Estimation ponctuelle et estimation par intervalle de confiance

1 Moyenne aléatoire, variance aléatoire, variance aléatoire corrigée

Le théorème suivant introduit la moyenne aléatoire et de la variance aléatoire et donne quelques propriétés classiques de ces estimateurs.

Théorème 1 (Moyenne aléatoire et variance aléatoire)**.** *Soit $(X_1, \dots, X_n)$ une suite de variables indépendantes et de même loi d'espérance $\mu \in \mathbb{R}$ et de variance $\sigma^2 \geq 0$.*

La moyenne aléatoire : *La variable aléatoire $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, appelée moyenne aléatoire, vérifie : $\mathbb{E}(\overline{X}) = \mu$ et $\mathrm{var}(\overline{X}) = \frac{\sigma^2}{n}$.*

La variance aléatoire : *La variable aléatoire $S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$, appelée variance aléatoire, vérifie : $\mathbb{E}(S^2) = \frac{n-1}{n}\sigma^2$.*

Notons que la variable aléatoire $\overline{X^r} = \frac{1}{n}\sum_{i=1}^n X_i^r$, appelée moment aléatoire d'ordre $r \in \mathbb{N}^*$, généralise la moyenne aléatoire qui est le moment aléatoire d'ordre $r = 1$. Nous utiliserons la notion de moment aléatoire pour introduire la méthode des moments.

Démonstration. Très clairement, l'espérance de la moyenne aléatoire satisfait

$$\mathbb{E}(\overline{X}) = \frac{1}{n}(\mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n)) = \mu.$$

Par ailleurs, en utilisant les propriétés classiques de la variance, on obtient

$$\mathrm{var}(\overline{X}) = \frac{1}{n^2}(\mathrm{var}(X_1) + \cdots + \mathrm{var}(X_n)) = \frac{\sigma^2}{n}.$$

Concernant l'espérance de la variance aléatoire, on a

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2 &= \frac{1}{n}\sum_{i=1}^n \left((X_i - \mu) - (\overline{X} - \mu)\right)^2, \\
&= \frac{1}{n}\sum_{i=1}^n \left((X_i - \mu)^2 - 2(X_i - \mu)(\overline{X} - \mu) + (\overline{X} - \mu)^2\right), \\
&= \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 - 2(\overline{X} - \mu)\frac{1}{n}\sum_{i=1}^n (X_i - \mu) + (\overline{X} - \mu)^2, \\
&= \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 - 2(\overline{X} - \mu)^2 + (\overline{X} - \mu)^2, \\
&= \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 - (\overline{X} - \mu)^2.
\end{aligned}$$

Ainsi, en prenant l'espérance de S^2 et en utilisant la variance de la moyenne aléatoire on en déduit que

$$\mathbb{E}(S^2) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i - \mu)^2 - \mathbb{E}(\overline{X} - \mu)^2 = \frac{1}{n}\sum_{i=1}^{n} \mathrm{var}(X_i) - \mathrm{var}(\overline{X}) = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2.$$

□

On peut remarquer que l'espérance de S^2 n'est pas égale σ^2 (on dit que S^2 est un estimateur biaisé de σ^2). À l'inverse, la variance aléatoire corrigée S^2_{corr} définie par

$$S^2_{\mathrm{corr}} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

vérifie $\mathbb{E}(S^2_{\mathrm{corr}}) = \sigma^2$ (on dit que S^2_{corr} est un estimateur sans biais de σ^2).

2 Estimation ponctuelle

2.1 Modélisation statistique

Définition 1. *On appelle n-échantillon d'une loi de probabilité $\mathbb{P}_X$ une suite $(X_1, \ldots, X_n)$ de variables aléatoires réelles indépendantes et de même loi $\mathbb{P}_X$. On appelle n-échantillon expérimental une réalisation $(x_1, \ldots, x_n) \in \mathbb{R}^n$ du n-échantillon.*

Dans le cadre de la statistique paramétrique la loi $\mathbb{P}_X$ dépend d'un paramètre inconnu $\theta \in \Theta \subset \mathbb{R}^k$. C'est le cas, par exemple, lorsque le $n-$échantillon suit :

— Une loi binomiale de paramètre inconnu $p \in]0, 1[$.

