Questions posées en statistique:

① De combien ----> estimation

② Est-ce que le modèle est en adéquation avec mes données ----> Test

Faire des stats revient à répondre aux questions précédentes:

des données, observations, data.

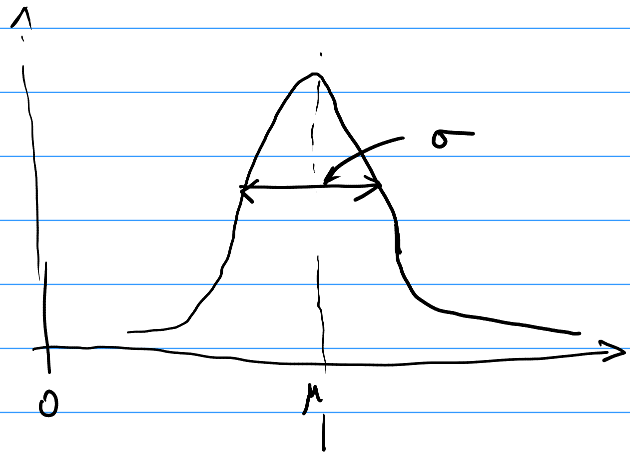
Objet math qui imite

mes données. $\xrightarrow{\text{" on modélise " / " on pose " un modèle}}$ " un mécanisme génératif de ces données " $\longrightarrow$ " loi de proba "

" loi de proba " $\xrightarrow{\text{laquelle}}$ " estimation " $\xrightarrow{\text{Est-ce que}}$ " test "

X "durée"

$$X \leadsto \mathcal{N}(\mu, \sigma^2)$$



$$\mathbb{E}(X) = \mu$$

$$Var(X) = \sigma^2.$$

$X_1, X_2, \dots, X_{10}$

$$\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i = \frac{X_1 + \dots + X_{10}}{10}$$ nous renseigne sur μ

$$S^2 = \frac{1}{(10-1)} \sum_{i=1}^{10} (X_i - \bar{X})^2$$ nous renseigne sur σ^2.

$X_1, \dots, X_{10}$

$$S = f(X_1, \dots, X_{10}) = \boxed{18}$$

X : "temps de passage à la caisse".

$$X \leadsto \mathcal{E}(\lambda)$$

sa densité $f(x) = \lambda e^{-\lambda x}$ $\quad x > 0$

$\lambda > 0$

$$\mathbb{E}(X) = \frac{1}{\lambda}. \qquad Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

$$= \frac{1}{\lambda^2}$$

X_1 : " temps de passage du client 1 "

X_2 : " " " " " 2 "

$$\boxed{T} = X_1 + X_2$$

$$\rightsquigarrow \mathcal{E}(\lambda) \perp \rightsquigarrow \mathcal{E}(\lambda)$$

$$\underline{X_1 \rightsquigarrow \mathcal{E}(\lambda_1)}, \ \underline{X_2 \rightsquigarrow \mathcal{E}(\lambda_2)}, \ldots\ldots, \underline{X_M \rightsquigarrow \mathcal{E}(\lambda_M)}$$

Loi $\left(\min(X_1, \ldots\ldots, X_M)\right) \rightsquigarrow \mathcal{E}(\lambda_1 + \cdots + \lambda_M)$

$$X_1 \rightsquigarrow \mathcal{E}(\lambda), \ \ldots\ldots, \ X_n \rightsquigarrow \underline{\mathcal{E}(\lambda)}$$

$$M_n = \min(X_1, \ldots\ldots, X_n)$$

$$M_n \rightsquigarrow \mathcal{E}(\underbrace{\lambda + \cdots + \lambda}_{n \text{ fois}}) \sim \mathcal{E}(n\lambda). \ ??$$

On cherche la loi de M_n ?

Soit $\underline{x > 0}$

$$\underbrace{F_{M_n}(x) = \mathbb{P}(M_n \leq x)}_{\text{fonction de répartition.}}$$

$$\mathbb{P}(M_n \leq x) = \mathbb{P}\left[\min(X_1, \dots, X_n) \leq x\right]$$

$$= 1 - \mathbb{P}\left[\min(X_1, \dots, X_n) > x\right]$$

$$= 1 - \mathbb{P}\left[\underset{\text{et}}{X_1 > x ,} \ \underset{\text{et}}{X_2 > x ,} \ \underset{\text{et}}{\dots ,} \ X_n > x\right]$$

$$= 1 - \mathbb{P}\left[(X_1 > x) \underset{\text{et}}{\cup} (X_2 > x) \underset{\text{et}}{\cup} \dots \underset{\text{et}}{\cup} (X_n > x)\right]$$

$$= 1 - \left[\mathbb{P}(X_1 > x) \times \mathbb{P}(X_2 > x) \times \dots \times \mathbb{P}(X_n > x)\right]$$

