$\boxed{Ex1}$

1) on a $E[\tilde{m}_n] = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} \underbrace{E[X_k]}_{km} = m$ donc $\tilde{m}_n$ est un estimateur non biaisé de m

$$Var(\tilde{m}_n) = \frac{1}{n^2} \sum_{k=1}^{n} \frac{1}{k^2} \underbrace{Var(X_k)}_{\sigma^2} = \frac{\sigma^2}{n^2} \sum_{k=1}^{n} \frac{1}{k^2}$$

les variables X_k sont indépendants

Mais $\sum_{k=1}^{n} \frac{1}{k^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{n^2} < n$ donc $Var(\tilde{m}_n) < \frac{\sigma^2}{n}$

ce qui signifie qu'on a $\boxed{\lim_{n \to +\infty} Var(\tilde{m}_n) = 0}$. Comme

$\tilde{m}_n$ estimateur non biaisé de m
$Var(\tilde{m}_n) \xrightarrow[n \to +\infty]{} 0$
alors $\boxed{\tilde{m}_n \text{ est un estimateur convergent de } m}$

2) La vraisemblance de $(X_1, \dots, X_n)$ s'écrit

$$L(x; m) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - im)^2}{2\sigma^2}\right)$$

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

avec $x = (x_1, \dots, x_n)$. Donc la log-vraisemblance est

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

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

$$\iff \frac{1}{\cancel{2}\sigma^2} \left[\sum_{i=1}^{n} i x_i - m \sum_{i=1}^{n} i^2\right] \geq 0$$

$$\iff m \leq \frac{\sum_{k=1}^{n} k x_k}{\sum_{k=1}^{n} k^2} = \boxed{\frac{6}{n(n+1)(2n+1)} \sum_{k=1}^{n} k x_k = \hat{m}_{MV}}$$

On a donc le tableau de variation

m	$-\infty$		$\hat{m}_{MV}$		$+\infty$
$\frac{\partial \ln L}{\partial m}$		+	0	−	
$\ln L$		$\nearrow$		$\searrow$	

$\hat{m}_{MV}$ est donc le maximum global unique de la vraisemblance

On en déduit l'estimateur du max de vraisemblance de m

$$\boxed{\hat{m}_{MV} = \frac{6}{n(n+1)(2n+1)} \sum_{k=1}^{n} k X_k}$$

3) $$E[\hat{m}_{MV}] = \frac{6}{n(n+1)(2n+1)} \sum_{k=1}^{n} k \underbrace{E(X_k)}_{km} = \boxed{m}$$

$$Var[\hat{m}_{MV}] = \left[\frac{6}{n(n+1)(2n+1)}\right]^2 \sum_{k=1}^{n} k^2 \sigma^2$$

$$= \boxed{\frac{6\sigma^2}{n(n+1)(2n+1)}} \xrightarrow[n \to +\infty]{} 0$$

donc $\hat{m}_{MV}$ estimateur convergent de m

Comme $Var[\hat{m}_{MV}] \sim \frac{6\sigma^2}{2n^3} = \boxed{\frac{3\sigma^2}{n^3}}$

La variance de $\hat{m}_{MV}$ tend vers 0 lorsque $n \to +\infty$ avec une vitesse très rapide de l'ordre de $\frac{1}{n^3}$. Cette vitesse "s'explique" par le fait que les variables $X_1, \dots, X_n$ n'ont pas la même loi. Cette propriété ne serait pas vérifiée si $(X_1, \dots, X_n)$ était un échantillon

4) On a vu $$\frac{\partial \ln L}{\partial m} = \frac{1}{\sigma^2}\left[\sum_{k=1}^{n} k x_k - m \sum_{k=1}^{n} k^2\right]$$

donc $$\frac{\partial^2 \ln L}{\partial m^2} = -\frac{1}{\sigma^2}\sum_{k=1}^{n} k^2 = -\frac{1}{\sigma^2}\frac{n(n+1)(2n+1)}{6}$$

d'où $$E\left[\frac{\partial^2 \ln L}{\partial m^2}\right] = -\frac{1}{\sigma^2}\frac{n(n+1)(2n+1)}{6}$$

La borne de Cramer-Rao pour un estimateur non-biaisé du paramètre m est

$$\boxed{BCR(m) = \frac{6\sigma^2}{n(n+1)(2n+1)}}$$

$\hat{m}_{MV}$ estimateur non biaisé de m et $Var[\hat{m}_{MV}] = BCR(m)$ donc <u>$\hat{m}_{MV}$ est l'estimateur efficace de m. que l'on préfèrera à $\tilde{m}_n$</u>

5) La loi a posteriori du paramètre m possède la densité

$$p(m|x_1,\dots,x_n) \underset{\text{proportionnel à}}{\propto} p(x_1,\dots,x_n|m)\, p(m)$$

avec $p(m) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{1}{2\sigma_0^2}(m-m_0)^2\right)$

et

$$p(x_1,\dots,x_n|m) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k - km)^2\right]$$

On en déduit

$$\ln p(m|x_1,\dots,x_n) = cste - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-km)^2 - \frac{1}{2\sigma_0^2}(m-m_0)^2$$

