ECARTYPE .PEARSON (Notes!D3:D27)	5,6	5,0	4,6	4,9	5,0	4,4	4,6	4,1	3,1	3,9	3,8	2,7	3,3	3,7	2,5	ECARTYPEP
26	=ECARTYPE .STANDARD (Notes!D3:D27)	5,7	5,2	4,7	5,0	5,1	4,4	4,7	4,2	3,2	4,0	3,9	2,7	3,4	3,8	2,6	ECARTYPE
27	=STDEVPA (Notes!D3:D27)	6,0	5,4	4,6	5,5	5,3	5,0	5,3	4,1	3,1	3,9	3,8	3,4	3,3	4,0	2,5	
28	=STDEVA (Notes!D3:D27)	6,1	5,5	4,7	5,6	5,4	5,1	5,4	4,2	3,2	4,0	3,9	3,5	3,4	4,1	2,6	

La colonne *B* expose la syntaxe des formules entrées en *D4:D28*, ces dernières ayant ensuite été recopiées dans la plage *E4:V28*. Toutes utilisent les données stockées dans la feuille `Notes` présentée figure 13-1.

Quartiles et centiles

Quartiles

QUARTILE.INCLURE et *QUARTILE.EXCLURE* divisent la population en quatre quarts et renvoient les valeurs marquant ces séparations (lignes *12* et *13* de la figure 13-10).

	B	C	D	F	G	H	J	K	L	N	O	P	R	S	T
1	Maths			Physique			Français			Histoire			Langues		
2	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
4	2	2	5	2	2	7	3	2	4	4	2	7	5	2	7
5	2	3	6	3	3	7	4	3	7	6	2	8	6	4	7
6	3	3	6	5	4	7	5	4	8	6	3	8	7	4	7
7	5	4	6	6	4	7	5	4	9	7	4	8	7	4	8
8	6	4	8	6	4	8	6	5	9	7	6	9	7	4	9
9	7	5	9	6	5	8	6	6	9	7	6	9	7	5	9
10	7	6	9	7	5	9	8	6	9	8	7	9	9	6	10
11	7	8	9	8	8	10	9	7	11	8	7	9	9	7	10
12	11	10	11	8	8	10	9	7	11	8	7	9	9	9	10
13	11	10	11	10	8	10	11	7	11	10	9	11	9	9	10
14	11	11	12	11	9	12	11	10	11	10	9	12	10	9	11
15	11	11	14	11	9	13	12	10	11	11	9	12	10	10	11
16	12	11	14	12	9	14	12	11	12	11	10	12	10	10	11
17	13	11	14	12	10	14	13	11	12	12	11	12	11	11	11
18	14	12	15	13	11	16	13	11	13	12	11	12	13	11	12
19	14	12	16	13	12	16	14	12	14	14	11	13	13	11	12
20	14	12	17	13	12	17	14	12	14	15	11	13	14	12	13
21	18	14	17	15	15	17	15	13	14	15	12	14	14	12	13
22	18	14	18	17	15	17	16	13	14	15	12	14	14	12	13
23	19	17	18	17	16	17	17	13	14	15	12	14	14	12	13
24	19	17	18	17	17	18	17	13	14	15	13	15	14	13	14
25	20	18	18	18	17	19	17	14	15	16	14	15	15	13	14
26	20	18	18	19	17	19	18	15	15	16	14	15	15	13	14
27	20	18	19		19	19		15	16	16	15	16	16	16	16
28			19					16	18	17	15		16		16

Figure 13–11
Ici, les quinze variables sont triées dans l'ordre croissant. La feuille qui contient ce tableau s'appelle « NotesTri ».

Dans le tableau de la figure 13-11, on a fait ressortir sur fond de couleur, en caractères verts les plus petites valeurs, en caractères mauves les plus grandes, en caractères noirs les valeurs à partir desquelles sont calculés les premier et troisième quartiles, en blanc sur fond rouge les valeurs à partir desquelles est calculé le deuxième quartile (autrement dit, la médiane).

La différence entre *QUARTILE.INCLURE* et *QUARTILE.EXCLURE* réside dans la prise en compte ou non des valeurs extrêmes de la population. Pour le deuxième quartile, leurs résultats sont identiques, mais pour le premier et le troisième, *QUARTILE.EXCLURE* renvoie des valeurs légèrement plus excentrées.

REPRÉSENTATION Les boîtes à moustaches

Les fonctions *MIN*, *MAX*, *MEDIANE* et *QUARTILE.EXCLURE* peuvent être utilisées pour construire des boîtes à moustaches. Ces dernières donnent une représentation de la dispersion d'une population en enfermant la moitié des valeurs les plus centrales dans des boîtes et en donnant, par des traits recouvrant les premier et dernier quartiles, l'étendue des valeurs extrêmes de la population.

	B	D	E	F	H	I	J	L	M	N	P	Q	R	T	U	V
2		Maths			Physique			Français			Histoire			Langues		
3		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
6	=MEDIANE(NotesTri!B4:B28)	11,5	11,0	14,0	11,0	9,0	13,5	12,0	11,0	12,0	11,0	10,0	12,0	10,0	10,0	11,0
8	=QUARTILE.EXCLURE(NotesTri!B4:B28;1)	7,0	5,3	9,0	6,0	5,0	8,3	6,0	6,0	9,0	7,5	6,5	9,0	8,0	5,3	9,5
10	=QUARTILE.EXCLURE(NotesTri!B4:B28;3)	18,0	14,0	18,0	15,0	15,0	17,0	15,0	13,0	14,0	15,0	12,0	14,0	14,0	12,0	13,0
12	=MAX(NotesTri!B4:B28)	20	18	19	19	19	19	18	16	18	17	15	16	16	16	16
14	=MIN(NotesTri!B4:B28)	2	2	5	2	2	7	3	2	4	4	2	7	5	2	7

Figure 13–12 Ce tableau utilise les fonctions MIN, MAX, MEDIANE et QUARTILE.EXCLURE pour obtenir les valeurs nécessaires à la construction des boîtes à moustaches.

À la figure 13-13, les losanges bleus représentent les notes trimestrielles de chaque élève par matière. Comme plusieurs points se trouvent regroupés autour d'une même note, ils forment des amas plus ou moins importants. Les extrémités de chaque boîte marquent le début du deuxième quartile et la fin du troisième. Le trait rouge apparaissant au sein de chaque boîte représente la médiane. Les traits verts matérialisent les plus petites valeurs et les violets, les plus grandes. Les quelques points qui apparaissent sur l'axe des abscisses correspondent aux valeurs de texte Abs, représentées comme des valeurs nulles.

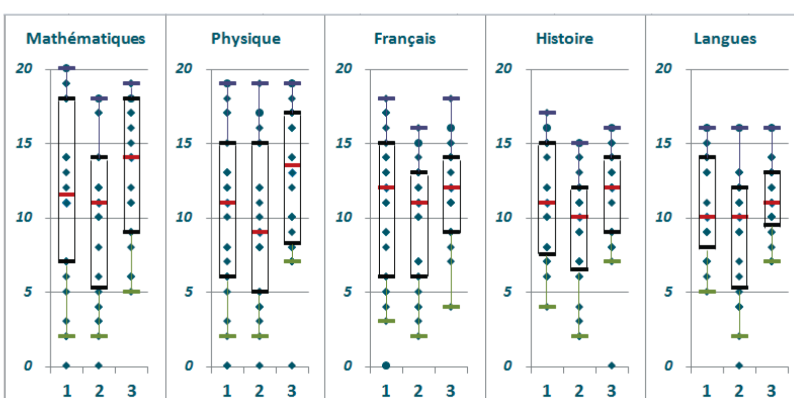


Figure 13–13 Ces quinze boîtes à moustaches donnent une idée plus précise de la dispersion des notes dans les diverses matières.

Vous remarquerez que cette représentation montre que les notes du troisième trimestre en langues sont beaucoup plus concentrées autour de la médiane que les notes de mathématiques du premier trimestre.

Centiles

On peut travailler avec une division plus fine de la population en utilisant *CENTILE.INCLURE* et *CENTILE.EXCLURE* (lignes 15 et 16 de la figure 13-10). En entrant par exemple 20%, 40%, 60% et 80% en deuxième argument de quatre fonctions *CENTILE.INCLURE*, on divise la population en cinq (au lieu de quatre avec les quartiles). Le deuxième argument peut prendre n'importe quelle valeur entre 0 et 1.

Mesure des écarts à la moyenne

Une troisième série de fonctions mesure divers types d'écarts entre les valeurs de la variable et la moyenne. Plus les résultats renvoyés sont grands, plus la dispersion est importante.

Tableau 13–3 Écarts à la moyenne

Fonction	Description
ECART.MOYEN	Cette fonction calcule la valeur absolue des écarts entre les notes et la moyenne. Elle en fait la somme qu'elle divise par le nombre de notes et renvoie le résultat obtenu (ligne 18 de la figure 13-10). Cet indicateur est facile à comprendre, mais il ne présente pas des propriétés mathématiques très intéressantes. De ce fait, on utilise plutôt la variance et sa racine carrée, l'écart-type.
VAR.P.N et *VAR.S*	Cette fonction (ligne 20 de la figure 13-10) fait le même calcul que la fonction *ECART.MOYEN*, mais en travaillant à partir du carré des écarts à la moyenne. *VAR.S* (ligne 21 de la figure 13-10) fait un calcul similaire, mais au lieu de diviser par n (population totale de la variable), elle divise par n-1.
ECARTYPE.PEARSON et *ECARTYPE.STANDARD*	Ces fonctions (lignes 25 et 26 de la figure 13-10) correspondent respectivement à la racine carrée de *VAR.P.N* et à celle de *VAR.S*.

Calculée à partir du carré des distances, la variance est très influencée par les valeurs aberrantes (extrêmes). Aussi, avant de faire ce calcul, il est important d'identifier ces dernières afin de ne pas les prendre en compte.

Par ailleurs, si l'on utilise un échantillon aléatoire pour estimer la variance d'une population, on risque de sous-estimer cette dernière car la dispersion d'un échantillon a de fortes chances d'être inférieure à celle de la population dont il est issu. C'est pourquoi, dans le cas d'un échantillon, au lieu de diviser la somme des carrés des écarts à la moyenne par n, on divise par n-1. On parle alors d'une variance sans biais. En résumé, si vous réalisez vos calculs à partir des valeurs d'une population globale, il faudra utiliser les fonctions *VAR.P.N* et *ECARTYPE.PEARSON*, alors que si vous travaillez à partir d'un échantillon dans la perspective d'estimer la variance et l'écart-type d'une population globale, il faudra utiliser les fonctions *VAR.S* et *ECARTYPE.STANDARD*.

Les quatre fonctions *VARPA*, *VARA*, *STDEVPA* et *STDEVA* (lignes 22, 23, 27 et 28 de la figure 13-10) font respectivement les mêmes calculs que les quatre fonctions précédentes, mais en prenant en compte les valeurs logiques (`1` pour `VRAI` et `0` pour `FAUX`) et textuelles (valeur `0`).

ALLER PLUS LOIN **Symétrie d'une distribution**

Lorsqu'on étudie la forme d'une distribution, outre sa dispersion, on s'intéresse également à son caractère plus ou moins symétrique et plus ou moins aplati. La fonction *COEFFICIENT.ASYMETRIE* proposée par Excel caractérise le degré d'asymétrie d'une distribution par rapport à sa moyenne. Excel 2013 propose une deuxième fonction *COEFFICIENT.ASYMETRIE.P* qui calcule la même chose que *COEFFICIENT.ASYMETRIE*, mais sur les données d'une population totale, alors qu'avec *COEFFICIENT.ASYMETRIE*, le calcul est effectué sur les données d'un échantillon.

Le coefficient d'aplatissement de Pearson ou Kurtosis, quant à lui, indique le degré d'aplatissement d'une distribution. Étant donné que la fonction *KURTOSIS* d'Excel fait référence à une loi normale, cette notion sera développée vers la fin du chapitre, quand la loi normale aura été présentée.

$$\frac{1}{(n-1)(n-2)}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma_é}\right)^3$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma_p}\right)^3$$

Figure 13–14 La première formule renvoie le coefficient d'asymétrie d'une distribution à partir des données d'un échantillon. L'écart-type utilisé au dénominateur est un écart-type standard (calculé sur un échantillon). Le x surmonté d'une barre désigne la moyenne des valeurs. La seconde formule renvoie le coefficient d'asymétrie d'une distribution à partir des données d'une population totale. L'écart-type utilisé au dénominateur est un écart-type Pearson (calculé sur une population totale).

Une asymétrie positive traduit un décalage de la moyenne vers les plus grandes valeurs (la moyenne est supérieure à la médiane), alors qu'une asymétrie négative traduit la situation inverse (la moyenne est inférieure à la médiane). Lorsque le coefficient d'asymétrie est proche de zéro, moyenne et médiane sont confondues. La figure 13-15 illustre ces trois situations.

	B	D	F	H
2	Notes	Phys. 2	Hist. 2	Lang. 3
4	1	0	0	0
5	2	1	2	0
6	3	1	1	0
7	4	3	1	0
8	5	2	0	0
9	6	0	2	0
10	7	0	3	3
11	8	3	0	1
12	9	3	3	2
13	10	1	1	4
14	11	1	4	4
15	12	2	3	2
16	13	0	1	4
17	14	0	2	3
18	15	2	2	0
19	16	1	0	2
20	17	3	0	0
21	18	0	0	0
22	19	1	0	0
23	20	0	0	0
25	Moyenne	9,96	9,28	11,24
27	Médiane	9,00	10,00	11,00
29	Coef. Asym. Éch.	0,19	-0,42	0,02
31	Coef. Asym. Pop.	0,18	-0,39	0,02

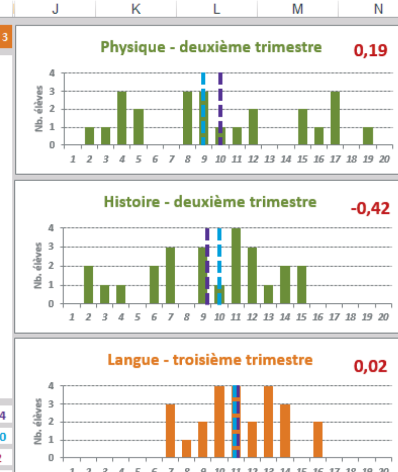


Figure 13–15
Les résultats en langues du troisième trimestre constituent une distribution presque symétrique. Au deuxième trimestre, les résultats de physique traduisent une asymétrie positive et ceux d'histoire, une asymétrie négative.

> Pour construire les résultats de la figure 13-15, on a utilisé la fonction *NB.SI* dans la plage *D4:H23*, *MOYENNE* dans la plage *D25:H25*, *MEDIANE* dans la plage *D27:H27*, *COEFFICIENT.ASYMETRIE* dans la plage *D29:H29* et *COEFFICIENT.ASYMETRIE.P* (**nouveauté Excel 2013**) dans la plage *D31:H31*. Les valeurs utilisées sont toujours celles de la feuille `Notes`. On a utilisé le violet pour la moyenne et le bleu pour la médiane, couleurs reprises dans les graphiques. En rouge apparaît la valeur du coefficient d'asymétrie (*COEFFICIENT.ASYMETRIE*) de chacune des trois distributions.

Ordonner les valeurs

Dans les sections précédentes, pour ordonner les valeurs nous avons eu recours au tri. Or Excel fournit quatre fonctions rendant les mêmes services sans déplacer les données physiquement : *EQUATION.RANG*, *MOYENNE.RANG*, *RANG.POURCENTAGE.INCLURE* et *RANG.POURCENTAGE.EXCLURE*. Malgré de légères variantes quant à la nature de leurs résultats, toutes renvoient une valeur qui reflète la place de chaque individu dans la population totale.

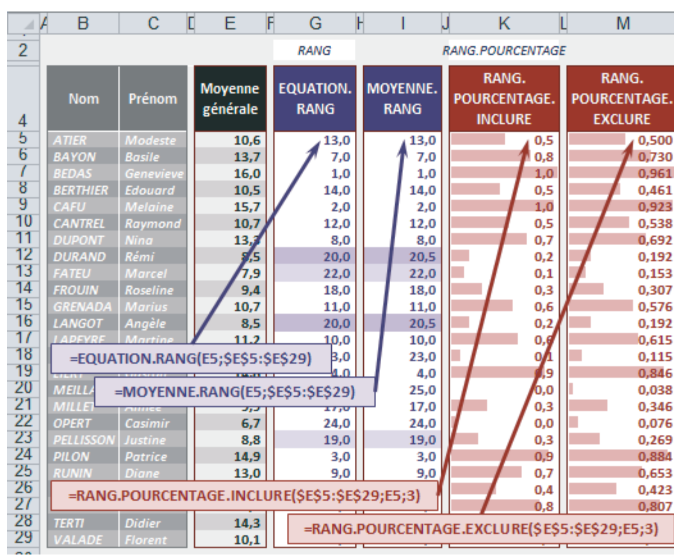


Figure 13–16
Les quatre cadres présentent la syntaxe des formules entrées en ligne 5. Elles ont ensuite été recopiées dans l'ensemble du tableau. En ligne 2, on trouve deux formes anciennes des fonctions conservées dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Tableau 13–4 Ordonner les valeurs

Fonction	Description
EQUATION.RANG et *MOYENNE.RANG*	Ces deux fonctions renvoient un nombre correspondant à la place de chaque individu dans la population totale. Si vous ne précisez pas le troisième argument, Excel attribue le rang `1` à la plus grande valeur. Si vous souhaitez qu'il inverse sa position, il faut saisir `VRAI` en troisième argument. Excel attribue alors le rang `1` à la plus petite valeur. La fonction *EQUATION.RANG* traite les ex-æquo. Deux élèves ont `8,5` de moyenne générale : la fonction renvoie deux fois la position `20`, puis passe directement à `22`. La fonction *MOYENNE.RANG* traite également les ex-æquo mais en opérant un lissage : elle renvoie donc deux fois la valeur `20,5`. Les trois rangs successifs sont donc : `19`, `20,5` et `22`, alors qu'avec *EQUATION.RANG*, ils sont `19`, `20` et `22`.
RANG.POURCENTAGE .INCLURE et *RANG.POURCENTAGE .EXCLURE*	La logique de ces deux fonctions est de ramener à `1` la plus grande valeur de la variable et à `0` la plus petite, puis d'exprimer toutes les autres proportionnellement en tenant compte de la valeur et de la position. Avec *RANG.POURCENTAGE.INCLURE*, la plus grande valeur correspond à `1` et la plus petite à `0`. *RANG.POURCENTAGE.EXCLURE* renvoie un nombre légèrement inférieur à `1` pour la plus grande valeur et légèrement supérieur à `0` pour la plus petite. Le dernier argument sert à régler la précision du résultat. En indiquant `3`, on obtient des nombres à trois décimales. Dans l'exemple présenté figure 13-16, c'est le cas, mais pour améliorer la lisibilité, on a appliqué un format de nombre qui réduit l'affichage à une décimale (pour la colonne *K*). En revanche, le format conditionnel qui calcule les barres proportionnelles aux résultats utilise les valeurs précises à trois décimales renvoyées par la fonction.

Liaison entre deux variables quantitatives

Grâce aux indices présentés dans la section précédente, nous avons pu approfondir, individuellement, notre connaissance des quinze variables quantitatives présentées figure 13-1. L'étape suivante consiste à chercher s'il existe une relation entre elles. Pour que cette recherche ait vraiment un intérêt, nous avons créé cinq nouvelles variables en considérant la moyenne annuelle par élève et par matière. Parallèlement à cela, nous avons récupéré les notes obtenues à l'examen blanc de fin d'année (voir la figure 13-17).

Figure 13–17
Dix nouvelles variables quantitatives (cinq examens blancs et cinq moyennes annuelles par matière). Ce tableau se trouve également sur la feuille Notes.

	A	B	X	Y	Z	AA	AB	AD	AE	AF	AG	AH
1			Moyenne 3 trimestres					Examen blanc				
2	Nom	Prénom	M	P	F	H	L	M	P	F	H	L
3	ATIER	Modeste	6,0	5,7	14,0	14,3	13,0	10	8	12	15	13
4	BAYON	Basile	18,0	16,0	12,3	10,7	11,7	16	16	10	12	12
5	BEDAS	Genevieve	18,3	18,3	14,3	14,7	14,3	19	17	11	14	14
6	BERTHIER	Edouard	10,7	9,7	11,0	11,0	10,3	12	11	10	13	11
7	CAFU	Melaine	18,3	17,0	13,5	14,0	14,5	16	18	13	16	14
8	CANTREL	Raymond	11,3	10,0	12,3	9,7	10,0	14	11	11	12	11
9	DUPONT	Nina	13,3	12,0	13,0	14,0	14,3	11	14	11	13	13
10	DURAND	Rémi	12,3	9,3	6,3	7,0	7,7	15	12	7	9	9
11	FATEU	Marcel	8,0	8,3	6,7	8,0	8,3	9	12	7	10	10
12	FROUIN	Roseline	3,7	5,5	14,0	12,3	10,0	5	11	12	13	11
13	GRENADA	Marius	11,3	11,7	10,7	10,3	9,7	11	14	15	11	12
14	LANGOT	Angèle	12,3	9,0	7,0	6,3	8,0	13	11	9	12	10
15	LAPEYRE	Martine	7,3	6,7	14,0	14,7	13,3	12	11	13	13	12
16	LEGRAND	Véronique	4,3	6,0	8,7	8,0	7,0	13	11	11	11	11
17	LIERT	Gaston	15,5	19,0	16,0	14,3	12,3	15	20	16	17	11
18	MEILLAC	Bernadette	4,5	5,3	6,0	6,7	7,0	8	11	5	11	10
19	MILLET	Aimée	13,0	13,3	8,0	7,7	7,3	11	14	7	11	11
20	OPERT	Casimir	6,3	6,0	8,0	6,7	6,3	9	11	7	11	9
21	PELLISSON	Justine	4,5	3,7	10,7	11,3	12,3	8	10	10	10	12
22	PILON	Patrice	12,7	16,7	16,0	15,0	14,0	12	18	15	16	14
23	RUNIN	Diane	16,3	14,0	10,5	11,5	11,7	17	13	10	14	12
24	TALAMON	Sylvain	18,7	17,0	3,0	6,0	6,7	19	16	6	10	9
25	TEROIN	Pascal	17,0	14,0	14,0	13,7	13,7	16	15	11	15	13
26	TERTI	Didier	14,3	14,0	16,3	14,0	13,0	13	13	15	17	12
27	VALADE	Florent	16,7	17,3	5,0	5,0	6,3	16	18	4	8	9

COMPRENDRE **Expliquer graphiquement une variable A à partir d'une variable B**

Vérifions dans quelle mesure la note obtenue à l'examen blanc s'explique par la moyenne annuelle.

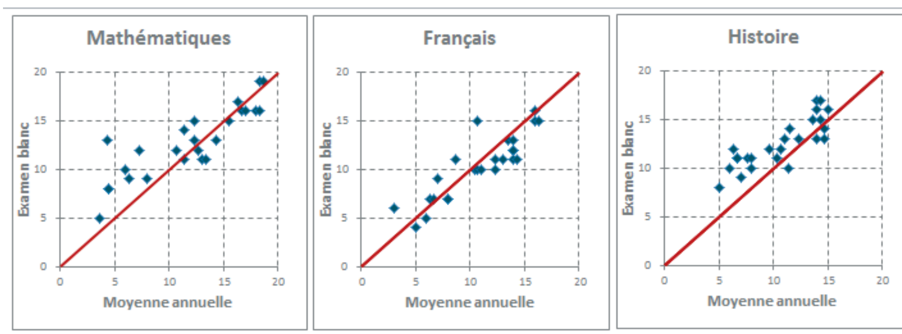


Figure 13–18 Ces trois graphiques en nuages de points appliqués aux mathématiques, au français et à l'histoire permettent de comparer assez précisément les valeurs du contrôle continu et celles de l'examen blanc.

Dans les trois graphiques de la figure 13-18, chaque point représente la performance d'un élève dans une des matières. Les points situés sous la bissectrice correspondent aux élèves qui ont mieux travaillé tout au long de l'année qu'à l'examen blanc, alors que ceux qui apparaissent au-dessus reflètent une meilleure performance à l'examen.

On constate qu'en français, les élèves ont eu tendance à rater leur examen blanc, alors qu'en histoire, les résultats ont été plutôt meilleurs que tout au long de l'année.

Utiliser le coefficient de corrélation

Le coefficient de corrélation est un indice numérique qui mesure la liaison entre deux variables. Un seul nombre résume forcément très mal une situation complexe. Néanmoins, le coefficient de corrélation fournit rapidement une valeur utile pour comparer plusieurs situations.

Qu'est-ce qu'un coefficient de corrélation ?

La figure 13-19 propose la définition du coefficient de corrélation (première formule). Comme cette dernière fait appel à la covariance, on en donne également la définition (deuxième formule). À l'image de la variance qui fait la somme des carrés des écarts à la moyenne divisé par le nombre d'individus, la covariance fait la somme des produits des écarts à la moyenne des deux variables, qu'elle divise par le nombre d'individus. Elle permet d'évaluer le sens de variation de deux variables et de qualifier leur indépendance. Une autre façon d'énoncer cette définition est de dire que le coefficient de corrélation est le rapport de la somme des produits deux à deux des variables centrées réduites sur le nombre d'individus. Il s'agit donc d'un nombre compris entre `-1` et `1`. Calculé à partir des données centrées réduites, il est indépendant des moyennes et des écarts-types des deux variables.

Figure 13–19
Le coefficient de corrélation (ρ) est le rapport de la covariance des deux variables sur le produit de leurs écarts-types (Pearson).

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \, \sigma_Y}$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})$$

Comprendre Variable centrée réduite

Quatre des dix variables présentées figure 13-17 ont été soumises au même traitement. On a soustrait de chaque valeur la moyenne de la variable, puis on a divisé le résultat par l'écart-type (Pearson). Les valeurs ainsi obtenues définissent quatre nouvelles variables qualifiées de centrées réduites.

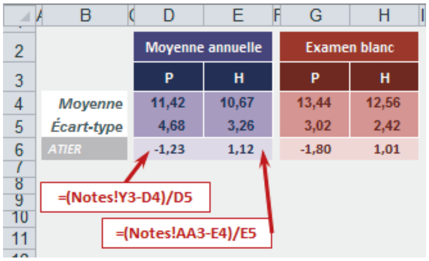


Figure 13–20 Détail du calcul, pour le premier élève de la liste, fournissant les premières valeurs des quatre variables centrées réduites.

Il se trouve qu'Excel fournit la fonction *CENTREE.REDUITE* qui renvoie directement ce résultat à partir de ses trois arguments (valeur, moyenne et écart-type). La figure 13-21 en donne une illustration. Ce tableau est situé sur une feuille nommée `CentrRed`.

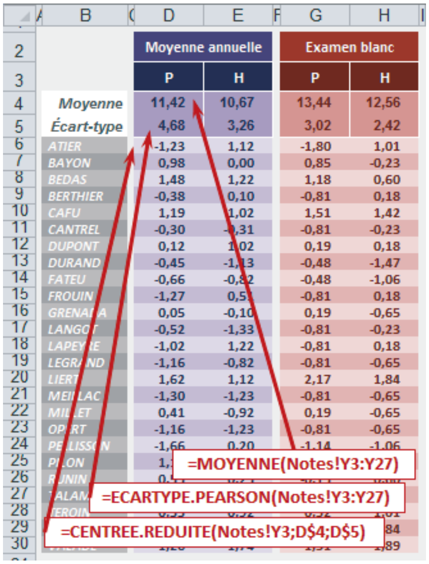


Figure 13–21 Mise en œuvre de la fonction CENTREE.REDUITE pour transformer les valeurs des quatre variables étudiées (physique et histoire en moyenne annuelle et examen blanc).

Les variables centrées réduites présentent l'avantage d'être complètement indépendantes des unités de départ. Elles autorisent ainsi des comparaisons entre populations de natures diverses : on peut comparer les résultats d'une société avec ceux d'une filière, on peut rapprocher deux groupes d'individus différents, et ainsi de suite.

La moyenne d'une variable centrée réduite est égale à 0 et son écart-type est égal à 1. Par conséquent, l'unité de mesure d'une variable centrée réduite est l'écart-type. Une note égale à 1 signifie qu'elle est située à un écart-type de la moyenne, une note égale à 2, qu'elle est située à deux écarts-types, etc.

Calculer le coefficient de corrélation

Excel fournit la fonction *COEFFICIENT.CORRELATION* qui, à partir des données de deux variables, renvoie directement leur coefficient de corrélation. Cette fonction calcule ce coefficient à partir des valeurs brutes de chaque variable (vous n'avez pas besoin de calculer au préalable les deux variables centrées réduites correspondantes). Les cinq coefficients affichés dans la plage *D3:H3* de la figure 13-22 ont été obtenus ainsi (la fonction *COEFFICIENT.CORRELATION* utilise les données présentées figure 13-17).

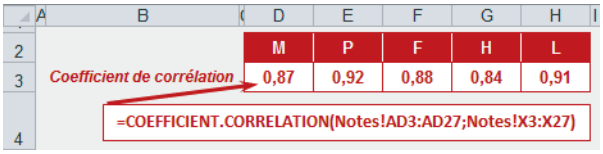


Figure 13–22
Mise en œuvre de la fonction COEFFICIENT.CORRELATION à partir des données stockées dans la feuille Notes (figure 13-17).

ALLER PLUS LOIN Diverses approches du coefficient de corrélation

Les diverses fonctions fournies par Excel permettent d'appréhender le coefficient de corrélation de plusieurs façons. Les calculs proposés ici s'appliquent aux notes obtenues en physique. Les deux variables étudiées sont donc la moyenne annuelle et les notes obtenues à l'examen blanc pour cette matière.

Dans la cellule *E8* de la figure 13-23, on a fait la somme des produits deux à deux des deux variables centrées réduites. Vous pouvez vérifier que le rapport de cette somme sur le nombre d'individus donne bien le même résultat que la fonction *COEFFICIENT.CORRELATION*.

Dans la plage *E14:E18*, on a appliqué la définition du coefficient de corrélation donnée figure 13-19. On a donc d'abord calculé la covariance à l'aide de la fonction *COVARIANCE.PEARSON*, puis l'écart-type des deux variables. Le rapport de la covariance sur le produit des deux écarts-types redonne bien le même résultat que la fonction *COEFFICIENT.CORRELATION*.

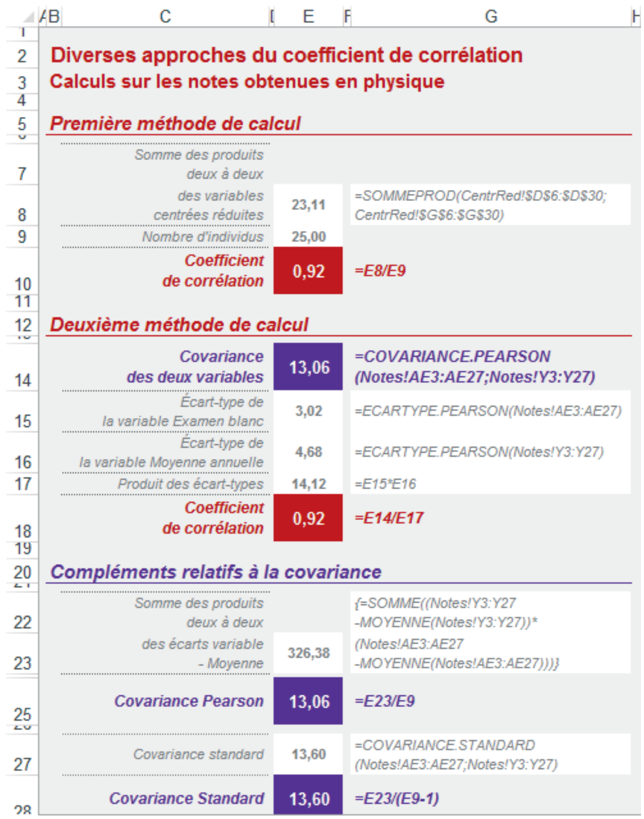


Figure 13–23 Détail de quelques calculs pour mieux comprendre coefficient de corrélation et covariance. La colonne G affiche la syntaxe des formules entrées en E8:E28.

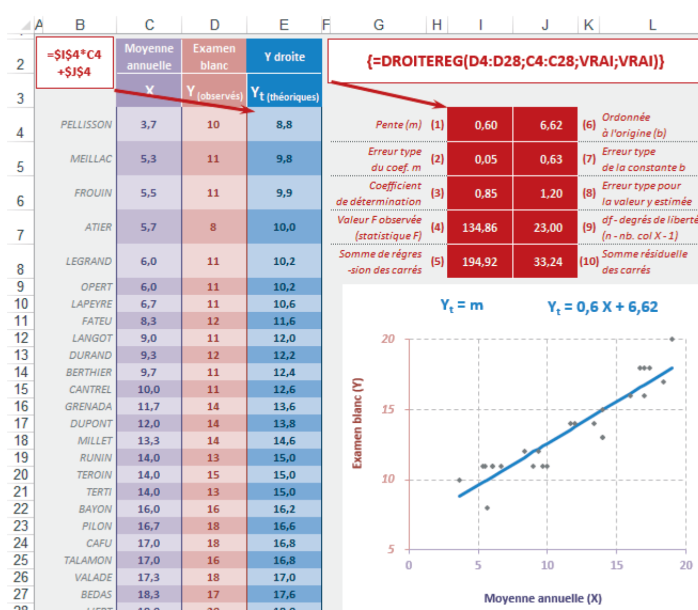
La définition de la covariance donnée figure 13-19 correspond à la fonction Excel *COVARIANCE.PEARSON* (au dénominateur du rapport figure le nombre d'individus). Or, précédemment, quand nous avions présenté variance et écart-type, nous avions fait allusion au fait qu'il existait deux couples de fonctions (*VAR.P.N* et *ECARTYPE.PEARSON* quand on travaille à partir des données d'une population et *VAR.S* et *ECARTYPE.STANDARD* quand on travaille à partir des données d'un échantillon). Cette nuance existe également avec la covariance, pour laquelle Excel fournit les fonctions *COVARIANCE.PEARSON* (n au dénominateur) et *COVARIANCE.STANDARD* (n−1 au dénominateur). Les formules proposées dans la plage *E23:E28* détaillent les calculs correspondant à ces deux natures de covariances. Faites attention car la cellule *E23* contient une formule matricielle. Il faut donc la valider correctement (voir le chapitre 4).
En outre, Excel fournit une troisième fonction, *COVARIANCE*, qui est l'ancienne forme de la fonction *COVARIANCE.PEARSON*. Elle est conservée pour assurer la compatibilité avec les versions antérieures.

> **BON À SAVOIR Fonction PEARSON**
>
> Le coefficient de corrélation d'échantillonnage de Pearson donné par la fonction *PEARSON* renvoie exactement la même valeur que le coefficient de corrélation. L'algorithme qui définit le coefficient de Pearson est un rapport semblable dans lequel on aurait multiplié le numérateur et le dénominateur par `n` (le nombre d'observations).

Utiliser la régression

Vous avez sans doute remarqué que les graphiques présentés figure 13-18 montraient des nuages de points prenant des formes d'ellipses allongées ; il ne vous aura pas échappé non plus que les coefficients de corrélation calculés dans la section précédente sont tous assez proches de `1` (ils varient entre `0,84` et `0,92`). Ces deux observations laissent supposer que les variables `Moyenne annuelle` et `Examen blanc` sont assez étroitement liées. Si les points ne dessinaient pas de forme particulière et si les coefficients de corrélation étaient plus proches de zéro, nous pourrions émettre davantage de doutes sur cette éventuelle liaison.

Il n'est donc pas aberrant de chercher dans quelle mesure la note à l'examen blanc (`Y`) peut être expliquée par la moyenne annuelle (`X`) à travers une relation du type $Y_t = mX + b$. Il s'agit de l'équation de la droite de pente `m` et d'ordonnée à l'origine `b`. Comme c'est en physique que le coefficient de corrélation est le plus élevé, c'est pour cette matière que nous allons mettre en place la droite de régression représentée figure 13-24.



	B	C	D	E
2	=$I$4*C4 +$J$4	Moyenne annuelle	Examen blanc	Y droite
3		X	Y (observés)	Yt (théoriques)
4	PELLISSON	3,7	10	8,8
5	MEILLAC	5,3	11	9,8
6	FROUIN	5,5	11	9,9
7	ATIER	5,7	8	10,0
8	LEGRAND	6,0	11	10,2
9	OPERT	6,0	11	10,2
10	LAPEYRE	6,7	11	10,6
11	FATEU	8,3	12	11,6
12	LANGOT	9,0	11	12,0
13	DURAND	9,3	12	12,2
14	BERTHIER	9,7	11	12,4
15	CANTREL	10,0	11	12,6
16	GRENADA	11,7	14	13,6
17	DUPONT	12,0	14	13,8
18	MILLET	13,3	14	14,6
19	RUNIN	14,0	13	15,0
20	TEROIN	14,0	15	15,0
21	TERTI	14,0	13	15,0
22	BAYON	16,0	16	16,2
23	PILON	16,7	18	16,6
24	CAFU	17,0	18	16,8
25	TALAMON	17,0	16	16,8
26	VALADE	17,3	18	17,0
27	BEDAS	18,3	17	17,6
28	LIERT	19,0	20	18,0

{=DROITEREG(D4:D28;C4:C28;VRAI;VRAI)}

G	H	I	J	K	L
Pente (m)	(1)	0,60	6,62	(6)	Ordonnée à l'origine (b)
Erreur type du coef. m	(2)	0,05	0,63	(7)	Erreur type de la constante b
Coefficient de détermination	(3)	0,85	1,20	(8)	Erreur type pour la valeur y estimée
Valeur F observée (statistique F)	(4)	134,86	23,00	(9)	df - degrés de liberté (n - nb. col X - 1)
Somme de régres-sion des carrés	(5)	194,92	33,24	(10)	Somme résiduelle des carrés

Figure 13–24
La fonction DROITEREG renvoie une matrice de dix valeurs (I4:J8) qui permettent de définir l'équation de la droite de régression et d'en nuancer la représentativité.

Régression simple

On cherche la droite qui constituera la meilleure image stylisée du nuage de points gris représenté figure 13-24. Bien évidemment, à moins qu'en des circonstances exceptionnelles les points du nuage s'alignent parfaitement sur la droite, le résultat obtenu ne décrit qu'imparfaitement la réalité. La fonction *DROITEREG* utilisée ici pour définir les paramètres `m` et `b` de l'équation de la droite de régression renvoie donc toute une série de valeurs statistiques pour mesurer la pertinence de la modélisation obtenue.

À la figure 13-24, on a repris en colonnes *C* et *D* la moyenne annuelle et les résultats à l'examen blanc pour la physique. La fonction *DROITEREG* a ensuite été entrée dans la plage *I4:J8*. Comme elle renvoie plusieurs résultats simultanément, il s'agit d'une fonction matricielle qui doit être traitée comme telle, c'est-à-dire validée en pressant simultanément les touches *Ctrl+Maj+Entrée*, en ayant, préalablement à sa saisie, sélectionné la plage *I4:J8*.

Sa syntaxe est donnée dans le cadre situé dans le coin supérieur droit de la figure. Les deux premiers arguments précisent les plages occupées par les données des variables `Y` et `X`. Le troisième argument, s'il est égal à `FAUX`, indique que `b` doit être forcé à `0` (on cherche alors une droite d'équation $Y_t = mX$). S'il est égal à `VRAI` ou omis, `b` est calculé « normalement ». Lorsque le quatrième argument est `VRAI`, toutes les valeurs statistiques sont renvoyées (il faut juste, avant la saisie de la fonction, avoir sélectionné une plage suffisamment grande pour qu'Excel ait la place de les afficher). Dans le cadre d'une régression simple, la fonction peut renvoyer jusqu'à dix valeurs. Leur rôle est indiqué de part et d'autre de la plage *I4:J8*.

Tableau 13–5 Résultats renvoyés par la fonction DROITEREG

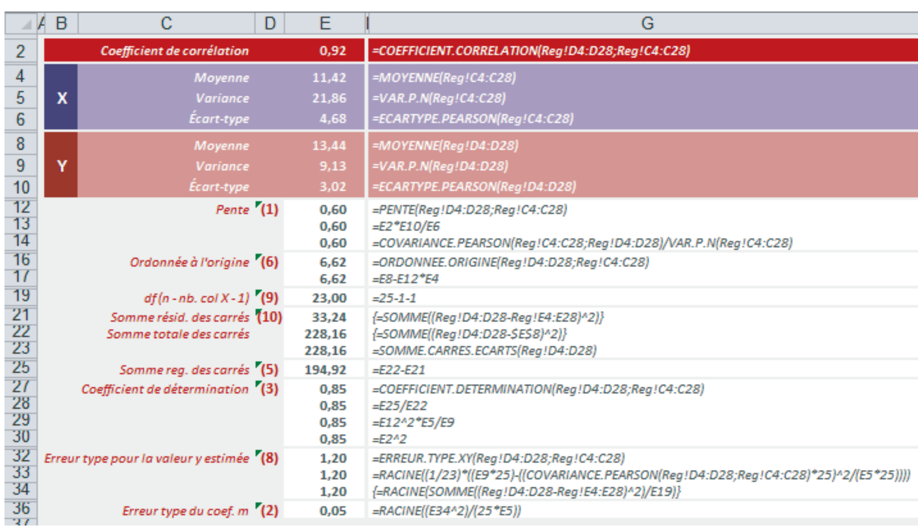
Résultat	Utilisation
`Pente (1)` et `ordonnée à l'origine (6)`	Les deux premiers résultats retournés par la fonction sont la pente (`0,6`) et l'ordonnée à l'origine (`6,62`) de la droite de régression : $Y_t = 0,6\ X + 6,62$. Cette équation a été utilisée en colonne *E* pour trouver la variable *Y* « théorique » qui donne la valeur des points de la droite pour les `x` connus. La formule entrée en *E4* est indiquée en rouge dans le cadre blanc. Elle a ensuite été recopiée dans la plage *E5:E28*.
`Coefficient de détermination (3)`	Cet indice mesure la similitude entre les valeurs `y` estimées et les valeurs `y` réelles. Dans notre exemple, il est égal à `0,85`. Il est équivalent au coefficient de corrélation élevé au carré.
`Erreurs types (2), (7)` et `(8)`	*DROITEREG* renvoie trois mesures d'erreur. L'`erreur type pour la valeur y estimée` (`1,2` dans notre exemple) mesure le degré d'erreur dans la prévision de `y` à partir de `x`. Les deux autres résultats (`Erreur type du coefficient m`:`0,05` et `Erreur type de la constante b`:`0,63`) mesurent la fiabilité des deux paramètres `m` et `b` qui participent à la construction de l'équation de la droite de régression.

