$\boxed{\text{Cours du 9/11/2022}}$

Ex $\boxed{X_i \sim N(m, \sigma^2)}$ σ^2 connue

$\theta = m$

$$\boxed{\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i}$$

Biais $E[\hat{\theta}_n] - \theta = b_n(\theta)$

$$b_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \underbrace{E[X_i]}_{m} - m = 0$$

$\boxed{\hat{\theta}_n \text{ est non biaisé}}$

Variance

$$\text{Var}(\hat{\theta}_n) = \frac{1}{n^2}\text{Var}\left(\sum X_i\right)$$

$$\underset{X_1,\dots,X_n \text{ ind}}{=} \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{\text{var } X_i}_{\sigma^2}$$

$$= \boxed{\frac{\sigma^2}{n}}$$

$$\tilde{\theta}_n = \frac{2}{n(n+1)}\sum_{i=1}^{n}(i X_i)$$

$$E(\tilde{\theta}_n) = \frac{2}{n(n+1)}\sum_{i=1}^{n} i\,\underbrace{E[X_i]}_{m}$$

$$= \frac{2m}{n(n+1)}\underbrace{\sum_{i=1}^{n} i}_{\frac{n(n+1)}{2}} = m$$

$\boxed{\tilde{\theta}_n \text{ est un estimateur non biaisé de } m}$

$$\text{Var}(\tilde{\theta}_n) = \frac{4}{n^2(n+1)^2}\sum_{i=1}^{n}(i^2\sigma^2)$$

Rappel

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

donc $$\text{var}(\tilde{\theta}_n) = \frac{4\sigma^2}{6}\,\frac{\cancel{n}(\cancel{n+1})(2n+1)}{n^2(n+1)^2}$$

$$= \frac{2\sigma^2}{3}\,\underbrace{\frac{2n+1}{n(n+1)}}_{\sim \frac{2}{n}}$$

$$\approx \boxed{\frac{4}{3}\frac{\sigma^2}{n}} > \frac{\sigma^2}{n}$$

Exemple 2 $\quad X_i \sim N(m, \sigma^2)$ avec σ^2 inconnus

$$\theta = \begin{pmatrix} m \\ \sigma^2 \end{pmatrix}$$

$$\hat\theta = \begin{pmatrix} \hat\theta_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar X \\ \boxed{\hat\theta_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar X)^2} \end{pmatrix}$$

Rq $\quad \frac{1}{n}\sum (X_i - \bar X)^2 = \frac{1}{n}\sum X_i^2 - \left(\frac{1}{n}\sum X_i\right)^2$

Biais de $\hat\theta$

$$\begin{bmatrix} E[\hat\theta_1] - \theta_1 \\ E[\hat\theta_2] - \theta_2 \end{bmatrix} \in \mathbb{R}^2 = \begin{bmatrix} E[\hat\theta_1] - m \\ E[\hat\theta_2] - \sigma^2 \end{bmatrix}$$

on a vu $E[\hat\theta_1] = m$ donc $\boxed{E[\hat\theta_1] - m = 0}$

$$E[\hat\theta_2] \overset{?}{=} \frac{1}{n}\sum_{i=1}^{n} E\Big[\underbrace{(X_i - m + m - \bar X)^2}_{(X_i - m)^2 + (m - \bar X)^2 + 2(X_i - m)(m - \bar X)}\Big]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\Big\{\underbrace{E[(X_i - m)^2]}_{\sigma^2} + \underbrace{E[(m - \bar X)^2]}_{Var \bar X = E[(\bar X - \underbrace{E(\bar X)}_{m})^2] = \frac{\sigma^2}{n}} \boxed{-2E[(X_i - m)(\bar X - m)]}$$

$$\boxed{} \overset{?}{=} -2E\left[(X_i - m)\left(\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} m\right)\right]$$

$$-\frac{2}{n} E\left[\sum_{k=1}^{n}(X_i - m)(X_k - m)\right]$$

$$-\frac{2}{n}\sum_{k=1}^{n}\underbrace{E[(X_i - m)(X_k - m)]}_{Cov(X_i, X_k) = \begin{cases} \sigma^2 & k = i \\ 0 & k \neq i \end{cases}} = \boxed{-\frac{2\sigma^2}{n}}$$

On en déduit

$$E[\hat{\theta}_2] = \sigma^2 + \frac{\sigma^2}{n} - 2\frac{\sigma^2}{n} = \boxed{\sigma^2\left(\frac{n-1}{n}\right)}$$

Cette analyse permet de proposer un nouvel estimateur

$$\tilde{\theta}_2 = \underbrace{\frac{n}{n-1}}\hat{\theta}_2 \Rightarrow E[\tilde{\theta}_2] = \sigma^2$$

donc $\boxed{\tilde{\theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

$$\hat{\sigma}^2 = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 & \text{si } m \text{ ~~in~~ est connue} \\ \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 & \text{si } m \text{ est inconnue} \end{cases}$$

$$\bar{x} = \frac{1}{n}\sum x_i$$

__Vraisemblance__ $\quad X_i \sim N(m, \sigma^2)$ $\quad$ σ^2 connue $\quad \boxed{\theta = m \in \mathbb{R}}$

$$L(x_1, \dots, x_n ; \theta) = p(x_1, \dots, x_n ; \theta)$$

$$\underset{X_1, \dots, X_n \text{ ind}}{=} \prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x_i - m)^2}{2\sigma^2}\right]$$

$$= \boxed{\frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2\right]}$$

$$L(x_1, \dots, x_n ; \theta) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma}\sum_{i=1}^{n}(x_i - m)^2\right]$$

$$\boxed{\ln L = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2}$$

$$\frac{\partial \ln L}{\partial m} = -\frac{1}{\cancel{2}\sigma^2}\sum_{i=1}^{n}\cancel{2}(x_i-m)\cancel{(-1)} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-m)$$

$$\frac{\partial^2 \ln L}{\partial m^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(-1) = \frac{-n}{\sigma^2}$$

$$E\left[\frac{\partial^2 \ln L}{\partial m^2}\right] = E\left[\frac{-n}{\sigma^2}\right] = -\frac{n}{\sigma^2}$$

Pour un estimateur non biaisé de m, la borne de Cramér-Rao est

$$BCR(m) = \frac{(1+0)^2}{-\left[\frac{-n}{\sigma^2}\right]} = \boxed{\frac{\sigma^2}{n}}$$ //

On a vu tout à l'heure que

$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$ est un estimateur de m non biaisé

et de variance $Var(\hat{\theta}) = \frac{\sigma^2}{n} = BCR(m)$

Cours du 16/11/2022

$$E[(\hat{\theta}-\theta)^2] = E[(\hat{\theta}-E[\hat{\theta}] + \overbrace{E[\hat{\theta}]-\theta}^{b_n(\theta)})^2]$$

$$= \underbrace{E[(\hat{\theta}-E[\hat{\theta}])^2]}_{Var\,\hat{\theta}} + b_n^2(\theta) + 2b_n(\theta)\underbrace{E[\hat{\theta}-E[\hat{\theta}]]}_{0}$$

$$A = cov(\hat{\theta}) = \begin{pmatrix} Var\,\hat{a} & cov(\hat{a},\hat{b}) & \times \\ \times & Var\,\hat{b} & \times \\ \times & \times & Var\,\hat{\sigma}^2 \end{pmatrix} \geq B = I_n^{-1}(\theta) \quad \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & \\ & & b_{33} \end{pmatrix}$$

$$\hat{\theta} = \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{\sigma}^2 \end{pmatrix}$$

σ^{2},

$$x^T (A - B)\, x \geqslant 0 \quad \forall x \in \mathbb{R}^3$$

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \underbrace{(1\ 0\ 0) \begin{pmatrix} \operatorname{Var} \hat{a} - b_{11} & \times & \times \\ \times & \operatorname{Var} \hat{b} - b_{22} & \times \\ \times & \times & \operatorname{Var} \hat{\sigma}^2 - b_{33} \end{pmatrix}}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \geqslant 0$$

$$(1\ 0\ 0) \begin{pmatrix} \operatorname{Var} \hat{a} - b_{11} \\ \times \\ \times \end{pmatrix}$$

$$\Rightarrow \quad \operatorname{Var} \hat{a} - b_{11} \geqslant 0 \iff \boxed{\operatorname{Var} \hat{a} \geqslant b_{11}}$$

$$x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \boxed{\operatorname{Var} \hat{b} \geqslant b_{22}}$$

$$x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \boxed{\operatorname{Var} \hat{\sigma}^2 \geqslant b_{33}}$$

Example $\boxed{X_i \sim N(m, \sigma^2) \qquad \theta = \begin{pmatrix} m \\ \sigma^2 \end{pmatrix}}$ $\boxed{\begin{matrix} \operatorname{Var} \hat{m} \geqslant ? \\ \operatorname{Var} \hat{\sigma}^2 \geqslant ? \end{matrix}}$

$$I_n(\theta) = \begin{pmatrix} \underbrace{E\left[-\frac{\partial^2 \ln L}{\partial m^2}\right]}_{-\frac{n}{\sigma^2}} & E\left[-\frac{\partial^2 \ln L}{\partial \sigma^2 \partial m}\right] \times \\ \times & \underbrace{E\left[-\frac{\partial^2 \ln L}{\partial (\sigma^2)^2}\right]}_{?} \end{pmatrix}$$

$$\ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - m)^2$$

$$\frac{\partial \ln L}{\partial \sigma^2} = 0 - \underbrace{\frac{n}{2} \frac{1}{\sigma^2}} + \frac{1}{2} \frac{1}{\sigma^4} \sum_{i=1}^{n} (x_i - m)^2$$

$$\frac{\partial^2 \ln L}{\partial(\sigma^2)^2} = \frac{n}{2\sigma^4} + \frac{1}{2}\left(\frac{-2}{\sigma^6}\right)\sum_{i=1}^{n}(x_i - m)^2$$

$$E[\;] = \frac{n}{2\sigma^4} \underbrace{- \frac{1}{\sigma^6}\sum_{i=1}^{n} \underbrace{E((x_i-m)^2)}_{\sigma^2}}_{-\frac{n}{\sigma^4} = -\frac{2n}{2\sigma^4}} = \boxed{\frac{-n}{2\sigma^4}}$$

$$\frac{\partial^2 \ln L}{\partial m \,\partial \sigma^2} = 0 + \frac{1}{2\sigma^4}\sum_{i=1}^{n} 2(x_i - m)(-1)$$

