Sandie Souchet

Travaux Dirigés de Statistique

Licence 2 MIASHS

Université Paris 1

2022-2023

Feuille TD 1

Rappels de probabilité

Exercice 1

Soit X la var continue de densité f donnée par

$$f_X(x) = \begin{cases} kx \exp\left(-\frac{1}{2}x^2\right) & x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

1. Déterminer k pour que f_X soit bien une densité de probabilité
2. Déterminer la fonction de répartition de X
3. Calculer $\mathbb{E}\left(X^2\right)$ et $\mathbb{V}\left(X^2\right)$
4. Soit $Y = X^2$. On admet que Y est également une var continue.
   (a) Exprimer F_Y la fdr de Y en fonction de F_X celle de X. Expliciter F_Y. En déduire la loi de Y
   (b) Calculer $P\left(Y \in]-\infty, -2] \cup [3, +\infty[\right)$

Exercice 2

Soit X une var telle que $X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right)$. On admet que $Y = \frac{X-m}{\sigma}$ et $Z = Y^2$ sont également des var continues.

1. Exprimer F_Y la fonction de répartition de Y en fonction de F_X celle de X. F_Y peut-elle être explicitée ?
2. En déduire f_Y la densité de probabilité de Y en fonction de f_X celle de X, puis expliciter f_Y et déduire la loi de Y
3. Exprimer F_Z la fonction de répartition de Z en fonction de F_Y celle de Y. F_Z peut-elle être explicitée ?
4. En déduire f_Z la densité de probabilité de Z en fonction de f_Y celle de Y, puis expliciter f_Z et déduire la loi de Z
5. On suppose dans cette question que $m = 3$ et $\sigma^2 = 4$, calculer $P\left(X \in]2, 6]\right)$

Exercice 3

Soit X une var telle que $X \hookrightarrow \mathcal{E}(2)$. On admet que $Y = \exp(-2\,X)$ est également une var continue.

Déterminer la loi de Y

Exercice 4

Soit $X_1, \cdots, X_n$ n var indépendantes et de même loi μ de fonction de répartition F

On définit $M_n = \max\{X_1, \cdots, X_n\}$ et $I_n = \min\{X_1, \cdots, X_n\}$

1. Exprimer F_{M_n} la fonction de répartition de M_n en fonction de F (on demande une démonstration du résultat de cours)
2. Exprimer F_{I_n} la fonction de répartition de I_n en fonction de F (on demande une démonstration du résultat de cours)
3. On suppose dans cette question que $\mu = \mathcal{E}(\lambda)$. Déterminer F_{I_n} et déduire la loi de I_n dans ce cas.
4. On suppose dans cette question que $\mu = \mathcal{U}(]0, 1[)$. Déterminer F_{M_n} et déduire la loi de M_n dans ce cas.
5. On suppose dans cette question que $\mu = \mathcal{P}(r, c)$. Déterminer F_{I_n} et déduire la loi de I_n dans ce cas.

FEUILLE TD 2

QUALITÉ DES ESTIMATEURS

Exercice 5

Soient $X_1, \cdots, X_n$ n variables aléatoires et $a_1, \cdots, a_n,\ b,\ n+1$ réels.

On suppose que, pour tout $(k,j) \in \{1, \cdots, n\} \times \{1, \cdots, n\}$, $\mathbb{C}\text{ov}\big(X_k,\ X_j\big)$ existe.

1. Montrer que

$$\mathbb{V}\left(\sum_{k=1}^{n} a_k\ X_k\ +\ b\right) = \sum_{k=1}^{n}\ a_k^2\ \mathbb{V}\left(X_k\right)\ + 2\sum_{k=1}^{n-1}\ \sum_{j=k+1}^{n}\ a_k\ a_j\ \mathbb{C}\text{ov}\big(X_k,\ X_j\big)$$

2. On suppose dans la suite de l'exercice que les variables sont indépendantes et de même loi d'espérance m et de variance σ^2

   (a) Calculer l'espérance et la variance de $\overline{X}_n = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$

