Examen de Proba Stat 2024

Exercice 1 : $P(X_1 = 1) = p \quad P(X_1 = 0) = 1 - p$

1.1 : $\mathbb{E}(X_1) = 0.(1-p) + 1 \times p = p \qquad \mathbb{E}(\mathrm{Var}(X_1)) = \mathbb{E}((X_1-p)^2) = p(1-p)^2 + (1-p)(-p)^2 = p(1-p)$

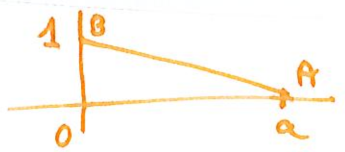
1.2 : $\mathbb{E}(\overline{X}) = \frac{1}{n}\mathbb{E}\left(\sum X_k\right) = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(X_k) = \frac{np}{n} = p$

$\mathrm{Var}(\overline{X}) = \frac{1}{n^2}\mathrm{var}\left(\sum_{i=1}^{n} X_i\right) \underset{\text{indépendance}}{=} \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{n\,p(1-p)}{n^2} = \frac{p(1-p)}{n}$.

1.3 : D'après l'approximation centrale $\frac{\overline{X}-p}{\sqrt{p(1-p)}}\sqrt{400}$ suit une loi proche d'une $\mathcal{N}(0,1)$

Donc $\overline{X}$ suit une loi proche d'une $\mathcal{N}\left(\frac{1}{5} ; \left(\frac{2}{5} \times \frac{1}{20}\right)^2\right) = \mathcal{N}\left(\frac{1}{5} ; \left(\frac{1}{50}\right)^2\right)$

Exercice 2 : Aire (OAB) : $\frac{a}{2}$; Equation de (AB) : $y = \frac{1}{a}(a-x)$.



2.2 : $P((X,Y) \in \mathbb{R}^2) = 1$ or $P((X,Y)\in\mathbb{R}^2) = \iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = \iint_{(OAB)} K\,dx\,dy = \frac{Ka}{2}$

donc $K = \frac{2}{a}$.

2.3 : $P(X > \frac{a}{2}$ et $Y > \frac{1}{2}) = 0$ mais $P(X > \frac{a}{2}) \neq 0$ et $P(Y > \frac{1}{2}) \neq 0$

on peut aussi calculer la covariance de (X,Y) mais c'est plus long.

2.4 : $f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy$ Cela se démontre $F_X(x) = P(X \leq x) = P(X \leq x ; Y \in \mathbb{R}) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(x,y)\,dy\,dx$

donc $f_X(x) = 0$ si $x \notin [0;a]$ et si $x \in [0;a]$, $f_X(x) = \int_0^{\frac{1}{a}(a-x)} K\,dy = \frac{2}{a^2}(a-x)$

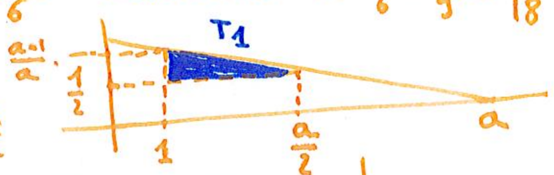
2.5 : $\mathbb{E}(X) = \int_{\mathbb{R}} x f_X(x)\,dx = \int_0^a \frac{2}{a^2}(a-x)\,x\,dx = \frac{2}{a^2}\left[a\frac{x^2}{2} - \frac{x^3}{3}\right]_0^a = 2a\left(\frac{1}{2}-\frac{1}{3}\right) = \frac{2a}{6} = \frac{a}{3}$

$\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 f_X(x)\,dx = \frac{2}{a^2}\int_0^a a x^2 - x^3\,dx = \frac{2}{a^2}\left(\frac{a^4}{3} - \frac{a^4}{4}\right) = \frac{a^2}{6} \qquad \mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{a^2}{6} - \frac{a^2}{9} = \frac{a^2}{18}$.

