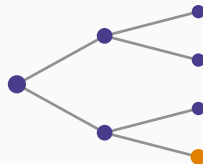


Probabilités

Chapitre 18 : La loi normale



La loi qui domine tout

Une seule table pour toutes

La normale au travail

Ce qu'il faut retenir

La loi qui domine tout

Loi normale

X suit une loi **normale** — ou de Laplace-Gauss — d'espérance m et d'écart-type σ , notée $X \sim \mathcal{N}(m; \sigma)$, si sa densité est

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}$$

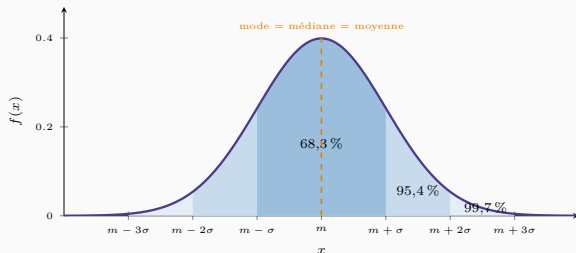
Cette formule ne sert jamais

On ne l'intègre pas — elle n'a pas de primitive élémentaire. On ne la calcule pas à la main.

Retenez-en une seule chose : elle ne dépend de x qu'à travers $\frac{x-m}{\sigma}$. C'est ce qui rendra une table unique suffisante.

La courbe et ses trois nombres

la loi normale et la règle des trois écarts-types



Pourquoi elle est partout. Non par commodité, mais par théorème : dès qu'une grandeur résulte de l'addition de nombreuses causes indépendantes et de faible poids individuel, sa loi tend vers la normale. C'est le théorème central limite — et c'est le sujet du chapitre 21.

Deux paramètres, et c'est tout

m place la cloche, σ règle sa largeur. La forme, elle, ne change jamais.

Symétrique autour de m : le coefficient d'asymétrie est nul, l'aplatissement vaut 3.

Les trois nombres à savoir par cœur
68 — 95 — 99,7.

Ils permettent de juger n'importe quel écart de tête, sans table ni calculatrice.

Cinq faits

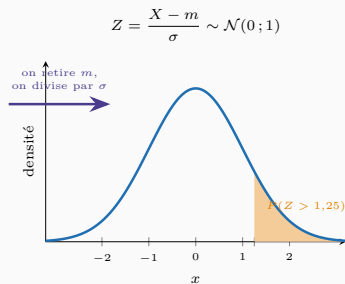
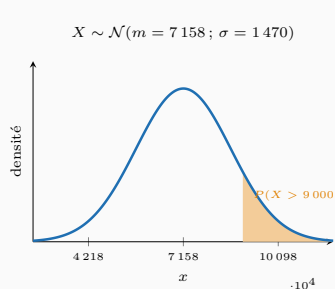
1. Elle est **symétrique** autour de m : mode, médiane et moyenne coïncident.
2. m la **déplace**, σ la **resserre ou l'étale**; la forme reste la même.
3. Ses points d'inflexion sont en $m - \sigma$ et $m + \sigma$.
4. Son coefficient d'asymétrie est nul, son aplatissement vaut 3.
5. Elle est définie sur $\mathbb{R}$ tout entier, mais l'essentiel de son aire tient dans $[m - 3\sigma ; m + 3\sigma]$.

La règle des trois écarts-types

68 % — 95 % — 99,7 %

Ces trois nombres, appris une fois, permettent de juger n'importe quel écart de tête. Un résultat à plus de trois écarts-types de la moyenne n'arrive que trois fois sur mille : **c'est presque toujours le signe d'une anomalie, non du hasard.**

Une seule table pour toutes



Les deux aires colorées sont égales. Centrer et réduire ne déforme rien : cela ne fait que changer d'unité de mesure.

$$\frac{9\,000 - 7\,158}{1\,470} = 1,25 \quad \text{et la table donne } P(Z > 1,25) = 1 - 0,8944 = 0,1056.$$

Une seule table suffit donc pour toutes les lois normales du monde
— et z n'est rien d'autre que la *cote* z de la statistique descriptive.

Centrer et réduire

Si $X \sim \mathcal{N}(m; \sigma)$, alors

$$Z = \frac{X - m}{\sigma} \sim \mathcal{N}(0; 1)$$

$\mathcal{N}(0; 1)$ est la loi normale **centrée réduite**, et sa fonction de répartition se note Π .

Pourquoi cela fonctionne

Retirer m **translate** la courbe sans la déformer. Diviser par σ **change l'unité de mesure**.

Aucune aire n'est modifiée : on a seulement remplacé les dirhams par des écarts-types. C'est un changement d'unité, pas un changement de loi.

Ce que z signifie

La cote z , déjà rencontrée

$$z = \frac{x - m}{\sigma} \quad \text{se lit : « à combien d'écarts-types de la moyenne »}$$

Un salaire, une note, un rendement

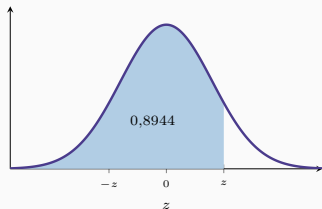
Un salaire de 9 000 dirhams dans une population de moyenne 7 158 et d'écart-type 1 470 :

$$z = \frac{9\,000 - 7\,158}{1\,470} = 1,25$$

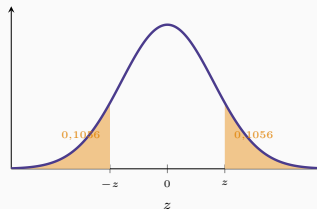
Ce salaire est à 1,25 écart-type au-dessus de la moyenne. Cette phrase est comparable d'une entreprise à l'autre, d'un pays à l'autre — le montant brut ne l'est pas.