— Une loi normale dont les paramètres inconnues sont $\mu \in \mathbb{R}$ et $\sigma^2 > 0$

— Une loi de Poisson de paramètre inconnu $\lambda > 0$.

L'objectif est d'estimer θ à partir de l'échantillon expérimental. Les notions d'estimateur ou de statistique n'ont pas de définition très formelle. Une statistique ou un estimateur est juste une variable aléatoire $g(X_1, ..., X_n)$ qui est une fonction du $n-$échantillon $(X_1, \ldots, X_n)$.

2.2 Une méthode intuitive d'estimation : la méthode des moments

Très souvent, le paramètre inconnu θ de la loi $\mathbb{P}_X$ de l'échantillon correspond, à une fonction près, à un moment de cette loi. Il est alors naturel d'utiliser le moment aléatoire pour estimer ce paramètre. Plus précisément, on considère un $n-$échantillon $(X_1, \ldots, X_n)$ de loi $\mathbb{P}_X$ qui dépend d'un paramètre inconnu $\theta \in \Theta \subset \mathbb{R}$. Si $\theta = g(\mathbb{E}(X_1^r))$ avec $r \in \mathbb{N}^*$ alors il est naturel, pour estimer θ, de substituer le moment $\mathbb{E}(X_1^r)$ par la moment aléatoire d'ordre r ($\overline{X^r} = \frac{1}{n}\sum_{i=1}^{n} X_i^r$), ce qui fournit l'estimateur des moments : $\hat{\theta}^{\mathrm{m}} = g(\overline{X^r})$. En pratique, le paramètre θ est souvent une fonction du moment d'ordre $r = 1$ (l'espérance) donc l'estimateur des moments dépend souvent de la moyenne aléatoire.

Exemple 1. *Quelques exemples d'estimateurs des moments sont donnés ci-dessous*

— *Lorsque $\mathbb{P}_X$ est une loi de Bernoulli $\mathcal{B}(p)$ de paramètre inconnu $p \in [0, 1]$ alors $p = \mathbb{E}(X_1)$ ainsi la méthode des moments fournit comme estimateur la moyenne aléatoire : $\hat{p}^{\mathrm{m}} = \overline{X}$.*

— *Lorsque $\mathbb{P}_X$ est une loi de exponentielle $\mathcal{E}(\lambda)$ de paramètre inconnu $\lambda > 0$ alors $\lambda = \frac{1}{\mathbb{E}(X_1)}$ ainsi la méthode des moments fournit comme estimateur l'inverse de la moyenne aléatoire : $\hat{\lambda}^{\mathrm{m}} = \frac{1}{\overline{X}}$.*

— *Lorsque $\mathbb{P}_X$ est une loi uniforme $\mathcal{U}(0, \theta)$ de paramètre inconnu $\theta > 0$ alors $\theta = 2\mathbb{E}(X_1)$ ainsi la méthode des moments fournit comme estimateur le double de la moyenne aléatoire : $\hat{\theta}^{\mathrm{m}} = 2\overline{X}$.*

Dans la suite, nous introduisons une méthode d'estimation bien plus générale que la méthode des moments : l'estimation via la méthode du maximum de vraisemblance.

