Correction feuille de TD 6

2IMACS, 2011 – 2012

$\boxed{\text{Exc. 1}}$ $X \sim \mathcal{N}(m, \sigma^2)$.

1) $\sigma = 20$

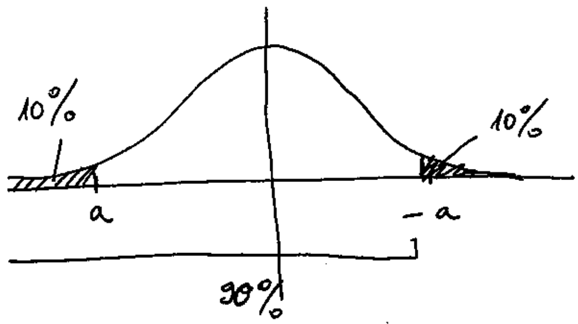
$\mathbb{P}(X \geq 750) \geq 90\% \iff \mathbb{P}(X < 750) \leq 10\%$

$Z = \dfrac{X-m}{\sigma} \sim \mathcal{N}(0,1)$

On veut $\mathbb{P}\left(\dfrac{X-m}{\sigma} < \dfrac{750-m}{\sigma}\right) \leq 10\%$

$\iff \mathbb{P}\left(Z < \underbrace{\dfrac{750-m}{20}}_{a}\right) \leq 10\%$, avec $Z \sim \mathcal{N}(0,1)$

$\Rightarrow a < 0$



$\mathbb{P}(Z < a) = \mathbb{P}(Z > -a)$ par symétrie

$\Rightarrow \mathbb{P}(Z \leq -a) = 1 - \mathbb{P}(Z > -a) \geq 90\%$

$F_Z(-a) \geq 0{,}9 \Rightarrow -a \geq 1{,}29$

avec la table de la loi $\mathcal{N}(0,1)$

$\Rightarrow \dfrac{m-750}{20} \geq 1{,}29 \Rightarrow \boxed{m \geq} 750 + 20 \times 1{,}29 = \boxed{775{,}8 \text{ ml}}$

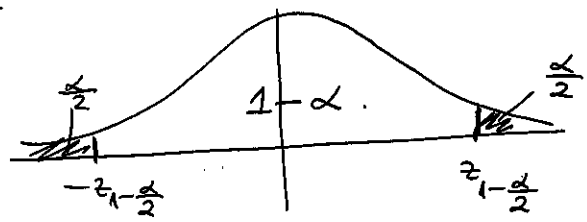
2) $\overline{X}_n = \dfrac{X_1 + \ldots + X_n}{n} \sim \mathcal{N}\left(m, \dfrac{\sigma^2}{n}\right)$, si X_i sont indép. de loi $\mathcal{N}(m, \sigma^2)$

$\Rightarrow \dfrac{\sqrt{n}(\overline{X}_n - m)}{\sigma} \sim \mathcal{N}(0,1)$

a) $\boxed{\sigma \text{ connu}}$, $\sigma = 20$

Si $Z \sim \mathcal{N}(0,1) \Rightarrow \mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$

avec $z_{1-\frac{\alpha}{2}}$ le quantile d'ordre $1-\frac{\alpha}{2}$ de la $\mathcal{N}(0,1)$



car $F_Z\left(z_{1-\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2} \Rightarrow \mathbb{P}\left(Z > z_{1-\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$

et $\mathbb{P}\left(Z < -z_{1-\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$ (par symétrie) donc $\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\overline{X}_n - m)}{\sigma} \leq z_{1-\frac{\alpha}{2}} \right) = 1-\alpha$

$$\Leftrightarrow \mathbb{P}\Big(m \in \underbrace{\left[\overline{X}_n \pm \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right]}_{IC_{1-\alpha}(m)} \Big) = 1-\alpha .$$

intervalle de confiance pour m de niveau de confiance $1-\alpha$

Ici $1-\alpha = 90\%$ $\Rightarrow 1-\frac{\alpha}{2} = 95\% \Rightarrow z_{1-\frac{\alpha}{2}} \cong 1{,}65$
$\alpha = 10\%$

$$\Rightarrow IC_{90\%}(m) = \left[\overline{X}_n \pm \frac{\sigma}{\sqrt{n}} \cdot 1{,}65\right]$$