$X \sim \mathcal{E}(\lambda)$ $\begin{cases} f(x) = \lambda e^{-\lambda x} & \text{densité} \\ F(x) = 1 - e^{-\lambda x} & \text{fonction de répartition} \\ S(x) = e^{-\lambda x} & \text{" " survie} \end{cases}$

$F(x) = \mathbb{P}(X \leq x)$ et $S(x) = \mathbb{P}(X > x)$

$$= 1 - \Big(\underbrace{e^{-\lambda x} \times \dots \times e^{-\lambda x}}_{n \text{ fois}}\Big) = 1 - e^{-n\lambda x}$$

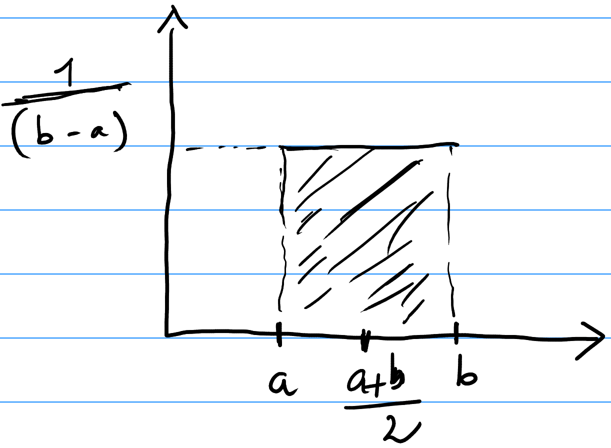
donc $F_{M_n}(x) = 1 - e^{-n\lambda x}$ donc $M_n \sim \mathcal{E}(n\lambda)$

$\underline{M_n} = \underline{\min(X_1, \dots, X_n)}$

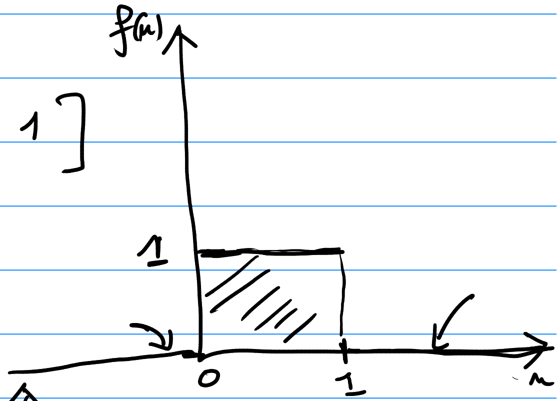
$U_1 \sim$ Uniform$[0,1]$, ..., $U_n \sim$ Uniforme $[0,1]$

Petit rappel (loi uniforme) :

$U \sim$ Uniforme $[a,b]$

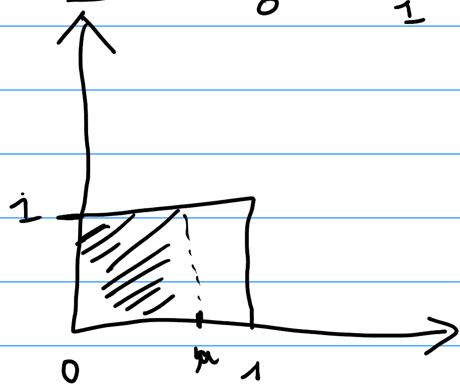


$\mathbb{E}(U) = \frac{a+b}{2}$



$U_1, \dots, U_n \xrightarrow{iid}$ Uniforme $[0,1]$

$$f(x) = \begin{cases} 0 \text{ si } x < 0 \\ 1 \text{ si } 0 \leq x \leq 1 \\ 0 \text{ si } x \geq 1 \end{cases}$$



$$F(x) = \begin{cases} 0 \text{ si } x < 0 \\ x \text{ si } 0 \leq x < 1 \\ 1 \text{ si } x \geq 1 \end{cases}$$

