EXAMEN STATISTIQUE - 1SN

Mardi 27 Novembre 2019 (14h-15h30)

Partiel sans document (Une feuille A4 recto-verso autorisée)

Exercice 1 : Estimation (10 points)

On considère n observations $x_1, ..., x_n$ issues d'un échantillon $(X_1, ..., X_n)$ distribué suivant la même loi de densité

$$f(x_i; \boldsymbol{\theta}) = \frac{\beta}{\alpha} \exp\left(\beta x_i - \frac{1}{\alpha} e^{\beta x_i}\right), \quad x_i \in \mathbb{R}$$

avec $\alpha > 0$, $\beta > 0$ et $\boldsymbol{\theta} = (\alpha, \beta)^T$. On cherche tout d'abord à estimer le paramètre α à partir des observations $x_1, ..., x_n$ (β est connu dans les 5 premières questions de cet exercice), puis on s'intéresse à l'estimation du vecteur $\boldsymbol{\theta} = (\alpha, \beta)^T$ dans la dernière question de cet exercice.

1. Déterminer l'estimateur du maximum de vraisemblance du paramètre α noté $\widehat{\alpha}_{\text{MV}}$.
2. Montrer que $Y_i = e^{\beta X_i}$ suit une loi gamma dont on déterminera les paramètres. En s'aidant des tables, déterminer la moyenne et la variance de la variable aléatoire Y_i.
3. L'estimateur $\widehat{\alpha}_{\text{MV}}$ est-il un estimateur sans biais et convergent du paramètre α ?
4. Déterminer la borne de Cramer-Rao pour un estimateur non biaisé du paramètre α. L'estimateur $\widehat{\alpha}_{\text{MV}}$ est-il l'estimateur efficace du paramètre α ?
5. On suppose que α est muni d'une loi a priori inverse gamma IG(b, a) définie par

$$p(\alpha) \propto \frac{1}{\alpha^{a+1}} \exp\left(-\frac{b}{\alpha}\right) \mathbb{I}_{\mathbb{R}^+}(\alpha)$$

   où a et b sont deux paramètres connus et $\mathbb{I}_{\mathbb{R}^+}$ est la fonction indicatrice sur $\mathbb{R}^+$ (telle que $\mathbb{I}_{\mathbb{R}^+}(\alpha) = 1$ si $\alpha > 0$ et $\mathbb{I}_{\mathbb{R}^+}(\alpha) = 0$ sinon).

   - Montrer que l'estimateur du maximum a posteriori du paramètre α s'écrit

   $$\widehat{\alpha}_{\text{MAP}} = c_1(n)\widehat{\alpha}_{\text{MV}} + c_2(n)$$

     où $c_1(n)$ et $c_2(n)$ sont deux fonctions de n, a, b que l'on déterminera. Déterminer les limites de $c_1(n)$ et de $c_2(n)$ lorsque $n \to \infty$ et commenter le résultat obtenu.
   - Montrer que la loi de $\alpha | x_1, ..., x_n$ est une loi inverse gamma dont on déterminera les paramètres. En déduire l'estimateur MMSE de α.

   On suppose désormais que α et β sont deux paramètres inconnus que l'on cherche à estimer à partir de l'échantillon $(X_1, ..., X_n)$ et on pose $\boldsymbol{\theta} = (\alpha, \beta)^T$.

6. Montrer que l'estimateur du maximum de vraisemblance de $\boldsymbol{\theta} = (\alpha, \beta)^T$ construit à partir de l'observation des variables aléatoires X_i s'obtient comme la solution du système suivant

$$\alpha = \frac{1}{n} \sum_{i=1}^{n} \exp(\beta x_i), \quad g(\beta; x_1, ..., x_n) = \frac{1}{\beta} + \frac{1}{n} \sum_{i=1}^{n} x_i - \frac{\sum_{i=1}^{n} x_i \exp(\beta x_i)}{\sum_{i=1}^{n} \exp(\beta x_i)} = 0.$$

   Expliquer succinctement comment on peut trouver la solution de ce système.

