Correction Feuille 4
IC3 Probas - Stats

Exo. 1 $\begin{cases} X \sim \mathcal{P}(\lambda) \\ Y \sim \mathcal{P}(\mu) \end{cases}$ $X \perp\!\!\!\perp Y$. $\begin{cases} U \sim \mathcal{N}(0;1) \\ V \sim \mathcal{N}(0;1) \end{cases}$ $U \perp\!\!\!\perp V$

1) $Z = X+Y$ à valeurs dans $\mathbb{N}$. Soit $k \in \mathbb{N}$.

$$\mathbb{P}(Z=k) = \mathbb{P}(X+Y=k) = \mathbb{P}\left(\bigcup_{l=0}^{k} \{X=k-l ; Y=l\}\right)$$

car union disjointe

$$= \sum_{l=0}^{k} \mathbb{P}(X=k-l ; Y=l)$$

$$= \sum_{l=0}^{k} \mathbb{P}(X=k-l)\,\mathbb{P}(Y=l) \quad \text{car } X \perp\!\!\!\perp Y$$

$$= e^{-(\lambda+\mu)} \sum_{l=0}^{k} \frac{\lambda^{k-l}}{(k-l)!} \cdot \frac{\mu^{l}}{l!}$$

$$= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{l=0}^{k} C_k^l \lambda^{k-l} \mu^{l} = \frac{e^{-(\lambda+\mu)} (\lambda+\mu)^k}{k!}$$

donc $Z \sim \mathcal{P}(\lambda+\mu)$.

2) $T = \frac{U^2+V^2}{2}$, à valeurs dans $\mathbb{R}_+$ - Soit $t \geq 0$.

$$\mathbb{P}(T<t) = \mathbb{P}(U^2+V^2<2t) = \iint_{\{u^2+v^2<2t\}} e^{-\frac{(u^2+v^2)}{2}} \frac{du\,dv}{2\pi} \quad \text{car } U \perp\!\!\!\perp V$$

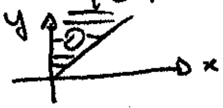
$$= \int_{\theta=0}^{2\pi} \int_{r=0}^{\sqrt{2t}} r e^{-r^2/2} \,dr \frac{d\theta}{2\pi} = \int_0^{\sqrt{2t}} r e^{-r^2/2} dr = 1 - e^{-\frac{(\sqrt{2t})^2}{2}}$$

$$= 1 - e^{-t}$$

Ainsi, T est une v.a.r. continue dont la densité est donnée par

$$f(t) = \begin{cases} e^{-t} & \text{si } t \geq 0 \\ 0 & \text{sinon} \end{cases} \quad \text{donc } T \sim \mathcal{E}xp(1).$$

Exo. 2 $f(x;y) = \begin{cases} e^{-y} & \text{si } (x,y) \in D \\ 0 & \text{sinon} \end{cases}$ avec $D = \{(x,y) \in \mathbb{R}^2 : y \geq x\}$



1) f densité sur $\mathbb{R}^2 \iff \begin{cases} f \geq 0 & \text{ok} \\ \iint_{\mathbb{R}^2} f(x;y)\,dx\,dy = 1 \end{cases}$

$$\iint_{\mathbb{R}^2} f(x;y)\,dx\,dy = \iint_D e^{-y}\,dx\,dy = \int_{x=0}^{+\infty} \int_{y=x}^{+\infty} e^{-y}\,dy\,dx = 1.$$

2) Soit f_X (resp. f_Y) la densité de X (resp. Y).

- si $x<0$ alors $f_X(x)=0$; si $y<0$ alors $f_Y(y)=0$.

- si $x \geq 0$: $f_X(x) = \int_{-\infty}^{+\infty} f(x;y)\,dy = \int_x^{+\infty} e^{-y}\,dy = e^{-x}$

$$\text{donc } f_X(x) = \begin{cases} e^{-x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

- si $y \geq 0$: $f_Y(y) = \int_{-\infty}^{+\infty} f(x;y)\,dx = \int_0^y e^{-y}\,dx = y e^{-y}$

$$\text{donc } f_Y(y) = \begin{cases} y e^{-y} & \text{si } y \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Donc on a clairement que $f(x;y) \neq f_X(x) f_Y(y)$ pour au moins un x et un y donc pas d'indépendance entre X et Y.

<u>Exo.3</u> Soit $X_i = \begin{cases} 1 & \text{si la ième personne va voter ce candidat} \\ 0 & \text{sinon} \end{cases}$
$(i = 1, 2, \dots, N)$

Les (X_i) sont $\perp\!\!\!\perp$ et suivent la loi $Ber(p)$: Ainsi, par la LGN,

$$\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i \xrightarrow[N \to +\infty]{(P)} \mathbb{E}[X_1] = p.$$

Par ailleurs, le TLC nous dit que $\sqrt{N}\left(\frac{\bar{X}_N - p}{\sigma}\right) \xrightarrow[N\to+\infty]{(loi)} Z \sim \mathcal{N}(0;1)$.