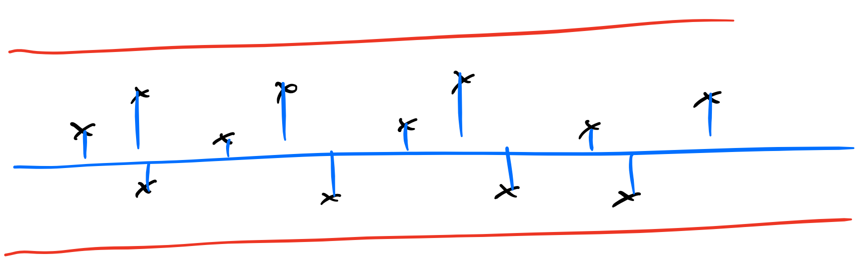
$$E[\;] = -\frac{1}{\sigma^4}\sum_{i=1}^{n}\left(E[x_i] - m\right) = \boxed{0}$$

$$I_n(\theta) = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{bmatrix}$$

$$I_n^{-1}(\theta) = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{bmatrix}$$

$$\operatorname{Var} \hat{m} \geq \frac{\sigma^2}{n}$$

$$\operatorname{Var} \widehat{\sigma^2} \geq \frac{2\sigma^4}{n}$$



Estimateur du max de vraisemblance

Ex 1 $\quad X_i \sim \mathcal{N}(m, \sigma^2) \qquad \theta = \begin{pmatrix} m \\ \sigma^2 \end{pmatrix}$

$$\ln L(x_1, \dots, x_n; \theta) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2$$

$$\frac{\partial \ln L}{\partial m} = 0 \iff -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i - m)(-1) = 0$$

$$\iff \sum x_i - nm = 0 \iff \boxed{m = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = 0 \iff -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - m)^2 = 0$$

$$\iff n\sigma^2 = \sum_{i=1}^{n}(x_i - m)^2$$

$$\iff \boxed{\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2}$$

L'estimateur du max de vraisemblance de θ vérifie

$$\begin{cases} m = \frac{1}{n}\sum_{i=1}^{n} x_i \\ \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 \end{cases} \Rightarrow \boxed{\begin{array}{l} \hat{m} = \frac{1}{n}\sum_{i=1}^{n} X_i \\ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{m})^2 \end{array}}$$

Exemple 1

$X_i \sim \mathcal{P}(\lambda) \qquad \theta = (\lambda)$

$$P[X_i = x_i] = \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \qquad x_i \in \mathbb{N}$$

Quel est l'estimateur du maximum de vraisemblance de λ ?

vraisemblance

$$L(x_1,\dots,x_n;\lambda) = \prod_{i=1}^{n}\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right)$$

$$= \boxed{\frac{\lambda^{\sum x_i}}{\prod x_i!}e^{-n\lambda}}$$

log.vraisemblance

$$\ln L = \left(\sum x_i\right)\ln\lambda - \ln\left(\prod x_i!\right) - n\lambda$$

$$\frac{\partial \ln L}{\partial \lambda} = \frac{\sum x_i}{\lambda} + 0 - n = 0$$

$$\boxed{\lambda = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

$$\boxed{\hat{\lambda}_{MV} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

Invariance Fonctionnelle

$X_i \sim N(m, \sigma^2)$

$$\hat{m} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{m})^2$$

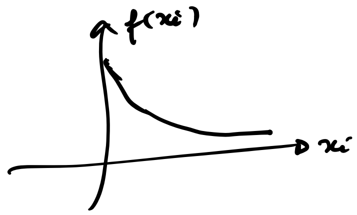
Quel est l'estimateur du maximum de vraisemblance de σ ?

$\sigma = \sqrt{\sigma^2}$

$= h(\sigma^2)$ $\Rightarrow$ $\boxed{\hat{\sigma}_{MV} = h(\hat{\sigma}^2_{MV})}$

ici $\boxed{\hat{\sigma}_{MV} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{m})^2}}$

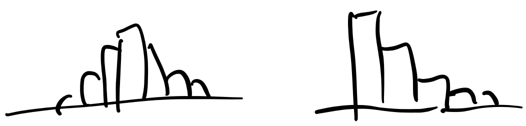
Méthode des moments.



$X_i \sim \Gamma(\theta, v)$

$$f(x_i) = \frac{\theta^v}{\Gamma(v)} e^{-\theta x_i} x_i^{v-1} \qquad x_i > 0$$

$\begin{pmatrix} \theta \\ v \end{pmatrix}$ Estimateurs de θ et de v ?



$$\begin{cases} E[X_i] = m_1 = \dfrac{v}{\theta} \\ E[X_i^2] = m_2 = \dfrac{v}{\theta^2} + \dfrac{v^2}{\theta^2} \end{cases} \Rightarrow \begin{cases} v = \theta m_1 \\ m_2 = \dfrac{\theta m_1}{\theta^2} + \dfrac{\theta^2 m_1^2}{\theta^2} \end{cases}$$

$$\Uparrow$$

$$\mathrm{Var}\, X_i + E^2(X_i)$$

$$\Rightarrow \begin{cases} \dfrac{1}{\theta} = \dfrac{m_2 - m_1^2}{m_1} \Rightarrow \boxed{\theta = \dfrac{m_1}{m_2 - m_1^2}} \\ v = \theta m_1 = \boxed{\dfrac{m_1^2}{m_2 - m_1^2} = v} \end{cases}$$

$$E[X_i] \simeq \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$E[X_i^2] \simeq \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Dans cet exemple

$$\hat{\theta}_{MO} = \frac{\frac{1}{n}\sum X_i}{\frac{1}{n}\sum X_i^2 - \left(\frac{1}{n}\sum X_i\right)^2}$$

$$\hat{v}_{MO} = \frac{\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2}{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum X_i\right)^2}$$

n

Exemple 2 $\quad x_i \sim N(m, \sigma^2) \qquad \theta = \begin{pmatrix} m \\ \sigma^2 \end{pmatrix}$

$$\hat{m}_{MV} = \frac{1}{n}\sum_{i=1}^{n} x_i$$
$$\hat{\sigma}^2_{MV} = \frac{1}{n}\sum_{i=1}^{n} (x_i - \hat{m})^2$$

$$\begin{cases} E[x_i] = m \\ E[x_i^2] = \sigma^2 + m^2 \end{cases} \Rightarrow \boxed{\begin{aligned} m &= m_1 \\ \sigma^2 &= m_2 - m_1^2 \end{aligned}}$$

$$\boxed{\begin{aligned} \hat{m}_{MO} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ \hat{\sigma}^2_{MO} &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \end{aligned}}$$

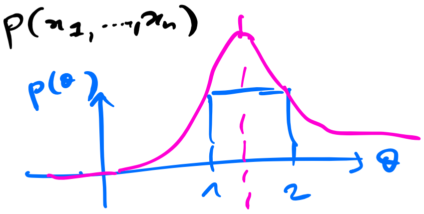
Cours du 23/11/2022

Bayes $\quad P(A|B) = \dfrac{P(B|A)\,P(A)}{P(B)} \qquad A = \theta, \quad B = (x_1, \dots, x_n)$

vraisemblance $\mathcal{L}$

$$\underbrace{p(\theta \,|\, \overbrace{x_1, \dots, x_n}^{B})}_{\text{loi a posteriori}} = \frac{p(x_1, \dots, x_n \,|\, \theta)\; p(\theta)}{p(x_1, \dots, x_n)}$$

loi a priori sur θ



Estimateur MAP = estimateur du maximum a posteriori

$$P(B) = \int \underbrace{p(A, B)}_{P(B|A)\,p(A)} \, dA$$

$$\boxed{p(x_1, \dots, x_n) = \int p(x_1, \dots, x_n \,|\, \theta)\, p(\theta)\, d\theta}$$

$$\boxed{p(\theta \,|\, x_1, \dots, x_n) \propto p(x_1, \dots, x_n \,|\, \theta)\, p(\theta)}$$

$X_i \sim N(\theta,\sigma^2) \Rightarrow$ vraisemblance $\displaystyle\prod_{i=1}^{n} \underbrace{f(x_i;\theta)}_{\frac{1}{\sqrt{2\pi\sigma^2}}\exp -\frac{(x_i-\theta)^2}{2\sigma^2}}$

$$\boxed{p(x_1,\dots,x_n|\theta) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2\right]}$$

vraisemblance

Loi a priori $\quad \theta \sim N(\mu,\nu^2) \Rightarrow \boxed{p(\theta) = \frac{1}{\sqrt{2\pi\nu^2}}\exp\left[-\frac{(\theta-\mu)^2}{2\nu^2}\right]}$

Estimateur MAP
$$\hat{\theta}_{MAP} = \arg\max_{\theta}\left[p(x_1,\dots,x_n|\theta)\,p(\theta)\right]$$
$$= \arg\max_{\theta}\left[\ln p(x_1,\dots,x_n|\theta) + \ln p(\theta)\right]$$

$$\ln p = K - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2 - \frac{1}{2\nu^2}(\theta-\mu)^2$$

$$\frac{\partial \ln p}{\partial\theta} = 0 \iff 0 \cancel{-}\frac{1}{\cancel{2}\sigma^2}\sum_{i=1}^{n}\cancel{2}(x_i-\theta)(\cancel{-1}) - \frac{1}{\cancel{2}\nu^2}\cancel{2}(\theta-\mu) = 0$$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}x_i - \frac{n\theta}{\sigma^2} - \frac{\theta}{\nu^2} + \frac{\mu}{\nu^2} = 0$$

$$\theta\underbrace{\left(\frac{n}{\sigma^2}+\frac{1}{\nu^2}\right)}_{\frac{n\nu^2+\sigma^2}{\sigma^2\nu^2}} = \frac{1}{\sigma^2}\sum_{i=1}^{n}x_i + \frac{\mu}{\nu^2}$$

$X_i|\theta \sim N(\theta,\sigma^2)$

$\theta \sim N(\mu,\nu^2)$

$$\boxed{\theta = \frac{n\nu^2}{n\nu^2+\sigma^2}\,\frac{1}{n}\sum_{i=1}^{n}x_i + \frac{\sigma^2}{n\nu^2+\sigma^2}\,\mu}$$

$$\boxed{\hat{\theta}_{MAP} = \frac{n\nu^2}{n\nu^2+\sigma^2}\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right) + \frac{\sigma^2}{n\nu^2+\sigma^2}\,\mu}$$

Loi de $\theta | x_1, \dots, x_n$: $\boxed{p(\theta | x_1, \dots, x_n) \propto \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \theta)^2 - \frac{1}{2v^2} (\theta - \mu)^2 \right]}$

$$\exp(a\theta^2 + b\theta + c) \propto \exp\left(- \frac{(\theta - m_\theta)^2}{2\sigma_\theta^2} \right)$$

$$\exp\left[-a \left(\theta^2 + \frac{b}{a} \theta + \frac{c}{a} \right) \right]$$