Tableau 13–5 Résultats renvoyés par la fonction DROITEREG (suite)

Résultat	Utilisation
`Somme de régression des carrés (5)` et `somme résiduelle des carrés (10)`	Ces deux valeurs peuvent être utilisées pour calculer le coefficient de détermination. La `somme résiduelle des carrés` correspond à la somme du carré des écarts entre la variable Y observée et la variable Y théorique. La `somme de régression des carrés` correspond à la différence entre la `somme résiduelle des carrés` et la `somme totale des carrés` (cette dernière étant égale au produit de la variance par le nombre d'individus).
`Valeur F (4)` et `df, degrés de liberté (9)`	Ces deux derniers paramètres (`134,86` et `23`) seront abordés dans la section traitant de la régression multiple.

COMPRENDRE Relations entre les divers résultats

DROITEREG renvoie un bloc de huit valeurs statistiques, mais quelques fonctions d'Excel permettent d'en obtenir certaines individuellement. Par ailleurs, il peut être intéressant de connaître les relations qui existent entre les divers résultats obtenus. La figure 13-25 présente les différentes méthodes disponibles pour obtenir les résultats donnés par la fonction *DROITEREG*, ainsi que les relations liant certains d'entre eux.



	B	C	D	E	G
2		Coefficient de corrélation		0,92	=COEFFICIENT.CORRELATION(Reg!D4:D28;Reg!C4:C28)
4	X	Moyenne		11,42	=MOYENNE(Reg!C4:C28)
5		Variance		21,86	=VAR.P.N(Reg!C4:C28)
6		Écart-type		4,68	=ECARTYPE.PEARSON(Reg!C4:C28)
8	Y	Moyenne		13,44	=MOYENNE(Reg!D4:D28)
9		Variance		9,13	=VAR.P.N(Reg!D4:D28)
10		Écart-type		3,02	=ECARTYPE.PEARSON(Reg!D4:D28)
12		Pente	(1)	0,60	=PENTE(Reg!D4:D28;Reg!C4:C28)
13				0,60	=E2*E10/E6
14				0,60	=COVARIANCE.PEARSON(Reg!C4:C28;Reg!D4:D28)/VAR.P.N(Reg!C4:C28)
16		Ordonnée à l'origine	(6)	6,62	=ORDONNEE.ORIGINE(Reg!D4:D28;Reg!C4:C28)
17				6,62	=E8-E12*E4
19		df (n - nb. col X - 1)	(9)	23,00	=25-1-1
21		Somme résid. des carrés	(10)	33,24	{=SOMME((Reg!D4:D28-Reg!E4:E28)^2)}
22		Somme totale des carrés		228,16	{=SOMME((Reg!D4:D28-$E$8)^2)}
23				228,16	=SOMME.CARRES.ECARTS(Reg!D4:D28)
25		Somme reg. des carrés	(5)	194,92	=E22-E21
27		Coefficient de détermination	(3)	0,85	=COEFFICIENT.DETERMINATION(Reg!D4:D28;Reg!C4:C28)
28				0,85	=E25/E22
29				0,85	=E12^2*E5/E9
30				0,85	=E2^2
32		Erreur type pour la valeur y estimée	(8)	1,20	=ERREUR.TYPE.XY(Reg!D4:D28;Reg!C4:C28)
33				1,20	=RACINE((1/23)*((E9*25)-((COVARIANCE.PEARSON(Reg!D4:D28;Reg!C4:C28)*25)^2/(E5*25))))
34				1,20	{=RACINE(SOMME((Reg!D4:D28-Reg!E4:E28)^2)/E19)}
36		Erreur type du coef. m	(2)	0,05	=RACINE((E34^2)/(25*E5))

Figure 13–25 Les formules de la plage G12:G36 fournissent individuellement certains résultats relatifs à la droite de régression. Les cellules E21, E22 et E34 contiennent des formules matricielles. Il faut donc les valider en pressant les touches Ctrl+Maj+Entrée.

La syntaxe de toutes les formules entrées en *E2:E36* est donnée en colonne G.

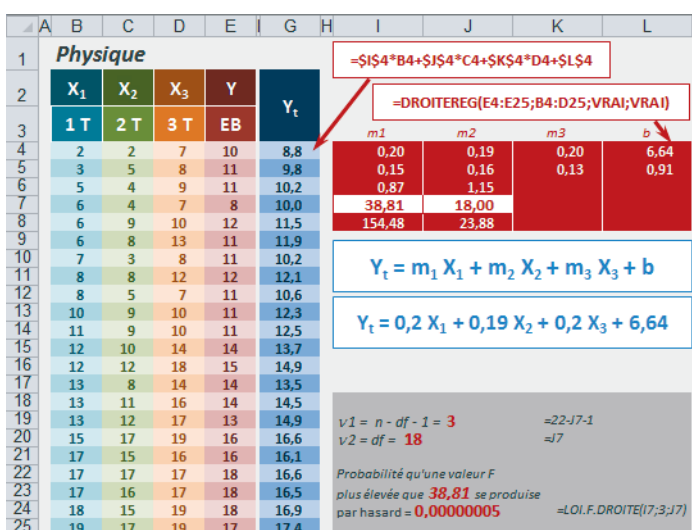
- La plage *E12:E14* présente trois formules ayant pour résultat la pente de la droite de régression. Parmi elles, on trouve la fonction *PENTE* qui renvoie directement `0,6` à partir des valeurs des deux variables Y et X. La pente peut également être obtenue en faisant le produit du coefficient de corrélation par l'écart-type de Y, divisé par l'écart-type de X.

- La plage *E16:E17* contient deux formules donnant l'ordonnée à l'origine de la droite de régression. Parmi elles, la fonction *ORDONNEE.ORIGINE* renvoie directement `6,62` à partir des valeurs des deux variables Y et X.
- La plage *E21:E25* présente les trois résultats : `somme totale des carrés`, `somme résiduelle des carrés` et `somme de régression des carrés` déjà détaillés plus haut. Notons qu'Excel fournit la fonction *SOMME.CARRES.ECARTS* qui renvoie directement la somme totale des carrés.
- La plage *E27:E30* présente quatre formules pour obtenir le coefficient de détermination. Parmi elles, la fonction *COEFFICIENT.DETERMINATION* renvoie directement `0,85` à partir des valeurs des deux variables Y et X. La formule de la cellule *E30* permet de vérifier qu'il s'agit bien du coefficient de corrélation élevé au carré.
- La plage *E32:E34* donne trois méthodes pour obtenir l'`erreur type pour la valeur y estimée`. Parmi elles, la fonction *ERREUR.TYPE.XY* renvoie directement `1,2` à partir des valeurs des deux variables Y et X.
- La formule de la cellule *E36* montre comment l'`erreur type du coefficient m` peut être obtenue à partir de l'`erreur type pour la valeur y estimée` et la variance de X.

Régression multiple

Dans la section précédente, à travers l'équation `Yt = m X + b`, nous avons tenté d'expliquer une variable Y à partir d'une variable X unique.

Nous allons à nouveau mettre en œuvre la fonction *DROITEREG*, mais pour tenter cette fois-ci d'expliquer, à travers l'équation $Y_t = m_1 X_1 + m_2 X_2 + m_3 X_3 + b$, une variable Y (les notes obtenues à l'examen blanc) à partir de plusieurs variables X (les trois notes trimestrielles). Nous poursuivons notre étude en nous intéressant toujours à la physique.



Physique

	X_1 1 T	X_2 2 T	X_3 3 T	Y EB	Y_t
4	2	2	7	10	8,8
5	3	5	8	11	9,8
6	5	4	9	11	10,2
7	6	4	7	8	10,0
8	6	9	10	12	11,5
9	6	8	13	11	11,9
10	7	3	8	11	10,2
11	8	8	12	12	12,1
12	8	5	7	11	10,6
13	10	9	10	11	12,3
14	11	9	10	11	12,5
15	12	10	14	14	13,7
16	12	12	18	15	14,9
17	13	8	14	14	13,5
18	13	11	16	14	14,5
19	13	12	17	13	14,9
20	15	17	19	16	16,6
21	17	15	16	16	16,1
22	17	17	17	18	16,6
23	17	16	17	18	16,5
24	18	15	19	18	16,9
25	19	17	19	17	17,4

=$I$4*B4+$J$4*C4+$K$4*D4+$L$4

=DROITEREG(E4:E25;B4:D25;VRAI;VRAI)

m1	m2	m3	b
0,20	0,19	0,20	6,64
0,15	0,16	0,13	0,91
0,87	1,15		
38,81	18,00		
154,48	23,88		

$Y_t = m_1 X_1 + m_2 X_2 + m_3 X_3 + b$

$Y_t = 0{,}2\ X_1 + 0{,}19\ X_2 + 0{,}2\ X_3 + 6{,}64$

v1 = n - df - 1 = 3 =22-J7-1
v2 = df = 18 =J7

Probabilité qu'une valeur F plus élevée que 38,81 se produise par hasard = 0,00000005 =LOI.F.DROITE(I7;3;J7)

Figure 13–26
La fonction DROITEREG renvoie une matrice de quatorze valeurs (I4:L8) qui définissent l'équation $Y_t = 0{,}2\ X_1 + 0{,}19\ X_2 + 0{,}2\ X_3 + 6{,}64$.

Les paramètres renvoyés répondent à la même logique que celle décrite pour la régression simple, mais ils sont adaptés à la régression multiple.

> **EN PRATIQUE** **Soigner la taille de la matrice**
>
> Avant d'entrer la fonction, il faut sélectionner une plage suffisante (*I4:L8*) pour qu'Excel ait la place d'afficher tous les résultats (le nombre de lignes reste inchangé, c'est le nombre de colonnes qui évolue). Avec trois variables `X`, il faut sélectionner quatre colonnes (trois pour renvoyer les coefficients m_1, m_2 et m_3 et la quatrième pour renvoyer la valeur de `b`). Sur la deuxième ligne de la matrice, la fonction renvoie également quatre erreurs types (trois s'appliquant aux coefficients m_1, m_2, m_3 et la quatrième concernant `b`). Les six autres cellules (*I6:J8*) affichent les mêmes paramètres que dans le cadre d'une régression simple. Les cellules *K6:L8* inutilisées par *DROITEREG* affichent normalement des valeurs d'erreur `#N/A`, mais pour améliorer la lisibilité de la figure, on les a masquées avec une mise en forme conditionnelle. Si l'utilisateur ne sélectionne pas une matrice de taille suffisante, Excel ne renverra qu'une partie des résultats attendus.

Les deux paramètres `Valeur F observée` (*I7* = `38,81`) et `df` (*J7* = `18`) peuvent être utilisés pour tester la fiabilité du coefficient de détermination. Plus ce dernier est proche de `1`, plus la corrélation entre les variables `X` et la variable `Y` est forte et plus l'équation de régression peut être utilisée avec confiance pour prévoir une valeur `y` quelconque.

Dans notre exemple, le coefficient de détermination est égal à `0,87`. `0,87` étant assez proche de `1`, on est tenté de penser que l'équation trouvée peut fournir un bon modèle de prévision. Néanmoins, comme pour tout indicateur statistique, il est précieux de l'associer à une probabilité qui quantifiera le risque couru en faisant confiance au modèle. Les valeurs `F` et `df` sont là pour nous aider dans cette tâche.

La `Valeur F observée` renvoyée par *DROITEREG* est une valeur butoir à mettre en perspective avec la loi de Fisher pour évaluer la probabilité d'obtenir une valeur `F` supérieure par hasard. Pour rapprocher `38,81` des valeurs renvoyées par la loi de Fisher, vous pouvez vous reporter aux tables statistiques (qui, avec les degrés de liberté `3` et `18` et pour un risque assumé α égal à `5 %`, donnent la valeur `3,16`) ou utiliser la fonction Excel *LOI.F.DROITE* (voir la fin de ce chapitre sur les distributions théoriques).

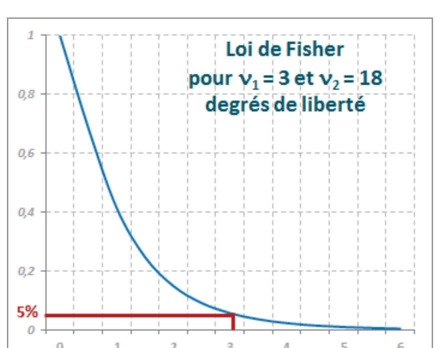


Figure 13–27
Pour tracer cette courbe, on a utilisé la fonction Excel LOI.F.DROITE avec $\nu_1 = 3$ et $\nu_2 = 18$ degrés de liberté. Toutes les valeurs de F supérieures à 3,16 correspondent à un coefficient significatif avec un risque d'erreur inférieur à 5 %.

COMPRENDRE Degrés de liberté

Le calcul de *LOI.F.DROITE* (loi de Fisher) s'effectue en fonction des degrés de liberté ν_1 et ν_2. Dans le cadre de l'étude des lois théoriques, nous reviendrons un peu plus loin sur la notion de degré de liberté. Pour l'instant, sachez simplement que la loi de Fisher prend en compte les deux paramètres ν_1 et ν_2 pour renvoyer des valeurs différentes. Dans le cadre de notre droite de régression, les valeurs à retenir sont ν_1 `= nb. Individus - df - 1 = 22 - 18 - 1 = 3` et ν_2 `= df = 18`.
`df` est calculé à partir du nombre d'observations (`22` dans notre exemple, puisqu'on a supprimé les données des élèves pour lesquels il manquait une note trimestrielle), auquel on soustrait le nombre de variables `X` (`3` dans notre exemple), puis `1`. On a donc `df = 22 - 3 - 1 = 18`.

À la figure 13-27, on a matérialisé la limite correspondant au risque habituellement accepté de 5 %. Il correspond bien à ce que vous pouvez lire dans les tables statistiques, à savoir une valeur qui tourne autour de `3,16`.

`LOI.F.DROITE(38,81;3;18)` renvoie `0,00000005` qui est une probabilité très faible (ce qui est cohérent avec le graphique puisque `38,81` se trouve bien au-delà de `6`, dernière valeur représentée). Vous pouvez donc en conclure que « l'explication » de la note à l'examen blanc par les trois notes trimestrielles à travers l'équation $Y_t = 0{,}2\ X_1 + 0{,}19\ X_2 + 0{,}2\ X_3 + 6{,}64$ n'est pas sans fondement.

ALLER PLUS LOIN Fonction LOGREG

Si, au lieu de dessiner une droite, votre nuage de points suit une courbe exponentielle, vous obtiendrez une meilleure modélisation en utilisant la fonction *LOGREG*. Cette dernière renvoie la même matrice de paramètres que *DROITEREG* mais, dans ce cas, les coefficients décrivent l'équation affichée figure 13-28.

$$Y_t = b\, m_1^{X_1}\, m_2^{X_2}\ \ldots\ m_n^{X_n}$$

Figure 13–28 Équation de la courbe de régression construite à partir des valeurs renvoyées par la fonction LOGREG.

Les méthodes utilisées pour tester l'équation obtenue sont les mêmes que pour la fonction DROITEREG. Elles sont calquées sur le modèle linéaire suivant : $\ln(Y) = \ln(b) + X_1 \ln(m_1) + X_2 \ln(m_2) + \ldots + X_n \ln(m_n)$. Ainsi, les erreurs types renvoyées dans la deuxième ligne de la matrice résultat doivent être rapprochées de ln(m1), ln(m2), … ln(b) et non de m1, m2 ou b.

Faire des prévisions

La régression étudiée dans les sections précédentes permet de modéliser une variable à travers une équation. La plupart du temps, l'objectif de ce modèle est de faire des prévisions. Comme on l'a fait dans la plage *E4:E28* de la figure 13-24, il est possible d'appliquer l'équation à diverses valeurs de `X` pour obtenir toutes les valeurs `Y` théoriques souhaitées. Néanmoins, Excel propose une autre panoplie de fonctions capables de renvoyer directement ces valeurs théoriques.

À *DROITEREG* et *LOGREG* correspondent respectivement les fonctions *TENDANCE* et *CROISSANCE*. Il s'agit de fonctions matricielles qu'il convient de valider en pressant simultanément les touches *Ctrl+Maj+Entrée*.

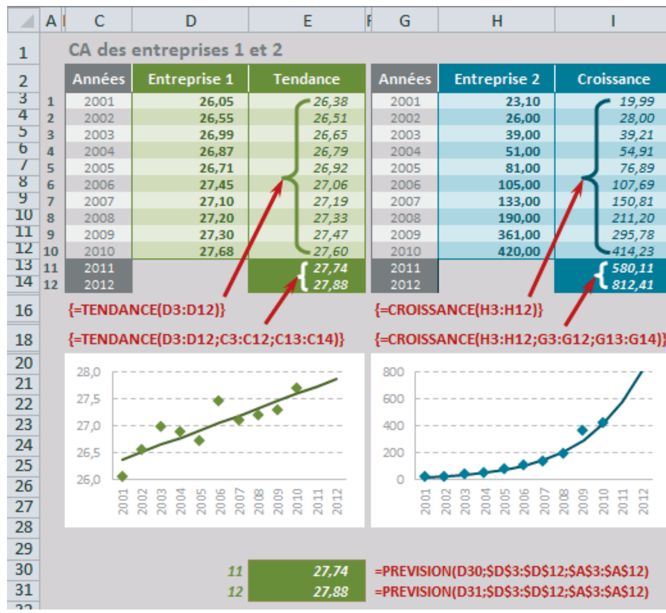


Figure 13–29
Les fonctions TENDANCE et CROISSANCE calculent des valeurs prévisionnelles à partir de trois matrices de données : les Y connus, les X connus et les X pour lesquels on souhaite la prévision.

Tableau 13–6 Faire des prévisions

Fonction	Description
TENDANCE et *PREVISION*	La plage *D3:D12* contient les chiffres d'affaires de l'entreprise 1 sur les dix dernières années. La fonction *TENDANCE* a été entrée une première fois dans la plage *E3:E12* pour calculer les valeurs Y théoriques correspondant à la droite de régression. Sa syntaxe apparaît en rouge sous le tableau. Elle a été entrée une deuxième fois dans la plage *E13:E14* en précisant les deuxième et troisième arguments qui désignent les valeurs de X connues et celles pour lesquelles on souhaite une prévision. Si vous souhaitez obtenir un seul Y prévisionnel, vous pouvez également utiliser la fonction *PREVISION* (plage *E30:E31* de la figure 13-29). Dans ce cas, il faut simplement veiller à ce que les valeurs de X soient exprimées sous la forme 1, 2, 3, etc. (plage *A3:A14* de la figure 13-29). La syntaxe des fonctions est donnée en *G30* et *G31*.
CROISSANCE	La plage *H3:H12* contient les chiffres d'affaires de l'entreprise 2 sur les dix dernières années. Sa croissance présentant un caractère exponentiel, c'est la fonction *CROISSANCE* qui semblait la mieux adaptée pour modéliser sa distribution. Elle a été entrée deux fois, d'abord avec un argument dans la plage *I3:I12*, puis en précisant les deuxième et troisième arguments (les X connus et ceux pour lesquels on souhaite une prévision) dans la plage *I13:I14*.

Distributions théoriques

COMPRENDRE Fonction de densité

La figure 13-30 affiche six représentations des variables étudiées au début de ce chapitre. Pour chacune, on a tracé, à partir de l'histogramme, le polygone des fréquences qui donne une approximation continue de la variable.

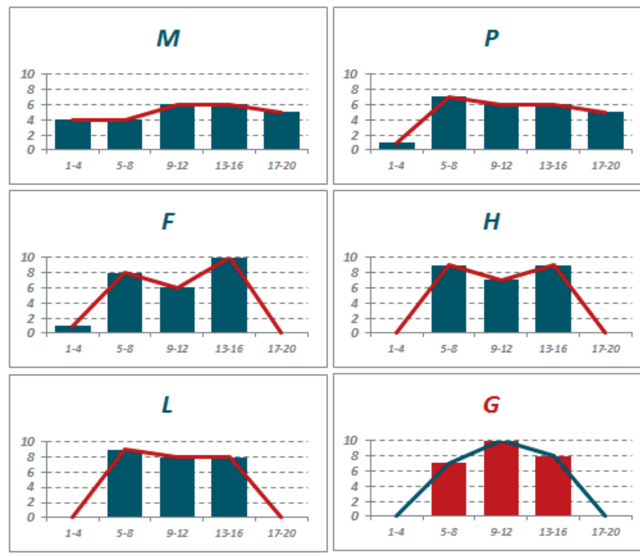


Figure 13–30 Représentation de la distribution des cinq variables « Moyenne par matière » et de la variable « Moyenne générale ».

En théorie, la distribution d'une variable continue est déterminée par une fonction réelle appelée fonction de densité. Cette dernière prend des valeurs positives dans l'intervalle associé à l'ensemble des modalités de la variable et des valeurs nulles ailleurs. La fonction de densité sert à calculer la proportion d'individus pour lesquels la variable prend une valeur entre deux nombres quelconques a et b.

Cette proportion est égale à l'aire située sous la courbe définie par la fonction de densité, entre les nombres a et b. L'ensemble des modalités de la variable (en théorie de -∞ à +∞ et en pratique, le vecteur pour lequel la variable est définie et non nulle) correspond à 100 % de la population. L'aire sous la courbe est donc égale à 1. La figure 13-31 illustre une forme possible de fonction de densité.

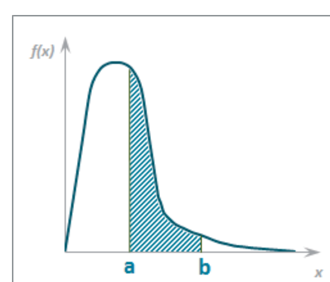


Figure 13–31 Forme possible d'une fonction de densité. L'aire hachurée de bleu représente la proportion d'individus pour lesquels la variable prend une valeur entre a et b.

Dans les sections précédentes, nous avons abordé les fonctions statistiques (moyennes, écarts-types, quartiles, etc.) servant à étudier de façon détaillée les variables quantitatives observées.

Pour de nombreuses raisons, il peut être utile de disposer de distributions théoriques. Ces dernières permettent de :

- caractériser une distribution observée en constatant qu'elle est très proche d'une distribution théorique de référence ;
- évaluer la vraisemblance d'hypothèses, en confrontant la distribution observée et la distribution théorique attendue ;
- calculer une probabilité approchée en utilisant une distribution théorique, jumelle de celle du phénomène observé (probabilités associées à un indicateur statistique).

Ces entités « magiques » sont proposées par la théorie des probabilités, dans laquelle la grandeur étudiée s'appelle variable aléatoire et la distribution associée s'appelle loi de la variable aléatoire. Excel propose une quarantaine de fonctions couvrant une douzaine de lois.

Variable aléatoire

Il s'agit d'une variable à partir de laquelle on peut déduire une distribution de fréquences probables (ou distribution de probabilités), par opposition à une variable statistique à partir de laquelle on établit des distributions de fréquences observées.

Tout comme pour une variable statistique, on peut calculer les caractéristiques de tendance centrale et de dispersion d'une variable aléatoire. Les plus usuelles sont la moyenne et la variance. La variance ne change pas de nom, mais la moyenne est appelée espérance, pour illustrer le fait qu'elle représente une moyenne possible et non une moyenne de valeurs observées.

Lois de probabilités discrètes

Les variables discrètes prennent un nombre fini de valeurs différentes. La distribution d'une variable discrète observée est caractérisée par le pourcentage d'individus associé à chaque modalité. Avec une variable discrète aléatoire, ces pourcentages sont remplacés par des probabilités : nombres compris entre `0` et `1` et dont la somme vaut `1`.

Loi binomiale

La loi binomiale est la variable aléatoire discrète la plus utilisée en statistiques.

La loi binomiale modélise ce que l'on appelle un tirage « avec remise » par opposition au tirage « sans remise ». Son espérance est égale à `np`. Dans l'exemple présenté figure 13-33, on a bien `4 × 0,5 = 2` qui correspond au calcul obtenu cellule *H22* (syntaxe de la formule en *J22*). Sa variance est égale à `npq`. Dans l'exemple présenté figure 13-33, on a bien `4 × 0,5 × 0,5 = 1` qui correspond au calcul obtenu cellule *H23* (syntaxe de la formule en *J23*).

COMPRENDRE Les fondements de cette loi

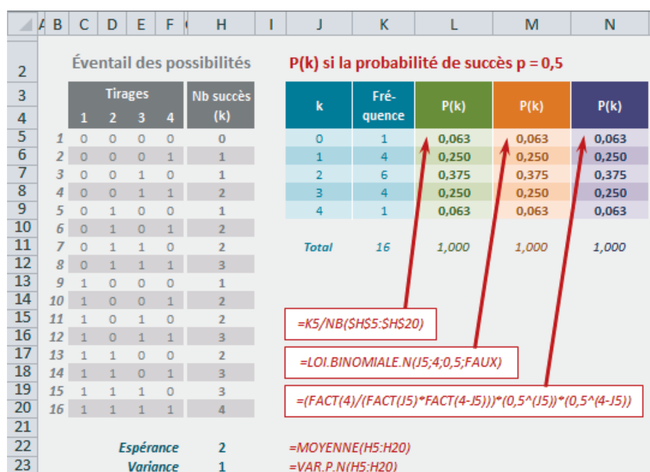
Une expérience aléatoire est dite épreuve de Bernoulli si l'ensemble de ses résultats peut se résumer à deux états portant le nom de succès et d'échec (lancer de pièce, réponse ou non-réponse à un questionnaire, etc.). À partir de la probabilité de succès, notée p, on déduit la probabilité d'échec, q = 1 - p. Prenons l'exemple d'une pièce de monnaie jetée en l'air n fois. Elle va retomber k fois du côté face (que l'on considère arbitrairement comme k succès) et n - k fois du côté pile (n - k échecs).
La loi binomiale est parfaitement adaptée à ce genre de problématique. Elle donne la probabilité d'obtenir k succès, en d'autres termes de connaître P(k). Les règles mathématiques de calcul des probabilités aboutissent à la formule exprimée figure 13-32.

Figure 13–32 On appelle loi binomiale l'ensemble des valeurs de P(k). En réalité, cette formule définit plutôt une famille de lois, chacune étant déterminée par une valeur de n et de p.

$$P(k) = \frac{n!}{k!\,(n-k)!}\,p^k(1-p)^{n-k}$$

Pour mieux comprendre le lien entre les épreuves de Bernoulli et cette formule, voici un exemple concret. Considérons une pièce de monnaie parfaitement équilibrée, c'est-à-dire ayant une probabilité p = 0,5 de tomber du côté face et, donc, une probabilité q = 1 - p = 0,5 de tomber du côté pile. Imaginons maintenant quatre lancers successifs de cette pièce et considérons tous les résultats possibles. Le tableau gris situé à gauche de la figure 13-33 liste les 16 situations envisageables. En colonne *C*, on a indiqué les résultats du premier lancer, en colonne *D*, les résultats du second lancer et ainsi de suite. La valeur 1 symbolise le côté face et 0 le côté pile. En colonne *H*, on a simplement dénombré le nombre de fois où le côté face a été obtenu, c'est-à-dire k, le nombre de succès.

Figure 13–33 Approche concrète de la loi binomiale à travers l'exemple de quatre lancers successifs d'une pièce de monnaie.



Éventail des possibilités

	Tirages 1	2	3	4	Nb succès (k)
1	0	0	0	0	0
2	0	0	0	1	1
3	0	0	1	0	1
4	0	0	1	1	2
5	0	1	0	0	1
6	0	1	0	1	2
7	0	1	1	0	2
8	0	1	1	1	3
9	1	0	0	0	1
10	1	0	0	1	2
11	1	0	1	0	2
12	1	0	1	1	3
13	1	1	0	0	2
14	1	1	0	1	3
15	1	1	1	0	3
16	1	1	1	1	4
				Espérance	2
				Variance	1

=MOYENNE(H5:H20)
=VAR.P.N(H5:H20)

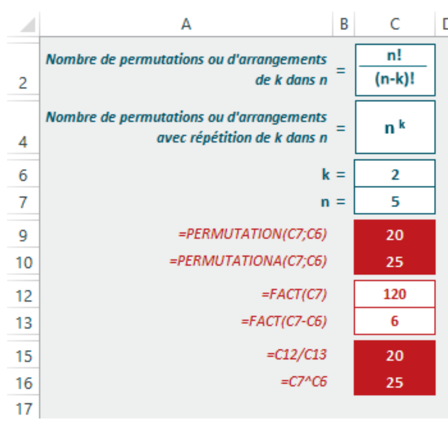
P(k) si la probabilité de succès p = 0,5

k	Fréquence	P(k)	P(k)	P(k)
0	1	0,063	0,063	0,063
1	4	0,250	0,250	0,250
2	6	0,375	0,375	0,375
3	4	0,250	0,250	0,250
4	1	0,063	0,063	0,063
Total	16	1,000	1,000	1,000

=K5/NB($H$5:$H$20)
=LOI.BINOMIALE.N(J5;4;0,5;FAUX)
=(FACT(4)/(FACT(J5)*FACT(4-J5)))*(0,5^(J5))*(0,5^(4-J5))

Dans le tableau de droite, on a listé en colonne *J* toutes les valeurs possibles de k. Ces dernières vont de 0 (4 fois côté pile) à 4 (4 fois côté face). En colonne *K*, on a calculé les fréquences obtenues dans le tableau gris pour chaque valeur de k. En colonne *L*, on a simplement fait le rapport entre ces fréquences et le nombre total de situations possibles, ce qui donne la probabilité d'obtenir chaque valeur de k, c'est-à-dire P(k). Dans les colonnes suivantes (*M* et *N*), on a appliqué la fonction Excel *LOI.BINOMIALE.N* et l'algorithme qui la sous-tend (figure 13-32) afin de vérifier que les trois méthodes de calcul renvoyaient bien la même probabilité.
Il faut bien entendu insister sur le fait que ces chiffres supposent que la pièce est parfaitement équilibrée (p = 0,5). Dans le cas inverse, il faudrait utiliser une autre valeur de p. Si la pièce était déséquilibrée du côté pile, il faudrait utiliser une valeur p < 0,5 et, dans le cas inverse, une valeur p > 0,5.

Plus n est grand et plus le calcul de P(k) est lourd. Généralement, au-delà d'un certain seuil (dès que $np > 5$ et $nq > 5$) on fait appel à la loi normale de même moyenne (np) et même variance (npq). Lorsque p est très petit ($p < 0,1$), on utilise la loi de Poisson qui dépend d'un paramètre unique noté λ, réel strictement positif ($\lambda = np$).

> **OUPS Combinaisons, arrangements, permutations**
>
> Ce petit aparté concerne le classement un peu étrange de certaines fonctions dans Excel 2010 et Excel 2013. Dans le chapitre 12, nous avions présenté les fonctions *COMBIN* et *COMBINA*, toutes deux rangées dans les fonctions Mathématiques ! La fonction *COMBIN* s'appuie sur l'algorithme `n! / (k! (n - k)!)` (combinaison de k éléments dans n) qui correspond bien au début de la formule affichée figure 13-32.
>
> Dans le cadre de la présentation de cette fonction, nous avions évoqué la notion d'arrangements qui calcule la même chose que la fonction *COMBIN*, mais en tenant compte de l'ordre des k éléments, l'algorithme devenant alors `n! / (n - k)!`. Or, Excel offre (mais cette fois-ci dans la catégorie Statistiques !) la fonction *PERMUTATION* qui s'appuie sur cet algorithme. Sous Excel 2013, vous bénéficiez en plus de la fonction *PERMUTATIONA* qui renvoie les arrangements de k éléments dans n, mais avec répétitions, selon l'algorithme n^k.
>
> 
>
>
> **Figure 13–34** Mise en œuvre des fonctions PERMUTATION et PERMUTATIONA (pour ceux qui ont Excel 2013). La syntaxe des formules entrées en C9:C16 est présentée en A9:A16.

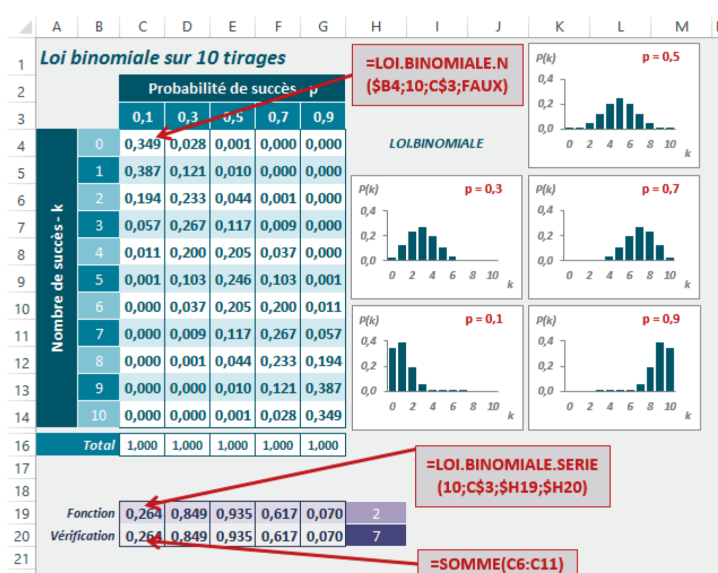
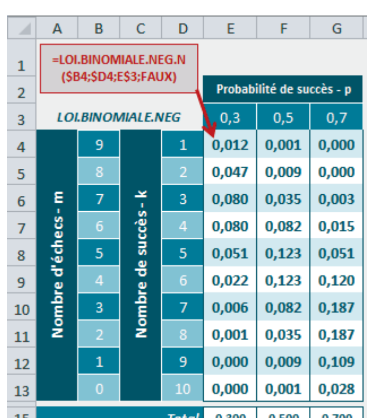


Figure 13–35 Mise en œuvre de la fonction LOI.BINOMIALE.N. Si vous travaillez sous Excel 2013, vous disposez d'une nouvelle fonction, LOI.BINOMIALE.SERIE.

Tableau 13–7 Loi binomiale

Fonction	Description
LOI.BINOMIALE.N	La plage *C4:G14* de la figure 13-35 donne un exemple d'utilisation de la fonction *LOI.BINOMIALE.N* pour `n = 10` et `p = 0,1`, `p = 0,3`, `p = 0,5`, `p = 0,7` et `p = 0,9`. Les cinq représentations graphiques correspondantes apparaissent juste à côté. *LOI.BINOMIALE* indiquée en italique est l'ancienne forme de la fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.
LOI.BINOMIALE .SERIE	Utilisée avec seulement ses trois premiers arguments, cette fonction renvoie exactement les mêmes résultats que la fonction *LOI.BINOMIALE.N*. Si vous précisez le quatrième argument, vous définissez une fourchette de succès dont vous cherchez la probabilité. Dans la figure 13-35, cette fonction est testée en ligne *19*, avec *2* et *7* comme troisième et quatrième arguments. La fonction entrée en ligne *20* montre que calculer une probabilité à l'aide de *LOI.BINOMIALE.SERIE* revient à faire une somme des probabilités renvoyées par *LOI.BINOMIALE.N*. Dans l'exemple proposé figure 13-35, on constate que la probabilité retournée par la fonction *LOI.BINOMIALE.SERIE* pour la fourchette `[2,7]` correspond bien à la somme des probabiltés (renvoyées par *LOI.BINOMIALE.N*) d'avoir 2, 3, 4, 5, 6 ou 7 succès. **Nouveauté Excel 2013**.
LOI.BINOMIALE .NEG.N	La plage *E4:G13* de la figure 13-36 propose une application de la fonction *LOI.BINOMIALE.NEG.N* pour `p = 0,3`, `p = 0,5` et `p = 0,7`. Cette fonction renvoie la probabilité de passer par `m` échecs avant d'obtenir `k` succès. *LOI.BINOMIALE.NEG* indiquée en italique est l'ancienne forme de la fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.



	A	B	C	D	E	F	G
1	=LOI.BINOMIALE.NEG.N ($B4;$D4;E$3;FAUX)						
2					Probabilité de succès - p		
3	*LOI.BINOMIALE.NEG*				0,3	0,5	0,7
4	Nombre d'échecs - m	9	Nombre de succès - k	1	0,012	0,001	0,000
5		8		2	0,047	0,009	0,000
6		7		3	0,080	0,035	0,003
7		6		4	0,080	0,082	0,015
8		5		5	0,051	0,123	0,051
9		4		6	0,022	0,123	0,120
10		3		7	0,006	0,082	0,187
11		2		8	0,001	0,035	0,187
12		1		9	0,000	0,009	0,109
13		0		10	0,000	0,001	0,028
15				*Total*	0,300	0,500	0,700

Figure 13–36
Mise en œuvre de la fonction LOI.BINOMIALE.NEG.N. Cette fonction applique la formule indiquée figure 13-37.

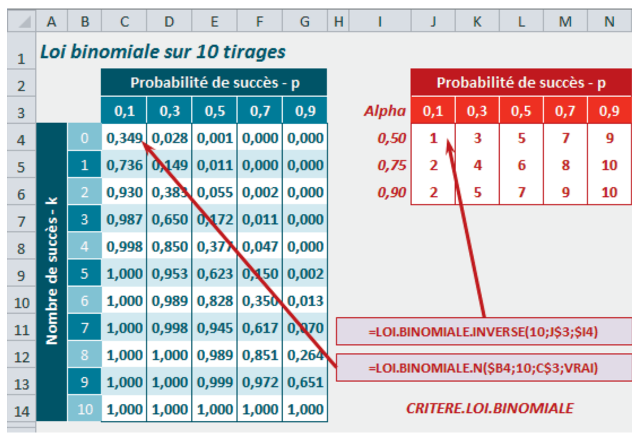
Figure 13–37
Formule définissant la fonction LOI.BINOMIALE.NEG.N. Ici, k représente toujours le nombre de succès et m, le nombre d'échecs.

$$P(k) = \frac{(k+m-1)!}{(k-1)!\,m!}\,p^k(1-p)^m$$

Tableau 13–8 Loi binomiale

Fonction	Description
LOI.BINOMIALE .INVERSE	La fonction *LOI.BINOMIALE.INVERSE* renvoie la plus petite valeur de `k` pour laquelle la distribution binomiale cumulée est supérieure ou égale à un critère (`alpha`). La plage *C4:G14* de la figure 13-38 accueille à nouveau la fonction *LOI.BINOMIALE.N*, mais cette fois-ci, elle affiche les probabilités cumulées. La plage *J4:N6* affiche plusieurs résultats possibles de la fonction *LOI.BINOMIALE .INVERSE*, toujours pour `10` tirages, `5` probabilités de succès différentes et `3` valeurs `alpha`. *CRITERE.LOI.BINOMIALE* est l'ancienne forme de cette fonction. Elle est conservée pour assurer la compatibilité avec les versions antérieures.

Figure 13–38
Mise en œuvre de la fonction LOI.BINOMIALE.INVERSE.



Pratique **Fonction PROBABILITE**

À partir du moment où vous avez réuni dans un tableau les modalités d'une variable et les probabilités correspondantes, vous pouvez utiliser la fonction *PROBABILITE* pour renvoyer la probabilité d'une modalité ou d'une classe de modalités.

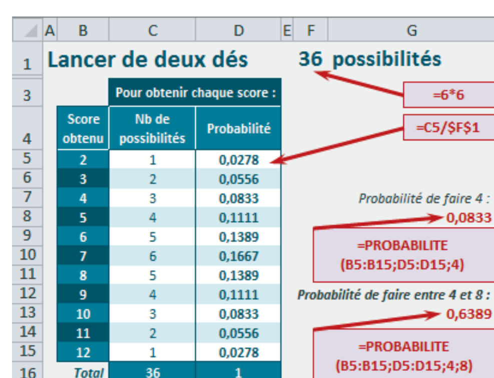


Figure 13–39
Mise en œuvre de la fonction PROBABILITE.

L'exemple utilisé pour illustrer cette fonction est une variable qui associe l'ensemble des scores qu'il est possible d'obtenir en lançant deux dés (`11` entiers compris entre `2` et `12`) avec la probabilité associée à chaque modalité. On part du principe que les dés ne sont pas pipés. On a donc une probabilité de `1/6` d'obtenir chaque face. Le nombre total de lancers possibles est donné par le produit entré en *F1* (`= 6 × 6`).

