# ECOLE DES HAUTES ETUDES COMMERCIALES

**Module** : Statistique de Gestion
**Niveau** : 1$^{ère}$ année Master (TC)
**Groupes** : 1, 2, 7 et 8

**Année Académique** : 2013/2014
**Enseignant** : KHERRI Abdenacer
**Site web** : www.sg-ehec.jimdo.com


Support pédagogique de cours N° 03 :

# Estimation

**Plan du cours :**



## 1. **Introduction** :

La statistique est l'ensemble des méthodes scientifiques à partir desquelles on recueille, organise, résume, présente et analyse des données, et qui permettent d'en tirer des conclusions et de prendre des décisions judicieuses.

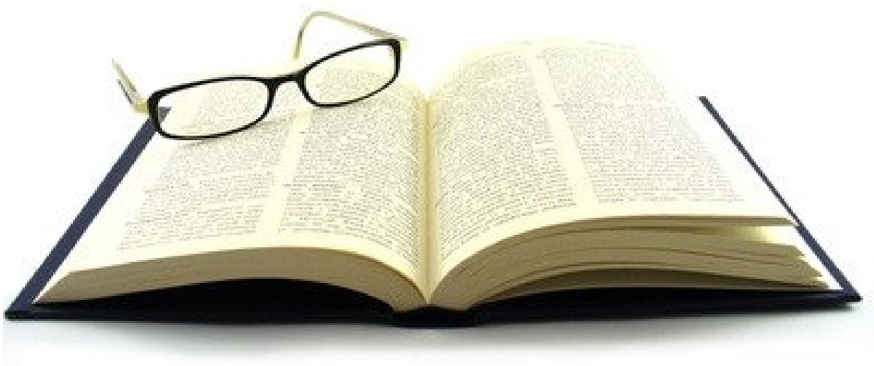
Au lieu d'examiner l'ensemble des données possible qu'on appelle encore "***la population***", en pratique, on en étudie une toute petite partie appelée "***échantillon***". À partir des résultats mesurés sur cet échantillon, nous essayons d'induire des conclusions valables pour l'entièreté de la population : c'est la partie de la statistique que l'on appelle "***statistique inductive***" ou "***statistique inférentielle***". De manière générale, l'inférence statistique est l'étude des conclusions que l'on peut tirer d'un échantillon pour une population dont l'échantillon est issu, ainsi que le degré de précision des conclusions.

Dans ce cadre, les problèmes qui se posent sont ceux de l'estimation des paramètres (moyenne, variance, écart-type et proportion) d'une population à partir des échantillons issus de cette même population.

Pour que les conclusions soient valables, il faut que l'échantillon soit représentatif de la population. Cela signifie qu'il doit être prélevé d'une manière aléatoire, c'est-à-dire que tous les éléments de la population ont la même probabilité d'être choisis.

## 2. **Terminologie** :

| Terme | Définition |
|---|---|
| **Estimation** | Une estimation est une valeur particulière prise par un estimateur. |
| **Estimateur** | Statistique définie à partir d'un échantillon (fonction des $X_i$ de l'échantillon) qui permet d'estimer un paramètre. |
| **Estimation ponctuelle** | La valeur de l'estimateur calculée à partir de l'échantillon spécifique qui a été observé. |
| **Estimation par intervalle de confiance** | Au lieu d'estimer le paramètre par une seule valeur, on préfèrera donner un intervalle de valeurs pour celui-ci. On pourra ainsi fixer un niveau de confiance à notre estimation et déterminer le degré de précision ou la marge d'erreur qui lui est associée. |
| **Intervalle de confiance** | Intervalle qui permet de définir une marge d'erreur entre les résultats d'un sondage et un relevé de la population. |
| **Marge d'erreur** | Valeur ± ajoutée ou soustraite à l'estimation ponctuelle pour construire l'intervalle de confiance d'un paramètre de la population. |
| **$\sigma$ connu** | Cas où des données historiques ou d'autres informations fournissent une valeur de l'écart-type de la population avant tout échantillonnage. La procédure d'estimation par intervalle utilise cette valeur de $\boldsymbol{\sigma}$ dans le calcul de la marge d'erreur. |
| **$\sigma$ inconnu** | Cas le plus courant caractérisé par l'absence bonne base d'estimation de l'écart-type de la population avant l'échantillonnage. La procédure d'estimation par intervalle utilise l'écart-type de l'échantillon $\boldsymbol{s}$ pour calculer la marge d'erreur. |
| **Tirage exhaustif** | Tirage sans remise. |
| **Tirage non exhaustif** | Tirage avec remise. |
| **Grands échantillons** | La taille de l'échantillon est supérieure ou égale à 30. |
| **Petits échantillons** | La taille de l'échantillon est inférieure à 30. |
| **Biaisé et non-biaisé** | "Biais" désigne un écart entre la valeur d'un paramètre et la valeur estimée de ce paramètre. |



