Examen, *Janvier 2018*

- Les formules employées doivent être rappelées sur la copie.
- Aucun document n'est autorisé.
- Les résultats doivent être présentés 4 chiffres après la virgule.

Exercice #1 (5 points)

Lors d'essais d'accès à l'Internet à partir du réseau de l'IHEC, il est constaté que 90% des essais permettent une connexion correcte. Un enseignant a besoin de se connecter une fois par jour. Si X désigne le nombre d'essais nécessaires, alors X suit une loi géométrique de paramètre $p = 0.90$ et notée $X \sim G(p)$ telle que

$$P(X = x) = p(1-p)^{x-1}, \text{ pour tout } x \in N^{\star}.$$

1. Si la probabilité de se connecter p est inconnue, donner l'estimateur de p par la méthode du maximum de vraisemblance, pour un échantillon de $n-$observations de X.
2. Montrer que cet estimateur est sans biais et convergent.
3. Pour n assez grande, quelle est la loi de l'inverse de cet estimateur?

Exercice #2 (5 points)

$X_1, X_2, \cdots, X_n$ étant des variables aléatoires i.i.d. selon une loi uniforme telle que $X_i \sim U(\alpha, \beta)$, avec α et β sont deux paramètres inconnus.

1. Formuler la densité de probabilité de la loi $U(\alpha, \beta)$.
2. Montrer que $E(X_i) = \frac{\alpha+\beta}{2}$.
3. Sachant que $V(X_i) = \frac{(\beta-\alpha)^2}{12}$, chercher les estimateurs $\hat{\alpha}$ et $\hat{\beta}$ de α et β respectivement par la méthode des moments.

Exercice #3 (10 points)

Le siège vacant dans un vol d'avion induit une perte de revenu pour la compagnie aérienne. Pour cette raison une grande compagnie veut estimer le nombre moyen des sièges vides par vol durant l'année dernière. Pour y répondre, 225 enregistrements sur les vols sont choisis aléatoirement et le nombre de sièges vides est donné pour chaque vol de l'échantillon. Les statistiques dérivées de l'échantillon sont

$$\bar{x} = 11.6 \text{ sièges et } s = 4.1 \text{ sièges.}$$

1. Estimer la moyenne μ par un intervalle de confiance des sièges vides par vol avec $\alpha = 0.10$.
2. Interpréter cet intervalle.
3. Expliquer la logique de toute hypothèse considérée.
4. Si un siège vide coûte $185DT$ pour $100km$ de vol pour la compagnie et si la destination moyenne des vols de cette compagnie est $1850kms$, évaluer la perte moyenne par vol.

Examen, *Janvier 2018*

Exercice #1 (5 points)

1. La densité de l'échantillon observé est la fonction de vraisemblance

1
$$L(x_1, x_2, \cdots, x_n, p) = p^n(1-p)^{\sum_{i=1}^n x_i - n}.$$

Sa forme logarithmique est

$$LogL(x_1, x_2, \cdots, x_n, p) = nLogp + Log(1-p)\sum_{i=1}^n x_i - nLog(1-p).$$

La dérivée par rapport au paramètre à estimer est

$$\frac{\partial LogL(x_1, x_2, \cdots, x_n, p)}{\partial p} = \frac{n}{p} - \frac{\sum_{i=1}^n x_i}{1-p} + \frac{n}{1-p} = 0.$$

La résolution donne

1
$$\hat{P} = \frac{n}{\sum_{i=1}^n x_i},$$

comme estimateur du maximum de vraisemblance pour p.

2. Etant donnée que

1
$$E(\hat{P}) = \frac{n}{\sum_{i=1}^n E(x_i)} = \frac{n}{n(1/p)} = p,$$

$\hat{P}$ est un estimateur sans biais de p. Par ailleurs, il est possible de procéder autrement en disant que $\hat{P}$ est l'estimateur du maximum de vraisemblance pour p donc $\hat{P}$ est sans biais et convergent.
1

3. L'inverse de cet estimateur est

$$\frac{1}{\hat{P}} = \bar{X}.$$

Pour n assez grande, $\bar{X}$ obéit à la loi normale telle que

1
$$\bar{X} \sim N\left(\frac{1}{p}, \frac{1-p}{np^2}\right).$$

Exercice #2 (5 points)

1. Si $X_i \sim U(\alpha, \beta)$, avec α et β sont deux paramètres inconnus, alors

1
$$f(x) = \begin{cases} \frac{1}{\beta-\alpha}, & \forall x \in [\alpha, \beta] \\ 0, & \text{sinon.} \end{cases}$$