La probabilité d'obtenir 2 est renvoyée par la formule entrée en *D5*. Sa syntaxe est indiquée en rouge dans le cadre gris. La plage *C5:C15* donne, pour chaque score, le total des combinaisons permettant de l'obtenir. La formule *D5* a été recopiée dans la plage *D6:D15*.
Les cellules *G8* et *G13* donnent respectivement la probabilité d'obtenir 4 et celle d'obtenir un score compris entre 4 et 8 (cumul des cellules *D7:D11*). La syntaxe des formules correspondantes apparaît dans les cadres gris.

Loi hypergéométrique

Contrairement à la loi binomiale, la loi hypergéométrique modélise des tirages « sans remise ».

COMPRENDRE Tirages avec ou sans remise

Pour comprendre l'importance de cette notion de remise, prenons un cas concret. Imaginons un lycée réunissant 1 000 élèves (750 filles et 250 garçons). À partir de la population totale de ce lycée, on souhaite obtenir un échantillon de 10 individus réunissant 5 garçons. Quelle est la probabilité de réussir la construction de cet échantillon dans les proportions souhaitées ?
Il s'agit bien d'un tirage sans remise, car une fois un individu choisi, on ne peut pas le choisir une nouvelle fois. Si l'on représente les trois premières étapes de la construction de cet échantillon, on obtient la figure 13-40.

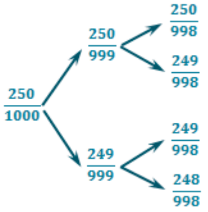


Figure 13–40 Probabilité de tirer un garçon dans les étapes 1 à 3 de la construction de l'échantillon.

Pour le premier tirage, on a une probabilité de `250/1 000` d'obtenir un garçon. Cependant, dès la deuxième étape, la situation se complique. En effet, si le premier tirage était une fille, on a, à la deuxième étape, la probabilité de `250/999` de tirer un garçon. En revanche, si le premier tirage était un garçon, cette probabilité devient `249/999`. La situation à la troisième étape se complique à nouveau et il en est ainsi jusqu'à la dixième étape.
Néanmoins, pour simplifier les calculs, on peut considérer que les valeurs `250/999` et `249/999` sont assez proches. De ce fait, on les assimile et on considère qu'à chaque étape, la probabilité d'obtenir un garçon est égale à `1/4`. On est donc ramené à la logique d'un tirage avec remise. Ainsi, dans un souci de simplification, à la place de la loi hypergéométrique, on utilise souvent une approximation par la loi binomiale. Nous verrons dans l'exemple présenté figure 13-43 que cette dernière est d'autant mieux adaptée que la population est très grande et l'échantillon très petit.

Elle dépend de trois paramètres : la taille de la population totale (`N`), le nombre d'individus (parmi les `N`) dotés de la propriété étudiée (`K`) et la taille de l'échantillon (`n`). Par analogie avec la loi binomiale, on note `p = K/N` (dans notre exemple, `p` est bien égal à `250/1 000`, c'est-à-dire `0,25`). Son espérance est égale à `np` (dans notre exemple, `10 × 0,25`, c'est-à-dire, `2,5`) et sa variance à (`npq`) `[( N - n )/( N - 1 )]` (dans notre exemple, `1,86`. Voir la figure 13-43).

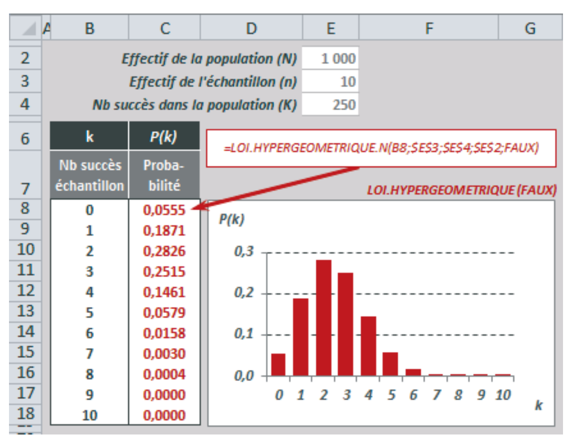
Figure 13–41
Formule définissant la fonction LOI.HYPERGEOMETRIQUE.N.

$$P(k) = \frac{\dfrac{K!}{k!\,(K-k)!}\;\dfrac{(N-K)!}{(n-k)!\,(N-K-n+k)!}}{\dfrac{N!}{n!\,(N-n)!}}$$

k : *Nombre de succès de l'échantillon*
n : *Effectif de l'échantillon*
K : *Nombre de succès de la population*
N : *Effectif de la population*

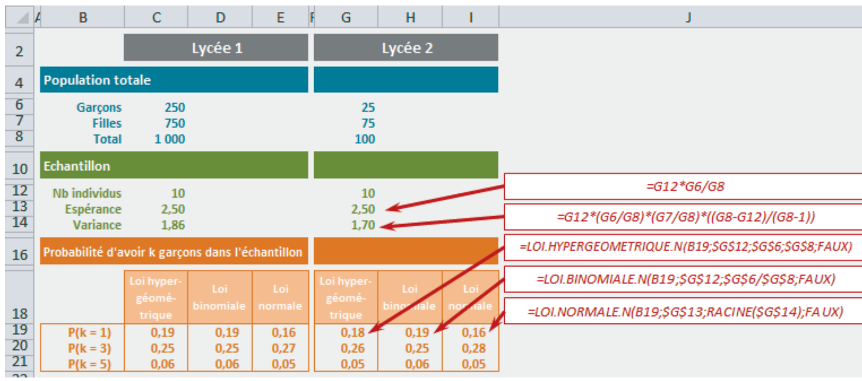
La fonction *LOI.HYPERGEOMETRIQUE.N* renvoie la probabilité d'obtenir `k` succès dans un échantillon (`n`) sachant que la population (`N`) connaît elle-même `K` succès pour cette même propriété. La figure 13-42 montre un exemple d'utilisation de cette fonction. Le taux de succès de la population est de `1/4` et sa représentation (pour `k = 1 à 10`) apparaît en rouge dans le graphique. La formule dont la syntaxe apparaît en rouge dans le cadre blanc a été entrée en `C8`, puis recopiée dans toute la colonne. *LOI.HYPERGEOMETRIQUE* est l'ancienne forme de cette fonction (avec le dernier argument égal à `FAUX`). Elle est conservée pour assurer la compatibilité avec les versions antérieures.

Figure 13–42
Mise en œuvre de la fonction LOI.HYPERGEOMETRIQUE.N.



COMPRENDRE Approximations de la loi hypergéométrique

En reprenant l'exemple proposé en début de section, à savoir la création d'un échantillon de 10 élèves à partir de la population totale d'un lycée, on a mis au point la figure 13-43.



	Lycée 1			Lycée 2		
Population totale						
Garçons	250			25		
Filles	750			75		
Total	1 000			100		
Echantillon						
Nb individus	10			10		
Espérance	2,50			2,50		
Variance	1,86			1,70		
Probabilité d'avoir k garçons dans l'échantillon						
	Loi hyper-géomé-trique	Loi binomiale	Loi normale	Loi hyper-géomé-trique	Loi binomiale	Loi normale
P(k = 1)	0,19	0,19	0,16	0,18	0,19	0,16
P(k = 3)	0,25	0,25	0,27	0,26	0,25	0,28
P(k = 5)	0,06	0,06	0,05	0,05	0,06	0,05

Figure 13–43 Approximations de la loi hypergéométrique par la loi binomiale et la loi normale.

La plage *C6:C14* reprend les données du problème initial et la plage *C19:C21* utilise la fonction *LOI.HYPERGEOMETRIQUE.N* pour renvoyer la probabilité d'obtenir 1, 3 ou 5 garçons dans un échantillon de 10 individus. On trouve, dans la plage *D19:D21*, une approximation de ces résultats avec la fonction *LOI.BINOMIALE.N* qui utilise pour p, la valeur K / N (250/1000). Dans la plage *E19:E21*, on a réalisé une deuxième approximation en utilisant la loi normale, avec comme espérance et variance, celles de la loi hypergéométrique (*C13* et *C14*). Dans la partie droite du tableau, on a fait exactement les mêmes calculs, mais en partant d'une population totale de 100 individus (au lieu de 1 000).
On constate qu'avec une population importante (1 000) et un échantillon modeste (10), l'approximation réalisée avec la loi binomiale est assez fiable. En revanche, avec une population plus modeste (100), cette approximation devient légèrement moins bonne, mais reste toutefois proche des résultats renvoyés par la loi hypergéométrique. L'échantillon choisi est trop petit pour que les approximations réalisées avec la loi normale soient de bonne qualité (il faudrait un minimum de 21 individus pour respecter le seuil $np > 5$).

Loi de Poisson

La loi de Poisson est définie par la formule présentée figure 13-44. Elle ne dépend que d'un paramètre, λ, l'espérance, réel strictement positif égal à np.

λ correspond également à la variance de cette loi.

Figure 13–44
Formule définissant la fonction LOI.POISSON.N.

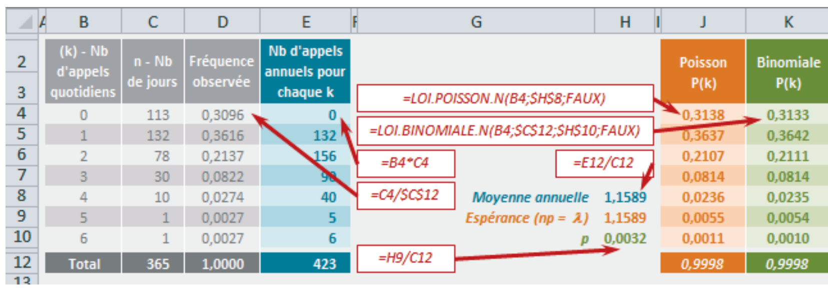
$$P(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

k : *Nombre de succès*
λ=np : *Espérance*

Elle est utilisée pour étudier les événements rares, cette rareté se traduisant par le fait que la probabilité que chaque événement survienne est faible (accidents, suicides d'enfants, mutations biologiques…).

EN PRATIQUE **Modéliser les appels téléphoniques d'un ermite**

Pour l'aborder plus en détail, prenons un exemple. Sur une année, on a comptabilisé les appels téléphoniques quotidiens reçus par un ermite (ces valeurs varient de 0 à 6). On les a consignées dans un tableau (partie gauche de la figure 13-45). En colonne *D*, on a pu en déduire les fréquences observées et, en colonne *E*, le nombre annuel d'appels pour chaque classe (pour chaque valeur de k), la somme donnant le total d'appels annuels.
En *H8*, on a fait le rapport du nombre total d'appels sur le nombre de jours pour connaître la moyenne d'appels quotidiens (1,1589).



	B	C	D	E	G	H	J	K
2-3	(k) - Nb d'appels quotidiens	n - Nb de jours	Fréquence observée	Nb d'appels annuels pour chaque k			Poisson P(k)	Binomiale P(k)
4	0	113	0,3096	0			0,3138	0,3133
5	1	132	0,3616	132			0,3637	0,3642
6	2	78	0,2137	156			0,2107	0,2111
7	3	30	0,0822	90			0,0814	0,0814
8	4	10	0,0274	40	Moyenne annuelle	1,1589	0,0236	0,0235
9	5	1	0,0027	5	Espérance (np = λ)	1,1589	0,0055	0,0054
10	6	1	0,0027	6	p	0,0032	0,0011	0,0010
12	Total	365	1,0000	423			0,9998	0,9998

Figure 13–45 Modélisation par une loi de Poisson et une loi binomiale des appels téléphoniques quotidiens reçus par un ermite.

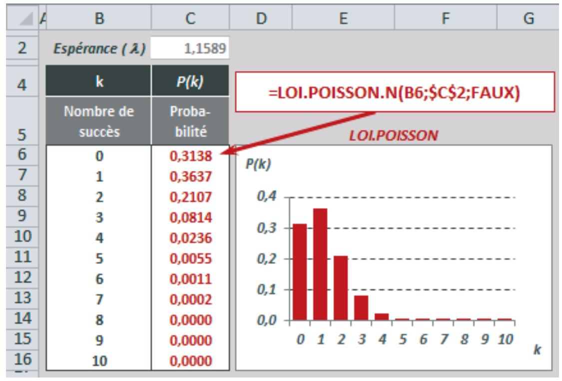
Pour modéliser cette distribution par une loi de Poisson, il faut d'abord vérifier que les conditions de cette modélisation sont remplies. Ces dernières sont :

- une valeur de $n > 50$ (ce qui est le cas puisque $n = 365$) ;
- une valeur de $np < 10$. Pour modéliser la distribution, on va adopter la moyenne de la variable observée (*H8*) comme espérance de la loi de Poisson : $\lambda = np = 1{,}1589$, qui est bien inférieure à 10 ;
- une valeur de $p < 0{,}1$. Connaissant np (1,1589) et n (365), on peut en déduire $p = np / n$, c'est-à-dire le contenu de la cellule *H10* (0,0032). p est bien inférieure à 0,1.

On peut donc utiliser la fonction *LOI.POISSON.N* avec les diverses valeurs de k (*B4:B10*) comme premier argument et 1,1589 comme valeur de λ. Exception faite des valeurs marginales ($k = 5$ et $k = 6$), les probabilités retournées dans la plage *J4:J10* reflètent assez bien les proportions observées dans la réalité. Dans la plage *K4:K10*, on a fait la même chose, mais en utilisant la fonction *LOI.BINOMIALE.N* avec $n = 365$ et $p = 0{,}0032$. À nouveau, on observe que les probabilités retournées sont assez fidèles à la réalité (en excluant toujours les deux valeurs marginales $k = 5$ et $k = 6$).

À la figure 13-46, la fonction *LOI.POISSON.N* a été mise en œuvre pour $\lambda = 1{,}1589$. Entrée une première fois dans la cellule *C6*, la formule dont la syntaxe apparaît en rouge dans le cadre blanc a été ensuite recopiée dans la plage *C7:C16*. Le graphique illustre les probabilités associées à chaque valeur de k. *LOI.POISSON* est l'ancienne forme de cette fonction. Elle est conservée pour assurer la compatibilité avec les versions antérieures.

Figure 13–46
Mise en œuvre de la fonction LOI.POISSON.N.



Lois de probabilités continues

Les variables dites « continues » prennent généralement toutes les valeurs d'un intervalle donné, borné ou non, de l'ensemble des nombres réels. L'ensemble des valeurs possibles est dit « domaine de variation ».

COMPRENDRE Variables continues et intervalles

En statistique descriptive, les valeurs prises par une variable continue observée (`X`) correspondent chacune à un petit intervalle. On obtient ainsi une fonction de densité `f(x)`, définie pour chaque valeur de `x` (en réalité, chaque mini-intervalle autour de `x`) du domaine de variation. La probabilité d'appartenir à un intervalle `[a, b]` est égale à la surface délimitée par la courbe au-dessus de cet intervalle.

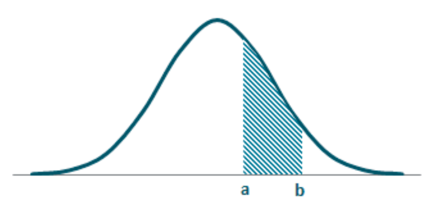


Figure 13–47 La surface hachurée sous la courbe représente la probabilité pour que x prenne sa valeur entre a et b.

La surface délimitée par la courbe toute entière (somme des probabilités de tous les événements) vaut `1`.

SYNTAXE Fonction de densité de probabilité, P(X = x) ou fonction de répartition, P(X < x)

Les fonctions présentées dans cette section et dans les sections suivantes utilisent pour la plupart un argument `Cumulative` qui peut prendre la valeur logique `VRAI` ou `FAUX`.

- En indiquant `FAUX`, on fait appel à la fonction de densité de probabilité. Cette dernière renvoie `P(X = x)`. Il s'agit de la probabilité que `X` prenne sa valeur « autour de » `x`, généralement dans l'intervalle `[x-0,5 , x+0,5]`. Le graphique correspondant pour la loi normale apparaît figure 13-48.
- En indiquant `VRAI`, on fait appel à la fonction de répartition. Cette dernière renvoie `P(X < x)`. Il s'agit de la probabilité que `X` prenne sa valeur entre `-x` et `x` ou, en d'autres termes, du cumul des probabilités de toutes les valeurs situées entre `-x` et `x`. Le graphique correspondant pour la loi normale apparaît figure 13-49.

Loi normale

Une variable aléatoire réelle `X` suit une loi normale d'espérance `m` et d'écart-type σ si elle admet pour densité de probabilité la fonction présentée figure 13-48 (deuxième cadre blanc). On note souvent cette variable `N(m,σ)`. Sa courbe représentative a une forme de cloche. La densité est surtout importante autour de la moyenne, puis décroît de façon symétrique d'autant plus rapidement que l'écart-type est petit.

> **USAGE Loi normale**
>
> La loi normale est sans doute la plus utile des lois de probabilités théoriques. En effet, elle permet de modéliser beaucoup de distributions statistiques observées et de décrire nombre de phénomènes aléatoires. Elle est aussi très souvent mise à contribution au sein de tests statistiques pour évaluer la fiabilité de certains résultats. Enfin, elle est régulièrement utilisée comme approximation de certaines lois (comme la loi binomiale quand `n > 30`).

Figure 13–48
Mise en œuvre de la fonction LOI.NORMALE.N. Le graphique illustre les valeurs de la plage C4:C14, qui affiche les résultats de LOI.NORMALE.N non cumulative (fonction de densité de probabilité) pour une espérance de 5, et un écart-type de 1,5.

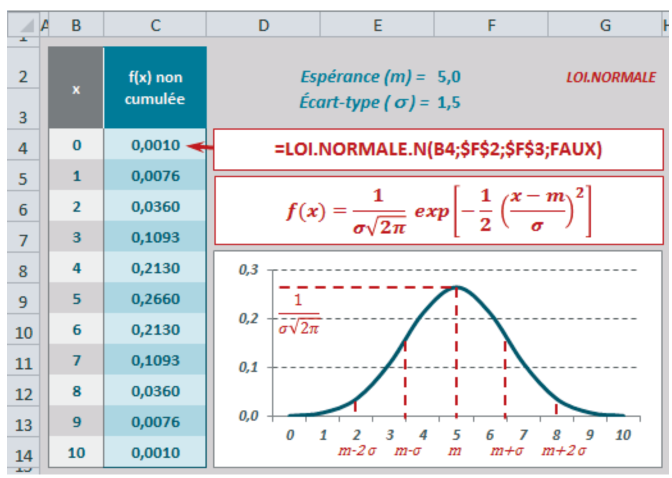


Tableau 13–9 Loi normale

Fonction	Description
LOI.NORMALE.N	`X` étant une variable aléatoire continue suivant une loi normale d'espérance `m` et d'écart-type σ, *LOI.NORMALE.N* renvoie, si le quatrième argument de la fonction est positionné sur `FAUX`, `P(X = x)`, probabilité que la variable `X` prenne la valeur `x` ou, plus précisément, probabilité que la variable `X` prenne sa valeur dans l'intervalle `[x-0,5 , x+0,5]`. La figure 13-48 illustre un exemple d'application de cette fonction pour `m = 5` et `σ = 1,5`. Sa syntaxe apparaît en rouge dans le premier cadre blanc. *LOI.NORMALE* correspond à l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et Excel 2013 pour assurer la compatibilité avec les versions antérieures.

Tableau 13–9 Loi normale (suite)

Fonction	Description
LOI.NORMALE .INVERSE.N	La figure 13-49 présente la mise en œuvre de deux fonctions : *LOI.NORMALE.N* dans sa version cumulée (plage *C4:C14* dont les valeurs sont illustrées dans le graphique) et *LOI.NORMALE.INVERSE.N* (plage *D4:D14*). La première renvoie `F(x)` à partir de `x`, c'est-à-dire la probabilité que `X` prenne sa valeur dans l'intervalle `[-x,x[`, alors que la deuxième fait l'opération inverse et renvoie `x` à partir de `F(x)`. Les formules, entrées d'abord en *C4* et *D4*, ont été recopiées dans la colonne. Leur syntaxe apparaît en rouge dans les deux cadres blancs. *LOI.NORMALE.INVERSE* correspond à l'ancienne forme de la fonction. Elle est conservée dans Excel 2010 et Excel 2013 pour assurer la compatibilité avec les versions antérieures.

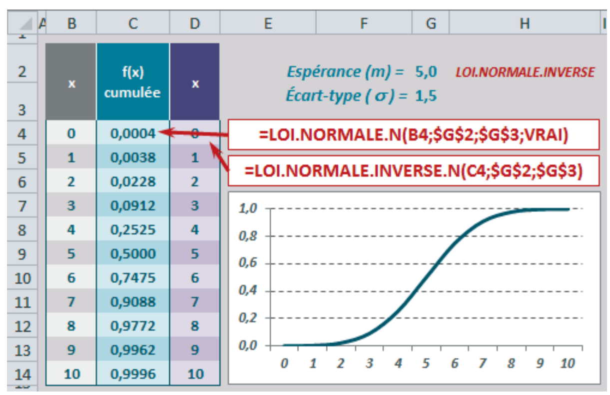


Figure 13–49
Mise en œuvre de la fonction LOI.NORMALE.N dans sa version cumulée et de la fonction LOI.NORMALE.INVERSE.N.

Loi normale centrée réduite

Une loi normale centrée réduite est une loi normale d'espérance nulle (centrée) et d'écart-type `1` (réduite). Sa fonction de densité de probabilité, par convention appelée `Z`, est présentée à la figure 13-50, dans le premier cadre blanc.

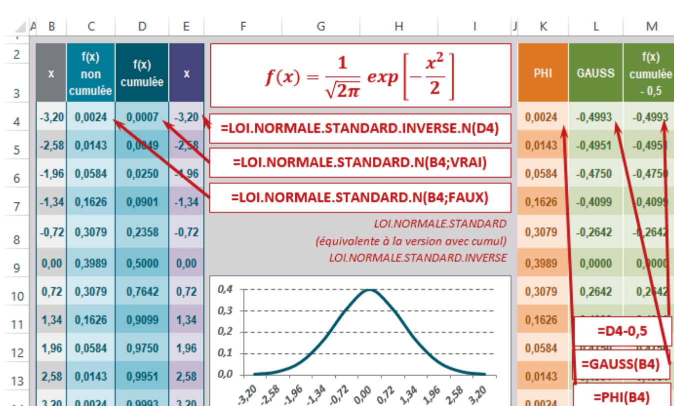


Figure 13–50
Mise en œuvre des fonctions LOI.NORMALE.STANDARD.N (dans ses deux versions : cumulative et non cumulative) et LOI.NORMALE.STANDARD. INVERSE.N. En relation avec la loi normale centrée réduite, Excel 2013 propose les deux nouvelles fonctions PHI et GAUSS.

> **Rappel Variable centrée réduite**
>
> Pour mieux comprendre cette notion, consultez, plus haut dans ce chapitre, l'aparté intitulé « Variable centrée réduite ».

Tableau 13–10 Loi normale centrée réduite

Fonction	Description
LOI.NORMALE.STANDARD.N	`Z` étant une variable aléatoire continue suivant une loi normale centrée réduite, *LOI.NORMALE.STANDARD.N* renvoie, si son deuxième argument est positionné sur `FAUX`, `P(Z = z)` (probabilité que la variable `Z` prenne la valeur `z`) et, si son deuxième argument est positionné sur `VRAI`, `P(Z < z)`. La figure 13-50 illustre un exemple d'application de cette fonction. Sa syntaxe apparaît en rouge dans les deux derniers cadres blancs. La plage *C4:C14* donne les résultats de la fonction dans sa version non cumulative : `P(Z = z)` (valeurs représentées sur le graphique) et la plage *D4:D14* dans sa version cumulative : `P(Z < z)`. *LOI.NORMALE.STANDARD* correspond à l'ancienne forme de cette fonction (avec cumul). Elle est conservée dans Excel 2010 et Excel 2013 pour assurer la compatibilité avec les versions antérieures.
LOI.NORMALE.STANDARD.INVERSE.N	La plage *E4:E14* de la figure 13-50 affiche les résultats de la fonction *LOI.NORMALE.STANDARD.INVERSE.N*. À partir de `F(x)`, probabilité de la fonction de répartition, *LOI.NORMALE.STANDARD.INVERSE.N* renvoie la première valeur de `x` respectant `P(X < x)`. Sa syntaxe apparaît en rouge dans le deuxième cadre blanc. *LOI.NORMALE.STANDARD.INVERSE* correspond à l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et Excel 2013 pour assurer la compatibilité avec les versions antérieures.
PHI	Excel 2013 propose une nouvelle fonction : *PHI*. Cette dernière utilise un argument pour lequel elle renvoie une valeur équivalente au résultat de la fonction *LOI.NORMALE.STANDARD.N* dans sa version non cumulative. De fait, on constate que la colonne *C* et la colonne *K* de la figure 13-50 affichent bien les mêmes résultats. La syntaxe de la formule entrée en *K4* apparaît dans le dernier cadre, à droite. **Nouveauté Excel 2013**.
GAUSS	Excel 2013 propose une nouvelle fonction : *GAUSS*. Cette dernière utilise un argument pour lequel elle renvoie une valeur équivalente au résultat de la fonction *LOI.NORMALE.STANDARD.N* dans sa version cumulative, diminué de `0,5`. De fait, on constate que les colonnes *L* (*GAUSS*) et *M* (*LOI.NORMALE.STANDARD.N* dans sa version cumulative - 0,5) de la figure 13-50 affichent bien les mêmes résultats. La syntaxe des formules entrées en *L4* et *M4* apparaît dans les deux premiers cadres, à droite. **Nouveauté Excel 2013**.

Utiliser la loi normale centrée réduite

La figure 13-51 illustre l'utilisation de cette loi. L'aire sous la première courbe représente la probabilité que `Z` prenne une valeur inférieure ou égale à `2`. L'aire sous la deuxième courbe (équivalente à la première) représente la probabilité que `Z` prenne une valeur supérieure ou égale à `2`. Et enfin, l'aire sous la troisième courbe représente la probabilité que `Z` prenne une valeur dans l'intervalle `[-2,-1]`. Pour chacune, la formule entrée en ligne *11* voit sa syntaxe exposée en ligne 9. En ligne *13*, on a formalisé la probabilité recherchée.

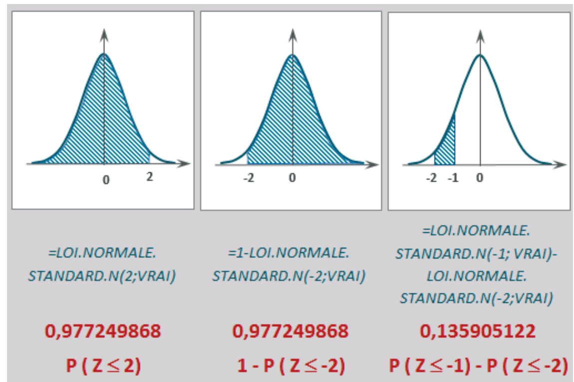
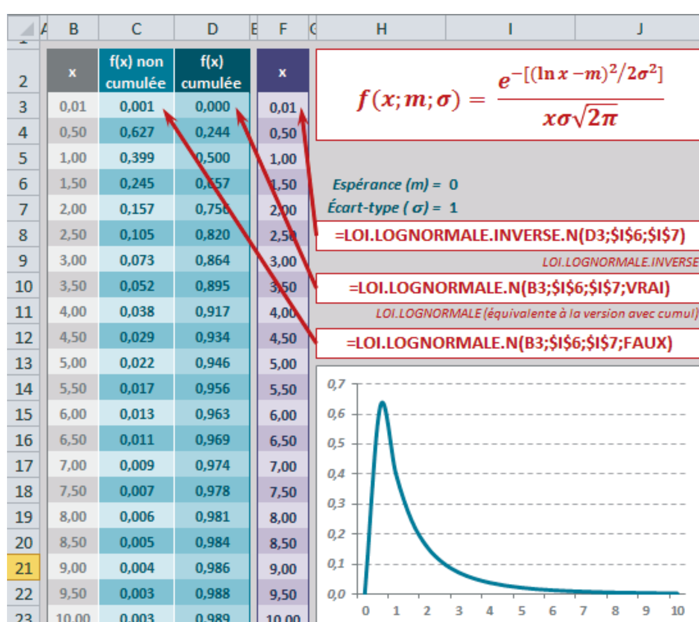


Figure 13–51
Illustration de trois calculs de probabilité à partir d'une variable aléatoire suivant une loi normale centrée réduite.

Loi log-normale

Une variable aléatoire réelle `X` suit une loi log-normale d'espérance `m` et d'écart-type σ si elle admet pour densité de probabilité la fonction présentée figure 13-52 (premier cadre blanc).



	x	f(x) non cumulée	f(x) cumulée	x
3	0,01	0,001	0,000	0,01
4	0,50	0,627	0,244	0,50
5	1,00	0,399	0,500	1,00
6	1,50	0,245	0,657	1,50
7	2,00	0,157	0,756	2,00
8	2,50	0,105	0,820	2,50
9	3,00	0,073	0,864	3,00
10	3,50	0,052	0,895	3,50
11	4,00	0,038	0,917	4,00
12	4,50	0,029	0,934	4,50
13	5,00	0,022	0,946	5,00
14	5,50	0,017	0,956	5,50
15	6,00	0,013	0,963	6,00
16	6,50	0,011	0,969	6,50
17	7,00	0,009	0,974	7,00
18	7,50	0,007	0,978	7,50
19	8,00	0,006	0,981	8,00
20	8,50	0,005	0,984	8,50
21	9,00	0,004	0,986	9,00
22	9,50	0,003	0,988	9,50
23	10,00	0,003	0,989	10,00

Figure 13–52
Mise en œuvre des fonctions LOI.LOGNORMALE.N et LOI.LOGNORMALE.INVERSE.N.

Loi log-normale

Une variable aléatoire X suit une loi log-normale de paramètres m (espérance) et σ (écart-type) si la variable Y = ln(X) suit une loi normale de paramètres m et σ. Exprimée à l'aide des fonctions Excel, cette relation donne donc : `LOI.LOGNORMALE.N(x;m;σ;VRAI)=LOI.NORMALE.N(LN(x);m;σ;VRAI)` ou encore `=LOI.NORMALE.STANDARD.N([LN(x) - m] / σ;VRAI)`.

LOI.LOGNORMALE.N renvoie, si son quatrième argument est positionné sur FAUX, P(X = x) (probabilité que la variable X prenne la valeur x) et, si son quatrième argument est positionné sur VRAI, P(X < x). La figure 13-52 illustre un exemple d'application de cette fonction pour m = 0 et σ = 1. Sa syntaxe apparaît en rouge dans les deux derniers cadres blancs. La plage *C3:C23* donne les résultats de la fonction dans sa version non cumulative P(X = x) (valeurs représentées sur le graphique) et la plage *D3:D23* dans sa version cumulative P(X < x).

La figure 13-52 présente également la mise en œuvre de la fonction *LOI.LOGNORMALE.INVERSE.N* (plage *F3:F23*). Elle renvoie x à partir de F(x), probabilité de la fonction de répartition. Sa syntaxe apparaît en rouge dans le deuxième cadre blanc. *LOI.LOGNORMALE* est l'ancienne forme de *LOI.LOGNORMALE.N* cumulée et *LOI.LOGNORMALE.INVERSE* est celle de *LOI.LOGNORMALE.INVERSE.N*. Toutes deux sont conservées dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

USAGE Loi log-normale

La modélisation par la loi log-normale est bien adaptée aux variables strictement positives, suivant une distribution asymétrique avec un allongement vers les valeurs élevées. Ces distributions sont fréquentes dans le domaine biologique (par exemple le poids des personnes) ou économique (la répartition des revenus).

Elle est également bien adaptée pour l'étude des variables qu'il est possible de décomposer en une multitude de variables plus petites, indépendantes. Si l'on s'intéresse par exemple au temps de parcours d'un itinéraire, on peut le décomposer en plusieurs temps de parcours élémentaires, chacun correspondant à un petit morceau de l'itinéraire initial. Si l'on ajuste chaque composante par une loi log-normale, le temps de parcours global peut lui-même être approché par une loi log-normale.

Loi Gamma

Histoire Fonctions eulériennes

En 1755, Euler publie un traité de calcul différentiel et intégral où l'on rencontre les fonctions dites aujourd'hui eulériennes. Parmi elles, la plus connue est sans doute l'intégrale eulérienne de seconde espèce, appelée fonction Gamma.

La loi Gamma est une loi de probabilité dont la portée est très vaste. En effet, des phénomènes réels très divers peuvent être approchés par une fonction Gamma. Son domaine de prédilection est une distribution à valeurs positives, fortement asymétrique et dotée d'une queue de distribution à décroissance rapide. En pratique, elle est souvent utilisée dans le domaine des assurances, pour décrire les phénomènes de durée de vie ou évaluer le temps écoulé entre deux sinistres.

Comprendre Fondements mathématiques

La fonction Gamma d'Euler est définie par l'intégrale présentée figure 13-53.

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$$

Figure 13–53 Fonction Gamma d'Euler

On montre mathématiquement que lorsqu'une distribution aléatoire suit la loi Gamma, elle admet pour densité de probabilité la fonction $f(x;\alpha;\beta)$, dont la définition, exprimée sous sa forme la plus générale, apparaît figure 13-54.

$$f(x;\alpha;\beta) = \frac{1}{\beta^{\alpha}\,\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$$

$$\alpha > 0 \text{ et } \beta > 0$$

Figure 13–54 Définition de la densité de probabilité d'une variable suivant une loi Gamma.

Tableau 13–11 Loi Gamma

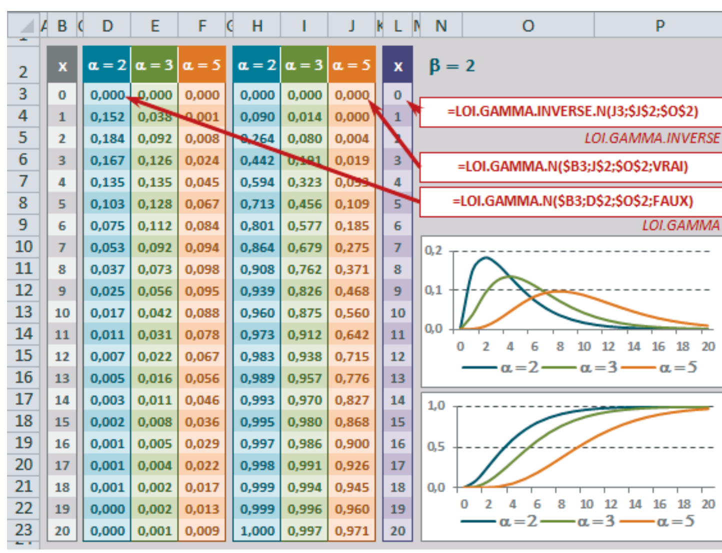
Fonction	Description
LOI.GAMMA.N	*LOI.GAMMA.N* renvoie, si son quatrième argument est positionné sur FAUX, P(X = x) et, si son quatrième argument est positionné sur VRAI, P(X < x). La figure 13-55 illustre un exemple d'application de cette fonction pour $\beta = 2$ avec $\alpha = 2$, $\alpha = 3$ et $\alpha = 5$. Sa syntaxe apparaît en rouge dans les deuxième et troisième cadres blancs. La plage *D3:F23* donne les résultats de la fonction dans sa version non cumulative P(X = x) (valeurs représentées sur le premier graphique) et la plage *H3:J23* dans sa version cumulative P(X < x) (valeurs représentées sur le deuxième graphique). *LOI.GAMMA* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Tableau 13–11 Loi Gamma (suite)

Fonction	Description
LOI.GAMMA .INVERSE.N	La plage *L3:L23* de la figure 13-55 affiche les résultats de la fonction *LOI.GAMMA.INVERSE.N*. À partir de `F(x)`, probabilité de la fonction de répartition, *LOI.GAMMA.INVERSE.N* renvoie la première valeur de `x` respectant `P(X < x)`. Sa syntaxe apparaît en rouge dans le premier cadre blanc. *LOI.GAMMA.INVERSE* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

ASTUCE Formats de nombre

Pour éviter d'entrer les valeurs de α « en dur » et disposer néanmoins de libellés explicites, on a appliqué aux cellules *D2*, *E2*, *F2*, *H2*, *I2* et *J2* le format de nombre `"a = "0`. Cette astuce permet d'avoir dans les cellules concernées uniquement des valeurs numériques (2, 3 et 5) qui peuvent être utilisées dans les formules et participer aux calculs.



	x	α = 2	α = 3	α = 5	α = 2	α = 3	α = 5	x
3	0	0,000	0,000	0,000	0,000	0,000	0,000	0
4	1	0,152	0,038	0,001	0,090	0,014	0,000	1
5	2	0,184	0,092	0,008	0,264	0,080	0,004	[illegible]
6	3	0,167	0,126	0,024	0,442	[illegible]	0,019	3
7	4	0,135	0,135	0,045	0,594	0,323	[illegible]	4
8	5	0,103	0,128	0,067	0,713	0,456	0,109	5
9	6	0,075	0,112	0,084	0,801	0,577	0,185	6
10	7	0,053	0,092	0,094	0,864	0,679	0,275	7
11	8	0,037	0,073	0,098	0,908	0,762	0,371	8
12	9	0,025	0,056	0,095	0,939	0,826	0,468	9
13	10	0,017	0,042	0,088	0,960	0,875	0,560	10
14	11	0,011	0,031	0,078	0,973	0,912	0,642	11
15	12	0,007	0,022	0,067	0,983	0,938	0,715	12
16	13	0,005	0,016	0,056	0,989	0,957	0,776	13
17	14	0,003	0,011	0,046	0,993	0,970	0,827	14
18	15	0,002	0,008	0,036	0,995	0,980	0,868	15
19	16	0,001	0,005	0,029	0,997	0,986	0,900	16
20	17	0,001	0,004	0,022	0,998	0,991	0,926	17
21	18	0,001	0,002	0,017	0,999	0,994	0,945	18
22	19	0,000	0,002	0,013	0,999	0,996	0,960	19
23	20	0,000	0,001	0,009	1,000	0,997	0,971	20

Figure 13–55
Mise en œuvre des fonctions LOI.GAMMA.N et LOI.GAMMA.INVERSE.N.

POUR LES CURIEUX Comment passe-t-on de la fonction Gamma d'Euler à f(x;α;β) ?

Partons de l'intégrale eulérienne de seconde espèce présentée figure 13-53. Pour une valeur alpha (α) donnée, la fonction Gamma prend la valeur présentée figure 13-56, à partir de laquelle on peut déduire l'équation de la figure 13-57.

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$$

Figure 13–56 Valeur de la fonction Gamma pour x = α.

$$\int_0^\infty \frac{1}{\Gamma(\alpha)} e^{-t} t^{\alpha-1} dt = 1$$

Figure 13–57 Autre forme de l'équation présentée figure 13-56.

L'équation affichée figure 13-57 n'est autre que l'intégrale de la densité de probabilité d'une variable aléatoire de valeurs réelles positives. Elle caractérise la distribution Gamma dite « de base ». α représente ici un paramètre qui influe sur la forme de la courbe de la distribution. En effet, selon la valeur de α, cette équation décrit trois grandes familles de courbes :

- `α > 1` : la distribution adopte une forme de cloche asymétrique.
- `α = 1` : la distribution est exponentielle.
- `α < 1` : la distribution est monotone décroissante.

Dans l'optique de conférer une plus grande souplesse à cette distribution, il est possible de façonner la relation exposée figure 13-57 de manière à introduire un deuxième paramètre positif Bêta (β). Partons du principe que la variable aléatoire `T` suit la distribution Gamma de base. À l'aide d'une nouvelle variable aléatoire `X = βT` (`T = X/β`), l'équation de la figure 13-57 devient celle de la figure 13-58.

$$\int_0^\infty \frac{1}{\Gamma(\alpha)} \frac{x^{\alpha-1}}{\beta^{\alpha-1}\beta} e^{-\frac{x}{\beta}} t^{\alpha-1} dx = 1$$

Figure 13–58 Autre forme de l'équation présentée figure 13-57 avec T = X/β.

En effectuant les calculs, on retrouve bien la fonction de densité de probabilité `f(x;α;β)` présentée figure 13-54. Le paramètre β est un paramètre d'échelle ou de dispersion. La fonction `f(x;α;β)` caractérise la distribution Gamma dite « complète ».

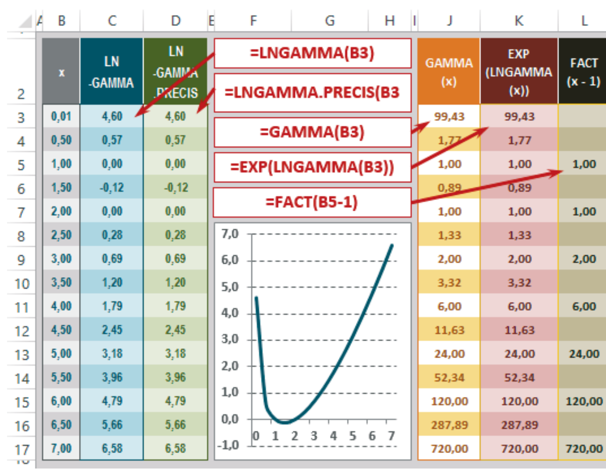
À partir de l'équation générale présentée figure 13-54 et en attribuant aux paramètres α et β des valeurs particulières, on obtient des formes spécifiques de `f(x;α;β)`.