$a = \frac{1}{2\sigma^2}$ $\quad \boxed{\exp\left[- \frac{(\theta - m_\theta)^2}{2\sigma_\theta^2} \right]}$

coeff de θ^2 $\quad + \frac{1}{\cancel{2}\sigma_\theta^2} = + \frac{n}{\cancel{2}\sigma^2} + \frac{1}{\cancel{2}v^2}$

$$\frac{1}{\sigma_\theta^2} = \frac{nv^2 + \sigma^2}{\sigma^2 v^2}$$

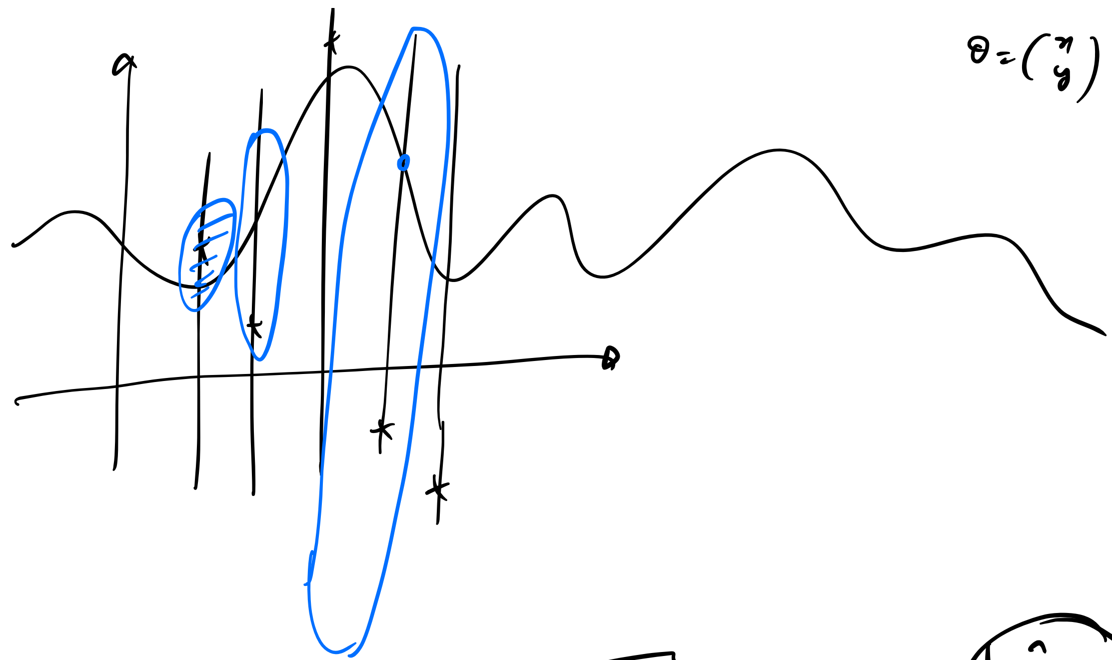
$$\boxed{\sigma_\theta^2 = \frac{\sigma^2 v^2}{nv^2 + \sigma^2}}$$

coeff de θ $\quad \cancel{+} \frac{1}{\cancel{2}\sigma_\theta^2} (-\cancel{2} m_\theta) = \cancel{+} \frac{1}{\cancel{2}\sigma^2} \left(-\cancel{2} \sum_{i=1}^{n} x_i \right) \cancel{+} \frac{1}{\cancel{2}v^2} (-\cancel{2}\mu)$

$$m_\theta = \frac{\sigma_\theta^2}{\sigma^2} \sum_{i=1}^{n} x_i + \frac{\sigma_\theta^2}{v^2} \mu$$

i.e $\quad \boxed{m_\theta = \frac{v^2}{nv^2 + \sigma^2} \sum_{i=1}^{n} x_i + \frac{\sigma^2}{nv^2 + \sigma^2} \mu}$

ici $\quad \boxed{E[\theta | x_1, \dots, x_n] = m_\theta = \hat{\theta}_{MMSE} = \hat{\theta}_{MAP}}$



$\theta = \begin{pmatrix} x \\ y \end{pmatrix}$

Intervalle de Confiance $\boxed{\theta \in \mathbb{R}}$

$x_i \sim N(m, \sigma^2)$
intervalle de confiance pour $\theta = m$

Étape 1 : on cherche un estimateur $\hat{\theta}$

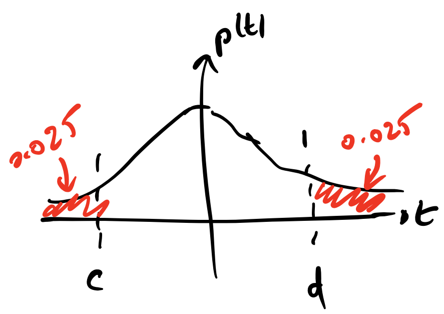
$\hat{\theta}_{MV}$
$\hat{\theta}_{MO}$
$\hat{\theta}_{MAP}$
$\hat{\theta}_{MMSE}$

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Étape 2 : on va chercher la loi de l'estimateur

$$\hat{\theta} \sim N\left(m, \frac{\sigma^2}{n}\right)$$

Étape 3 : on en déduit une statistique T de loi indépendante des paramètres θ



$$T = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i - m}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

$$P[c < T < d] = 0.95$$

$c \overset{?}{=}$

$F_{N(0,1)}(c) = 0.025$

$c = F^{-1}(0.025)$

$d = -c = F^{-1}(0.95 + 0.025)$

Étape 4 : on en déduit l'intervalle de confiance

$$c < T = \frac{\frac{1}{n}\sum x_i - m}{\sqrt{\sigma^2/n}} < d \quad \text{avec proba } 0{,}95$$

$$c\frac{\sigma}{\sqrt{n}} < \frac{1}{n}\sum x_i - m < d\frac{\sigma}{\sqrt{n}}$$

$$\boxed{\underbrace{\frac{1}{n}\sum x_i - \frac{d\sigma}{\sqrt{n}}}_{a} < m < \underbrace{\frac{1}{n}\sum x_i - c\frac{\sigma}{\sqrt{n}}}_{b}}$$

TD du 23/11/2022

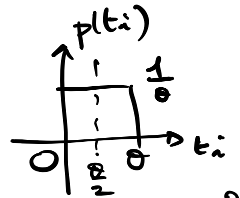
Biais $b_n(\theta) = E[\hat\theta] - \theta$

Variance $\sigma_n(\theta) = E[(\hat\theta - E[\hat\theta])^2] = \mathrm{Var}\,\hat\theta$

Erreur Quadratique moyenne

$$E[(\hat\theta - \theta)^2] = b_n^2(\theta) + \sigma_n(\theta)$$

1)



$$E[T_i] = \int_0^\theta \frac{t_i}{\theta}\,dt_i = \frac{1}{\theta}\frac{\theta^2}{2} = \boxed{\frac{\theta}{2}}$$

$$\mathrm{Var}\,T_i = E[T_i^2] - E^2[T_i] \qquad \Big|\ E[T_i^2] = \int_0^\theta \frac{1}{\theta}t_i^2\,dt_i = \frac{1}{\theta}\frac{\theta^3}{3} = \frac{\theta^2}{3}$$

$$= \frac{\theta^2}{3} - \frac{\theta^2}{4} = \boxed{\frac{\theta^2}{12}}$$

2)

$$\overline{T} = \frac{1}{n}\sum_{i=1}^{n} T_i$$

$$E[\overline{T}] = \frac{1}{n}\sum_{i=1}^{n}\frac{\theta}{2} = \boxed{\frac{\theta}{2}}$$

$$\mathrm{Var}[\overline{T}] = \mathrm{var}\left(\frac{1}{n}\sum T_i\right) = \frac{1}{n^2}\mathrm{var}\left(\sum T_i\right) \underset{T_1,\dots,T_n}{=} \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}\,T_i = \frac{\mathrm{Var}[T_1]}{n}$$

$$= \boxed{\frac{\theta^2}{12n}}$$

$\overline{T}$ est un estimateur biaisé de θ

Mais on peut enlever le biais en posant $\hat\theta_1 = 2\overline{T}$

$$E[\hat\theta_1] = 2E(T) = 2 \times \frac{\theta}{2} = \theta$$

$\boxed{\hat\theta_1 \text{ estimateur non biaisé de } \theta}$

$$Var(\hat\theta_1) = Var(2\overline{T}) = 4\, Var\,\overline{T} = \boxed{\frac{\theta^2}{3n}}$$

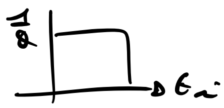
$$\left|\begin{array}{l} Var\,\hat\theta_1 \xrightarrow[n\to+\infty]{} 0 \\ \underbrace{E[\hat\theta_1] - \theta}_{} = 0 \xrightarrow[n\to+\infty]{} 0 \end{array}\right. \Rightarrow \boxed{\hat\theta_1 \text{ est un estimateur convergent de } \theta}$$

Remarque $E[X_i] = \frac{\theta}{2} \Rightarrow \theta = 2\, E[X_i]$

et $\hat\theta_1 = \frac{2}{n}\sum_{i=1}^{n} X_i$

$\boxed{\text{C'est l'estimateur des moments de } \theta \text{ !!}}$

3) Estimateur du max de vraisemblance

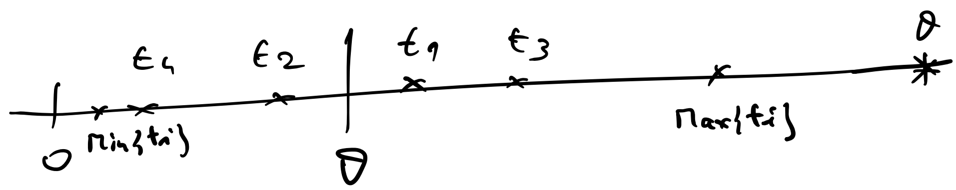


$$f(t_i) = \frac{1}{\theta}\mathbb{1}_{]0,\theta[}(t_i)$$

$$\hat\theta_{MV} = \underset{\theta}{\arg\max}\ L(t_1, \dots, t_n ; \theta)$$

$$L(t_1, \dots, t_n) = \frac{1}{\theta^n}\prod_{i=1}^{n}\mathbb{1}_{]0,\theta[}(t_i)$$

$$\mathbb{1}_{]0,\theta[}(t_i) = 1 \iff 0 < t_i < \theta \quad \forall i$$



$$\boxed{\hat\theta_{MV} = \underset{i\in\{1,\dots,n\}}{Max}\{T_i\}}$$

$\boxed{\text{Soit } Y_n = \sup_i T_i}$

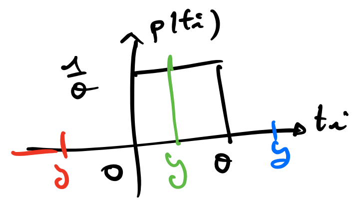
Déterminer la fonction de répartition de Y_n en remarquant que
$$y_n < y \iff t_i < y \quad \forall i = 1, \dots, n$$

