UNIVERSITÉ de Picardie Jules Verne

Biostatistiques
Partie 2

diane.zarzoso.lacoste@u-picardie.fr

V22-23

Plan du cours

Partie 2. Introductions aux statistiques

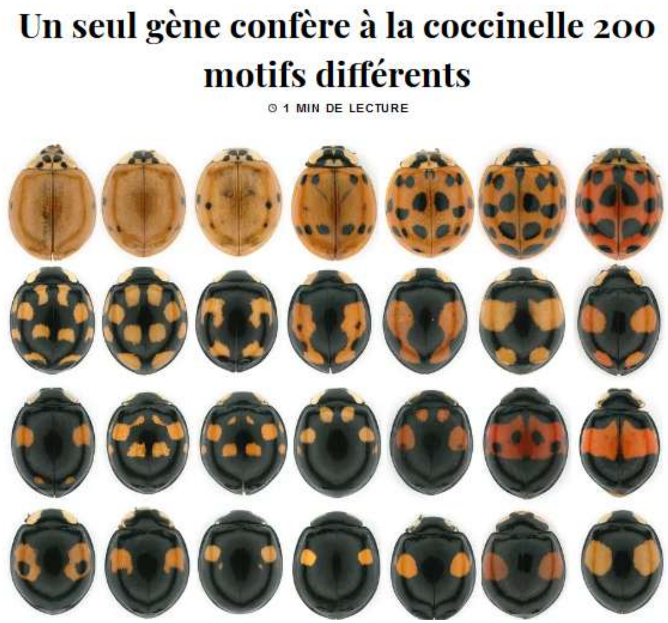
1. De la variabilité dans les sciences de la vie

2. Statistiques descriptives

3. Vers les statistiques inférentielles

- Notion d’estimation
- Intervalles de confiance
- Introduction aux tests statistiques
- S’orienter dans les tests statistiques

De la variabilité dans les sciences de la vie



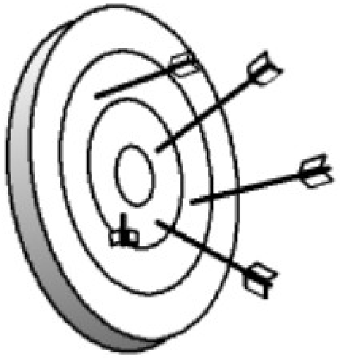
Coccinelle arlequin

Même lorsqu'on les estime biologiquement équivalents,

2 individus ne seront jamais parfaitement identiques

→ Variabilité interindividuelle

(sous de multiples influences : facteurs environnementaux locaux, histoire, comportements, (épi)génétique, …)



Même si on répète une mesure avec un appareil fiable et dans les mêmes conditions, on n'obtiendra jamais des résultats parfaitement identiques

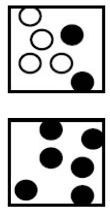
→ Erreur de mesure (aléatoire)

Fluctuation d’échantillonnage

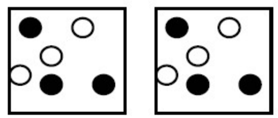
Tous les individus biologiques étant différents, il n’y aura jamais 2 échantillons identiques !

Ex : échantillonnage de 6 souris avec 2 phénotypes possibles A ou B.

A: 2 échantillons, même très différents, ne proviennent pas nécessairement de deux populations différentes.



B : 2 échantillons, même très semblables, ne proviennent pas nécessairement de deux populations semblables.



Alors, comment s'en sort-on ??

Quand on **compare** 2 moyennes / proportions issus d'une expérience de labo ou d'observations de terrain, il y aura **toujours une différence** entre elles = **Variabilité**

- **1 part** de cette variabilité (voire la totalité!) sera inévitablement due au **hasard** (fluctuations d'échantillonnage) = **variabilité résiduelle**

- éventuellement 1 **part** de cette variabilité sera due à un **effet réel de VI**
  → celle que l'on cherche à **expliquer** / mettre évidence dans étude

➔ Indispensable d'évaluer la **fiabilité** de ces moyennes / proportions en calculant leur **intervalle de confiance** (IC)

➔ Puis calculer la **probabilité** que le hasard à lui seule puisse expliquer une **différence** au moins aussi grande que celle **observée** entre échantillons ou avec une référence (= p-value)

Statistiques = seul moyen d'effectuer ces vérifications de façon objective et reproductible

Plan du cours

Partie 2. Introductions aux statistiques

1. De la variabilité dans les sciences de la vie

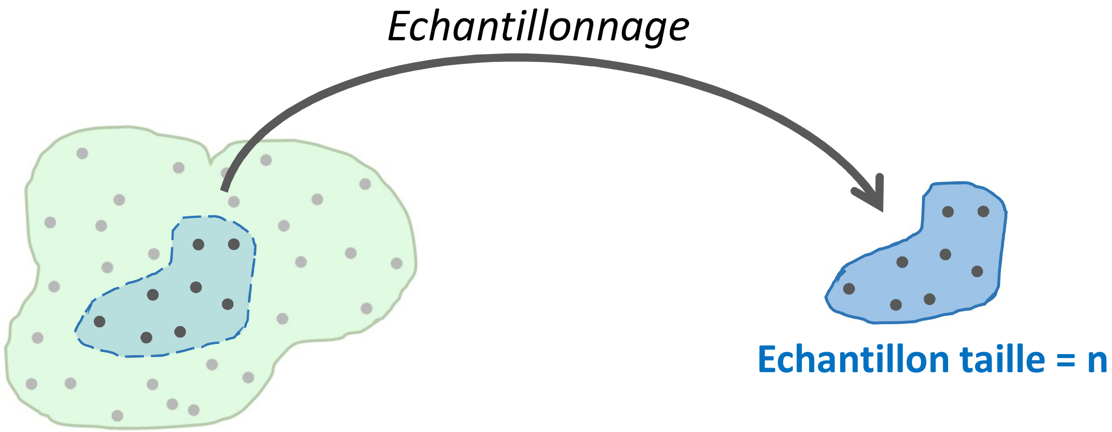
2. Statistiques descriptives

3. Vers les statistiques inférentielles

- Notion d’estimation
- Intervalles de confiance
- Introduction aux tests statistiques
- S’orienter dans les tests statistiques

Statistiques descriptives

*J'ai tiré aléatoirement un échantillon de taille **n** d'1 population statistique, et mesuré une variable aléatoire (X) sur les individus statistiques….*

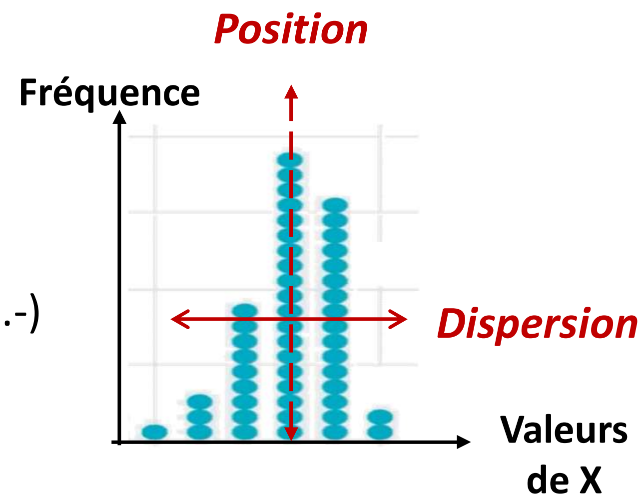


Et maintenant, je fais quoi ??

Statistiques descriptives

1) Représenter graphiquement la distribution de la VA :

Répartition de la fréquence des valeurs prises par une VA



- Forme? Symétrie? (loi de probabilité ? -Normale? Poisson? Binomiale?..-)
- Présence de points aberrants (outliers), erreurs de saisie ?

2) Résumer les statistiques (= caractéristiques, indicateurs numériques) de l'échantillon:

- **Position :** Quelles sont la ou les valeurs centrales autour desquelles se regroupent les données ?

  Moyenne ? Médiane ? Mode ?

- **Dispersion :** Quel est l'étalement des données autour des valeurs centrales.

  Variance ? Ecart type ?

