TD1 : Estimation Ponctuelle

Master 1 Data Science et Intelligence Artificielle

Ibrahima SY

10/12/2021

Exercice 1 :

Dans le cadre d'une étude sur la santé au travail, on a interrogé au hasard 500 salariés de différents secteurs et de différentes régions de France. 145 d'entre eux déclarent avoir déjà subi un harcèlement moral au travail.

1. Identifier la population, la variable, son type et son/ses paramètre(s).
2. Donner une estimation ponctuelle de la proportion de salariés ayant déjà subi un harcèlement moral au travail

Exercice 2 :

En vue de réaliser un programme de rééducation, des chercheurs ont soumis un questionnaire de neuropsychologie cognitive à 150 enfants dyslexiques tirés au sort. Le questionnaire comporte 20 questions et les chercheurs ont recueilli pour chaque enfant dyslexique le nombre x_i de bonnes réponses. Les résultats ainsi récoltés sont tels que :

$$\sum_1^n x_i = 1502 \; ; \quad \sum_1^n x_i^2 = 19486$$

1. Identifier la population, la variable, son type et son/ses paramètre(s).
2. Donner une estimation ponctuelle du nombre moyen de bonnes réponses dans la population étudiée.
3. Donner une estimation ponctuelle de l'écart-type de la variable

Exercice 3 :

On étudie la caractéristique X d'une population, qui suit une de bernouilli $\mathcal{B}(p)$. Afin d'estimer le paramètre de cette loi, on fait un sondage de taille n. Soit $(X_1, X_2, \ldots, X_n)$ l'échantillon aléatoire associé.

1. Monter que la statistque $\sum_{i=1}^n X_i$ est exhaustive pour p, puis interpreter
2. Déterminer l'estimateur du maximum de vraisemblance T_n de p
3. Déterminer le score $S(X, \theta)$
4. Calculer l'information de Fisher apportée par un n-échantillon issu de X sur le paramètre p
5. Montrer T_n est un estimateur sans de p, puis determiner la variance de cette estimateur.
6. Calculer le risque quadratique de T_n.
7. T_n est-il convergent ?
8. Calculer la borne FDCR relative à ce paramètre ?
9. Estimateur est-il efficace ?

Exercice 4 :

On étudie la caractéristique X d'une population, que l'on sait être de loi exponentielle $\mathcal{E}(\lambda)$. Afin d'estimer le paramètre λ de cette loi, on fait un sondage de taille n. Soit $(X_1, X_2, \ldots, X_n)$ l'échantillon aléatoire associé.

1. Monter que la statistque $\sum_{i=1}^{n} X_i$ est exhaustive pour λ, puis interpreter
2. Déterminer l'estimateur du maximum de vraisemblance T_n de λ
3. Déterminer le score $S(X, \theta)$
4. Calculer l'information de Fisher apportée par un n-échantillon issu de X sur le paramètre λ
5. Montrer T_n est un estimateur asymptotiquement sans bias λ, puis determiner la variance de cette estimateur.
6. Calculer le risque quadratique de T_n.
7. T_n est-il convergent ?
8. Calculer la borne FDCR relative à ce paramètre ?
9. Estimateur est-il efficace ?

Rappel :

- densité de la loi gamma :
$$f(x; \alpha, \beta) = x^{\alpha-1} \frac{\beta^{\alpha} \, \mathrm{e}^{-\beta x}}{\Gamma(\alpha)} \text{ pour } x > 0$$
- Proprietés de la Fonction gamma :
- $\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt, \alpha > 0$
- $\Gamma(1) = 1$
- $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$
- $\Gamma(n) = (n-1)!, n \in \mathbb{N}$

Exrecice 5 :

On étudie la caractéristique X d'une population, que l'on sait être de loi normale $\mathcal{N}(\mu, \sigma^2)$. Afin d'estimer les paramètres de cette loi, on fait un sondage de taille n. Soit $(X_1, X_2, \ldots, X_n)$ l'échantillon aléatoire associé.

1. On veut estimer le paramètre μ. Déterminer l'estimateur du maximum de vraisemblance T_n de μ et étudier ses propriétés. Le fait que σ soit ou non connu modifie-t-il le résultat?
2. On suppose μ connu et on veut estimer σ . Déterminer l'estimateur du maximum de vraisemblance T_n^2 de σ^2 et étudier ses propriétés. Calculer la borne FDCR relative à ce paramètre. Conclusion?
3. En déduire un estimateur $\hat{\sigma}_n$ de σ. Cet estimateur peut-il être sans biais?
4. Dans le cas où μ est inconnu, donner un estimateur S_n^2 de σ^2 et étudier ses propriétés.