**Université Abou Bekr Belkaïd, Tlemcen.**

Département de mathématiques

Contrôle continu : **Statistiques Mathématiques -1h30**

**Master 1 : Statistiques et Probabilités Approfondies – 2021/2022**

**Exercice 1.**

**Partie I.**

1. Soit **X** une v.a de loi paramétrée $\mathbf{P}_\theta$ayant pour densité de probabilité

$$f(x,\theta)=\theta\alpha^{\theta}x^{-(\theta+1)}, x\geq\alpha>0.$$

Où $\alpha$ est une constante donnée.Pour le n-échantillon $(X_1,X_2,\ldots,X_n)$ issu de la v.a, construire un estimateur $\widehat{\theta_n}$pour le paramètre $\theta$.

2. En déduire une statistique exhaustive, associée au modèle, que l'on notera T(X). Dire si $\widehat{\theta_n}$ peut s'écrire en fonction de la statistique T(X). Conclure en justifiant.

**Partie II.**

1. On considère le n-échantillon $(X_1, X_2,\ldots, X_n)$ dont les v.a $(X_i, 1\leq i\leq n)$ obéissent à une loi exponentielle de paramètre $\alpha$inconnu. On cherche à construire un estimateur S(X) pour le paramètre $e^{-\alpha}$, pour cela, on propose un nouvel échantillon

$(Y_1, Y_2,\ldots, Y_n)$ où pour $1\leq i\leq n$ :

$$Y_i=\begin{cases}1 & si\quad X_i>1\\ 0 & \quad sinon\end{cases}$$

1. Construire un estimateur pour le paramètre $e^{-\alpha}$. Est-il sans biais ?
2. Calculer le risque associé à cette estimation.

**Exercice 2.**

**Partie I.**

Un petit avion commercial, peut accueillir 30 passagers. Une étude effectuée a confirmé que 20% des clients parmi 30 ayant réservé ne se présentent pas à l'heure de l'embarquement et cela pose un problème pour la compagnie aérienne chargée de ce vol. On propose la v.a **Z**qui représente le nombre de clients qui ne viennent pas sachant qu'ils ont déjà réservé.

1. Trouver la loi de la v.a **Z**et préciser ses paramètres de dispersion.
2. À un niveau de confiance $1-\alpha=95\%$, proposer un intervalle de confiance permettant l'estimation **du nombre de clients** à prévoir.

**Partie II.**

À une boulangerie, on désire acheter 100 baguettes de pain dont le poids moyen est de 248.5 grammes et une variance de 25 grammes. On veut savoir si le poids est considérablement respecté, autrement dit, on veut valider si une baguette de pain pèse 250 grammes (ellepeut peser moins).

À niveau égale à 3% , construire un test classique associé à nos hypothèses.

M1 - Stat. Proba/App — Corrigé-type - Epreuve Contrôle-Continu (1)

2021-2022

**Exercice 01:** (10 pts)

**Partie I:** $X \sim \mathbb{P}_\theta$ de d.d.p $f(x,\theta) = \theta \alpha^\theta x^{-(\theta+1)}$, $x \geq \alpha$, $\alpha > 0$ ($\alpha$ donnée)

1. Construire un estimateur $\hat{\theta}_n(X)$ pour $\theta$: Soit le néchantillon $(X_1, \dots, X_n)$ issu de X.

Calculons d'abord $\mathbb{E}(X_i)$

$$\mathbb{E}(X_i) = \theta \alpha^\theta \int_{\mathbb{R}} x_i \cdot x_i^{-(\theta+1)} \mathbb{1}_{[\alpha,+\alpha[}(x_i)\, dx_i$$

$$= \theta \alpha^\theta \int_\alpha^{+\alpha} x_i^{-\theta} dx_i = \theta \alpha^\theta \left[ -\frac{\alpha^{-\theta+1}}{1-\theta} \right] = \frac{\alpha\theta}{\theta-1}, \theta > 0\ (\theta \neq 1), \alpha > 0$$

Par la méthode des moment, on peut estimer convenablement le paramètre $\theta$ comme suit:

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i = \frac{\alpha\theta}{\theta-1}$$, par conséquent

$$\boxed{\hat{\theta}_n(X) = \frac{\overline{X}_n}{\overline{X}_n - \alpha}, \alpha > 0}$$ (0,5 pts)

2. Le fait que $\mathbb{P}_\theta$ appartient aux familles exponentielles, nous permet de conclure une statistique exhaustive que l'on notera $T(X)$.

En effet: La vraisemblance s'écrit sous la formes

$$\mathcal{L}(x,\theta) = \prod_{i=1}^n f(x_i,\theta) = \underbrace{(\theta\alpha^\theta)^n}_{a(\theta)} e^{\underbrace{-(\theta+1)}_{b(\theta)} \underbrace{\sum_{i=1}^n \ln(x_i)}_{T(X)}} \underbrace{\prod_{i=1}^n \mathbb{1}_{[\alpha,+\alpha[}(x_i)}_{h(x)}$$

$$x_i^{-(\theta+1)} = e^{-(\theta+1)\ln(x_i)}, x_i \geq \alpha > 0.$$

→ Remarque que $\hat{\theta}_n(X)$ ne s'écrit pas en f^ct de la statistique exhaustive $T(X)$, par suite, le théorème de R-Blackwell permet