## 3. Symboles utilisés :

| Symbole | Signification |
|---|---|
| $\sigma$ | L'écart-type de la population |
| $\sigma^2$ | La variance de la population |
| $\sigma^*$ | L'écart-type estimé de la population |
| $s^2$ | Variance biaisée d'un échantillon |
| $\grave{s}^2$ | Variance non biaisée d'un échantillon |
| $s$ | Ecart-type biaisé d'un échantillon |
| $\grave{s}$ | Ecart-type non biaisé d'un échantillon |
| $\theta$ | Valeur estimée d'un paramètre |
| $\Theta$ | Estimateur |
| $\mu$ | La moyenne de la population |
| $\alpha$ | Le coefficient de risque |
| $1-\alpha$ | Le coefficient de confiance |
| $Z_{\alpha/2}$ | La valeur critique (l'écart réduit) |

## 4. Formules de la variance et de l'écart-type d'un échantillon :

- **La variance biaisée :** $s^2 = \frac{\sum(x_i-\bar{x})^2}{n} = \frac{\sum x_i{}^2}{n} - \bar{x}^2$
- **La variance non biaisée :** $\grave{s}^2 = \frac{\sum(x_i-\bar{x})^2}{n-1} = \frac{\sum x_i{}^2}{n-1} - \frac{n}{n-1}\bar{x}^2$
- **L'écart-type biaisé :** $s = \sqrt{s^2}$
- **L'écart-type non biaisée :** $\grave{s} = \sqrt{\grave{s}^2}$

## 5. **Estimation** :

Pour bien comprendre le sens du mot "***estimation***", on va citer 3 définitions différentes ensuite on donne une définition de synthèse.

***Déf (01)*** : Action d'estimer, de déterminer une valeur[1].

***Déf (02)*** : Recherche de la valeur d'un ou de plusieurs paramètres d'une loi statistique à partir d'observations ou de sondages sur un ou plusieurs échantillons d'une population[2].

***Déf (03)*** : une estimation est une valeur particulière prise par un estimateur[3].

Donc on peut dire que :

**L'estimation est le procédé par lequel on détermine les valeurs inconnues des paramètres de la population à partir des données de l'échantillon. Pour cela, on utilise des distributions théoriques, c'est à dire des variables aléatoires dont on connait les lois de probabilité.**

## 6. **Estimateur** :

Un estimateur est une statistique permettant d'évaluer un paramètre inconnu relatif à une loi de probabilité (comme son espérance ou sa variance). Il peut par exemple servir à estimer certaines caractéristiques d'une population totale à partir de données obtenues sur un échantillon.

### 6.1. **Définition :**

Soient $\boldsymbol{X_1, X_2, \dots, X_i, \dots, X_n}$, $\boldsymbol{n}$ réalisations indépendantes de la variable aléatoire $\boldsymbol{X}$ (discrète ou continue) et $\boldsymbol{\theta}$ un paramètre associé à la loi de probabilité suivi par $\boldsymbol{X}$, un **estimateur** du paramètre $\boldsymbol{\theta}$ est une variable aléatoire $\boldsymbol{\Theta}$ fonction des $\boldsymbol{X_i}$ : $\boldsymbol{\Theta = f(X_1, X_2, \dots, X_i, \dots, X_n)}$