Avec les observations on obtient

$$IC_{90\%}(m) = \left[760 \pm \frac{20}{\sqrt{51}} \times 1{,}65\right] = \underline{[755{,}38 \; ; 764{,}62]}$$

b) Pour $\boxed{\sigma \text{ inconnu}}$ on utilise le fait que

$$\frac{\sqrt{n}(\overline{X}_n - m)}{S_n} \sim \mathcal{T}(n-1)$$

loi de Student à $n-1$ degrés de liberté

où $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$

est la variance empirique
(estimateur consistant et sans biais de la variance)

$$\Rightarrow \mathbb{P}\left(-t_{n-1;1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\overline{X}_n - m)}{S_n} \leq t_{n-1;1-\frac{\alpha}{2}} \right) = 1-\alpha$$

(comme pour la la loi $\mathcal{N}(0,1)$)

le quantile d'ordre $1-\frac{\alpha}{2}$ de la loi $\mathcal{T}(n-1)$.

$$\Rightarrow \mathbb{P}\Big(m \in \underbrace{\left[\overline{X}_n \pm \frac{S_n}{\sqrt{n}} t_{n-1;1-\frac{\alpha}{2}}\right]}_{IC_{1-\alpha}(m)} \Big) = 1-\alpha$$

intervalle de confiance pour m de niveau de conf. $1-\alpha$

Application numérique :

$\overline{x}_n = 760$, $s_n = 24{,}2$, $n = 51$, $t_{50;95\%} = 1{,}6759$ (avec la table à la fin du poly)

$$\Rightarrow IC_{90\%}(m) = \underline{[754{,}32 \; ; 765{,}68]}$$

$\boxed{\text{Exc. 2}}$

1) $\mathbb{P}\left(|\hat{p}_n - p| \leq 1\%\right) \geq 95\%$, $\hat{p}_n = \dfrac{X_1 + \dots + X_n}{n}$, avec X_i indép. de loi $\mathrm{Ber}(p)$.

a) $$\mathbb{E}(X_i) = p \implies \mathbb{E}(\hat{p}_n) \overset{\text{linéarité}}{=} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = p$$

$$\implies \mathbb{P}\left(|\hat{p}_n - p| \leq 1\%\right) = \mathbb{P}\left(|\hat{p}_n - \mathbb{E}(\hat{p}_n)| \leq 1\%\right) = 1 - \underbrace{\mathbb{P}\left(|\hat{p}_n - \mathbb{E}(\hat{p}_n)| > 1\%\right)}_{\leq \frac{\mathrm{Var}(\hat{p}_n)}{(0{,}01)^2}}$$

inég. de Tchebychev

donc $\mathbb{P}\left(|\hat{p}_n - p| \leq 1\%\right) \geq 1 - \mathrm{Var}(\hat{p}_n) \cdot 10^4$.

or $\mathrm{Var}(\hat{p}_n) = \dfrac{1}{n^2} \sum \mathrm{Var}(X_i) = \dfrac{1}{n^2} n p(1-p) = \dfrac{p(1-p)}{n}$.

d'où $\mathbb{P}\left(|\hat{p}_n - p| \leq 1\%\right) \geq 1 - \dfrac{p(1-p)}{n} \cdot 10^4$.