2. Par définition

$$E(X) = \int_{\alpha}^{\beta} x f(x) dx = \left[\frac{x^2}{2(\beta - \alpha)} \right]_{\alpha}^{\beta} \quad \text{donc} \quad E(X) = \frac{\alpha + \beta}{2}.$$

1

De même,

$$E(X^2) = \int_{\alpha}^{\beta} x^2 f(x) dx = \left[\frac{x^3}{3(\beta - \alpha)} \right]_{\alpha}^{\beta}$$

donc $E(X^2) = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} = \frac{\beta^2 + \alpha\beta + \alpha^2}{3}$. Il s'ensuit

$$V(X) = E\left(X^2\right) - [E(X)]^2 = \frac{\beta^2 + \alpha\beta + \alpha^2}{3} - \frac{(\alpha + \beta)^2}{4} = \frac{(\beta - \alpha)^2}{12}.$$

3. Lorsque $E(X^2) = V(X) + [E(X)]^2 = \frac{(\beta-\alpha)^2}{12} + \frac{(\alpha+\beta)^2}{4}$, l'estimation par la méthode des moments suppose

1

$$\begin{cases} \bar{X} = \frac{\alpha+\beta}{2}, \\ \frac{1}{n} \sum_{i=1}^{n} X_i^2 = \frac{(\beta-\alpha)^2}{12} + \frac{(\alpha+\beta)^2}{4} \end{cases}$$

Si S désigne l'écart type empirique tel que $S = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}$, le système précédent devient

1

$$\begin{cases} \alpha + \beta = 2\bar{X}, \\ \beta - \alpha = S\sqrt{12} \end{cases}$$

1

ce qui donne $\hat{\alpha} = \bar{X} - S\sqrt{3}$ et $\hat{\beta} = \bar{X} + S\sqrt{3}$.

Exercice #3 (10 points)

1. Les données concernent un échantillon d'enregistrements de 225 observations. Si X est une variable aléatoire désignant "*le nombre de sièges vides pour chaque vol de l'échantillon*" dont la loi est $N(\mu, \sigma^2)$, telles que la moyenne μ et la variance σ^2 sont des paramètres inconnus.

1

Le fait d'avoir $X_i \sim N(\mu, \sigma^2)$ permet de déduire que $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ d'où

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1),$$

où $n = 225$. Or, l'écart type σ est inconnu, l'utilisation de son estimateur s telle que

1

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

donne une fonction pivotale

$$Z = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1).$$

La table de Student ne fournit pas une valeur correspondant à un degré de liberté assez grand comme $n - 1 = 224$. Donc pour procéder, la distributiobn de Student est approchée par la loi normale standard $N(0,1)$,

$$Z = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim N(0,1).$$

Les valeurs z_1 et z_2 se déterminent par la probabilité $P(z_1 \leq Z \leq z_2) = 1 - \alpha$, sachant que le risque de ne pas avoir les valeurs de Z dans cet intervalle est 10% d'où $1 - \alpha = 0.90$. Alors cette probabilité peut se transformer en $P(Z \leq z) = 0.95$ telle que $z_1 = -z_2 = -z$ et la table de la loi $N(0,1)$ donne $z = 1.645$. Par conséquent, l'intervalle de confiance est la résolution des inégalités $z_1 \leq Z \leq z_2$ donc

$$-1.645 \leq \frac{\bar{X} - \mu}{s/\sqrt{n}} \leq 1.645,$$

d'où l'intervalle recherché est

$$IC_{10\%}(\mu) = \bar{X} \pm 1.645 \frac{s}{\sqrt{n}}.$$

Pour $\bar{X} = 11.6$ et $s_1^2 = 16.81$ cet intervalle devient

$$IC_{10\%}(\mu) = \bar{X} \pm 1.645 \frac{s}{\sqrt{n}} = [11.15, 12.05].$$

2. L'intervalle ainsi construit traduit que chaque vol de la compagnie étudiée contient un nombre allant de 11.15 à 12.05 sièges vides.

3. L'hypothèse de normalité de distribution est vérifiée car la taille de l'échantillon utilisé est assez grande.

4. Si un siège vide coûte $185DT$ pour $100km$ de vol pour la compagnie et si la destination moyenne des vols de cette compagnie est $1850kms$, la perte moyenne par vol est une valeur en DT de l'intervalle $[38160.875, 41241.125]$.