Exercice 2 : Tests Statistiques (10 points)

On considère n observations $x_1, ..., x_n$ issues d'un échantillon $(X_1, ..., X_n)$ distribué suivant la même loi de Weibull de densité

$$f(x_i; \boldsymbol{\theta}) = \frac{\beta x_i^{\beta-1}}{\alpha} \exp\left(-\frac{x_i^\beta}{\alpha}\right) \mathbb{I}_{\mathbb{R}^+}(x_i),$$

où $\beta > 0$ est un paramètre connu, $\alpha > 0$, et où $\mathbb{I}_{\mathbb{R}^+}(x)$ est la fonction indicatrice sur $\mathbb{R}^+$. L'objectif de cet exercice est d'étudier un test statistique basé sur les observations $x_1, ..., x_n$ qui permet de déterminer si $\alpha = \alpha_0$ ou si $\alpha = \alpha_1 < \alpha_0$. On considère donc le test d'hypothèses

$$H_0 : \alpha = \alpha_0, \quad H_1 : \alpha = \alpha_1 \qquad \text{avec } \alpha_1 < \alpha_0.$$

1. Montrer que la statistique de test du théorème de Neyman-Pearson est

$$T_n = \sum_{i=1}^{n} X_i^\beta$$

et indiquer la région critique de ce test. Représenter cette région critique pour $\beta = 2$ et $n = 2$.

2. Montrer que $Y_i = \frac{2}{\alpha} X_i^\beta$ suit une loi du χ^2 à deux degrés de liberté, i.e., $Y_i \sim \chi_2^2$. En déduire la loi de $U_n = \frac{2}{\alpha} T_n$ sous les deux hypothèses H_0 et H_1.

3. On note $F_{2n}(x)$ la fonction de répartition d'une loi du χ_{2n}^2. Exprimer les risques de première et seconde espèce α et β en fonction du seuil du test de Neyman-Pearson noté S, de F_{2n} et de α_0 et α_1. En déduire les caractéristiques opérationnelles du récepteur (courbes COR) pour ce test et analyser les performances en fonction des valeurs de α_0 et α_1. Donner l'allure de ces courbes pour plusieurs valeurs du couple (α_0, α_1).

4. On désire vérifier que l'hypothèse d'une loi de Weibull pour les variables aléatoires $X_1, ..., X_n$ est correcte. Pour cela, on construit les variables $Y_i = \frac{2}{\alpha} X_i^\beta$ et on cherche à vérifier si ces variables suivent une loi du χ_2^2 à l'aide d'un test de Kolmogorov. Expliquer le principe de ce test (on précisera notamment comment on calcule la statistique de ce test et comment le seuil est calculé à partir d'un risque α donné).

LOIS DE PROBABILITÉ CONTINUES

m : moyenne σ^2 : variance **F. C. :** fonction caractéristique

LOI	**Densité de probabilité**	**m**	σ^2	**F. C.**
Uniforme	$f(x) = \frac{1}{b-a}$ $x \in]a, b[$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\dfrac{e^{itb} - e^{ita}}{it(b-a)}$
Gamma $\Gamma(\theta, \nu)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\theta x} x^{\nu-1}$ $\theta > 0,\ \nu > 0$ $x \geq 0$	$\dfrac{\nu}{\theta}$	$\dfrac{\nu}{\theta^2}$	$\dfrac{1}{\left(1 - i\frac{t}{\theta}\right)^\nu}$
Inverse gamma $\mathrm{IG}(\theta, \nu)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\frac{\theta}{x}} \frac{1}{x^{\nu+1}}$ $\theta > 0,\ \nu > 0$ $x \geq 0$	$\frac{\theta}{\nu-1}$ si $\nu > 1$	$\frac{\theta^2}{(\nu-1)^2(\nu-2)}$ si $\nu > 2$	(*)
Première loi de Laplace	$f(x) = \frac{1}{2} e^{-\|x\|}$	0	2	$\dfrac{1}{1+t^2}$
Normale $\mathcal{N}(m, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Khi2 χ^2_ν $\Gamma\left(\frac{1}{2}, \frac{\nu}{2}\right)$	$f(x) = k e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}$ $k = \dfrac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}$ $\nu \in \mathbb{N}^*,\ x \geq 0$	ν	2ν	$\dfrac{1}{(1-2it)^{\frac{\nu}{2}}}$
Cauchy $c_{\lambda,\alpha}$	$f(x) = \dfrac{1}{\pi\lambda\left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0,\ \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda\|t\|}$
Beta	$f(x) = k x^{a-1}(1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0,\ x \in]0, 1[$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)