Si on considère $\boldsymbol{n}$ observations : $\boldsymbol{x_1, x_2, \dots, x, \dots, x_n}$, l'estimateur $\boldsymbol{\Theta}$ fournira une estimation de $\boldsymbol{\theta}$ notée également $\boldsymbol{\hat{\theta}}$ :
$\boldsymbol{\hat{\theta} = f(x_1, x_2, \dots, x_i, \dots, x_n)}$

1 . http://www.le-dictionnaire.com/definition.php?mot=estimation
2 . http://www.larousse.fr/dictionnaires/francais/estimation
3 . http://fr.wikipedia.org/wiki/Estimation

L'estimation d'un paramètre inconnu, noté $\boldsymbol{\theta}$ est fonction des observations résultant d'un échantillonnage aléatoire simple de la population.

L'estimateur est donc une nouvelle variable aléatoire construite à partir des données expérimentales et dont la valeur se rapproche du paramètre que l'on cherche à connaître.

**L'estimateur de $\boldsymbol{\theta}$ est une variable aléatoire $\boldsymbol{\Theta}$ dont la distribution de probabilité s'appelle la distribution d'échantillonnage du paramètre $\boldsymbol{\theta}$.**

**L'estimateur $\boldsymbol{\Theta}$ admet donc une espérance $\boldsymbol{E(\Theta)}$ et une variance $\boldsymbol{V(\Theta)}$.**

## 6.2. <u>Propriétés</u> :

### 6.2.1. Convergence :

L'estimateur $\boldsymbol{\Theta}$ doit tendre vers la valeur réelle du paramètre $\boldsymbol{\theta}$ lorsque le nombre d'individus étudié augmente. On dit que "l'estimateur est convergent".

Si $\forall \boldsymbol{\varepsilon} > \mathbf{0}$ $\boldsymbol{P}(|\boldsymbol{\Theta} - \boldsymbol{\theta}| > \boldsymbol{\varepsilon}) \rightarrow \mathbf{0}$ lorsque $\boldsymbol{n} \rightarrow \infty$

Ceci équivaut à dire qu'en limite $\boldsymbol{\Theta} \rightarrow \boldsymbol{\theta}$ lorsque $\boldsymbol{n} \rightarrow \infty$

### 6.2.2. Biais d'un estimateur :

Le biais d'un estimateur noté $\boldsymbol{B(\Theta)}$ est la différence moyenne entre sa valeur et celle du paramètre qu'il estime. Le biais doit être égal à **0** pour avoir un bon estimateur.

$$\boldsymbol{B(\Theta) = B(\Theta - \theta) = B(\Theta) - B(\theta) = B(\Theta) - \theta = 0}$$

Ainsi l'estimateur sera "sans biais" si son espérance est égale à la valeur du paramètre de la population.

$\boldsymbol{B(\Theta) = \theta}$

### 6.2.3. Variance d'un estimateur :

Si deux estimateurs sont convergents et sans biais, le plus efficace est celui qui a "la variance la plus faible" car ses valeurs sont en moyenne plus proches de la quantité estimée.

$\boldsymbol{V(\Theta) = E(\Theta - E(\Theta))^2}$

## 7. Types d'estimation :

La distribution exacte d'une variable $X$ modélisant le caractère qui intéresse le statisticien est généralement partiellement connue. Souvent la loi de $X$ dépend d'un paramètre inconnu, on cherche à se faire une idée sur ce paramètre à partir des données observées sur l'échantillon.

Attribuer au paramètre une valeur numérique unique est une ***estimation ponctuelle***, mais quelles sont les chances pour que cette estimation ponctuelle soit exacte ?

Plutôt que d'estimer un paramètre à l'aide d'un seul nombre, il arrive fréquemment que l'on fasse l'estimation en donnant un intervalle de valeurs, Un ***intervalle de confiance*** est défini de telle sorte que l'on puisse affirmer avec un degré de confiance fixé que le paramètre visé se trouve dans cet intervalle.

