Fiche TD complète

1 Estimation ponctuelle

Exercice 1.

Soient $X_1, \ldots, X_n$ i.i.d. de loi de Poisson de moyenne λ. Déterminez le MLE pour le paramètre λ.

Exercice 2.

Soient $X_1, \ldots, X_n$ i.i.d. $\text{Bin}(1, p)$. Déterminez le MLE pour le paramètre p.

Exercice 3.

Soit $(X_1, \ldots, X_n)$ un échantillon aléatoire simple issu d'une population de densité

$$f_\theta(x) = \begin{cases} \dfrac{1}{\theta} e^{-\frac{1}{\theta}(x-\gamma)} \text{ si } x > \gamma \\ 0 \text{ sinon} \end{cases}$$

où $\theta > 0$. Déterminez une estimation des paramètres θ et γ par la méthode du MLE.

Exercice 4.

Soit $X_1, \ldots, X_n$ un échantillon aléatoire simple issu d'une population de densité

$$f_\theta(x) = \begin{cases} \frac{\theta}{1-\theta} x^{\frac{2\theta-1}{1-\theta}} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

où $1/2 < \theta < 1$. Déterminez le MLE pour θ.

Exercice 5.

Les éléments d'une population possèdent un caractère X qui suit une loi de densité

$$f_\theta(x) = \frac{\theta + 1}{2}(1 - |x|)^\theta, x \in (-1, 1)$$

où on suppose le paramètre $\theta > -1$. On en extrait un échantillon simple $X_1, \ldots, X_n$. Déterminez l'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ.

Exercice 6.

Les éléments d'une population possèdent un caractère X qui suit une loi de densité

$$f_\theta(x) = \frac{\sqrt{\theta}}{\sqrt{2\pi}} e^{-\theta x^2/2}$$

où $\theta > 0$. Pour étudier le paramètre θ, on a effectué une suite de n expériences indépendantes qui ont donné les réalisations $x_1, \ldots, x_n$ de n v.a. $X_1, \ldots, X_n$ i.i.d. de même loi que X.

1. Déterminez un estimateur $\hat{\theta}$ du paramètre θ par la méthode du maximum de vraisemblance.
2. $\hat{\theta}$ est-il exhaustif ?
3. Calculez la moyenne et la variance de $\hat{\theta}$. Déduisez-en un estimateur $\hat{\theta}_1$ de θ non biaisé. Quelle est la variance de $\hat{\theta}_1$? Est-il convergent ?

Exercice 7.

Les éléments d'une population possèdent un caractère X qui suit une loi de densité

$$f_\theta(x) = \frac{2}{\sqrt{\pi}\theta^{3/2}} x^2 e^{-x^2/\theta}$$

où $\theta > 0$. Une suite de n expériences indépendantes a donné les valeurs $x_1, \ldots, x_n$.

1. Déterminez un estimateur $\hat{\theta}$ du paramètre θ par la méthode du maximum de vraisemblance.
2. Examinez les qualités suivantes de $\hat{\theta}$: efficacité, biais, convergence, exhaustivité.

Exercice 8.

Les éléments d'une population possèdent un caractère X qui suit une loi de probabilité dont la densité est

$$f_{a,\theta}(x) = \begin{cases} \theta e^{-\theta(x-a)} \text{ si } x \geq a \\ 0 \text{ sinon} \end{cases}$$

où $\theta, a > 0$. Une suite de n expériences indépendantes a donné les valeurs $x_1, \ldots, x_n$.

Inférence sur a, en supposons que θ est connu. Proposez un estimateur $\hat{a}$ de a par la méthode du maximum de vraisemblance.

2 Tests statistique

Exercice 9.

On veut savoir si la résistance moyenne de composants produits dans une usine est 400Ω. On considère que la distribution des résistances est normale, et on mesure pour 16 composants les valeurs $392, 396, 386, 389, 388, 387, 403, 397, 401, 391, 400, 402, 394, 406, 406, 400$.

(a) Donner les estimations ponctuelles des moyenne et variance.

(b) Peut-on considérer, au seuil de signification $\alpha = 5\%$, que le lot respecte la norme de 400 ? Même question avec un seuil de $\alpha = 1\%$.

Exercice 10.