C'est une fonction quadratique de m qui admet un maximum global unique vérifiant

$$\frac{\partial \ln p}{\partial m} = 0 \iff \frac{1}{\sigma^2}\sum_{k=1}^{n} k(x_k - km) - \frac{1}{\sigma_0^2}(m-m_0) = 0$$

Soit $\frac{1}{\sigma^2}\sum_{k=1}^{n} k x_k + \frac{m_0}{\sigma_0^2} = m\left[\frac{1}{\sigma_0^2} + \sum_{k=1}^{n}\frac{k^2}{\sigma^2}\right]$

Pour simplifier les notations, on note $S_n = \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$

On en déduit

$$\sum_{k=1}^{n} k x_k + \frac{m_0}{\sigma_0^2} = m\left[\frac{1}{\sigma_0^2} + \frac{S_n}{\sigma^2}\right]$$

$$= m\left[\frac{\sigma^2 + \sigma_0^2 S_n}{\sigma^2\sigma_0^2}\right]$$

d'où $\boxed{m = \frac{\sigma_0^2}{\sigma^2+\sigma_0^2 S_n}\sum_{k=1}^{n} k x_k + \frac{\sigma^2}{\sigma^2+\sigma_0^2 S_n} m_0}$

On a donc

$$\boxed{\hat{m}_{MAP} = \frac{\sigma_0^2 S_n}{\sigma^2+\sigma_0^2 S_n}\hat{m}_{MV} + \frac{\sigma^2}{\sigma^2+\sigma_0^2 S_n} m_0 \quad \text{avec } \beta(n) = \frac{\sigma_0^2 S_n}{\sigma^2+\sigma_0^2 S_n}}$$

On remarque que $\beta(n) \xrightarrow[n\to+\infty]{} 1$ ce qui signifie que $\hat{m}_{MAP}$ et $\hat{m}_{MV}$ ont des comportements asymptotiques similaires, ce qui est un résultat très général

quand $n \to 0$, on a $\beta(n) \approx 0$ donc $\hat{m}_{MAP}$ se comporte comme la moyenne a priori m_0

$\boxed{\text{Ex2}}$ 1) Le théorème de Neyman Pearson indique que

H_0 est rejetée si $$\frac{p(x_1,\dots,x_n\,;\,m_1)}{p(x_1,\dots,x_n\,;\,m_0=0)} > \text{seuil}$$

c'est à dire qu'on rejette H_0 si

$$\frac{\exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k - k m_1)^2\right]}{\exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n} x_k^2\right]} > \text{seuil}$$

soit en prenant le logarithme

$$-\frac{1}{2\sigma^2}\left[\sum_{k=1}^{n} x_k^2 - 2m_1\sum_{k=1}^{n} k x_k + \sum_{k=1}^{n} k^2 m_1^2\right] + \frac{1}{2\sigma^2}\sum_{k=1}^{n} x_k^2 > \text{seuil}$$

soit $$\frac{m_1}{\sigma^2}\sum_{k=1}^{n} k x_k > \text{seuil}$$

Comme $\boxed{m_1 < 0}$ on en déduit

Rejet de H_0 si $\displaystyle\sum_{k=1}^{n} k X_k < S$

La statistique de test est donc $T_n = \sum_{k=1}^{n} k X_k$ et la région critique du test est

$$\left\{(x_1,\dots,x_n) \,/\, \sum_{k=1}^{n} k x_k < S\right\}$$

2) Comme $X_1, \dots, X_n$ sont des variables aléatoires indépendantes, $\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ est un vecteur gaussien et $\sum_{k=1}^{n} k X_k = [1\ 2 \cdots n] \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = A \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ est une variable aléatoire gaussienne car A est une matrice de rang 1.

On en déduit

$$\boxed{\begin{array}{ll} \text{sous } H_0 & T_n \sim (0, \sigma^2 S_n) \quad \text{avec } S_n = \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \\ \text{sous } H_1 & T_n \sim (S_n m_1, \sigma^2 S_n) \end{array}}$$

3) $$\alpha = P[\text{Rejeter } H_0 \mid H_0 \text{ vraie}] = P[T_n < S \mid m = 0]$$
$$= P\left[\frac{T_n}{\sqrt{S_n}\sigma} < \frac{S}{\sqrt{S_n}\sigma} \,\middle|\, \frac{T_n}{\sqrt{S_n}\sigma} \sim N(0,1)\right]$$

donc $$\boxed{\alpha = F\left[\frac{S}{\sqrt{S_n}\,\sigma}\right]}$$

$$\beta = P[\text{Rejeter } H_1 \mid H_1 \text{ vraie}] = 1 - P[T_n < S \mid m = m_1]$$
$$= 1 - P\left[\frac{T_n - m_1 S_n}{\sqrt{S_n}\sigma} < \frac{S - m_1 S_n}{\sqrt{S_n}\sigma} \,\middle|\, m = m_1\right]$$

d'où $$\boxed{\beta = 1 - F\left[\frac{S - m_1 S_n}{\sqrt{S_n}\,\sigma}\right]}$$