- En positionnant α à `1`, la fonction `f(x;α;β)` n'est autre que la densité de probabilité d'une distribution exponentielle (de paramètre `λ = 1 / β`). Exprimée à l'aide des fonctions Excel, cette propriété donne la relation suivante : `=LOI.GAMMA.N(x;1;β;FAUX)=LOI.EXPONENTIELLE.N(x;1/β;FAUX)`.
- Lorsque `α = n/2` (avec `n` entier) et `β = 2`, `f(x;α;2)` représente la distribution du Khi-deux à `n` degrés de liberté. Exprimée à l'aide des fonctions Excel, cette propriété, donne la relation suivante : `LOI.GAMMA.N(x;n/2;2;FAUX)=LOI.KHIDEUX.N(x;n;FAUX)`

Tableau 13–12 Loi Gamma

Fonction	Description
LNGAMMA et *LNGAMMA .PRECIS*	La fonction *LNGAMMA* renvoie le logarithme népérien de la fonction Gamma. Figure 13-59, les résultats de la fonction *LNGAMMA* affichés en *C3:C17* correspondent à la courbe représentée sur le graphique. Dans la plage *D3:D17*, on trouve les résultats de la fonction *LNGAMMA.PRECIS*… qui ressemblent comme deux gouttes d'eau à ceux de la colonne *C* !
GAMMA	Excel 2013 propose une nouvelle fonction, *GAMMA*. Cette dernière utilise un argument, dont elle renvoie l'exponentielle de *LNGAMMA*. Ses résultats apparaissent dans la colonne *J* de la figure 13-59. En *K3*, on a entré `=EXP(LNGAMMA(B3))` afin de vérifier que cette définition de la fonction renvoyait bien des valeurs équivalentes à la fonction *GAMMA*. Cette formule a été ensuite recopiée dans la colonne *K*. En colonne *L*, on a entré une formule vérifiant que lorsque `x` est un nombre entier, on a la relation `=EXP(LNGAMMA(x))=FACT(x - 1)`. **Nouveauté Excel 2013**.

Figure 13–59
Mise en œuvre des fonctions LNGAMMA et LNGAMMA.PRECIS et, pour ceux qui travaillent sous Excel 2013, présentation de la fonction GAMMA.



	x	LN GAMMA	LN GAMMA PRECIS	GAMMA (x)	EXP (LNGAMMA (x))	FACT (x - 1)
3	0,01	4,60	4,60	99,43	99,43	
4	0,50	0,57	0,57	1,77	1,77	
5	1,00	0,00	0,00	1,00	1,00	1,00
6	1,50	-0,12	-0,12	0,89	0,89	
7	2,00	0,00	0,00	1,00	1,00	1,00
8	2,50	0,28	0,28	1,33	1,33	
9	3,00	0,69	0,69	2,00	2,00	2,00
10	3,50	1,20	1,20	3,32	3,32	
11	4,00	1,79	1,79	6,00	6,00	6,00
12	4,50	2,45	2,45	11,63	11,63	
13	5,00	3,18	3,18	24,00	24,00	24,00
14	5,50	3,96	3,96	52,34	52,34	
15	6,00	4,79	4,79	120,00	120,00	120,00
16	6,50	5,66	5,66	287,89	287,89	
17	7,00	6,58	6,58	720,00	720,00	720,00

Loi Bêta

La loi Bêta est caractérisée par deux paramètres de forme α et β (ils influencent le tracé de la courbe). Sa fonction de densité de probabilité est présentée figure 13-60 sous sa forme standard, c'est-à-dire pour x compris entre 0 et 1, et sous sa forme plus générale, c'est-à-dire pour x compris entre a et b, figure 13-61.

Figure 13–60
Forme standard (0 ≤x ≤1) de la fonction de densité de probabilité de la loi Bêta.
Γ est la fonction Gamma d'Euler.

$$f(x,\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}$$

$$\alpha > 0 \text{ et } \beta > 0$$

$$0 \le x \le 1$$

Figure 13–61
Forme plus générale ($a \leq x \leq b$) de la fonction de densité de probabilité de la loi Bêta.

$$f(x,\alpha,\beta,a,b) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}(x-a)^{\alpha-1}(b-x)^{\beta-1}(b-a)^{-(\alpha+\beta-1)}$$

$$\alpha > 0 \text{ et } \beta > 0$$
$$a \leq x \leq b$$

CORRESPONDANCES Loi Gamma et loi Bêta

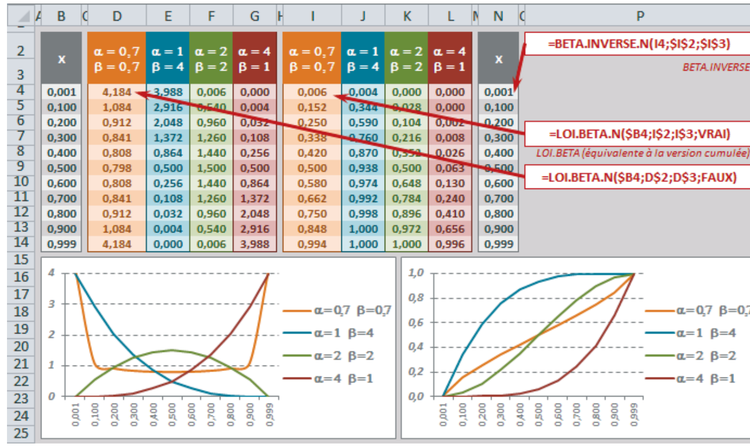
Si X et Y sont indépendamment distribuées selon une loi Gamma, de paramètres (α,θ) et (β,θ), alors X / (X + Y) est distribuée selon une loi Bêta de paramètres (α,β).

Concrètement, la loi Bêta peut être mise à contribution dans des domaines très variés. Elle sert par exemple à modéliser des audiences radiophoniques dans le but d'optimiser une campagne publicitaire. Elle est également souvent utilisée en gestion de projets, pour calculer la durée probable d'une tâche élémentaire et définir des scénarios optimistes et pessimistes.

Tableau 13–13 Loi Bêta

Fonction	Description
LOI.BETA.N	*LOI.BETA.N* renvoie, si son quatrième argument est positionné sur FAUX, P(X = x)et, si son quatrième argument est positionné sur VRAI, P(X < x). La figure 13-62 illustre un exemple d'application de cette fonction pour quatre couples de valeurs α et β. Sa syntaxe apparaît en rouge dans les deuxième et troisième cadres blancs. Elle a été utilisée ici sans que l'on ait précisé les cinquième et sixième arguments (optionnels) qui correspondent aux a et b de la forme générale de la fonction de densité. Il s'agit donc de la forme standard de la fonction ; c'est pourquoi x prend ses valeurs entre 0 et 1. La plage *D4:G14* donne les résultats de la fonction dans sa version non cumulative P(X = x) (valeurs représentées sur le graphique de gauche) et la plage *I4:L14* dans sa version cumulative P(X < x) (valeurs représentées sur le graphique de droite). *LOI.BETA* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.
BETA.INVERSE.N	La plage *N4:N14* de la figure 13-62 affiche les résultats de la fonction *BETA.INVERSE.N*. À partir de F(x), probabilité de la fonction de répartition, *BETA.INVERSE.N* renvoie la première valeur de x respectant P(X < x). Sa syntaxe apparaît en rouge dans le premier cadre blanc. *BETA.INVERSE* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Figure 13–62
Mise en œuvre des fonctions LOI.BETA.N et BETA.INVERSE.N.



x	α = 0,7 β = 0,7	α = 1 β = 4	α = 2 β = 2	α = 4 β = 1	α = 0,7 β = 0,7	α = 1 β = 4	α = 2 β = 2	α = 4 β = 1	x
0,001	4,184	3,988	0,006	0,000	0,006	0,004	0,000	0,000	0,001
0,100	1,084	2,916	0,540	0,004	0,152	0,344	0,028	0,000	0,100
0,200	0,912	2,048	0,960	0,032	0,250	0,590	0,104	0,002	0,200
0,300	0,841	1,372	1,260	0,108	0,338	0,760	0,216	0,008	0,300
0,400	0,808	0,864	1,440	0,256	0,420	0,870	0,352	0,026	0,400
0,500	0,798	0,500	1,500	0,500	0,500	0,938	0,500	0,063	0,500
0,600	0,808	0,256	1,440	0,864	0,580	0,974	0,648	0,130	0,600
0,700	0,841	0,108	1,260	1,372	0,662	0,992	0,784	0,240	0,700
0,800	0,912	0,032	0,960	2,048	0,750	0,998	0,896	0,410	0,800
0,900	1,084	0,004	0,540	2,916	0,848	1,000	0,972	0,656	0,900
0,999	4,184	0,000	0,006	3,988	0,994	1,000	1,000	0,996	0,999

Loi de Weibull

Une variable aléatoire réelle X suit une loi de Weibull de paramètres α (paramètre de forme) et β (paramètre de temps), si elle admet pour densité de probabilité la fonction présentée figure 13-63.

Figure 13–63
Loi de Weibull : f(x,α,β), sa fonction de densité de probabilité, P(X = x) et F(x,α,β), sa fonction de répartition, P(X < x).

$$f(x,\alpha,\beta) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$$

$$F(x,\alpha,\beta) = 1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$$

$\alpha > 0$ et $\beta > 0$

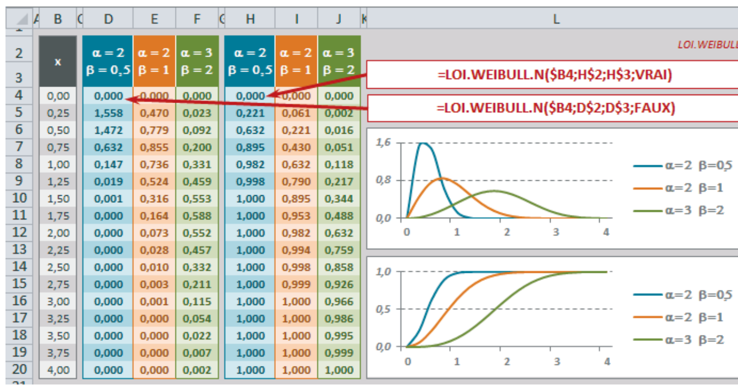
Il n'y a pas « une » loi de Weibull, mais toute une famille de lois, correspondant à diverses valeurs de α et β. Parmi celles-ci, on distingue la loi exponentielle (avec $\alpha = 1$ et $\beta = 1 / \lambda$) ou la loi de Rayleigh (avec $\alpha = 2$). Lorsque α est compris entre 1,5 et 2 ou 3 et 3,6, on obtient une loi log-normale. La loi de Weibull est un cas particulier de la loi Gamma.

EN PRATIQUE À quoi sert-elle ?

Dans les entreprises, cette loi est très utilisée pour les contrôles de fiabilité. Ainsi, $\alpha < 1$ correspond à un matériel qui se bonifie avec le temps (matériel en rodage), $\alpha = 1$ correspond à un matériel sans usure et $\alpha > 1$, à un matériel qui se dégrade avec le temps (la plupart de nos produits de consommation !). Cette loi est également utilisée pour étudier les problèmes dits « de valeurs extrêmes » comme la survenue de crues exceptionnelles dans une rivière.

LOI.WEIBULL.N renvoie, si son quatrième argument est positionné sur `FAUX`, `P(X = x)` et, si son quatrième argument est positionné sur `VRAI`, `P(X < x)`. La figure 13-64 illustre un exemple d'application de cette fonction pour trois couples de valeurs α et β. Sa syntaxe apparaît en rouge dans les cadres blancs.

Figure 13–64
Mise en œuvre de la fonction LOI.WEIBULL.N.



x	α = 2 β = 0,5	α = 2 β = 1	α = 3 β = 2	α = 2 β = 0,5	α = 2 β = 1	α = 3 β = 2
0,00	0,000	0,000	0,000	0,000	0,000	0,000
0,25	1,558	0,470	0,023	0,221	0,061	0,002
0,50	1,472	0,779	0,092	0,632	0,221	0,016
0,75	0,632	0,855	0,200	0,895	0,430	0,051
1,00	0,147	0,736	0,331	0,982	0,632	0,118
1,25	0,019	0,524	0,459	0,998	0,790	0,217
1,50	0,001	0,316	0,553	1,000	0,895	0,344
1,75	0,000	0,164	0,588	1,000	0,953	0,488
2,00	0,000	0,073	0,552	1,000	0,982	0,632
2,25	0,000	0,028	0,457	1,000	0,994	0,759
2,50	0,000	0,010	0,332	1,000	0,998	0,858
2,75	0,000	0,003	0,211	1,000	0,999	0,926
3,00	0,000	0,001	0,115	1,000	1,000	0,966
3,25	0,000	0,000	0,054	1,000	1,000	0,986
3,50	0,000	0,000	0,022	1,000	1,000	0,995
3,75	0,000	0,000	0,007	1,000	1,000	0,999
4,00	0,000	0,000	0,002	1,000	1,000	1,000

Sur la figure 13-64, la plage *D4:F20* donne les résultats de la fonction dans sa version non cumulative `P(X = x)` (valeurs représentées sur le premier graphique) et la plage *H4:J20* dans sa version cumulative `P(X < x)` (valeurs représentées sur le deuxième graphique). *LOI.WEIBULL* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Loi exponentielle

Une variable aléatoire réelle `X` suit une loi exponentielle de paramètre λ si elle admet pour densité de probabilité la fonction présentée figure 13-65.

Figure 13–65
Loi exponentielle : f(x,λ), sa fonction de densité de probabilité, P(X = x) et F(x,λ), sa fonction de répartition, P(X < x).

$$f(x,\lambda) = \lambda e^{-\lambda x}$$

$$F(x,\lambda) = 1 - e^{-\lambda x}$$

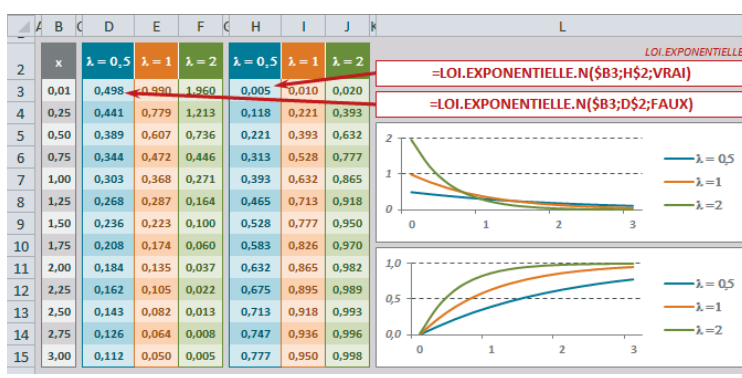
$$\lambda > 0$$

La loi exponentielle ne dépend que d'un seul paramètre : λ, l'ordonnée à l'origine de la courbe de densité de probabilité. Celui-ci peut représenter le nombre de fois où un événement est survenu durant un laps de temps donné. Quand il vaut `1`, on parle de loi exponentielle standard. La loi exponentielle est une forme particulière de la loi de Weibull pour $\alpha = 1$ et $\beta = 1 / \lambda$.

En pratique À quoi sert-elle ?

Cette loi sert à modéliser les variables caractérisées par une absence de mémoire ou par une durée de vie sans vieillissement. En d'autres termes, le comportement d'une telle variable dans un intervalle donné dépend uniquement de la longueur de cet intervalle et non de sa position sur l'axe des temps.
Pour donner un exemple concret, si l'on étudie la fréquentation d'un lieu (boutique, guichet d'information, etc.) et si l'analyse d'un échantillon de « clients » observés montre que la fréquentation ne dépend pas de l'heure à laquelle elle est mesurée, alors le temps séparant la survenue de deux clients suit une loi exponentielle.
Cette loi est également régulièrement utilisée pour modéliser la durée de vie de la radioactivité. Le paramètre λ, représentant alors l'inverse de l'espérance de vie de l'atome, s'appelle la constante de désintégration. La médiane (temps nécessaire pour que la population voie son effectif diminuer de 50 %) s'appelle la demi-vie ou période.

LOI.EXPONENTIELLE.N renvoie `P(X = x)` si son troisième argument est positionné sur `FAUX` et `P(X < x)` si son troisième argument est positionné sur `VRAI`. La figure 13-66 illustre un exemple d'application de cette fonction pour trois valeurs λ. Sa syntaxe apparaît en rouge dans les cadres blancs. La plage *D3:F15* donne les résultats de la fonction dans sa version non cumulative `P(X = x)` (valeurs représentées sur le premier graphique) et la plage *H3:J15* dans sa version cumulative `P(X < x)` (valeurs représentées sur le deuxième graphique). *LOI.EXPONENTIELLE* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et Excel 2013 pour assurer la compatibilité avec les versions antérieures.



	B	D	E	F	H	I	J
2	x	λ = 0,5	λ = 1	λ = 2	λ = 0,5	λ = 1	λ = 2
3	0,01	0,498	0,990	1,960	0,005	0,010	0,020
4	0,25	0,441	0,779	1,213	0,118	0,221	0,393
5	0,50	0,389	0,607	0,736	0,221	0,393	0,632
6	0,75	0,344	0,472	0,446	0,313	0,528	0,777
7	1,00	0,303	0,368	0,271	0,393	0,632	0,865
8	1,25	0,268	0,287	0,164	0,465	0,713	0,918
9	1,50	0,236	0,223	0,100	0,528	0,777	0,950
10	1,75	0,208	0,174	0,060	0,583	0,826	0,970
11	2,00	0,184	0,135	0,037	0,632	0,865	0,982
12	2,25	0,162	0,105	0,022	0,675	0,895	0,989
13	2,50	0,143	0,082	0,013	0,713	0,918	0,993
14	2,75	0,126	0,064	0,008	0,747	0,936	0,996
15	3,00	0,112	0,050	0,005	0,777	0,950	0,998

Figure 13–66
Mise en œuvre de la fonction LOI.EXPONENTIELLE.N.

Loi du Khi-deux

Une variable aléatoire réelle `Y` suit une loi du Khi-deux à ν degrés de liberté, si elle admet pour densité de probabilité la fonction présentée figure 13-67.

Figure 13–67
Fonction de densité de la loi du Khi-deux.

$$f_\nu(y) = \frac{1}{2^{\frac{\nu}{2}}\,\Gamma\left(\frac{\nu}{2}\right)}\, y^{\frac{\nu}{2}-1}\, e^{-\frac{y}{2}}$$

$\nu \geq 1$ (entier)
pour tout y positif
Γ est la fonction gamma d'Euler

COMPRENDRE D'où vient cette fonction ?

On considère ν variables aléatoires $(X_1, X_2, \ldots, X_\nu)$ indépendantes et suivant toutes une loi normale centrée réduite. On s'intéresse à la variable Y, somme du carré de ces ν variables indépendantes (figure 13-68).

Figure 13–68 Y est une variable construite à partir de la somme des carrés de ν variables indépendantes suivant toutes une loi normale centrée réduite.

$$Y = X_1^2 + X_2^2 + \ldots + X_\nu^2$$

De savantes démonstrations montrent que si deux variables indépendantes suivent respectivement deux lois Gamma de paramètres (α_1, β) et (α_2, β), leur somme suit une loi Gamma de paramètres $(\alpha_1 + \alpha_2, \beta)$. D'autres savantes démonstrations généralisent ce résultat à la somme de ν variables pour en conclure que la variable Y présentée quelques lignes plus haut suit une loi Gamma de paramètres $(\nu / 2, 2)$.
En appliquant ces paramètres à la définition de la loi Gamma, on retrouve la fonction de densité présentée figure 13-67 et on peut en conclure que Y suit bien une loi du Khi-deux à ν degrés de liberté.

La loi du Khi-deux dépend du paramètre ν, appelé `nombre de degrés de liberté` (`ddl` en abrégé). Il exprime le nombre de composantes indépendantes de Y.

COMPRENDRE Degrés de liberté

En statistiques, « degrés de liberté » désigne le nombre de valeurs aléatoires non déterminées par une équation. Un petit exemple s'impose : si l'on cherche deux nombres x et y dont la somme est 8, l'équation $x + y = 8$ ne permet de déterminer aucun des deux nombres, mais si x est choisi arbitrairement, y est déterminé (et inversement). À travers une telle équation, vous mettez en jeu deux variables aléatoires (X,Y), mais vous ne disposez que d'un degré de liberté.

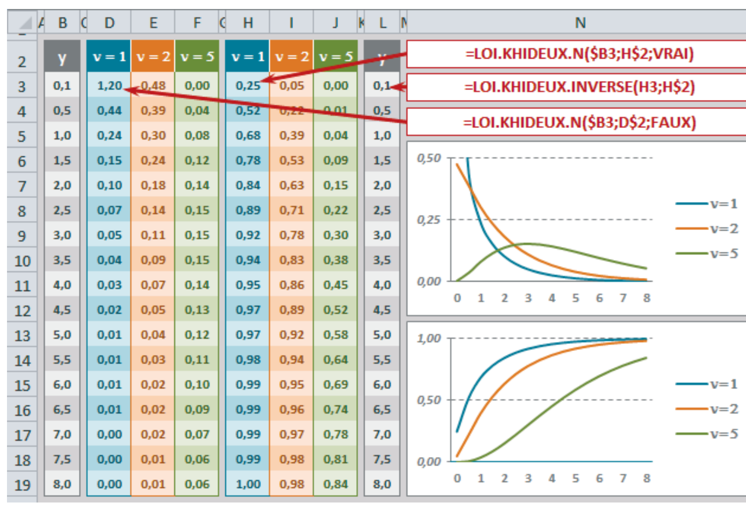
L'espérance d'une loi du Khi-deux est égale à ν et sa variance à 2ν. Lorsque ν est « grand » ($\nu > 100$), la loi du Khi-deux peut être approchée par une loi normale d'espérance ν et de variance 2ν.

EN PRATIQUE À quoi sert-elle ?

Cette loi est mise à contribution à travers deux tests : le test d'ajustement du Khi-deux et le test d'indépendance. Deux exemples sont donnés, vers la fin de ce chapitre, dans la section traitant des tests statistiques. Le premier exemple cherche à déterminer si la clientèle d'un magasin se répartit équitablement entre ses différentes caisses et le deuxième veut savoir si l'âge a un effet sur le fait d'être propriétaire ou locataire.

Tableau 13–14 Loi du Khi-deux

Fonction	Description
LOI.KHIDEUX.N	Si son troisième argument est positionné sur FAUX, cette fonction renvoie `P(Y = y)` et si son troisième argument est positionné sur VRAI, `P(Y < y)`. La figure 13-69 illustre un exemple d'application de cette fonction pour trois valeurs de ν. Sa syntaxe apparaît en rouge dans les premier et troisième cadres blancs. La plage *D3:F19* donne les résultats de la fonction dans sa version non cumulative `P(Y = y)` (valeurs représentées sur le premier graphique) et la plage *H3:J19* dans sa version cumulative `P(Y < y)` (valeurs représentées sur le deuxième graphique).
LOI.KHIDEUX .INVERSE	La plage *L3:L19* de la figure 13-69 affiche les résultats de la fonction *LOI.KHIDEUX.INVERSE*. À partir de `F(y)`, probabilité de la fonction de répartition, *LOI.KHIDEUX.INVERSE* renvoie la première valeur de y respectant `P(Y < y)`. Sa syntaxe apparaît en rouge dans le deuxième cadre blanc.



	B	D	E	F	H	I	J	L
2	y	v = 1	v = 2	v = 5	v = 1	v = 2	v = 5	y
3	0,1	1,20	0,48	0,00	0,25	0,05	0,00	0,1
4	0,5	0,44	0,39	0,04	0,52	0,22	0,01	0,5
5	1,0	0,24	0,30	0,08	0,68	0,39	0,04	1,0
6	1,5	0,15	0,24	0,12	0,78	0,53	0,09	1,5
7	2,0	0,10	0,18	0,14	0,84	0,63	0,15	2,0
8	2,5	0,07	0,14	0,15	0,89	0,71	0,22	2,5
9	3,0	0,05	0,11	0,15	0,92	0,78	0,30	3,0
10	3,5	0,04	0,09	0,15	0,94	0,83	0,38	3,5
11	4,0	0,03	0,07	0,14	0,95	0,86	0,45	4,0
12	4,5	0,02	0,05	0,13	0,97	0,89	0,52	4,5
13	5,0	0,01	0,04	0,12	0,97	0,92	0,58	5,0
14	5,5	0,01	0,03	0,11	0,98	0,94	0,64	5,5
15	6,0	0,01	0,02	0,10	0,99	0,95	0,69	6,0
16	6,5	0,01	0,02	0,09	0,99	0,96	0,74	6,5
17	7,0	0,00	0,02	0,07	0,99	0,97	0,78	7,0
18	7,5	0,00	0,01	0,06	0,99	0,98	0,81	7,5
19	8,0	0,00	0,01	0,06	1,00	0,98	0,84	8,0

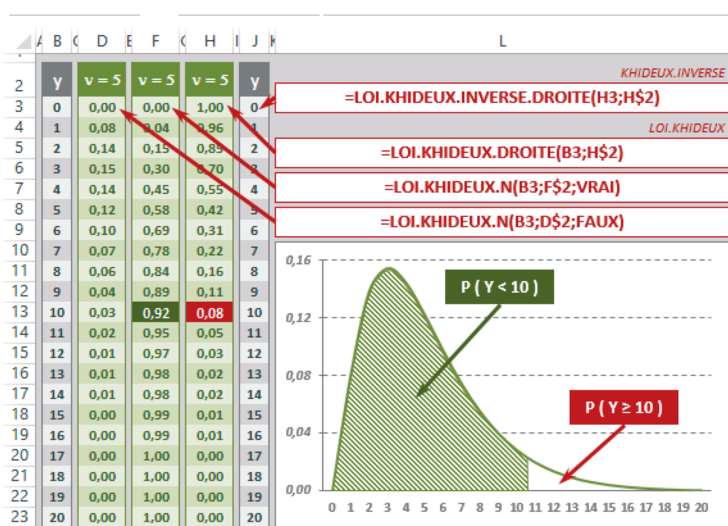
Figure 13–69
Mise en œuvre des fonctions LOI.KHIDEUX.N et LOI.KHIDEUX.INVERSE.

Tableau 13–15 Loi du Khi-deux

Fonction	Description
LOI.KHIDEUX .DROITE	La fonction *LOI.KHIDEUX.N*, dans sa version cumulée, renvoie `P(Y < y)`, c'est-à-dire la probabilité que la variable Y prenne une valeur inférieure à y. Or, que ce soit pour un test d'ajustement ou d'indépendance (voir l'aparté précédent), on n'a pas besoin de connaître `P(Y < y)`, mais plutôt `P(Y ≥ y)`. Il est bien évident que l'une peut se déduire de l'autre à travers la relation suivante : `P(Y ≥ y) = 1 - P(Y < y)`. Cependant, Excel, dans sa grande bonté, propose la fonction *LOI.KHIDEUX.DROITE* qui renvoie directement le résultat. La plage *D3:D23* de la figure 13-70 donne les résultats de la fonction *LOI.KHIDEUX.N* pour 5 degrés de liberté dans sa version non cumulative (valeurs ayant permis de tracer la courbe du graphique) et la plage *F3:F23* donne les résultats de la même fonction dans sa version cumulative.

Tableau 13–15 Loi du Khi-deux (suite)

Fonction	Description
	La plage *H3:H23* donne les résultats de la fonction *LOI.KHIDEUX.DROITE* pour 5 degrés de liberté. Sa syntaxe apparaît en rouge dans le deuxième cadre blanc. On constate aisément que ses résultats ajoutés à ceux de la plage *F3:F23* donnent toujours 1. On a symbolisé cette relation dans le graphique en prenant 10 comme valeur de y et en hachurant en vert la surface sous la courbe correspondant à la valeur de la cellule *F13*. La partie blanche, quant à elle, correspond à la valeur de la cellule *H13* (la somme de *F13* et *H13* faisant bien 1, mesure de la surface totale sous la courbe). *LOI.KHIDEUX* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.
LOI.KHIDEUX .INVERSE.DROITE	La plage *J3:J23* de la figure 13-70 affiche les résultats de la fonction *LOI.KHIDEUX.INVERSE.DROITE*. À partir de `1 - F(y)` (`1 -` probabilité de la fonction de répartition), *LOI.KHIDEUX.INVERSE.DROITE* renvoie la première valeur de `y` respectant `P(Y ≥ y)`. Sa syntaxe apparaît en rouge dans le premier cadre blanc. *KHIDEUX.INVERSE* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.



y	v = 5	v = 5	v = 5	y
0	0,00	0,00	1,00	0
1	0,08	0,04	0,96	1
2	0,14	0,15	0,85	2
3	0,15	0,30	0,70	3
4	0,14	0,45	0,55	4
5	0,12	0,58	0,42	5
6	0,10	0,69	0,31	6
7	0,07	0,78	0,22	7
8	0,06	0,84	0,16	8
9	0,04	0,89	0,11	9
10	0,03	0,92	0,08	10
11	0,02	0,95	0,05	11
12	0,01	0,97	0,03	12
13	0,01	0,98	0,02	13
14	0,01	0,98	0,02	14
15	0,00	0,99	0,01	15
16	0,00	0,99	0,01	16
17	0,00	1,00	0,00	17
18	0,00	1,00	0,00	18
19	0,00	1,00	0,00	19
20	0,00	1,00	0,00	20

Figure 13–70
Mise en œuvre des fonctions LOI.KHIDEUX.DROITE et LOI.KHIDEUX.INVERSE.DROITE.

Loi de Student

Une variable aléatoire réelle `T` suit une loi de Student à ν degrés de liberté si elle admet pour densité de probabilité la fonction présentée figure 13-71.

Figure 13–71
Fonction de densité de la loi de Student.

$$f_\nu(t) = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}}$$

$\nu \geq 1$ (entier)
Γ est la fonction Gamma d'Euler

COMPRENDRE D'où vient cette fonction ?

On considère deux variables aléatoires indépendantes X et Y. X suit une loi normale centrée réduite et Y, une loi du Khi-deux à ν degrés de liberté. On s'intéresse à la variable T issue du rapport de X et Y, présenté figure 13-72.

Figure 13–72 T est une variable construite à partir du rapport de X qui suit une loi normale centrée réduite et de la racine carrée de Y sur ν, Y suivant une loi du Khi-deux à ν degrés de liberté.

$$T_\nu = \frac{X}{\sqrt{\frac{Y}{\nu}}}$$

De savantes démonstrations montrent que l'on peut passer du rapport présenté figure 13-72 à celui de la figure 13-73, le numérateur et le dénominateur étant deux variables indépendantes suivant deux lois Gamma de paramètres 1/2 et ν/2.

Figure 13–73 $X^2/2$ et Y/2 sont deux variables aléatoires indépendantes suivant deux lois Gamma de paramètres 1/2 et ν/2.

$$\frac{T_\nu^2}{\nu} = \frac{\frac{X^2}{2}}{\frac{Y}{2}}$$

On en conclut que T_ν^2 / ν suit une loi Bêta de paramètres 1/2 et ν/2, ce qui (… toujours à l'issue de savants calculs) donne la définition de la fonction de densité présentée figure 13-71.

La loi de Student dépend du paramètre ν, appelé nombre de degrés de liberté. Lorsque $\nu > 1$, l'espérance est égale à 0 et, lorsque $\nu > 2$, la variance vaut $\nu / (\nu - 2)$.

EN PRATIQUE À quoi sert-elle ?

À travers ce que l'on appelle « Test de Student », cette loi sert essentiellement à comparer les moyennes de deux populations. Pour y parvenir, on utilise les moyennes de deux petits échantillons tirés de ces populations, pour calculer une statistique appelée « T de Student », censée suivre une loi de Student. La comparaison de ce T et de la valeur pour laquelle la fonction de répartition de Student renvoie une probabilité d'erreur acceptable (généralement 5 %) permet d'en déduire l'égalité (ou la différence) des moyennes des deux populations initiales (voir l'exemple développé à la fin de cet ouvrage dans la section traitant des tests statistiques).

Tableau 13–16 Loi de Student

Fonction	Description
LOI.STUDENT.N	Si son troisième argument est positionné sur `FAUX`, cette fonction renvoie `P(T = t)` et, si son troisième argument est positionné sur `VRAI`, elle renvoie `P(T < t)`. La figure 13-74 illustre un exemple d'application de cette fonction pour trois valeurs de `v`. Sa syntaxe apparaît en rouge dans les deux derniers cadres blancs. La plage *D3:F11* donne les résultats de la fonction dans sa version non cumulative `P(T = t)` (valeurs représentées sur le graphique de gauche) et la plage *H3:J11* dans sa version cumulative `P(T < t)` (valeurs représentées sur le graphique de droite).
LOI.STUDENT .INVERSE.N	La plage *L3:L11* de la figure 13-74 affiche les résultats de la fonction *LOI.STUDENT.INVERSE*. À partir de `F(t)`, probabilité de la fonction de répartition, *LOI.STUDENT.INVERSE* renvoie la première valeur de `t` respectant `P(T < t)`. Sa syntaxe apparaît en rouge dans le premier cadre blanc.

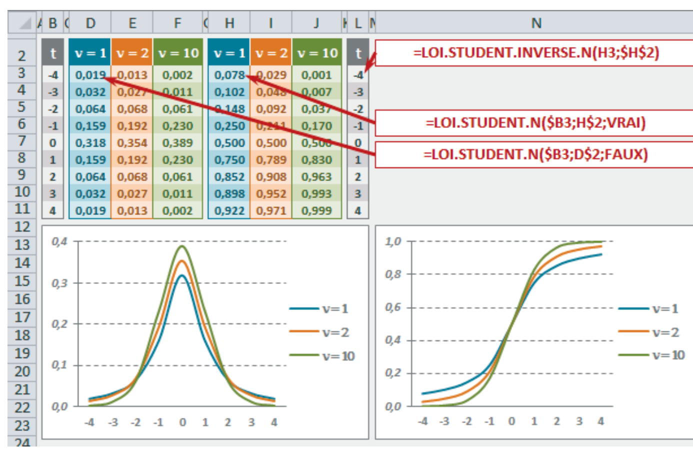


Figure 13–74
Mise en œuvre des fonctions LOI.STUDENT.N et LOI.STUDENT.INVERSE.N.

Tableau 13–17 Loi de Student

Fonction	Description
LOI.STUDENT .DROITE	La fonction *LOI.STUDENT.N*, dans sa version cumulée, renvoie `P(T < t)`. Or, pour un test de Student, on n'a pas besoin de connaître `P(T < t)`, mais plutôt `P(T ≥ t)` (queue de courbe à droite, renvoyée par *LOI.STUDENT.DROITE*) ou encore `P(T < -t) + P(T ≥ t)` (cumul des deux queues de courbe à droite et à gauche, renvoyé par *LOI.STUDENT.BILATERALE*). La plage *D3:D13* de la figure 13-75 donne les résultats de la fonction *LOI.STUDENT.N* pour `5` degrés de liberté dans sa version non cumulative (valeurs ayant permis de tracer la courbe des deux graphiques) et la plage *F3:F13* donne les résultats de la même fonction, dans sa version cumulative. Leur syntaxe apparaît dans les deuxième et troisième cadres blancs. La plage *H3:H13* donne les résultats de la fonction *LOI.STUDENT.DROITE* pour `5` degrés de liberté. Sa syntaxe apparaît en rouge dans le premier cadre blanc.

Tableau 13–17 Loi de Student (suite)

Fonction	Description
	On constate aisément que ses résultats ajoutés à ceux de la plage *F3:F13* donnent toujours `1`. On a symbolisé cette relation dans le premier graphique en prenant `2` comme valeur de `t` et en hachurant en vert la surface sous la courbe correspondant à la valeur de la cellule *F10*. La partie orange, quant à elle, correspond à la valeur de la cellule *H10* (la somme de *F10* et *H10* faisant bien `1`, mesure de la surface totale sous la courbe). *LOI.STUDENT* en mode unilatéral (c'est-à-dire avec son troisième argument égal à `1`) est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.
LOI.STUDENT .BILATERALE	La plage *J8:J13* donne les résultats de la fonction *LOI.STUDENT.BILATERALE* pour `5` degrés de liberté. Sa syntaxe apparaît en rouge dans le dernier cadre blanc. On constate aisément que ses résultats sont le double de ceux renvoyés par la fonction *LOI.STUDENT.DROITE* (la fonction est symétrique). On a symbolisé les deux queues de courbe correspondant au cumul des deux probabilités renvoyé par la fonction, en prenant `2` comme valeur de `t` et en hachurant en rouge la surface sous la courbe correspondant à la valeur de la cellule *J10*. *LOI.STUDENT* en mode bilatéral (c'est-à-dire avec son troisième argument égal à `2`) est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.
LOI.STUDENT .INVERSE .BILATERALE	La plage *L8:L13* de la figure 13-75 affiche les résultats de la fonction *LOI.STUDENT.INVERSE.BILATERALE*. À partir de `[1 - F(t)] + F(-t)` [`1` - probabilité de la fonction de répartition pour `t`] + probabilité de la fonction de répartition pour `-t`, *LOI.STUDENT.INVERSE.BILATERALE* renvoie la première valeur de `t` respectant `P(T ≥ t) + P(T < -t)`. Sa syntaxe apparaît en rouge dans le quatrième cadre blanc. *LOI.STUDENT.INVERSE* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

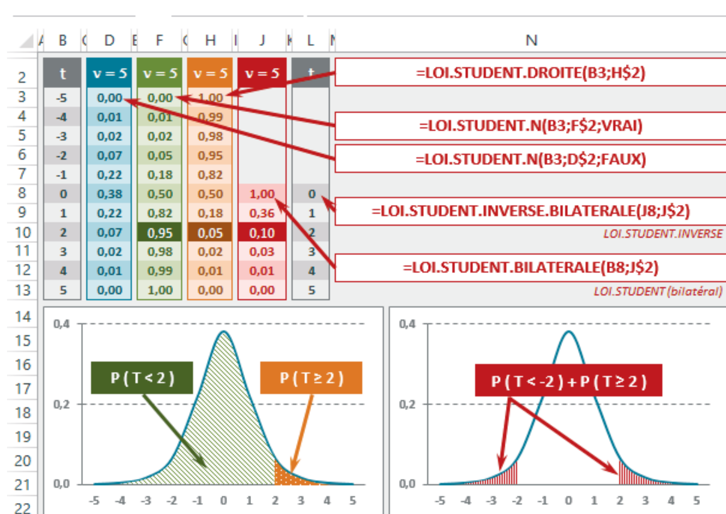
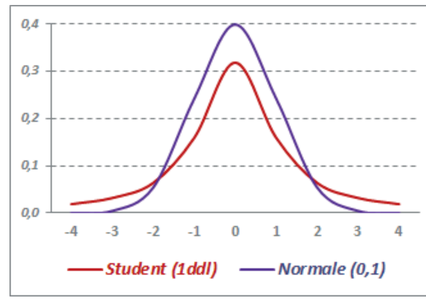


Figure 13–75
Mise en œuvre des fonctions LOI.STUDENT.DROITE, LOI.STUDENT.BILATERALE et LOI.STUDENT.INVERSE .BILATERALE.

CORRESPONDANCES Loi de Student, loi normale et loi de Fisher

Lorsque ν croît, la distribution de Student converge vers une loi normale. En pratique, les deux distributions sont très proches dès que $\nu > 30$.

Figure 13–76 Représentation graphique de la fonction de densité de probabilité de la loi normale centrée réduite et de la loi de Student à un degré de liberté.



Plus ν grandit, plus le sommet de la courbe rouge (Student) se rapproche de celui de la courbe violette (Normale) et plus sa base se resserre.
Le carré d'une variable aléatoire distribuée selon une loi de Student à ν degrés de liberté est distribué selon une loi de Fisher à $\nu_1 = 1$ et $\nu_2 = \nu$ degrés de liberté.

Loi de Fisher-Snedecor

Une variable aléatoire réelle X suit une loi de Fisher-Snedecor à ν_1 et ν_2 degrés de liberté, si elle admet pour densité de probabilité la fonction présentée figure 13-77.

Figure 13–77
Fonction de densité de la loi de Fisher-Snedecor.

$$f_{\nu_1,\nu_2}(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_2}{2}\right)\Gamma\left(\frac{\nu_1}{2}\right)} \left(\nu_1^{\frac{\nu_1}{2}}\ \nu_2^{\frac{\nu_2}{2}}\right) \left(\frac{x^{\frac{\nu_1}{2}-1}}{(\nu_1 x + \nu_2)^{\frac{\nu_1+\nu_2}{2}}}\right)$$

Γ est la fonction Gamma d'Euler
$\nu_1 \geq 1$ et $\nu_2 \geq 1$ (entiers)

COMPRENDRE D'où vient cette fonction ?

On considère deux variables aléatoires indépendantes Y_1 et Y_2, toutes deux suivant une loi du Khi-deux à respectivement ν_1 et ν_2 degrés de liberté. On s'intéresse à la variable X issue du rapport de Y_1 et Y_2, présentée figure 13-78.

Figure 13–78 X est une variable construite à partir du rapport des variables Y_1 et Y_2, toutes deux suivant une loi du Khi-deux à, respectivement, ν_1 et ν_2 degrés de liberté.

$$X = \frac{Y_1/\nu_1}{Y_2/\nu_2}$$

En passant par des relations intermédiaires faisant intervenir une loi Bêta de paramètres $\nu_1/2$ et $\nu_2/2$, de savantes démonstrations aboutissent à la définition de la fonction de densité de la loi de Fisher-Snedecor présentée figure 13-77.

La loi de Fisher-Snedecor dépend des paramètres ν_1 et ν_2, appelés nombre de degrés de liberté (voir, plus haut, l'aparté sur les degrés de liberté).

Lorsque $\nu_2 > 2$, l'espérance vaut $\nu_2 / (\nu_2 - 2)$. Lorsque $\nu_2 > 4$, la variance est égale à $2\nu_2^2(\nu_1 + \nu_2 - 2) / [\nu_1(\nu_2 - 2)^2 (\nu_2 - 4)]$.

