Statistiques

* **Estimation Ponctuelle :**

$\theta \begin{cases} m \\ \sigma \\ \sigma^2 \end{cases}$ $X_1, X_2, \ldots, X_n$ v.a

$T_n = f(X_1, X_2, \ldots X_n)$ Un estimateur

→ Caractéristiques principales d'un estimateur :

$T_n \xrightarrow[n \to +\infty]{P} \theta$ converge

$b(T_n) = E(T_n) - \theta = 0$

$E[T_n - (E(T_n))^2] = V(T_n) - b(T_n)^2$ $\underset{0}{\shortparallel}$ "si l'estimateur est sans biais"

* **Estimation par intervalle :**

$P[a < \theta < b] = \underbrace{1-\alpha}$ θ ? a ? b ?

Nv de risque ← Nv de confiance

→ **Théorème Central limite**

Soit $X_1, \ldots X_n$ n v.a indépendantes et suivants la $\hat{m}$ loi $L(m, \sigma)$ (L est une loi quelconque)

si $n \longrightarrow +\infty$: $\bar{X}_n = \frac{1}{n} \sum_{i=1} X_i \sim N(m, \sigma/\sqrt{n})$

⚠ Dans la réalité $n \geq 30$ pour appliquer T.C.L.

* $X \cap Y = \emptyset$

→ $V(\alpha X + \beta Y) = \alpha^2 V(X) + \beta^2 V(Y)$

→ $V(\bar{X}_n) = \frac{n\sigma^2}{n^2} = \sigma^2/n$

→ Cas 1 : $X \sim N(m, \sigma)$

le but : Estimation de la moyenne m

• σ est connue

$X_1, X_2, \ldots, X_n \sim N(m, \sigma)$

$\bar{X}_n = \frac{1}{n} \sum_i X_i \sim N(m, \sigma/\sqrt{n})$

$$\boxed{\bar{X}_n - \mu_\alpha \frac{\sigma}{\sqrt{n}} < m < \bar{X}_n + \mu_\alpha \cdot \frac{\sigma}{\sqrt{n}}}$$

• σ est inconnue

$a = \bar{X}_n - \mu_\alpha S^*/\sqrt{n}$

$b = \bar{X}_n + \mu_\alpha S^*/\sqrt{n}$

$$\boxed{a < m < b}$$

→ **Théorème :**

Soient $X_1, X_2, \ldots, X_n$ n v.a indépendantes et suivant loi $N(m, \sigma)$: $R_n = \frac{\bar{X}_n - m}{S^*/\sqrt{n}} = \frac{\bar{X}_n - m}{S/\sqrt{n-1}}$

$R_n \sim T_{(n-1)}$

[illegible] va L_n est dite de $\chi^2_{(n)}$ si :

$L_n = \sum_{i=1}^{n} U_i^2$ où $\forall i$ $U_i \sim N(0,1)$ et indépendantes

- Une va R_n est dite de Student à n degrés de liberté

$$R_n = \frac{nT}{\sqrt{L_n}} \quad \text{avec } T \sim N(0,1) \text{ et } L_n \sim \chi^2_{(n)}$$

$$\frac{nT}{\sigma^2} = \sum_i \left(\frac{x_i - m}{\sigma}\right)^2 \sim N(0,1) \quad \text{d'où } \frac{nT}{\sigma^2} \sim \chi^2_{(n)}$$

<u>Le but</u> : Estimation de σ^2

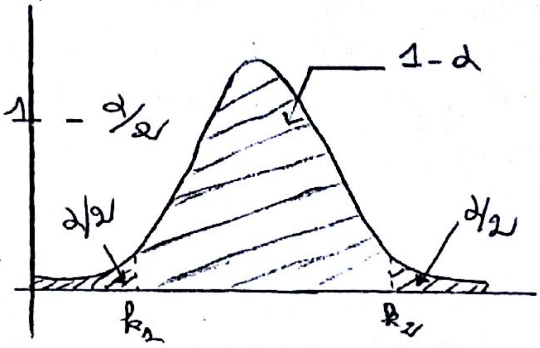
* <u>si m connue</u> : $\boxed{T_n = \frac{1}{n} \sum_{i=1}^{n} (x_i - m)^2}$

* L'hypothèse de risque symétrique *

$$P\left[\frac{nT_n}{\sigma^2} < k_1\right] = \alpha/2 \Rightarrow k_1$$

$$P\left[\frac{nT_n}{\sigma^2} > k_2\right] = \alpha/2 \Rightarrow P\left[\frac{nT_n}{\sigma^2} < k_2\right] = 1 - \alpha/2$$

$$\Downarrow k_2$$



l'intervalle de confiance :

$$\frac{nt}{k_1} < \sigma^2 < \frac{nt}{k_2} \quad \text{avec } t \text{ valeur prise par } T$$

* <u>si m inconnue</u> :

on utilise : $\frac{nS^2}{\sigma^2} \sim \chi^2_{(n-1)}$

<u>Démo</u> :

$$\frac{nS^2}{\sigma^2} = \sum_{i=1}^{n-1} U_i^2 = R_n - L_n \quad \text{où } R_n \sim \chi^2_{(n)} \text{ et } L_n \sim \chi^2_{(1)}$$

avec <u>plusieurs conditions</u>

- $X_i \sim N(m, \sigma)$

$$\boxed{P\left[\frac{nS^2}{l_2} < \sigma^2 < \frac{nS^2}{l_1}\right] = 1 - \alpha}$$

$$\boxed{S_n^{*2} = \frac{n}{n-1} S_n^2}$$

Définitions:

