Statistiques appliquées (L3)
Groupes 14 et 15

Interrogation écrite n°1 *bis*

Rappel:
Si U suit une loi normale $N(0,1)$ alors $P(U \leq 1.96) = 0.975$ et $P(U \leq 1.645) = 0.95$

Vrai ou faux (2 points)

Répondez par vrai ou faux sans donner de justification. Une bonne réponse vaut $+0.5$ et une mauvaise réponse -0.25

1. La variance de la somme de deux variables aléatoires est toujours égale à la somme de leurs variances.

2. La somme de deux variables aléatoires normales indépendantes suit toujours une loi normale.

2. La convergence quadratique implique la convergence en probabilité.

3. La convergence presque sûre implique la convergence en probabilité.

Exercice 1 (12 points)

Première partie

Un sondage relatif à la consommation de cigarettes est effectué auprès d'un échantillon de $n = 200$ personnes tirées aléatoirement dans une très large population. Nous nous intéressons à estimer la proportion p de fumeurs dans la population.

1- Quelle loi suit la variable X_i indiquant pour chaque personne si elle est fume ou pas? Pourquoi peut-on considérer que les variables X_i sont indépendantes et identiquement distribuées?

2- Donner un estimateur convergent de la proportion p. Quelle est la loi de cet estimateur? Donner son espérance et sa variance.

3- Sachant que la proportion de fumeurs observée dans l'échantillon est de 0.35, construire un intervalle de confiance bilatéral à 95% pour la proportion p. Nous rappelons qu'il est possible d'utiliser l'approximation d'une loi binomiale $B(n,p)$ par une loi normale si la condition $np(1-p) > 15$ est satisfaite.

Deuxième partie

Nous nous proposons maintenant d'estimer la consommation moyenne journalière d'un fumeur. Nous tirons un échantillon de $n = 200$ personnes dans une très large population de fumeurs. Nous supposons que les variables Y_i correspondant à la consommation journalière des individus $i = 1, 2, .., n$ sont indépendantes et identiquement distribuées selon un loi normale d'espérance m et de variance connue $\sigma^2 = 16$

1- Quel est le meilleur estimateur de m? Pourquoi cet estimateur est-t-il convergent et sans biais? Quelle est sa loi?

2- Sachant que la consommation journalière moyenne observée dans l'échantillon est de 12 cigarettes, donner un intervalle de confiance bilatéral à 90% pour m.

3- Quelle est la différence fondamentale entre cette estimation par intervalle de confiance et celle de la première partie?

Exercice 2 (6 points)

Soit X une variable aléatoire suivant une loi de densité:

$$f(x) = \begin{cases} e^{\theta - x} & \text{si} \quad x \geq \theta \\ 0 & \text{sinon} \end{cases}$$

On suppose disposer d'un échantillon i.i.d de grande taille n tiré dans cette loi.

1- On admettra que $E(X) = \theta + 1$ et $V(X) = 1$.

1-a- Sans avoir recours à l'estimateur du maximum de vraisemblance, donner un estimateur $\hat{\theta}$ convergent de θ. Est-il sans biais? Déterminer sa variance.

1-b- Déterminer un intervalle de confiance à 95% du paramètre θ (en fonction de n et $\hat{\theta}$)

2- Calculer la fonction de vraisemblance $L(\theta, x_1, .., x_n)$ de l'échantillon et la log-vraisemblance de l'échantillon $l(\theta, x_1, .., x_n)$.

Question bonus: déterminer l'estimateur du maximum de vraisemblance de θ.