### 7.1. L'estimation ponctuelle :

A partir de l'examen d'un échantillon, on essaye d'obtenir une information quantitative sur des paramètres de la population à estimer (généralement $\boldsymbol{\mu}$ et $\boldsymbol{\sigma}$)
$\boldsymbol{\Theta}$ est un estimateur de $\boldsymbol{\theta}$ si $\boldsymbol{\Theta}$ converge en moyenne vers $\boldsymbol{\theta}$ :
$E(\boldsymbol{\Theta}) \rightarrow \theta$ et $V(\boldsymbol{\Theta}) \rightarrow 0$ quand $n \rightarrow \infty$
Un estimateur $\boldsymbol{\Theta}$ de $\boldsymbol{\theta}$ est dit sans biais si : $\mathbf{E(\Theta) = \theta}$
Par exemple $\overline{X}$ est un estimateur ponctuel non biaisé de $\boldsymbol{\mu}$ car $\mathbf{E}(\overline{X}) = \mu_{\overline{X}} = \mu$
$S^2$ est estimateur biaisé de $\sigma^2$ car $E(S^2) = \frac{n-1}{n}\sigma^2$
Si on utilise la variance $\grave{S}^2 = \frac{n}{n-1}S^2$ donc $E(\grave{S}^2) = \sigma^2$ dans ce cas on dit que $\grave{S}^2$ est un estimateur sans biais de $\sigma^2$

#### 7.1.1. Estimation d'une moyenne :

La "moyenne arithmétique" constitue le meilleur estimateur de $\boldsymbol{\mu}$, espérance de la loi de probabilité de la variable aléatoire $X$ :

$$\mu = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

#### 7.1.2. Estimation d'une variance :

- $\mu$ **connu :**

La "variance observée" constitue le meilleur estimateur de $\boldsymbol{\sigma}^2$, variance de la loi de probabilité de la variable aléatoire $X$ lorsque l'espérance $\boldsymbol{\mu}$ est connue :

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{\sum {x_i}^2}{n} - \mu^2$$

- **$\mu$ inconnu :**

Le meilleur estimateur de $\boldsymbol{\sigma}^2$, variance de la loi de probabilité de la variable aléatoire $\boldsymbol{X}$ lorsque l'espérance $\boldsymbol{\mu}$ est inconnue est :

$$\sigma^2 = \frac{n}{n-1} S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

**<u>Exemple</u> :** Lors d'un concours radiophonique, on note $\boldsymbol{X}$ le nombre des réponses reçues chaque jour, on suppose que $\boldsymbol{X \to N(\mu, \sigma)}$. Durant **10** jours on a obtenu :

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | **200** | **240** | **190** | **150** | **220** | **180** | **170** | **230** | **210** | **210** |

- Calculer la moyenne $\overline{x}$, l'écart-type biaisé $s$ et l'écart-type non biaisé $\grave{s}$
- Donner une estimation ponctuelle de la moyenne $\boldsymbol{\mu}$ et de l'écart-type $\boldsymbol{\sigma}$.

**<u>Réponse</u> :**

### 7.1.3. Estimation d'une fréquence (proportion) :

Soit le schéma de Bernoulli dans lequel le caractère **A** correspond au succès. On not $\boldsymbol{\pi}$ e la proportion des individus de la "population" possédant le caractère **A**. La valeur de ce paramètre étant inconnu, on cherche à estimer la fréquence $\boldsymbol{\pi}$ à partir des données observables sur un échantillon.

A chaque échantillon non exhaustif de taille $\boldsymbol{n}$, on associe l'entier $\boldsymbol{k}$, nombre d'individus possédant le caractère **A**.

Soit $\boldsymbol{K}$ une variable aléatoire discrète suivant une loi binomiale **B(*n*,*p*)** et pour laquelle on souhaite estimer "la proportion $\boldsymbol{\pi}$ ".

La "proportion observée" du nombre de succès observé dans un échantillon de taille $\boldsymbol{n}$ constitue le meilleur estimateur de $\boldsymbol{\pi}$ :

$$\boldsymbol{\pi = p = \frac{k}{n}}$$

**<u>Exemple (01)</u> :**

On a prélevé au hasard, dans une population de pièces **100** unités, Sur ces **100** unités, **20** sont défectueuses. Estimer la proportion de la population dans ce cas.