→ Il est bien claire que de conclure l'existance d'un estimateur $T^*$ plus efficace que $\hat{\theta}_n(X)$

tel que: $\underbrace{T^*}= \mathbb{E}\left(\hat{\theta}_n(X) / T(X)\right)$

qui a de variance minimale. (0,5 pts)

## Partie II :

1. Considérons $(X_1, \dots, X_n)$ telle que $X_i$ sont i.i.d suivant $\mathcal{E}(\alpha)$, $\alpha$ inconnue. $\alpha > 0$

Pour construire un estimateur $S(X)$ pour le paramètre $e^{-\alpha}$, on définit la v.a

$$Y_i = \mathbb{1}_{\{X_i > 1\}}, \quad i \in \{1, \dots, n\}$$

Sachant que $\forall i \in \{\overline{1,n}\}$, $Y_i \sim \mathcal{B}(p)$, où $p = \mathbb{P}(Y_i = 1) = \mathbb{P}(X_i > 1)$
$= 1 - F_{X_i}(1) = 1 - (1 - e^{-\alpha}) = \underbrace{e^{-\alpha}}_{\text{inconnu}}$

par conséquent $p = e^{-\alpha}$ est inconnu. Or $\mathbb{E}(Y_i) = p$. Un estimateur classique qu'on puisse proposer est $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

$\overline{Y}_n$ est effectivement sans biais pour $e^{-\alpha}$, puisque $\mathbb{E}(\overline{Y}_n) = \mathbb{E}(Y_i) = e^{-\alpha}$.

0,5 pt

2. Le risque associé à cet estimation est :

$$R(\overline{Y}_n, p) = \mathrm{Var}(\overline{Y}_n) + \cancel{\text{biais}^2}_{=0}$$

$$= \frac{1}{n} e^{-\alpha}(1 - e^{-\alpha})$$

Conclusion : pour n assez grand, $R(\overline{Y}_n, p) \xrightarrow[n \to \infty]{} 0$ 0,5 pt

## Exercice 02 : 10 pts

### Partie (I) :

$n = 30$ passagers. Soit Z v.a qui indique le nombre de passagers qui ne viennent pas

1. $Z \sim \mathcal{B}(n, p)$, $n = 30$, $p = 20\%$, $\mathbb{E}(Z) = n \cdot p = 6$, $\mathrm{Var}(Z) = np(1-p) = 4.8$

01 pt

2. En se donnant un niveau de confiance $1 - \alpha = 95\%$ $(\alpha = 5\%)$ l'intervalle de confiance I permettant l'estimation (encadrement) du nombre de client à prévoir est décrit comme suit :

Soit $p_1 = 1 - p$ proportion des clients qui se présentent pas.

$p_1$ : proportion des clients qui sont prévus

D'où, le nombre à prévoir sera donné par :

$$N_{\text{à prévoir}} = n \cdot p_1$$

Par suite, nous allons construire un intervalle de confiance pour $p$ ensuite passer à $p_1 = 1-p$ et enfin à $n_{\text{à prévoir}}$.

Ainsi, d'après le cours :

$$I_{(\text{pour } p)} = \left[ \overline{X}_n - 1.96\sqrt{\frac{\overline{X}_n(1-\overline{X}_n)}{n}} \, , \, \overline{X}_n + 1.96\sqrt{\frac{\overline{X}_n(1-\overline{X}_n)}{n}} \right]$$

où $\mathbb{P}(p \in I) = 1-\alpha = 0.95$, $\overline{X}_n = \hat{p}_n = 0.2$, $n = 30$

Par conséquent : $I_{(\text{pour } p)} = [0.0574, 0.342]$, donc

$I_{(\text{pour } p_1)} = [0.658, 0.942]$ et alors

$I_{\text{pour } (n_{\text{à prévoir}})} = [20, 28]$ car $30 \times 0.658 = 19.74 \simeq 20$ passagers et $30 \times 0.942 = 28.26 \simeq 28$ passagers

0.4 pts

Partie (II) :

$\alpha = 3\%$, le test effectué est celui qui défend :

$H_0 : m = m_0 = 250\,(gr)$ ; contre $H_1 : m < m_0 = 250$

. Sachant que la moyenne (poids moyen) observée est : $\overline{X}_n(\omega) = \overline{x}_n = 248.5\,(gr)$

$\sigma = 5\,(gr)$

. On désigne par la v.a $X$ : poids des baguettes.

$n = 100$ baguettes (assez grand pour appliquer le T.L.C)

Donc $\sqrt{n}\left(\frac{\overline{X}_n - m}{\sigma}\right) \underset{n \text{ assez grand}}{\overset{Loi}{=}} \mathcal{N}(0,1)$

. La région de rejet de l'hypothèse $H_0$ est : $R = \{X / \overline{X}_n < k\}$

. Sous $H_0$ : $\mathbb{P}_{m_0}(R) = 3\% = 0.03$ donc $\mathbb{P}\left(N < \underbrace{10\left(\frac{k-250}{5}\right)}_{2(k-250) = q_{1-\alpha}}\right) = 0.03$ où $N \sim \mathcal{N}(0,1)$

Puisque $F(q_{1-\alpha}) = 1 - F(q_{1-\alpha})$, $F$ f.d.r de $N$, donc d'après la table $q_{1-\alpha} = -1.88$ (quantile d'ordre 0.97)

donc $k = 249.06$, Ainsi $R = [0, 250.94[$

Puis $\overline{x}_n = 248.5 \in R$, alors $H_0$ est repoussée

(i.e baguette de pain pèse en moyenne moins que 250 gr)