   (b) En utilisant l'inégalité de Markov à l'ordre 2, montrer que $\overline{X}_n$ est un estimateur consistant pour m

   (c) Montrer que $S_n^2 = \frac{1}{n}\sum\limits_{k=1}^{n}\left(X_k - m\right)^2$ et $\widehat{S}_n^2 = \frac{1}{n-1}\sum\limits_{k=1}^{n}\left(X_k - \overline{X}_n\right)^2$ sont des estimateurs sans biais pour σ^2

Exercice 6

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi exponentielle de paramètre $\frac{1}{\theta}$

On définit

$$\widehat{\theta}_n^1 = X_n \quad \widehat{\theta}_n^2 = \frac{1}{2}\left(X_1 + X_n\right) \quad \widehat{\theta}_n^3 = \frac{1}{3}\left(X_1 + 2X_n\right) \quad \widehat{\theta}_n^4 = \frac{1}{n}\sum_{k=1}^{n} X_k \quad \widehat{\theta}_n^5 = \min\left\{X_1, \cdots, X_n\right\}$$

1. Déterminer la loi de $\widehat{\theta}_n^5$

2. Parmi ces cinq estimateurs, débiaiser ceux qui ne sont pas sans biais pour θ

3. Calculer alors le risque quadratique associé à chacun des estimateurs débiaisés. Lequel est le meilleur ?

Exercice 7

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi μ^θ avec θ inconnu et $t = (x_1, \cdots, x_n)$ une observation de ce vecteur.

1. Dans tous les cas suivants, déterminer si cela est possible une Statistique Exhaustive associée au modèle et l'Information de Fisher.
   (a) $\mu^\theta = \mathcal{N}(m, \theta)$
   (b) $\mu^\theta = \mathcal{G}\left(\frac{1}{\theta}\right)$
   (c) $\mu^\theta = \Gamma\left(\alpha, \frac{1}{\theta}\right)$
2. On suppose qu'il existe, pour chacun des modèles précédents, $\widehat{\theta}_n$ un estimateur efficace. Donner alors son espérance, sa variance et la variable dont il dépend.

Exercice 8

Soient T_1 et T_2 deux estimateurs d'un paramètre θ inconnu.

Ces estimateurs sont sans biais, de variance respective (connue) σ_1^2 et σ_2^2

De plus $\mathbb{C}\text{ov}\left(T_1, T_2\right) = \rho$

1. Déterminer deux réels a et b tels que $T = a\ T_1 + b\ T_2$ soit sans biais pour θ et de variance minimum (parmi tous les estimateurs de cette forme).
2. On suppose que les hypothèses de l'Inégalité de Cramer-Rao sont satisfaites pour le modèle associé à l'estimation de θ et on note $I(\theta)$ l'Information de Fisher associée au modèle.

   On suppose de plus que T_1 et T_2 sont tous deux efficaces
   (a) Que valent σ_1^2 et σ_2^2 ?
   (b) Que vaut alors T ? Calculer sa variance.
   (c) Comparer $\mathbb{V}(T_1)$ et $\mathbb{V}(T)$. En déduire que $\rho = \frac{1}{I(\theta)}$
   (d) Calculer $\mathbb{V}(T_1 - T_2)$. Que pouvez-vous en déduire ?

Feuille TD 3

Estimateurs Naturels

Exercice 9

Déterminer le meilleur estimateur sans biais pour σ^2 de la forme

$$S = a\,\widehat{S}^2_{n(X)} + b\,\widehat{S}^2_{d(Y)}$$

où $\widehat{S}^2_{n(X)}$ et $\widehat{S}^2_{d(Y)}$ sont les variances empiriques modifiées associées aux deux échantillons $T = (X_1, \cdots, X_n)$ de loi $\mathcal{N}\left(m_X, \sigma^2\right)$ et $V = (Y_1, \cdots, Y_d)$ de loi $\mathcal{N}\left(m_Y, \sigma^2\right)$ lorsque les deux moyennes sont inconnues et que les deux échantillons sont indépendants l'un de l'autre.