✍ **<u>Réponse</u> :**

……………………………………………………………………………………………

…………..…………………………………………………………………………………

..…………………………………………………………………………………..……

……………………………………………………………………………………………

……………………………………………………………………………………………

## 7.2. <u>Estimation par intervalle de confiance</u> :

L'**estimation par intervalle** associe à un échantillon aléatoire, **un intervalle** $[\boldsymbol{\theta_1}, \boldsymbol{\theta_2}]$ qui recouvre $\boldsymbol{\theta}$ avec une certaine probabilité.

Cet intervalle est appelé l'**intervalle de confiance** du paramètre $\boldsymbol{\theta}$ car la probabilité que $\boldsymbol{\theta}$ dont la valeur est inconnue se trouve compris entre $\boldsymbol{\theta_1}$ *et* $\boldsymbol{\theta_2}$ est égale à $\boldsymbol{1 - \alpha}$ **le coefficient de confiance** $\boldsymbol{P(\theta_1 < \theta < \theta_2) = 1 - \alpha}$

Son complément **α** correspond au **coefficient de risque** $\boldsymbol{P(\theta \notin [\theta_1, \theta_2]) = \alpha}$

Un intervalle de confiance indique **la précision d'une estimation** car pour un risque **α** donné, l'intervalle est d'autant plus grand que la précision est faible.

### 7.2.1. Estimation d'une moyenne :

On veut estimer la moyenne **μ** de la population à l'aide d'un échantillon aléatoire (si la taille de l'échantillon est grande $n \geq 30$ la distribution d'échantillonnage de la moyenne est normale quelle que soit la distribution de la population).

$\overline{X} \to N(\mu_{\overline{X}}, \sigma_{\overline{X}})$ avec :

$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ (si le tirage est non exhaustif)

$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ (si le tirage est exhaustif)

On a : $\frac{\overline{X}-\mu_{\overline{X}}}{\sigma_{\overline{X}}}$ suit une loi normale centrée réduite **N(0,1)**

On cherche un intervalle centré sur **μ** avec une probabilité égale à **α** : $P\left(\frac{\overline{X}-\mu_{\overline{X}}}{\sigma_{\overline{X}}} < Z_{\alpha/2}\right) = \alpha$

On peut distinguer entre les cas suivants :

- **$\sigma$ connue :** $\mu \in \left[\overline{x} - Z_{\alpha/2}\sigma_{\overline{X}} , \overline{x} + Z_{\alpha/2}\sigma_{\overline{X}}\right]$

**Exemple (01) :**

La taille moyenne d'un échantillon aléatoire de **40** personnes extrait d'une population de **780** individus est de **170 cm**, l'écart-type pour toute la population vaut **24 cm**. Trouver l'intervalle de confiance pour la taille moyenne de la population à **95 %**[1].

✍ **Réponse :**

…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………

1 . KHALDI Khaled, ***Méthodes statistiques et probabilités***, OPU, Alger, Algérie, 2000, P175.

…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………

- **$\boldsymbol{\sigma}$ inconnue ($\boldsymbol{n \geq 30}$) :**

$$\boldsymbol{\mu} \in \left[\overline{\boldsymbol{x}} - \boldsymbol{Z}\alpha_{/_2}\,, \overline{\boldsymbol{x}} + \boldsymbol{Z}\alpha_{/_2}\boldsymbol{\sigma}_{\overline{X}}\right]$$

avec

$$\boldsymbol{\sigma}_{\overline{X}} = \frac{\boldsymbol{S}}{\sqrt{\boldsymbol{n}-\boldsymbol{1}}}$$

**<u>Exemple (01)</u> :**

Cinq cent étudiants se présentent à un examen, un échantillon aléatoire de **38** notes donne une moyenne égale à **8,65** et un écart-type égal à **2,82**.

Trouvez l'intervalle de confiance pour la moyenne des notes de la population à[1] :

- 0,90
- 0,95
- 0,99

**<u>Réponse</u> :**

…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………

1 . KHALDI Khaled, ***Méthodes statistiques et probabilités***, Op.cit, P175.