LOIS DE PROBABILITÉ DISCRÈTES

$\mathbf{m}$: moyenne σ^2 : variance **F. C. :** fonction caractéristique

$p_k = P[X = k]$ $p_{1,...,m} = P[X_1 = k_1, ..., X_m = k_m]$

LOI	Probabilités	m	σ^2	F. C.
Uniforme	$p_k = \frac{1}{n}$ $k \in \{1, ..., n\}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\dfrac{e^{it}\left(1-e^{itn}\right)}{n\left(1-e^{it}\right)}$
Bernoulli	$p_1 = P[X = 1] = p$ $p_0 = P[X = 0] = q$ $p \in [0,1]$ $q = 1 - p$	p	pq	$pe^{it} + q$
Binomiale $B(n,p)$	$p_k = C_n^k p^k q^{n-k}$ $p \in [0,1]$ $q = 1 - p$ $k \in \{0, 1, ..., n\}$	np	npq	$\left(pe^{it} + q\right)^n$
Binomiale négative	$p_k = C_{n+k-1}^{n-1} p^n q^k$ $p \in [0,1]$ $q = 1 - p$ $k \in \mathbb{N}$	$n\dfrac{q}{p}$	$n\dfrac{q}{p^2}$	$\left(\dfrac{p}{1-qe^{it}}\right)^n$
Multinomiale	$p_{1,...,m} = \frac{n!}{k_1!...k_m!} p_1^{k_1} ... p_m^{k_m}$ $p_j \in [0,1]$ $q_j = 1 - p_j$ $k_j \in \{0, 1, \ldots, n\}$ $\sum_{j=1}^m k_j = n$ $\sum_{j=1}^m p_j = 1$	np_j	Variance : $np_j q_j$ Covariance : $-np_j p_k$	$\left(\sum_{j=1}^m p_j e^{it}\right)^n$
Poisson $P(\lambda)$	$p_k = e^{-\lambda}\dfrac{\lambda^k}{k!}$ $\lambda > 0$ $k \in \mathbb{N}$	λ	λ	$\exp\left[\lambda\left(e^{it}-1\right)\right]$
Géometrique	$p_k = pq^{k-1}$ $p \in [0,1]$ $q = 1 - p$ $k \in \mathbb{N}^*$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	$\dfrac{pe^{it}}{1-qe^{it}}$

Ex1

1) $f(x_1,\dots,x_n;\alpha) = \prod_{i=1}^{n}\left[\frac{\beta}{\alpha} e^{\beta x_i} \exp\left(-\frac{1}{\alpha} e^{\beta x_i}\right)\right]$

$= \left(\frac{\beta}{\alpha}\right)^n \left(\prod_{i=1}^{n} e^{\beta x_i}\right) \exp\left(-\frac{1}{\alpha}\sum_{i=1}^{n} e^{\beta x_i}\right)$

La log-vraisemblance est donc

$\ln f(x_1,\dots,x_n;\alpha) = K - n\ln\alpha - \frac{1}{\alpha}\sum_{i=1}^{n} e^{\beta x_i}$ $\left(\text{soit } K = \ln\left(\beta^n \prod_{i=1}^{n} e^{\beta x_i}\right)\right)$

où K est une fonction indépendante de α

d'où $\frac{\partial \ln f}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha^2}\sum_{i=1}^{n} e^{\beta x_i} \geq 0 \iff \alpha \leq \frac{1}{n}\sum_{i=1}^{n} e^{\beta x_i}$