Exercice 10

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{E}\left(\frac{1}{\lambda}\right)$

On veut estimer $\theta = \lambda^2$ et, pour cela, on introduit les deux estimateurs suivants

$$\widehat{\theta}_n = \left(\overline{X}_n\right)^2 \qquad \widehat{S}^2_n = \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k - \overline{X}_n\right)^2$$

1. Expliquer le choix de ces deux estimateurs.
2. Ces estimateurs sont-il sans biais ? Asymptotiquement sans biais ?
3. Ces estimateurs sont-ils consistants ?
4. Déterminer la loi de $Y_n = n\,\overline{X}_n$ puis calculer $\mathbb{E}\left[Y_n^4\right]$
5. Calculer le risque quadratique de $\widehat{\theta}_n$

Exercice 11

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{P}(\lambda)$

On veut estimer $\theta = \lambda^2$

1. Donner $\mathbb{E}(X_1)$ et calculer $\mathbb{E}[X_1(X_1 - 1)]$
2. Proposer deux estimateurs naturels pour θ
3. Ces deux estimateurs sont-ils sans biais ? Asymptotiquement sans biais ?
4. Ces deux estimateurs sont-ils consistants ?

Exercice 12

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{U}\left(]0, \theta[\right)$

On veut estimer θ et, pour cela, on introduit les deux estimateurs suivants

$$T_n^1 = \overline{X}_n \qquad T_n^2 = \max\left\{X_1, \cdots, X_n\right\}$$

1. Expliquer le choix de ces deux estimateurs.
2. Déterminer la loi de T_n^2 puis calculer $\mathbb{E}\left(T_n^2\right)$ et $\mathbb{V}\left(T_n^2\right)$
3. Débiaiser si nécessaire chacun des estimateurs
4. Les deux estimateurs sans biais obtenus sont-ils consistants ?
5. Lequel de ces deux estimateurs sans biais est le meilleur ?

Exercice 13

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{P}(\theta, 3)$

On veut estimer θ

1. Que représente θ pour la loi de l'échantillon ?
2. En déduire un estimateur sans biais pour θ
3. Que vaut $\mathbb{E}(X_1)$?
4. En déduire un estimateur sans biais pour θ
5. Ces deux estimateurs sont-ils consistants ?
6. Lequel est meilleur au sens du risque quadratique ?

FEUILLE TD 4
ESTIMATEUR DU MAXIMUM DE VRAISEMBLANCE

Exercice 14

Un agent d'assurance visite m clients par jour afin de leur proposer un contrat d'assurance vie.

Pour chaque client, il a une probabilité θ de "succès".

1. Quelle est la loi de X la var représentant le nombre de succès quotidiens de l'agent ?

2. θ est un paramètre inconnu que l'on se propose d'estimer à l'aide d'une observation $t = (x_1, \cdots, x_n)$ du n-échantillon $T = (X_1, \cdots, X_n)$ où, $\forall i = 1, \cdots, n$, X_i est le nombre de succès enregistrés le $i^{ème}$ jour.

   (a) Déterminer l'estimateur du maximum de vraisemblance de θ

   (b) Cet estimateur est-il sans biais ? Convergent ? Efficace ?

   (c) On a observé qu'ayant visité 30 clients par jour pendant 15 jours, l'agent a obtenu 225 souscriptions. En utilisant l'estimateur précédent, donner une estimation de θ

Exercice 15

Soit X une var continue dont la densité f est donnée par

$$f(x) = \begin{cases} \frac{2}{\theta^2}\, x & \text{si } x \in [0, \theta] \\ 0 & \text{sinon} \end{cases} \qquad \text{avec} \qquad \theta > 0$$

On admet que $\mathbb{E}(X) = \frac{2}{3}\theta$ et que $\mathbb{V}(X) = \frac{1}{18}\theta^2$

