Préparation Agrégation
2012 - 2013

Feuille d'exercices 2

Exercice 1.

Soit $(X_1, \dots, X_n)$ un échantillon de loi inconnue $F(x)$. On souhaite estimer F. On pose pour tout $x \in \mathbb{R}$

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{X_i \leq x}.$$

Montrer que $\hat{F}_n$ est un estimateur sans biais de F.

Exercice 2.

1) Soit $(X_1, \dots, X_n)$ n variables indépendantes de même loi exponentielle de paramètre $\lambda > 0$ ($f(x) = \lambda \exp(-\lambda x),\ x > 0$).

a) Déterminer la loi de $S = X_1 + \cdots + X_n$.

b) On pose $Y = \frac{X_1}{S}$. Montrer que les variables aléatoires S et Y sont indépendantes.

2) On considère n machines de même type qui fonctionnent de façon indépendante. On désigne par X_i $(i = 1, \dots, n)$ l'instant auquel survient la première panne de la i^{eme} machine. Ces variables aléatoires ont une loi exponentielle de paramètre inconnu λ. A partir de l'observation de n machines , on se propose d'estimer la probabilité que sur une machine de ce type la première panne se produise avant l'instant x, c'est-à-dire la quantité

$$q(\lambda) = P_\lambda(X_1 \leq x) = 1 - \exp(-\lambda x).$$

a) Montrer que $U = I_{[0,x]}(X)$ est un estimateur sans biais de $q(\lambda)$.

b) Soit $\hat{U} = E(U|S)$.

i) Montrer que $\hat{U}$ est un estimateur sans biais de $q(\lambda)$.

ii) Montrer que $Var(U) = Var(\hat{U}) + Var(U - \hat{(U)})$.

iii) En déduire que $\hat{U}$ est l'unique estimateur sans biais de $q(\lambda)$ de variance minimale.

Exercice 3.

On considère une variable aléatoire discrète X dont la loi de probabilité est définie par

$$\mathbb{P}(X = k) = \frac{\theta^{k-1}}{(1+\theta)^k} \qquad \text{pour } k \geq 1.$$

où θ est un paramètre inconnu.

1) Déterminer l'espérance et la variance de cette variable.

2) Déterminer l'estimateur du maximum de vraisemblance de θ et ses propriétés.

Exercice 4.

1) On considère un échantillon issu d'une population uniforme sur $\{1,\ 2, \dots,\ N\}$. Déterminer l'estimateur du maximum de vraisemblance de N et ses propriétés.

2) On considère un échantillon issu d'une variable aléatoire de loi $\mathcal{U}[0, \theta]$, pour $\theta > 0$. Déterminer l'estimateur du maximum de vraisemblance de θ et sa loi.

3) On considère un échantillon issu d'une variable aléatoire de loi $\mathcal{U}[-\theta, \theta]$, pour $\theta > 0$. Déterminer l'estimateur du maximum de vraisemblance de θ et sa loi.

Exercice 5.

On considère un échantillon issu d'une variable aléatoire de densité

$$f(x) = \frac{1}{\theta} e^{-\frac{1}{\theta}(x-\gamma)} 1_{\{x > \gamma\}},$$

où $\theta > 0$ et $\gamma \in \mathbb{R}$.

Déterminer l'estimateur du maximum de vraisemblance des paramètres θ et γ.

Déterminer un estimateur des paramètres θ et γ par la méthode des moments.

Exercice 6.

Pour évaluer rapidement les résultats obtenus par ses 200 élèves, un professeur décide de corriger quelques copies tirées au hasard. Il admet que les notes suivent une loi normale de variance 4.

1) Le professeur corrige un échantillon de 7 copies et trouve une moyenne de 11. Quel est l'intervalle de confiance à 95% de la moyenne des 200 copies ?

2) Combien de copies le professeur doit-il corriger s'il veut situer la moyenne dans un intervalle de confiance d'amplitude 2 avec un risque 5%.

3) En trouvant une moyenne égale à 11, combien de copies faut-il corriger pour pouvoir dire que la moyenne de tous les élèves est supérieure à 10, avec un risque de 1%.

Exercice 7.

Une laiterie produit des camemberts commercialisés sous la marque *Le gourmet de Lamartine*. La masse X, exprimée en gramme, d'un camembert tiré au hasard dans la production, est distribuée selon une loi normale.