2.3 La méthode du maximum de vraisemblance

Définition 2. *On considère un $n-$échantillon $(X_1, ..., X_n)$ d'une loi $\mathbb{P}_X$ paramétrée par $\theta \in \Theta \subset \mathbb{R}^k$.*

— *Lorsque $\mathbb{P}_X$ est une loi discrète alors, la vraisemblance est*

$$L(\theta, x_1, \ldots, x_n) = \prod_{i=1}^{n} \underbrace{\mathbb{P}(X_i = x_i, \theta)}_{=\mathbb{P}_X(x_i,\theta)}.$$

— *Lorsque $\mathbb{P}_X$ est une loi continue de densité $f_X(., \theta)$, la vraisemblance est définie par*

$$L(\theta, x_1, \ldots, x_n) = \prod_{i=1}^{n} f_X(x_i, \theta).$$

Dans le cas discret, la vraisemblance $L(\theta, x_1, \ldots, x_n)$ s'interprète comme la probabilité d'observer l'échantillon expérimental $(x_1, \ldots, x_n)$ lorsque θ est la vraie valeur du paramètre.

Définition 3. *Soit $(x_1, \ldots, x_n)$ un $n-$échantillon expérimental. Une estimation du maximum de vraisemblance est une valeur $\hat{\theta}^{\mathrm{mv}}$ telle que :*

$$L(\hat{\theta}^{\mathrm{mv}}, x_1, \ldots, x_n) = \sup \{L(\theta, x_1, \ldots, x_n) : \theta \in \Theta\}.$$

Sous réserve d'existence et d'unicité, la valeur expérimentale de l'estimateur du maximum de vraisemblance $\hat{\theta}^{\mathrm{mv}}$ est une fonction du $n-$échantillon expérimental notée $g(x_1, \ldots, x_n)$. L'estimateur du maximum de vraisemblance est la statistique $\hat{\theta}^{\mathrm{mv}} = g(X_1, \ldots, X_n)$. Très souvent on maximise le logarithme de la fonction de vraisemblance (appelée log-vraisemblance) transformant les produits en sommes afin de simplifier les calculs :

$$\ln(L(\theta, x_1, \ldots, x_n)) = \begin{cases} \sum_{i=1}^{n} \ln(\mathbb{P}_X(x_i, \theta)) & \text{dans le cas discret} \\ \sum_{i=1}^{n} \ln(f_X(x_i, \theta)) & \text{dans le cas continue} \end{cases}.$$

Exemple 2. *Soit $(X_1, \ldots, X_n)$ un $n-$échantillon d'une loi $\mathcal{B}(p)$ avec $p \in [0, 1]$. Pour $x_i \in \{0, 1\}$ on a $\mathbb{P}(X_i = x_i) = 1 - p$ si $x_i = 0$ ou $\mathbb{P}(X_i = x_i) = p$ si $x_i = 1$; ce que l'on reformule de façon concise par $\mathbb{P}(X_i = x_i) = p^{x_i}(1-p)^{1-x_i}$. Ainsi, pour un $n-$échantillon expérimental $(x_1, \ldots, x_n)$, avec $p \in]0, 1[$, la log-vraisemblance est définie par :*

$$\begin{aligned} \ln(L(p, x_1, \ldots, x_n)) &= \sum_{i=1}^{n} \ln\left(p^{x_i}(1-p)^{1-x_i}\right) = \sum_{i=1}^{n} \left(x_i \ln(p) + (1-x_i)\ln(1-p)\right), \\ &= \ln(p) \sum_{i=1}^{n} x_i + \ln(1-p)\left(n - \sum_{i=1}^{n} x_i\right). \end{aligned}$$

Donc, la dérivée de la log-vraisemblance vaut :

$$\frac{\partial L(p, x_1, \ldots, x_n)}{\partial p} = \frac{\sum_{i=1}^{n} x_i}{p(1-p)} - \frac{n}{1-p}.$$

En maximisant la log-vraisemblance, via la dérivée, on remarque que l'estimateur du maximum de vraisemblance est la moyenne aléatoire $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

3 Propriétés des estimateurs

Il y a de nombreuse façons de construire un estimateur pour un paramètre inconnu. A présent, nous allons introduire les notions de biais et d'erreur quadratique moyenne permettant de comparer les estimateurs.