On suppose θ inconnu et l'on se propose de l'estimer à l'aide d'une observation $t = (x_1, \cdots, x_n)$ du n-échantillon $T = (X_1, \cdots, X_n)$ de loi P_X

On définit pour cela les deux estimateurs suivants

$$M_n = \max\{X_1, \cdots, X_n\} \qquad T_n = \frac{3}{2\,n}\sum_{i=1}^{n} X_i$$

1. Montrer que M_n est l'estimateur du maximum de vraisemblance de θ

2. Déterminer la loi de M_n

3. T_n est-il sans biais ? Convergent ?

4. M_n est-il sans biais ? Convergent ?

5. M_n est-il meilleur que T_n en terme de risque quadratique ?

Exercice 16

Soit X une variable aléatoire continue dont la densité de probabilité est donnée par

$$f(x) = \begin{cases} \frac{1}{\theta}\, x^{-\left(\frac{1}{\theta}+1\right)} & \text{si } x \in [1, +\infty[\\ 0 & \text{sinon} \end{cases}$$

Le paramètre θ est inconnu et $\theta \in]0, +\infty[$

1. Montrer que F la fonction de répartition de la variable X est

$$F(x) = \begin{cases} 1 - x^{-\frac{1}{\theta}} & \text{si } x \in [1, +\infty[\\ 0 & \text{sinon} \end{cases}$$

2. On admet que $Y = \ln(X)$ est une var continue. Montrer que Y suit la loi $\mathcal{E}\left(\frac{1}{\theta}\right)$
3. En déduire que $\mathbb{E}[\ln(X)] = \theta$ et $\mathbb{V}[\ln(X)] = \theta^2$
4. On se propose d'estimer le paramètre θ à partir d'une observation $t = (x_1, \cdots, x_n)$ du n-échantillon $T = (X_1, \cdots, X_n)$ de loi P_X
   (a) Trouver $\widehat{\theta}_n$ l'estimateur du maximum de vraisemblance de θ
   (b) $\widehat{\theta}_n$ est-il sans biais ?
   (c) $\widehat{\theta}_n$ est-il convergent ?
   (d) $\widehat{\theta}_n$ est-il efficace ?

FEUILLE TD 5
INTERVALLES DE CONFIANCE

Exercice 17

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{N}\left(m, \sigma^2\right)$

1. On suppose σ^2 connue.
   (a) Déterminer pour m , $\widehat{\mathrm{I}}_n^{1-\alpha}(m)$, un IC exact, bilatéral et symétrique, au niveau $1-\alpha$
   (b) Déterminer n la taille minimale de l'échantillon pour que l'intervalle précédent soit de largeur inférieure ou égale à 10% de σ
2. On suppose σ^2 inconnue. Déterminer pour m , $\widehat{\mathrm{I}}_n^{1-\alpha}(m)$, un IC exact, bilatéral et symétrique, au niveau $1-\alpha$
3. On suppose m connue. Déterminer pour σ^2 , $\widehat{\mathrm{I}}_n^{1-\alpha}\left(\sigma^2\right)$, un IC exact, unilatéral, au niveau $1-\alpha$
4. On suppose m inconnue. Déterminer pour σ^2 , $\widehat{\mathrm{I}}_n^{1-\alpha}\left(\sigma^2\right)$, un IC exact, unilatéral, au niveau $1-\alpha$

Exercice 18

Soient $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{N}\left(m_x, \sigma_x^2\right)$ et $V = (Y_1, \cdots, Y_d)$ un d-échantillon de loi $\mathcal{N}\left(m_y, \sigma_y^2\right)$

On suppose que les deux échantillons sont indépendants entre eux et que tous les paramètres des deux lois sont inconnus.