Statistiques descriptives

Indicateurs de position :

- **Mode** : Valeur de X la plus fréquente (Mo)

- **Médiane** : Valeur de X telle que 50% des observations lui sont ≤, et 50% ≥ (Me)

  → Nécessite de classer les valeurs par ordre croissant, puis :

  - Si n impaire: $Me = x_{(n+1)/2}$ *(valeur pour le rang (n+1)/2)*

  ex: 50.4 58.8 ***61.1*** *63.4 63.6 → n = 5 et n+1/2 = 3 →* Me *= 3*^{e} *valeur =* ***61.1***

  - Si n paire : $Me = (x_{(n/2)} + x_{(n/2+1)}) / 2$ *(moyenne des 2 valeurs centrales)*

  ex: 50.4 ***58.8 61.1*** *63.4 → n = 4, n/2=2, n/2+1 = 3 →* Me *= moy. 2*^{e} *et 3*^{e} *valeur =* ***(58.8 + 61.1)/2 = 62.25***

- **Moyenne arithmétique :** Somme des valeurs de X divisée par l'effectif total ($Moy = \bar{x}$)

$$\bar{x} = \frac{\sum xi}{n}$$

Statistiques descriptives

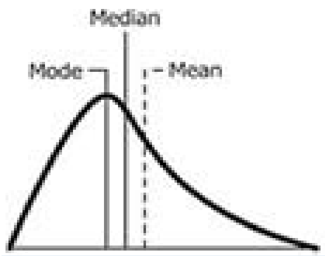
Lien avec la forme de la distribution des données (skweness) :

- **Symétrique :** valeurs également répartie de part et d'autre de la moyenne
  → $Mo = Me = Moy$ *(ex : loi Normale, Student)*

- **Asymétrie positive :** les valeurs les + fréquentes sont basses, queue vers valeurs élevées (à Droite)
  → $Mo < Me < Moy$ *(Normalisation :* $\sqrt{x}$, $\log_{10}(x)$, $1/x$ *)*

- ***Asymétrie négative :*** les valeurs les + fréquentes sont élevées, queue vers valeurs basses (à Gauche)
  → $Mo > Me > Moy$ *(Normalisation :* $\sqrt{K-x}$, $\log_{10}(K-x)$, $\frac{1}{K-x}$, avec K (constante) $= max(x+1)$)

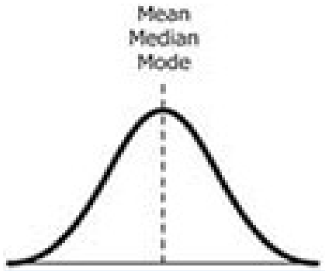
Asymétrie positive (queue à droite)



$Mo < Me <$ Moy

Normalisation

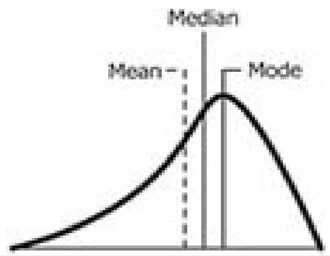
Symétrique



$Mo = Me =$ Moy

Normalisation

Asymétrie négative (queue à gauche)



$Mo > Me >$ Moy

$\boldsymbol{Transformation}$:
(asym. modérée à sévère)

$\sqrt{x}$, $\log_{10}(x)$, $1/x$

$\sqrt{K-x}$, $\log_{10}(K-x)$, $1/(K-x)$

Statistiques descriptives

Indicateurs de dispersion :

- **Etendue** : Différence entre les 2 valeurs les plus extrêmes.

$$E = max - min$$

- **Variance** : moyenne des carrés des écarts à la moyenne.

$$S^2 = \frac{\sum(xi - \bar{x})^2}{n}$$

- Toujours positive
- Plus elle est petite, plus les valeurs de la VA sont similaires et proches de $\bar{x}$

- **Ecart type** : racine carrée de la variance.

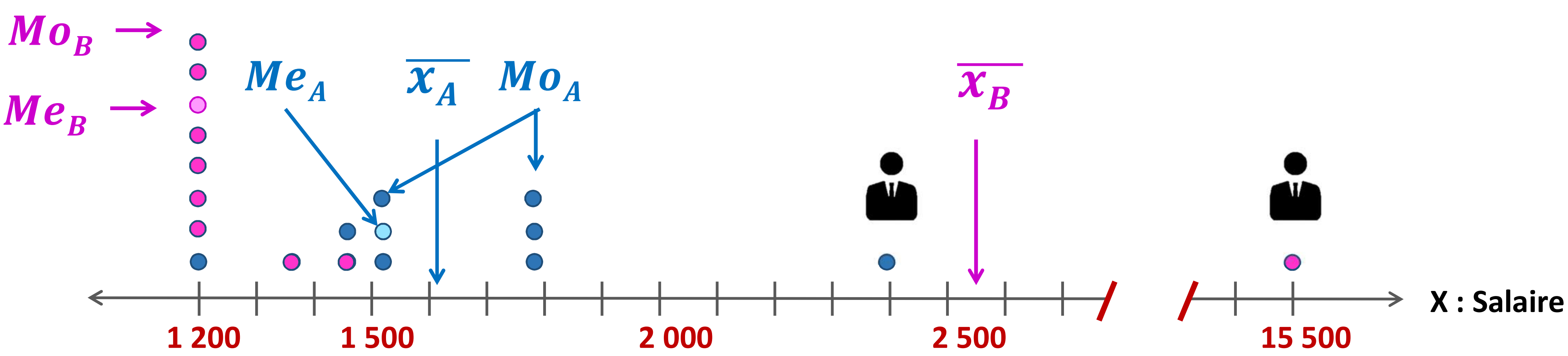
$$S = \sqrt{(S^2)}$$

- Même ordre de grandeur (unité) que la moyenne
- L'écart qu'on observera en moyenne entre une donnée prise au hasard dans l'échantillon et la moyenne des données.

Dans 2 entreprises concurrentes A et B, de chacune 11 personnes, le salaire mensuel net moyen est respectivement de 1611,8 et 2539,3 €.

1. Dans laquelle préféreriez-vous travailler?

A. Entreprise A ✓ **A** **B. Entreprise B**

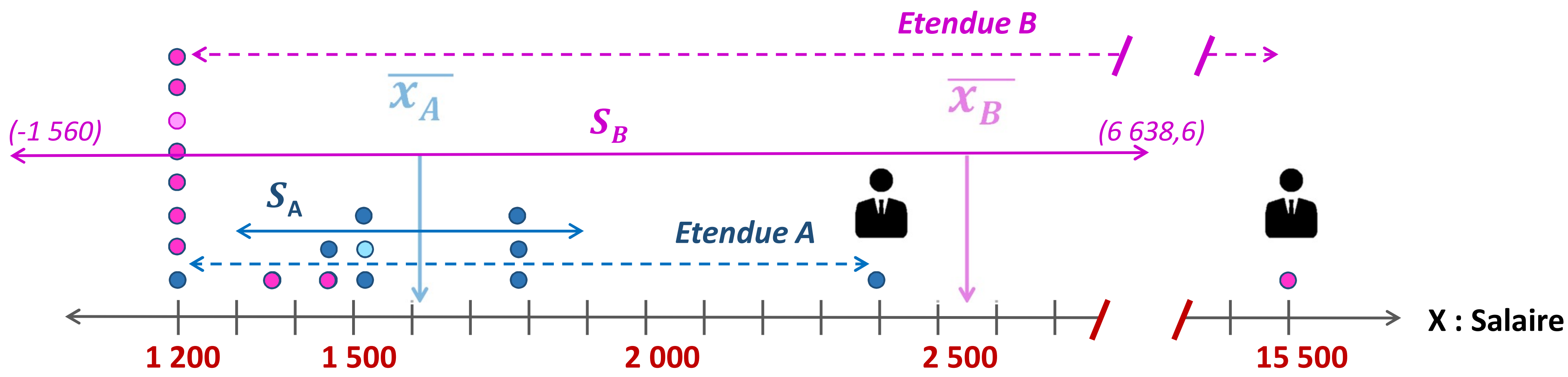


x_A = employés : 1 203, 1 356, 1 452, 1 452, 1 512, 1 512, 1512, 1 777, 1 777, 1 777, **patron :** 2 400

x_B = employés : 1 203, 1 203, 1 203, 1 203, 1 203, 1 203, 1 203, 1 203, 1 356, 1 452, **patron :** 15 500

2. Dans quelle entreprise la variance est-elle la plus grande ?

A. Entreprise A **B. Entreprise B** ✓ **B**



	Variance S^2	Ecart type S	Etendue E
Entreprise A	92 421,4	304,0	1 197
Entreprise B	16 804 340,6	4 099,3	14 294