3.1 Biais d'un estimateur

Définition 4. *Soit* $(X_1, \ldots, X_n)$ *un* $n-$*échantillon d'une loi* $\mathbb{P}_X$ *paramétrée par* $\theta \in \Theta \subset \mathbb{R}$ *et* $\widehat{\theta} = g(X_1, \ldots, X_n)$ *un estimateur. Le biais de l'estimateur* $\widehat{\theta}$ *pour* θ *est :* $B_\theta(\widehat{\theta}) = \mathbb{E}(\widehat{\theta}) - \theta$*. On dit qu'un estimateur est sans biais si pour tout* $\theta \in \Theta$ *on a* $B_\theta(\widehat{\theta}) = 0$.

Exemple 3. *Soit* $(X_1, \ldots, X_n)$ *un échantillon d'espérance* $\mu \in \mathbb{R}$ *et de variance* $\sigma^2 \geq 0$.
- *La moyenne aléatoire* $\overline{X}$ *est un estimateur sans biais de* μ.
- *La variance aléatoire* $S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$ *estime* σ^2 *avec un biais : en effet,* $\mathbb{E}(S^2) = \frac{n-1}{n}\sigma^2$.
- *La variance aléatoire corrigée* $S^2_{\text{corr}} = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$ *est un estimateur sans biais pour* σ^2.

Un petit mot d'humour approprié pour cette section :
Deux statisticiens sont à la chasse au canard... Un canard prend sont envol, le premier vise et tire, mais ça passe $5m$ au dessus. Le second vise et tire, mais ça passe $5m$ en dessous. Ils s'exclament : « On l'a eu ! ».

3.2 Erreur quadratique d'un estimateur

Définition 5. *Soit* $(X_1, \ldots, X_n)$ *un* $n-$*échantillon d'une loi* $\mathbb{P}_X$ *paramétrée par* $\theta \in \Theta \subset \mathbb{R}$ *et* $\widehat{\theta} = g(X_1, \ldots, X_n)$ *un estimateur. L'erreur quadratique moyenne de* $\widehat{\theta}$ *pour* θ *est :* $\mathrm{EQM}_\theta(\widehat{\theta}) = \mathbb{E}((\widehat{\theta} - \theta)^2)$.

Proposition 1. *Soit* $(X_1, \ldots, X_n)$ *un* $n-$*échantillon d'une loi* $\mathbb{P}_X$ *paramétrée par* $\theta \in \Theta \subset \mathbb{R}$ *et* $\widehat{\theta} = g(X_1, \ldots, X_n)$ *un estimateur. L'erreur quadratique moyenne de l'estimateur* $\widehat{\theta}$ *pour* θ *admet la décomposition biais-variance*[1] *:* $\mathrm{EQM}_\theta(\widehat{\theta}) = \mathrm{var}(\widehat{\theta}) + B_\theta(\widehat{\theta})^2$.

Démonstration.

$$\begin{aligned}
\mathrm{EQM}_\theta(\widehat{\theta}) &= \mathbb{E}((\widehat{\theta} - \theta)^2) = \mathbb{E}((\widehat{\theta} - \mathbb{E}(\widehat{\theta}) + \mathbb{E}(\widehat{\theta}) - \theta)^2). \\
&= \underbrace{\mathbb{E}((\widehat{\theta} - \mathbb{E}(\widehat{\theta}))^2)}_{=\mathrm{var}(\widehat{\theta})} + \underbrace{(\mathbb{E}(\widehat{\theta}) - \theta)^2}_{=B_\theta(\widehat{\theta})^2} + \underbrace{2\mathbb{E}(\widehat{\theta} - \mathbb{E}(\widehat{\theta}))(\mathbb{E}(\widehat{\theta}) - \theta))}_{=0}
\end{aligned}$$