On va choisir n t.q. $1 - \dfrac{p(1-p)}{n} \cdot 10^4 \geq 95\%$

$$\iff \frac{p(1-p)}{n} \cdot 10^4 \leq 0{,}05$$

$$\iff n \geq p(1-p) \cdot \frac{10^6}{5}$$

Mais p est inconnu, donc on va prendre le "pire" des cas :

$p(1-p) \leq \frac{1}{4}$ (démo par $p^2 - p + \frac{1}{4} = (p - \frac{1}{2})^2 \geq 0$)

↑ égalité pour $p = \frac{1}{2}$

$$n \geq \frac{1}{4} \cdot \frac{10^6}{5} = \frac{1}{2} \cdot 10^5 = \boxed{50000}.$$

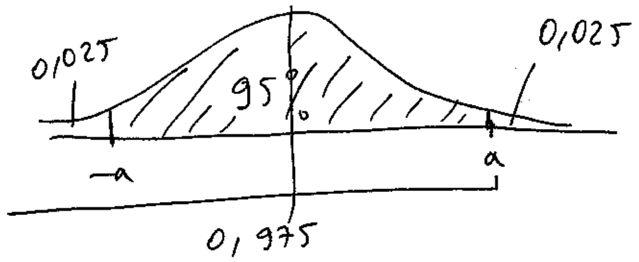
b) Par le TCL on a $\dfrac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{p(1-p)}} \xrightarrow[n \to \infty]{\text{loi}} \mathcal{N}(0,1)$.

car les X_i sont indép. de même loi et admettant une variance

$$\implies \mathbb{P}\left(|\hat{p}_n - p| \leq 1\%\right) = \mathbb{P}\left(\left|\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{p(1-p)}}\right| \leq \frac{\sqrt{n} \cdot 1\%}{\sqrt{p(1-p)}}\right) \underset{n \text{ grand}}{\approx} \mathbb{P}\left(|Z| \leq \frac{\sqrt{n} \cdot 1\%}{\sqrt{p(1-p)}}\right)$$

avec $Z \sim \mathcal{N}(0,1)$.

On va donc choisir n t.q. $\mathbb{P}\Big(|Z| \le \underbrace{\dfrac{1\%\sqrt{n}}{\sqrt{p(1-p)}}}_{a}\Big) \ge 95\%$



$$\mathbb{P}(-a < Z < a) \ge 95\% \Rightarrow \mathbb{P}(Z>a) + \underbrace{\mathbb{P}(Z<-a)}_{\substack{\parallel \\ \mathbb{P}(Z>a) \text{ par sym.}}} \le 5\%$$

$$\Rightarrow \mathbb{P}(Z>a) \le 0{,}025$$

$$\Rightarrow F_Z(a) = 1-\mathbb{P}(Z>a) \ge 0{,}975 \xRightarrow[\text{avec la table}]{} \underline{a \ge 1{,}96}.$$

$$\Rightarrow \frac{1\%\sqrt{n}}{\sqrt{p(1-p)}} \ge 1{,}96 \Rightarrow n \ge (1{,}96)^2 \cdot p(1-p) \cdot 10^4$$

Comme $p(1-p) \le \frac{1}{4}$,
on va prendre $n \ge (1{,}96)^2 \cdot \frac{1}{4} \cdot 10^4 = \boxed{9604}$.

(bcp. plus petite que la valeur donnée par Tchebychev ...)

2) $n = 100$, $\hat{p}_n = 51\%$

On a, par TCL : $\dfrac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{p(1-p)}} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$.

On a aussi, car $\hat{p}_n \xrightarrow[n\to\infty]{\mathbb{P}} p$ (estimateur consistant)

$$\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1).$$

$$\Rightarrow \mathbb{P}\left(-1{,}96 \le \frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} \le 1{,}96\right) \xrightarrow[n\to\infty]{} \mathbb{P}(-1{,}96 \le Z \le 1{,}96) = 0{,}95$$

la même valeur qu'avant
(le quantile d'ordre 0,975)
$z_{0,975}$

$$\Rightarrow \mathbb{P}\Big(p \in \overbrace{\Big[\hat{p}_n \pm \sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \cdot 1{,}96\Big]}^{IC_{0,95}(p)}\Big) \xrightarrow[n\to\infty]{} 0{,}95$$

