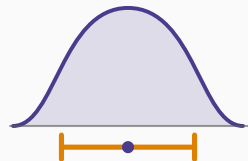


Statistique inférentielle

Chapitre 9 : Le principe de l'intervalle de confiance



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Pourquoi un intervalle

Ce que la confiance signifie

Le compromis fondamental

Ce qu'il faut retenir

Pourquoi un intervalle

Un nombre seul est une information incomplète

L'estimation ponctuelle est presque sûrement fausse

$\bar{x} = 7\,005,80$. La probabilité que ce nombre soit exactement m est nulle : $\bar{X}$ est une variable continue.

Annoncer une estimation sans indiquer sa précision, c'est donc communiquer une valeur dont on sait qu'elle est fausse, sans dire de combien.

Deux annonces, deux décisions

« Le salaire moyen est de 7 006 dirhams. »

« Le salaire moyen est compris entre 6 366 et 7 646 dirhams, avec une confiance de 95 %. »

La seconde permet de savoir si l'écart avec un concurrent à 7 200 est significatif. La première ne permet rien.

Intervalle de confiance

Un **intervalle de confiance** au niveau $1 - \alpha$ est un intervalle, calculé à partir de l'échantillon, qui contient le paramètre inconnu avec une probabilité $1 - \alpha$ **avant le tirage**.

α est le **risque**; $1 - \alpha$ est le **niveau de confiance**.

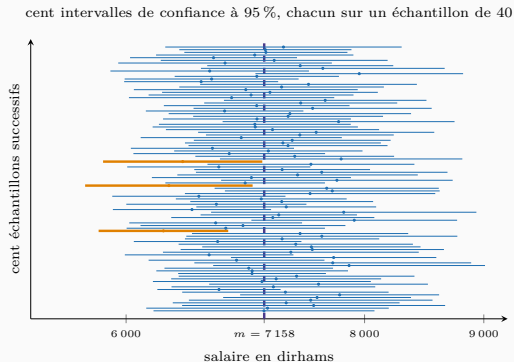
Les bornes sont aléatoires, pas le paramètre

Le paramètre m est fixe : il ne se promène pas.

Ce sont les **bornes** qui varient d'un échantillon à l'autre. L'intervalle est un filet mobile lancé vers une cible immobile.

Ce que la confiance signifie

Cent intervalles, cent échantillons



Ce que « 95 % » veut dire

Sur ces cent intervalles, **3** ratent la vraie moyenne. On en attendait cinq.

La confiance porte sur la méthode, pas sur l'intervalle.

Une fois l'échantillon tiré, $[6\,366 ; 7\,646]$ contient m ou ne le contient pas : il n'y a plus de hasard, donc plus de probabilité.

La phrase juste : « si je répétais l'opération, 95 fois sur cent mon intervalle contiendrait m ».

La phrase fausse, et la plus fréquente : « il y a 95 % de chances que m soit dans cet intervalle-là ».

Deux formulations

Correcte : « si je répétais l'opération, 95 intervalles sur cent contiendraient m . »

Incorrecte : « il y a 95 % de chances que m soit dans $[6\,366 ; 7\,646]$. »

Pourquoi la seconde est fausse

Une fois l'échantillon tiré, il n'y a plus de hasard. m vaut 7 158,26 : il est dans l'intervalle, ou il n'y est pas. La probabilité vaut 1 ou 0, et l'on ne sait simplement pas laquelle.

La confiance porte sur la méthode, pas sur le résultat particulier. C'est la même distinction qu'entre un estimateur et une estimation.

Ce qu'il faut dire en réunion

« Notre procédure d'enquête se trompe une fois sur vingt. Avec cet échantillon, elle donne [6 366 ; 7 646]. »

C'est rigoureux, compréhensible, et cela porte la responsabilité au bon endroit : sur la **méthode retenue**, qui est un choix de gestion, et non sur le nombre sorti.

Une seconde façon de le dire

Toutes les valeurs de l'intervalle sont **compatibles** avec les données observées, au seuil retenu. Celles qui sont dehors ne le sont pas.

Cette lecture prépare directement les tests d'hypothèse : le chapitre 16 montrera qu'un intervalle de confiance et un test bilatéral sont **le même objet**.

Le compromis fondamental

Les trois niveaux usuels

Confiance	Quantile	Demi-largeur (nos données)
90 %	1,645	± 523
95 %	1,960	± 624
99 %	2,576	± 820

On ne peut pas gagner sur les deux tableaux

Vouloir se tromper moins souvent oblige à élargir le filet.

À la limite, un intervalle de confiance 100 % serait $] -\infty ; +\infty[$: jamais faux, jamais utile. **Un intervalle très sûr et très étroit n'existe pas** — sauf en augmentant n .

Ce qui fait la largeur d'un intervalle

$$\text{demi-largeur} = \underbrace{z_{\alpha/2}}_{\text{confiance}} \times \underbrace{\frac{\sigma}{\sqrt{n}}}_{\text{dispersion et taille}}$$

Trois leviers : le **niveau de confiance**, la **dispersion** de la population, la **taille** de l'échantillon.

Un seul est à la main du gestionnaire

La dispersion σ est une donnée du problème : on ne la choisit pas.

Le niveau de confiance est une convention à laquelle personne ne dérogera. **Reste n — et c'est donc toujours par là que passe une décision de précision.** Le chapitre 12 en fera un calcul.

Ce qu'il faut retenir

Six points

1. Une estimation ponctuelle sans précision est une information incomplète.
2. Un **intervalle de confiance** contient le paramètre avec une probabilité $1 - \alpha$ **avant le tirage**.
3. Ce sont les **bornes** qui sont aléatoires; le paramètre est fixe.
4. La confiance porte sur la **méthode**, jamais sur l'intervalle particulier obtenu.
5. Plus de confiance signifie un intervalle plus large : les deux ne s'améliorent pas ensemble.
6. Trois leviers — confiance, dispersion, taille — dont un seul se décide : n .

La suite

Le principe est posé. Le chapitre 10 le met en formule pour le paramètre le plus courant, la moyenne — avec une difficulté qu'il faudra traiter : σ est inconnu.