4) Les courbes COR expriment $\Pi = 1 - \beta$ en fonction de α. Elles sont définies par

$$\Pi = 1 - \beta = F\left[\frac{S - m_1 S_n}{\sigma\sqrt{S_n}}\right] \quad \text{avec } S = \sqrt{S_n}\,\sigma\, F^{-1}(\alpha)$$

donc $$\boxed{\Pi = F\left[F^{-1}(\alpha) - \frac{m_1\sqrt{S_n}}{\sigma}\right]}$$

Comme $m_1 < 0$, on a $-m_1\sqrt{S_n} > 0$ donc si $\sigma \searrow$ $\frac{-m_1\sqrt{S_n}}{\sigma} \nearrow$ donc $\Pi \nearrow$ ce qui est logique, car plus la variance σ^2 est faible, plus il est facile de tester $m = 0$ contre $m = m_1$.

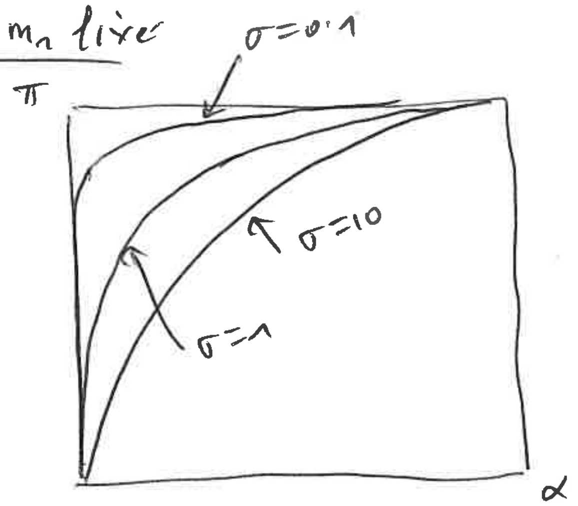
si $m_1 \searrow$ $\quad -\frac{m_1\sqrt{s_n}}{\sigma} \nearrow$ donc $\Pi \nearrow$, ce qui est aussi logique :

plus m_1 est "faible" ("grand" par valeur négative) plus il est facile
car $m_1 < 0$

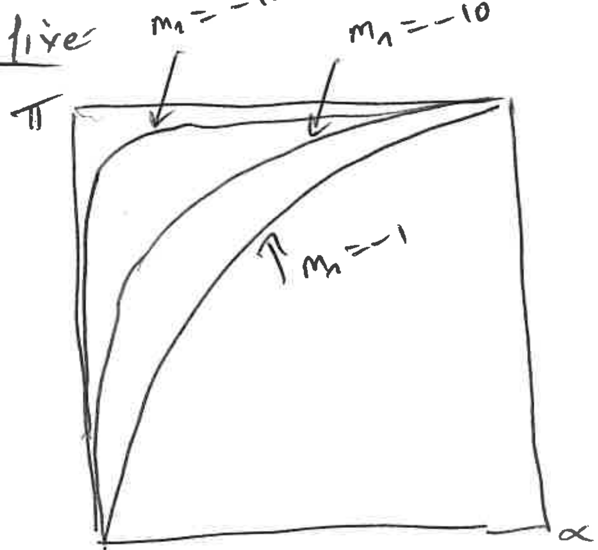
de distinguer $m = m_1$ de $m = 0$

Allure des courbes COR

à m_1 fixe



à σ fixe



5) Dans le cas de $K = 4$ classes, on découpe $\mathbb{R}$ en 4 intervalles
$]-\infty, -a]$ $]-a, 0]$ $]0, a]$ $]a, +\infty[$ équiprobables, i.e., tels
que
$$P[\,Y_k \in]-\infty, -a]\,] = F(-a) = \frac{1}{4}$$
d'où $\boxed{a = -F^{-1}(1/4)}$ 1pt

On va compter combien d'observations $y_1, \dots, y_n$ ~~tomb~~ appartiennent à chaque intervalle. Soit Z_1, Z_2, Z_3 et Z_4 les nombres d'observations dans les 4 intervalles $]-\infty, -a]$ $]-a, 0[$ $]0, a[$ et $[a, +\infty[$. On forme la statistique du test du χ^2

$$\phi = \sum_{k=1}^{K} \frac{(Z_k - n p_k)^2}{n p_k} \qquad (1)$$

avec $p_k = \frac{1}{4}$ et n le nombre total d'observations

On se fixe ensuite un risque α (par exemple $\alpha = 0.01$)

$$\alpha = P[\phi > s_\alpha \mid \phi \sim \chi^2_4] = 1 - F_4[s_\alpha]$$

où F_4 est la fonction de répartition d'une loi du χ^2_4

Soit $\boxed{s_\alpha = F_4^{-1}(1-\alpha)}$

Si $\phi > s_\alpha$ on rejette H_0 avec le risque α choisi et on en conclut que les variables aléatoires $Y_1, \dots, Y_4$ ne respectent pas l'hypothèse $X_k \sim N(km, \sigma^2)$.

si $\phi < s_\alpha$, on accepte H_0 avec le risque α choisi et donc on en conclut que l'hypothèse $X_k \sim N(km, \sigma^2)$ est correcte