CORRECTION PARTIEL PROBABILITÉS & STATISTIQUE

15/10/18

Exercice I

1) X est une v.a. à valeurs dans $\mathbb{N}^*$, donc $P(X>t)=1$ si $t<1$.

Soit $k\in\mathbb{N}^*$, si $t\in[k,k+1[$,

$$P(X>t)=P(X>k)=\sum_{i=k+1}^{+\infty}P(X=i)=\sum_{i=k+1}^{+\infty}p(1-p)^{i-1}$$

$$=p(1-p)^k\sum_{i=0}^{+\infty}(1-p)^i=p(1-p)^k\frac{1}{1-(1-p)}=(1-p)^k$$

car $|1-p|<1$

nb: $(X>k)=$ "On a au moins k échecs au départ"

donc il est normal d'avoir $P(X>k)=(1-p)^k$

2) $t\in\mathbb{R}$, $P(Z\leq t)=1-P(Z>t)=1-P(X>t \text{ et } Y>t)$

$=1-P(X>t)P(Y>t)$ par indépendance

$$=\begin{cases}0 & \text{si } t<1\\ 1-\left((1-p)^2\right)^k & \text{si } t\in[k,k+1[\text{ et } k\in\mathbb{N}^*\end{cases}$$

$(1-p)^2=1-2p+p^2=1-(2p-p^2)$

Donc d'après la question 1), Z a pour fonction de répartition la fonction de répartition de la loi $\mathcal{G}(2p-p^2)$.

Pourquoi ce résultat est intuitif :

X loi du premier succès d'une expérience A de proba de succès p
Y ——————————————— B ——————— p
) avec A et B deux expériences indépendantes.

Z loi du premier succès d'une des deux expériences A ou B.

Si on note S_A = "l'expérience A réussit" et S_B = "l'expérience B réussit"

$$P(S_A \cup S_B) = P(S_A) + P(S_B) - P(S_A \cap S_B) = p + p - p^2 = 2p - p^2$$

3) $t \in \mathbb{R}$, $\phi_T(t) = \mathbb{E}\left[e^{it(X+Y-2)}\right] = e^{-2it}\, \mathbb{E}\left[e^{itX}\right] \mathbb{E}\left[e^{itY}\right]$ par indépendance.

$$\mathbb{E}[e^{itX}] = \sum_{k=1}^{+\infty} e^{itk} p(1-p)^{k-1} = e^{it} p \sum_{k=0}^{+\infty} \left(e^{it}(1-p)\right)^k$$

$$= \frac{e^{it} p}{1 - e^{it}(1-p)} \quad \text{car } |e^{it}(1-p)| = |1-p| < 1.$$

Donc $\phi_T(t) = \dfrac{p^2}{\left(1 - e^{it}(1-p)\right)^2}$

<u>Exercice II</u>:

1) $x \in \mathbb{R}$, $f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)\,dy = 0$ si $x \leq 0$

Si $x > 0$

$$f_X(x) = \int_0^x \theta^2 e^{-\theta x}\,dy = \theta^2 e^{-\theta x}\left[y\right]_0^x = \theta^2 x e^{-\theta x}$$

Donc X v.a. à densité, de densité $f_X(x) = \theta^2 x e^{-\theta x} \mathbb{1}_{\mathbb{R}^+}(x)$

$y \in \mathbb{R}$, $f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)\,dy = 0$ si $y < 0$

Si $y \geq 0$

$$f_Y(y) = \int_y^{+\infty} \theta^2 e^{-\theta x}\,dx = \left[-\theta e^{-\theta x}\right]_y^{+\infty} = \theta e^{-\theta y}$$

Donc $Y \sim \mathcal{E}(\theta)$.

On a pas $f_{(X,Y)}(x,y) = f_X(x) f_Y(y)$ p.p donc X et Y ne sont pas indépendantes.