(intervalle de confiance asymptotique)

Application numérique :

$$IC_{0,95}(p) = \left[0,51 \pm \sqrt{\frac{0,51 \times 0,49}{100}} \times 1,96\right] = \underline{\left[0,412\ ;\ 0,608\right]}$$

3) pour $n = 10000$ on obtient

$$IC_{0,95}(p) = \underline{\left[0,5002\ ;\ 0,5198\right]}$$

("on est sûr à 95% que le candidat va gagner")

$\boxed{\text{Exc. 3}}$ $X \sim \text{Expo}\left(\frac{1}{\theta}\right)$.

1) $\mathbb{E}(X) = \theta$, $\text{Var}(X) = \theta^2$

2) Comme $\mathbb{E}(X) = \theta$, on va proposer comme estimateur la moyenne empirique $\overline{X_n} = \frac{X_1 + \ldots + X_n}{n}$ et on sait que c'est un estimateur

- consistant : car par la loi des grands nombres, $\overline{X_n} \xrightarrow[n\to\infty]{\mathbb{P}} \mathbb{E}(X_1) = \theta$
- sans biais, car $\mathbb{E}(\overline{X_n}) = \theta$.

3) Par le $\boxed{\text{TCL}}$, comme les X_i sont indép., de même loi et admettent une variance,

$$\frac{\sqrt{n}\,(\overline{X_n} - \mathbb{E}(X_1))}{\sqrt{\text{Var}(X_1)}} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$$

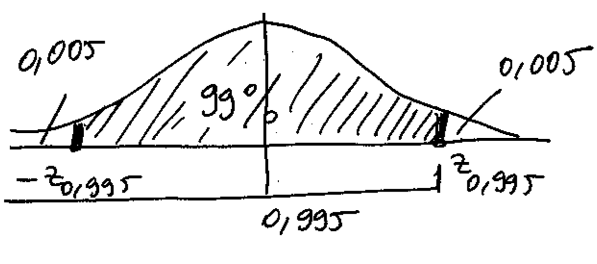
$$\Leftrightarrow \frac{\sqrt{n}\,(\overline{X_n} - \theta)}{\theta} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1).$$

4) $\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}\,(\overline{X_n} - \theta)}{\theta} \leq z_{1-\frac{\alpha}{2}}\right) \xrightarrow[n\to\infty]{} \underbrace{\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}\right)}_{= 1-\alpha}$

avec $Z \sim \mathcal{N}(0,1)$

Ici $1-\alpha = 99\%$; et avec la table on trouve

$$z_{1-\frac{\alpha}{2}} = z_{0,995} \simeq 2,58$$



$\Rightarrow \quad \mathbb{P}\left(-2{,}58 \leq \dfrac{\sqrt{n}\,(\overline{X}_n-\theta)}{\theta} \leq 2{,}58\right) \underset{n \text{ grand}}{\approx} 0{,}99.$

$\Rightarrow \quad \mathbb{P}\left(\dfrac{\sqrt{n}\,\overline{X}_n}{\sqrt{n}+2{,}58} \leq \theta \leq \dfrac{\sqrt{n}\,\overline{X}_n}{\sqrt{n}-2{,}58}\right) \simeq 0{,}99$

$$\boxed{IC_{0{,}99}(\theta) = \left[\frac{\sqrt{n}\,\overline{X}_n}{\sqrt{n}+2{,}58} \,;\, \frac{\sqrt{n}\,\overline{X}_n}{\sqrt{n}-2{,}58}\right]}$$

(intervalle de confiance asymptotique pour θ de niveau de conf. 0,99)

Exo. 4 $\quad X_1, \dots, X_n \sim \mathcal{U}([0,\theta])$.