> **En pratique À quoi sert-elle ?**
>
> À travers ce que l'on appelle « Test de Fisher-Snedecor », cette loi sert essentiellement à comparer les variances de deux populations. Pour y parvenir, on utilise les variances de deux petits échantillons tirés de ces populations, pour calculer une statistique appelée « F », censée suivre une loi de Fisher-Snedecor. La comparaison de ce `F` et de la valeur pour laquelle la fonction de répartition de Fisher-Snedecor renvoie une probabilité d'erreur acceptable (généralement 5 %) permet d'en déduire l'égalité (ou la différence) des variances des deux populations initiales (voir l'exemple développé à la fin de cet ouvrage dans la section traitant des tests statistiques).
> La statistique « F » est également utilisée pour tester la fiabilité du coefficient de détermination dans un calcul de régression (voir plus haut la section traitant de la régression multiple).

Tableau 13–18 Loi de Fisher-Snedecor

Fonction	Description
LOI.F.N	Si son quatrième argument est positionné sur `FAUX`, cette fonction renvoie `P(X = x)` et si son quatrième argument est positionné sur `VRAI`, elle renvoie `P(X < x)`. La figure 13-79 illustre un exemple d'application de cette fonction pour trois couples de valeurs ν_1 et ν_2. Sa syntaxe apparaît en rouge dans les deux derniers cadres blancs. La plage *D4:F16* donne les résultats de la fonction dans sa version non cumulative `P(X = x)` (valeurs représentées sur le graphique de gauche) et la plage *H4:J16* dans sa version cumulative `P(X < x)` (valeurs représentées sur le graphique de droite).
INVERSE.LOI.F.N	La plage *L4:L16* de la figure 13-79 affiche les résultats de la fonction *INVERSE.LOI.F.N*. À partir de `F(x)`, probabilité de la fonction de répartition, *INVERSE.LOI.F.N* renvoie la première valeur de `x` respectant `P(X < x)`. Sa syntaxe apparaît en rouge dans le premier cadre blanc.

Figure 13–79
Mise en œuvre des fonctions LOI.F.N et INVERSE.LOI.F.N.

Tableau 13–19 Loi de Fisher-Snedecor

Fonction	Description
LOI.F.DROITE	La fonction *LOI.F.N*, dans sa version cumulée, renvoie `P(X < x)`, c'est-à-dire la probabilité que la variable `X` prenne une valeur inférieure à `x`. Or, pour un test de Fisher-Snedecor, on n'a pas besoin de connaître `P(X < x)`, mais plutôt, `P(X ≥ x)` (queue de courbe à droite, renvoyée par *LOI.F.DROITE*). La plage *D4:D16* de la figure 13-80 donne les résultats de la fonction *LOI.F.N* pour 10 et 40 degrés de liberté dans sa version non cumulative (valeurs ayant permis de tracer la courbe du graphique) et la plage *F4:F16* donne les résultats de la même fonction dans sa version cumulative. Leur syntaxe apparaît dans les troisième et quatrième cadres blancs. La plage *H4:H16* donne les résultats de la fonction *LOI.F.DROITE* pour 10 et 40 degrés de liberté. Sa syntaxe apparaît en rouge dans le deuxième cadre blanc. On constate aisément que ses résultats ajoutés à ceux de la plage *F4:F16* donnent toujours 1. On a symbolisé cette relation dans le graphique en prenant 2 comme valeur de `x` et en hachurant en vert la surface sous la courbe correspondant à la valeur de la cellule *F12*. La partie orange, quant à elle, correspond à la valeur de la cellule *H12* (la somme de *F12* et *H12* faisant bien 1, mesure de la surface totale sous la courbe). *LOI.F* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.
INVERSE.LOI.F.DROITE	La plage *J4:J16* de la figure 13-80 affiche les résultats de la fonction *INVERSE.LOI.F.DROITE*. À partir de `1 - F(x)` (1 - probabilité de la fonction de répartition pour `x`), *INVERSE.LOI.F.DROITE* renvoie la première valeur de `x` respectant `P(X ≥ x)`. Sa syntaxe apparaît en rouge dans le premier cadre blanc. *INVERSE.LOI.F* est l'ancienne forme de cette fonction. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

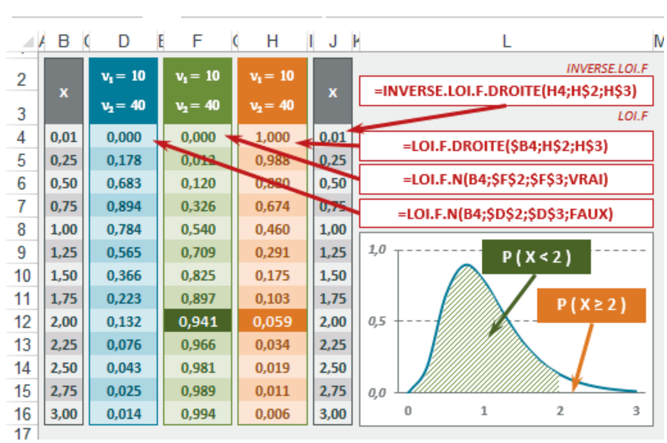


Figure 13–80
Mise en œuvre des fonctions LOI.F.DROITE et INVERSE.LOI.F.DROITE.

Indicateurs, tests et intervalles de confiance

Les fonctions afférentes aux tests, intervalles et indicateurs sont présentées ici, car elles font toutes référence aux distributions théoriques abordées dans les sections précédentes.

Aplatissement d'une courbe

> **PRÉCISION** **Pourquoi présenter cet indicateur (Kurtosis) ici ?**
>
> Le Kurtosis est un coefficient indiquant le degré d'aplatissement d'une courbe à l'aune de celui de la loi normale. À ce titre, sa place serait davantage au début de ce chapitre, parmi les indicateurs destinés à mieux cerner la forme d'une distribution (dispersion, coefficient d'asymétrie, etc.). Toutefois, la fonction *KURTOSIS* d'Excel faisant référence à la loi normale, il semblait plus intéressant de la présenter une fois tous les détails sur cette loi connus.

La fonction *KURTOSIS* utilise les valeurs d'une variable pour renvoyer un coefficient.

- Si ce coefficient est proche de zéro, on en conclut que le degré d'aplatissement de la courbe de la variable étudiée est proche de celui de la loi normale.
- Si ce coefficient est négatif, sa courbe est plus aplatie.
- Si ce coefficient est positif, sa courbe est plus pointue.

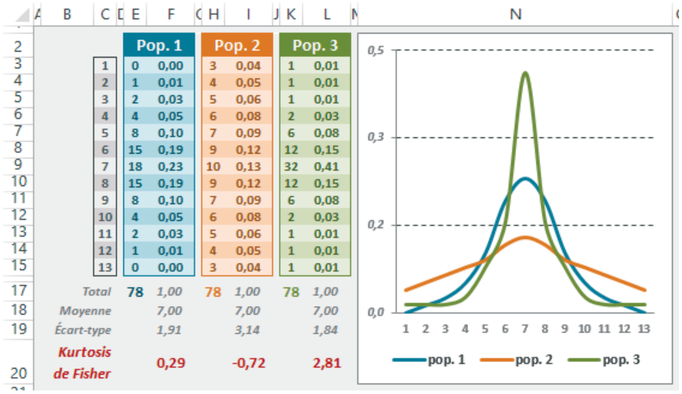
Pour illustrer ces différentes situations, on a mesuré la consommation annuelle de champagne (en nombre de bouteilles) dans trois échantillons de 78 ménages tirés de trois populations distinctes.

Figure 13–81
Nombre de bouteilles de champagne consommées annuellement par 78 ménages issus de 3 populations différentes.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
2		Population 1							Population 2							Population 3						
3		2	3	3	4	7	4		1	1	1	2	7	2		3	4	4	5	7	5	
4		4	4	5	5	7	5		2	2	3	3	7	3		5	5	5	5	7	11	
5		5	5	5	5	7	5		3	3	4	4	7	4		10	10	9	9	7	9	
6		12	11	11	10	7	10		4	4	4	5	7	5		9	9	9	6	7	6	
7		10	10	9	9	7	9		5	5	5	5	7	5		6	6	6	6	7	6	
8		9	9	9	9	7	9		13	13	13	12	7	12		6	6	6	6	7	6	
9		6	6	6	6	7	6		12	12	11	11	7	11		8	8	8	8	7	8	
10		6	6	6	6	7	6		11	11	10	10	7	10		8	8	8	8	7	8	
11		6	6	6	6	7	7		10	10	10	9	7	9		8	8	12	13	7	7	
12		8	8	8	8	7	8		9	9	9	9	8	9		7	7	7	7	7	7	
13		8	8	8	8	6	8		6	6	6	6	6	6		7	7	7	7	7	7	
14		8	8	8	8	7	7		6	6	6	8	7	8		7	7	7	7	7	7	
15		7	7	7	7	8	7		8	8	8	8	8	8		7	7	1	2	7	7	

Les plages *B3:G15*, *I3:N15* et *P3:U15* ont été nommées respectivement `Population1`, `Population2` et `Population3`.

À partir des valeurs de ces trois variables, on a calculé la fréquence des diverses modalités (`1` à `13`) ainsi que la moyenne, l'écart-type et le Kurtosis (figure 13-82).



	Pop. 1		Pop. 2		Pop. 3	
1	0	0,00	3	0,04	1	0,01
2	1	0,01	4	0,05	1	0,01
3	2	0,03	5	0,06	1	0,01
4	4	0,05	6	0,08	2	0,03
5	8	0,10	7	0,09	6	0,08
6	15	0,19	9	0,12	12	0,15
7	18	0,23	10	0,13	32	0,41
8	15	0,19	9	0,12	12	0,15
9	8	0,10	7	0,09	6	0,08
10	4	0,05	6	0,08	2	0,03
11	2	0,03	5	0,06	1	0,01
12	1	0,01	4	0,05	1	0,01
13	0	0,00	3	0,04	1	0,01
Total	78	1,00	78	1,00	78	1,00
Moyenne		7,00		7,00		7,00
Écart-type		1,91		3,14		1,84
Kurtosis de Fisher		0,29		-0,72		2,81

Figure 13–82
Écart-type, moyenne et Kurtosis des trois variables étudiées.

Pour calculer les indicateurs de la première variable, on a utilisé les formules suivantes :

- en *E3*, `=NB.SI(Population1;C3)`, formule recopiée ensuite dans la plage *E4:E15* ;
- en *E17*, `=SOMME(E3:E15)` et *F17*, `=SOMME(F3:F15)` ;
- en *F18*, `=MOYENNE(Population1)` ;
- en *F19*, `=ECARTYPE.STANDARD(Population1)` ;
- en *F20*, `=KURTOSIS(Population1)`.

Les trois courbes du graphique de la figure 13-82 ont été tracées à partir des plages *F3:F15*, *I3:I15* et *L3:L15*. On observe bien que la courbe de la variable correspondant à la population 3, dotée d'un Kurtosis de `2,81`, est assez pointue, alors que celle correspondant à la population 2, dotée d'un Kurtosis de `-0,71` est relativement aplatie. Celle qui correspond à la population 1, quant à elle, ressemble beaucoup à une distribution normale (Kurtosis de `0,29`).

COMPRENDRE Comment le Kurtosis est-il calculé ?

Le coefficient de Kurtosis est obtenu à partir de la moyenne des écarts des données à la moyenne, élevés à la puissance 4. Pour obtenir ensuite un nombre sans dimension, on divise le résultat par le carré de la variance (première formule présentée en rouge figure 13-83).

Cette formule, appliquée aux valeurs des trois variables, renvoie `3,11`, `2,19` et `5,42`. Ces coefficients diffèrent de ceux renvoyés par la fonction *KURTOSIS*. En effet, en statistiques, on manipule deux types de coefficients d'aplatissement : le Kurtosis et le Kurtosis de Fisher.

Le premier correspond à la première formule de la figure 13-83 et renvoie un coefficient qu'il s'agit ensuite de comparer à `3` (étalon correspondant au Kurtosis d'une distribution normale).

Le second correspond à la dernière formule de la figure 13-83, constituée de trois composantes dont le détail est fourni par les formules qui apparaissent en gris. Le Kurtosis de Fisher renvoie un coefficient qu'il s'agit ensuite de comparer à `0` (étalon correspondant au Kurtosis de Fisher d'une distribution normale). La fonction *KURTOSIS* d'Excel renvoie donc un Kurtosis de Fisher.

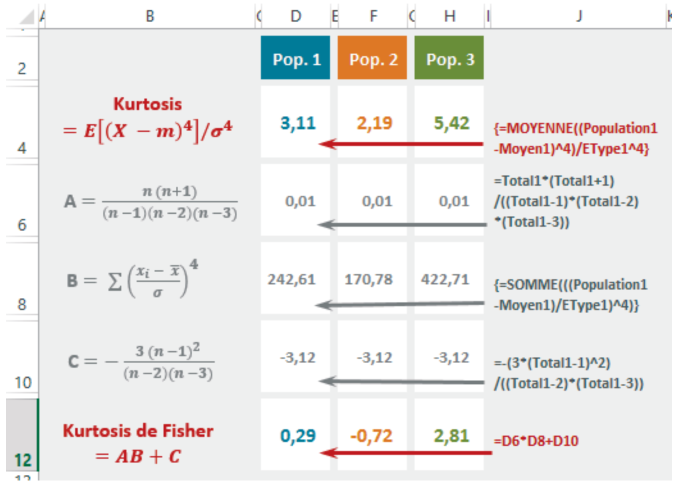


Figure 13–83 Décomposition du calcul d'un Kurtosis et d'un Kurtosis de Fisher.

Dans la figure 13-82 :

- les cellules *E17*, *H17* et *K17* ont été nommées `Total1`, `Total2` et `Total3` ;
- les cellules *F18*, *I18* et *L18* ont été nommées `Moyen1`, `Moyen2` et `Moyen3` ;
- les cellules *F19*, *I19* et *L19* ont été nommées `EType1`, `EType2` et `Etype3`.

Ces noms ont été utilisés pour construire les formules de la figure 13-83, dont la syntaxe est donnée en colonne *J*.

- En ligne *4*, on a calculé le Kurtosis en appliquant la formule indiquée en rouge colonne *B* (attention, il s'agit d'une formule matricielle, donc à valider avec les touches *Ctrl+Maj+Entrée*). Les résultats correspondent au Kurtosis des trois variables.
- En ligne *6*, on a calculé la première composante, A, du Kurtosis de Fisher en appliquant la formule indiquée en gris colonne *B*.
- En ligne *8*, on a calculé la deuxième composante, B, du Kurtosis de Fisher en appliquant la formule indiquée en gris colonne *B* (attention, il s'agit d'une formule matricielle, donc à valider avec les touches *Ctrl+Maj+Entrée*).
- En ligne *10*, on a calculé la troisième composante, C, du Kurtosis de Fisher en appliquant la formule indiquée en gris colonne *B*.
- Enfin, en ligne *12*, on a assemblé les trois composantes A, B et C pour calculer le Kurtosis de Fisher qui renvoie bien la même valeur que la fonction *KURTOSIS* d'Excel.

Tests d'hypothèses

L'étude de certaines populations, trop coûteuse, nécessite l'utilisation de techniques d'échantillonnage. Quelques tests ont été mis au point pour estimer ensuite dans quelle mesure les résultats obtenus sur les échantillons peuvent être étendus à la population. Ces tests permettent également de comparer plusieurs sous-populations, ou de vérifier l'adéquation à une loi de probabilité des données observées.

Un test sert à éprouver une hypothèse : on confronte une hypothèse H_0 (hypothèse à vérifier) à une hypothèse H_1 (contre-hypothèse). Pour mieux comprendre cette notion, voici quelques exemples d'hypothèses H_0 : « L'âge n'a pas d'effet sur le fait d'être locataire ou propriétaire », ou encore « La moyenne de la population 1 est égale à celle de la population 2 ». L'objectif du test est de savoir si, pour un niveau de confiance donné, on doit rejeter ou non H_0. Le risque d'erreur (α) couramment utilisé est `5 %` (on vise donc un niveau de confiance de `95 %`). Le risque de rejeter H_0 à tort est dit « de première espèce » et celui de l'accepter à tort est dit « de seconde espèce ».

> **TECHNIQUE Test unilatéral ou bilatéral (qualifier H_1)**
>
> Pour appréhender cette distinction, il faut bien comprendre que la formulation d'un test dépend énormément de la personne intéressée par son résultat. Prenons l'exemple d'un fabriquant de distributeurs de boissons automatiques. À travers un test, on souhaite vérifier la moyenne de `10 ml` par café annoncée par ce fabriquant. Quel que soit le point de vue du testeur, l'hypothèse H_0 sera $\mu = 10$. En revanche, H_1 sera formulée différemment pour un technicien ou un consommateur. Pour un technicien, H_1 sera $\mu \neq 10$ (test bilatéral) alors que pour un consommateur, H_1 sera $\mu < 10$ (test unilatéral). En effet, le technicien est intéressé par tout écart sur la moyenne annoncée : quantités trop petites (qui peuvent à la longue altérer la qualité du produit distribué) ou trop grandes (si le gobelet est insuffisamment grand et les débordements fréquents, la machine peut éventuellement tomber en panne plus souvent), alors que le consommateur est surtout motivé par le fait de ne pas être lésé. Il est donc essentiellement intéressé par les quantités inférieures à la moyenne annoncée (il n'est pas intéressé par la quantification des gobelets ayant reçu davantage de café).
>
> En résumé, on peut dire que lorsque les valeurs du paramètre étudié sous H_1 sont toutes plus grandes ou toutes plus petites que la valeur du paramètre sous H_0, le test est dit unilatéral ($H_0 : \mu = 10$ versus $H_1 : \mu > 10$ ou versus $H_1 : \mu < 10$), alors qu'avec $H_0 : \mu = 10$ versus $H_1 : \mu \neq 10$, le test est dit bilatéral.

On distingue les tests de conformité (comparaison d'un paramètre à une norme), les tests d'homogénéité (égalité entre deux paramètres observés), les tests d'adéquation à une distribution statistique et les tests d'indépendance (validation d'une liaison entre deux caractères).

Test de Student

Ce test a été conçu pour aider à comparer les moyennes de deux populations à partir des données de deux petits échantillons d'individus (n_1 et n_2) tirés au hasard et distribués selon une loi normale. Pour comprendre le fonctionnement de ce test, nous avons construit l'exemple présenté figure 13-84.

- La plage *B5:B33*, nommée `Poids1`, contient le poids des individus issus de la ville 1.
- La plage *F5:F33*, nommée `Poids2`, contient le poids des individus issus de la ville 2.
- La plage *I5:I33*, nommée `Poids3`, contient le poids des individus issus de la ville 3.

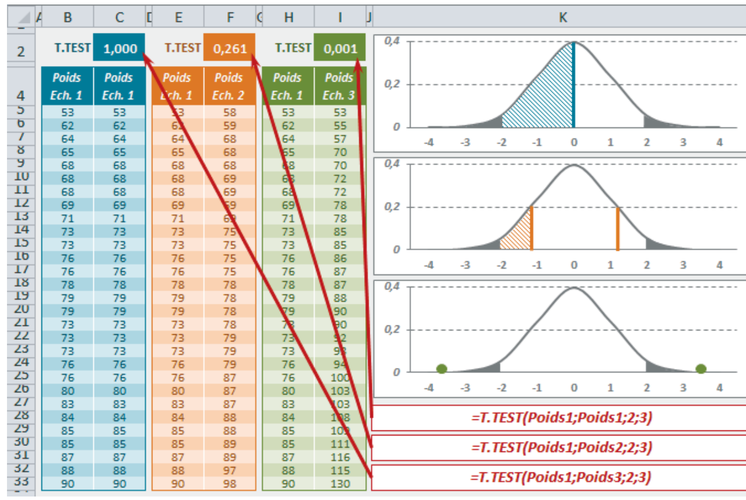


Figure 13–84
Poids des 29 individus de trois échantillons tirés de villes différentes.

À partir de ces trois échantillons, on souhaite tester les deux hypothèses H_0 suivantes, avec un risque d'erreur acceptable de 5 % :

- `Moyenne du poids des individus de la ville 1 = Moyenne du poids des individus de la ville 2` ;
- `Moyenne du poids des individus de la ville 1 = Moyenne du poids des individus de la ville 3`.

Pour mettre en œuvre ce test, Excel propose la fonction *T.TEST*. La cellule *F2* contient la formule `=T.TEST(Poids1;Poids2;2;3)` pour tester la première hypothèse ; la cellule *I2* contient la formule `=T.TEST(Poids1;Poids3;2;3)` pour tester la seconde hypothèse. À titre de référence, on a entré en *C2* la formule `=T.TEST(Poids1;Poids1;2;3)` qui donne `1`, probabilité que l'hypothèse H_0, `Moyenne du poids des individus de la ville 1 = Moyenne du poids des individus de la ville 1` soit vraie (… ce qui, bien entendu est un test inutile puisqu'il s'agit de la même ville, mais permet de mieux comprendre la nature du résultat renvoyé).

DÉTAIL Troisième et quatrième arguments

Les deux hypothèses H_1 sont : `Moyenne population ville 1 ≠ Moyenne population ville 2` et `Moyenne population ville 1 ≠ Moyenne population ville 3`. Nous sommes donc dans un test bilatéral ; c'est pourquoi le troisième argument est positionné sur 2 (voir l'aparté « Test unilatéral ou bilatéral »). S'il s'agissait d'un test unilatéral, il serait positionné sur `1`.
Les échantillons 1 et 2 n'ont pas la même variance (voir la figure 13-86) et les échantillons 1 et 3 non plus. C'est pourquoi le quatrième argument est positionné sur 3 (s'ils avaient eu la même variance, il aurait été positionné sur 2).

Le résultat de la cellule *F2* est `0,261` (`26,1 %`). Il signifie que l'on est encore au-dessus du seuil de `5 %` à partir duquel on a décidé de rejeter l'hypothèse H_0. On en conclut donc que la moyenne des poids est la même pour les populations des villes 1 et 2. Le résultat de la cellule *I2* est `0,001` (`0,1 %`). Il signifie que l'on est tombé sous le seuil de `5 %` à partir duquel on a décidé de rejeter l'hypothèse H_0. On en conclut donc que la moyenne des poids de la population de la ville 1 est différente de celle de la ville 3 (… davantage de fast-foods dans la ville 3 peut-être ?).

COMPRENDRE Sur quelle formule est basée la fonction T.TEST ?

La formule sur laquelle est fondée la fonction *T.TEST* est présentée figure 13-85. Elle utilise la taille, la moyenne et l'écart-type de deux échantillons dont la distribution suit une loi normale. De savantes démonstrations montrent que cette statistique suit une loi de Student à $n_1 + n_2 - 2$ degrés de liberté.

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{(n_1-1)\sigma_1^2+(n_2-1)\sigma_2^2}{n_1+n_2-2}\right)}} \sim T_{n_1+n_2-2}$$

n_1, n_2 Taille des deux échantillons
$\overline{X}_1, \overline{X}_2$ Moyenne des deux échantillons
σ_1, σ_2 Écart-type des deux échantillons
$T_{n_1+n_2-2}$ *Loi de Student à $n_1 + n_2$ -2 degrés de liberté*

Figure 13–85 Calcul de la statistique sur laquelle est fondée la fonction T.TEST.

Les lignes *4* à *9* de la figure 13-86 calculent la taille, la moyenne et la variance des trois échantillons (la syntaxe des formules entrées en colonne *F* apparaît en colonne *I*).

	B	C	D	E	F	H	I
2			Villes 1-1	Villes 1-2	Villes 1-3	Risque erreur	
4	*Taille 1er éch.*	n_1	29	29	29		=NB(Poids1)
5	*Taille 2è éch.*	n_2	29	29	29		=NB(Poids3)
6	*Moyenne 1er éch.*	$\overline{X}_1$	75,00	75,00	75,00		=MOYENNE(Poids1)
7	*Moyenne 2è éch.*	$\overline{X}_2$	75,00	77,79	88,86		=MOYENNE(Poids3)
8	*Variance 1er éch.*	S_1	73,07	73,07	73,07		=VAR.S(Poids1)
9	*Variance 2è éch.*	S_2	73,07	102,46	355,84		=VAR.S(Poids3)
11		*t*	0,000	1,135	3,604	2,002	=-ABS((F6-F7)/RACINE(((1/F4)+(1/F5))*(((F4-1)*F8)+((F5-1)*F9))/(F4+F5-2)))
12	*P(T > |t|)*		1,00	0,26	0,001	0,050	=LOI.STUDENT.BILATERALE(F11;F4+F5-2)

Figure 13–86 À partir de la définition du test, calcul de la statistique T pour les trois échantillons.

La ligne *11* présente le résultat de la statistique `T` appliquée aux valeurs des échantillons 1-1, 1-2 et 1-3. La ligne *12* calcule la probabilité `P(T > |t|)`, selon une loi de Student à `(29 + 29 - 2) = 56` degrés de liberté, en utilisant pour `t` la valeur calculée en ligne *11*. Pour comprendre l'utilisation de la fonction *LOI.STUDENT.BILATERALE*, consultez la section réservée à l'étude de la loi de Student. La cellule *H11* correspond à la statistique `T` (`2,002`) qui donne la valeur de seuil pour accepter ou rejeter H_0, c'est-à-dire `5 %` (`0,05`).
Les trois graphiques de la figure 13-84 donnent une illustration des trois probabilités renvoyées par *T.TEST*. On a représenté en gris les deux queues de courbe correspondant au seuil d'acceptation ou de rejet des hypothèses H_0. On constate que les valeurs de `t` matérialisées par les barres bleues (villes 1-1) et orange (villes 1-2) sont bien à l'intérieur du seuil, alors que la valeur de `t` pour les villes 1-3, représentée par un cercle vert, est située à l'extérieur de la zone d'acceptation.

TEST.STUDENT est l'ancienne forme de la fonction *T.TEST*. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Test sur la moyenne

Ce test a été conçu pour mesurer la validité de la moyenne annoncée pour une population, à l'aune de celle d'un échantillon de valeurs observées, tirées de cette population. Les valeurs de cet échantillon doivent être distribuées selon une loi normale et la mesure de la représentativité de leur moyenne réalisée avec la fonction *Z.TEST*.

Pour comprendre le fonctionnement de ce test, nous avons construit l'exemple présenté figure 13-87. Les valeurs qui apparaissent dans la plage *B2:K11* représentent un échantillon de `100` taux de nitrates mesurés sur plusieurs jours, en différents points d'un circuit de distribution. Le fournisseur s'étant engagé sur un taux moyen de 11 ne devant pas être dépassé, on cherche à vérifier l'hypothèse H_0 : `µ = 11`, avec un risque d'erreur acceptable de `0,05`. L'hypothèse H_1 est donc : `µ > 11`.

Figure 13–87 Échantillon de 100 taux de nitrates observés sur plusieurs points d'un circuit de distribution.

	B	C	D	E	F	G	H	I	J	K	M	N	O
2	9,8	10,5	10,7	10,9	11,0	11,1	11,3	11,4	11,5	11,8			
3	10,1	10,5	10,7	10,9	11,0	11,1	11,3	11,4	11,6	11,8			
4	10,2	10,5	10,8	10,9	11,0	11,1	11,3	11,4	11,6	11,8			
5	10,3	10,5	10,8	10,9	11,0	11,1	11,3	11,4	11,6	11,8			
6	10,3	10,6	10,8	10,9	11,0	11,1	11,3	11,4	11,6	11,9			
7	10,4	10,6	10,8	10,9	11,0	11,1	11,3	11,4	11,7	11,9	Seuil acceptable	µ	11,0
8	10,4	10,6	10,8	10,9	11,0	11,1	11,3	11,4	11,7	11,9			
9	10,4	10,6	10,8	10,9	11,0	11,2	11,3	11,4	11,7	12,0			
10	10,4	10,7	10,9	10,9	11,0	11,2	11,3	11,5	11,7	12,0	Test	Z	0,022
11	10,4	10,7	10,9	10,9	11,1	11,2	11,3	11,5	11,7	12,3			

La plage *B2:K11* a été nommée `Nitrates`. La cellule *O7*, qui contient le taux moyen annoncé par le fournisseur a été nommée `Seuil`. La cellule *O10* contient la formule `=Z.TEST(Nitrates;Seuil)`. Elle permet de tester l'hypothèse H_0. Son résultat est `0,022`. Cette valeur étant inférieure à `0,05`, on peut rejeter l'hypothèse H_0 et mettre en doute la fiabilité du fournisseur quant à la qualité de l'eau distribuée.

COMPRENDRE Sur quelle formule est basée la fonction Z.TEST ?

La formule sur laquelle est fondée la fonction *Z.TEST* est présentée figure 13-88. Elle utilise la moyenne théorique de la population ainsi que la taille et la moyenne de l'échantillon, dont la distribution suit une loi normale. Le σ qui apparaît au dénominateur est en priorité celui de la population si on le connaît ; toutefois, on ne le connaît généralement pas et on utilise celui ayant été calculé à partir des valeurs de l'échantillon. De savantes démonstrations montrent que cette statistique suit une loi normale centrée réduite.

Figure 13–88 Calcul de la statistique sur laquelle est fondée Z.TEST.

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

$\bar{X}$ Moyenne de l'échantillon
μ Moyenne de la population
σ Écart-type de la population
n Taille de l'échantillon

À la figure 13-89, en *D2*, *D4* et *D6*, on a calculé la taille, la moyenne et l'écart-type de l'échantillon. En *I2* figure la formule renvoyant la valeur de *Z* (pour cet échantillon) à partir de ces trois calculs partiels. La colonne *J* affiche la syntaxe des formules entrées en colonne *I*.

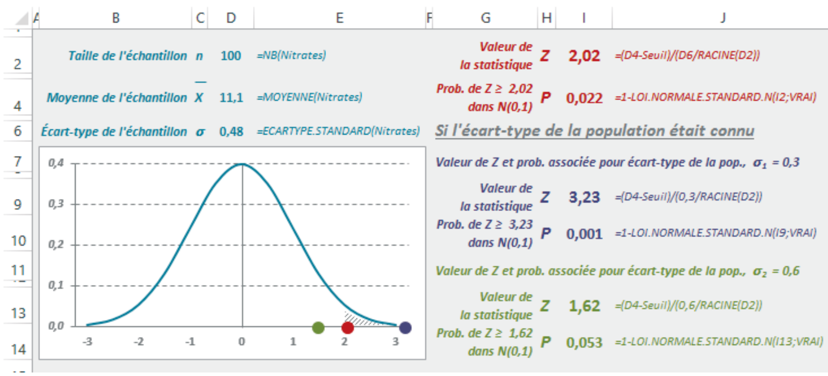


Figure 13–89 *À partir de la définition du test, calcul de la statistique Z pour l'échantillon étudié.*

Pour notre échantillon de `100` valeurs et pour une moyenne de la population annoncée de `11`, `Z` vaut `2,02`. Il faut donc calculer la probabilité que `Z` dépasse cette valeur selon la loi normale centrée réduite. En utilisant la fonction *LOI.NORMALE.STANDARD.N* dans sa version cumulée (voir la section consacrée à l'étude de la loi normale), la probabilité retournée en *I4* est `0,022`, valeur identique à celle renvoyée par la fonction *Z.TEST*.

La fonction *Z.TEST* accepte un troisième argument facultatif qui est l'écart-type de la population lorsque ce dernier est connu (on a vu que lorsqu'il ne l'est pas, c'est l'écart-type de l'échantillon qui est utilisé). Afin de voir l'effet de cette valeur sur la probabilité retournée, on a calculé deux autres statistiques `Z` en prenant un écart-type de la population inférieur à celui de l'échantillon (`0,3`) et un écart-type supérieur (`0,6`). Dans le premier cas, la probabilité retournée est encore bien plus petite que précédemment (on est donc encore davantage amené à rejeter l'hypothèse H_0). Dans le deuxième cas, la probabilité retournée est plus grande et atteint une valeur très légèrement supérieure à 5 %, ce qui nous autoriserait à ne pas rejeter H_0.

Le graphique de la figure 13-89 donne une illustration des trois probabilités calculées (avec l'écart-type de l'échantillon, un écart-type pour la population de `0,3` et un autre de `0,6`). On a représenté en gris la queue de courbe correspondant au seuil d'acceptation ou de rejet de l'hypothèse H_0). On a matérialisé par des cercles de couleur la position des valeurs de `Z` correspondant aux trois calculs. Les pastilles rouge et mauve se situent bien dans la zone de rejet alors que la pastille verte est placée dans la zone d'acceptation.

TEST.Z est l'ancienne forme de la fonction *Z.TEST*. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Test de Fisher-Snedecor

À partir de deux échantillons indépendants (n_1, n_2) distribués selon une loi normale, le test de Fisher-Snedecor indique si, avec un risque d'erreur acceptable, les variances des populations dont ils sont issus peuvent être considérées comme identiques. La statistique utilisée ici est construite à partir du rapport des variances des deux échantillons (la plus importante étant placée au dénominateur). De savantes démonstrations montrent que cette statistique suit une loi de Fisher-Snedecor à $n_1 - 1$ et $n_2 - 1$ degrés de liberté (voir la section réservée à l'étude de la loi de Fisher-Snedecor).

Pour comprendre ce test, nous avons construit l'exemple présenté figure 13-90. À partir de deux échantillons de 50 femmes et 50 hommes habitant tous la même ville, on souhaite savoir, pour l'ensemble de la ville, si la variance de la variable « Taille » pour la population masculine est la même que celle de la population féminine. La plage réunissant les tailles de l'échantillon féminin (*C2:L6*) a été nommée `Femmes` et celle qui réunit les tailles de l'échantillon masculin (*C8:L12*) a été nommée `Hommes`.

Figure 13–90
Taille de 50 hommes et 50 femmes constituant deux échantillons dont les valeurs suivent une loi normale.

	B	C	D	E	F	G	H	I	J	K	L
2	Femmes	147	155	158	160	161	163	164	165	167	171
3		150	156	158	160	162	163	164	166	167	175
4		151	154	158	160	162	163	164	166	168	172
5		152	156	158	160	162	163	165	166	168	174
6		153	155	158	157	159	163	166	169	169	180
8	Hommes	162	171	173	175	176	177	182	180	183	186
9		163	171	173	175	176	177	179	181	183	187
10		165	172	174	175	177	178	179	181	184	187
11		167	172	174	175	176	178	179	181	184	188
12		169	172	174	175	177	178	179	185	191	189
14					Test F	0,81					
15					=F.TEST(Femmes;Hommes)						

L'hypothèse H_0 que nous souhaitons évaluer avec *F.TEST* est donc « variance de la variable `Taille des hommes habitant la ville` = variance de la variable `Taille des femmes habitant la ville` ». La formule `=F.TEST(Femmes;Hommes)` entrée en *G14* renvoie `0,81` (`81 %`), qui est bien supérieur à `5 %` (pourcentage d'erreur généralement retenu pour considérer comme vraie l'hypothèse H_0). On peut donc affirmer que la variance de la taille des hommes de cette ville est la même que celle de la taille des femmes.

COMPRENDRE Sur quelle formule est basée la fonction F.TEST ?

Le test de Fisher-Snedecor étant calculé à partir du rapport des variances de deux échantillons, on a entré, dans la figure 13-91, en *D2* et *D6*, les formules calculant ces deux variances. La syntaxe des formules apparaît en colonne *F*. La cellule *D10* affiche la valeur de ce rapport, `0,93` (on place toujours la variance la plus grande au dénominateur).

On utilise ensuite *LOI.F.N*, présentée dans la section réservée à l'étude de la loi de Fisher-Snedecor. Avec les degrés de liberté `49` et `49` ($n_1 - 1$ et $n_2 - 1$), on obtient bien `0,81` (ce qui est conforme au résultat renvoyé par la fonction *F.TEST*). On utilise ici `2*LOI.F.N` car on travaille avec une probabilité bilatérale (l'hypothèse H_1 étant $\sigma_F \neq \sigma_h$, le test est bilatéral).

	B	D	F
2	σ_f	44,34	=VAR.S(Femmes)
4	n_f	50,00	=NB(Femmes)
6	σ_h	41,44	=VAR.S(Hommes)
8	n_h	50,00	=NB(Hommes)
10	$f\ (\sigma_h / \sigma_h)$	0,93	=D6/D2
12	$P\ (f_{49,49} \geq 0{,}93)$	0,81	=2*LOI.F.N(D10;D8-1;D4-1;VRAI)
14	$P\ (f_{3000,3000} \geq 0{,}93)$	0,06	=2*LOI.F.N(D10;3000;3000;VRAI)

Figure 13–91 *À* partir de la définition du test, calcul de la statistique F pour l'échantillon étudié.

Mesurer la validité de la statistique F à l'aune de la loi de Fisher-Snedecor permet de tenir compte de la taille de l'échantillon utilisé. De manière un peu simpliste, on peut dire qu'une légère différence observée sur la variance de deux petits échantillons devient de moins en moins acceptable avec de grands échantillons. Avec les deux variances de notre exemple (41,44 et 44,34), on peut dire que le même rapport de variances (0,93) observé sur deux échantillons de 3 000 individus nous rapprocherait de la probabilité critique de 5 % (voir les calculs effectués dans la cellule *D14*). À partir d'un échantillon de 3 350 individus, nous serions amenés à rejeter l'hypothèse H_0.

TEST.F est l'ancienne forme de la fonction *F.TEST*. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Test d'ajustement du Khi-deux

Ce test a été conçu pour vérifier que les valeurs observées dans un échantillon se distribuaient selon une loi normale (condition sur laquelle reposent les tests de Student, du Z et de Fisher). Plus généralement, le test d'ajustement du Khi-deux permet de confronter l'hypothèse H_0, « la variable étudiée obéit à la distribution théorique spécifiée », à la contre-hypothèse H_1, « la variable étudiée n'obéit pas à la distribution théorique spécifiée ». Pour mettre en œuvre ce test, Excel propose la fonction *CHISQ.TEST*.

Concrètement, pour mettre en œuvre ce test, il faut regrouper les données de l'échantillon en n classes et comparer les effectifs observés dans chacune des classes (O_i) avec les effectifs espérés théoriquement (T_i) pour que H_0 soit vraie. On peut alors calculer la statistique d'ajustement dont la formule est donnée figure 13-92.

Figure 13–92
Statistique d'ajustement du Khi-deux.

$$\sum_{i=0}^{n} \frac{(O_i - T_i)^2}{T_i} \sim \chi^2_{n-1-r}$$

CRITÈRES Conditions d'utilisation de ce test

On démontre qu'à partir d'une taille d'échantillon supérieure ou égale à 30 et avec des effectifs espérés tous supérieurs ou égaux à 5, la statistique d'ajustement obéit, sous H_0, à une loi du Khi-deux à n-1-r degrés de liberté (r représentant le nombre de paramètres qu'il a fallu définir pour pouvoir calculer les effectifs théoriques).

Pour comprendre le fonctionnement de ce test, nous avons construit l'exemple présenté figure 13-93. Un restaurateur propose dix plats. Il souhaite savoir si le choix de ses clients se répartit équitablement entre ces dix plats. Il procède à deux observations : la première un mercredi midi et la seconde un samedi soir. La plage *C3:D12* contient les résultats de ces deux observations (à chaque fois, il a servi 250 clients).

Figure 13–93
Un restaurateur observe les choix de ses clients relativement aux dix plats proposés dans sa carte.

	B	C	D	F	H
2	N° plat	Choix observés 1	Choix observés 2	Choix théoriques	
3	1	26	31	25,0	=TotPlats/TotChoix
4	2	24	24	25,0	
5	3	31	31	25,0	
6	4	18	18	25,0	
7	5	26	26	25,0	
8	6	15	13	25,0	
9	7	35	36	25,0	
10	8	21	18	25,0	
11	9	24	21	25,0	
12	10	30	32	25,0	
14	10	250	250	250	=SOMME(Théoriques)
16		0,154			=CHISQ.TEST(Observés1;Théoriques)
17			0,017		=CHISQ.TEST(Observés2;Théoriques)
18					

La plage *C3:C12* a été nommée `Observés1`, la plage *D3:D12* `Observés2`, la cellule *B14* `TotChoix`, la cellule *C14* `TotPlats` et la plage *F3:F12* `Théoriques`. Dans cette dernière, le restaurateur a calculé les effectifs théoriques correspondant à l'hypothèse H_0 : « le choix des clients se répartit équitablement sur les dix plats ». La syntaxe de la formule utilisée dans cette plage est donnée cellule *H3*.

CRITÈRES Vérifier que l'on est bien dans les conditions d'utilisation du test

Avant d'aller plus loin, on vérifie bien que l'exemple se prête au test d'ajustement du Khi-deux :
- la taille de l'échantillon est supérieure à `30` (on a observé `250` individus) ;
- tous les effectifs espérés sont supérieurs ou égaux à `5` (ils valent tous `25`).

On peut donc procéder au test.

En *C16*, on a entré la formule `=CHISQ.TEST(Observés1;Théoriques)` qui renvoie `0,154` et en *D17*, la formule `=CHISQ.TEST(Observés2;Théoriques)` qui renvoie `0,017`. En fixant à 5 % le risque d'accepter H_0 à tort, on peut en conclure que le mercredi midi, H_0 est vraie (le choix des clients se répartit équitablement entre les dix plats) car `0,154 > 0,05`, alors qu'en considérant les résultats du samedi soir, il faut rejeter H_0 (`0,017 < 0,05`).

COMPRENDRE Sur quels calculs est basée la fonction CHISQ.TEST ?