2) Soit $\phi : \mathbb{R}^2 \to \mathbb{R}^+$ mesurable

③

$$\mathbb{E}[\phi(X-Y,Y)] = \iint_{\mathbb{R}^2} \phi(x-y,y) f_{(X,Y)}(x,y)\,dx\,dy = \iint_{\mathcal{D}} \phi(x-y,y)\,\theta^2 e^{-\theta x}\,dx\,dy$$

avec $\mathcal{D} = \{(x,y) \in \mathbb{R}^2 \,/\, 0 < y < x\}$

On veut faire un changement de variable. On a $0<y<x \iff \begin{cases} x-y>0 \\ y>0 \end{cases}$

Donc $\varphi : \mathcal{D} \to (\mathbb{R}^{+*})^2$, $\varphi^{-1} : (\mathbb{R}^{+*})^2 \to \mathcal{D}$

$(x,y) \mapsto (x-y, y)$ $\qquad (u,v) \mapsto (u+v, v)$

φ est un $\mathcal{C}^1$-difféomorphisme. $\quad \mathrm{Jac}\,\varphi^{-1} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1$

D'après la formule du changement de variable on a

$$\mathbb{E}[\phi(X-Y,Y)] = \iint_{(\mathbb{R}^{+*})^2} \phi(u,v)\,\theta^2 e^{-\theta(u+v)}\,du\,dv = \iint_{\mathbb{R}^2} \phi(u,v)\,\theta e^{-\theta u} \mathbb{1}_{u>0}\, \theta e^{-\theta v} \mathbb{1}_{v>0}\,du\,dv$$

Donc ~~(X,Y) est~~ $(X-Y,Y)$ est un vecteur à densité, de densité

$$f_{(X-Y,Y)}(u,v) = \theta e^{-\theta u} \mathbb{1}_{u>0}\, \theta e^{-\theta v} \mathbb{1}_{v>0}$$

en particulier $X-Y$ et $Y \sim \mathcal{E}(\theta)$, et $Y \perp\!\!\!\perp (X-Y)$.

Exercice III

1) X_1 est une v.a positive donc $\mathbb{E}[X_1^a]$ existe toujours mais peut être infini

$$\mathbb{E}[X_1^a] = \int_{\mathbb{R}} x^a \frac{1}{\theta} x^{\frac{1}{\theta}-1} \mathbb{1}_{]0,1[}(x)\,dx = \int_0^1 \frac{x^{\frac{1}{\theta}+a-1}}{\theta}\,dx$$

D'après le critère de Riemann, cette intégrale est infinie lorsque $\frac{1}{\theta}+a-1 \leq -1$ i.e. $\frac{1}{\theta}+a \leq 0$

Si $\frac{1}{\theta}+a>0$,

$$\mathbb{E}[X_1^a] = \left[\frac{1}{\theta}\,\frac{x^{\frac{1}{\theta}+a}}{\frac{1}{\theta}+a}\right]_0^1 = \frac{1}{1+a\theta}$$

2) On est dans le cadre de v.a. i.i.d. à densité, donc

$$V(\theta; x_1, \dots x_n) = \prod_{i=1}^{n} f_\theta(x_i) = \frac{1}{\theta^n}\prod_{i=1}^{n}\left(x_i^{\frac{1}{\theta}-1}\right) \quad \theta\in\mathbb{R}^{+*}, \text{ les } x_i \text{ sont dans }]0,1[.$$

$$h(\theta) = \ln V(\theta; x_1, \dots x_n) = -n\ln\theta + \left(\frac{1}{\theta}-1\right)\sum_{i=1}^{n}\ln x_i$$

h est $\mathcal{C}^\infty$ sur $\mathbb{R}^{+*}$

$$h'(\theta) = -\frac{n}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^{n}\ln x_i \qquad h'(\theta^*)=0 \iff \theta^* = -\frac{\sum_{i=1}^{n}\ln x_i}{n}$$

de plus $h'(\theta)<0$ si $\theta>\theta^*$ et $h'(\theta)>0$ si $0<\theta^*$

donc ~~θ~~ h maximale en l'unique point ~~θ^*~~ $\theta^* = -\frac{\sum_{i=1}^{n}\ln x_i}{n}$ estimation du maximum de vraisemblance.

estimateur du maximum de vraisemblance $\hat{\theta}_n^{MV} = -\frac{\sum_{i=1}^{n}\ln X_i}{n}$ (à valeurs dans $\mathbb{R}^{+*}$).

~~3)~~ Exercice IV

1) T v.a. discrète à valeurs dans $\{0,1\}$
Z ——— $\{-1,0,1\}$ } donc (Z,T) est un vecteur aléatoire discret.