En particulier $\mathbb{P}\left(\left|\sqrt{N}\left(\frac{\bar{X}_N - p}{\sigma}\right)\right| \geq \varepsilon\right)$ est proche de $\mathbb{P}(|Z| \geq \varepsilon)$ dès que N est assez grand (en pratique $N \geq 30$), $\forall \varepsilon > 0$.

D'où

$$\mathbb{P}(|\bar{X}_N - p| \geq 1\%) = \mathbb{P}\left(\left|\sqrt{N}\left(\frac{\bar{X}_N - p}{\sigma}\right)\right| \geq \frac{0{,}01\sqrt{N}}{\sigma}\right)$$

$$\underset{(TLC)}{\approx} \mathbb{P}\left(|Z| \geq \frac{0{,}01\sqrt{N}}{\sigma}\right)$$

$$= 2\left(1 - F\left(\frac{0{,}01\sqrt{N}}{\sigma}\right)\right)$$

où F est la fct de répartition de Z.

donc $\mathbb{P}(|\bar{X}_N - p| \geq 0{,}01) \leq 0{,}05 \iff 1 - F\left(\frac{0{,}01\sqrt{N}}{\sigma}\right) \leq 0{,}025.$

$$\iff 0{,}975 \leq F\left(\frac{0{,}01\sqrt{N}}{\sigma}\right)$$

$$\underset{(table)}{\iff} 1{,}96 \leq \frac{0{,}01\sqrt{N}}{\sigma}$$

$$\iff N \geq 38416\,\sigma^2$$

Or $X_1 \sim Ber(p)$ donc $\sigma^2 = Var(X_1) = p(1-p) \leq \frac{1}{4}\ \forall p \in]0,1[$

donc $\mathbb{P}(|\bar{X}_N - p| \geq 0{,}01) \leq 0{,}05 \iff N \geq 9604.$

(2) EXO.4

$$S_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar{X}_N)^2$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}X_i^2 + \frac{N}{N-1}(\bar{X}_N)^2 - \frac{2N}{N-1}(\bar{X}_N)^2$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}X_i^2 - \frac{N}{N-1}(\bar{X}_N)^2$$

$$= \frac{N}{N-1}\cdot\frac{1}{N}\sum_{i=1}^{N}Y_i - \frac{N}{N-1}(\bar{X}_N)^2 \quad \text{avec } Y_i = X_i^2$$

Par la LGN : $\bar{X}_N \xrightarrow[N\to+\infty]{P} m$ et $\frac{1}{N}\sum_{i=1}^{N}Y_i \xrightarrow[N\to+\infty]{P} \mathbb{E}(Y_1) = \mathbb{E}(X_1^2) = \sigma^2+m^2$

(car les (Y_i) sont $\perp\!\!\!\perp$ et de même loi)

En utilisant l'indication, on obtient que :

$$S_N^2 = \frac{N}{N-1}\cdot\frac{1}{N}\sum_{i=1}^{N}Y_i - \frac{N}{N-1}(\bar{X}_N)^2$$

$$\xrightarrow[N\to+\infty]{P} 1\times(\sigma^2+m^2) - 1\times m^2 = \sigma^2.$$

donc S_N^2 est un estimateur de σ^2.

EXO.5 IC à 95% pour la proportion $p \in]0,1[$, paramètre inconnu

Données observées sur l'échantillon de taille $N=200$: $\bar{x}_N = 0,51$.

donc $IC_{0,95}(p) = \left[\bar{x}_N \pm z_{0,975}\times\sqrt{\frac{\bar{x}_N(1-\bar{x}_N)}{N}}\right]$

$$= \left[0,51 \pm 1,96\times\frac{\sqrt{0,51\times0,49}}{10}\right] = [0,42\,;\,0,60]$$

Si $N=1000$, l'erreur est + petite donc l'IC sera plus resserré.

$$IC_{0,95}(p) = \left[0,51 \pm 1,96\times\frac{\sqrt{0,51\cdot0,49}}{31,6}\right] = [0,48\,;\,0,54]$$

Prix à payer pour cette précision : + de personnes à interroger.

EXO.6

1) Par la LGN, la moyenne empirique $\bar{X}_N = \frac{1}{N}\sum_{i=2}^{N}X_i$ est un estimateur de m. Les (X_i) étant $\perp\!\!\!\perp$ et de loi $\mathcal{N}(m;\sigma^2)$, alors $\bar{X}_N \sim \mathcal{N}(m;\frac{\sigma^2}{N})$

2) $IC_{0,95}(m) = \left[\bar{x}_N \pm z_{0,975}\times\frac{\sigma}{\sqrt{N}}\right] = \left[\bar{x}_N \pm 1,96\times\frac{1}{\sqrt{N}}\right]$

3) On cherche N minimal tel que $2\cdot\frac{1,96}{\sqrt{N}} \leq 0,3$

$$\iff N \geq 171.$$