On en déduit le tableau de variations suivant

α	0		$\frac{1}{n}\sum_{i=1}^{n} e^{\beta x_i}$		$+\infty$
$\frac{\partial \ln f}{\partial \alpha}$		$+$	0	$-$	
$\ln f$		$\nearrow$		$\searrow$	

et donc $\boxed{\hat{\alpha}_{MV} = \frac{1}{n}\sum_{i=1}^{n} e^{\beta X_i}}$ 1pt

2) $Y_i = e^{\beta X_i} \iff X_i = \frac{1}{\beta}\ln Y_i$ est un changement de variables bijectif de $\mathbb{R}$ dans $\mathbb{R}^+$. La densité de Y_i est

$$g(y_i) = \frac{\beta}{\alpha} y_i \exp\left(-\frac{1}{\alpha} y_i\right) \underbrace{\left|\frac{dx_i}{dy_i}\right|}_{\frac{1}{\beta y_i}} \qquad y_i > 0$$

soit $\boxed{g(y_i) = \frac{1}{\alpha}\exp\left(-\frac{y_i}{\alpha}\right) \mathbb{1}_{\mathbb{R}^+}(y_i)}$ 1pt

c'est une loi gamma $\Gamma\left(\frac{1}{\alpha}, 1\right)$ donc $\boxed{E[Y_i] = \alpha}$ et $\boxed{\operatorname{Var} Y_i = \alpha^2}$

1pt

3) On a vu que $\hat{\alpha}_{MV} = \frac{1}{n}\sum_{i=1}^{n} e^{\beta x_i}$ donc $\boxed{\hat{\alpha}_{MV} = \frac{1}{n}\sum_{i=1}^{n} y_i}$ (2)

On en déduit $E[\hat{\alpha}_{MV}] = E[y_i] = \boxed{\alpha}$

et $Var[\hat{\alpha}_{MV}] = \frac{Var\, y_i}{n} = \boxed{\frac{\alpha^2}{n}}$

$\hat{\alpha}_{MV}$ est donc un estimateur non biaisé de α. Puisque $Var[\hat{\alpha}_{MV}] \xrightarrow[n\to+\infty]{} 0$, il est convergent

1pt

4) On a vu $\frac{\partial \ln f}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha^2}\sum_{i=1}^{n} e^{\beta x_i}$

donc $\frac{\partial^2 \ln f}{\partial \alpha^2} = \frac{n}{\alpha^2} - \frac{2}{\alpha^3}\sum_{i=1}^{n} y_i$

On en déduit $E\left[\frac{\partial^2 \ln f(x_1,\dots,x_n;\alpha)}{\partial \alpha^2}\right] = \frac{n}{\alpha^2} - \frac{2n}{\alpha^2} = \boxed{\frac{-n}{\alpha^2}}$

La borne de Cramér-Rao d'un estimateur non biaisé de α est donc

$$\boxed{BCR = \frac{\alpha^2}{n}}$$

1pt

5) La loi a posteriori de α s'écrit

$$p(\alpha | x_1,\dots,x_n) \propto p(x_1,\dots,x_n;\alpha)\, p(\alpha)$$

$$\propto \frac{1}{\alpha^{n+a+1}} \exp\left[-\frac{1}{\alpha}\left(b + \sum_{i=1}^{n} e^{\beta x_i}\right)\right]$$

proportionnel à

L'estimateur MAP est la valeur de α qui maximise $p(\alpha|X_1,\dots,X_n)$. Mais $p(\alpha|X_1,\dots,X_n)$ a la même forme que la vraisemblance avec les changements suivants

- $\frac{1}{\alpha^n} \to \frac{1}{\alpha^{n+a+1}}$

$$\exp\left(-\frac{1}{\alpha}\sum_{i=1}^{n} e^{\beta x_i}\right) \to \exp\left[-\frac{1}{\alpha}\left(b + \sum_{i=1}^{n} e^{\beta x_i}\right)\right]$$

On en déduit

1pt

$$\boxed{\hat{\alpha}_{MAP} = \frac{1}{n+a+1}\left[b + \sum_{i=1}^{n} e^{\beta x_i}\right]}$$

Rq : on peut aussi résoudre $\frac{\partial \ln p(\alpha \mid x_1 \ldots x_n)}{\partial \alpha} = 0$, ce qui conduit au même résultat

On en déduit

$$\boxed{\hat{\alpha}_{MAP} = \underbrace{\frac{n}{n+a+1}}_{C_1(n)} \hat{\alpha}_{MV} + \underbrace{\frac{b}{n+a+1}}_{C_2(n)}}$$

On remarque $C_1(n) \xrightarrow[n \to +\infty]{} 1$ et $C_2(n) \xrightarrow[n \to +\infty]{} 0$.