Lire la table

ce que donne la table : $\Pi(z) = P(Z \leq z)$



la symétrie : $\Pi(-z) = 1 - \Pi(z)$



z	0,00	0,03	0,05	0,07
1,1	0,8643	0,8708	0,8749	0,8790
1,2	0,8849	0,8907	0,8944	0,8980
1,3	0,9032	0,9082	0,9115	0,9147

Comment on lit la table. La *ligne* donne le dixième, la *colonne* le centième : 1,2 et 0,05 se croisent sur 1,25.

La table ne donne que la gauche, et que le positif. Deux réflexes suffisent :

- à droite : $P(Z > z) = 1 - \Pi(z)$;
- négatif : $\Pi(-z) = 1 - \Pi(z)$;
- entre deux bornes : $\Pi(b) - \Pi(a)$.

Les trois valeurs qui reviennent dans tout le cursus : $z = 1,645$ (95 % à gauche), $z = 1,96$ (95 % au centre), $z = 2,33$ (99 % à gauche).

Tout se ramène à une seule lecture

Question	Formule
$P(Z \leq z)$	$\Pi(z)$, lecture directe
$P(Z > z)$	$1 - \Pi(z)$
$P(Z \leq -z)$	$1 - \Pi(z)$
$P(a \leq Z \leq b)$	$\Pi(b) - \Pi(a)$

Le calcul complet sur les salaires

$$P(X > 9\,000) = P(Z > 1,25) = 1 - \Pi(1,25) = 1 - 0,8944 = 0,1056$$

Environ 10,6 % des salariés dépassent 9 000 dirhams.

Les quantiles usuels

α	z tel que $\Pi(z) = \alpha$	Usage
0,90	1,282	
0,95	1,645	seuil unilatéral à 5 %
0,975	1,960	intervalle bilatéral à 95 %
0,99	2,326	Value at Risk à 99 %

Le 1,96 que vous verrez partout

$$P(-1,96 \leq Z \leq 1,96) = 0,95$$

Tout intervalle de confiance à 95 % contient ce nombre. Il vient d'ici, et de nulle part ailleurs.

La normale au travail

Quand n est grand

Si $X \sim \mathcal{B}(n; p)$ avec $np \geq 5$ et $n(1 - p) \geq 5$, alors

$$X \approx \mathcal{N}(np; \sqrt{np(1 - p)})$$

Le portefeuille du chapitre 15

$m = 8, \sigma = 2,77$. Probabilité d'au moins 14 défauts :

$$z = \frac{13,5 - 8}{2,77} = 1,99 \quad P(X \geq 14) \approx 1 - \Pi(1,99) = 0,0233$$

Valeur exacte : 0,0312. L'écart vient de ce que n n'est pas si grand, et p pas si proche de 0,5.

Pourquoi 13,5 et non 14

On remplace une loi discrète par une loi continue : les bâtons deviennent des rectangles de largeur 1, centrés sur les entiers.

Le rectangle du 14 commence donc à 13,5. Oublier cette correction fausse systématiquement le résultat, d'autant plus que n est petit.

Une approximation qui a fait son temps

Cette technique servait à éviter des calculs impossibles à la main. Aujourd'hui, R et Python donnent la binomiale exacte en une instruction.

Elle reste néanmoins à connaître : elle explique *pourquoi* les grands portefeuilles se comportent normalement, et cette raison, elle, ne périmera pas.

La somme de normales est normale

Si $X \sim \mathcal{N}(m_1; \sigma_1)$ et $Y \sim \mathcal{N}(m_2; \sigma_2)$ sont **indépendantes**, alors

$$X + Y \sim \mathcal{N}\left(m_1 + m_2; \sqrt{\sigma_1^2 + \sigma_2^2}\right)$$

Ce que cela permet en finance

Un portefeuille est une somme de positions. Si les rendements sont normaux, **le rendement du portefeuille l'est aussi** — et son risque se calcule sans simulation.

C'est la raison technique pour laquelle l'hypothèse de normalité s'est imposée en gestion d'actifs. Ses limites — les rendements réels ont des queues plus épaisses — ont fait tout le sujet des cours de risques financiers.

Ce qu'il faut retenir

Sept points

1. $\mathcal{N}(m; \sigma)$ est symétrique, en cloche, entièrement déterminée par deux paramètres.
2. 68 %, 95 %, 99,7 % : la règle des trois écarts-types.
3. $Z = \frac{X - m}{\sigma}$ ramène toute normale à $\mathcal{N}(0; 1)$: **une seule table suffit**.
4. z se lit « à combien d'écarts-types de la moyenne » — c'est la cote z de la descriptive.
5. La table ne donne que $\Pi(z)$; symétrie et complément fournissent tout le reste.
6. 1,645, 1,96, 2,326 : les trois quantiles à retenir.
7. Une somme de normales indépendantes est normale.

La suite

Trois lois se déduisent de la normale par des opérations simples — carrés et rapports. Elles ne servent qu'aux tests statistiques, mais elles y sont indispensables.

Le chapitre 19 les présente, puis dresse la **carte complète** des lois du cours.