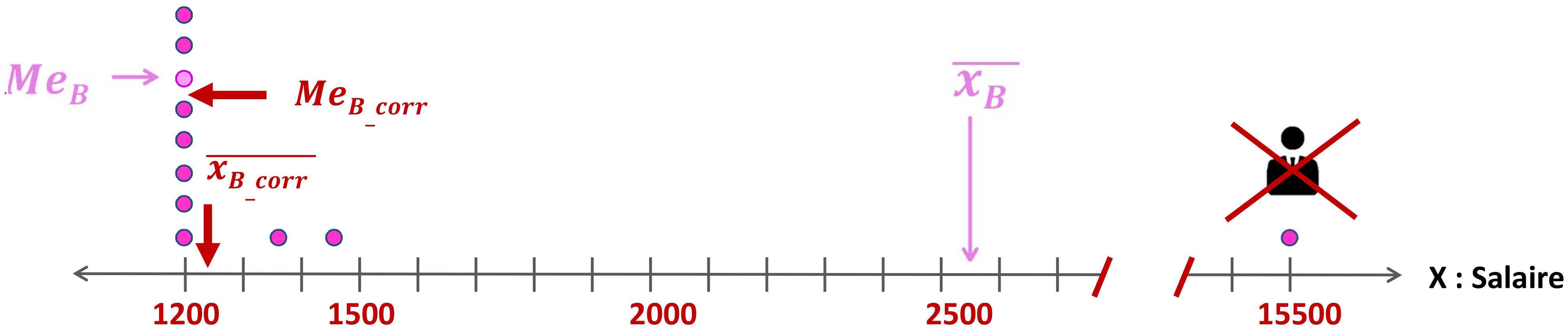
3. Quel serait l'effet de la démission du patron B sur le salaire moyen et médian de son entreprise ?

A. $\overline{x_B}$ diminue et Me_B augmente

B. Me_B diminue et $\overline{x_B}$ augmente

C. Les 2 diminuent, mais Me_B davantage

D. Les 2 diminuent, mais $\overline{x_B}$ davantage ✓ D

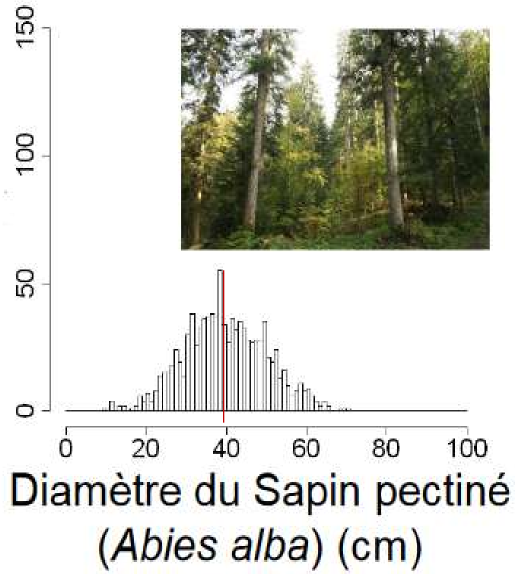


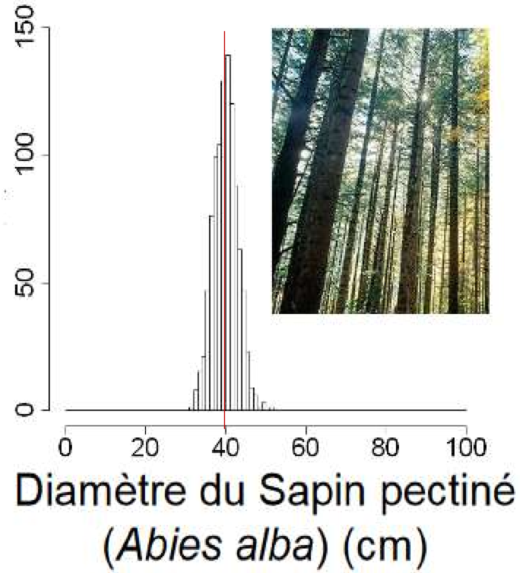
x_B = **employés :** 1 203, 1 203, 1 203, 1 203, 1 203, 1 203, 1 203, 1 203, 1 356, 1 452, ~~**patron :** 15 500~~

A retenir de cet exemple : Une moyenne ne suffit pas !!!

- $\bar{x}$ = indicateur de position le plus utilisé, mais pas toujours le plus pertinent (cf. *Me*)
- $\bar{x}$ est plus sensible aux valeurs extrêmes que ***Me***
- $\bar{x}$ est très peu informative si fournie sans indication de dispersion des données (S^2, S)

Deux échantillons peuvent avoir la même moyenne mais des variances très différentes !





Plan du cours

Partie 2. Introductions aux statistiques

1. De la variabilité dans les sciences de la vie

2. Statistiques descriptives

3. Vers les statistiques inférentielles

- **Notion d'estimation**
- Intervalles de confiance
- Introduction aux tests statistiques
- S'orienter dans les tests statistiques

Notion d'estimation

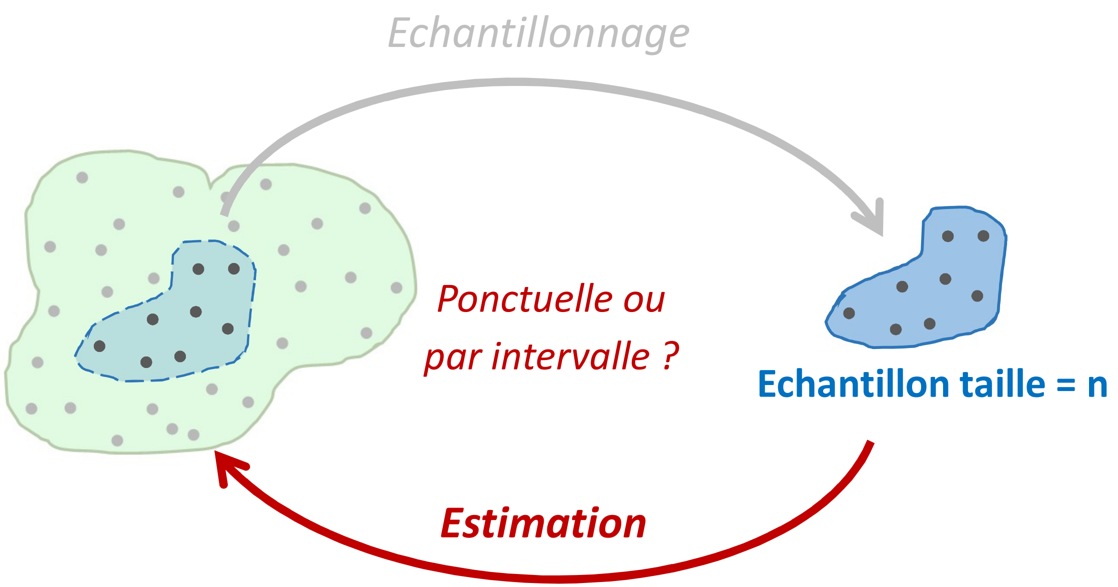
J'ai calculé les statistiques de mon échantillon et veux maintenant ***estimer*** *les paramètres de la population statistique….*

Paramètres (inconnus) :

μ ?

σ^2 ?

σ ?