Donc $\hat{\alpha}_{MAP}$ et $\hat{\alpha}_{MV}$ ont le même comportement asymptotiquement $(n \to +\infty)$, ce qui est un résultat classique

1pt

La loi a posteriori de $\alpha \mid x_1, \ldots, x_n$ a été déterminée précédemment

$$p(\alpha \mid x_1, \ldots, x_n) \propto \frac{1}{\alpha^{a+n+1}} \exp\left[-\frac{1}{\alpha}\left(b + \sum_{i=1}^{n} e^{\beta x_i}\right)\right]$$

C'est une loi inverse gamma : $IG\left(a+n, b + \sum_{i=1}^{n} e^{\beta x_i}\right)$

1pt

On en déduit d'après la table

$$\hat{\alpha}_{MMSE} = E[\alpha \mid x_1, \ldots, x_n]$$

$$\boxed{\hat{\alpha}_{MMSE} = \frac{b + \sum_{i=1}^{n} e^{\beta x_i}}{a+n-1}}$$

6- L'estimateur du maximum de vraisemblance de $\theta = \binom{\alpha}{\beta}$ vérifie (3 bis)

$$\frac{\partial \ln f(x_1,\dots,x_n;\theta)}{\partial \alpha} = 0 \quad \text{et} \quad \frac{\partial \ln f(x_1,\dots,x_n;\theta)}{\partial \beta} = 0$$

avec $\ln f(x_1,\dots,x_n;\theta) = n\ln\beta - n\ln\alpha + \sum_{i=1}^{n} \beta x_i$

$$- \frac{1}{\alpha}\sum_{i=1}^{n} e^{\beta x_i}$$

Donc

$$\frac{\partial \ln f}{\partial \alpha} = 0 \iff \frac{-n}{\alpha} + \frac{1}{\alpha^2}\sum_{i=1}^{n} e^{\beta x_i} = 0$$

$$\iff \boxed{\alpha = \frac{1}{n}\sum_{i=1}^{n} e^{\beta x_i}} \quad (*)$$

$$\frac{\partial \ln f}{\partial \beta} = 0 \iff \frac{n}{\beta} + \sum_{i=1}^{n} x_i - \frac{1}{\alpha}\sum_{i=1}^{n} x_i e^{\beta x_i} = 0$$

En divisant par n et en remplaçant α par son expression $(*)$ on obtient

$$\boxed{\frac{1}{\beta} + \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{\sum_{i=1}^{n}\left(x_i e^{\beta x_i}\right)}{\sum_{i=1}^{n} e^{\beta x_i}} = 0}$$

2pts

Pour résoudre ce système, on peut chercher une solution de l'équation $g(\beta; x_1,\dots,x_n) = 0$ à l'aide d'un algorithme de recherche de zéro comme l'algorithme de Newton-Raphson ce qui conduit à $\hat{\beta}_{MV}$ - On remplace alors cette expression de $\hat{\beta}_{MV}$ dans $(*)$ pour obtenir

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

1pt

1) D'après le théorème de Neyman-Pearson, on rejette H_0 si

$$\frac{f(x_1,\dots,x_n;\alpha_1)}{f(x_1,\dots,x_n;\alpha_0)} > \text{seuil} \iff \left(\frac{1}{\alpha_0}-\frac{1}{\alpha_1}\right)\sum_{i=1}^{n} x_i^{\beta} > \text{seuil}$$