1) $\mathbb{E}(X_i) = \dfrac{\theta}{2}$, $\quad \mathrm{Var}(X_i) = \dfrac{\theta^2}{12}$ (faits dans l'exo sur la loi uniforme, exo 1, TD 3)

2) $\hat{\theta}_n = 2\,\overline{X}_n$

- $\hat{\theta}_n$ estimateur consistant : m.q. $\hat{\theta}_n \xrightarrow[n\to\infty]{\mathbb{P}} \theta$.

Soit $\varepsilon > 0$; $\mathbb{P}\left(|\hat{\theta}_n - \theta| > \varepsilon\right) = \mathbb{P}\left(|2\overline{X}_n - \theta| > \varepsilon\right)$

$= \mathbb{P}\left(\left|\overline{X}_n - \dfrac{\theta}{2}\right| > \dfrac{\varepsilon}{2}\right) \xrightarrow[n\to\infty]{} 0$

car $\overline{X}_n \xrightarrow[n\to\infty]{\mathbb{P}} \dfrac{\theta}{2} = \mathbb{E}(X_1)$ par la LGN.

- $\hat{\theta}_n$ estimateur non-biaisé : $\mathbb{E}(\hat{\theta}_n) = 2\,\mathbb{E}(\overline{X}_n) \underset{\text{linéarité}}{=} 2\,\mathbb{E}(X_1) = 2 \cdot \dfrac{\theta}{2} = \underline{\theta}$.

3) Par le TCL, $\overline{X}_n \underset{n \text{ grand}}{\overset{\text{loi}}{\approx}} \mathcal{N}\left(m, \dfrac{\sigma^2}{n}\right)$,

car les X_i sont indép de même loi et admettant une variance

avec $m = \mathbb{E}(X_1) = \dfrac{\theta}{2}$, $\sigma^2 = \mathrm{Var}(X_1) = \dfrac{\theta^2}{12}$

$\Rightarrow \overline{X}_n \underset{n \text{ grand}}{\overset{\text{loi}}{\approx}} \mathcal{N}\left(\dfrac{\theta}{2}, \dfrac{\theta^2}{12n}\right)$

$\Rightarrow \hat{\theta}_n \underset{n \text{ grand}}{\overset{\text{loi}}{\approx}} \mathcal{N}\left(\theta, 4 \cdot \dfrac{\theta^2}{12n}\right) = \boxed{\mathcal{N}\left(\theta, \dfrac{\theta^2}{3n}\right)}$

4) On a $\hat{\theta}_n \overset{loi}{\approx} \mathcal{N}\left(\theta, \frac{\theta^2}{3n}\right) \Rightarrow \frac{\hat{\theta}_n - \theta}{\sqrt{\frac{\theta^2}{3n}}} \approx \mathcal{N}(0,1)$

$$\Rightarrow \mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\hat{\theta}_n - \theta}{\frac{\theta}{\sqrt{3n}}} \leq z_{1-\frac{\alpha}{2}}\right) \approx \mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

avec $Z \sim \mathcal{N}(0,1)$.

Ici $1-\alpha = 0{,}95 \Rightarrow z_{1-\frac{\alpha}{2}} = z_{0{,}975} = 1{,}96$

$$\Rightarrow \mathbb{P}\left(-1{,}96 \leq \frac{\sqrt{3n}\left(\hat{\theta}_n - \theta\right)}{\theta} \leq 1{,}96\right) \underset{n \text{ grand}}{\approx} 0{,}95$$

$$\Rightarrow \mathbb{P}\left(\frac{\hat{\theta}_n}{1 + \frac{1{,}96}{\sqrt{3n}}} \leq \theta \leq \frac{\hat{\theta}_n}{1 - \frac{1{,}96}{\sqrt{3n}}}\right) \approx 0{,}95$$

$$\boxed{IC_{0{,}95}(\theta) = \left[\frac{\hat{\theta}_n}{1 + \frac{1{,}96}{\sqrt{3n}}} ; \frac{\hat{\theta}_n}{1 - \frac{1{,}96}{\sqrt{3n}}}\right]}$$

intervalle de confiance asympt. au niveau 0,95