En déduire la loi de Y_n, sa moyenne et sa variance

—

$$P[Y_n < y] = P[T_1 < y, T_2 < y, \dots, T_n < y]$$

$$= \prod_{i=1}^{n} P[T_i < y] \qquad T_i \sim \mathcal{U}(]0,\theta[)$$



$$\text{donc} \quad P[T_i < y] = \begin{cases} 0 & y \leq 0 \\ 1 & y \geq \theta \\ \int_0^y \frac{1}{\theta}\, dt_i & y \in]0,\theta[\end{cases}$$

$$\text{donc} \quad \boxed{P[T_i < y] = \frac{y}{\theta}\, \mathbb{1}_{]0,\theta[}(y)} \qquad y \in]0,\theta[$$

$$\text{Donc} \quad P[Y_n < y] = \prod_{i=1}^{n} \begin{cases} 0 & y \leq 0 \\ \frac{y}{\theta} & y \in]0,\theta[\\ 1 & y \geq \theta \end{cases}$$

$$= \boxed{\begin{cases} 0 & y \leq 0 \\ \left(\frac{y}{\theta}\right)^n & y \in]0,\theta[\\ 1 & y \geq \theta \end{cases}}$$

La densité de Y_n est donc

$$\boxed{p(y_n) = \begin{cases} \frac{n y^{n-1}}{\theta^n} & y \in]0,\theta[\\ 0 & \text{sinon} \end{cases}}$$

Moyenne $\quad E[Y_n] = \int_0^\theta y \times \frac{n y^{n-1}}{\theta^n}\, dy = \frac{n}{\theta^n}\left[\frac{y^{n+1}}{n+1}\right]_0^\theta = \boxed{\frac{n\theta}{n+1}}$

Variance $\quad \mathrm{Var}\, Y_n = E[Y_n^2] - E^2[Y_n]$

$$E(Y_n^2) = \int_0^\theta y^2 \frac{n y^{n-1}}{\theta^n}\, dy = \frac{n}{\theta^n}\left[\frac{y^{n+2}}{n+2}\right]_0^\theta$$

$$= \frac{n}{n+2}\theta^2$$

$$\mathrm{Var}\,Y_n = \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 = \frac{n\theta^2\left[(n+1)^2 - n(n+2)\right]}{(n+2)(n+1)^2}$$

donc $\boxed{\mathrm{Var}\,Y_n = \frac{n}{(n+2)(n+1)^2}\theta^2}$

Y_n est un estimateur biaisé de θ car $E[Y_n] = \frac{n}{n+1}\theta \neq \theta$

On enlève le biais en construisant $\boxed{\hat{\theta}_2 = \frac{n+1}{n} Y_n = \frac{n+1}{n} \sup_i T_i}$

Par construction, $\hat{\theta}_2$ est un estimateur non biaisé de θ !

Montrer que $\hat{\theta}_2$ est un estimateur convergent de θ
Choix entre $\hat{\theta}_1$ et $\hat{\theta}_2$? Pourquoi ?

$$\mathrm{var}\,\hat{\theta}_2 = \frac{(n+1)^2}{n^2}\mathrm{Var}\,Y_n = \frac{\theta^2}{n(n+2)} \xrightarrow[n\to+\infty]{} 0 \quad \sim \boxed{\frac{\theta^2}{n^2}}$$

$$E[\hat{\theta}_2] - \theta_2 = 0 \xrightarrow[n\to+\infty]{} 0$$

$\Rightarrow \boxed{\hat{\theta} \xrightarrow[n\to+\infty]{} \theta}$ estimateur convergent

$$\begin{cases} E[\hat{\theta}_1] - \theta_1 = 0 \\ \mathrm{var}\,\hat{\theta}_1 = \frac{\theta^2}{3n} \end{cases}$$

On préfère $\hat{\theta}_2$ qui converge plus vite que $\hat{\theta}_1$

Exo 3

$$f(x_i) = \begin{cases} \beta e^{\beta(\alpha - x_i)} & x \in]\alpha, +\infty[\\ 0 & \text{sinon } (x \leq \alpha) \end{cases}$$

1) Estimateur du max de vraisemblance de $\omega = \frac{1}{\beta}$? (α connu)
Montrer que cet estimateur est sans biais et convergent
Est-il efficace

2) même travail avec α (β connu)

Rq

$Y_i = X_i - \alpha \sim \mathcal{E}(\beta)$

$$g(y_i) = \begin{cases} \beta e^{-\beta y_i} & y_i > 0 \\ 0 & \text{sinon} \end{cases}$$

$E[Y_i] = \frac{1}{\beta} = \omega$ Table

$\mathrm{Var}(Y_i) = \frac{1}{\beta^2} = \omega^2$

Loi exponentielle ou loi gamma

principe d'invariance fonctionnelle

$$\widehat{f(\theta)} = f(\hat{\theta})$$

$$\hat{\omega} = \frac{1}{\hat{\beta}}$$

Méthode 2

i) Vraisemblance $\omega = \frac{1}{\beta}$

Méthode 1

$$L(x_1,\dots,x_n;\omega) = \frac{1}{\omega^n} e^{\sum_{i=1}^{n} \frac{1}{\omega}(\alpha - x_i)}$$

$$\ln L = -n\ln\omega + \frac{1}{\omega}\sum_{i=1}^{n}(\alpha - x_i)$$

$$\frac{\partial \ln L}{\partial \omega} = 0 \iff \boxed{\frac{-n}{\omega} - \frac{1}{\omega^2}\sum_{i=1}^{n}(\alpha - x_i) = 0}$$

donc $\omega = \frac{1}{n}\sum_{i=1}^{n}(x_i - \alpha)$

$$\boxed{\hat{\omega}_{MV} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \alpha)}$$

$E[X_i - \alpha]$

$Var(X_i - \alpha) = Var\, X_i$

Biais $E[\hat{\omega}_{MV}] - \omega$

$$E[\hat{\omega}_{MV}] = \frac{1}{n}\sum_{i=1}^{n}\underbrace{E[Y_i]}_{\omega} = \omega$$ donc $\boxed{\hat{\omega}_{MV} \text{ estimateur non biaisé de } \omega}$

Variance

$$Var\,\hat{\omega}_{MV} = \frac{1}{n^2}\sum_{i=1}^{n}\underbrace{Var(X_i - \alpha)}_{Var\, Y_i} = \frac{Var[Y_1]}{n} = \frac{\omega^2}{n}$$

$Var\,\hat{\omega} = \frac{\omega^2}{n} \xrightarrow[n\to+\infty]{} 0$

$E[\hat{\omega}] - \omega = 0 \xrightarrow[n\to+\infty]{} 0$

donc $\boxed{\hat{\omega}_{MV} \text{ estimateur convergent}}$

Efficacité $\boxed{Var\,\hat{\omega}_{MV} = BCR}$?

$$\boxed{BCR = \frac{[1 + b_n'(\omega)]^2}{E\left[-\frac{\partial^2 \ln L(X_1,\dots,X_n;\omega)}{\partial \omega^2}\right]}}$$

ici $b_n(\omega) = 0 \Rightarrow$ $\boxed{\text{Numérateur} = 1}$

$$\frac{\partial \ln L}{\partial \omega} = \frac{-n}{\omega} - \frac{1}{\omega^2}\sum_{i=1}^{n}(\alpha - x_i)$$

$$\frac{\partial^2 \ln L}{\partial \omega^2} = \frac{n}{\omega^2} + \frac{2}{\omega^3}\sum_{i=1}^{n}(\alpha - x_i)$$

$$E\left[\frac{\partial^2 \ln L}{\partial \omega^2}\right] = \frac{n}{\omega^2} + \frac{2}{\omega^3}\sum_{i=1}^{n} -E[\underbrace{x_i - \alpha}_{y_i}]$$

$$= \frac{n}{\omega^2} + \frac{-2n\omega}{\omega^3} = -\frac{n}{\omega^2}$$

$$\boxed{BCR = \frac{1}{-\left(\frac{-n}{\omega^2}\right)} = \frac{\omega^2}{n}}$$

$\operatorname{Var}\hat{\omega}_{MV} = BCR$

$\hat{\omega}_{MV}$ non biaisé

donc $\hat{\omega}_{MV}$ est l'estimateur efficace de ω

Cours du 30/11/2022

H_0 pas d'anomalie = fonctionnement normal

H_1 anomalie = problème

$\alpha = P[$ Rejeter $H_0 \mid H_0$ vraie$] = P[$ décider qu'on a une anomalie $\mid$ on est dans un régime normal$]$

$= PFA$ = probabilité de fausse alarme

$\beta = P[$ Rejeter $H_2 \mid H_1$ vraie$] = P[$ décider qu'on est en fonctionnement normal $\mid$ on a une anomalie$]$

$= PND$ probabilité de non détection

$\pi = 1 - \beta = P[$ Accepter $H_2 \mid H_1$ vraie$] = P[$ décider qu'on a une anomalie $\mid$ on a une anomalie$]$

$= P_D$ = probabilité de détection

Exemple $\quad X_i \sim N(m, \sigma^2)$ $\quad$ H_0 $\quad \underline{m = m_0}$ (moyenne "petite")

$\underline{\sigma^2 \text{ connue}}$ $\quad$ H_1 $\quad m = m_1 > m_0$ (moyenne "grande")

$\left[\frac{1}{n} \ \frac{1}{n} \ \cdots \ \frac{1}{n}\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

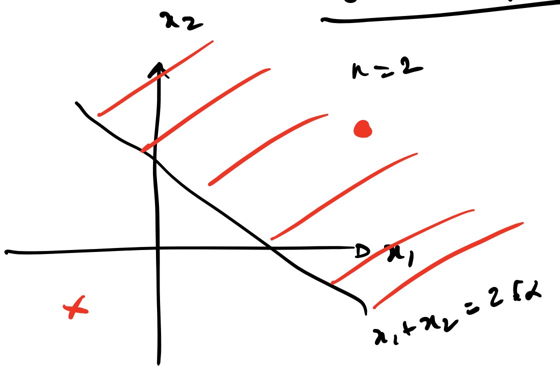
Rejet de H_0 si $\quad T = \frac{1}{n}\sum_{i=1}^{n} x_i > \text{Seuil} = S_\alpha$

Statistique de test $\quad T = \frac{1}{n}\sum_{i=1}^{n} x_i$

Région critique du test $\quad \left\{ (x_1, \ldots, x_n) \,/\, \frac{1}{n}\sum_{i=1}^{n} x_i > S_\alpha \right\}$

$n = 2$

$x_1 + x_2 > 2 S_\alpha$



Question : comment calculer α et β ?

$$\alpha = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}]$$
$$= P\left[\frac{1}{n}\sum_{i=1}^{n} x_i > S_\alpha \,\middle|\, m = m_0\right]$$
$$= 1 - \underbrace{P\left[\frac{1}{n}\sum_{i=1}^{n} x_i \leq S_\alpha \mid m = m_0\right]}_{\text{Fonction de répartition.}}$$

$$U = \frac{\frac{1}{n}\sum x_i - m}{\sqrt{\sigma^2/n}} \leq \frac{S_\alpha - m}{\sqrt{\sigma^2/n}}$$