Un fabricant se vante de proposer des tubes à essai d'une durée de vie supérieure à $2000h$ de chauffage. A l'aide d'un échantillon de 100 tubes testés, on estime la durée de vie moyenne à 1975h, avec un écart-type de $130h$. Peut-on affirmer, au risque 5%, que le fabriquant ment ?

Exercice 11.

Un fabricant annonce que la masse d'un composant de l'un de ses produits est de $75mg$. Les mesures pour le vérifier étant coûteuses, trois seulement sont réalisées, dont les résultats sont $70, 72$ et $74mg$. Peut-on, au risque de 5% de se tromper, dénoncer la publicité du fabriquant ?

Exercice 12.

Un laboratoire pharmaceutique désire étudier les effets secondaires potentiels d'un médicament sur le taux de cholestérol des patients. Cent volontaires sains sont donc choisis pour tester le médicament.

(a) Avant l'expérience, le taux de cholestérol moyen de ces volontaires est de $2.02 \pm 0.2g/l$. Le taux de cholestérol moyen dans la population étant de 2 g/l, vérifier que cet échantillon est représentatif au risque 5

(b) Après un mois de traitement, seuls 97 volontaires reviennent faire un test. Leur taux moyen de cholestérol est passé à $2.09 g/l$ avec un écart-type d'échantillon de $0.25 g/l$.

La différence est-elle significative au risque 5% ? Au risque 1% ?

Exercice 13.

Pour étudier un nouvel alliage métallique, on a soumis un échantillon aléatoire de 16 tiges aux essais pour obtenir les résistances suivantes en kg/cm^2 :
$1895, 1920, 1886, 1890, 1864, 1880, 1875, 1915, 1850, 1927, 1910, 1912, 1886, 1903, 1854, 1880.$

On suppose la résistance distribuée normalement.

(a) Estimer par intervalle avec un niveau de confiance de 95%, la résistance moyenne à la rupture.

(b) Avant l'introduction de ce nouvel alliage la résistance moyenne à la rupture des tiges était de $1840 kg/cm^2$. Que peut-on conclure des essais effectués avec le nouvel alliage ?

Exercice 14.

Les habitants d'une région aéroportuaire se plaignent que le bruit des avions dépasse la limite autorisée de 80 décibels en moyenne imposée par la législation. On admet que l'intensité du bruit causé par les avions est une variable aléatoire X de loi gaussienne d'espérance μ et de variance 64 .

On mesure un échantillon journalier de $\mathrm{n} = 16$ variables aléatoires indépendantes $\mathrm{X}_1, \ldots, \mathrm{X}_\mathrm{n}$ de l'intensité du bruit, et on effectue le test statistique suivant.

$$\begin{cases} \mathrm{H}_0 : \mu = \mu_0 = 80 \text{ décibels} \\ \mathrm{H}_1 : \mu = \mu_1 = 85 \text{ décibels} \end{cases}$$

1. Expliciter les risques de première et deuxième espèces. De quel point de vue est fait ce test ? Celui des habitant ou celui des responsables de l'aéroport ?
2. Quelle variable de décision faut-il choisir et quelle est sa loi ?
3. Calculer le seuil de la région critique pour un risque $\alpha = 5\%$.
4. Calculer la puissance du test.
5. Enoncer les règles de décision avec les probabilités d'erreur.
6. La moyenne calculée sur l'échantillon est $\bar{x} = 83$ décibels. Les habitants ont-ils raison de se plaindre ? Le test d'hypothèses ainsi établi leur est-il favorable ou défavorable ?
7. Combien faudrait-il faire de relevés journaliers, pour que le risque de deuxième espèce soit de 5% ?
8. Quelle serait alors le seuil de décision ?

Exercice 15.

Sur un échantillon de 900 naissances, on constate qu'il y a 470 garçons. Un généticien décide d'utiliser ces données pour effectuer le test suivant relatif aux proportions p et 1 -p de naissances respectivement masculines et féminines :

$$\mathrm{H}_0 : \mathrm{p} = 0.5$$
$$\mathrm{H}_1 : \mathrm{p} = 0.55$$

1) Construire un test pour ces hypothèses avec un risque $\alpha = 5\%$. Peut-on être satisfait du test ? Si non comment peut-on l'améliorer ?