Statistiques (observées) :

$\overline{x}$

S^2

S

Estimation ponctuelle de paramètre de position

- **Moyenne arithmétique :**

inconnue

$\mu = \cancel{(\sum x_i)\ /N}$ si inconnue, **estimée** par $\bar{x} = (\sum x_i)\ /n$

pour une population **Estimation** *pour un échantillon*

Notion « distribution d'échantillonnage $\bar{M}$ » : estimateur sans biais de μ *(Licence)*

La loi des Grands Nombres : $\bar{M} \rightarrow \mu$ quand **n** $\rightarrow$ **+∞**

→ *Plus la taille **n** de l'échantillon sera grande, plus la moyenne calculée sur cet échantillon tendra vers la véritable moyenne μ de la population.*

En pratique on considère 1 échantillon comme grand si n ≥ 30

Estimation ponctuelle de paramètre de dispersion

- **Variance** :

inconnue

$$\sigma^2 = \cancel{\frac{\sum(xi - \mu)^2}{N}} \xleftarrow{\text{Estimation}} S^2 = \frac{\sum(xi - \bar{x})^2}{n}$$

population *échantillon*

Variance estimée/corrigée S^{*2} (ou $\hat{\sigma}^2$) *:* $$S^{*2} = \frac{\sum(Xi - \bar{X})^2}{n - 1}$$

→ Sans cette **correction**, S^2 sous-estime systématiquement σ^2 (devient négligeable pour n>30).
→ **NB:** *R fait systématiquement cette correction.*

- **Ecart type :**

inconnue

$$\sigma = \cancel{\sqrt{(\sigma^2)}} \xleftarrow{\text{Estimation}} S = \sqrt{(S^2)}$$

population *échantillon*

Ecart type estimé/corrigé S^* (parfois notée $\hat{\sigma}$): $$S^* = \sqrt{(S^{*2})}$$

En bref :

Population (N)

Echantillon (n)

Dispersion

Paramètres : souvent inconnues

Variance σ²

$$\frac{1}{N}\sum_{i=1}^{N}(xi-\mu)^2$$

Ecart type σ

$$\sqrt{\sigma^2}$$

$S^2 < S^{*2} \approx \sigma^2$

Vrai si n > 30

$$S^{*2} = \frac{1}{n-1}\sum_{i=1}^{n}(xi-\bar{x})^2$$

correction

Statistiques :

Variance s²

$$\frac{1}{n}\sum_{i=1}^{n}(xi-\bar{x})^2$$

Ecart type s

$$\sqrt{s^2}$$

Estimation

Position

Moyenne μ

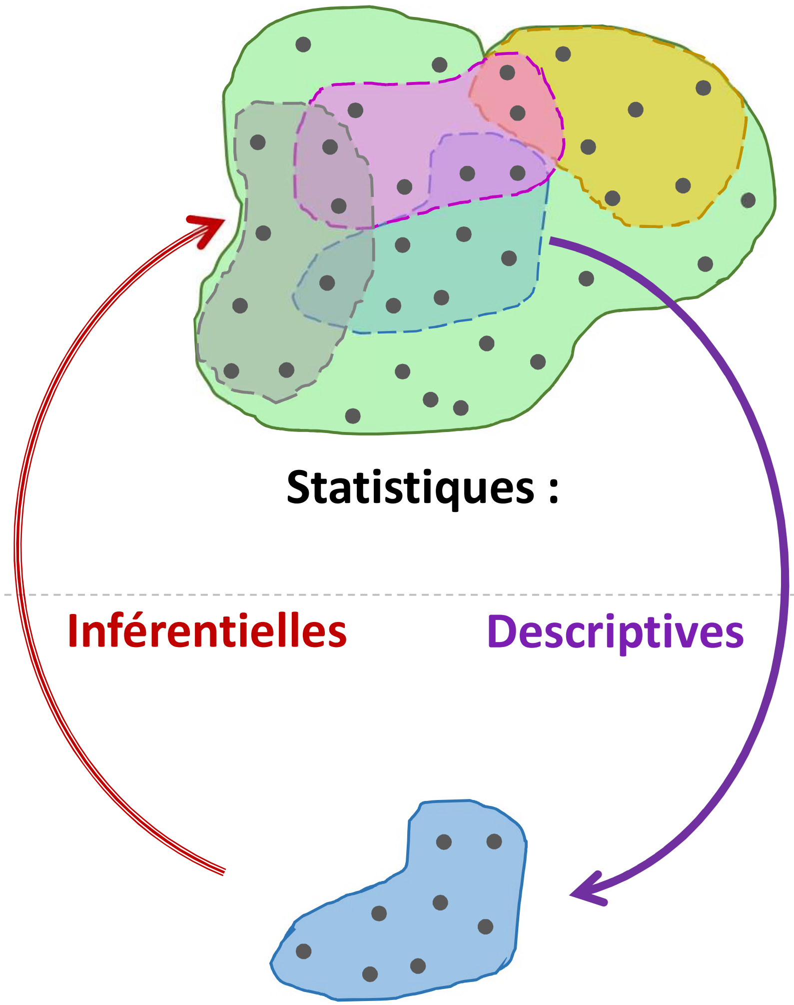
$$\frac{1}{N}\sum_{i=1}^{N} xi$$

$\bar{x} \approx \mu$

Moyenne $\bar{x}$

$$\frac{1}{n}\sum_{i=1}^{n} xi$$

Population taille = N



Echantillon taille = n

Plan du cours

Partie 2. Introductions aux statistiques

1. De la variabilité dans les sciences de la vie

2. Statistiques descriptives

3. Vers les statistiques inférentielles

- Notion d'estimation
- **Intervalles de confiance**
- Introduction aux tests statistiques
- S'orienter dans les tests statistiques

Intervalles de confiance

J'ai calculé les statistiques (paramètres) de mon échantillon, je m'en suis servi pour pouvoir estimer les paramètres (inconnus) de la population dont l'échantillon est issu (avec correction pour la variance)….

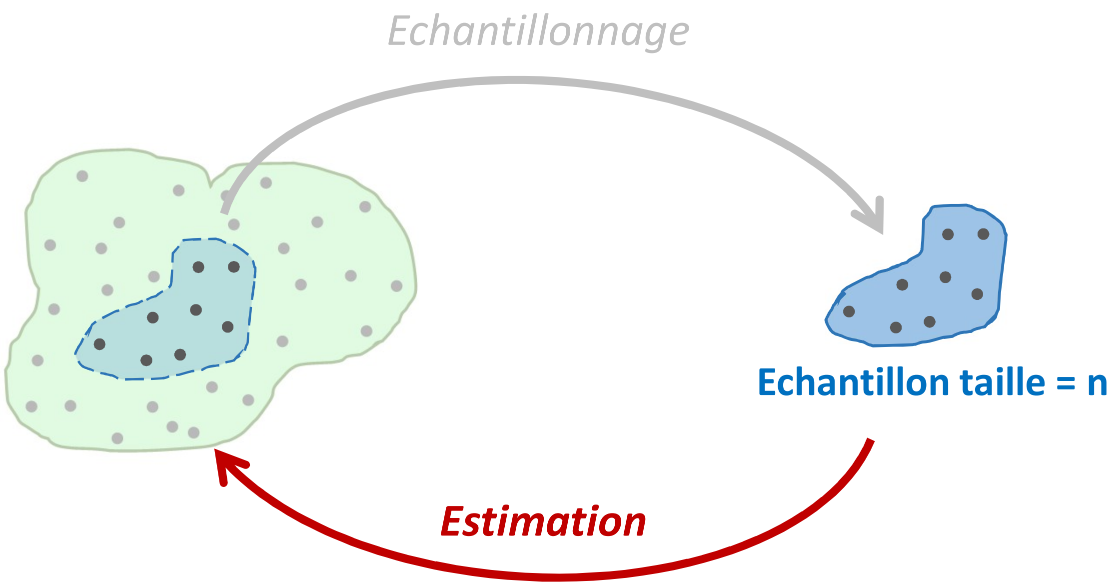
*Quelle est la **précision** de ces estimations ?* → *toujours les assortir d'**intervalles de confiance***

Paramètres (inconnus) :

$\mu \leftarrow \overline{x}$

$\sigma^2 \leftarrow S^{*2}$

$\sigma \leftarrow S^*$



Statistiques (observées) :