$\sim N(0,1)$

$$X_i \sim N(m, \sigma^2) \Rightarrow \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(m, \frac{\sigma^2}{n}\right)$$

$$\left(Var\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2} Var\left(\sum X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{Var X_i}_{\sigma^2} = \frac{n\sigma^2}{n^2} \right.$$

$x_1, \dots, x_n$ iid $\qquad = \frac{\sigma^2}{n}$

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du$$

fonction de la répartition de la loi normale $N(0,1)$

$$\boxed{\frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(m, \frac{\sigma^2}{n}\right) \iff \frac{T-m}{\sqrt{\sigma^2/n}} \sim N(0,1)}$$

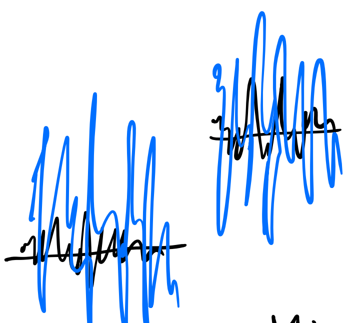
on centre et on réduit

$$\alpha = 1 - P\left[\underbrace{\frac{\frac{1}{n}\sum X_i - m}{\sqrt{\sigma^2/n}}}_{U} \leq \frac{S_\alpha - m}{\sqrt{\sigma^2/n}} \right]_{m = m_0}$$

$$\boxed{\alpha = 1 - F\left[\frac{S_\alpha - m_0}{\sqrt{\sigma^2/n}} \right]} \qquad F\left(\sqrt{n} \, \frac{S_\alpha - m_0}{\sqrt{\sigma^2}} \right)$$

$U \sim N(0,1)$

Rq : α dépend du seuil noté S_α, de m_0 (valeur de θ sous H_0), n nombre de données, et de σ^2 (variance des données)



Calcul de β

$$\beta = P[\text{Rejeter } H_1 \mid H_1 \text{ vraie}]$$

$$= P\left[T = \frac{1}{n}\sum_{i=1}^{n} X_i < S_\alpha \mid m = m_1 \right]$$

$$= P\left[\frac{T - m}{\sqrt{\sigma^2/n}} < \frac{S_\alpha - m}{\sqrt{\sigma^2/n}} \mid m = m_1 \right]$$

$$\beta = F\left[\frac{S_\alpha - m_1}{} \right]$$

$$\sqrt{\sigma^2/n}\ \Big)$$

En pratique, pour faire le test, il faut se fixer soit α, soit β.

En stat, on se fixe $\alpha \begin{array}{l}\nearrow 0{,}01 \\ \searrow 0{,}05\end{array}$, on en déduit le seuil S_α et ensuite on calcule β

on se fixe α

$$\begin{cases} \alpha = 1 - F\left(\dfrac{S_\alpha - m_0}{\sqrt{\sigma^2/n}}\right) & (1) \Rightarrow S_\alpha \\ \beta = F\left(\dfrac{S_\alpha - m_1}{\sqrt{\sigma^2/n}}\right) & (2) \Rightarrow \beta \end{cases}$$

Valeur du seuil

$$F\left(\frac{S_\alpha - m_0}{\sqrt{\sigma^2/n}}\right) = 1 - \alpha$$

$$\Rightarrow \quad \frac{S_\alpha - m_0}{\sqrt{\sigma^2/n}} = F^{-1}(1-\alpha)$$

$$\boxed{S_\alpha = m_0 + \sqrt{\sigma^2/n}\ F^{-1}(1-\alpha)}$$

$\boxed{\text{Courbe COR}}$ $\qquad \Pi = f(\alpha)$

$$\beta = F\left(\frac{s_\alpha - m_1}{\sqrt{\sigma^2/n}}\right)$$

$$\boxed{\Pi = 1 - \beta = 1 - F\left(\boxed{\frac{m_0 - m_1}{\sqrt{\sigma^2/n}}} + F^{-1}(1-\alpha)\right)}$$

$$\frac{\sqrt{n}}{\sigma}(m_0 - m_1) = -\frac{\sqrt{n}}{\sigma}\underbrace{(m_1 - m_0)}_{>0}$$

$\boxed{\text{NEYMAN-PEARSON}}$

$X_i \sim N(m, \sigma^2)$

H_0 $\quad m = m_0$

H_1 $\quad m = m_1 > m_0$

Rejet de H_0 si
$$\frac{\cancel{\frac{1}{(2\pi\sigma^2)^{n/2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m_1)^2\right]}{\cancel{\frac{1}{(2\pi\sigma^2)^{n/2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m_0)^2\right]} > s_\alpha$$

On teste équivalent si

Rejet de H_0 si $-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\cancel{x_i^2}-2m_1 x_i+m_1^2)+\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\cancel{x_i^2}-2m_0 x_i+m_0^2) > k_\alpha$

$$n\frac{m_0^2-m_1^2}{2\sigma^2}+\sum x_i\underbrace{\left(\frac{m_1-m_0}{\sigma^2}\right)}_{>0} > k_\alpha$$

$m_1 > m_0$

Test équivalent

$$\boxed{\text{Rejet de } H_0 \text{ si } \sum_{i=1}^{n} x_i > U_\alpha}$$

T

TD du 3/11/2022

Exo 1

1) La vraisemblance s'écrit

$$L(t_1,\dots,t_n;\theta)=\prod_{i=1}^{n} f_T(t_i)=\theta^n e^{-\theta\sum_{i=1}^{n} t_i}$$

L'estimateur du maximum de vraisemblance est

$$\hat{\theta}_{MV}=\underset{\theta}{\arg\max}\, L(t_1,\dots,t_n;\theta)=\underset{\theta}{\arg\max}\, \ln L(t_1,\dots,t_n;\theta)$$

Comme d'habitude, il est plus simple de travailler avec la log vraisemblance qui est maximale pour θ vérifiant

$$\frac{\partial \ln L}{\partial\theta}=0 \iff \frac{\partial}{\partial\theta}\left[n\ln\theta-\theta\sum_{i=1}^{n} t_i\right]=0$$

$$\iff \frac{n}{\theta} - \sum_{i=1}^{n} t_i = 0$$

$$\iff \theta = \frac{n}{\sum_{i=1}^{n} t_i} \qquad (1)$$

On en déduit
$$\boxed{\hat{\theta}_{MV} = \frac{n}{\sum_{i=1}^{n} T_i}}$$

Rq : si on demande de vérifier que (1) est un bien le maximum de vraisemblance, on peut étudier le signe de $\frac{\partial \ln L}{\partial \theta}$ et faire un tableau de variations :

$$\frac{\partial \ln L}{\partial \theta} \geq 0 \iff \frac{n}{\theta} - \sum_{i=1}^{n} t_i > 0 \iff \theta < \frac{n}{\sum t_i}$$

d'où

θ	0		$n/\sum t_i$		$+\infty$
$\frac{\partial \ln L}{\partial \theta}$		$+$	0	$-$	
$\ln L$		$\nearrow$		$\searrow$	

$\theta = \frac{n}{\sum t_i}$ est donc bien le maximum unique de la vraisemblance.

2) Estimateur MAP

$$\hat{\theta}_{MAP} = \arg\max_{\theta} p(\theta \mid t_1, \dots, t_n)$$

$$= \arg\max_{\theta} \frac{p(t_1, \dots, t_n \mid \theta)\, p(\theta)}{p(t_1, \dots, t_n)}$$

Comme $p(t_1, \dots, t_n)$ est indépendante de θ, on peut travailler avec $p(t_1, \dots, t_n \mid \theta)\, p(\theta)$.

On a $p(t_1, \dots, t_n \mid \theta)\, p(\theta) = \theta^n e^{-\theta \sum t_i} \times \lambda e^{-\lambda \theta}$

$\theta(\sum t_i + \lambda)$

$$\propto \theta^n e^{-\theta(\boxed{\sum t_i + \lambda})}$$

La vraisemblance est $p(t_1, \dots, t_n ; \theta) = \theta^n e^{-\theta \boxed{\sum t_i}}$

donc pour passer de la vraisemblance à la loi a posteriori, il suffit de remplacer $\sum t_i$ par $\sum t_i + \lambda$ d'où

$$\boxed{\hat{\theta}_{MAP} = \frac{n}{\sum_{i=1}^{n} T_i + \lambda}}$$

Estimateur MMSE

$$\hat{\theta}_{MMSE} = E[\theta \mid t_1, \dots, t_n]$$

qui est la moyenne de la loi a posteriori. Donc, on va chercher la loi a posteriori et déterminer sa moyenne à l'aide des tables de loi.

Comme $p(\theta \mid t_1, \dots, t_n) \propto \theta^n e^{-\theta(\sum t_i + \lambda)}$

On reconnait une loi gamma $G\left(n+1, \sum_{i=1}^{n} t_i + \lambda\right)$

dont la moyenne est (voir table) $\dfrac{n+1}{\sum t_i + \lambda}$

On en déduit

$$\boxed{\hat{\theta}_{MMSE} = \frac{n+1}{\sum_{i=1}^{n} T_i + \lambda}}$$

Quand n est grand, on remarque que $\hat{\theta}_{MAP}$ se comporte comme $\hat{\theta}_{MV}$. On oublie la loi a priori lorsqu'on a beaucoup d'observations !