Sa loi

$Y = \max(U_1, \dots, U_n)$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\max(U_1, \dots, U_n) \leq y)$$

$$\mathbb{P}(\max(U_1, \dots, U_n) \leq y) = \mathbb{P}(\text{"}U_1 \leq y\text{" et "}U_2 \leq y\text{" et } \dots \text{ et "}U_n \leq y\text{"})$$

$$= \mathbb{P}\left[(U_1 \leq y) \cap (U_2 \leq y) \cap \dots \cap (U_n \leq y)\right]$$

$$\underset{\text{Indep}}{=} \underbrace{\mathbb{P}[U_1 \leq y]}_{F_U(y)} \times \underbrace{\mathbb{P}[U_2 \leq y]}_{F_U(y)} \times \dots \times \underbrace{\mathbb{P}[U_n \leq y]}_{F_U(y)}$$

$$= \left[F_U(y)\right]^n = [y]^n = F_Y(y)$$

$$f_Y(y) = \frac{d}{dy}\left[F_Y(y)\right] = n\,y^{n-1}$$

Y est à valeur dans $[0,1]$.

$$\int_0^1 f_Y(y)\,dy = \int_0^1 n y^{n-1}\,dy = \left[y^n\right]_0^1 = 1.$$

$$f(x) = \lambda e^{-\lambda x}$$

$$\boxed{f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}}$$

$$\theta = \frac{1}{\lambda}.$$

$X \sim \Gamma(\alpha, \beta)$ si elle est positive :

$$f(x) = \frac{x^{\alpha-1} \beta^{\alpha} e^{-\beta x}}{\Gamma(\alpha)} =$$

$\Gamma(\alpha) \rightarrow$??

si $\alpha = 1$

$$f(x) = \beta e^{-\beta x}$$

$\Gamma(\alpha) =$ " fonction Gamma d'Euler ".

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx.$$

$$\Gamma(1) = \int_0^{+\infty} e^{-x} dx = 1$$

si α est un nombre entier :

$$\Gamma(\alpha) = (\alpha - 1)\,\Gamma(\alpha - 1) = (\alpha-1)(\alpha-2)\,\Gamma(\alpha-2)$$

$$= \dots \dots 2\,\underbrace{\Gamma(1)}_{1}$$

$$= (\alpha - 1)!$$

Erreur quadratique moyenne : MSE

$$\mathbb{E}\left[(\hat{\theta} - \theta)^2\right]$$

$\hat{\theta}$: estimateur

θ : paramètre estimé.

$$\mathbb{E}\left[(\hat{\theta}-\theta)^2\right] = \mathbb{E}\left[\left(\underbrace{\hat{\theta}-\mathbb{E}(\hat{\theta})}_{a} + \underbrace{\mathbb{E}(\hat{\theta})-\theta}_{b}\right)^2\right]$$

$$= \mathbb{E}\left[\underbrace{\left(\hat{\theta}-\mathbb{E}(\hat{\theta})\right)^2}_{a^2} + \underbrace{2\left(\hat{\theta}-\mathbb{E}(\hat{\theta})\right)\left(\mathbb{E}(\hat{\theta})-\theta\right)}_{2ab} + \underbrace{\left(\mathbb{E}(\hat{\theta})-\theta\right)^2}_{b^2}\right]$$

$$= \underbrace{\mathbb{E}\left[\left(\hat{\theta}-\mathbb{E}(\hat{\theta})\right)^2\right]}_{Var(\hat{\theta})} + \mathbb{E}\left[2\left(\hat{\theta}-\mathbb{E}(\hat{\theta})\right)\left(\mathbb{E}(\hat{\theta})-\theta\right)\right] + \mathbb{E}\left[\underbrace{\left(\mathbb{E}(\hat{\theta})-\theta\right)^2}_{\underbrace{\left(\mathbb{E}(\hat{\theta})-\theta\right)^2}_{(\text{biais})^2}}\right]$$

Nous avons :

$$\mathbb{E}\left[2\left(\hat{\theta}-\mathbb{E}(\hat{\theta})\right)\left(\overbrace{\mathbb{E}(\hat{\theta})-\theta}^{\text{biais}}\right)\right] = 2\,\text{biais}\ \mathbb{E}\left[\hat{\theta}-\mathbb{E}(\hat{\theta})\right]$$

$$= 2\,\text{biais}\left(\mathbb{E}(\hat{\theta}) - \mathbb{E}(\mathbb{E}(\hat{\theta}))\right)$$

$$= 2\,\text{biais}\left(\mathbb{E}(\hat{\theta}) - \mathbb{E}(\hat{\theta})\right) = 0$$

Ainsi :

$$\mathbb{E}\left[(\hat{\theta}-\theta)^2\right] = Var(\hat{\theta}) + \underbrace{\left(\mathbb{E}(\hat{\theta})-\theta\right)^2}_{(\text{biais})^2}$$