La formule sur laquelle est fondée la fonction *CHISQ.TEST* est présentée figure 13-92. À la figure 13-94, on l'a appliquée en *D2* et *D3* aux deux échantillons observés (attention, il s'agit de formules matricielles qu'il faut donc valider en pressant les touches *Ctrl+Maj+Entrée*).

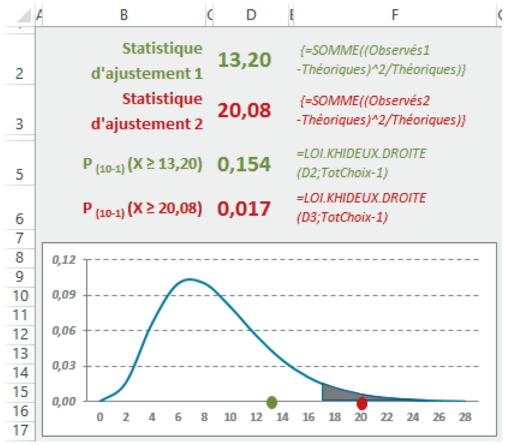


Figure 13–94
Calcul de la statistique sur laquelle est fondée CHISQ.TEST.

Comme on l'a annoncé dans l'introduction, cette statistique suit une loi du Khi-deux à `n-1-r` degrés de liberté. Les effectifs théoriques ayant pu être directement calculés sans fixer de paramètre quelconque, `r = 0`. Il faut donc mesurer les deux statistiques calculées en *D2* et *D3* à l'aune d'une loi du Khi-deux à 9 degrés de liberté (`10-1`).
Pour tracer la courbe du graphique, on a utilisé la fonction *LOI.KHIDEUX.N* (sous sa forme non cumulative) avec 9 degrés de liberté. La surface sous la courbe correspondant à la probabilité P_9 `(X > 16,92) = 0,05`, c'est-à-dire la zone de rejet de l'hypothèse H_0, a été grisée. On a utilisé en *D5* la formule `=LOI.KHIDEUX.DROITE(D2;TotChoix-1)` pour obtenir la probabilité correspondant à la première statistique et, en *D6*, la formule `=LOI.KHIDEUX.DROITE(D3;TotChoix-1)` pour obtenir la probabilité correspondant à la deuxième statistique (voir la section consacrée à l'étude de la loi du Khi-deux). Les deux valeurs retournées sont bien les mêmes que celles de la fonction *CHISQ.TEST*.
Sur le graphique, on a matérialisé par des points de couleur la valeur des deux statistiques. On observe bien que la première (`13,20`) se trouve dans la zone d'acceptation, alors que la seconde (`20,08`) est située dans la zone de rejet de l'hypothèse H_0.

Test d'indépendance du Khi-deux

Le test du Khi-deux peut également servir à mesurer l'indépendance de deux caractères pris par une même population. En d'autres termes, ce test indique si la valeur prise pour l'un des caractères influence celle qui est prise pour l'autre. Il est souvent utilisé pour croiser des tranches d'âge et des niveaux de revenus, ou encore des niveaux de scolarité et des types de sports ou de voyages et, même, le sexe des individus et leurs opinions politiques. Le test d'indépendance utilise la même statistique que le test d'ajustement, c'est-à-dire celle qui est basée sur la formule présentée figure 13-92.

Pour comprendre le fonctionnement de ce test, nous avons construit l'exemple présenté figure 13-96. À partir d'un échantillon de 2 500 individus, on cherche à savoir si le fait d'être propriétaire ou locataire est lié à l'âge. À partir des données observées (plage *D3:E6* nommée `Observés`), on a fait les totaux en ligne et en colonne (plages *F3:F7* et *D7:F7*). Dans le tableau vert, on a utilisé ces totaux pour construire les valeurs théoriques de la plage *D9:E12* (nommée `Théo`). La syntaxe de la formule entrée en *D9* apparaît en *H9*. Elle a ensuite été recopiée dans la plage *D9:E12*.

EN PRATIQUE Calculer le tableau des effectifs théoriques

Concrètement, pour mettre en œuvre ce test, il faut regrouper les données de la variable correspondant au premier caractère en M classes, celles de la variable correspondant au deuxième caractère en N classes, puis calculer les effectifs croisés O_{mn} (voir figure 13-95).

Effectifs observés

		Caractère B						
		B_1	B_2	...	B_n	...	B_N	Total
Caractère A	A_1	O_{11}	O_{12}		O_{1n}		O_{1N}	$O_{1\bullet}$
	A_2	O_{21}	O_{22}		O_{2n}		O_{2N}	$O_{2\bullet}$
	...							...
	A_m	O_{m1}	O_{m2}		O_{mn}		O_{mN}	$O_{m\bullet}$
	...							...
	A_M	O_{M1}	O_{M2}		O_{mn}		O_{MN}	$O_{M\bullet}$
	Total	$O_{\bullet 1}$	$O_{\bullet 2}$	...	$O_{\bullet n}$	...	$O_{\bullet N}$	O_{MN}

Effectifs théoriques

		Caractère B						
		B_1	B_2	...	B_n	...	B_N	Total
Caractère A	A_1	$(O_{1\bullet} * O_{\bullet 1}) / O_{MN}$	$(O_{1\bullet} * O_{\bullet 2}) / O_{MN}$		$(O_{1\bullet} * O_{\bullet n}) / O_{MN}$		$(O_{1\bullet} * O_{\bullet N}) / O_{MN}$	$O_{1\bullet}$
	A_2	$(O_{2\bullet} * O_{\bullet 1}) / O_{MN}$	$(O_{2\bullet} * O_{\bullet 2}) / O_{MN}$		$(O_{2\bullet} * O_{\bullet n}) / O_{MN}$		$(O_{2\bullet} * O_{\bullet N}) / O_{MN}$	$O_{2\bullet}$
	...							...
	A_m	$(O_{m\bullet} * O_{\bullet 1}) / O_{MN}$	$(O_{m\bullet} * O_{\bullet 2}) / O_{MN}$		$(O_{m\bullet} * O_{\bullet n}) / O_{MN}$		$(O_{m\bullet} * O_{\bullet N}) / O_{MN}$	$O_{m\bullet}$
	...							...
	A_M	$(O_{M\bullet} * O_{\bullet 1}) / O_{MN}$	$(O_{M\bullet} * O_{\bullet 2}) / O_{MN}$		$(O_{M\bullet} * O_{\bullet n}) / O_{MN}$		$(O_{M\bullet} * O_{\bullet N}) / O_{MN}$	$O_{M\bullet}$
	Total	$O_{\bullet 1}$	$O_{\bullet 2}$	...	$O_{\bullet n}$	...	$O_{\bullet N}$	O_{MN}

Figure 13–95
Tableaux des effectifs observés et théoriques nécessaires au calcul d'un test d'indépendance du Khi-deux.

Une fois ce premier tableau établi (plage *D5:I10*), il faut calculer les T_{mn}, effectifs théoriques obtenus à partir du produit des totaux en lignes et en colonnes, ramené à l'effectif total (plage *D16:I21*). La suite se déroule comme pour le test d'ajustement, en utilisant comme valeurs observées les données du premier tableau et comme valeurs théoriques celles du deuxième tableau. La valeur de la statistique obtenue doit être mesurée à l'aune d'une loi du Khi-deux à $(M-1) \times (N-1)$ degrés de liberté.

		Propriétaires	Locataires	Total
Observations	20-29	240	260	500
	30-44	375	375	750
	45-59	431	319	750
	60 et +	268	232	500
	Total	1 314	1 186	2 500
V. théoriques	20-29	262,8	237,2	500
	30-44	394,2	355,8	750
	45-59	394,2	355,8	750
	60 et +	262,8	237,2	500
	Total	1 314	1 186	2 500

=SOMME(D3:E3)
=SOMME(F3:F6)
=$F3*D$7/$F$7
CHISQ.TEST = 0,0035
=CHISQ.TEST(D3:E6;D9:E12)

Figure 13–96
Le tableau gris réunit les données observées et le tableau vert, les valeurs théoriques.

CRITÈRES Vérifier que l'on est bien dans les conditions d'utilisation du test

Avant d'aller plus loin, on vérifie bien que l'exemple se prête au test d'indépendance du Khi-deux :

- la taille de l'échantillon est supérieure à 30 (on a observé 2 500 individus) ;
- tous les effectifs espérés sont supérieurs ou égaux à 5 (ils valent au minimum 237).

On peut donc procéder au test.

En *H12*, on a entré la formule `=CHISQ.TEST(Observés;Théo)` qui renvoie `0,0035`. En fixant à 5 % le risque d'accepter H_0 à tort (H_0 = « L'âge n'a pas d'influence sur le fait d'être propriétaire ou locataire »), on peut en conclure que l'âge a une influence, car `0,0035 < 0,05`. Il faut donc rejeter H_0.

COMPRENDRE Détail du calcul

À la figure 13-97, on a appliqué la formule de calcul de la statistique du Khi-deux en *C2* (attention, il s'agit d'une formule matricielle, qu'il faut donc valider en pressant les touches *Ctrl+Maj+Entrée*).

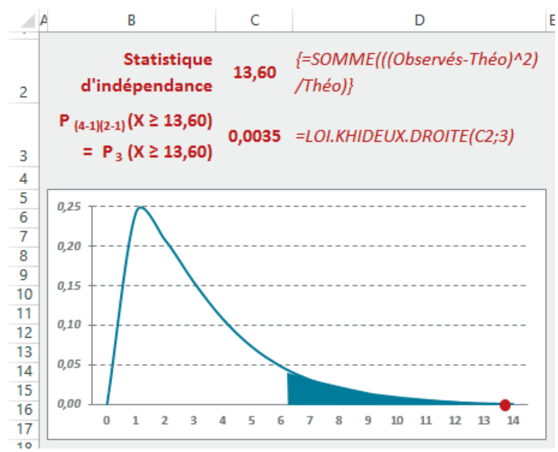


Figure 13–97
Calcul de la statistique sur laquelle est fondée CHISQ.TEST.

Dans le cadre d'un test d'indépendance, cette statistique suit une loi du Khi-deux à `(M-1) × (N-1) = (4-1) × (2-1) = 3` degrés de liberté. Il faut donc mesurer la statistique calculée en *C2* à l'aune d'une loi du Khi-deux à 3 degrés de liberté.

Pour tracer la courbe du graphique, on a utilisé la fonction *LOI.KHIDEUX.N* (sous sa forme non cumulative) avec 3 degrés de liberté. La surface sous la courbe correspondant à la probabilité P_3 `(X ≥ 7,8) = 0,05`, c'est-à-dire la zone de rejet de l'hypothèse H_0, a été grisée. En *C3*, on a utilisé la formule `=LOI.KHIDEUX.DROITE(C2;3)` pour obtenir la probabilité correspondant à cette statistique (voir la section consacrée à l'étude de la loi du Khi-deux). La valeur retournée correspond bien au résultat de *CHISQ.TEST*.

Sur le graphique, on a matérialisé par un point rouge la valeur de la statistique (`13,60`). On observe qu'elle se situe dans la zone de rejet de l'hypothèse H_0.

TEST.KHIDEUX est l'ancienne forme de la fonction *CHISQ.TEST*. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Intervalles de confiance

L'objet de nombreuses études statistiques est de faire des estimations sur une population à partir d'observations réalisées sur un échantillon. La valeur (`V`) obtenue à l'issue des calculs n'est jamais une certitude. En revanche, si la variable étudiée (sur l'échantillon) suit une distribution théorique connue (loi normale ou de Student), on peut définir un intervalle dans lequel `V` a `x %` de chances de se trouver. On appelle ce dernier « intervalle de confiance », et `x %` représente le niveau de confiance.

α (niveau de signification) représente la probabilité d'erreur, c'est-à-dire la probabilité que `V` ne se situe pas dans l'intervalle de confiance. En pratique, on prend souvent `α = 0,01` ou `α = 0,05`, ce qui donne un niveau de confiance égal à `1-α` de `99 %` ou `95 %`.

Évaluer et encadrer la moyenne d'une population

Un constructeur garantit une machine dix ans. On étudie deux échantillons de `25` et `100` machines dont on mesure la durée de vie. À la suite de cela, on calcule la moyenne obtenue sur chaque échantillon afin d'en extrapoler une moyenne « théorique » pour l'ensemble des machines fournies par le fabriquant. C'est autour de cette moyenne « théorique » que l'on souhaite définir un intervalle de confiance. Pour calculer cet intervalle de confiance, Excel offre deux fonctions : *INTERVALLE.CONFIANCE.NORMAL* et *INTERVALLE.CONFIANCE.STUDENT*. Si votre échantillon est suffisamment grand (`> 100`), optez pour la première, sinon choisissez la seconde.

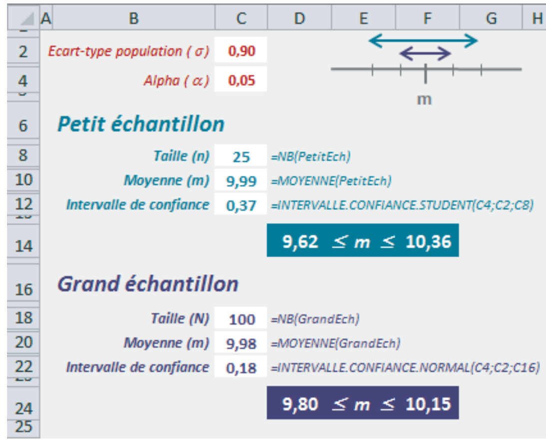
	A	B	C	D	E	F
2		6,8	8,9	9,8	10,4	11,1
3		7,6	9,1	9,8	10,4	11,2
4		8,2	9,3	10,1	10,5	11,9
5		8,8	9,6	10,2	10,9	12,3
6		8,8	9,7	10,4	10,9	13,0
8		6,8	9,1	9,8	10,2	10,9
9		7,6	9,1	9,8	10,2	10,9
10		7,7	9,3	9,8	10,2	11,2
11		8,1	9,3	9,8	10,3	11,2
12		8,2	9,3	9,8	10,3	11,2
13		8,2	9,3	9,8	10,3	11,2
14		8,6	9,4	9,8	10,3	11,2
15		8,6	9,4	9,8	10,4	11,2
16		8,6	9,4	9,9	10,4	11,3
17		8,6	9,4	9,9	10,4	11,3
18		8,6	9,6	9,9	10,7	11,3
19		8,6	9,6	10,0	10,7	11,3
20		8,6	9,6	10,0	10,7	11,3
21		8,8	9,6	10,0	10,7	11,3
22		8,8	9,6	10,1	10,7	11,9
23		8,8	9,6	10,1	10,7	11,9
24		8,8	9,6	10,1	10,9	12,1
25		8,8	9,6	10,1	10,9	12,1
26		9,1	9,7	10,1	10,9	12,3
27		9,1	9,7	10,1	10,9	13

Figure 13–98
Le premier tableau (bleu) réunit les durées de vie des 25 machines constituant le premier échantillon et le second tableau (violet), celles des 100 machines constituant le second échantillon.

Le premier tableau (plage *B2:F6*) a été nommé `PetitEch` et le second (plage *B8:F27*) `GrandEch`. Dans la figure 13-99, les cellules *C2* et *C4* donnent les paramètres nécessaires au calcul des fonctions *INTERVALLE.CONFIANCE.NORMAL* et *INTERVALLE.CONFIANCE.STUDENT*. Il s'agit de l'écart-type de la population (si vous ne le connaissez pas, utilisez celui de l'échantillon) et de la valeur de α (`5 %`), dont on déduit le niveau de confiance de l'intervalle trouvé (`1-α = 95 %`).

On a donc utilisé *INTERVALLE.CONFIANCE.STUDENT* (`0,37` en *C12*) pour encadrer la moyenne trouvée à partir du petit échantillon et *INTERVALLE.CONFIANCE.NORMAL* (`0,18` en *C22*) pour encadrer la moyenne trouvée à partir du grand échantillon. On obtient l'intervalle en soustrayant et en ajoutant cette valeur à la moyenne.

Figure 13–99
L'intervalle de confiance encadrant la moyenne à partir du petit échantillon est à peu près deux fois plus grand que l'intervalle calculé à partir du grand échantillon.



COMPRENDRE Les fondements du calcul

Dans les deux cas, le calcul est basé sur le rapport de l'écart-type de la population sur la racine carrée de la taille de l'échantillon. La différence se situe au niveau de la loi de probabilité utilisée pour renvoyer la valeur correspondant à un risque d'erreur assumé de `0,05`.

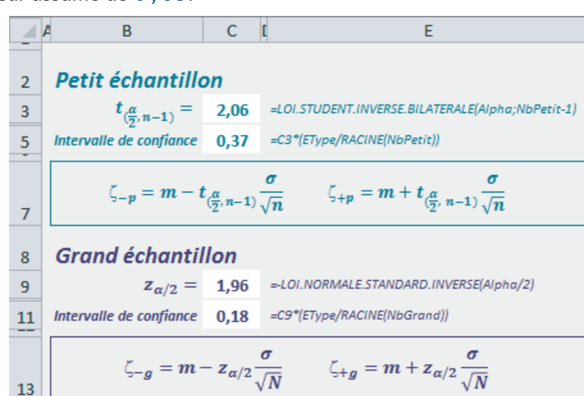


Figure 13–100 Calcul de la valeur de x pour que P(X < |x|) = 95 % à partir de la *loi normale* centrée réduite (1,96) et de la loi de Student (2,06).

En *C3*, on a calculé la valeur de `t`, pour qu'à l'aune d'une loi de Student à `n-1` (`24`) degrés de liberté, on ait `P(T < |t|) = 95 %`. La cellule affiche `2,06`, ce qui signifie que `P (-2,06 < T < 2,06) = 95 %`. En *C9*, on a calculé la valeur de `z` pour qu'à l'aune d'une loi normale centrée réduite, on ait `P(Z < |z|) = 95 %.` La cellule affiche `1,96`, ce qui signifie que `P (-1,96 < Z < 1,96) = 95 %`. Pour obtenir l'intervalle de confiance, il suffit d'appliquer les formules présentées en lignes *7* et *13*. Vous constatez que vous obtenez bien les mêmes résultats qu'avec les fonctions *INTERVALLE.CONFIANCE.NORMAL* et *INTERVALLE.CONFIANCE.STUDENT*.

INTERVALLE.CONFIANCE est l'ancienne forme de la fonction *INTERVALLE.CONFIANCE.NORMAL*. Elle est conservée dans Excel 2010 et 2013 pour assurer la compatibilité avec les versions antérieures.

Calculer et encadrer le coefficient de corrélation

COMPRENDRE La transformation de Fisher

La transformation de Fisher est une fonction qui transforme une distribution asymétrique en une distribution se rapprochant d'une loi normale (voir figure 13-101).

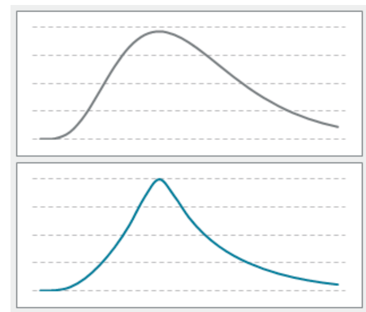
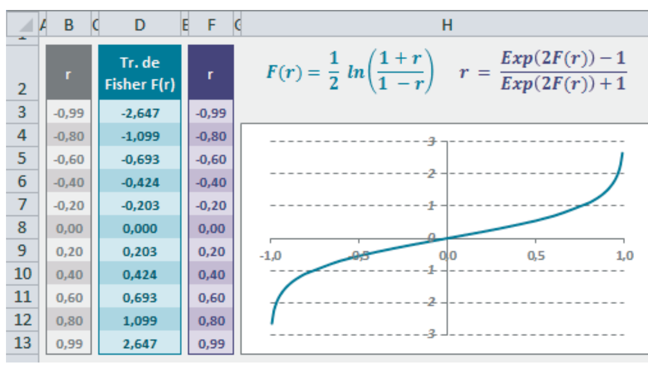


Figure 13–101 Distribution de Fisher (LOI.F.N) à 11 et 80 degrés de liberté (première courbe) passée par le filtre d'une transformation de Fisher (FISHER) (seconde courbe).

La fonction *FISHER*, disponible dans Excel, assure cette transformation (voir figure 13-102).



	r	Tr. de Fisher F(r)	r
3	-0,99	-2,647	-0,99
4	-0,80	-1,099	-0,80
5	-0,60	-0,693	-0,60
6	-0,40	-0,424	-0,40
7	-0,20	-0,203	-0,20
8	0,00	0,000	0,00
9	0,20	0,203	0,20
10	0,40	0,424	0,40
11	0,60	0,693	0,60
12	0,80	1,099	0,80
13	0,99	2,647	0,99

$$F(r) = \frac{1}{2}\,ln\left(\frac{1+r}{1-r}\right) \qquad r = \frac{Exp(2F(r))-1}{Exp(2F(r))+1}$$

Figure 13–102 Mise en œuvre des fonctions FISHER et FISHER.INVERSE.

La courbe représentée sur le graphique de la figure 13-102 correspond aux données de la plage *D3:D13*. La cellule *D3* contient la formule `=FISHER(B3)` qui a ensuite été recopiée dans la plage *D4:D13*. L'algorithme correspondant aux calculs effectués par la fonction *FISHER*, `F(r)`, apparaît en bleu, au sommet de la figure 13-102.

Excel fournit également la fonction *FISHER.INVERSE*, qui permet de faire la transformation inverse. La cellule *F3* contient la formule `=FISHER.INVERSE(D3)` qui a ensuite été recopiée dans la plage *F4:F13*. L'algorithme correspondant aux calculs effectués par la fonction *FISHER.INVERSE*, `r`, apparaît en violet, au sommet de la figure 13-102.

En statistiques, la transformation de Fisher est essentiellement utilisée pour réaliser un encadrement du coefficient de corrélation (ρ) de deux variables X et Y distribuées selon une loi normale.

À partir de l'établissement scolaire ayant déjà servi d'exemple au début de ce chapitre, on a tiré un nouvel échantillon aléatoire de `40` élèves pour lequel on a calculé la moyenne annuelle des contrôles continus réalisés en physique et consigné les notes obtenues à l'examen blanc pour cette même matière (figure 13-103). La plage *B5:F12* a été nommée `ExamenBlanc` et la plage *H5:L12* `MoyenneAn`.

Figure 13–103
Physique : moyenne annuelle des contrôles continus et notes obtenues à l'issue de l'examen blanc.

	B	C	D	E	F	H	I	J	K	L
2	Notes de physique									
4	Examen blanc					Moyenne annuelle				
5	8	9	9	10	10	8,5	9,5	9,5	9,5	10,5
6	10	10	10	10	10	10,5	9,5	10,0	10,5	10,5
7	10	10	11	11	11	9,5	8,9	11,6	11,5	11,5
8	11	11	11	11	11	10,7	10,7	10,5	10,2	10,5
9	11	11	11	11	11	11,0	11,3	11,2	10,5	10,0
10	11	11	11	12	12	11,5	11,0	10,2	13,0	11,5
11	12	12	12	12	12	11,5	12,5	11,4	13,0	12,2
12	12	12	13	13	14	11,5	11,5	13,5	12,5	13,5

On dispose donc de deux variables X et Y distribuées normalement et dont on calcule le coefficient de corrélation. On veut déduire de ce premier résultat le coefficient de corrélation entre ces mêmes variables, mais pour l'ensemble des élèves de l'établissement. Pour cela, on prend le coefficient calculé à partir de l'échantillon, mais associé à un encadrement. Ainsi, on peut dire que le coefficient de corrélation pour l'ensemble de l'établissement se situe dans une plage de valeurs comprises dans l'intervalle $[\rho_-, \rho_+]$ (voir la figure 13-105). À condition que les deux variables quantitatives X et Y soient distribuées selon une loi normale et que l'échantillon aléatoire ait une taille suffisante (>30), la transformation de Fisher permet de définir cet encadrement.

COMPRENDRE Complexité de l'encadrement de ρ

Réaliser un encadrement du coefficient de corrélation (ρ) n'est pas une tâche facile car la distribution de ρ est complexe dès qu'il s'éloigne de 0. En passant par la fonction *FISHER*, on transforme la distribution de ρ en une nouvelle variable, F(ρ), qui suit approximativement une loi normale (μ,σ) dès que l'échantillon est suffisamment grand (>30).

Figure 13–104 Espérance et variance de la loi F(r), n correspondant à la taille de l'échantillon.

$$\mu = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right) + \frac{r}{2(n-1)}$$

$$\sigma^2 = \frac{1}{n-3}$$

En quelque sorte, cette loi sert de « sas ». Grâce à elle, on peut construire un intervalle de confiance autour de F(ρ) et en déduire, par transformation inverse, un intervalle de confiance autour de ρ.

À partir des valeurs de notre exemple, la figure 13-105 (en colonne G, syntaxe des formules entrées en colonne *E)* présente une application pratique de la théorie exposée dans la section précédente. Cette application se déroule en cinq phases :

1. Détermination du niveau de confiance de l'encadrement à définir.
2. Calcul du coefficient de corrélation à partir des données de l'échantillon.
3. Application de la transformation de Fisher à ce coefficient.
4. Encadrement de la valeur obtenue à l'étape 3.
5. À partir de l'encadrement calculé à l'étape 4, définition de l'encadrement du coefficient de corrélation.

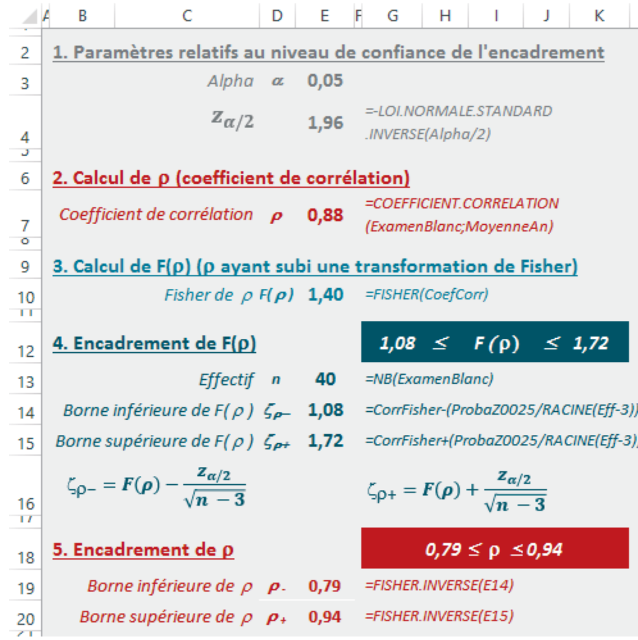


Figure 13–105
Un calcul en cinq étapes pour obtenir l'encadrement du coefficient de corrélation.

Dans la cellule *E3*, nommée `Alpha`, on a entré le risque acceptable pour la définition de notre encadrement `(5 %)`. Cette valeur signifie que la probabilité pour que le coefficient de corrélation des variables `X` et `Y` sur l'ensemble des élèves ne se trouve pas dans l'intervalle `[0,79, 0,94]` est de `0,05`. Dans la cellule *E4*, nommée `ProbaZ0025`, on a utilisé la fonction *LOI.NORMALE.STANDARD.INVERSE* pour calculer la valeur de `z` correspondant à `P(Z < |z|) = 0,95`, probabilité associée à la loi normale centrée réduite. La formule entrée en cellule *E4* utilise `Alpha/2` pour prendre en compte les deux queues de courbe. En effet, on a bien `=LOI.NORMALE.STANDARD.N(-1,96;VRAI) = 0,025`, et `=1-LOI.NORMALE.STANDARD.N(1,96;VRAI) = 0,025`, la somme des deux donnant bien les `0,05` fixés en *E3*.

Dans la cellule *E7*, nommée `CoefCorr`, on a calculé le coefficient de corrélation à partir des valeurs de l'échantillon, puis en *E10* nommée `CorrFisher`, on a appliqué la transformation de Fisher à ce coefficient. En *E13*, nommée `Eff`, apparaît l'effectif de l'échantillon (`n`). Cette valeur est utilisée en *E14* et *E15*, cellules dans lesquelles on a calculé les limites inférieure et supérieure de l'encadrement de `F(ρ)`. La syntaxe générale des formules utilisées pour le calcul de cet encadrement est donnée au niveau de la ligne *16*. Enfin, en *E19* et *E20*, on a utilisé la fonction *FISHER.INVERSE* pour déduire, des résultats obtenus à l'étape *4*, l'encadrement de `ρ`.

ALLER PLUS LOIN **Quelques outils complémentaires**

Si vous n'êtes pas satisfait des outils statistiques proposés à travers les 104 fonctions détaillées dans ce chapitre, vous pouvez toujours explorer l'utilitaire d'analyse, qui n'est pas affiché par défaut dans le ruban. Pour y accéder, il faut d'abord l'installer :

1. Sélectionnez *Fichier>Options>Compléments*.
2. Au bas de la boîte de dialogue, vérifiez que c'est bien *Compléments Excel* qui apparaît dans la liste déroulante et cliquez sur *Atteindre*.
3. Dans la boîte de dialogue, cochez les cases *Analysis ToolPak* et *Analysis ToolPak – VBA*, puis cliquez sur *OK*.

Le bouton d'accès à l'utilitaire d'analyse apparaît maintenant dans le ruban, au niveau de l'onglet *Données (Analyse>Utilitaire d'analyse)*. Pour l'utiliser, il suffit de cliquer sur ce bouton, puis de choisir, dans la liste des 19 outils proposés sur l'écran d'accueil, celui que vous souhaitez mettre en œuvre. Vous remarquerez que de nombreuses fonctionnalités offertes dans cet utilitaire trouvent leur équivalent parmi les 104 fonctions présentées dans ce chapitre (test d'égalité des variances, etc.).

Du côté des ingénieurs 14

D'un intérêt réel pour quelques-uns, les fonctions d'ingénierie présenteront certainement un caractère plus anecdotique pour la majorité. En dehors de toute considération professionnelle, elles permettront à certains de se replonger avec délice (ou avec horreur) dans les souvenirs de lycée et offriront peut-être le petit rafraîchissement indispensable pour mieux aider leur progéniture.

SOMMAIRE

- Nombre complexe
- Nombre binaire, octal, hexadécimal
- Fonctions de Bessel
- Fonctions d'erreur
- Fonction de conversion

MOTS-CLÉS

- Bessel
- Binaire
- Complexe
- Conjugué
- Conversion
- Décimal
- ERF
- Hexadécimal
- Imaginaire
- Kronecker
- Module
- Octal
- Réel

Les fonctions d'ingénierie proposées dans Excel 2010 couvrent des domaines très différents, mais, avec 54 spécimens, Excel ne prétend pas fournir la panoplie complète du parfait ingénieur. Il s'agit plutôt de proposer un petit échantillon des fonctions les plus couramment utilisées et qui, sans l'aide d'un tableur, nécessiteraient des calculs fastidieux. 26 d'entre elles concernent les calculs sur les nombres complexes, 8 correspondent à des fonctions spéciales (fonctions de Bessel et fonctions ERF). Les autres facilitent les conversions délicates.

Nombres complexes

Depuis le XVII^e siècle, la question de la racine carrée des nombres négatifs tourmentait les mathématiciens. Habitués à raisonner sur des nombres dont le carré était toujours positif, leurs convictions les plus profondes étaient ébranlées par des nombres capables de donner un résultat négatif lorsqu'ils étaient multipliés par eux-mêmes. Cependant, il était bien utile d'accepter d'écrire `i² = -1` pour résoudre des équations algébriques telles que `x² + 1 = 0`.

CULTURE Une représentation astucieuse

Au début du XVIII^e siècle, deux mathématiciens amateurs furent les premiers à élaborer une représentation « acceptable » de ce mystérieux nombre `i`.

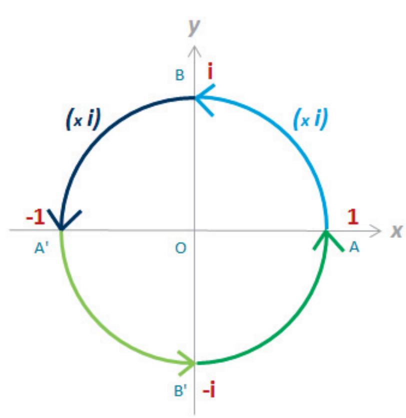


Figure 14–1 Représentation du nombre i comme opérateur de rotation.

La longueur `+1` est représentée par le segment `OA` et la longueur `-1` par le segment `OA'`. On envisage le passage de l'un à l'autre par une rotation d'un demi-tour dans le sens inverse des aiguilles d'une montre. On décide, arbitrairement, de noter cette transformation `= (+1) × (-1) = (-1)`, ce qui équivaut à traduire algébriquement l'opération de rotation d'un demi-tour par `(-1)`. Deux demi-tours successifs correspondent à un tour complet, ce qui peut se noter `= (-1) × (-1) = (+1)`.
Si l'on appelle `(+i)` l'opérateur correspondant à un quart de tour, deux quarts de tour successifs étant équivalents à un demi-tour, on a `(+i) × (+i) = (-1)`, que l'on peut écrire `i² = (-1)`. Le nombre `i` représente donc une rotation d'un quart de tour, c'est-à-dire le point `B` sur l'axe `Oy`.

Notation d'un nombre complexe

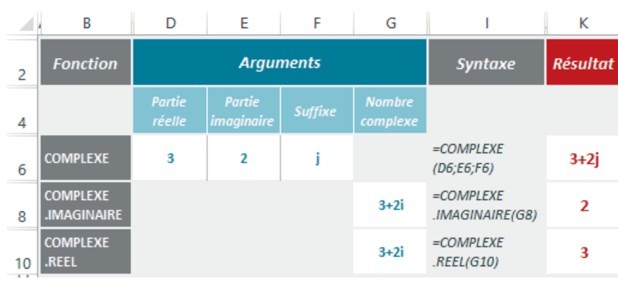
Un nombre complexe `(a, b)` est noté, par convention, `z = a + bi` (`z` est un vecteur et ne représente pas un nombre réel). `a` s'appelle la partie réelle du nombre complexe, `bi` la partie imaginaire. Un nombre complexe dans lequel `a = 0` se réduit à `z = bi` et est appelé nombre imaginaire pur. Un nombre complexe dans lequel `b = 0` se réduit à sa partie réelle (c'est un nombre réel). Les règles de calcul appliquées dans le corps des complexes sont celles du calcul algébrique ordinaire, avec la convention `i² = -1`.

PARAMÈTRE Expression de la partie imaginaire d'un nombre complexe

Parfois, on utilise `j` à la place de `i`.

Excel propose trois fonctions relatives à la notation d'un nombre complexe.

Figure 14–2
Mise en œuvre des fonctions COMPLEXE, COMPLEXE.IMAGINAIRE et COMPLEXE.REEL.



Fonction	Arguments				Syntaxe	Résultat
	Partie réelle	Partie imaginaire	Suffixe	Nombre complexe		
COMPLEXE	3	2	j		=COMPLEXE (D6;E6;F6)	3+2j
COMPLEXE .IMAGINAIRE				3+2i	=COMPLEXE .IMAGINAIRE(G8)	2
COMPLEXE .REEL				3+2i	=COMPLEXE .REEL(G10)	3

Tableau 14–1 Fonctions relatives à la notation d'un nombre complexe

Fonction	Description
COMPLEXE	Cette fonction renvoie un nombre complexe. Elle utilise trois arguments. Le premier représente la partie réelle du nombre complexe, le second sa partie imaginaire et le troisième indique s'il faut utiliser `i` ou `j` pour caractériser la partie imaginaire ; si l'on ne précise rien, c'est `i` qui est utilisé par défaut.
COMPLEXE .IMAGINAIRE	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la partie imaginaire.
COMPLEXE.REEL	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la partie réelle.

COMPRENDRE Représentation géométrique d'un nombre complexe

La méthode de Gauss consiste à considérer deux axes perpendiculaires, `Ox` et `Oy`, auxquels on rapporte les différents points du plan. Un point `A` de coordonnées réelles `a` et `b` est l'image d'un être mathématique appelé nombre complexe : `z = a + bi`, ou affixe de `A`. Tout point de l'axe des `x` correspond à un nombre réel `z = a` ; tout point de l'axe `Oy` à un nombre imaginaire pur `z = bi`.

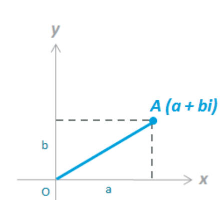


Figure 14–3 Représentation géométrique d'un nombre complexe.

Opérations simples sur les nombres complexes

Les nombres complexes peuvent être additionnés, soustraits, multipliés. Il faut juste respecter quelques conventions.

RAPPEL Corps des nombres complexes

Les règles de calcul sont celles du calcul algébrique ordinaire. Pour les opérations de base, on a les relations illustrées figure 14-4. Les calculs sont facilités par :

- la relation particulière qu'entretient un nombre complexe avec son conjugué ;
- la représentation trigonométrique des nombres complexes.

Égalité	$z = z'$	$a = a'$; $b = b'$
Addition	$z + z'$	$(a + a') + (b + b')i = A + Bi$
Multiplication	zz'	$(aa' - bb') + (ab' + ba')i = A + Bi$
Division	$z : z' = zz'^{-1}$	$\frac{aa' + bb'}{a'^2 + b'^2} + \frac{ba' - ab'}{a'^2 + b'^2} i$
Module de z (valeur absolue)	$\|z\|$	$\sqrt{a^2 + b^2}$

Figure 14–4 Règles de calcul avec les nombres complexes.

Complexes conjugués

Si `z = a + bi` et `z' = a - bi`, `z` et `z'` sont dits conjugués. Dans ce cas, on a `z + z' = 2a` et `zz' = a² + b²`. Cette relation particulière permet de simplifier bon nombre de calculs sur les complexes, en particulier la division d'un nombre complexe par un autre.

Tableau 14–2 Module et conjugué d'un nombre complexe

Fonction	Description
COMPLEXE .CONJUGUE	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le conjugué.
COMPLEXE .MODULE	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le module.

RAPPEL Module d'un nombre complexe

Le produit d'un nombre complexe et de son conjugué est le nombre réel ou nul `a² + b²`. On appelle module de `z`, noté `|z|`, la racine carrée de ce produit.

Figure 14–5
Les deux fonctions COMPLEXE.CONJUGUE et COMPLEXE.MODULE utilisent la notion de conjugué d'un nombre complexe.

	B	D	F	H
2	*Fonction*	*Arguments*	*Syntaxe*	*Résultat*
4		*Nombre complexe*		
6	COMPLEXE .CONJUGUE	3+2i	*=COMPLEXE .CONJUGUE(D6)*	3-2i
8	COMPLEXE .MODULE	3+2i	*=COMPLEXE .MODULE(D8)*	3,605551275

Représentation trigonométrique des nombres complexes

On fait correspondre à tout point `M` de coordonnées `a` et `b`, un nombre complexe `z = a + bi`.

COMPRENDRE Représentation trigonométrique des nombres complexes

On appelle `ρ` la grandeur géométrique du vecteur `OM` et `θ` l'angle de l'axe `Ox` avec le vecteur `OM`.

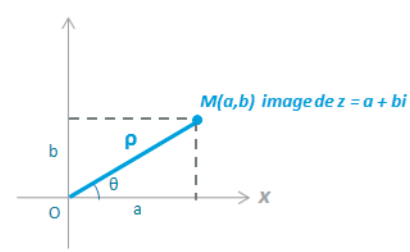


Figure 14–6 Représentation trigonométrique d'un nombre complexe.

Grâce aux relations élémentaires de trigonométrie et aux propriétés du triangle rectangle, on peut écrire un certain nombre de relations entre les deux composantes du nombre complexe et le sinus et le cosinus de l'angle `θ`.

$$a = \rho \cos \theta \qquad \cos \theta = \frac{a}{\sqrt{a^2 + b^2}}$$

$$b = \rho \sin \theta$$

$$\rho = \sqrt{a^2 + b^2} \qquad \sin \theta = \frac{b}{\sqrt{a^2 + b^2}}$$

Figure 14–7 Relations entre les composantes d'un nombre complexe et les caractéristiques de l'angle qui lui correspond.

`ρ` est le module de `z` et `θ` en est l'argument (voir aparté). Ainsi, le nombre `z = a + bi` peut s'écrire sous la forme trigonométrique : `z = ρcosθ + iρsinθ= ρ(cosθ + isinθ)`.

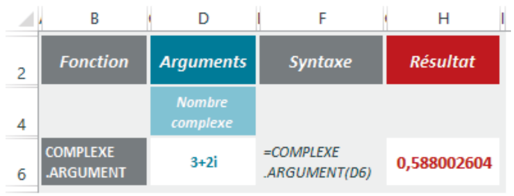


Figure 14–8
L'argument d'un complexe (arg z) est l'angle qui correspond à sa représentation trigonométrique.

Tableau 14–3 Argument d'un nombre complexe

Fonction	Description
COMPLEXE .ARGUMENT	Cette fonction utilise un seul argument, un nombre complexe, et renvoie, en radians, la mesure de l'angle correspondant à sa représentation trigonométrique.

Calculs de base avec les nombres complexes

En appliquant les règles de calcul énoncées dans la figure 14-4, on peut additionner, multiplier et même diviser des nombres complexes. Les fonctions *COMPLEXE.SOMME* et *COMPLEXE.PRODUIT* acceptent un nombre variable d'arguments. Chaque argument pouvant être une plage, elles permettent de travailler avec un très grand nombre de complexes. En revanche, les fonctions *COMPLEXE.DIFFERENCE* et *COMPLEXE.DIV* ne peuvent faire un calcul que sur deux complexes à la fois.

