Examen : Octobre 2017.

durée : 3h

Seul document autorisé : une feuille manuscrite recto-verso.
En particulier : calculatrice, téléphone, ordinateur portable, polycopié de cours, feuilles de TD et corrections sont interdits.

Les Exercices 1 et 2 sont indépendants.

Les réponses aux questions doivent toutes être justifiées. En particulier les résultats du cours utilisés pour construire la réponse doivent être cités.

Exercice 1 (Loi géométrique)**:**
La loi géométrique est couramment utilisée pour modéliser des temps d'attente dans un cadre discret. C'est la loi du premier succès dans un suite de tirages aléatoires selon une loi de Bernoulli de paramètre θ. Autrement dit, Si $(U_i)_{i\geq 1} \overset{i.i.d.}{\sim} \mathcal{B}er(\theta)$, $\theta \in]0,1[$, alors $X = \min\{i \geq 1 : U_i = 1\}$ suit une loi géométrique de paramètre θ. On pourra utiliser les résultats suivants sur les distributions usuelles :

1. **Loi géométrique** : Une variable aléatoire X suit un loi géométrique de paramètre $\theta > \in]0,1[$, notée $\mathcal{G}(\theta)$, si elle est à valeurs dans $\mathbb{N}^*$, avec

$$\forall x \in \mathbb{N} \setminus \{0\}, \quad \mathbb{P}_\theta(\{X = x\}) = \theta(1-\theta)^{x-1}$$

Son espérance et sa variance sont respectivement

$$\mathbb{E}_\theta(X) = \frac{1}{\theta} \quad ; \quad \mathbb{V}\mathrm{ar}_\theta(X) = \frac{1-\theta}{\theta^2}.$$

2. Une **loi Beta** de paramètres $a > 0$, $b > 0$, notée $\mathcal{B}eta(a,b)$ est une loi continue sur le segment $]0,1[$, de densité donnée par

$$f_{a,b}(x) = \mathbb{1}_{]0,1[}(x)\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}.$$

Si $Z \sim \mathcal{B}eta(a,b)$, alors

$$\mathbb{E}(Z) = \frac{a}{a+b}.$$

1. Par quelle mesure le modèle $\{\mathcal{G}(\theta), \theta \in]0,1[\}$ est-il dominé ? On note $p_\theta(\,\cdot\,)$ la densité de $\mathcal{G}(\theta)$ par rapport à cette mesure.
   — Donner l'expression de p_θ.
   — On considère maintenant un échantillon i.i.d. de taille n, $X = (X_1, \dots, X_n)$ où $X_i \sim \mathcal{G}(\theta)$. Donner l'expression de la log-vraisemblance $\log p_\theta^{\otimes n}(X)$.
2. **Maximum de vraisemblance :** Donner l'expression de l'estimateur du maximum de vraisemblance pour le paramètre θ, pour un échantillon de taille n. On note $\widehat{\theta}_{MV} = \widehat{\theta}_{MV}(X)$ cet estimateur.
3. **Risque quadratique :** On cherche maintenant à estimer le paramètre d'intérêt $g(\theta) = \mathbb{E}_\theta(X_1) = \frac{1}{\theta}$. On considère l'estimateur

$$\widehat{g}_n(X) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Quel est son risque quadratique ? Est-ce un estimateur efficace de $g(\theta)$? On admettra que le modèle est régulier.

Approche bayésienne : En l'absence d'information a priori sur θ, on se donne un prior π uniforme sur l'intervalle $]0,1[$. Noter que π est alors une loi $\mathcal{B}eta(1,1)$. Dans la suite on note $\boldsymbol{\theta}$ la variable aléatoire associée à π.

4. Montrer que pour $x = (x_1, \dots, x_n)$, $x_i > 0$, la loi a posteriori $\pi(\,\cdot\,|x)$ est une loi $\mathcal{B}eta$ dont on précisera les paramètres.
5. Quelle est l'espérance a posteriori de $\boldsymbol{\theta}$, sachant les observations $X = x = (x_1, \dots, x_n)$? On note $\widehat{\theta}_{EP}$ cet estimateur.
6. Soit $\theta_0 \in]0,1[$ le 'vrai' paramètre, c'est-à-dire tel que $X_i \overset{\text{i.i.d.}}{\sim} \mathcal{G}(\theta_0), i = 1, \dots, n$. Monter que les estimateurs $\widehat{\theta}_{MV}$ et $\widehat{\theta}_{EP}$ convergent presque sûrement vers θ_0 lorsque $n \to \infty$.

Exercice 2 (Loi exponentielle translatée)**:**
La loi exponentielles est un modèle classique de temps d'attente dans un cadre continu. On s'intéresse à une variante où l'observation est la somme d'une quantité positive fixe θ (le temps d'attente minimal) et d'une variable de loi exponentielle de paramètre 1 : Autrement dit on considère un échantillon $X_1, \dots, X_n$ n de variables aléatoires indépendantes identiquement distribuées de densité :

$$p_\theta(x) = \exp(\theta - x)\mathbb{1}_{\{x \geq \theta\}} \ , \theta > 0 \ .$$

On cherche dans cet exercice à estimer le paramètre *de translation* θ.

1. **Méthode des moments :** Proposer un estimateur sans biais de θ par la méthode des moments, utilisant le moment d'ordre 1 de la loi des X_i : on le notera $\widehat{\theta}_n$.
2. Calculer le risque quadratique de l'estimateur $\widehat{\theta}_n$.
3. **Estimateur alternatif :** On considère à présent l'estimateur $\tilde{\theta}_n$ défini par $\tilde{\theta}_n(X) = \inf_{1 \leq i \leq n} X_i$. Calculer sa loi.
4. L'estimateur $\tilde{\theta}_n$ est-il sans biais ?
5. Calculer le risque quadratique de l'estimateur $\tilde{\theta}_n$.
6. Au vu des résultats précédents, quel estimateur de θ proposeriez vous ?
7. On ne demande pas si les estimateurs construits sont efficaces car les hypothèses du théorème de Cramér-Rao ne s'appliquent pas : le modèle n'est pas régulier. Pourquoi ?
8. **Intervalle de confiance :** Soit $\delta \in]0,1[$.
   — Déterminer $x = x(n, \theta, \delta)$ tel que $\mathbb{P}_\theta(\tilde{\theta}_n > x(n, \theta, \delta)) = \delta$.
   — En déduire un intervalle de confiance de niveau $1 - \delta$ pour θ (c'est-à-dire contenant θ avec probabilité $1 - \delta$), de type $I = [a_n, \tilde{\theta}]$, où a_n est à déterminer.
   — Application numérique : exprimer a_n en fonction de $\tilde{\theta}_n$, pour $n = 10$. On donne : $\ln(0.05) \approx -3$.