2.6 : $P(X > 1$ et $Y > \frac{1}{2}) =$ aire du triangle $(T_1) \times K$

$= \frac{1}{2}\left(\frac{a-1}{a} - \frac{1}{2}\right)\left(\frac{a}{2} - 1\right) \times K = \frac{1}{2}\left(\frac{a-2}{2a}\right)\left(\frac{a-2}{2}\right)\frac{2}{a}$

$= \left(\frac{a-2}{2a}\right)^2$



$P(Y > \frac{1}{2}) = K \times \frac{1}{2} \times \frac{a}{2} \times \frac{1}{2} = \frac{1}{4} \qquad P(X > 1 \mid Y > \frac{1}{2}) = \left(\frac{a-2}{a}\right)^2$.

Exercice 3 :

3.1 : Les résidus ne semblent pas dépendre des abscisses et ne pas être liés, ils sont bien rapprochés de 0 et une distribution normale est possible, le modèle semble donc acceptable. Seul point un peu étrange la vitesse ultrasonique est déterministe dans le modèle or on peut penser que la mesure de la vitesse est aussi entachée d'une erreur aléatoire.

3.2 : $H_0 : \beta_1 = 0$ et $H_1 : \beta_1 \neq 0 \qquad \beta_1 = 0$ correspond à "non corrélées".

Si H_0 est vrai la variable aléatoire F suit une loi de Fisher de paramètre $(1;367)$ or ici $F(\omega) = 2936$ si $F \sim \mathcal{F}(1;367)$ $\mathbb{P}(F > 2936) = 6,42.10^{-177}$ extrêmement faible, donc si les hypothèses de la regression sont vérifiées il n'y a aucun doute on refuse H_0, on peut alors affirmer que vitesse ultrasonique et log de la résistance sont corrélées.

3.3 : $I_{95} = [1,042 ; 1,121]$

3.4 : $\log R = -1,42 + 1,082V$ d'où $R \simeq e^{-1,42+1,082V} \simeq 0,242\, e^{1,082V}$

3.5 : Sous les hypothèses de la regression

$T_n = \sqrt{n}\,\frac{\hat{y}_{n+1} - y_{n+1}}{\hat{\sigma}\sqrt{1+\left(\frac{x_{n+1}-\overline{x}}{\sigma_x}\right)^2}}$ suit une loi de Student à $n-2$ ddl.

ici $\hat{y}_{n+1} = -1,42 + 1,082 \times 4 \simeq 2,908 \qquad I_{95} = \left[\hat{y}_{n+1} - 1,96\frac{\hat{\sigma}}{\sqrt{n}}\sqrt{1+\left(\frac{x_{n+1}-\overline{x}}{\sigma_x}\right)^2} ; \hat{y}_{n+1} + 1,96\frac{\hat{\sigma}}{\sqrt{n}}\sqrt{}\right]$

avec $\hat{\sigma} \simeq 0,139 \quad n = 369 \quad x_{n+1} = 4 \quad \overline{x} = 4,343 \quad \sigma_x = 0,3634$

on obtient alors $I_{95}(\omega) = [2,888 ; 2,928]$ qui est un intervalle de confiance pour le log de la resistance, on a donc pour la resistance $J_{95} = [e^{2,888} ; e^{2,928}] \simeq [17,95 ; 18,69]$. C'est à dire que si notre echantillon fait partie des 95% les plus proche de la moyenne la valeur moyenne de $\log(R)$ appartient à $[2,888 ; 2,928]$

3.6 : A la variance dûe à l'echantillon, il faut rajouter la variance due à notre nouvel essai, ces variances s'ajoutent par indépendance

$\sigma_1^2 \simeq \hat{\sigma}^2 + \frac{\hat{\sigma}^2}{n}\left(1 + \left(\frac{x_{n+1}-\overline{x}}{\sigma_x}\right)^2\right)$ d'où $I'_{95} \simeq [2,635 ; 3,18]$

$J'_{95} \simeq [13,9 ; 24,1]$