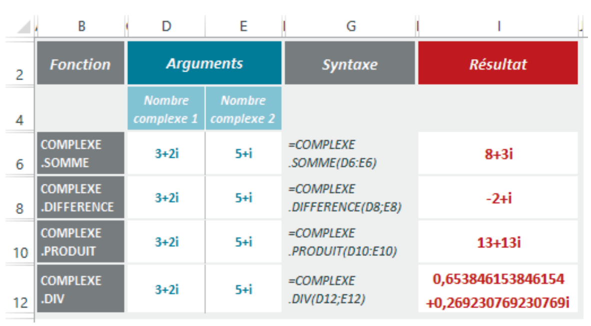


Figure 14–9
Les quatre opérations de base appliquées aux nombres complexes.

Tableau 14–4 Calculs de base sur les complexes

Fonction	Description
COMPLEXE.SOMME	Cette fonction renvoie la somme des nombres complexes contenus dans ses arguments.
COMPLEXE. DIFFERENCE	Cette fonction utilise deux arguments, deux nombres complexes. Elle retranche le deuxième du premier et renvoie le résultat de cette soustraction.
COMPLEXE.PRODUIT	Cette fonction renvoie le produit des nombres complexes contenus dans ses arguments.
COMPLEXE.DIV	Cette fonction utilise deux arguments, deux nombres complexes. Elle divise le premier par le deuxième et renvoie le quotient qui résulte de cette opération.

Puissance et racine d'un nombre complexe

Toujours en appliquant les règles de calcul énoncées plus haut, on peut calculer la puissance énième ou la racine carrée d'un nombre complexe. Les formules qui sous-tendent ce calcul sont indiquées dans la figure 14-10.

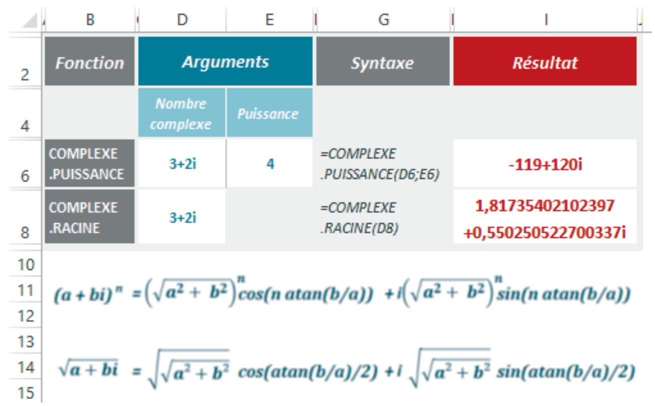


Figure 14–10
Élever un nombre complexe à la puissance n et calculer sa racine carrée.

Tableau 14–5 Puissance et racine d'un nombre complexe

Fonction	Description
COMPLEXE .PUISSANCE	Cette fonction utilise deux arguments. Le deuxième, un nombre décimal, sert à élever le premier, un nombre complexe, à une certaine puissance.
COMPLEXE.RACINE	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la racine carrée.

Fonctions circulaires appliquées aux nombres complexes

Le corps des complexes permet également de définir un sinus, un cosinus, une tangente, une cotangente, une sécante, une cosécante, un sinus hyperbolique, un cosinus hyperbolique, une sécante hyperbolique et une cosécante hyperbolique. Les formules de calcul sont fournies dans les figures 14-11, 14-12, 14-13, 14-14 et 14-15.

Sinus et cosinus d'un nombre complexe

Tableau 14–6 Sinus et cosinus d'un nombre complexe

Fonction	Description
COMPLEXE.COS	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le cosinus.
COMPLEXE.SIN	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le sinus.

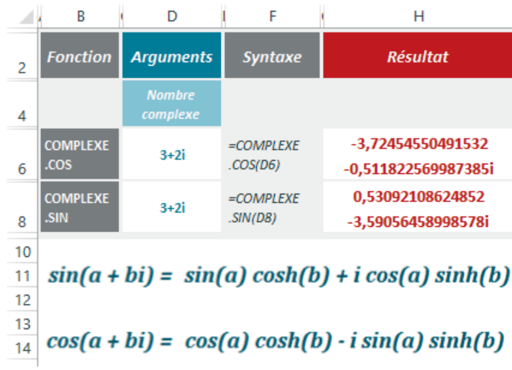


Figure 14–11
Sinus et cosinus d'un nombre complexe.

Tangente et cotangente d'un nombre complexe

Tableau 14–7 Tangente et cotangente d'un nombre complexe

Fonction	Description
COMPLEXE.TAN	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la tangente. **Nouveauté Excel 2013**.
COMPLEXE.COT	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la cotangente. **Nouveauté Excel 2013**.

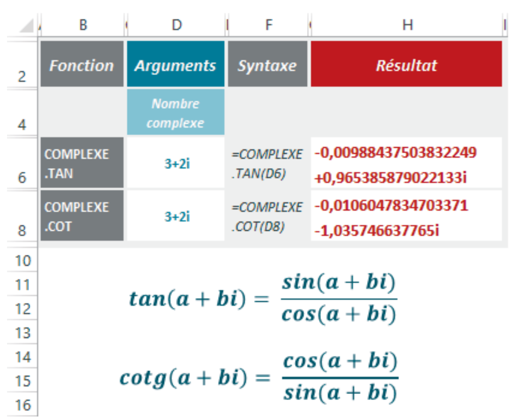


Figure 14–12
Tangente et cotangente d'un nombre complexe.

Sécante et cosécante d'un nombre complexe

Tableau 14–8 Sécante et cosécante d'un nombre complexe

Fonction	Description
COMPLEXE.SEC	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la sécante. **Nouveauté Excel 2013**.

Tableau 14–8 Sécante et cosécante d'un nombre complexe (suite)

Fonction	Description
COMPLEXE.CSC	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la cosécante. **Nouveauté Excel 2013**.

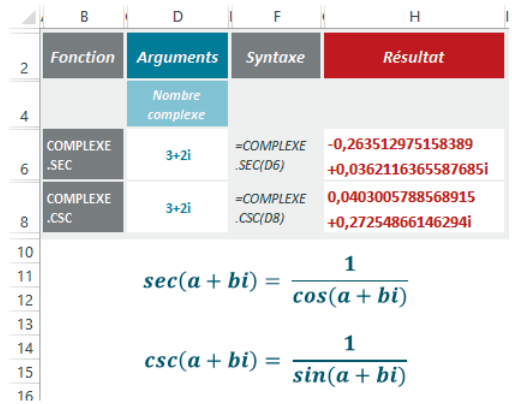


Figure 14–13
Sécante et cosécante d'un nombre complexe.

Sinus et cosinus hyperboliques d'un nombre complexe

Tableau 14–9 Sinus et cosinus hyperboliques d'un nombre complexe

Fonction	Description
COMPLEXE.SINH	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le sinus hyperbolique. **Nouveauté Excel 2013**.
COMPLEXE.COSH	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le cosinus hyperbolique. **Nouveauté Excel 2013**.

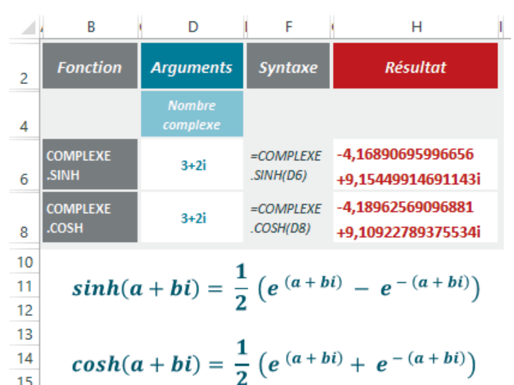


Figure 14–14
Sinus et cosinus hyperboliques d'un nombre complexe.

Sécante et cosécante hyperboliques d'un nombre complexe

Tableau 14–10 Sécante et cosécante hyperboliques d'un nombre complexe

Fonction	Description
COMPLEXE.SECH	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la sécante hyperbolique. **Nouveauté Excel 2013**.
COMPLEXE.CSCH	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie la cosécante hyperbolique. **Nouveauté Excel 2013**.

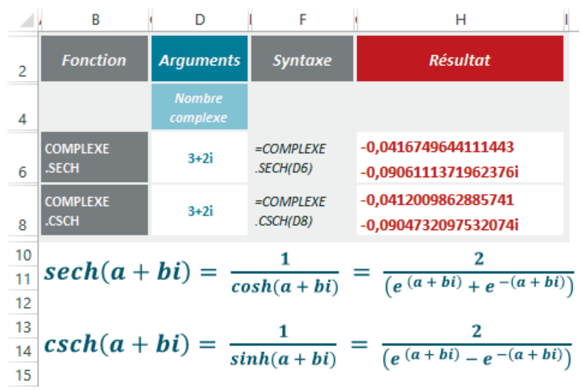


Figure 14–15
Sécante et cosécante hyperboliques d'un nombre complexe.

Exponentielle et logarithme d'un nombre complexe

On peut calculer l'exponentielle d'un nombre complexe, ainsi que son logarithme. Les formules de calcul sont fournies à la figure 14-16.

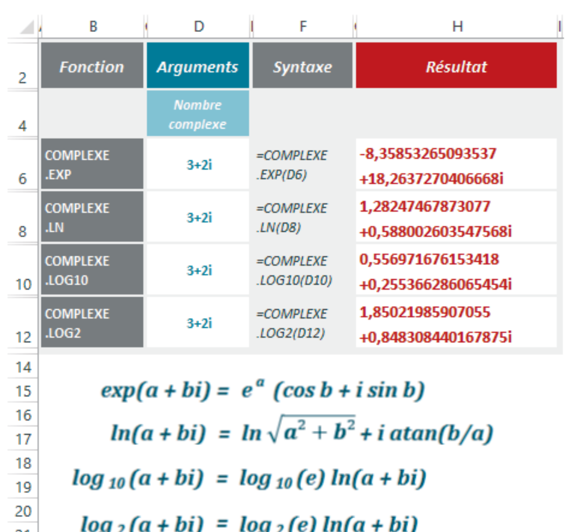


Figure 14–16
Exponentielle et logarithme d'un nombre complexe.

Tableau 14–11 Exponentielle et logarithme d'un nombre complexe

Fonction	Description
COMPLEXE.EXP	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie l'exponentielle.
COMPLEXE.LN	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le logarithme népérien.
COMPLEXE.LOG10	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le logarithme de base 10.
COMPLEXE.LOG2	Cette fonction utilise un seul argument, un nombre complexe, dont elle renvoie le logarithme de base 2.

Nombre binaire, octal, décimal, hexadécimal

On nomme couramment bit (de l'anglais *binary digit*) les chiffres de la numération binaire. Ceux-ci ne peuvent prendre que deux valeurs, notées par convention 0 et 1. Dans le système binaire, 2 s'écrit 10, 3 s'écrit 11, 4 s'écrit 100, et ainsi de suite. Les nombres ainsi construits sont très simples… mais inévitablement très longs ! Les microprocesseurs des ordinateurs ne comprennent que le langage binaire (soit le courant électrique passe, soit il ne passe pas), mais l'esprit humain appréhende plus facilement des nombres plus courts. C'est pourquoi les bases octale et hexadécimale, toutes deux multiples de la base deux, sont couramment employées en informatique pour ces raisons pratiques.

DÉTAIL Afficher les zéros non significatifs

Toutes les fonctions prennent au moins un argument (le nombre à convertir), mais celles qui convertissent vers le binaire, l'octal ou l'hexadécimal en acceptent un deuxième. Ce dernier précise la taille finale du nombre souhaité, ce qui déclenche l'affichage, devant le nombre, des zéros non significatifs nécessaires pour parvenir au résultat désiré.

	B	C	D	E
2	X	Nb. car.	Syntaxe	Résultat
4	10	6	=BINOCT(B4;C4)	000002
5	100 000	10	=BINOCT(B5;C5)	0000000040
6	10	6	=BINHEX(B6;C6)	000002
7	111 111	10	=BINHEX(B7;C7)	000000003F
8	10	6	=DECOCT(B8;C8)	000012
9	100 000	10	=DECOCT(B9;C9)	0000303240
10	10	6	=DECHEX(B10;C10)	00000A
11	111 111	10	=DECHEX(B11;C11)	000001B207
12	0012	6	=OCTHEX(B12;C12)	00000A
13	0000303240	10	=OCTHEX(B13;C13)	00000186A0
14	000A	6	=HEXOCT(B14;C14)	000012
15	000001B207	10	=HEXOCT(B15;C15)	0000331007
16	12	6	=OCTBIN(B16;C16)	001010
17	777	10	=OCTBIN(B17;C17)	0111111111
18	A	6	=HEXBIN(B18;C18)	001010
19	1FF	10	=HEXBIN(B19;C19)	0111111111
20	10	6	=DECBIN(B20;C20)	001010
21	511	10	=DECBIN(B21;C21)	0111111111

Figure 14–17
Fonctions de conversion vers le binaire, l'octal ou l'hexadécimal avec un deuxième argument non nul.

Excel fournit une douzaine de fonctions qui convertissent automatiquement les nombres d'une base à l'autre. N'oubliez pas de consulter le chapitre 12 dans lequel sont exposées les deux fonctions *BASE* et *DECIMAL*, qui facilitent également les conversions entre la base décimale et les autres.

Système binaire et système décimal

Excel fournit trois fonctions de conversion à partir du système binaire et trois autres à partir du système décimal.

TECHNIQUE Les nombres négatifs

Les nombres binaires sont exprimés sur 10 bits. Excel utilise le bit de poids fort pour indiquer le signe (il le fait passer à 1 lorsque le nombre est négatif). Ensuite, il exprime le nombre négatif sur les neuf autres bits, de manière à ce que la somme du positif et du négatif donne bien 1 000 000 000.

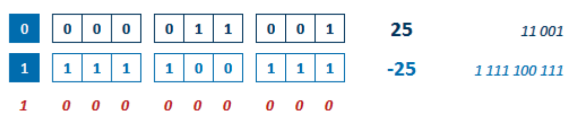


Figure 14–18 Illustration de l'expression d'un nombre négatif en système binaire. Lorsqu'on additionne 25 (000 011 001) et -25 ((1) 111 100 111), on obtient bien 1 000 000 000.

Le raisonnement dans les bases 8 et 16 est exactement le même, sauf que le bit de poids fort se trouve sur la trentième ou la quarantième position.

Tableau 14–12 Convertir depuis le système binaire ou décimal

Fonction	Description
BINDEC	Convertit un nombre binaire en nombre décimal.
BINOCT	Convertit un nombre binaire en nombre octal (30 bits).
BINHEX	Convertit un nombre binaire en nombre hexadécimal (40 bits).
DECBIN	Convertit un nombre décimal en nombre binaire (10 bits).
DECOCT	Convertit un nombre décimal en nombre octal (30 bits).
DECHEX	Convertit un nombre décimal en nombre hexadécimal (40 bits).

Figure 14–19
Conversion des nombres binaires et des nombres décimaux.

X	=BINOCT(x)	=BINHEX(x)	=BINDEC(x)
0	0	0	0
1	1	1	1
10	2	2	2
11	3	3	3
100	4	4	4
101	5	5	5
1 000	10	8	8
10 000	20	10	16
100 000	40	20	32
1 000 000	100	40	64
10 000 000	200	80	128
100 000 000	400	100	256
111 111 111	777	1FF	511
1 000 000 000	7777777000	FFFFFFFE00	-512
1 000 000 001	7777777001	FFFFFFFE01	-511
1 000 000 010	7777777002	FFFFFFFE02	-510
1 111 100 000	7777777740	FFFFFFFFE0	-32
1 111 111 110	7777777776	FFFFFFFFFE	-2

X	=DECOCT(x)	=DECHEX(x)	=DECBIN(x)
0	0	0	0
1	1	1	1
2	2	2	10
3	3	3	11
10	12	A	1010
50	62	32	110010
100	144	64	1100100
500	764	1F4	111110100
511	777	1FF	111111111
-512	7777777000	FFFFFFFE00	1000000000
-511	7777777001	FFFFFFFE01	1000000001
-500	7777777014	FFFFFFFE0C	1000001100
-100	7777777634	FFFFFFFF9C	1110011100
-50	7777777716	FFFFFFFFCE	1111001110
-10	7777777766	FFFFFFFFF6	1111110110
-3	7777777775	FFFFFFFFFD	1111111101
-2	7777777776	FFFFFFFFFE	1111111110
-1	7777777777	FFFFFFFFFF	1111111111

CLIN D'ŒIL Le système décimal

Le système décimal est un système de numération utilisant la base dix, très ancienne. Elle découle d'un choix naturel, dicté par le nombre des doigts des deux mains.

Système octal et système hexadécimal

L'hexadécimal est un système de numération « positionnel » en base 16. Il utilise 16 symboles, en général les chiffres arabes pour les dix premiers chiffres et les lettres A à F pour les six suivants. Le système octal est quelquefois utilisé en calcul à la place de l'hexadécimal. Il possède le double avantage de ne pas requérir de symbole supplémentaire pour ses chiffres et d'être une puissance de deux pour pouvoir regrouper les chiffres du nombre binaire. Excel fournit trois fonctions de conversion à partir du système octal et trois autres à partir du système hexadécimal.

ASTUCE Passer de l'une à l'autre

Pour trouver facilement l'expression d'un nombre octal ou hexadécimal, il suffit de regrouper les chiffres du nombre exprimé en base 2 : pour la base octale (2^3), on fait des paquets de trois à partir de la droite et pour la base hexadécimale (2^4), on fait des paquets de quatre.

Tableau 14–13 Convertir depuis le système octal ou hexadécimal

Fonction	Description
OCTDEC	Convertit un nombre octal en nombre décimal.
OCTBIN	Convertit un nombre octal en nombre binaire (10 bits).

Tableau 14–13 Convertir depuis le système octal ou hexadécimal (suite)

Fonction	Description
OCTHEX	Convertit un nombre octal en nombre hexadécimal (40 bits).
HEXBIN	Convertit un nombre hexadécimal en nombre binaire (10 bits).
HEXOCT	Convertit un nombre hexadécimal en nombre octal (30 bits).
HEXDEC	Convertit un nombre hexadécimal en nombre décimal.

Figure 14–20
Conversion d'un nombre octal et d'un nombre hexadécimal.

X	=OCTDEC(x)	=OCTHEX(x)	=OCTBIN(x)	X	=HEXDEC(x)	=HEXOCT(x)	=HEXBIN(x)
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	10	2	2	2	10
3	3	3	11	3	3	3	11
4	4	4	100	4	4	4	100
5	5	5	101	5	5	5	101
10	8	8	1000	8	8	10	1000
20	16	10	10000	10	16	20	10000
40	32	20	100000	20	32	40	100000
100	64	40	1000000	40	64	100	1000000
200	128	80	10000000	80	128	200	10000000
400	256	100	100000000	100	256	400	100000000
777	511	1FF	111111111	1FF	511	777	111111111
7777777000	-512	FFFFFFFE00	1000000000	FFFFFFFE00	-512	7777777000	1000000000
7777777001	-511	FFFFFFFE01	1000000001	FFFFFFFE01	-511	7777777001	1000000001
7777777002	-510	FFFFFFFE02	1000000010	FFFFFFFE02	-510	7777777002	1000000010
7777777740	-32	FFFFFFFFE0	1111100000	FFFFFFFFE0	-32	7777777740	1111100000
7777777776	-2	FFFFFFFFFE	1111111110	FFFFFFFFFE	-2	7777777776	1111111110

CLIN D'ŒIL Le système octal

Le décompte octal pourrait avoir été utilisé dans le passé à la place du décompte décimal, en comptant soit les trous entre les doigts, soit les doigts sans le pouce.

Opérations binaires

Si vous travaillez avec Excel 2013, vous disposez de nouvelles fonctions qui convertissent deux entiers positifs au format décimal en nombres binaires, pour réaliser des calculs de type `ET` ou `OU` et renvoyer le résultat au format décimal.

Opérations binaires de type ET, OU et OU EXCLUSIF

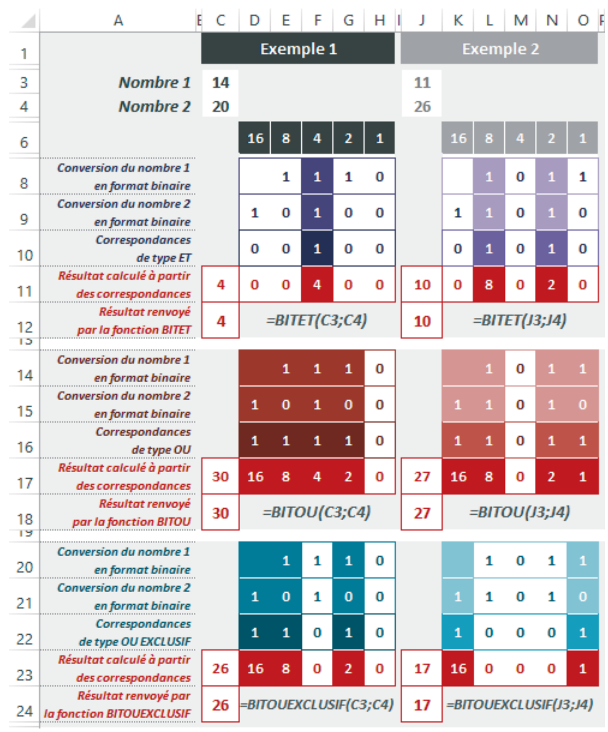
Tableau 14–14 Opérations binaires ET, OU et OU EXCLUSIF

Fonction	Description
BITET	Cette fonction utilise deux arguments, deux nombres entiers positifs au format décimal et strictement inférieurs à 281 474 976 710 656 (soit 2^{48}), dont elle renvoie une opération binaire ET. **Nouveauté Excel 2013**.

Tableau 14–14 Opérations binaires ET, OU et OU EXCLUSIF (suite)

Fonction	Description
BITOU	Cette fonction utilise deux arguments, deux nombres entiers positifs au format décimal et strictement inférieurs à 281 474 976 710 656 (soit 2^{48}), dont elle renvoie une opération binaire OU. **Nouveauté Excel 2013**.
BITOUEXCLUSIF	Cette fonction utilise deux arguments, deux nombres entiers positifs au format décimal et strictement inférieurs à 281 474 976 710 656 (soit 2^{48}), dont elle renvoie une opération binaire OU EXCLUSIF. **Nouveauté Excel 2013**.

Figure 14–21
Mise en œuvre des fonctions BITET, BITOU et BITOUEXCLUSIF.



La figure 14-21 illustre le fonctionnement et l'utilisation des fonctions *BITET*, *BITOU* et *BITOUEXCLUSIF*. À travers deux exemples, les divers schémas montrent comment ces fonctions établissent les trois types de correspondances logiques pour renvoyer les résultats au format décimal.

Opérations binaires destinées à décaler les bits

Tableau 14–15 Décaler les bits des représentations binaires des nombres

Fonction	Description
BITDECALD	Cette fonction utilise deux arguments. Le premier, un nombre entier positif au format décimal strictement inférieur à 281 474 976 710 656 (soit 2^{48}) désigne le nombre à transformer. Le second, un nombre entier dont la valeur absolue doit être inférieure ou égale à 53, désigne le nombre de bits de décalage à appliquer à la représentation binaire du premier argument. Si le second argument est positif, le décalage s'opère vers la droite (le nombre entré en premier argument devient plus petit). Si le second argument est négatif, le décalage s'opère vers la gauche et la fonction adopte la logique de la fonction *BITDECALG*. **Nouveauté Excel 2013**.
BITDECALG	Cette fonction utilise deux arguments. Le premier, un nombre entier positif au format décimal strictement inférieur à 281 474 976 710 656 (soit 2^{48}) désigne le nombre à transformer. Le second, un nombre entier dont la valeur absolue doit être inférieure ou égale à 53, désigne le nombre de bits de décalage à appliquer à la repésentation binaire du premier argument. Si le second argument est positif, le décalage s'opère vers la gauche et la fonction ajoute autant de zéros à droite de la représentation binaire du nombre que le décalage en impose (le nombre entré en premier argument devient plus grand). Si le second argument est négatif, le décalage s'opère vers la droite et la fonction adopte la logique de la fonction *BITDECALD*. **Nouveauté Excel 2013**.

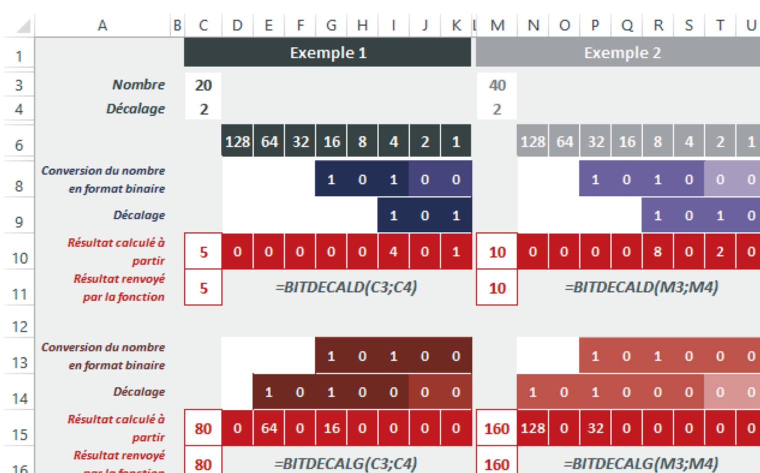


Figure 14–22
Mise en œuvre des fonctions BITDECALD et BITDECALG.

La figure 14-22 illustre le fonctionnement et l'utilisation des fonctions *BITDECALD* et *BITDECALG*. À travers deux exemples, les divers schémas montrent comment ces fonctions agissent sur les nombres entrés en premier argument, pour opérer des décalages sur les bits de droite ou de gauche de leur représentation binaire.

Fonctions de Bessel

Les fonctions de Bessel sont utilisées pour étudier de nombreux phénomènes physiques comme la propagation d'une onde électromagnétique dans un conducteur filaire, la vibration d'une membrane circulaire, la propagation de la chaleur et encore bien d'autres manifestations en mécanique quantique et en physique nucléaire.

COMPRENDRE **Histoire et fondements mathématiques**

Tous ces phénomènes physiques peuvent être décrits par une équation différentielle de second ordre de la forme indiquée figure 14-23.

$$x^2 f''(x) + xf'(x) + (x^2 - n^2)f(x) = 0$$

n entier relatif

Figure 14–23
Équation différentielle de second ordre permettant de décrire les phénomènes physiques énumérés au début de cette section.

C'est en étudiant les oscillations du fil pesant que Daniel Bernoulli (1700 - 1782) établit une première forme de cette équation. Cependant, il a fallu attendre Friedrich Wilhelm Bessel (1784 - 1846) pour que ses solutions soient étudiées de façon approfondie. Ces dernières prendront le nom de fonctions de Bessel.
La résolution d'une telle équation différentielle nécessite une bonne dose d'inspiration, l'emploi de plusieurs outils mathématiques de bon niveau et beaucoup d'opiniâtreté. Aussi, nous nous bornerons à rappeler brièvement les résultats (nous laisserons pudiquement de côté les démonstrations).

Deux familles de fonctions de Bessel

Il existe deux familles de fonctions, solutions de l'équation différentielle présentée figure 14-23.

Fonctions de Bessel, dites de première espèce

Notées $J_n(x)$, d'ordre n, elles sont définies en $x = 0$. Leur équation est donnée figure 14-24.

$$J_n(x) = \frac{x^n}{2^n} \sum_{k=0}^{k=+\infty} \frac{(-1)^k}{k!\,(n+k)!} \left(\frac{x}{2}\right)^{2k}$$

Figure 14–24
Équation des fonctions de Bessel dites de première espèce.

Récréation **Pour les curieux**

Le lecteur courageux amateur de casse-tête peut remarquer que l'équation donnée figure 14-25 est une solution particulière de l'équation de Bessel d'ordre 0 présentée figure 14-26.

$$J_0(x) = \frac{1}{\pi} \int_0^{\pi} cos\ (x\ sin\theta)\ d\theta$$

Figure 14–25 Solution de l'équation de Bessel d'ordre 0.

$$x^2 f''(x) + x f'(x) + x^2 f(x) = 0$$

Figure 14–26 Équation de Bessel d'ordre 0.

En utilisant une intégration par parties et en s'appuyant sur le développement en série entière de la fonction cosinus, il pourra montrer que $J_0(x)$ correspond à l'équation présentée figure 14-27.

$$J_0(x) = \sum_{n=0}^{n=+\infty} \frac{(-1)^n}{4^n\ (n!)^2}\ x^{2n}$$

Figure 14–27 Équation de $J_0(x)$.

À cette étape, le lecteur ne sera que sur la première marche de la montagne à gravir. La fonction $J_n(x)$ est une solution générale de l'équation présentée figure 14-23, avec toutefois une particularité : n est un entier relatif.

$$x^2 f''(x) + x f'(x) + (x^2 - \nu^2) f(x) = 0$$

Figure 14–28 Équation équivalente à celle de la figure 14-23, pour ν réel quelconque.

En élargissant l'étude à l'équation de la figure 14-28 où cette fois ν est un réel quelconque, on obtient l'expression générale de la fonction de Bessel de première espèce $J_\nu(x)$ présentée figure 14-29.

$$J_\nu(x) = \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{k=+\infty} \frac{(-1)^k}{k!\ \Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{2k}$$

$$\Gamma(x) = \int_0^{+\infty} t^{x-1}\ e^{-t}\ dt$$

Figure 14–29 Expression générale de la fonction de Bessel de première espèce, où Γ désigne la fonction gamma d'Euler.

Fonctions de Bessel, dites de deuxième espèce

On peut remarquer que les fonctions J_ν et $J_{-\nu}$ ne sont pas proportionnelles (par exemple, leur comportement diffère au voisinage de 0). Le couple (J_ν, $J_{-\nu}$) constitue donc une base de l'espace des solutions de l'équation de Bessel. Les fonctions de Bessel dites de deuxième espèce et notées $Y_\nu(x)$, d'index ν, ne sont pas définies en $x = 0$. Leur équation est présentée figure 14-30.

Figure 14–30
Équation des fonctions de Bessel dites de deuxième espèce.

$$Y_\nu(x) = \frac{\cos(\nu\pi)\ J_\nu(x)\ - J_{-\nu}(x)}{\sin(\nu\pi)}$$

Lorsque ν est entier $(\nu = n)$, on définit la fonction de Bessel de deuxième espèce d'ordre n par la fonction $Y_n(x)$ présentée figure 14-31.

Figure 14–31
Équation des fonctions de Bessel dites de deuxième espèce quand ν est entier.

$$Y_n(x) = \lim_{\nu \to n} Y_\nu(x) = \lim_{\nu \to n} \frac{\cos(\nu\pi)\, J_\nu(x) - J_{-\nu}(x)}{\sin(\nu\pi)}$$

Extension des fonctions de Bessel dans le plan complexe

De même qu'il a été procédé à l'extension de l'équation initiale avec l'entier n vers un réel ν quelconque, il est possible d'imaginer l'extension de l'index ν à un nombre complexe. Cet artifice permet de résoudre l'équation de Bessel dite « modifiée » (figure 14-32).

Figure 14–32
Équation des fonctions de Bessel dites modifiées.

$$x^2 y'' + x\, y' - (x^2 + \nu^2) y = 0$$

En posant $t = ix$ et avec le changement de fonction $y(x) = z(ix)$, l'équation présentée figure 14-32 se réduit à l'équation classique (figure 14-33).

Figure 14–33
Autre forme de l'équation des fonctions de Bessel dites modifiées.

$$t^2 z''(t) + t z'(t) + (t^2 - \nu^2) z(t) = 0$$

Les solutions de cette équation sont les fonctions de Bessel dites modifiées (appelées aussi « à argument imaginaire ») de la forme présentée figure 14-34.

Figure 14–34
Équation des fonctions de Bessel à argument imaginaire.

$$I_\nu(x) = e^{\frac{-i\pi\nu}{2}}\ J_\nu(i\,x)$$

Ces fonctions se décomposent en série entière sous la forme présentée figure 14-35.

Figure 14–35
Décomposition en série entière des fonctions de Bessel à argument imaginaire.

$$I_\nu(x) = \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{k=+\infty} \frac{1}{k!\ \Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{2k}$$

Fonctions de Bessel proposées par Excel

Excel propose une implémentation des quatre grandes familles des fonctions de Bessel que nous venons d'aborder :

- fonctions de Bessel dites de première espèce : *BESSELJ* ;
- fonctions de Bessel dites de deuxième espèce : *BESSELY* ;
- fonctions de Bessel dites modifiées : *BESSELI* et *BESSELK*.

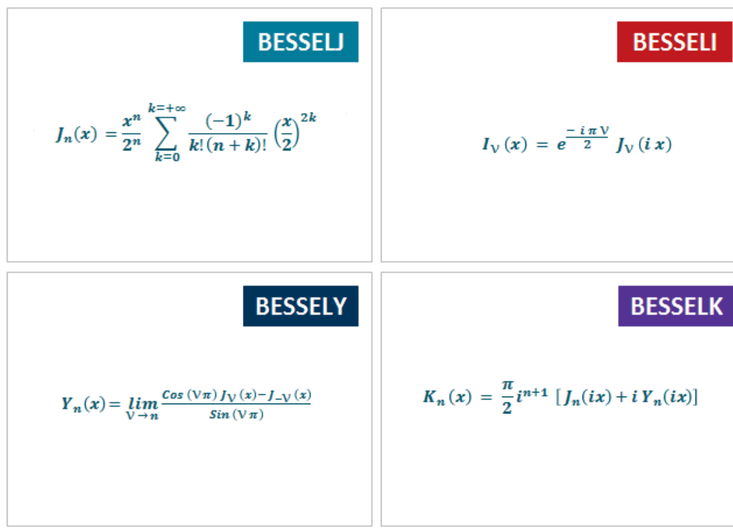


Figure 14–36
Équation des quatre fonctions de Bessel proposées par Excel.

Tableau 14–16 Fonctions de Bessel

Fonction	Description
BESSELJ	Renvoie la fonction de Bessel $J_n(x)$ (voir la figure 14-36).
BESSELY	Renvoie la fonction de Bessel $Y_n(x)$, également appelée fonction de Weber ou fonction de Neumann (voir la figure 14-36).
BESSELI	Renvoie la fonction de Bessel modifiée $I_\nu(x)$ qui correspond à une extension de la fonction de Bessel pour des arguments imaginaires (voir la figure 14-36).
BESSELK	Renvoie la fonction de Bessel modifiée $K_n(x)$ qui correspond à une extension des fonctions de Bessel J_n et Y_n pour des arguments imaginaires (voir la figure 14-36).

Figure 14–37
Quelques valeurs de x pour n = 0, 1 ou 2.

	B	C	D	E	G	H	I	J	L	M	N	O	Q	R	S	T
2		=BESSELJ(x;n)				=BESSELY(x;n)				=BESSELI(x;n)				=BESSELK(x;n)		
4	X	n = 0	n = 1	n = 2	X	n = 0	n = 1	n = 2	X	n = 0	n = 1	n = 2	X	n = 0	n = 1	n = 2
5	0	1,00	0,00	0,00	1	0,09	-0,78	-1,65	0,0	1,00	0,00	0,00	0,4	1,11	2,18	12,04
6	1	0,77	0,44	0,11	2	0,51	-0,11	-0,62	0,3	1,02	0,15	0,01	0,5	0,92	1,66	7,55
7	2	0,22	0,58	0,35	3	0,38	0,32	-0,16	0,6	1,09	0,31	0,05	0,6	0,78	1,30	5,12
8	3	-0,26	0,34	0,49	4	-0,02	0,40	0,22	0,9	1,21	0,50	0,11	0,7	0,66	1,05	3,66
9	4	-0,40	-0,07	0,36	5	-0,31	0,15	0,37	1,2	1,39	0,71	0,20	0,8	0,57	0,86	2,72
10	5	-0,18	-0,33	0,05	6	-0,29	-0,18	0,23	1,5	1,65	0,98	0,34	0,9	0,49	0,72	2,08
11	6	0,15	-0,28	-0,24	7	-0,03	-0,30	-0,06	1,8	1,99	1,32	0,53	1,0	0,42	0,60	1,62
12	7	0,30	0,00	-0,30	8	0,22	-0,16	-0,26	2,1	2,45	1,75	0,78	1,1	0,37	0,51	1,29
13	8	0,17	0,23	-0,11	9	0,25	0,10	-0,23	2,4	3,05	2,30	1,13	1,2	0,32	0,43	1,04
14	9	-0,09	0,25	0,14	10	0,06	0,25	-0,01	2,7	3,84	3,02	1,61	1,3	0,28	0,37	0,85
15	10	-0,25	0,04	0,25	11	-0,17	0,16	0,20	3,0	4,88	3,95	2,25	1,4	0,24	0,32	0,70
16	11	-0,17	-0,18	0,14	12	-0,23	-0,06	0,22	3,3	6,24	5,18	3,10	1,5	0,21	0,28	0,58
17	12	0,05	-0,22	-0,08	13	-0,08	-0,21	0,05	3,6	8,03	6,79	4,25	1,6	0,19	0,24	0,49
18	13	0,21	-0,07	-0,22	14	0,13	-0,17	-0,15	3,9	10,37	8,91	5,80	1,7	0,17	0,21	0,41
19	14	0,17	0,13	-0,15	15	0,21	0,02	-0,20	4,2	13,44	11,71	7,87	1,8	0,15	0,18	0,35

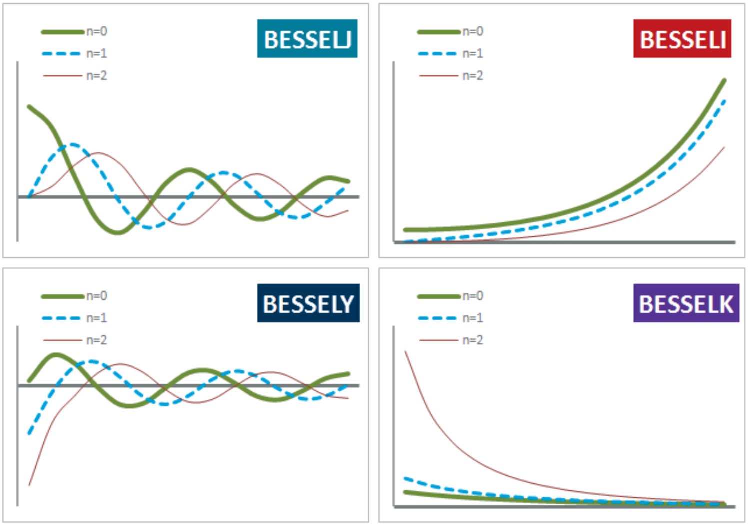


Figure 14–38
Représentation graphique des quatre fonctions de Bessel proposées par Excel, avec les valeurs calculées figure 14-37.

Fonctions d'erreur

Ces fonctions sont particulièrement utilisées pour calculer des probabilités d'erreur, notamment dans le domaine des communications. Pour comprendre leur genèse, il faut commencer par s'intéresser aux variables aléatoires « gaussiennes » qui ont la particularité de posséder une distribution de probabilité (voir le chapitre 13) décrivant de façon précise de nombreux phénomènes aléatoires réels (traitement du signal, prévision des crues, etc.). Les gaussiennes sont régies par la loi dite « normale » (voir le chapitre 13) et admettent pour densité de probabilité la fonction dont l'expression générale est donnée figure 13-48. Par la suite, nous travaillerons sur les lois normales centrées réduites qui sont des lois normales d'espérance nulle (`m = 0`) et d'écart-type égal à `1` (`σ = 1`) (voir la figure 13-50).

RAPPEL Densité de probabilité

Cette notion a été largement abordée dans le cadre du chapitre 13. Rappelons simplement que $P(X = x_0)$ donne la probabilité que la variable aléatoire X prenne la valeur x_0. Rappelons également que la somme des probabilités de tous les événements possibles est toujours égale à `1`. Dans le cas spécifique d'une loi normale gaussienne centrée réduite, cela se traduit par la relation présentée figure 14-39.

$$\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty} e^{\left(\frac{-t^2}{2}\right)}\, dt = 1$$

Figure 14–39 La somme des probabilités de tous les événements possibles est toujours égale à 1.

Fonction d'erreur ERF

L'expression « fonction d'erreur » véhicule, par certains côtés, un message… susceptible d'induire en erreur ! En effet, l'idée qu'elle suggère à première vue (la quantification d'une probabilité d'erreur) est à peu près opposée à ce qu'elle détermine réellement. La fonction d'erreur, notée communément `ERF(x)`, représente la probabilité qu'une variable aléatoire X prenne ses valeurs dans l'intervalle `[-x, +x]`. En d'autres termes, dans le cas d'une loi normale centrée réduite et en adoptant une phraséologie volontairement simpliste mais pédagogique, la fonction d'erreur calcule la probabilité que les événements les plus probables se réalisent.

La fonction d'erreur n'est autre que la somme des probabilités `P(t)` de tous les événements `t`, situés dans l'intervalle `[-x, +x]`. La loi de probabilité étant continue, cette somme s'exprime selon la formule présentée figure 14-40.

Figure 14–40
Définition de la fonction ERF.

$$erf(x) = \frac{1}{\sqrt{(2\pi)}}\int_{-x}^{x} e^{\left(\frac{-t^2}{2}\right)}\, dt$$

La fonction d'erreur `ERF(x)` quantifie tout simplement l'aire située sous la courbe dessinée par la loi de probabilité et délimitée par les valeurs `-x` et `+x`.