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

- **$\sigma$ inconnue ($n < 30$) :**

$$\boldsymbol{\mu} \in \left[\overline{\boldsymbol{x}} - \boldsymbol{t}_{\alpha/2}\boldsymbol{\sigma}_{\overline{X}}\, , \overline{\boldsymbol{x}} + \boldsymbol{t}_{\alpha/2}\boldsymbol{\sigma}_{\overline{X}}\right]$$

**<u>Exemple (01)</u> :**

Les notes de statistique d'une promotion nombreuses d'étudiants sont distribuées normalement. De cette promotion on extrait un échantillon de **9** notes. La note moyenne est **9,55** avec un écart-type de **3,65**.

Trouvez l'intervalle de confiance pour la moyenne des notes de la population à[1] :

- 0,90
- 0,95
- 0,99

**<u>Réponse</u> :**

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

1 . KHALDI Khaled, ***Méthodes statistiques et probabilités***, Op.cit, P176.

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

### 7.2.2. Estimation d’une variance :

On suppose que la distribution de la population est normale. On sait que la variable $\frac{(n-1)S^2}{\sigma^2}$ suit une loi de $\chi^2$ (khi-deux) à degrés $(n-1)$ de liberté.

On sait que :

$$P\left(\chi^2\left[\frac{\alpha}{2};n-1\right] \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2\left[1-\frac{\alpha}{2};n-1\right]\right) = 1-\alpha$$

On prend l’inverse des parties de l’inéquation, on obtient (avec le changement de la direction de l’inéquation) :

$$P\left(\frac{1}{\chi^2\left[1-\frac{\alpha}{2};n-1\right]} \leq \frac{\sigma^2}{(n-1)S^2} \leq \frac{1}{\chi^2\left[\frac{\alpha}{2};n-1\right]}\right) = 1-\alpha$$

On multiplie les parties de l’inéquation par $(n-1)S^2$, on obtient :

$$P\left(\frac{(n-1)S^2}{\chi^2\left[1-\frac{\alpha}{2};n-1\right]} \leq \frac{(n-1)S^2\sigma^2}{(n-1)S^2} \leq \frac{(n-1)S^2}{\chi^2\left[\frac{\alpha}{2};n-1\right]}\right) = 1-\alpha$$

Donc :

$$P\left(\frac{(n-1)S^2}{\chi^2\left[1-\frac{\alpha}{2};n-1\right]} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2\left[\frac{\alpha}{2};n-1\right]}\right) = 1-\alpha$$

On désigne les points $\boldsymbol{a}$ et $\boldsymbol{b}$ sachant que :

$$\boldsymbol{a} = \chi^2\left[\frac{\alpha}{2}; \boldsymbol{n}-\mathbf{1}\right] \quad \textbf{et} \quad \boldsymbol{b} = \chi^2\left[\mathbf{1}-\frac{\alpha}{2}; \boldsymbol{n}-\mathbf{1}\right]$$

On détermine les valeurs de $\boldsymbol{a}$ et $\boldsymbol{b}$ à partir de la lecture de la table de la loi $\boldsymbol{\chi^2}$ à $\boldsymbol{n-1}$ degrés de liberté.

<u>**Exemple (01)**</u> **:**

Echantillon aléatoire de taille $\mathbf{20}$ prélevé d'une population qui suit la loi normale, cet échantillon a donné une variance de $\boldsymbol{s^2 = 15}$.

Trouvez l'intervalle de confiance pour la variance de la population à seuil de **95%** [1] :

<u>**Réponse**</u> **:**

…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………
…………………………………………………………………………………………………

1. محمد صبحي أبو صالح، ***الطرق الإحصائية***، دار اليازوري العلمية للنشر و التوزيع، عمان، الأردن، 2009، ص321.

### 7.2.3. Estimation d'une proportion :

$$\boldsymbol{\pi} \in \left[\boldsymbol{p} - \boldsymbol{Z}_{\alpha/2}\sqrt{\frac{\boldsymbol{pq}}{\boldsymbol{n}}}\ ,\boldsymbol{p} + \boldsymbol{Z}_{\alpha/2}\sqrt{\frac{\boldsymbol{pq}}{\boldsymbol{n}}}\right]$$

**<u>Exemple (01)</u> :**

Un échantillon aléatoire de **150** individus sur les **10.000** personnes d'une population a montré que **27** individus possédaient un certain caractère[1].