$P(Z=0, T=1) = P(XY=0, X=1) = P(X=0, X=1) = 0$

De même $P(Z=1, T=0) = P(Z=-1, T=0) = 0$

$P(Z=-1, T=1) = P(X=1, Y=-1) = p(1-r)$ par indépendance de X et Y.

$P(Z=0, T=0) = P(X=0) = 1-p$

$P(Z=1, T=1) = P(X=1, Y=1) = pr$

$\mathbb{E}[T] = \mathbb{E}[X] = p.$

$\mathbb{E}[Z] = \mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y] = p(2r-1)$ par indépendance

$\cancel{\mathbb{E}[T^2]}$ $Var(T) = Var(X) = p(1-p)$

$\mathbb{E}[Z^2] = \mathbb{E}[X^2Y^2] = \mathbb{E}[X^2] = p$

Donc $Var(Z) = p - p^2(2r-1)^2$

$\mathbb{E}[ZT] = \mathbb{E}[X^2Y] = \mathbb{E}[XY] = p(2r-1)$

donc $Cov(Z,T) = p(2r-1) - p^2(2r-1) = p(1-p)(2r-1)$

$$Var(T,Z) = \begin{pmatrix} p(1-p) & p(1-p)(2r-1) \\ p(1-p)(2r-1) & p-p^2(2r-1)^2 \end{pmatrix}$$

3) $\mathbb{P}(Z=-1) = \mathbb{P}(Z=-1, T=0) + \mathbb{P}(Z=-1, T=1) = p(1-r)$

De même $\mathbb{P}(Z=0) = \mathbb{P}(Z=0, T=0) + \mathbb{P}(Z=0, T=1) = (1-p)$

$\mathbb{P}(Z=1) = \dots = pr.$

4) Nous sommes dans le cadre de v.a. discrètes.

$$V(p,r;z_1,\dots,z_n) = \prod_{i=1}^{n} \mathbb{P}_{p,r}(Z_i = z_i) = p^{n_{-1}}(1-r)^{n_{-1}}(1-p)^{n_0}p^{n_1}r^{n_1}$$

avec $n_k = \#\{i \in [\![1,n]\!] \,/\, z_i = k\}$

en particulier $n_{-1} + n_0 + n_1 = n$

Remarquons que si $z_1 = \dots = z_n = 0$ alors $n_1 = n_{-1} = 0$ et donc $V(p,r;0,\dots,0)$ ne dépend pas de r ! Ainsi la fonction est maximale (en r) pour tout $r \in]0,1[$, il n'y aura pas unicité de l'EMV. On peut prendre par exemple $1/2$ (pour $\hat{r}$).

$$h(p,r) = \ln V(p,r;x_1,\dots,x_n) = \underbrace{(n_{-1}+n_1)\ln p + n_0 \ln(1-p)}_{h_1(p)} + \underbrace{n_{-1}\ln(1-r) + n_1 \ln r}_{h_2(r)}$$

On peut se contenter de maximiser h_1 et h_2 !

h_1 et h_2 sont $\mathcal{C}^\infty$ sur $]0,1[$.

$$h_1'(p) = \frac{n_{-1}+n_1}{p} - \frac{n_0}{1-p} \quad \text{et } h_1'(p^*) = 0 \iff p^* = \frac{n_{-1}+n_1}{n}$$

De plus $h_1''(p) < 0$ la fonction est concave.

$$h_2'(r) = -\frac{n_{-1}}{1-r} + \frac{n_1}{r} \qquad h_2'(r^*) = 0 \iff r^* = \frac{n_1}{n_1+n_{-1}}$$

(on a traité à part le cas $n_1+n_{-1}=0$)

De plus $h_2''(r) < 0$ la fonction est concave. (encore une fois si $n_1+n_{-1} \neq 0$)

Donc on a un estimateur du maximum de vraisemblance

$$\begin{cases} \hat{p}_n^{MV} = \dfrac{N_1+N_{-1}}{n} \\ \hat{r}_n^{MV} = \begin{cases} \dfrac{N_1}{N_1+N_{-1}} & \text{si } N_1+N_{-1} > 0 \\ 1/2 & \text{sinon} \end{cases} \end{cases}$$

et on a non unicité si $N_1+N_{-1} = 0$!