Déterminer pour $\frac{\sigma_x^2}{\sigma_y^2}$, $\widehat{\mathrm{I}}_{n+d}^{1-\alpha}\left(\frac{\sigma_x^2}{\sigma_y^2}\right)$, un IC exact, bilatéral et symétrique, au niveau $1-\alpha$

Exercice 19

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi μ^θ dont la fdr est donnée par

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \left(\frac{x}{\theta}\right)^2 & \text{si } x \in [0, \theta] \\ 1 & \text{si } x > \theta \end{cases} \qquad \text{avec} \qquad \theta > 0$$

1. Déterminer la loi de $\frac{1}{\theta}\max\{X_1, \cdots, X_n\}$
2. Déterminer pour θ , $\widehat{\mathrm{I}}_n^{1-\alpha}(\theta)$, un IC exact, bilatéral et symétrique, au niveau $1-\alpha$

Exercice 20

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{G}\left(\frac{1}{\theta}\right)$ avec $\theta > 1$

Déterminer pour θ , $\widehat{\mathrm{I}}_n^{1-\alpha}(\theta)$, un IC asymptotique, unilatéral, au niveau $1-\alpha$

Exercice 21

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{P}(\theta, 3)$

1. Déterminer la loi de $\frac{1}{\theta}\min\{X_1, \cdots, X_n\}$
2. Déterminer pour θ , $\widehat{\mathrm{I}}_n^{1-\alpha}(\theta)$ un IC exact, bilatéral et symétrique, au niveau $1-\alpha$
3. Déterminer pour θ , $\widetilde{\mathrm{I}}_n^{1-\alpha}(\theta)$ un IC asymptotique, bilatéral et symétrique, au niveau $1-\alpha$
4. On observe T et on obtient, pour $n = 1000$, $t = (x_1, \cdots, x_n)$ tel que

$$\sum_{k=1}^{n} x_k = 1513 \quad \text{et} \quad \min(t) = 1.0001$$

   Comparer les deux intervalles précédents pour $\alpha = 0.05$

Exercice 22

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{E}\left(\frac{1}{\theta}\right)$

1. Déterminer la loi de $\frac{1}{\theta}\sum_{k=1}^{n} X_k$
2. Déterminer pour θ , $\widehat{\mathrm{I}}_n^{1-\alpha}(\theta)$, un IC exact, unilatéral, au niveau $1-\alpha$
3. Déterminer pour θ , $\widetilde{\mathrm{I}}_n^{1-\alpha}(\theta)$, un IC asymptotique, unilatéral, au niveau $1-\alpha$

Feuille TD 6
Tests Paramétriques

Exercice 23

Soit $T = (X_1)$ un 1-échantillon de loi discrète P_T^θ telle que $\mathcal{T} = \{1, 2, 3, 4, 5, 6, 7\}$

x_k	1	2	3	4	5	6	7
$P_T^{\theta_0}(\{x_k\})$	0.01	0.02	0.03	0.05	0.05	0.07	0.77
$P_T^{\theta_1}(\{x_k\})$	0.03	0.09	0.1	0.1	0.2	0.18	0.3

On veut tester $(\mathrm{H}_0) = \{\theta = \theta_0\}$ contre $(\mathrm{H}_1) = \{\theta = \theta_1\}$

1. Déterminer tous les tests (i.e. toutes les zones de rejet associées) ayant un risque de première espèce de 10%.
2. Parmi tous les tests précédents, déterminer celui qui est le plus puissant.

Exercice 24

On sait qu'une boite contient soit trois boules rouges et cinq boules noires, soit cinq boules rouges et trois boules noires. On tire, successivement et avec remise, trois boules. Si on obtient moins de trois boules rouges, on décide que la boite contient trois boules rouges et cinq noires.

1. Formaliser le cadre statistique en définissant un vecteur aléatoire approprié.
2. Déterminer le test en explicitant sa zone de rejet
3. Calculer les risques de première et seconde espèce associés à ce test.

Exercice 25

Soit $T = (X_1)$ un 1-échantillon de loi $\mathcal{U}(]\theta, \theta + 1[)$

On veut tester $(\mathrm{H}_0) = \{\theta = 0\}$ contre $(\mathrm{H}_1) = \{\theta = \frac{1}{2}\}$

Proposer un test avec un risque de première espèce de 10% et calculer le risque de seconde espèce associé.