$\overline{x}$

S^2

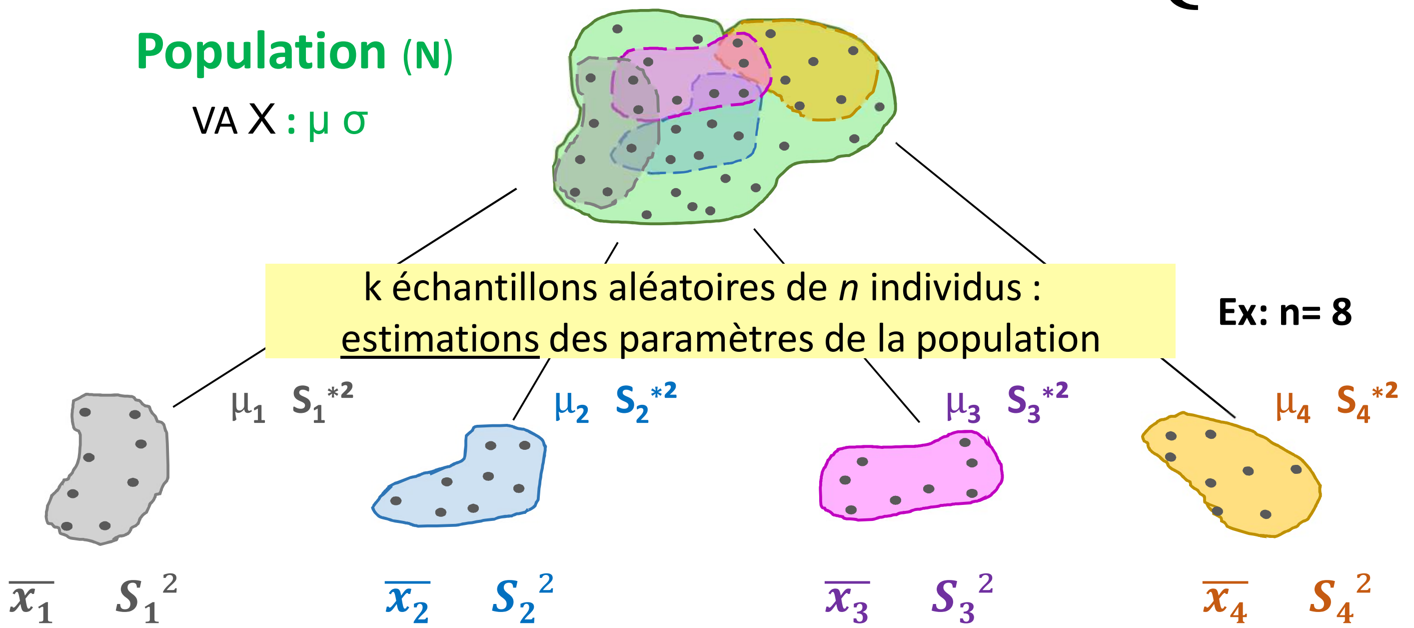
S

Notion de distribution d'échantillonnage d'un paramètre

1 grande variété d'échantillons de taille n (« n-échantillons ») peuvent être tirés au hasard (avec remise) de la même grande à population.

Si on note la série de moyennes $\overline{\boldsymbol{xi}}$ calculées sur ***k*** échantillons aléatoire de taille n

→ nouvelle VA M = **Distribution d'échantillonnage des moyennes** $M \left\{ \overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_4}, \ldots, \overline{x_k} \right\}$

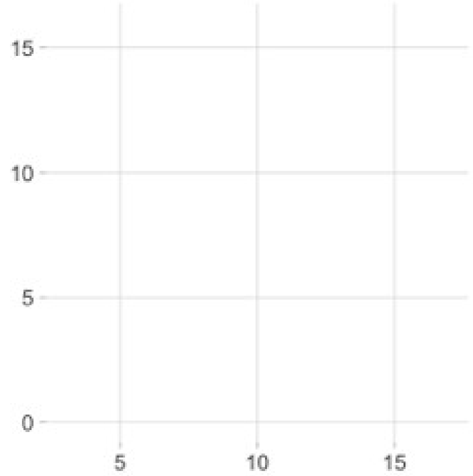


➔ La moyenne $\overline{\boldsymbol{M}}$ de M sera proche de μ.

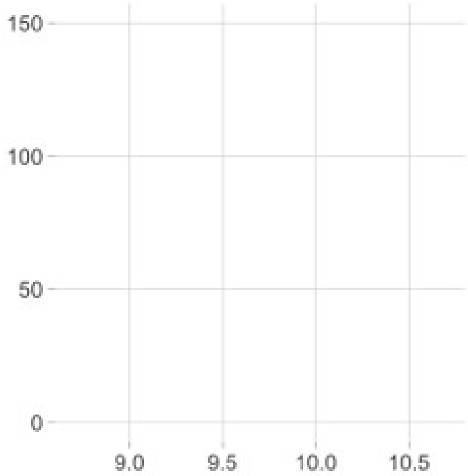
Distribution d’échantillonnage des moyennes et théorème central limite

La preuve en image…

n = 50



Distribution de la variable dans l’échantillon



Distribution d’échantillonnage de la moyenne

https://tinystats.github.io

Distribution d’échantillonnage des moyennes et théorème central limite

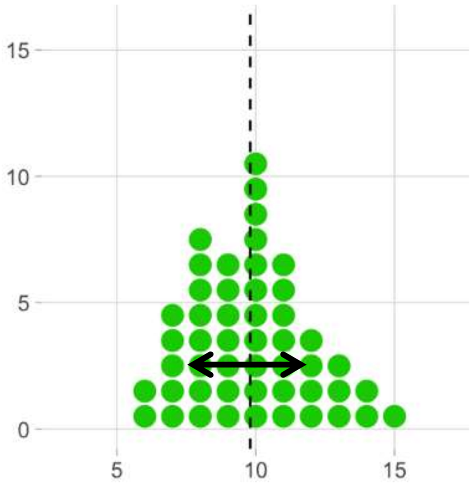
La preuve en image…

Moyenne de l’échantillon

$\overline{x_{1000}}$

n = 50

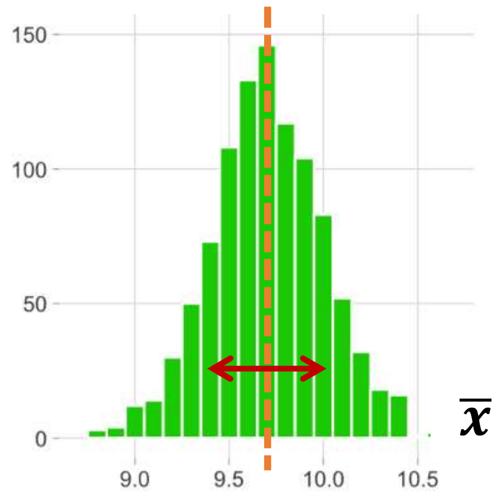
Ecart-type de l’échantillon S



$\overline{x}_{1000} = 9.8$

Moyenne des moyennes

$\bar{M}$ = μ



n = 1000

Erreur Standard de la moyenne

$\sigma/\sqrt{n}$

Distribution de la variable dans l’échantillon (quelconque)

Distribution d’échantillonnage de la moyenne (→ Normale)

https://tinystats.github.io

Distribution d'échantillonnage des moyennes et théorème central limite

Théorème Central Limite :

1. Quand on échantillonne un grand nombre de fois la population, la **distribution d'échantillonnage des moyennes** **(M)** tend vers une loi Normale $\sim N\ (\mu, \sigma/\sqrt{n})$,

et ce quelle que soit la distribution de la population d'origine!!

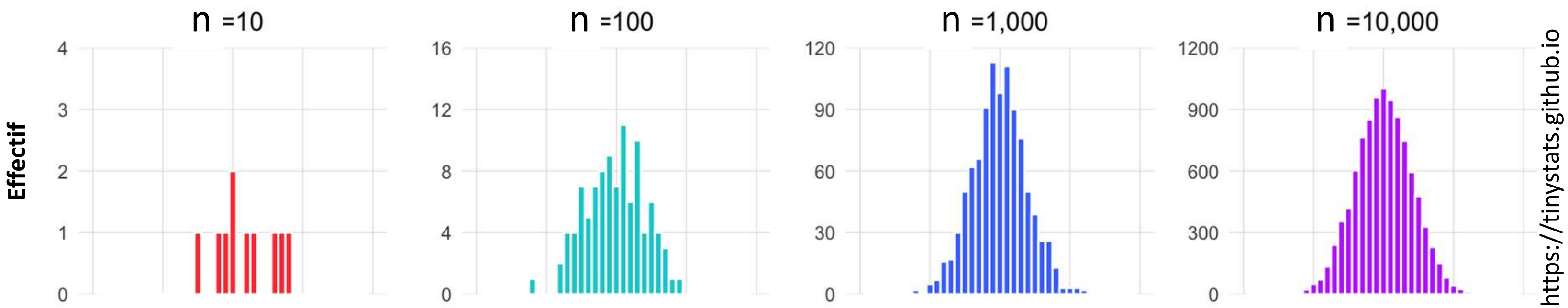
L' **erreur standard** $(\sqrt{\sigma^2/n} = \sigma/\sqrt{n}) =$ écart type des moyennes $\bar{x}$ autour de $\bar{M} (= \mu)$

sera *n* fois plus petit que l'écart type de X dans l'échantillon.