Rq
$$\frac{\hat{\theta}_{MAP}}{\hat{\theta}_{MV}} = \frac{n}{\sum t_i + d} \times \frac{\sum t_i}{n}$$

$$= \frac{1}{1 + \frac{d}{\sum t_i}} \xrightarrow[n \to +\infty]{} 1$$

donc $\frac{\hat{\theta}_{MAP}}{\hat{\theta}_{MV}} \xrightarrow[n \to +\infty]{} 1$

Exo 2

1) La vraisemblance est

$$L(z_1, z_2; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z_1-\theta)^2}{2\sigma^2}} \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z_2-\theta)^2}{2\sigma^2}}$$

$$= \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}\left[(z_1-\theta)^2 + (z_2-\theta)^2\right]\right\}$$

Comme $\theta \in \{0, 1\}$, l'estimateur du max de vraisemblance de θ vaut 0 si

$$L(z_1, z_2; \theta = 0) \geqslant L(z_1, z_2; \theta = 1)$$

On en déduit

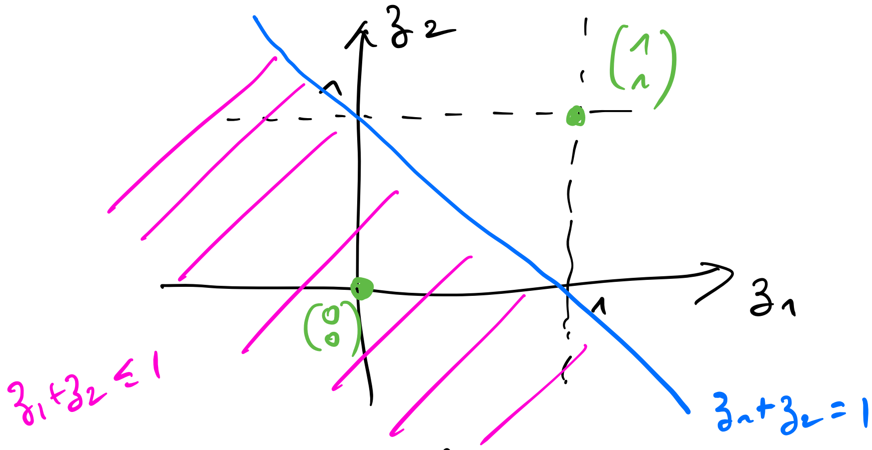
$$\hat{\theta}_{MV} = 0 \iff \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}(z_1^2 + z_2^2)\right] \geqslant \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}\left\{(z_1-1)^2 + (z_2-1)^2\right\}\right]$$

Soit

$$\hat{\theta}_{MV} = 0 \iff -\frac{1}{2\sigma^2}(z_1^2 + z_2^2) \geqslant -\frac{1}{2\sigma^2}\left(z_1^2 + z_2^2 - 2z_1 - 2z_2 + 2\right)$$

$$\iff 0 \geqslant \frac{z_1 + z_2 - 1}{\sigma^2}$$

$$\boxed{\hat{\theta}_{MV} = 0 \iff z_1 + z_2 \leq 1}$$



On choisit donc $\hat{\theta}_{MV}=0$ si le point $\begin{pmatrix}z_1\\z_2\end{pmatrix}$ est plus proche de $\begin{pmatrix}0\\0\end{pmatrix}$ que de $\begin{pmatrix}1\\1\end{pmatrix}$; ce qui est logique : Quand il n'y a pas de bruit $(e_1=e_2=0)$, on a

$$\begin{pmatrix}z_1\\z_2\end{pmatrix} = \begin{pmatrix}\theta\\\theta\end{pmatrix} \begin{cases} = \begin{pmatrix}0\\0\end{pmatrix} & \text{si } \theta = 0 \\ = \begin{pmatrix}1\\1\end{pmatrix} & \text{si } \theta = 1 \end{cases}$$

2) Estimateur MAP

$$\hat{\theta}_{MAP} = 0 \iff \underbrace{L(z_1,z_2;\theta=0)\, P[\theta=0]}_{\propto P[\theta=0 \mid z_1,z_2]} \geq \underbrace{L(z_1,z_2;\theta=1)\, P[\theta=1]}_{\propto P[\theta=1 \mid z_1,z_2]}$$

$\uparrow$
proportionnel à

Si $P[\theta=0] = P[\theta=1] = \frac{1}{2}$, on a $\boxed{\hat{\theta}_{MAP} = \hat{\theta}_{MV}}$

Par contre si $P[\theta=0] = p$ et $P[\theta=1] = 1-p$,

$$\hat{\theta}_{MAP}=0 \iff \frac{1}{2\pi\sigma^2}\exp\left[-\frac{1}{2\sigma^2}(z_1^2+z_2^2)\right] p \geq \frac{1}{2\pi\sigma^2}\exp\left[-\frac{1}{2\sigma^2}\{(z_1-1)^2+(z_2-1)^2\}\right] \times (1-p)$$

soit

$$\hat{\theta}_{MAP}=0 \iff \ln p - \frac{1}{2\sigma^2}(z_1^2+z_2^2) \geq \ln(1-p) - \frac{1}{2\sigma^2}\left[(z_1-1)^2+(z_2-1)^2\right]$$

$$\iff \ln\left(\frac{p}{1-p}\right) \geq -\frac{1}{2\sigma^2}\left[-2z_1-2z_2+2\right] = \frac{z_1+z_2-1}{\sigma^2}$$

d'où

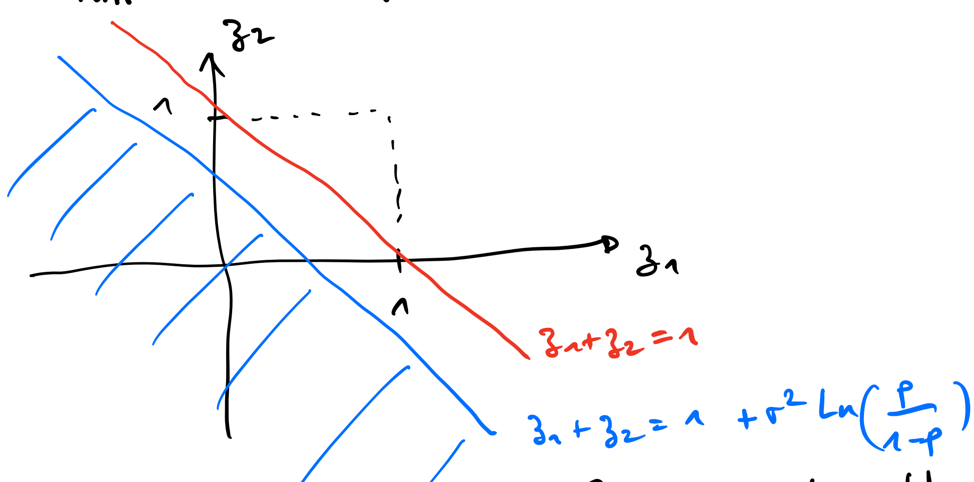
$$\boxed{\hat{\theta}_{MAP}=0 \iff z_1+z_2 \leq 1+\sigma^2\ln\left(\frac{p}{1-p}\right)}$$

- si $p<\frac{1}{2}$ (on a plus de chance d'avoir a priori $\theta=1$ que $\theta=0$)

alors $\frac{p}{1-p}<1$ donc $\ln\left(\frac{p}{1-p}\right)<0$, ce qui signifie

que $1+\sigma^2\ln\left(\frac{p}{1-p}\right)<1$. La zone associée

à $\hat{\theta}_{MAP}=0$ est représentée ci-dessous

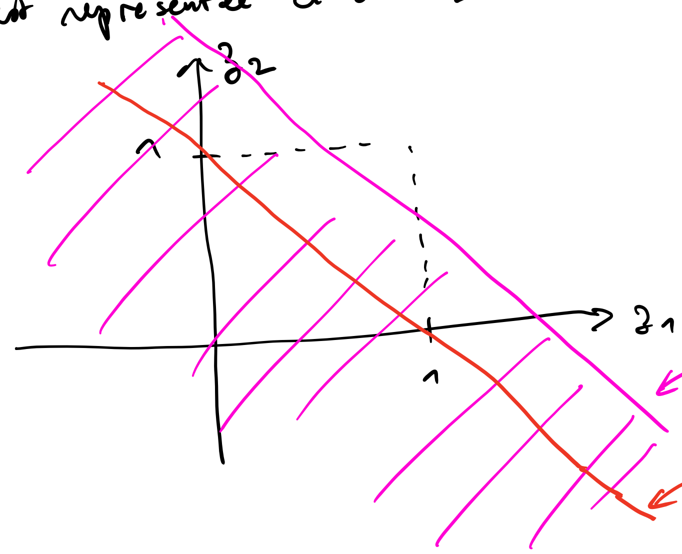


Comme $p=P[\theta=0]<\frac{1}{2}$, la zone associée à $\hat{\theta}_{MAP}=0$ est plus petite que celle délimitée par la droite $z_1+z_2=1$

- si $p>\frac{1}{2}$, on plus de chance a priori d'avoir $\theta=0$ que $\theta=1$

donc $\frac{p}{1-p}>1 \Rightarrow \ln\left(\frac{p}{1-p}\right)>0$, ce qui signifie

$1 + \sigma^2 \ln\left(\frac{p}{1-p}\right) > 1$. La zone associée à $\hat{\theta}_{MAP} = 0$ est représentée ci-dessous



Comme $p = P[\theta = 0] > \frac{1}{2}$, la zone associée à $\hat{\theta}_{MAP} = 0$ est plus grande que quand $p = \frac{1}{2}$ (délimitée par la droite $z_1 + z_2 = 1$)

$\boxed{\text{TD du } 6/12/2022}$

$\alpha = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}]$ PFA

$\beta = P[\text{Rejeter } H_1 \mid H_1 \text{ vraie}]$ PND

<u>NEYMAN PEARSON</u> <u>à α fixé</u>

On rejette H_0 si $\dfrac{L(x_1, \dots, x_n ; H_1)}{L(x_1, \dots, x_n ; H_0)} > \text{Seuil} = S_\alpha$

1) Rejet de H_0 si

$$\frac{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x_i^2}{2\sigma_1^2}\right)}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{x_i^2}{2\sigma_0^2}\right)} > \text{Seuil}$$

$$\frac{\dfrac{1}{(2\pi\sigma_1^2)^{n/2}} \exp\left(-\dfrac{1}{2\sigma_1^2}\sum_{i=1}^n x_i^2\right)}{\dfrac{1}{(2\pi\sigma_0^2)^{n/2}} \exp\left(-\dfrac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2\right)} > \text{seuil} = S_1$$

Un test équivalent est

Rejet de H_0 si $-\frac{n}{2}\ln(2\pi\sigma_1^2) - \frac{1}{2\sigma_1^2}\sum_{i=1}^n x_i^2 - \frac{n}{2}\ln(2\pi\sigma_0^2)$

$$+\frac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2 > S_2$$

Un test équivalent est

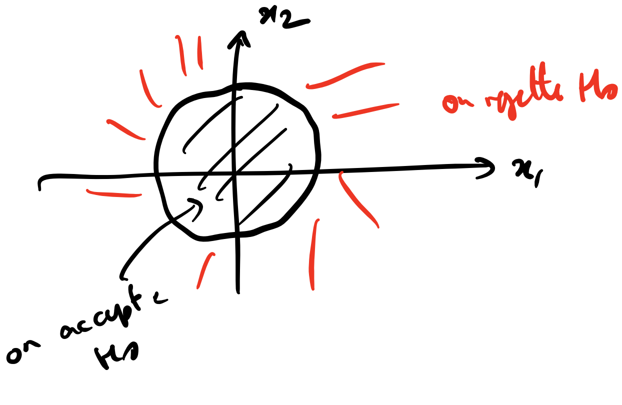
Rejet de H_0 si $\underbrace{\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)}_{c}\sum_{i=1}^n x_i^2 > S_3$

__1er cas__ $\boxed{\sigma_1 > \sigma_0}$ $\Leftrightarrow$ $\frac{1}{\sigma_1^2} < \frac{1}{\sigma_0^2}$ $\Leftrightarrow$ $c > 0$