Exercice 26

Soit $T = (X_1)$ un 1-échantillon de loi continue de densité donnée par

$$f^{\theta}(x) = \begin{cases} \frac{\theta}{(\theta+x)^2} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases} \quad \text{avec} \quad \theta > 0$$

On veut tester $(\mathsf{H}_0) = \{\theta = 1\}$ contre $(\mathsf{H}_1) = \{\theta = 2\}$

Trouver le test le plus puissant parmi tous les tests ayant un risque de première espèce égal à 5% et donner son risque de seconde espèce.

Exercice 27

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{U}(]0, \theta[)$

1. Déterminer la fdr de $\max\{X_1, \cdots, X_n\}$
2. On veut tester $(\mathsf{H}_0) = \{\theta = \theta_0\}$ contre $(\mathsf{H}_1) = \{\theta = \theta_1\}$ avec $\theta_0 < \theta_1$

   Proposer un test avec un risque de première espèce égal à α
3. Déterminer son risque de seconde espèce.

Exercice 28

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{L}(\theta)$

1. Déterminer la loi de $\sum_{k=1}^{n} |X_k|$
2. On veut tester $(\mathsf{H}_0) = \{\theta = \theta_0\}$ contre $(\mathsf{H}_1) = \{\theta = \theta_1\}$ avec $\theta_1 < \theta_0$

   Trouver le test le plus puissant parmi tous les tests ayant un risque de première espèce égal à α
3. Déterminer son risque de seconde espèce.

Exercice 29

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{L}(\theta)$

1. Lorsque n grand, on veut tester $(\mathsf{H}_0) = \{\theta = \theta_0\}$ contre $(\mathsf{H}_1) = \{\theta = \theta_1\}$ avec $\theta_1 < \theta_0$

   Proposer un test avec un risque de première espèce égal à α
2. Appliquer lorsque $n = 200$, $\theta_0 = 3$, $\theta_1 = 1$ et $\alpha = 0.05$
3. On veut maitenant tester $(\mathsf{H}_0) = \{\theta = \theta_0\}$ contre $(\mathsf{H}_1) = \{\theta < \theta_0\}$

   Proposer un test avec un risque de première espèce égal à α

Exercice 30

Dans une maternité, on observe sur une année la naissance de 528 enfants prématurés dont 289 garçons. Proposer un test pour savoir si les garçons ont autant de chances d'être prématurés que les filles ou plus de chances.

Exercice 31

Des chiots nés avec un poids de 2kg150 arrivent au poids de 2k800 en un mois lorsqu'ils sont nourris au lait maternel.

On veut savoir si le lait maternisé (non maternel donc) est aussi nourrissant et pour cela on obserse le gain de poids en gramme sur un mois de douze chiots

$$t = (550, 620, 540, 580, 650, 640, 600, 620, 590, 670, 620, 610)$$

En supposant que le gain de poids sur un mois suit la loi $\mathcal{N}\left(m, \sigma^2\right)$, peut-on conclure au seuil de 5% que le lait maternisé est aussi nourrissant que le lait maternel ou l'est moins ?

FEUILLE TD 7
TESTS DU KHI-DEUX

Exercice 32

Dans un service public, on s'intéresse à la durée du service fourni à chaque client.

On suppose que cette durée peut-être considérée comme une var.

L'observation de cette durée pour 100 clients a donné les résultats suivants

durée en mn	$]0,1]$	$]1,2]$	$]2,3]$	$]3,4]$	$]4,5]$	$]5,6]$	$]6,7]$	$]7,8]$	$]8,9]$	$]9,10]$
nombre de personnes	5	11	12	11	9	13	11	7	9	12

Peut-on considérer au seuil de 5% que l'observation précédente correspond à celle d'un 100-échantillon de loi $\mathcal{U}\,(]0,10[)$?