Distribution d’échantillonnage des moyennes et théorème central limite

Théorème Central Limite :

2. Cette tendance à la normalité est d’autant plus importante et resserrée autour de la moyenne réelle µ (faible erreur standard), que la taille n de l’échantillon est grande et tend vers l’infini (en pratique $n \geq 30$).



Très utile en statistique inférentielle → Loi Normale décrit les fluctuations des moyennes

(grands échantillons ou distribution N)

Rappels sur la loi Normale

Loi de probabilité la plus importante :

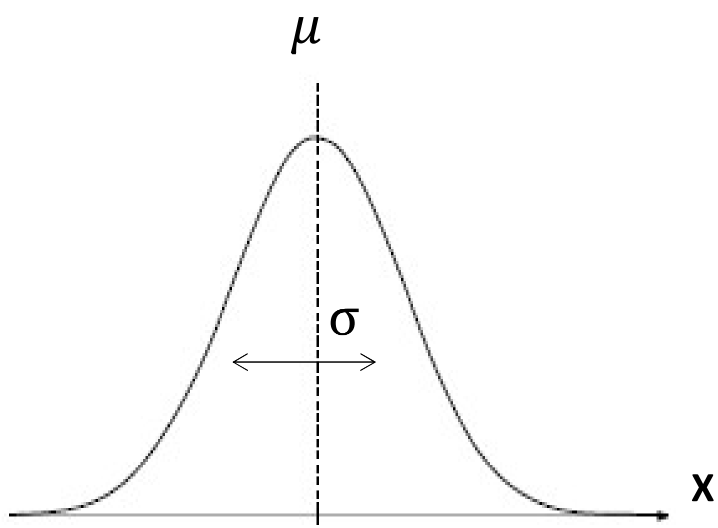
- S'applique à de nombreux phénomènes aléatoires naturels et expérimentaux.
- Raisons théorique et pratique (TCL, condition requise pour tests paramétriques).

Propriétés :

- Distribution/loi de proba « modèle » pour VA continues
- Fonction de densité = courbe en **cloche symétrique :**

- **Centrée** sur la moyenne $\boldsymbol{\mu}$ (moyenne = médiane =mode)
- Sa **largeur** dépend de l'écart type $\boldsymbol{\sigma}$

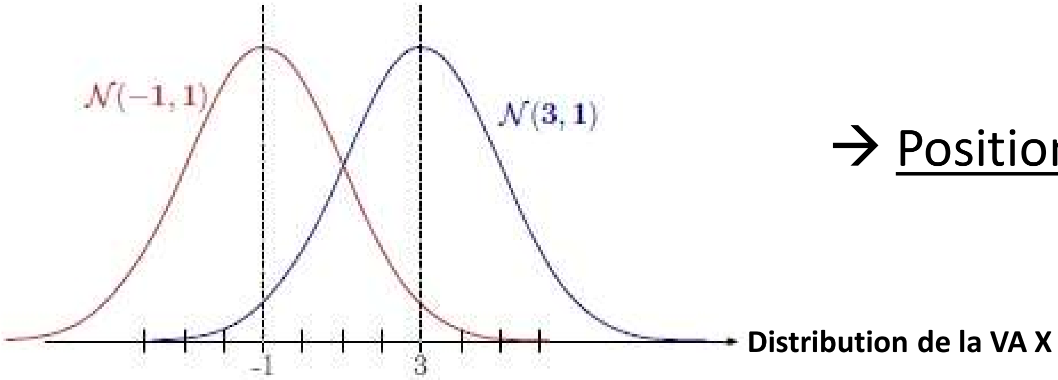
- Entièrement définie par ses 2 paramètres $N\ (\mu, \sigma)$
- Aire total sous la courbe = 1 (100% probabilité)



Rappels sur la loi Normale

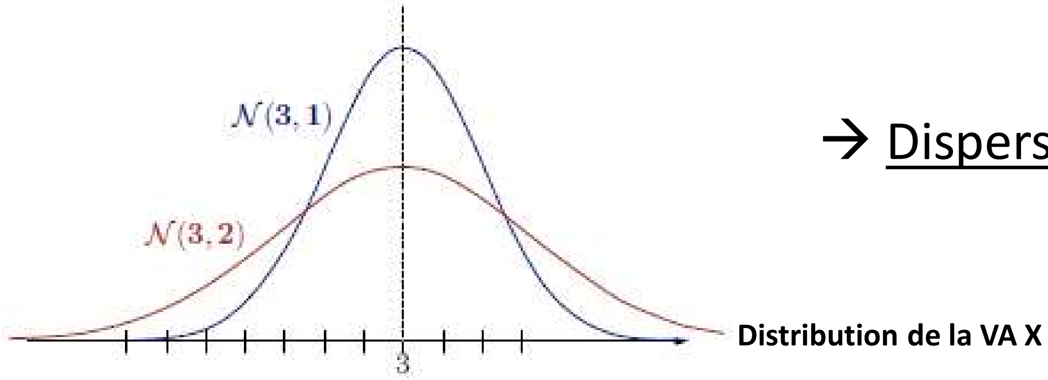
Les paramètres $N\,(\mu, \sigma)$:

Exemples de lois normales avec **moyennes différentes**, même écart-type :



→ Position relative courbes

Exemples de lois normales avec même moyenne, **écart-types différents** :



→ Dispersion relative courbes

Rappels sur la loi Normale

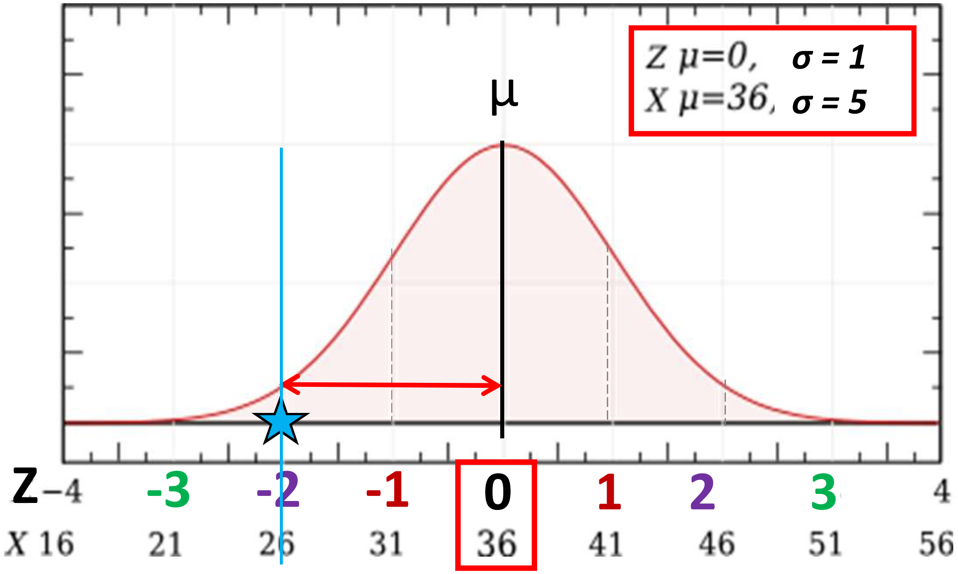
Pour faciliter l'étude et la comparaison de distributions *N* (quelles que soient les unités), on utilise la **Loi Normale Centrée Réduite LNCR** (cas particulier de loi N avec **μ = 0** et ***σ* = 1**).

Comment ? <u>Transformation</u> VA ***X*** ~ *N* (μ, σ) en nouvelle VA **Z** centrée-réduire ~ *N* $(0, 1)$.

$$Z_i = \frac{x_i - \mu_x}{\sigma_x}$$

Centrage

Réduction



Intérêt ? Permet mesurer « de combien d'écart types σ 1 valeur de la VA X s'écarte de la moyenne μ ».