Estimez l'intervalle de confiance de la proportion d'individus ayant le caractère dans la population avec un seuil de confiance $\boldsymbol{\alpha = 0,1}$ :

✍ **<u>Réponse</u> :**

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

1 . KHALDI Khaled, ***Méthodes statistiques et probabilités***, Op.cit, P178.

## 8. Synthèse :

| Table de l'écart-réduit | |
|---|---|
| Coefficient du risque (α) | Ecart-réduit ($Z\alpha/_2$) |
| α = 0,01 | $Z\alpha/_2 = 2{,}576$ |
| α = 0,05 | $Z\alpha/_2 = 1{,}960$ |
| α = 0,10 | $Z\alpha/_2 = 1{,}645$ |

| Paramètre à estimer | Estimateur |
|---|---|
| Moyenne de la population ($\mu$) | Moyenne de l'échantillon ($\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$) |
| Variance de la population ($\sigma^2$) | Variance de l'échantillon ($s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$) |
| Ecart-type de la population ($\sigma$) | Ecart-type de l'échantillon ($s = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$) |

Le tableau ci-dessous envisage tous les cas auxquels vous pouvez être confrontés[1].

| Paramètre à estimer | Loi de la population | | Statistique | Loi |
|---|---|---|---|---|
| Moyenne $\mu$ | Normale | $\sigma^2$ connu | $\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right)$ | N(0,1) |
| | | $\sigma^2$ inconnu | $\sqrt{n}\left(\frac{\bar{X}-\mu}{S}\right)$ | Student(n-1) |
| | Quelconque n > 30 | $\sigma^2$ connu | $\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right)$ | ~ N(0,1) |
| | | $\sigma^2$ inconnu | $\sqrt{n}\left(\frac{\bar{X}-\mu}{S}\right)$ | ~ N(0,1) |
| Variance $\boldsymbol{\sigma^2}$ | Normale | $\mu$ connu | $\sum\frac{(X_i-\mu)^2}{\sigma^2}$ | $\chi^2$ à n d.d.l. |
| | | $\mu$ inconnu | $\frac{(n-1)S^2}{\sigma^2}$ | $\chi^2$ à n-1 d.d.l. |
| Proportion $\boldsymbol{P}$ | n > 50 | | $\sqrt{n}\,\frac{F-p}{\sqrt{p(1-p)}}$ | ~ N(0,1) |

1 . http://www.iutbayonne.univ-pau.fr/~grau/2A/stat/cadre3.html

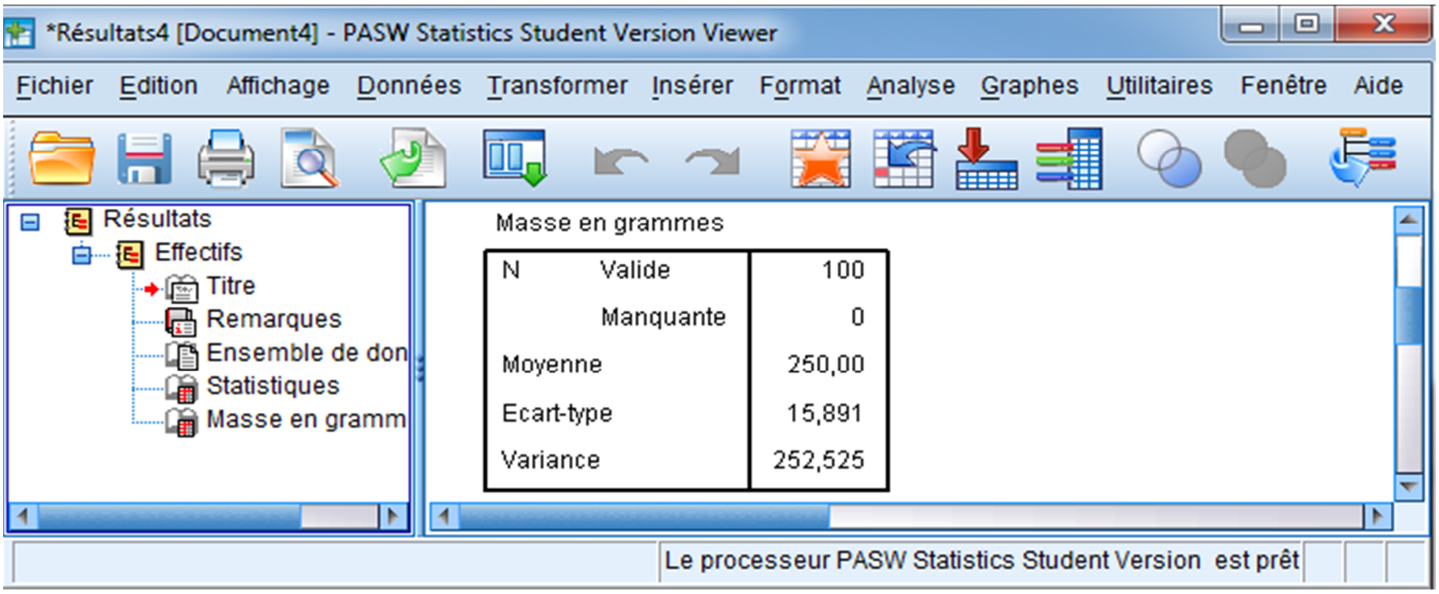
## 9. <u>Etude de cas</u> :