Puisque $\alpha_1 < \alpha_0$, on a $\frac{1}{\alpha_0}-\frac{1}{\alpha_1} < 0$ d'où

$$\boxed{\text{Rejet de } H_0 \text{ si } \sum_{i=1}^{n} x_i^{\beta} < S}$$

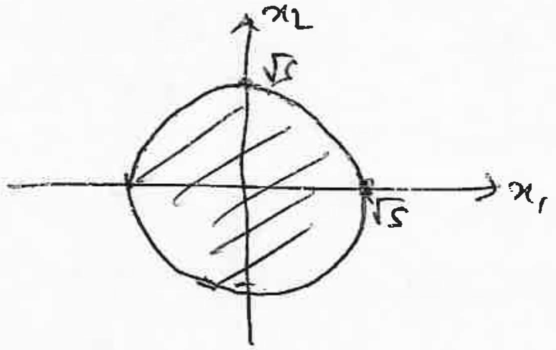
1pt

La statistique de test est donc $T_n = \sum_{i=1}^{n} x_i^{\beta}$ et la région critique est $\left\{(x_1,\dots,x_n) / \sum_{i=1}^{n} x_i^{\beta} < S\right\}$

Pour $n=2$ et $\beta=2$, cette région critique est définie par

$$\left\{(x_1,x_2) \;/\; x_1^2 + x_2^2 < S\right\}$$

qui est l'intérieur d'un disque de rayon $\sqrt{S}$



région critique du test

1pt

2) La loi de $Y_i = \frac{2}{\alpha} X_i^{\beta}$ est de densité

$$\pi(y_i) = \frac{\beta}{\alpha}\left[\left(\frac{\alpha}{2}y_i\right)^{1/\beta}\right]^{\beta-1} \exp\left(-\frac{\alpha}{2}y_i \times \frac{1}{\alpha}\right)\left|\frac{\alpha}{2\beta}\left(\frac{\alpha}{2}y_i\right)^{1/\beta-1}\right|$$

Car $Y_i = \frac{2}{\alpha} X_i^{\beta} \iff X_i = \left(\frac{\alpha}{2}Y_i\right)^{1/\beta}$ est un changement de variables bijectif de $\mathbb{R}^+$ dans $\mathbb{R}^+$ de Jacobien $\left|\frac{dx_i}{dy_i}\right| = \frac{1}{\beta}\times\frac{\alpha}{2}\left|\frac{\alpha}{2}y_i\right|^{\frac{1}{\beta}-1}$

On en déduit

$$\boxed{\pi(y_i) = \begin{cases} \frac{1}{2}\exp\left(-\frac{y_i}{2}\right) & y_i > 0 \\ 0 & y_i \le 0 \end{cases}}$$

1pt

qui est la densité d'une loi du χ^2_2.

La fonction caractéristique de $\frac{2}{\alpha} T_n$ s'écrit

$$\phi_{U_n}(u) = E\left[e^{iU_n u}\right] = E\left[e^{\frac{2i}{\alpha} u \sum_{i=1}^{n} x_i^\beta}\right]$$

$$= \prod_{k=1}^{n} E\left[e^{\frac{2iu}{\alpha} x_k^\beta}\right]$$

ind des variables x_i^β

$$= \prod_{k=1}^{n} \phi_{Y_k}(u)$$

La fonction caractéristique d'une loi du χ^2_2 est $\phi_{Y_k}(u) = \frac{1}{1-2it}$

donc

$$\phi_{U_n}(\alpha) = \frac{1}{(1-2it)^n}$$

1pt

qui est la fonction caractéristique d'une loi du χ^2_{2n}

donc $\boxed{U_n \sim \chi^2_{2n}}$ sous les 2 hypothèses H_0 et H_1

3) $\alpha = P[T_n < s \mid \alpha = \alpha_0]$ ($=$ P[Rejeter H_0 | H_0 vraie])

$$= P\left[U_n = \frac{2}{\alpha} T_n < \frac{2}{\alpha} s \mid \alpha = \alpha_0\right]$$

soit $\boxed{\alpha = F_{2n}\left[\frac{2s}{\alpha_0}\right]}$

1pt

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

$$= P[T_n > s \mid \alpha = \alpha_1]$$

$$= P\left[\frac{2}{\alpha} T_n > \frac{2}{\alpha} s \mid \alpha = \alpha_1\right]$$