Ex: si x_i *=26* ↔ Z_i *= -2*

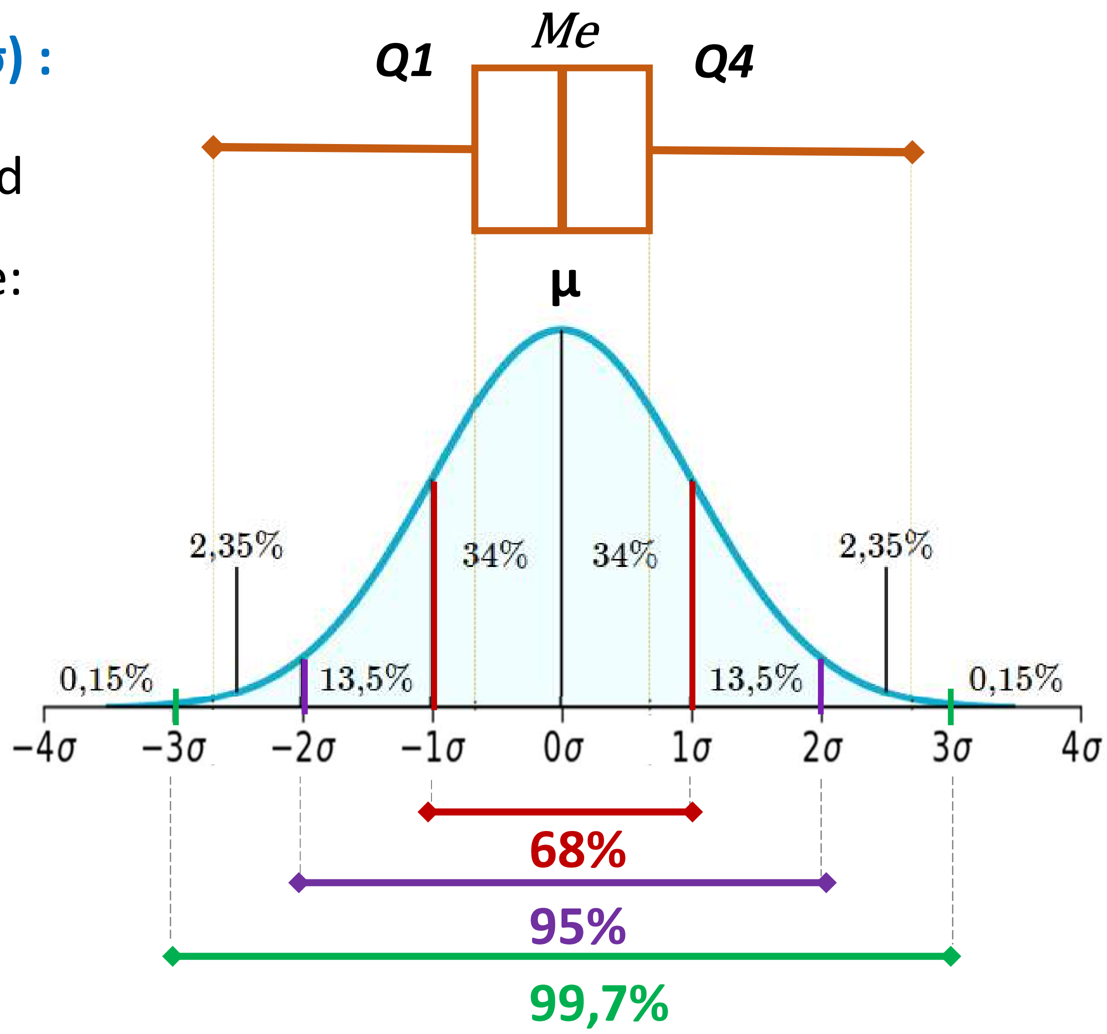
Puis utilise des tables de références pour estimer proba que X soit > ou < à 1 valeur donnée *(tables souvent unilatérales donnant l'aire sous courbe à droite ou à gauche d'1 valeur)*

Rappels sur la loi Normale

➔ Règle empirique d'une distribution N (μ, σ) :

Probabilité qu'1 observation tirée au hasard dans la population se trouve dans l'intervalle:

Intervalle autour de la moyenne :	**% des observations (= probabilités) :**
+/- 1 écart-type	≈ 68
+/- 2 écart-types	≈ 95
+/- 3 écart-types	≈ 99,7



A retenir pour estimation IC et tests d'hypothèses paramétriques.

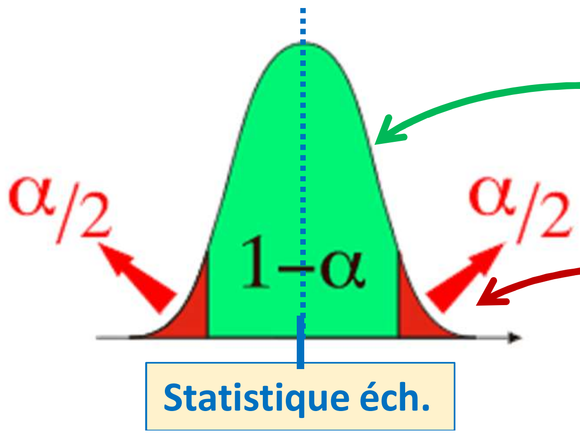
Intervalles de confiance

But : Donner la **précision de l'estimation** d'un paramètre inconnu de la population à partir d'un échantillon.

Comment ? Calculer **autour de chaque statistique observée** dans l'échantillon ($\overline{x}$, S^2, ...), un **intervalle contenant** très probablement la véritable valeur (inconnue) du **paramètre** que l'on étudie dans la grande **population** (μ, σ^2, ...), avec un **risque d'erreur consentie (α)**.

Interprétation : Plus cet intervalle est **petit**, meilleure (plus **précise**) est mon estimation.

EX: IC à 95% (0.95), avec un risque consenti d'erreur α = 5% (0.05) → on « accepte » de se tromper 5 fois sur 100

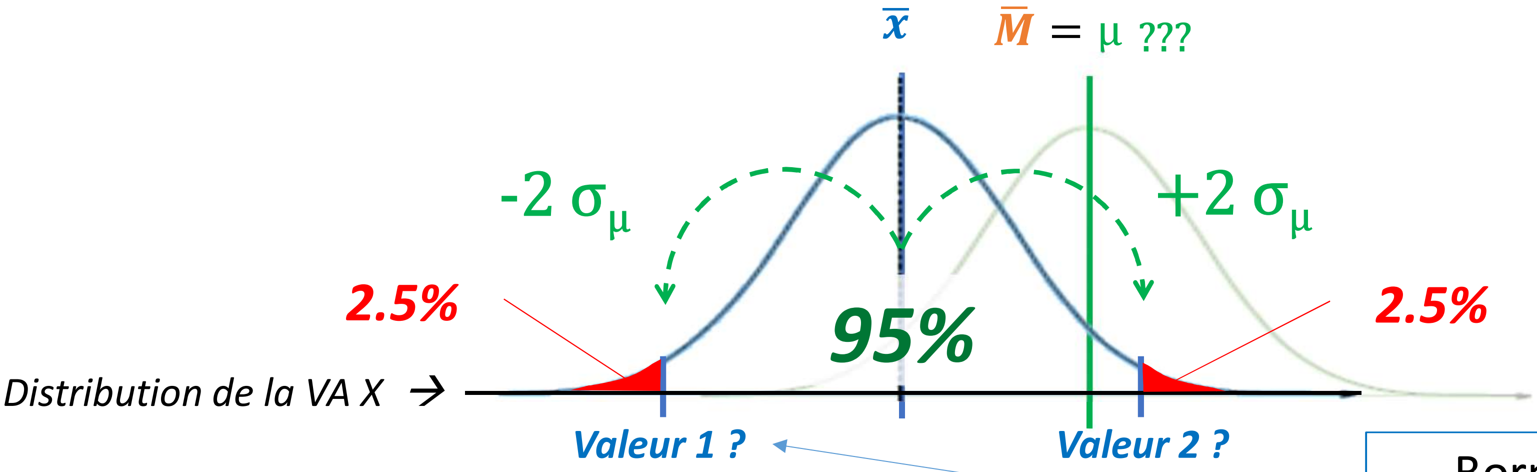


Le paramètre de la population a une probabilité de 95% (0.95) de se trouver dans cet intervalle (<u>niveau de confiance</u> = 1- α)…

…et 5% (<u>risque</u> α = 0.05) de chance de se trouver en dehors (donc 2.5% de chance d'un côté ou l'autre par symétrie)

Intervalles de confiance de la moyenne

Ex $IC_{0.95}[\mu]$ ***:*** L'idée est de fournir 1 **fourchette** de **valeurs** de la VA X, correspondant à 1 encadrement de **± 2 σ_μ** **autour** de $\bar{x}$ (observée sur l'éch.), de sorte à ce que 95 fois sur 100, elle **contienne** la vraie moyenne **μ** de la population (inconnue, mais qu'on cherche à estimer).