$\boxed{\text{Rejet de } H_0 \text{ si } \sum_{i=1}^n x_i^2 > S_\alpha}$

Stat de test $T = \sum_{i=1}^n X_i^2$

Région critique $\left\{(x_1, \ldots, x_n) \,/\, \sum_{i=1}^n x_i^2 > S_\alpha\right\}$

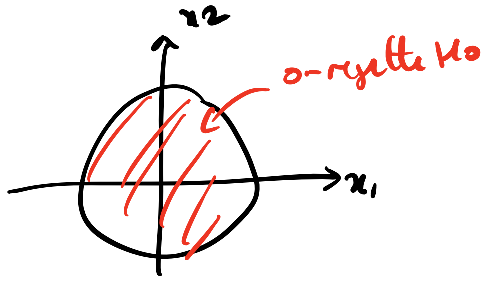


$2^{\text{ème}}$ cas. $\sigma_1 < \sigma_0$

Rejet de H_0 si $\sum_{i=1}^{n} x_i^2 < s_\alpha$

Statistique de test $T = \sum_{i=1}^{n} x_i^2$

Région critique $\{(x_1, \dots, x_n) / \sum_{i=1}^{n} x_i^2 < s_\alpha\}$



2) Calcul de β

Seuil s_α

$\alpha = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}]$

$= P[\sum x_i^2 > s_\alpha \mid \sigma^2 = \sigma_0^2]$

$X_i \sim N(0, \sigma^2)$ donc $\frac{X_i}{\sigma} \sim N(0,1)$

$= P[\underbrace{\frac{\sum x_i^2}{\sigma_0^2}}_{U} > \frac{s_\alpha}{\sigma_0^2} \mid \sigma^2 = \sigma_0^2]$ $\quad U \sim \chi^2_n$

$$\boxed{\alpha = \Phi_n\left(\frac{s_\alpha}{\sigma_0^2}\right)}$$

$\Rightarrow \frac{s_\alpha}{\sigma_0^2} = \Phi_n^{-1}(\alpha) \Rightarrow \boxed{s_\alpha = \sigma_0^2 \Phi_n^{-1}(\alpha)}$

$\beta = P[\text{Rejeter } H_1 \mid H_1 \text{ vraie}]$

$= P[\sum x_i^2 < s_\alpha \mid \sigma^2 = \sigma_1^2]$

$= P[\sum x_i^2 < s_\alpha \mid v \sim \chi^2_n]$

$$\boxed{\beta = 1-\phi_n\left(\frac{s_\alpha}{\sigma_1^2}\right)}$$

$$\boxed{\Pi = 1-\beta = \phi_n\left(\frac{s_\alpha}{\sigma_1^2}\right)}$$

3) Courbes COR $\qquad \Pi =$ Fonction de α

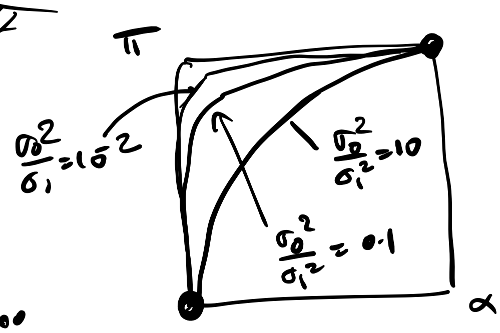
$$\Pi = \phi_n\left(\frac{s_\alpha}{\sigma_1^2}\right) = \boxed{\phi_n\left(\frac{\sigma_0^2}{\sigma_1^2}\,\phi_n^{-1}(\alpha)\right) = \Pi}$$

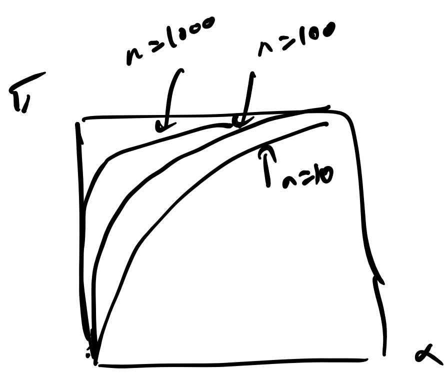
Courbes COR

Rq $\quad \phi_n(x) = \int_x^{+\infty} \text{—}\, dx$

ϕ_n est une fonction décroissante

Donc si $\frac{\sigma_0^2}{\sigma_1^2} \searrow$ alors $\phi_n \nearrow$





4) $T = \sum_{i=1}^{n} x_i^2$ $\qquad E(T) = n\,E[x_i^2] \quad (= nm)$

$x_i \sim N(0, \sigma^2)$

$E(x_i^2) = \operatorname{var} x_i + E^2[x_i] = \sigma^2 + 0 = \sigma^2$

$\boxed{m = \sigma^2}$

$\operatorname{Var}(T) = n \operatorname{Var}(x_i^2)$

$\operatorname{Var}(x_i^2) = \underbrace{E(x_i^4)}_{} - \underbrace{E^2[x_i^2]}_{\sigma^4} = \boxed{2\sigma^4}$

$\swarrow$

$\int x^4 p(x)\,dx = 3\sigma^4$

$$\boxed{T \sim N(n\sigma^2, 2n\sigma^4)}$$

Cours du 7/12/2022

NP

$x_1 = 1$

$x_2 = 5$

$x_3 = 1$

$x_4 = 2$

⋮

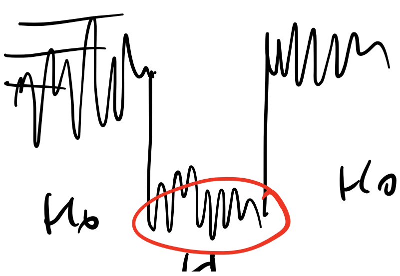
H_0 $\quad \boxed{\lambda = \lambda_0}$

H_1 $\quad \lambda = \boxed{\lambda_1 < \lambda_0}$

$X_i \sim P(\lambda) \qquad P[X_i = x_i] = \dfrac{\lambda^{x_i}}{x_i!} e^{-\lambda} \qquad x_i \in \mathbb{N}$

$\boxed{\lambda = E(X_i)} = \operatorname{Var}[X_i]$

(Tables)



Neyman Pearson?

à α fixé, le test qui minimise β est défini par

Rejet de H_0 si $\dfrac{L(x_1,\dots,x_n\,;H_1)}{L(x_1,\dots,x_n\,;H_0)} > S_\alpha$

si $\dfrac{P[X_1=x_1,\dots,X_n=x_n\,;\lambda=\lambda_1]}{P[\quad\quad\quad;\lambda=\lambda_0]} > S_\alpha$

si $\underbrace{\dfrac{\prod_{i=1}^{n}\left[\dfrac{\lambda_1^{x_i}}{\cancel{x_i}!}e^{-\lambda_1}\right]}{\prod_{i=1}^{n}\left[\dfrac{\lambda_0^{x_i}}{\cancel{x_i}!}e^{-\lambda_0}\right]}}$

$$\frac{\lambda_1^{\sum_{i=1}^{n} x_i}\, e^{-n\lambda_1}}{\lambda_0^{\sum x_i}\, e^{-n\lambda_0}} > S_\alpha$$

On rejette H_0 si $\underbrace{\sum_{i=1}^{n} x_i}_{T}\underbrace{(\ln \lambda_1 - \ln \lambda_0)}_{<0} > K_\alpha$

Donc

$$\boxed{\text{on rejette } H_0 \text{ si } \sum_{i=1}^{n} x_i < A_\alpha}$$

- stat de test $T = \sum_{i=1}^{n} X_i$

• Zone critique $\{(x_1,\dots,x_n) \,/\, \sum_{i=1}^{n} x_i < A_\alpha\}$

$\alpha = 0.01$ Comment trouver A_α ? β ?

$$\alpha \,(= 0{,}01) = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}]$$

$$= P[\underbrace{\sum_{i=1}^{n} X_i}_{T} < A_\alpha \mid \lambda = \lambda_0]$$

= Fonction de répartition de $T = \sum_{i=1}^{n} X_i$ au point A_α $= F_0(A_\alpha)$

$\boxed{\text{loi de } T = \sum_{i=1}^{n} X_i}$

$$\phi_T(t) = E[e^{itT}]$$

$$= E[\underbrace{e^{it\sum_{k=1}^{n} X_k}}_{\prod e^{itX_k}}] = \prod_{k=1}^{n} \underbrace{E[e^{itX_k}]}_{\phi_{X_k}(t)}$$

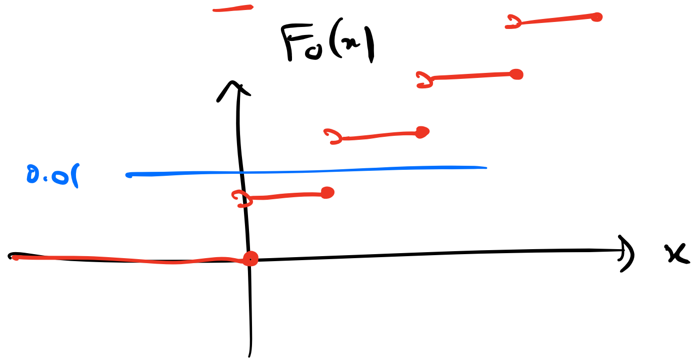
$$\phi_{X_k}(t) = \exp[\lambda(e^{it}-1)]$$

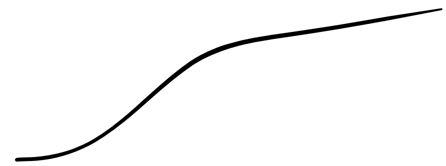
$$= \exp[n\lambda(e^{it}-1)]$$

c'est la fonction caractéristique d'une loi de Poisson de paramètre $n\lambda$

i.e $\boxed{T = \sum_{i=1}^{n} X_i \sim P(n\lambda)}$

$\alpha = F_0(A_\alpha)$ F_0 fonction de répartition d'une loi de Poisson $P(n\lambda_0)$





$F_0(A_\alpha) = 0.01$

La recherche de A_α est plus que pénible !!!