□

Exemple 4. *Soit* $X_1, \ldots, X_n$ *un échantillon d'une loi de moyenne* $\mu \in \mathbb{R}$ *et de variance* $\sigma^2 > 0$*. La moyenne aléatoire est un estimateur sans biais pour* μ *ainsi l'erreur quadratique moyenne de cet estimateur est :*

$$\mathrm{EQM}_\mu(\overline{X}) = \mathrm{var}(\overline{X}) = \frac{\sigma^2}{n}.$$

3.3 Propriétés asymptotiques d'un estimateur

Définition 6. *Soit* $(X_1, \ldots, X_n)$ *un n-échantillon d'une loi* $\mathbb{P}_X$ *paramétrée par* θ *et* $\widehat{\theta} = g(X_1, \ldots, X_n)$ *un estimateur. On dit que* $\widehat{\theta}$ *est un estimateur convergent de* θ *si* $\widehat{\theta}$ *converge en probabilité vers* θ.

Via la loi des grands nombres on montre que la moyenne aléatoire est un estimateur convergent pour l'espérance ; on montre également que la variance aléatoire et la variance aléatoire corrigée sont des estimateurs convergent pour la variance.

1. Lorsque θ est à paramètre à valeur dans $\mathbb{R}^d$ avec $d \geq 2$ (ce qui est exclu dans ce cours par simplification), le biais est un vecteur de $\mathbb{R}^d$ et la décomposition biais variance pour l'erreur quadratique moyenne est $\mathrm{EQM}_\theta(\widehat{\theta}) = \mathbb{E}(\|\widehat{\theta} - \mathbb{E}(\widehat{\theta})\|_2^2) + \|B_\theta(\widehat{\theta})\|_2^2$.

Définition 7. *Soit $(X_1, \ldots, X_n)$ un n-échantillon d'une loi $\mathbb{P}_X$ paramétrée par θ et $\widehat{\theta} = g(X_1, \ldots, X_n)$ un estimateur. On dit que $\widehat{\theta}$ est asymptotiquement normal si $\sqrt{n}(\widehat{\theta} - \theta)$ converge en loi vers une loi normale.*

D'après le théorème de la limite centrale, la moyenne aléatoire est un estimateur asymptotiquement normal.

4 Estimation par intervalle de confiance

La valeur estimée d'un paramètre ne donne pas d'information sur la précision de cette estimation. Dans cette partie, nous allons construire des intervalles de confiance (dont les bornes sont aléatoires) qui contiennent, avec une probabilité contrôlée, la paramètre d'intérêt.

4.1 Lois satellites de la loi normale et résultats préliminaires

Définition 8 (Loi du khi-deux)**.** *Soit $Z_1, \ldots, Z_n$ une suite de variables aléatoires réelles indépendantes d'une loi normale centrée réduite $\mathcal{N}(0,1)$. La loi du khi-deux à n degrés de liberté, notée $\chi^2(n\,\mathrm{ddl})$, est la loi de la variable aléatoire D définie par :*

$$D = \sum_{i=1}^{n} Z_i^2.$$

Définition 9 (Loi de Student)**.** *Soit Z et D deux variables aléatoires réelles indépendantes où Z suit une loi $\mathcal{N}(0,1)$ et D suit une loi $\chi^2(n\,\mathrm{ddl})$. La loi de Student à n degrés de liberté, notée $St(n\,\mathrm{ddl})$, est la loi de la variable aléatoire $\widehat{\theta}$ définie par :*

$$T = \frac{Z}{\sqrt{D/n}}.$$

Le théorème suivant donne des résultats fondamentaux sur la moyenne aléatoire, la variance aléatoire et la variance aléatoire corrigée dans le cas gaussien.