On tire un échantillon simple de 17 camemberts que l'on pèse. On obtient les masses suivantes :

(250, 254, 254, 253, 256, 250, 257, 251, 253, 255, 250, 255, 252, 261, 252, 251, 255)

1) Déterminer la moyenne et la variance de l'échantillon.

2) Donner un intervalle de confiance à 95% de la masse moyenne μ de la production.

3) Mr Michel M, le responsable de fabrication des camemberts *Le gourmet de Lamartine* souhaite savoir quelle taille minimale donner à un échantillon pour obtenir un intervalle de confiance pour μ d'amplitude inférieure à 1, au niveau 95%.

En bon statisticien, vous indiquez à Mr M, que l'on ne peut pas répondre sans connaître la variance. Pourquoi ?

Mr M vous préxcise alors que la variance est $\sigma^2 = 6,25$. Répondre alors à sa question.

4) Le responsable de fabrication estime que plus de 15% des camemberts ont une masse supérieure à $257g$. On tire un échantillon de 200 camemberts. On constate que 40 d'entre eux vérifient l'hypothèse. Au vu de cet échantillon, peut-on conclure, au seuil de 5%, que le responsable de fabrication à raison.

Exercice 8.

Soit $(X_1, \ldots, X_n)$ un échantillon extrait d'une loi $\mathcal{N}(\mu, \sigma)$ et $0 \leq \alpha \leq 1$.

1) On suppose $\sigma^2 = 2.5$. Donner un intervalle de confiance bilatérale de μ au seuil α.

2) On suppose σ^2 inconnu. Pour l'estimer on utilise l'estimateur sans biais déduit de $S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2$. Déterminer un intervalle de confiance au seuil α de σ^2.

3) Soit U et V deux variables aléatoires indépendantes de loi $\mathcal{N}(0,1)$ et χ^2_n. Montrer que la variable aléatoire $\frac{U}{\sqrt{\frac{V}{n}}}$ admet pour densité :

$$f(t) = \frac{n^{\frac{n}{2}}\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})(t^2+n)^{\frac{n+1}{2}}}.$$

La variance étant toujours inconnue, donner un intervalle de confiance de μ au seuil α.

4) Application numérique ; on observe les valeurs :

8.52, 9.24, 10.51, 12.09, 9.49, 11.14, 10.95, 10.93, 4.97, 8.02

Calculer les intervalles de confiance des questions 1), 2) et 3) au seuil 0.05.

Exercice 9.

Désireux d'acheter une guitare électrique, vous avez collecté des prix dans une boutique :

390, 460, 650, 410, 270 et 780 euros. Vous cherchez θ_0, l'espérance du prix d'une guitare. Un de vos camarades, plus malin que vous, a collecté 300 annonces sur internet. Il obtient un prix moyen de 550 euros et un écart-type empirique de 300 euros.

1) A partir de deux échantillons supposés constitués d'observations indépendantes et identiquement distribuées proposez deux estimations de l'espérance du prix d'une guitare. Rappelez les propriétés de ces estimateurs.

2) Ayant suivi la préparation à l'agrégation, vous trouvez un moyen d'améliorer encore la précision de l'estimation de θ_0 en combinant votre estimation et celle de votre camarade. Il vous propose de calculer la moyenne arithmétique des deux estimations en les pondérant donc chacune par $0,5$: $\widehat{\theta}_* = 0,5\widehat{\theta}_1 + 0,5\widehat{\theta}_2$, où $\widehat{\theta}_1$ et $\widehat{\theta}_2$ représentent respectivement votre estimation de θ_0 et celle de votre camarade. Calculer la variance de cette nouvelle estimation de l'espérance du prix.

3) Votre professeur de statistique, encore plus malin que vous, vous dit qu'en choisissant mieux la pondération de ces deux estimations, on peut encore améliorer la précision du résultat. Posons :

$$\widehat{\theta_{**}} = a\widehat{\theta}_1 + (1-a)\widehat{\theta}_2.$$

Donnez, en fonction de a, la variance de ce nouvel estimateur de l'espérance du prix. Quelle est la valeur de a qui minimise cette variance ?

4) Construisez un intervalle de confiance asymptotique à 95% pour l'espérance du prix à partir des deux échantillons initiaux (avec $n = 6$ pour l'un et $n = 300$ pour l'autre), ainsi qu'à partir de l'échantillon complet regroupant les deux. Qu'observez-vous ?