1. Une v.a L_n est dite de $\chi^2_{(n)}$ si:

Test d'hypothèses :

* Contrôle de Qualité : Exemple

Test unilatéral : → le fait de chercher qu'une seule valeur critique

→ Quand l'hypothèse H_1 est sous forme

H_1 : $m > m^*$

$m < m^*$

- Un téléviseur doit avoir une durée de vie de :

$m = 10\,500\ h$ $\sigma = 100\ h$

la norme — la moyenne théorique — l'écart-type

* on prend un échantillon : $X_1, X_2, \dots, X_n$ avec $n = 50$

X_i : "Durée de vie du téléviseur n° i"

* on doit calculer : $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$,

* Le test d'hypothèse est composé de :
  - La taille de l'échantillon : $n = 50$
  - H_0 : "$m = 10500$" Conformité
  - H_1 : "$m < 10500$" Non Conformité
  - α risque de 1ère espèce.

Décision Prise		Réalité : H_0 vrai	Réalité : H_1 vrai
	H_0 acceptée	Décision correcte $1-\alpha$	Erreur de 2ème espèce β
	H_0 rejetée	Erreur de 1ère espèce α	Décision Correcte $1-\beta$

Probabilité

- β est une err[eur] qui consiste [à] fermer la chaine de production [...] si l'échant[illon] est confor[me]

* α ne dépasse pas 5% en Général.
* $\alpha = 0$ est irréalisable.
* Résolution du problème :

$n \geq 30$, d'après T.C.L $\bar{X} = \frac{1}{-}\sum X_i \sim N(m, \sigma/\sqrt{n})$

Définitions:

+ **Sous H_0 :** $m = 10500\,h$

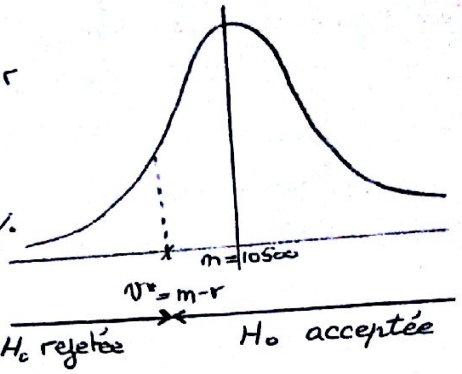
$\text{T.C.L}_{\text{sous } H_0}$ $\quad \bar{X} \sim N\left(10500, \frac{100}{\sqrt{50}}\right)$

$$\alpha = \mathbb{P}[\bar{X} < v^*]$$

$$= \mathbb{P}\left[\frac{\bar{X}-m}{\sigma/\sqrt{n}} < -\frac{r\sqrt{n}}{\sigma}\right] \quad \text{car } v^* = m - r$$

$$= \Pi\left(-\frac{r\sqrt{n}}{\sigma}\right) = 1 - \Pi\left(\frac{r\sqrt{n}}{\sigma}\right)$$

$$\Rightarrow \Pi\left(\frac{r\sqrt{n}}{\sigma}\right) = 1 - \alpha = 0{,}95 \quad \text{avec } \alpha = 5\%$$



$$\left.\begin{aligned}\Pi(1{,}64) &= 0{,}9495\\ \Pi(1{,}65) &= 0{,}9505\end{aligned}\right\} \frac{r\sqrt{n}}{\sigma} \simeq 1{,}645$$

$\Downarrow$

d'où r et ensuite $v^* = m - r$

> si $\bar{x}$ trouvé est supérieur à v^* alors le lot est acceptable, conforme.
> sinon, le lot est non conforme.

Test bilatéral :

- → chercher deux valeurs critiques.
- → Quand l'hypothèse H_1 est sous forme :

$$H_1 : m \neq m^*$$

Exemple : $n = 16$

$X_i \sim N(m, \sigma)$ où σ inconnue

$$\frac{\bar{X}_n - m}{S^*/\sqrt{n}} \sim T_{(n-1)}$$

$$\boxed{S^* = \left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right]^{1/2}}$$

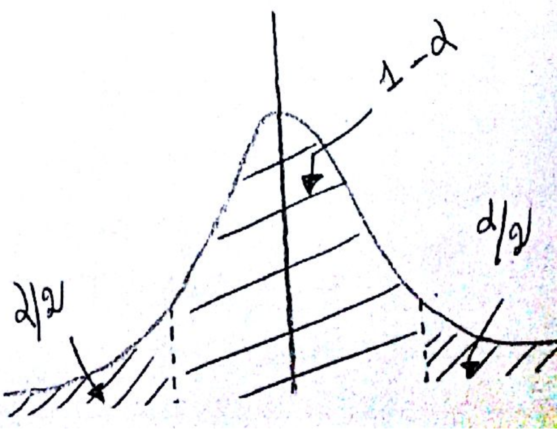
⇝ Le test d'hypothèse :

- $n = 16$
- H_0 : "$m = 400$" Conformité
- H_1 : "$m \neq 400$" Non Conformité
- $\alpha = 5\%$

+ **Sous H_0 :**

$$P[m - r < \bar{X} < m + r] = 1 - \alpha$$

$$P\Big[\underbrace{-\frac{r\sqrt{n}}{S^*}}_{-t_\alpha} < \frac{\bar{X}-m}{S^*/\sqrt{n}} < \underbrace{\frac{r\sqrt{n}}{S^*}}_{t_\alpha}\Big] = 0{,}95$$



D'après la table de student :

$$t_\alpha = 1{,}753 = \frac{r\sqrt{n}}{S^*} \Rightarrow r = \cdots$$