2) Ce généticien effectue une nouvelle étude sur un échantillon de même taille. II souhaite cette fois tester les hypothèses :

$$\mathrm{H}_0 : \mathrm{p} = 0.5$$
$$\mathrm{H}_1 : \mathrm{p} \neq 0.5$$

3 Intervalle de Confiance

Exercice 16. Des essais en laboratoire sur 20 lampes miniatures donnent les durées de vie suivantes, en heures :

$$451, 412, 412, 375, 407, 454, 375, 393, 355, 364, 414, 413, 345, 432, 392, 329, 439, 381, 451, 413.$$

On suppose la durée de vie distribuée normalement. Estimer par un intervalle de confiance 95% la durée de vie moyenne.

Exercice 17.

Une machine fabrique des billes métalliques dont le poids, mesuré en grammes, suit une loi normale. Nous prélevons au hasard 10 billes. Leurs poids sont

$$19,6; 20; 20,2; 20,1; 20; 19,9; 20; 20,3; 20,1; 19,8.$$

1. Quel est l'intervalle de confiance à 95% du poids des billes métalliques fabriquées ?
2. En réalité, l'écart-type σ de la population est connu et égal à 0,2. Quel est l'intervalle de confiance à 95% du poids des billes métalliques fabriquées ?

Exercice 18.

Voulant évaluer rapidement les résultats obtenus par ses 200 étudiants ingénieurs lors d'un partiel, un professeur décide de corriger quelques copies tirées au hasard. Il admet par ailleurs que les notes de ses étudiants 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 générale de ses étudiants dans un intervalle de confiance d'amplitude 2, avec un risque de 5%?
3. En trouvant une moyenne égale à 11, combien de copies le professeur devraitil corriger pour pouvoir dire, avec un risque de 1%, que la moyenne de tous les étudiants est supérieure à 10?

Exercice 19.

Une entreprise fabrique un certain type de composants électroniques dont la durée de vie X, exprimée en heures, est une variable aléatoire. Des mesures effectuées sur un échantillon aléatoire de taille 50 ont donné les résultats suivants :

$$\sum_{i=1}^{50} x_i = 60000; \quad \sum_{i=1}^{50} x_i^2 = 74 \times 10^6$$

1. Donner une estimation ponctuelle de la durée de vie moyenne des composants.
2. Donner une estimation ponctuelle de l'écart-type de cette durée de vie.
3. Donner l'intervalle de confiance à 95%, puis à 99% de cette durée de vie moyenne.

4. Quelle aurait du être la taille de l'échantillon pour que l'intervalle de confiance à 95% de la durée de vie moyenne des composants ait une amplitude de 60 heures ?

Exercice 20.

À la veille d'une consultation électorale, nous effectuons un sondage.

1. Dans un échantillon représentatif de 1000 personnes, 500 personnes déclarent vouloir voter pour X, 250 pour Y et 50 pour Z. Donner les intervalles de confiance à 95% et 99% du pourcentage de personnes ayant l'intention de voter X, Y ou Z.
2. Nous évaluons le pourcentage de personnes ayant l'intention de voter pour un quatrième candidat, H, à 17% ? Combien faut-il interroger de personnes pour obtenir un intervalle de confiance à 95% du pourcentage de personnes ayant l'intention de voter H, avec une précision de 1%?

Exercice 21.

On veut étudier la proportion p de gens qui vont au cinéma chaque mois. On prend donc un échantillon de taille $n = 100$. Soit N le nombre de personnes dans l'échantillon qui vont au cinéma mensuellement.

1. Quelle est la loi de N ? Par quelle loi peut-on l'approcher et pourquoi ? En déduire une approximation de la loi de $F = N/n$.
2. On observe une proportion f de gens qui vont chaque mois au cinéma. Donner la forme d'un intervalle de confiance pour p, de niveau de confiance $1 - \alpha$.
3. Applications numériques : $f = 0,1, 1 - \alpha = 90\%, 95\%, 98\%$.

Exercice 22.

On suppose que le poids d'un nouveau né est une variable normale d'écart-type égal à $0,5$ kg. Le poids moyen des 49 enfants nés au mois de Décembre 2019 dans l'hôpital de la ville a été de $3,6$ kg.

1. Déterminer un intervalle de confiance à 95% pour le poids moyen d'un nouveau né dans cet hôpital.
2. Quel serait le niveau de confiance d'un intervalle de longueur 0,1 kg centré en 3,6 pour ce poids moyen ?