| Etude de cas : Machine automatique |
|---|

Une machine automatique remplit des paquets dont la masse théorique doit être de **250g**. Les masses observées pour un échantillon de **100** paquets pris au hasard et avec remise à la sortie de la machine, ont donné les résultats suivants :

| Masse en grammes | Nombre de paquets |
|---|---|
| [ 215 - 225 [ | 7 |
| [ 225 - 235 [ | 11 |
| [ 235 - 245 [ | 19 |
| [ 245 - 255 [ | 26 |
| [ 255 - 265 [ | 18 |
| [ 265 - 275 [ | 13 |
| [ 275 - 285 [ | 6 |

Après l'analyse de ces données par **SPSS**, on a obtenu les valeurs de la moyenne, la variance et l'écart-type de cet échantillon, la figure suivante montre les résultats obtenus :




### <u>Travail à faire</u> :

1. Calculer la moyenne $\overline{x}$, l'écart-type biaisé $s$ et l'écart-type non biaisé $\hat{s}$ et comparer vos résultats avec ceux qui ont été obtenus par **SPSS**.
2. Soit $X$ la variable aléatoire qui, à un paquet prélevé au hasard et avec remise à la sortie de la machine, associe son poids en grammes. On suppose que $X$ suit une loi normale de paramètres $\mu$ et $\sigma$. À partir des résultats obtenus pour l'échantillon précédent, proposer une estimation ponctuelle de la moyenne $\mu$ et de l'écart-type $\sigma$.
3. Soit $\overline{X}$ la variable aléatoire qui, à chaque échantillon de **100** paquets prélevés au hasard et avec remise à la sortie de la machine, associe le poids moyen des paquets

de cet échantillon. On sait que $\overline{X}$ suit une loi normale de moyenne $\mu$ et d'écart-type $\frac{\sigma}{\sqrt{100}}$. En utilisant l'échantillon précédent et en prenant pour la valeur de $\sigma$ l'estimation ponctuelle obtenue à la 3$^{\text{ème}}$ question, déterminer un intervalle de confiance de la moyenne des poids des paquets au risque de **5 %**.

4. Démontrer que $\pi\left(Z_{\alpha/2}\right) = 1 - \frac{\alpha}{2}$
5. Déterminer un intervalle de confiance de la moyenne des poids des paquets au coefficient de confiance **75,4 %**.
6. Même question avec le coefficient de confiance **99 %**.
7. Calculer la longueur de chaque intervalle et commenter vos résultats.
8. Quelle doit être la taille minimale de l'échantillon pour connaitre, avec le coefficient de confiance **95 %**, la moyenne de la population à **2 grammes** près ?

**NB** : tous les résultats approchés seront arrondis à $10^{-2}$.

✍ **Réponse :**

.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................
.................................................................................................................................