(avec σ_μ = $\sigma/\sqrt{n}$ = erreur standard)

Bornes de l'IC à fournir en complément de la moyenne $\bar{x}$:
$IC_{95\%}$=[valeur 1 ; valeur 2]

Conditions :

- l'**échantillon** aléatoire est grand (**n ≥ 30** → TCL)
- et/ou la **distribution de *X* ~N** (ou est au moins unimodale et symétrique)

Intervalles de confiance de la moyenne

- **Cas 1: Grand échantillon (n $\geq$ 30)** et distribution quelconque de la VA X (TCL):

$$IC_{1-\alpha}[\mu] = \bar{x} \pm Z_{\alpha} * \frac{S}{\sqrt{n}}$$

Erreur standard

Statistique Z suit LNCR
(proche de 2 pour IC à 95%)

- **Cas 2: Petit échantillon** (10 ≤ n ≤ 30) **et distribution unimodale symétrique ou** très petit échantillon (n < 10) et distribution **normale** :

$$IC_{1-\alpha}[\mu] = \bar{x} \pm t_{n-1,\alpha} * \frac{S}{\sqrt{n}}$$

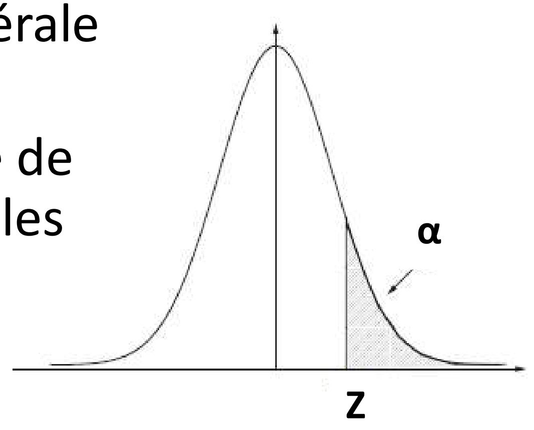
Statistique t suit loi Student
(proche de 2 pour IC à 95%)

NB : Si distribution en ”L”, ”J” ou ”U” → méthodes non paramétrique (ex: **”**bootstrap**”**)

Intervalles de confiance de la moyenne

La **valeur Z_α** se lit dans la table des valeurs de la **loi Normale CR** unilatérale

Attention : Comme on veut que le risque soit réparti de part et d'autre de l'IC, il faut lire la valeur **α/2** dans les tables construites en unilatéral (elles donnent l'aire sous la courbe à droite de Z).



Exemples : on fixe α, on le divise par 2 et on cherche le Z correspondant dans la table

Pour α = 0.05 (5%), lire α/2 = 0.025 → Z = 1.96

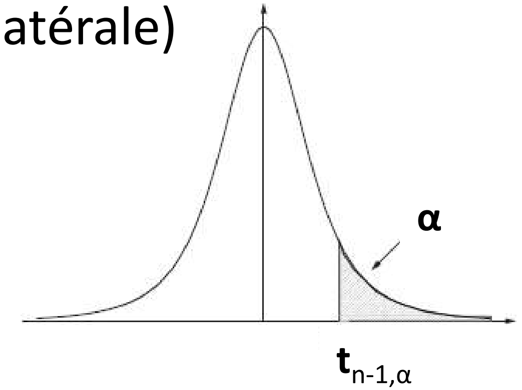
Pour α = 0.1 (10%), lire α/2 = 0.05 → Z = 1.6449

α	0.000	0.001	0.002	0.003	0.004	0.005	0.006
0.00	∞	3.0902	2.8782	2.7478	2.6521	2.5758	2.5121
0.01	2.3263	2.2904	2.2571	2.2262	2.1973	2.1701	2.1444
0.02	2.0537	2.0335	2.0141	1.9954	1.9774	1.9600	1.9431
0.03	1.8808	1.8663	1.8522	1.8384	1.8250	1.8119	1.7991
0.04	1.7507	1.7392	1.7279	1.7169	1.7060	1.6954	1.6849
0.05	1.6449	1.6352	1.6258	1.6164	1.6072	1.5982	1.5893
0.06	1.5548	1.5464	1.5382	1.5301	1.5220	1.5141	1.5063
0.07	1.4758	1.4684	1.4611	1.4538	1.4466	1.4395	1.4325
0.08	1.4051	1.3984	1.3917	1.3852	1.3787	1.3722	1.3658
0.09	1.3408	1.3346	1.3285	1.3225	1.3165	1.3106	1.3047
0.10	1.2816	1.2759	1.2702	1.2646	1.2591	1.2536	1.2481

Intervalles de confiance de la moyenne

La **valeur $t_{n-1,\alpha}$** se lit dans la table des valeurs de la **loi de Student** (unilatérale)

- Dépend de l'effectif (n) de l'échantillon
- On prendra la valeur à n-1 degré de liberté (ddl)



Exemple pour n = 5 : on fixe α, on le divise par 2 et on cherche le t correspondant, pour ddl = n -1

Pour α=0.05, lire α/2 = 0.025

Attention à bien lire dans les tables construites en unilatéral la valeur α/2 !

ddl = 5 - 1 = 4

n \ α	0.45	0.3	0.2	0.1	0.05	0.025	0.01
1	0.158	0.727	1.376	3.078	6.314	12.706	31.821
2	0.142	0.617	1.061	1.886	2.920	4.303	6.965
3	0.137	0.584	0.978	1.638	2.353	3.182	4.541
4	0.134	0.569	0.941	1.533	2.132	$\boxed{2.776}$	3.747
5	0.132	0.559	0.920	1.476	2.015	2.571	3.365

***NB:** +échantillon est grand, + distribution de statistique t → loi N, donc + valeur t → Z.*

Intervalles de confiance de la moyenne

Application 1 :

On a mesuré la taille de 40 étudiants de l'UPJV.

$\bar{x}$ = 178.0 cm

S^2 = 50.4

1. On veut calculer les intervalles de confiance à 95% et 99% de la moyenne (µ) de la pop. des étudiants de l'UPJV.

Quelle formule doit on utiliser ?

A : $$IC_{1-\alpha}[\mu] = \bar{x} \pm Z_{\alpha} * \frac{S}{\sqrt{n}}$$

B : $$IC_{1-\alpha}[\mu] = \bar{x} \pm t_{n-1,\alpha} * \frac{S}{\sqrt{n}}$$

Intervalles de confiance de la moyenne

Application 1 :

On a mesuré la taille de 40 étudiants de l'UPJV.

$\bar{x}$ = 178.0 cm

S^2 = 50.4

2. Quelles sont les valeurs de Z que l'on peut lire dans la table de la LNRC pour des seuils de risque de 5 et 1% ?

Pour IC 95 % :	Pour IC 99 % :
A : 1.6449	D : 1.2816
B : 1.9600	E : 2.3263
C : 2.5758	F : 2.5758

Intervalles de confiance de la moyenne

Application 1 :

On a mesuré la taille de 40 étudiants de l'UPJV.

$\bar{x}$ = 178.0 cm

S^2 = 50.4

3. Quelles sont les intervalles de confiance à 95 et 99 % de µ ?

Pour IC 95 % :

A : [79.22 ; 276.78]

B : [164.09 ; 191.91]

C : [177.65 ; 178.35]

D : [175.80 ; 180.20]

Pour IC 99 % :

E : [48.18 ; 307.82]

F : [159.71 ; 196.29]

G : [175.11 ; 180.89]

H : [177.54 ; 178.46]

Intervalles de confiance de la moyenne

Application 2 : Et si on réduit l'effectif ??

n = **10** étudiants

$\bar{x}$ = 178.0 cm

S^2 = 50.4

1. On veut calculer l'intervalle de confiance à 95 % de µ.

Quelle formule doit on utiliser ?

A : $$IC_{1-\alpha}[\mu] = \bar{x} \pm Z_{\alpha} * \frac{S}{\sqrt{n}}$$

B : $$IC_{1-\alpha}[\mu] = \bar{x} \pm t_{n-1,\alpha} * \frac{S}{\sqrt{n}}$$

Intervalles de confiance de la moyenne

Application 2 : Et si on réduit l'effectif ??

n = **10** étudiants

$\bar{x}$ = 178.0 cm

S^2 = 50.4

2. Quel est l'intervalle de confiance à 95 % de µ ?

Pour IC 95 % :

A : [177.60 ; 178.40]

B : [172.92 ; 183.08]

C : [177.67 ; 178.33]

D : [173.88 ; 182.12]