Théorème 2. *Soit $(X_1, \ldots, X_n)$ un n-échantillon d'une loi $\mathcal{N}(\mu, \sigma^2)$ alors :*

1. *La moyenne aléatoire $\overline{X}$ suit une loi $\mathcal{N}\left(\mu, \sigma^2/n\right)$.*
2. *La statistique*
$$\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} = \frac{nS^2}{\sigma^2} = \frac{(n-1)S^2_{\mathrm{corr}}}{\sigma^2}$$
suit une loi $\chi^2(n-1\,\mathrm{ddl})$.
3. *Les statistiques $\overline{X}$ et S^2 sont indépendantes.*
4. *La statistique*
$$\frac{\overline{X} - \mu}{\sqrt{S^2_{\mathrm{corr}}/n}}$$
suit une loi $St(n-1\,\mathrm{ddl})$.

4.2 Estimation par intervalle de confiance pour un échantillon gaussien

Théorème 3. *Soit $(X_1, \ldots, X_n)$ un n-échantillon d'une loi $\mathcal{N}(\mu, \sigma^2)$ et $\alpha \in]0,1[$.*

Intervalle de confiance pour la moyenne : variance connue *Soit $z_{1-\alpha/2}$ le $1-\alpha/2$ quantile d'une loi normale centrée réduite $\mathcal{N}(0,1)$ alors*

$$\mathbb{P}\left(\overline{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

Intervalle de confiance pour la moyenne : variance inconnue *Soit $t_{1-\alpha/2}$ le $1-\alpha/2$ quantile d'une loi $St(n-1\,\mathrm{ddl})$ alors*

$$\mathbb{P}\left(\overline{X}-t_{1-\alpha/2}\sqrt{\frac{S^2_{corr}}{n}}\leq \mu \leq \overline{X}+t_{1-\alpha/2}\sqrt{\frac{S^2_{corr}}{n}}\right)=1-\alpha.$$

Intervalle de confiance pour la variance *Soit $x_{\alpha/2}$ et $x_{1-\alpha/2}$ les $\alpha/2$ et $1-\alpha/2$ quantiles d'une loi $\chi^2(n-1\,\mathrm{ddl})$ alors*

$$\mathbb{P}\left(\frac{\sum_{i=1}^n(X_i-\overline{X})^2}{x_{1-\alpha/2}}\leq \sigma^2 \leq \frac{\sum_{i=1}^n(X_i-\overline{X})^2}{x_{\alpha/2}}\right)=1-\alpha.$$

Chacun des résultats donnés au théorème précédent donne un interval aléatoire (les bornes dépendent du $n-$échantillon $(X_1,\dots,X_n)$) qui contient le parmètre d'intérêt μ ou σ^2 avec une probabilité contrôlée au niveau $1-\alpha$. Ainsi, à partir d'un échantillon expérimental $(x_1,\dots,x_n)$ on construit les intervalles de confiance expérimentaux suivants :

— Lorsque la variance est connue, l'intervalle de confiance expérimental pour μ de niveau $1-\alpha$ est :

$$I_{1-\alpha}(\mu)=\left[\overline{X}^{exp}-z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\overline{X}^{\mathrm{exp}}+z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right],\ \text{avec } \overline{X}^{\mathrm{exp}}=\frac{1}{n}\sum_{i=1}^n x_i.$$

— Lorsque la variance est inconnue, l'intervalle de confiance expérimental pour μ de niveau $1-\alpha$ est :

$$I_{1-\alpha}(\mu)=\left[\overline{X}^{\mathrm{exp}}-t_{1-\alpha/2}\sqrt{\frac{S^{2,\mathrm{exp}}_{\mathrm{corr}}}{n}},\overline{X}^{\mathrm{exp}}+t_{1-\alpha/2}\sqrt{\frac{S^{2,\mathrm{exp}}_{\mathrm{corr}}}{n}}\right],\ \text{avec } S^{2,\mathrm{exp}}_{\mathrm{corr}}=\frac{1}{n-1}\sum_{i=1}^n(x_i-\overline{X}^{\mathrm{exp}})^2.$$