d'où

$$\boxed{\beta = 1 - F_{2n}\left[\frac{2s}{\alpha_1}\right]}$$

1pt

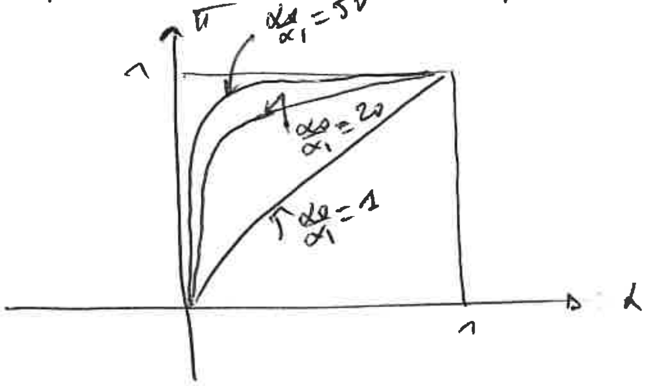
les Courbes COR expriment $\pi = 1-\beta$ en fonction de α - Puisque $\alpha = F_{2n}\left[\frac{2s}{\alpha_0}\right]$, on a $\boxed{s = \frac{\alpha_0}{2} F^{-1}_{2n}(\alpha)}$

d'où

$$\boxed{\pi = 1-\beta = F_{2n}\left[\frac{\alpha_0}{\alpha_1} F^{-1}_{2n}(\alpha)\right]}$$

1pt

F_{2n} est une fonction croissante comme toute fonction de répartition donc Π est une fonction croissante de $\frac{\alpha_0}{\alpha_1}$, ce qui peut s'illustrer comme suit



Plus $\frac{\alpha_0}{\alpha_1}$ est grand, meilleur est le test

4) La loi du χ^2_2 est de fonction de répartition

$$F_0(y) = \begin{cases} \int_0^y \frac{1}{2} e^{-u/2}\, du & u>0 \\ 0 & y<0 \end{cases}$$

c'est-à-dire $\boxed{F_0(y) = \left[1 - \exp(-\tfrac{y}{2})\right] \mathbb{1}_{\mathbb{R}^+}(y)}$ $(*)$

L'idée du test de Kolmogorov est d'évaluer la statistique de test

$$D_n = \sup_{x\in\mathbb{R}} \left|\hat{F}(x) - F_0(x)\right|$$

où $\hat{F}$ est la fonction de répartition empirique des données et F_0 est la fonction de répartition de la loi testée donnée par $(*)$. Comme vu en cours, D_n se calcule comme suit

$$D_n = \max_{i\in\{1,\dots,n\}} \left\{ \left|\hat{F}(x_i^{*+}) - F_0(x_i^*)\right|, \left|\hat{F}(x_i^{*-}) - F_0(x_i^*)\right| \right\}$$

où $x_1^*, \dots, x_n^*$ est l'ensemble des données $x_1, \dots, x_n$ rangées par ordre croissant, i.e., $x_1^* \leq x_2^* \leq \dots \leq x_n^*$

On obtient

$$\boxed{D_n = \max_{i\in\{1,\dots,n\}} \left\{ \left|\frac{i}{n} - 1 + \exp(-\tfrac{x_i^*}{2})\right|, \left|\frac{i-1}{n} - 1 + e^{-\frac{x_i^*}{2}}\right| \right\}}$$

On rejette l'hypothèse H_0 si $D_n > S_\alpha$ où S_α est déterminé comme suit

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

$$= P[D_n > S_\alpha \mid L = L_0]$$

$$= P[U_n = \sqrt{n} D_n > \sqrt{n} S_\alpha \mid U_n \text{ suit la loi de Kolmogorov de fonction de répartition } K]$$

$$= 1 - K(\sqrt{n} S_\alpha)$$

$$\Rightarrow \boxed{S_\alpha = \frac{1}{\sqrt{n}} K^{-1}(1-\alpha)}$$