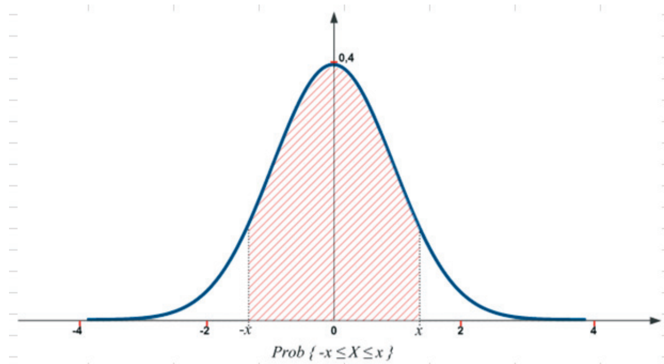


Figure 14–41
Aire correspondant à ERF(x).

COMPRENDRE Obtenir la formule générale de ERF

En tenant compte de la parité de la loi de probabilité et en procédant au changement de variable `z = t / 2`, on obtient la formule canonique de la fonction d'erreur.

$$erf(x) = \frac{2}{\sqrt{(\pi)}} \int_0^x e^{(-z^2)}\, dz$$

Figure 14–42 Formule canonique de la fonction d'erreur.

En dépit de l'arsenal mathématique dont jouissent les théoriciens des probabilités, il n'existe pas de primitive permettant de calculer exactement cette intégrale. En revanche, une valeur approchée de la fonction d'erreur peut être établie avec un haut degré de précision, en effectuant dans un premier temps une décomposition en série entière de la fonction e^{-z^2} et en procédant ensuite à une intégration de chaque terme. En passant rapidement sur ces calculs douloureux, nous arrivons à la formule générale utilisée pour calculer la fonction d'erreur (figure 14-43).

Figure 14–43
Formule générale utilisée pour calculer la fonction d'erreur.

$$erf(x) = \frac{2}{\sqrt{(\pi)}} \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)n!}\, x^{2n+1}$$

Fonction d'erreur complémentaire

De façon indirecte, la fonction d'erreur complémentaire, notée `ERFC(x)`, sert à déterminer la probabilité qu'une variable normale centrée réduite `X` dépasse la valeur `x`. La partie hachurée de la fonction *ERFC* représentée figure 14-44 est complémentaire de celle correspondant à la fonction d'erreur (figure 14-41).

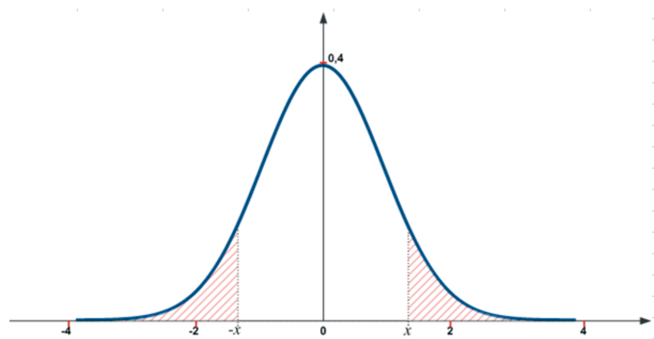


Figure 14–44
Aire correspondant à la fonction ERFC(x).

COMPRENDRE Obtenir la formule générale de ERFC

En tenant compte de la parité de la loi de probabilité et en procédant au même changement de variable que précédemment, la fonction d'erreur complémentaire est déterminée à l'aide de l'expression présentée figure 14-45.

Figure 14–45 Formule canonique de la fonction d'erreur complémentaire.

$$erfc(x) = \frac{2}{\sqrt{(\pi)}} \int_x^{+\infty} e^{\left(-z^2\right)}\, dz$$

En se rappelant que la somme des probabilités de tous les événements est égale à 1, on remarque que la fonction d'erreur complémentaire est liée à la fonction d'erreur par la relation présentée figure 14-46.

Figure 14–46 Relation liant les fonctions ERF et ERFC.

$$erfc(x) = 1 - erf(x)$$

Néanmoins, nous ne sommes pas encore au bout de nos peines, car l'information vraiment exploitable que l'on recherche est la probabilité de dépassement (en d'autres termes, on s'intéresse à la partie droite de la « queue de gaussienne »). Cette probabilité est déterminée par la fonction de Marcum, qui s'exprime à travers la formule présentée figure 14-47.

Figure 14–47 Fonction de Marcum.

$$Q(x) = \frac{1}{\sqrt{(2\pi)}} \int_x^{+\infty} e^{\left(-\frac{t^2}{2}\right)}\, dt$$

Sans calculer cette intégrale et en effectuant le changement de variable approprié, on vérifie aisément que les fonctions de Marcum et d'erreur complémentaire sont liées l'une à l'autre à travers la relation présentée figure 14-48. Ces fonctions sont particulièrement utilisées pour calculer des probabilités d'erreur, notamment dans le domaine des communications.

Figure 14–48 Relations liant les fonctions ERFC et les fonctions de Marcum.

$$erfc(x) = 2Q(\sqrt{2}\, x)$$

$$Q(x) = \frac{1}{2} erfc\left(\frac{x}{\sqrt{2}}\right)$$

Pour conclure ce rapide tour d'horizon, vous remarquerez qu'en raison des relations existant entre les fonctions d'erreur, d'erreur complémentaire et de Marcum, la fonction d'erreur fournit bel et bien une indication permettant, in fine, de calculer une probabilité d'erreur.

Fonctions d'erreur proposées par Excel

Excel propose quatre fonctions d'erreur : *ERF*, *ERF.PRECIS*, *ERFC* et *ERFC.PRECIS*.

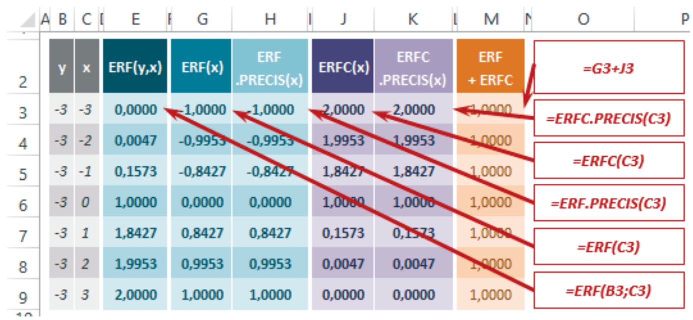


Figure 14–49
Mise en œuvre des fonctions ERF, ERF.PRECIS, ERFC et ERFC.PRECIS.

Tableau 14–17 Fonctions d'erreur

Fonction	Description
ERF	Cette fonction renvoie le résultat du calcul correspondant à l'équation présentée figure 14-40. Elle utilise deux arguments dont l'un est optionnel. Si vous ne précisez qu'un argument `(x)`, `ERF` calcule l'intégrale entre `-x` et `x`. Si vous précisez deux arguments `(y, x)`, `ERF` calcule l'intégrale entre `y` et `x`.
ERF.PRECIS	Cette fonction renvoie les mêmes résultats que la fonction `ERF`… ce qui jette un doute certain sur son utilité.
ERFC	Cette fonction renvoie le résultat du calcul correspondant à l'équation présentée figure 14-45. Son argument unique représente la limite inférieure pour le calcul de l'intégrale.
ERFC.PRECIS	Cette fonction renvoie les mêmes résultats que la fonction `ERFC`… ce qui jette également un doute certain sur son utilité.

Fonctions spéciales

Les trois dernières fonctions de la catégorie « Ingénieur » ne peuvent pas être associées à une famille particulière. On trouve *DELTA* et *SUP.SEUIL*, qui facilitent la comparaison entre deux valeurs, et *CONVERT*, qui convertit une valeur d'une unité dans une autre.

Comparer deux valeurs

Les fonctions *DELTA* et *SUP.SEUIL* servent à comparer deux nombres.

Figure 14–50
Mise en œuvre de la fonction DELTA.

	B	C	E	G
2	Arguments		Syntaxe	Résultat
4	Nombre 1	Nombre 2		
6	-9	-9	=DELTA(B6;C6)	1
8	-9	4	=DELTA(B8;C8)	0
10	1,5	1,5	=DELTA(B10;C10)	1
12	2,5	1,5	=DELTA(B12;C12)	0
14	10	10	=DELTA(B14;C14)	1

Tableau 14–18 Fonctions de comparaison

Fonction	Description
DELTA	Cette fonction utilise deux arguments, deux nombres. Elle teste leur égalité et renvoie 1 s'ils sont égaux. S'ils ne le sont pas, elle renvoie 0.
SUP.SEUIL	Cette fonction utilise deux arguments, deux nombres. Elle renvoie 1 si le premier argument (valeur testée) est supérieur ou égal au deuxième (seuil). Sinon, elle renvoie 0.

Figure 14–51
Mise en œuvre de la fonction SUP.SEUIL.

	B	C	E	G
2	Arguments		Syntaxe	Résultat
4	Nombre	Seuil		
6	-3	5	=SUP.SEUIL(B6;C6)	0
8	0	5	=SUP.SEUIL(B8;C8)	0
10	5	5	=SUP.SEUIL(B10;C10)	1
12	7	5	=SUP.SEUIL(B12;C12)	1
14	10	5	=SUP.SEUIL(B14;C14)	1

EN PRATIQUE À quoi sert la fonction DELTA ?

Si vous additionnez les résultats de plusieurs fonctions *DELTA*, vous obtenez le nombre de paires égales. Dans l'exemple présenté figure 14-50, la somme des valeurs de la plage *G6:G14* renvoie 3, correspondant effectivement au nombre de paires égales.
En mathématiques, le symbole de Kronecker (δ) est une fonction de deux variables qui est égale à 1 si celles-ci sont égales et 0 sinon.

EN PRATIQUE À quoi sert la fonction SUP.SEUIL ?

Si vous additionnez les résultats de plusieurs fonctions *SUP.SEUIL*, vous obtenez le nombre de valeurs supérieures à un seuil. Dans l'exemple présenté figure 14-51, la somme des valeurs de la plage *G6:G14* renvoie 3, correspondant effectivement au nombre de valeurs supérieures à 5.

Convertir les unités

La fonction *CONVERT* sert à convertir dans une unité donnée, un nombre exprimé dans une autre unité.

Figure 14–52
Mise en œuvre de la fonction CONVERT.

	B	C	D	F	H	J
2	*Arguments*			*Syntaxe*	*Résultat*	
4	*Nom-bre*	*De unité*	*À unité*			
6	1	mi	m	*=CONVERT (B6;C6;D6)*	1 609,34	*Miles en mètres*
8	1	day	mn	*=CONVERT (B8;C8;D8)*	1 440	*Jours en minutes*
10	0	C	F	*=CONVERT (B10;C10;D10)*	32	*Degrés Celsius en degrés Fahrenheit*

Tableau 14–19 Fonction de conversion

Fonction	Description
CONVERT	Cette fonction utilise trois arguments. Le premier est un nombre. Il est exprimé dans une unité (deuxième argument) et vous souhaitez le convertir dans une autre unité (troisième argument). Pour préciser les deuxième et troisième arguments, il faut choisir deux unités de la même famille et, surtout, respecter les majuscules et minuscules pour les saisir tels qu'ils apparaissent dans les figures 14-53, 14-54 et 14-55.

Figure 14–53
Convertir les poids, heures, pressions et forces.

Poids et masse	
Nom de l'unité	Libellé à utiliser dans les arguments
Gramme	"g"
Slug	"sg"
Livre masse	"lbm"
U (unité de masse atomique)	"u"
Once	"ozm"

Heure	
Nom de l'unité	Libellé à utiliser dans les arguments
Année	"yr"
Jour	"day"
Heure	"hr"
Minute	"mn"
Seconde	"sec"

Pression	
Nom de l'unité	Libellé à utiliser dans les arguments
Pascal	"Pa"
Atmosphère	"atm"
mm de mercure	"mmHg"

Force	
Nom de l'unité	Libellé à utiliser dans les arguments
Newton	"N"
Dyne	"dyn"
Livre force	"lbf"

Figure 14–54
Convertir les puissances, distances, températures et mesures liées au magnétisme.

Puissance	
Nom de l'unité	Libellé à utiliser dans les arguments
Cheval	"HP"
Watt	"W"

Magnétisme	
Nom de l'unité	Libellé à utiliser dans les arguments
Tesla	"T"
Gauss	"ga"

Distance	
Nom de l'unité	Libellé à utiliser dans les arguments
Mètre	"m"
Mille	"mi"
Mille nautique	"Nmi"
Pouce	"in"
Pied	"ft"
Yard	"yd"
Angstrom	"ang"
Pica (1/72 po)	"Pica"

Température	
Nom de l'unité	Libellé à utiliser dans les arguments
Degré Celsius	"C"
Degré Fahrenheit	"F"
Degré Kelvin	"K"

Figure 14–55
Convertir les mesures d'énergie et de capacité.

Énergie	
Nom de l'unité	Libellé à utiliser dans les arguments
Joule	"J"
Erg	"e"
Calorie (4,183991 J)	"c"
Calorie (4,186795 J)	"cal"
Électronvolt	"eV"
Cheval-heure	"HPh"
Watt-heure	"Wh"
Livre pied	"flb"
British Thermal Unit	"BTU"

Mesures de capacité	
Nom de l'unité	Libellé à utiliser dans les arguments
Cuillère à thé	"tsp"
Cuillère à soupe	"tbs"
Once fluide	"oz"
Tasse	"cup"
Pinte U.S.A.	"pt"
Pinte R.U.	"uk_pt"
Quart	"qt"
Gallon	"gal"
Litre	"l"

Annexe

Correspondances options Excel 2003 – Excel 2010

Affichage

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Afficher : Volet Office au démarrage	absent	absent
Afficher : Barre de formule	Affichage	Afficher>Barre de formule
Afficher : Barre d'état	absent	absent
Afficher : Fenêtres dans la barre des tâches	Fichier>Options	Options avancées>Afficher>Afficher toutes les fenêtres dans la barre des tâches
Commentaires : Aucun	Fichier>Options	Options avancées>Afficher>Aucun commentaire ou indicateur
Commentaires : Indicateur seul	Fichier>Options	Options avancées>Afficher>Indicateur seul, et commentaires au survol
Commentaires : Commentaire et indicateur	Fichier>Options	Options avancées>Afficher>Commentaires et indicateurs
Objets : Afficher tout	Fichier>Options	Options avancées>Afficher les options pour ce classeur>Pour les objets, afficher tout
Objets : Indicateurs de position	absent	absent
Objets : Masquer tout	Fichier>Options	Options avancées>Afficher les options pour ce classeur>Pour les objets, afficher rien (Masquer les objets)
Fenêtres : Sauts de page	Fichier>Options	Options avancées>Afficher les options pour cette feuille de calcul>Afficher les sauts de page

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Fenêtres : Formules	Fichier>Options	Options avancées>Afficher les options pour cette feuille de calcul>Formules dans les cellules au lieu de leurs résultats calculés
Fenêtres : Quadrillage	Fichier>Options	Options avancées>Afficher les options pour cette feuille de calcul>Afficher le quadrillage
Fenêtres : Couleur du quadrillage	Fichier>Options	Options avancées>Afficher les options pour cette feuille de calcul>Couleur du quadrillage
Fenêtres : En-têtes de ligne et de colonne	Fichier>Options	Options avancées>Afficher les options pour cette feuille de calcul>Afficher les en-têtes de ligne et de colonne
Fenêtres : Symboles du plan	Fichier>Options	Options avancées>Afficher les options pour cette feuille de calcul>Afficher les symboles du plan si un plan est appliqué
Fenêtres : Valeurs zéro	Fichier>Options	Options avancées>Afficher les options pour cette feuille de calcul>Afficher un zéro dans les cellules qui ont une valeur nulle
Fenêtres : Barre de défilement horizontale	Fichier>Options	Options avancées>Afficher les options pour ce classeur>Afficher la barre de défilement horizontale
Fenêtres : Barre de défilement verticale	Fichier>Options	Options avancées>Afficher les options pour ce classeur>Afficher la barre de défilement verticale
Fenêtres : Onglets de classeur	Fichier>Options	Options avancées>Afficher les options pour ce classeur>Afficher les onglets de classeur

Calcul

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Calcul : Automatique	Formules	Calcul>Options de calcul>Automatique
Calcul : Automatique sauf les tables	Formules	Calcul>Options de calcul>Automatique sauf dans les tables de données
Calcul : Sur ordre	Formules	Calcul>Options de calcul>Manuel
Calcul : Recalcul avant enregistrement	Fichier>Options	Formules>Mode de calcul>Recalculer le classeur avant de l'enregistrer
Calcul : Calculer maintenant	Formules	Calcul>Calculer maintenant
Calcul : Calculer document	Formules	Calcul>Calculer la feuille
Calcul : Itération	Fichier>Options	Formules>Mode de calcul>Activer le calcul itératif
Calcul : Nb maximal d'itérations	Fichier>Options	Formules>Mode de calcul>Nb maximal d'itérations
Calcul : Écart maximal	Fichier>Options	Formules>Mode de calcul>Écart maximal
Options de classeur : Mise à jour des références hors programme	Fichier>Options	Options avancées>Lors du calcul de ce classeur>Mise à jour des liaisons vers d'autres documents

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Options de classeur : Calcul avec la précision au format affiché	Fichier>Options	Options avancées>Lors du calcul de ce classeur>Définir le calcul avec la précision au format affiché
Options de classeur : Calendrier depuis 1904	Fichier>Options	Options avancées>Lors du calcul de ce classeur>Utiliser le calendrier depuis 1904
Options de classeur : Enregistrer les valeurs des liaisons externes	Fichier>Options	Options avancées>Lors du calcul de ce classeur>Enregistrer les valeurs des liaisons externes
Options de classeur : Accepter les étiquettes dans les formules	absent	absent

Modification

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Modification directe	Fichier>Options	Options avancées>Options d'édition>Modification directe
Glissement-déplacement de la cellule	Fichier>Options	Options avancées>Options d'édition>Glissement-déplacement de la cellule
Alerte avant remplacement	Fichier>Options	Options avancées>Options d'édition>Alerte avant remplacement
Déplacer la sélection après validation	Fichier>Options	Options avancées>Options d'édition>Déplacer la sélection après validation
Sens	Fichier>Options	Options avancées>Options d'édition>Sens
Décimale fixe	Fichier>Options	Options avancées>Options d'édition>Décimale fixe
Place	Fichier>Options	Options avancées>Options d'édition>Place
Couper, copier et trier les objets avec les cellules	Fichier>Options	Options avancées>Couper, copier et coller>Couper, copier et trier les objets avec les cellules
Confirmation de la mise à jour automatique des liens	Fichier>Options	Options avancées>Général>Confirmation de la mise à jour automatique des liens
Produire un retour animé	Fichier>Options	Options avancées>Général>Produire un retour animé
Saisie semi-automatique des valeurs de cellule	Fichier>Options	Options avancées>Options d'édition>Saisie semi-automatique des valeurs de cellule
Étendre les formules et formats de plage de données	Fichier>Options	Options avancées>Options d'édition>Étendre les formules et formats de plage de données
Activer la saisie automatique de pourcentage	Fichier>Options	Options avancées>Options d'édition>Activer la saisie automatique de pourcentage

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Afficher les boutons d'options de collage	Fichier>Options	Options avancées>Couper, copier et coller>Afficher le bouton Options de collage lorsqu'un contenu est collé
Afficher les boutons d'insertion	Fichier>Options	Options avancées>Couper, copier et coller>Afficher les boutons d'options d'insertion

Général

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Style de référence L1C1	Fichier>Options	Formules>Manipulation de formules>Style de référence L1C1
Ignorer les autres applications	Fichier>Options	Options avancées>Général>Ignorer les autres applications qui utilisent l'échange dynamique de données
Info-bulles de fonctions	Fichier>Options	Options avancées>Afficher>Afficher les info-bulles des fonctions
Liste des derniers fichiers utilisés	Fichier	Récent>Accéder rapidement à ce nombre de classeurs récents
Afficher la fenêtre des Propriétés	Fichier	Informations>Propriétés>Afficher le panneau de documents
Produire un retour sonore	Fichier>Options	Options avancées>Général>Produire un retour sonore
Zoom avec la roulette IntelliMouse	Fichier>Options	Options avancées>Options d'édition>Zoom avec la roulette IntelliMouse
Options Web	Fichier>Options	Options avancées>Général>Options Web
Options des services	Fichier	Partager>Publier vers Excel Services>Publier vers Excel Services>Options Excel Services
Nombre de feuilles de calcul par nouveau classeur	Fichier>Options	Général>Lors de la création de classeurs>Inclure ces feuilles
Police standard	Fichier>Options	Général>Lors de la création de classeurs>Utiliser cette police
Dossier par défaut	Fichier>Options	Enregistrement>Enregistrer les classeurs>Dossier par défaut
Au démarrage, ouvrir tous les fichiers du dossier	Fichier>Options	Options avancées>Général>Au démarrage, ouvrir tous les fichiers du dossier
Nom d'utilisateur	Fichier>Options	Général>Personnaliser votre copie de Microsoft Office

Transition

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Enregistrer les fichiers Excel sous	Fichier>Options	Enregistrement>Enregistrer les classeurs>Enregistrer les fichiers au format suivant

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Touche d'accès au menu Microsoft Office Excel	Fichier>Options	Options avancées>Compatibilité avec Lotus>Touche d'accès au menu Microsoft Excel
Touches alternatives de déplacement	Fichier>Options	Options avancées>Compatibilité avec Lotus>Touches alternatives de déplacement
Autre interprétation des formules	Fichier>Options	Options avancées>Paramètres de compatibilité avec Lotus>Autre interprétation des formules
Autre mode de saisie des formules	Fichier>Options	Options avancées>Paramètres de compatibilité avec Lotus>Autre mode de saisie des formules

Liste pers.

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Liste pers.	Fichier>Options	Options avancées>Général>Modifier les listes personnalisées

Graphique

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Graphique actif : Traitement des cellules vides Non tracées (laisse un vide)	Outils de graphique> Création	Données>Sélectionner des données>Cellules masquées et cellules vides>Relier les points de données par une courbe
Graphique actif : Traitement des cellules vides Valeur zéro	Outils de graphique> Création	Données>Sélectionner des données>Cellules masquées et cellules vides>Valeur zéro
Graphique actif : Traitement des cellules vides Interpolées	Outils de graphique> Création	Données>Sélectionner des données>Cellules masquées et cellules vides>Intervalles
Graphique actif : Tracer les cellules visibles seulement	Outils de graphique> Création	Données>Sélectionner des données>Cellules masquées et cellules vides>Afficher les données des lignes et colonnes masquées
Info-bulles de graphiques : Afficher les noms	Fichier>Options	Options avancées>Graphique>Afficher les noms des éléments
Info-bulles de graphiques : Afficher les valeurs	Fichier>Options	Options avancées>Graphique>Afficher les valeurs des points de données

Couleur

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Couleurs Modifier	Mise en page	Thèmes>Couleurs
Copier les couleurs de	Mise en page	Thèmes>Thèmes>Rechercher les thèmes

International

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Gestion des nombres : Séparateur de décimale	Fichier>Options	Options avancées>Options d'édition>Séparateur de décimale
Gestion des nombres : Séparateur de milliers	Fichier>Options	Options avancées>Options d'édition>Séparateur des milliers
Gestion des nombres : Utiliser les séparateurs système	Fichier>Options	Options avancées>Options d'édition>Utiliser les séparateurs système
Impression : Redimensionner A4/papier à lettres	Fichier>Options	Options avancées>Général>Ajuster le contenu aux formats papier A4 ou 8,5 X 11
De droite à gauche : Orientation par défaut de droite à gauche	Fichier>Options	Options avancées>Afficher>Orientation par défaut de droite à gauche
De droite à gauche : Orientation par défaut de gauche à droite	Fichier>Options	Options avancées>Afficher>Orientation par défaut de gauche à droite
De droite à gauche : Déplacement du curseur logique	Fichier>Options	Options avancées>Options d'édition>Déplacement du curseur logique
De droite à gauche : Déplacement du curseur visuel	Fichier>Options	Options avancées>Options d'édition>Déplacement du curseur visuel
De droite à gauche : Afficher la feuille active de droite à gauche	Fichier>Options	Options avancées>Afficher les options pour cette feuille de calcul>Afficher la feuille de droite à gauche

Options

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Paramètres : Enregistrer les informations de récupération automatique toutes les	Fichier>Options	Enregistrement>Enregistrer les classeurs>Enregistrer les informations de récupération automatique toutes les
Emplacement d'enregistrement de récupération automatique	Fichier>Options	Enregistrement>Enregistrer les classeurs>Emplacement du fichier de récupération automatique
Options de classeur : Désactiver la récupération automatique	Fichier>Options	Enregistrement>Exceptions de récupération automatique pour>Désactiver la récupératon automatique pour ce classeur uniquement

Vérification des erreurs

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Paramètres : activer la vérification des erreurs d'arrière-plan	Fichier>Options	Formules>Vérification des erreurs>Activer la vérification des erreurs en arrière-plan
Paramètres : couleur de l'indicateur d'erreur	Fichier>Options	Formules>Vérification des erreurs>Indiquer les erreurs à l'aide de cette couleur
Paramètres : Rétablir les erreurs ignorées	Fichier>Options	Formules>Vérification des erreurs>Rétablir les erreurs ignorées
Règles : Donne une valeur d'erreur	Fichier>Options	Formules>Règles de vérification des erreurs>Cellules dont les formules génèrent des erreurs
Règles : Date du texte avec des années à deux chiffres	Fichier>Options	Formules>Règles de vérification des erreurs>Cellules contenant des années à deux chiffres
Règles : Nombre stocké en tant que texte	Fichier>Options	Formules>Règles de vérification des erreurs>Nombres mis en forme en tant que texte ou précédés d'une apostrophe
Règles : Formule incohérente dans la zone	Fichier>Options	Formules>Règles de vérification des erreurs>Formules incohérentes avec d'autres formules de la zone
Règles : La formule omet des cellules dans la zone	Fichier>Options	Formules>Règles de vérification des erreurs>Cellules omises dans une formule appliquée à une zone
Règles : Des cellules déverrouillées contiennent des formules	Fichier>Options	Formules>Règles de vérification des erreurs>Formules dans des cellules déverrouillées
Règles : Formules faisant référence à des cellules vides	Fichier>Options	Formules>Règles de vérification des erreurs>Formules faisant référence à des cellules vides

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Règles : Erreur de validation des données de la liste	Fichier>Options	Formules>Règles de vérification des erreurs>Données incorrectes dans un tableau

Orthographe

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Langue du dictionnaire	Fichier>Options	Vérification>Lors de la correction orthographique dans les programmes Microsoft Office>Langue du dictionnaire
Ajouter les mots à	Fichier>Options	Vérification>Lors de la correction orthographique dans les programmes Microsoft Office>Dictionnaires personnels
Suggérer à partir du dictionnaire principal uniquement	Fichier>Options	Vérification>Lors de la correction orthographique dans les programmes Microsoft Office>Suggérer à partir du dictionnaire principal uniquement
Ignorer les mots en MAJUSCULES	Fichier>Options	Vérification>Lors de la correction orthographique dans les programmes Microsoft Office>Ignorer les mots en MAJUSCULES
Ignorer les mots contenant des chiffres	Fichier>Options	Vérification>Lors de la correction orthographique dans les programmes Microsoft Office>Ignorer les mots qui contiennent des chiffres
Ignorer les adresses Internet et les adresses de fichiers	Fichier>Options	Vérification>Lors de la correction orthographique dans les programmes Microsoft Office>Ignorer les chemins d'accès aux fichiers
Options de correction automatique	Fichier>Options	Vérification>Options de correction automatique>Options de correction automatique

Sécurité

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Paramètres de cryptage de fichiers pour ce classeur : Mot de passe pour la lecture	Fichier	Enregistrer sous>Outils>Options générales>Mot de passe pour la lecture
Paramètres de partage de fichiers pour ce classeur : Mot de passe pour la modification	Fichier	Enregistrer sous>Outils>Options générales>Mot de passe pour la modification

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Paramètres de partage de fichiers pour ce classeur : Lecture seule recommandée	Fichier	Enregistrer sous>Outils>Options générales>Lecture seule recommandée
Paramètres de partage de fichiers pour ce classeur : Signatures numériques	Insertion	Texte>Ligne de signature
Options de confidentialité : Supprimer les informations personnelles des propriétés de ce fichier à l'enregistrement	Fichier>Options	Centre de gestion de la confidentialité>Paramètres du Centre de gestion de la confidentialité>Paramètres spécifiques au document>Supprimer les informations personnelles des propriétés du fichier lors de l'enregistrement
Sécurité des macros : Sécurité des macros	Développeur	Code>Sécurité des macros

Correspondances commandes Excel 2003 – Excel 2010

Fichier

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Nouveau	Fichier	Nouveau
Ouvrir	Fichier	Ouvrir
Fermer	Fichier	Fermer
Enregistrer	Fichier	Enregistrer
Enregistrer sous	Fichier	Enregistrer sous
Enregistrer en tant que page Web	Fichier	Enregistrer sous>Type>Page Web (*.htm, *.html)
Enregistrer l'espace de travail	Affichage	Fenêtre>Enregistrer l'espace de travail
Recherche de fichiers	Fichier	Ouvrir>Partie supérieure de la boîte de dialogue
Autorisation	Fichier	Informations>Autorisations>Protéger le classeur
Aperçu de la page Web		Commande existante, mais non présente dans le ruban
Mise en page	Mise en page	Mise en page>Lanceur de boîte de dialogue
Zone d'impression	Mise en page	Mise en page>Zone d'impression
Aperçu avant impression	Fichier	Imprimer>Partie droite de la fenêtre

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Imprimer	Fichier	Imprimer>Partie gauche de la fenêtre
Envoyer vers	Fichier	Partager>Diverses commandes présentes dans la fenêtre de droite, d'autres existent mais ne sont pas proposées dans la fenêtre
Propriétés	Fichier	Informations>Propriétés>Propriétés avancées
Fichiers récents	Fichier	Récent
Quitter	Fichier	Quitter

Édition

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Annuler	Barre d'outils Accès rapide	Annuler
Répéter	Barre d'outils Accès rapide	Répéter
Couper	Accueil	Presse-papiers>Couper
Copier	Accueil	Presse-papiers>Copier
Presse-papiers	Accueil	Presse-papiers>Lanceur de boîte de dialogue
Coller	Accueil	Presse-papiers>Coller
Collage spécial	Accueil	Presse-papiers>Coller>Collage spécial
Coller comme lien hypertexte		Commande existante, mais non présente dans le ruban. Voir le chapitre 13.
Remplissage	Accueil	Édition>Remplissage
Effacer	Accueil	Édition>Effacer
Supprimer	Accueil	Cellules>Supprimer>Supprimer les cellules
Supprimer une feuille	Accueil	Cellules>Supprimer>Supprimer une feuille
Déplacer ou copier une feuille	Accueil	Cellules>Format>Déplacer ou copier une feuille
Rechercher	Accueil	Édition>Rechercher et sélectionner>Rechercher
Remplacer	Accueil	Édition>Rechercher et sélectionner>Remplacer
Atteindre	Accueil	Édition>Rechercher et sélectionner>Atteindre
Liaisons	Fichier	Information>Modifier les liens d'accès au fichier
Liaisons	Données	Connexions>Modifier les liens d'accès
Objet	Outils de l'objet	Correspond à l'onglet contextuel qui apparaît lorsqu'on sélectionne un objet

Affichage

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Normal	Affichage	Affichages classeur>Normal
Aperçu des sauts de page	Affichage	Affichages classeur>Aperçu des sauts de page
Volet Office	absent	absent
Barres d'outils	absent	absent
Barre de formule	Affichage	Afficher>Barre de formule
Barre d'état	absent	Mais peut quand même être masquée si l'on passe en plein écran (Affichage>Affichages classeur)
En-tête et pied de page	Insertion	Texte>En-tête et pied de page
Commentaires	Révision	Commentaires>Afficher tous les commentaires
Affichages personnalisés	Affichage	Affichages classeur>Affichages personnalisés
Plein écran	Affichage	Affichages classeur>Plein écran
Zoom	Affichage	Zoom>Zoom

Insertion

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Cellules	Accueil	Cellules>Insérer>Insérer des cellules
Lignes	Accueil	Cellules>Insérer>Insérer des lignes dans la feuille
Colonnes	Accueil	Cellules>Insérer>Insérer des colonnes dans la feuille
Feuille	Accueil	Cellules>Insérer>Insérer une feuille
Graphique	Insertion	Graphiques>Lanceur de boîte de dialogue
Caractères spéciaux	Insertion	Symboles>Symbole
Saut de page	Mise en page	Mise en page>Sauts de page>Insérer un saut de page
Fonction	Formules	Bibliothèque de fonctions>Insérer une fonction
Nom>Définir	Formules	Noms définis>Définir un nom>Définir un nom
Nom>Coller	Formules	Noms définis>Utiliser dans la formule
Nom>Créer	Formules	Noms définis>Créer à partir de la sélection
Nom>Appliquer	Formules	Noms définis>Définir un nom>Appliquer les noms
Nom>Étiquette	absent	absent
Commentaire	Révision	Commentaires>Nouveau commentaire
Image>Images clipart	Insertion	Illustrations>Images ClipArt
Image>À partir du fichier	Insertion	Illustrations>Image

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Image>À partir d'un scanneur ou d'un appareil-photo numérique	absent	absent
Image>Formes automatiques	Insertion	Illustrations>Formes
Image>WordArt	Insertion	Texte>WordArt
Image>Organigramme hiérarchique	Insertion	Illustrations>SmartArt
Diagramme	Insertion	Illustrations>SmartArt
Objet	Insertion	Texte>Objet
Lien hypertexte	Insertion	Liens>Lien hypertexte

Format

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Cellule	Accueil	Cellules>Format>Format de cellule
Ligne>Hauteur	Accueil	Cellules>Format>Hauteur de ligne
Ligne>Ajustement automatique	Accueil	Cellules>Format>Ajuster la hauteur de ligne
Ligne>Masquer	Accueil	Cellules>Format>Masquer & Afficher>Masquer les lignes
Ligne>Afficher	Accueil	Cellules>Format>Masquer & Afficher>Afficher les lignes
Colonne>Largeur	Accueil	Cellules>Format>Largeur de colonnes
Colonne>Ajustement automatique	Accueil	Cellules>Format>Ajuster la largeur de colonne
Colonne>Masquer	Accueil	Cellules>Format>Masquer & Afficher>Masquer les colonnes
Colonne>Afficher	Accueil	Cellules>Format>Masquer & Afficher>Afficher les colonnes
Colonne>Largeur standard	Accueil	Cellules>Format>Largeur par défaut
Feuille>Renommer	Accueil	Cellules>Format>Renommer la feuille
Feuille>Masquer	Accueil	Cellules>Format>Masquer & Afficher>Masquer la feuille
Feuille>Afficher	Accueil	Cellules>Format>Masquer & Afficher>Afficher la feuille
Feuille>Arrière-plan	Mise en page	Mise en page>Arrière plan
Feuille>Couleur d'onglet	Accueil	Cellules>Format>Couleur d'onglet
Mise en forme automatique	Accueil	Style>Mettre sous forme de tableau
Mise en forme conditionnelle	Accueil	Style>Mise en forme conditionnelle
Style	Accueil	Style>Styles de cellules

Outils

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Orthographe	Révision	Vérification>Orthographe
Bibliothèque de recherche	Révision	Vérification>Recherche
Vérification des erreurs	Formules	Audit de formules>Vérification des erreurs>Vérification des erreurs
Espace de travail partagé	Fichier	Partager>Enregistrer dans Share Point
Partager le classeur	Révision	Modifications>Partager le classeur
Suivi des modifications	Révision	Modifications>Suivi des modifications
Comparaison et fusion des classeurs		Commande existante, mais non présente dans le ruban. Voir le chapitre 13.
Protection>Protéger la feuille	Révision	Modifications>Protéger la feuille
Protection>Permettre aux utilisateurs de modifier des plages	Révision	Modifications>Permettre la modification des plages
Protection>Protéger le classeur	Révision	Modifications>Protéger le classeur
Protection>Protéger et partager le classeur	Révision	Modifications>Protéger et partager le classeur
Collaboration en ligne	absent	absent
Valeur cible	Données	Outils de données>Analyse de scénarios>Valeur cible
Gestionnaire de scénarios	Données	Outils de données>Analyse de scénarios>Gestionnaire de scénarios
Audit de formules>Repérer les antécédents	Formules	Audit de formules>Repérer les antécédents
Audit de formules>Repérer les dépendants	Formules	Audit de formules>Repérer les dépendants
Audit de formules>Repérer une erreur	Formules	Audit de formules>Repérer une erreur
Audit de formules>Supprimer toutes les flèches	Formules	Audit de formules>Supprimer toutes les flèches
Audit de formules>Évaluation de formule	Formules	Audit de formules>Évaluation de formule
Audit de formules>Afficher la fenêtre Espions	Formules	Audit de formules>Fenêtre espion
Audit de formules>Mode Audit de formules	Formules	Audit de formules>Afficher les formules
Audit de formules>Afficher la barre d'outils Audit de formules	absent	absent

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Macro>Macros	Affichage	Macros>Macros>Afficher les macros
Macro>Nouvelle macro	Affichage	Macros>Macros>Enregistrer une macro
Macro>Sécurité	Développeur	Code>Sécurité des macros
Macro>Visual Basic Editor	Développeur	Code>Visual Basic
Macro>Microsoft Script Editor	absent	absent
Macros complémentaires	Développeur	Compléments>Compléments
Options de correction automatique	Fichier	Options>Vérification>Options de correction automatique
Personnaliser>Barres d'outils	absent	absent
Personnaliser>Commandes	Fichier	Options>Personnaliser le ruban
Personnaliser>Options>Lister les noms de polices dans leur format de police	absent	absent
Personnaliser>Options>Afficher les Info-bulles	Fichier	Options>Standard>Style d'info-bulle
Options	Fichier	Options (voir le détail dans la section suivante)

Données

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Trier	Données	Trier et filtrer>Trier
Filtrer>Filtre automatique	Données	Trier et filtrer>Filtrer
Filtrer>Afficher tout	Données	Trier et filtrer>Effacer
Filtrer>Filtre élaboré	Données	Trier et filtrer>Avancé
Formulaire		Commande existante, mais non présente dans le ruban
Sous-totaux	Données	Plan>Sous-total
Validation	Données	Outils de données>Validation des données
Table	Données	Outils de données>Analyse de scénarios
Convertir	Données	Outils de données>Convertir
Consolider	Données	Outils de données>Consolider
Grouper et créer un plan>Masquer	Données	Plan>Masquer
Grouper et créer un plan>Afficher les détails	Données	Plan>Afficher les détails

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Grouper et créer un plan>Grouper	Données	Plan>Grouper>Grouper
Grouper et créer un plan>Dissocier	Données	Plan>Dissocier>Dissocier
Grouper et créer un plan>Plan automatique	Données	Plan>Grouper>Plan automatique
Grouper et créer un plan>Effacer le plan	Données	Plan>Dissocier>Effacer le plan
Grouper et créer un plan>Paramètres	Données	Plan>Lanceur de boîte de dialogue
Rapport de tableau croisé dynamique	Insertion	Tableaux>Tableau croisé dynamique
Données externes>Importer des données	Données	Données externes>Les divers boutons du groupe
Données externes>Nouvelle requête sur le Web	Données	Données externes>À partir du site Web
Données externes>Créer une requête	Données	Données externes>À partir d'autres sources>Provenance : Microsoft Query
Données externes>Modifier la requête		Commande existante, mais non présente dans le ruban
Données externes>Propriétés de la plage de données	Données	Connexions>Connexions>Propriétés
Données externes>Paramètres		Commande existante, mais non présente dans le ruban
Liste>Créer une liste	Insertion	Tableaux>Tableau
Liste>Redimensionner la liste	Outils de table> Création	Propriétés>Redimensionner le tableau
Liste>Ligne total	Outils de table> Création	Options de style de tableau>Ligne des totaux
Liste>Convertir en plage	Outils de table> Création	Outils>Convertir en plage
Liste>Publier la liste	Outils de table> Création	Données de table externe>Exporter>Exporter le tableau dans une liste SharePoint
Liste>Afficher la liste sur le serveur	Outils de table> Création	Données de table externe>Ouvrir dans le navigateur
Liste>Supprimer la liaison de la liste	Outils de table> Création	Données de table externe>Supprimer la liaison
Liste>Synchroniser la liste		Commande existante, mais non présente dans le ruban. Voir le chapitre 13.

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Liste>Ignorer les modifications et actualiser		Commande existante, mais non présente dans le ruban. Voir le chapitre 13.
Liste>Masquer la bordure des listes inactives	absent	absent
XML>Importer	Développeur	XML>Importer
XML>Exporter	Développeur	XML>Exporter
XML>Actualiser les données XML	Développeur	XML>Actualiser les données
XML>Source XML	Développeur	XML>Source
XML>Propriétés du mappage XML	Développeur	XML>Propriétés du mappage
XML>Modifier la requête	absent	absent
XML>Kits d'extension XML	Développeur	XML>Kits d'extension
Actualiser les données	Données	Connexions>Actualiser tout

Fenêtre

Excel 2003	Excel 2010 onglet	Excel 2010 commande
Nouvelle fenêtre	Affichage	Fenêtre>Nouvelle fenêtre
Réorganiser	Affichage	Fenêtre>Réorganiser tout
Comparer en côte à côte avec	Affichage	Fenêtre>Afficher côte à côte
Masquer	Affichage	Fenêtre>Masquer
Afficher	Affichage	Fenêtre>Afficher
Fractionner	Affichage	Fenêtre>Fractionner
Figer les volets	Affichage	Fenêtre>Figer les volets
Classeurs ouverts	Affichage	Fenêtre>Changement de fenêtre