— L'intervalle de confiance expérimental pour σ^2 de niveau $1-\alpha$ est :

$$I_{1-\alpha}(\sigma^2)=\left[\frac{\sum_{i=1}^n(x_i-\overline{X}^{\mathrm{exp}})^2}{x_{1-\alpha/2}},\frac{\sum_{i=1}^n(x_i-\overline{X}^{\mathrm{exp}})^2}{x_{\alpha/2}}\right].$$

4.3 Constructions d'intervalles de confiance asymptotiques

Théorème 4. *Soit $(X_1,\dots,X_n)$ un $n-$échantillon d'une loi $\mathcal{B}(p)$, $z_{1-\alpha/2}$ le $1-\alpha/2$ quantile d'une loi normale centrée réduite $\mathcal{N}(0,1)$ alors :*

$$\lim_{n\to+\infty}\mathbb{P}\left(\overline{X}-z_{1-\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\leq p\leq \overline{X}+z_{1-\alpha/2}\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\right)=1-\alpha.$$

Ce théorème donne un intervalle aléatoire (les bornes dépendent du $n-$échantillon $(X_1,\dots,X_n)$) qui contient le paramètre p avec une probabilité asymptotique de niveau $1-\alpha$. Ainsi, à partir d'un échantillon expérimental $(x_1,\dots,x_n)$ on construit l'intervalle de confiance expérimental suivant :

$$I_{1-\alpha}(p)=\left[\overline{X}^{\mathrm{exp}}-z_{1-\alpha/2}\sqrt{\frac{\overline{X}^{\mathrm{exp}}(1-\overline{X}^{\mathrm{exp}})}{n}},\overline{X}^{\mathrm{exp}}+z_{1-\alpha/2}\sqrt{\frac{\overline{X}^{\mathrm{exp}}(1-\overline{X}^{\mathrm{exp}})}{n}}\right]$$

Une règle d'usage pour utiliser un intervalle de confiance asymptotique est : $n\overline{X}^{\mathrm{exp}}(1-\overline{X}^{\mathrm{exp}})\geq 5$.

4.4 Intervalle de confiance asymptotique pour une moyenne

Théorème 5. *Soit $(X_1, \dots, X_n)$ un $n-$échantillon d'espérance μ et de variance σ^2 et $z_{1-\alpha/2}$ le $1-\alpha/2$ quantile d'une loi normale centrée réduite $\mathcal{N}(0,1)$ alors*

$$\lim_{n\to+\infty} \mathbb{P}\left(\overline{X} - z_{1-\alpha/2}\sqrt{\frac{S^2_{\text{corr}}}{n}} \leq \mu \leq \overline{X} + z_{1-\alpha/2}\sqrt{\frac{S^2_{\text{corr}}}{n}}\right) = 1-\alpha.$$

Ce théorème donne un intervalle aléatoire (les bornes dépendent du $n-$échantillon $(X_1, \dots, X_n)$) qui contient le paramètre μ avec une probabilité asymptotique de niveau $1-\alpha$. Ainsi, à partir d'un échantillon expérimental $(x_1, \dots, x_n)$ on construit l'intervalle de confiance expérimental suivant :

$$I_{1-\alpha}(\mu) = \left[\overline{X}^{\text{exp}} - z_{1-\alpha/2}\sqrt{\frac{S^{2,\text{exp}}_{\text{corr}}}{n}}, \overline{X}^{\text{exp}} + z_{1-\alpha/2}\sqrt{\frac{S^{2,\text{exp}}_{\text{corr}}}{n}}\right].$$

Cet intervalle de confiance de niveau asymptotique $1-\alpha$ reste valide lorsqu'on substitue la variance aléatoire corrigée S^2_{corr} par la variance aléatoire S^2. En effet, S^2_{corr} et S^2 sont des estimateurs convergents pour σ^2.