On approche la loi de $T = \sum_{i=1}^{n} X_i$ par une loi normale en vertu du théorème central limite

$$\sum_{i=1}^{n} X_i \underset{n \text{ grand}}{\approx} N(n\lambda, n\lambda)$$

$$E\left[\sum X_i\right] = \sum \underbrace{E(X_i)}_{\lambda} = n\lambda$$

$$Var\left[\sum X_i\right] \underset{X_1,\dots,X_n \text{ ind}}{=} \sum \underbrace{Var X_i}_{\lambda} = n\lambda$$

$$\alpha = P[T < A_\alpha \mid \lambda = \lambda_0]$$

$$\approx P[\quad X \quad < A_\alpha] \qquad X \sim N(n\lambda_0, n\lambda_0)$$

$$\approx P\left[\frac{X - n\lambda_0}{\sqrt{n\lambda_0}} < \frac{A_\alpha - n\lambda_0}{\sqrt{n\lambda_0}} \right]$$

$$\approx F\left(\frac{A_\alpha - n\lambda_0}{\sqrt{n\lambda_0}} \right)$$ F fonction de répartition de la loi $N(0,1)$

donc

$$\boxed{A_\alpha \approx F^{-1}(\alpha) \sqrt{n\lambda_0} + n\lambda_0}$$

TD du 7/12/2022

1) Rejet de H_0 si $$\frac{L(x_1, \dots, x_n ; H_1)}{L(x_1, \dots, x_n ; H_0)} > S_\alpha$$

si $$\frac{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x_i^2}{2\sigma_1^2}\right)}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{x_i^2}{2\sigma_0^2}\right)} > S_\alpha$$

si $$\left(\frac{\sigma_0^2}{\sigma_1^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n} x_i^2 + \frac{1}{2\sigma_0^2}\sum_{i=1}^{n} x_i^2\right) > S_\alpha$$

Rejet de H_0 si $$\sum_{i=1}^{n} x_i^2 \underbrace{\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right)}_{>0\,?\ \text{ou} <0\,?} > k_\alpha$$

$\boxed{\text{si } \sigma_1 > \sigma_0}$ alors $\frac{1}{\sigma_1^2} < \frac{1}{\sigma_0^2}$ donc $\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2} > 0$

donc $\boxed{\text{rejet de } H_0 \text{ si } T = \sum_{i=1}^{n} x_i^2 > S_\alpha}$ //

$\boxed{\text{si } \sigma_1 < \sigma_0}$

$\boxed{\text{rejet de } H_0 \text{ si } T = \sum_{i=1}^{n} x_i^2 < S_\alpha}$

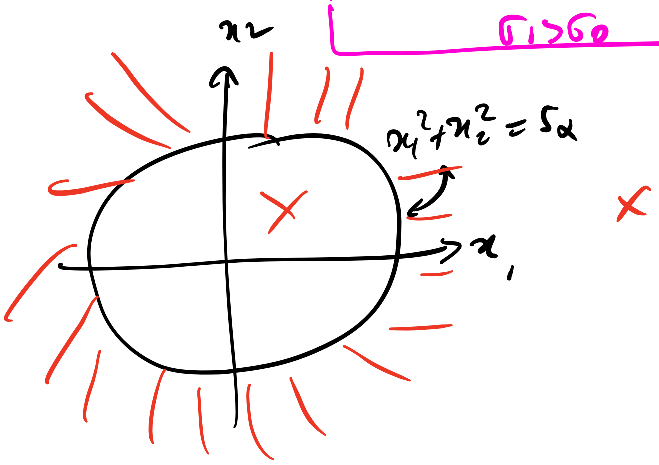
$\boxed{\sigma_1 > \sigma_0}$ $\Downarrow$

$\operatorname{Var} X = \underbrace{E(X^2)}_{} - \underbrace{E^2(X)}_{0}$

Stat de test $T = \sum_{i=1}^{n} x_i^2$

Région critique $\{(x_1 \ldots x_n) / \sum_{i=1}^{n} x_i^2 > S_\alpha\}$

$\sigma_1 > \sigma_0$



2) S_α?

$\beta = P[\text{Rejeter } H_1 \mid H_1 \text{ vraie}]$

$\alpha = P[\overset{\text{Accepter } H_1}{\text{Rejeter } H_0} \mid H_0 \text{ vraie}]$

$= P[\underbrace{\sum_{i=1}^{n} x_i^2}_{T} > S_\alpha \mid \sigma^2 = \sigma_0^2]$

Loi de T sous H_0

Rappel : si $z_1, \ldots, z_n \sim N(0,1)$ alors $\sum_{i=1}^{n} z_i^2 \sim \chi_n^2$

" | $z_1, \dots, z_n$ indépendants //

$$X_i \sim N(0, \sigma^2) \Longrightarrow \frac{X_i - 0}{\sigma} \sim N(0,1)$$

$$\boxed{\sum_{i=1}^{n} \left(\frac{X_i}{\sigma}\right)^2 \sim \chi^2_n}$$

$$\alpha = P\left[\frac{\sum_{i=1}^{n} X_i^2}{\sigma^2} > \frac{S_\alpha}{\sigma^2} \;\middle|\; \sigma^2 = \sigma_0^2 \right]$$

$$\boxed{\alpha = 1 - F_{\chi^2_n}\left[\frac{S_\alpha}{\sigma_0^2}\right]} \Longrightarrow 1 - \alpha = F_{\chi^2_n}\left[\frac{S_\alpha}{\sigma_0^2}\right]$$

$$= \phi\left[\frac{S_\alpha}{\sigma_0^2}\right]$$

$$F^{-1}(1-\alpha) = \frac{S_\alpha}{\sigma_0^2}$$

$$\boxed{S_\alpha = \sigma_0^2 F^{-1}_{\chi^2_n}(1-\alpha)} = \boxed{\sigma_0^2 \phi^{-1}(\alpha)} \quad (*)$$

Calcul de β

$$\beta = P[\text{Rejeter } H_1 \mid H_1 \text{ vraie}]$$

$$= P\left[\sum_{i=1}^{n} X_i^2 \leq S_\alpha \;\middle|\; \sigma^2 = \sigma_1^2 \right]$$

$$= P\left[\frac{\sum_{i=1}^{n} X_i^2}{\sigma^2} \leq \frac{S_\alpha}{\sigma^2} \;\middle|\; \sigma^2 = \sigma_1^2 \right]$$

$$\boxed{\beta = F_{\chi^2_n}\left(\frac{S_\alpha}{\sigma_1^2}\right)} \quad (*)$$

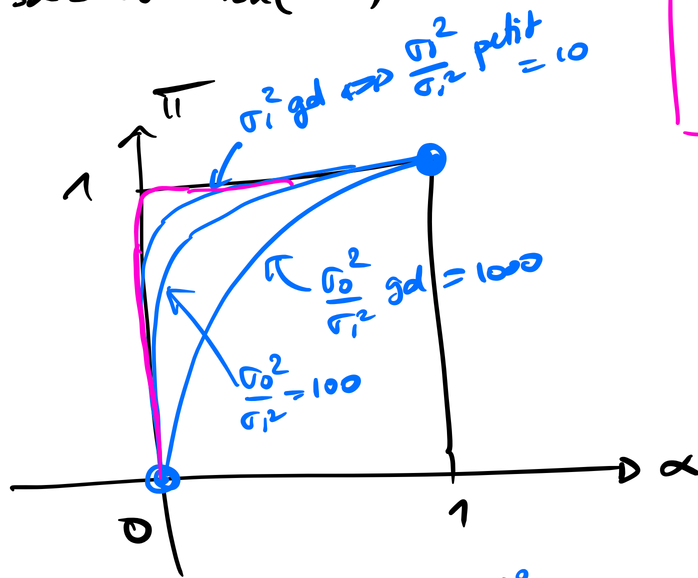
3) $\boxed{\text{COURBES COR}}$ ($\Pi = 1-\beta$ en fonction de α)

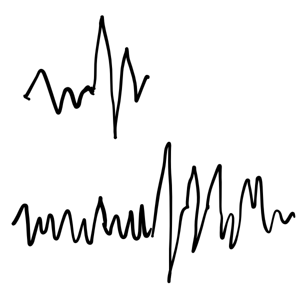
$$\begin{cases} \Pi = 1-\beta = 1 - \overset{=\Phi(}{F_{\chi^2_n}}\left(\dfrac{S_\alpha}{\sigma_1^2}\right) \\ S_\alpha = \sigma_0^2 F^{-1}_{\chi^2_n}(1-\alpha) \end{cases}$$

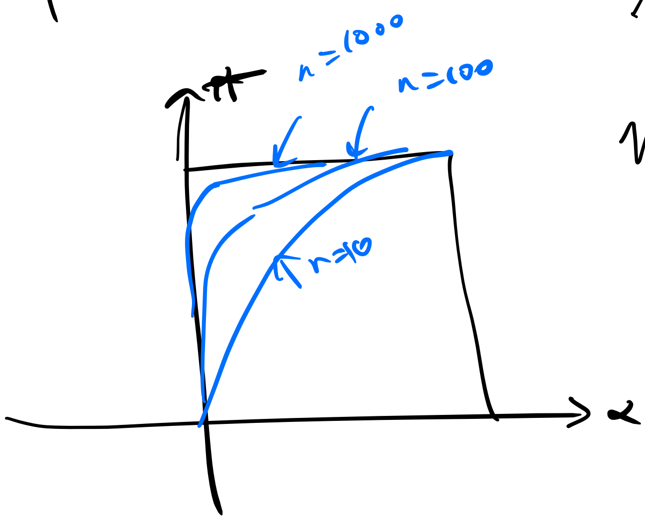
donc

$$\boxed{\Pi = 1 - F_{\chi^2_n}\left(\frac{\sigma_0^2}{\sigma_1^2} F^{-1}(1-\alpha)\right)}$$

$$\boxed{= \Phi_n\left(\frac{\sigma_0^2}{\sigma_1^2}\Phi_n^{-1}(\alpha)\right)}$$







4) $T = \sum_{i=1}^{n} x_i^2$ $\boxed{\text{loi asymptotique de } T}$

↳ quand n est grand $T \sim N(? , ?)$

$$E[T] = E\left[\sum_{i=1}^{n} x_i^2\right] = \sum_{i=1}^{n} \frac{E[x_i^2]}{\sigma^2}$$

$$Var(T) = Var\Big(\overbrace{\sum_{i=1}^{n} x_i^2}^{= n\sigma^2}\Big)$$

$$\overset{x_i \text{ iid}}{=} \sum_{i=1}^{n} var(x_i^2)$$

$$= n \, var(x_1^2) \; \cancel{= n\sigma^2}$$

$$Var(x_i^2) = E[x_1^4] - \underbrace{E[x_i^2]^2}_{\sigma^4}$$

$$\int u^4 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{u^2}{2\sigma^2}\right) du \underset{\text{par partie}}{=} 3\sigma^4$$

$$Var \; x_i^2 = 3\sigma^4 - \sigma^4 = \boxed{2\sigma^4}$$

donc $\boxed{T \sim N(n\sigma^2, 2n\sigma^4)}$

$$\alpha = P[\,T > s_\alpha \mid \sigma^2 = \sigma_0^2\,]$$

$$= P\Big(\underbrace{\frac{T - n\sigma^2}{\sqrt{2n\sigma^4}}}_{U} > \frac{s_\alpha - n\sigma^2}{\sqrt{2n\sigma^4}} \;\Big|\; U \sim N(0,1)\Big]$$

$$\boxed{\alpha = 1 - F\left(\frac{s_\alpha - n\sigma_0^2}{\sqrt{2n\sigma_0^4}}\right)}$$