Exercice 33

On lance 60 fois un dé et on note pour chaque lancer la face obtenue. Voici le tableau des résultats obtenus

face	1	2	3	4	5	6
nombre de fois	9	6	11	6	12	16

1. Décrire le cadre statistique associé.
2. Peut-on considérer au seuil de 5% que le dé est équilibré ?
3. A partir de quelle taille d'échantillon les fréquences observées auraient-elles conduit à considérer le dé pipé ?

Exercice 34

On a observé sur 100 ans le nombre de morts liés à une certaine maladie. Les observations sont les suivantes

nombre de morts	0	1	2	3	4
nombre d'années	54	32	11	2	1

1. Décrire le cadre statistique associé.
2. Peut-on considérer au seuil de 10% que le nombre de morts annuel suit une loi de Poisson ?

Exercice 35

On observe les 30 données indépendantes suivantes

$$0.5\ 1.7\ 0.1\ 0.5\ 1.6\ 0.2\ 1.6\ 0.6\ 1.2\ 1.2\ 1.1\ 0.7\ 0.5\ 2.2\ 2.1\ 0.1\ 0.6\ 0.7\ 0.3\ 0.1$$
$$0.5\ 1.8\ 2.8\ 1.6\ 3.3\ 0.2\ 0\ 1.5\ 0.9\ 0.7$$

1. Décrire le cadre statistique associé.
2. Peut-on considérer au seuil de 10% que ces 30 données représentent l'observation d'un 30-échantillon de loi exponentielle ?

Exercice 36

On relève le taux de cholestérol de 20 patients dont 10 mangent une pomme par jour (groupe A) et 10 n'en mangent pas (groupe B). Voici les résultats

taux du groupe A	7.30	7.35	6.99	7.36	7.35	7.35	7.34	7.38	7.37	7.42
taux du groupe B	7.51	7.4	7.34	7.39	7.35	7.36	7.37	7.37	7.38	7.41

1. Décrire le cadre statistique associé.
2. Peut-on considérer au seuil de 5% que le fait de manger une pomme par jour joue un rôle sur le taux de cholestérol ?

Exercice 37

On relève les diamètres de 20 pièces fabriquées par une machine A et les diamètres de 30 pièces fabriquées par une machine B.

Voici les relevés pour la machine A

$$101.5\ 104.1\ 100.5\ 96.9\ 103.3\ 104.8\ 104.3\ 95.3\ 101.7\ 95.1\ 95.7\ 103.6\ 98.1\ 98.4\ 101.3\ 98.1\ 98.6\ 105.2\ 99.4\ 96.6$$

Voici les rélevés pour la machine B

$$96.9\ 97.2\ 94.5\ 87.1\ 89.2\ 94.4\ 90.4\ 88.3\ 87.2\ 97.6\ 92.6\ 99.6\ 99\ 94.2\ 96.5\ 89.8\ 94\ 92.5\ 96.6\ 94.9\ 95.2\ 101\ 97.5\ 93.4\ 93\ 97.8\ 92.1\ 96.3\ 88.3\ 94.9$$

1. Formaliser le cadre statistique.
2. Peut-on considérer au seuil de 5% que le diamètre des pièces fabriquées par la machine A suit la loi $\mathcal{N}(100, 4)$
3. Peut-on considérer au seuil de 5% que le diamètre des pièces fabriquées par la machine B suit une loi Normale.
4. Peut-on considérer au seuil de 5% que le diamètre des pièces fabriquées par la machine A et le diamètre des pièces fabriquées par la machine B suivent la même loi.

Exercices supplémentaires

à faire en autonomie

Exercice 38 Complément de la feuille de TD1

Soit X la var de loi μ. Calculer la fonction de répartition de X lorsque

1. $\mu = \mathcal{U}\left(]a,b[\right)$
2. $\mu = \beta\left(3,2\right)$
3. $\mu = \Gamma\left(2,1\right)$
4. $\mu = \mathcal{B}\left(3,p\right)$

Exercice 39 Complément de la feuille de TD1

Soit X la var de loi μ. Calculer $\mathbb{E}(X)$ et $\mathbb{V}(X)$ lorsque

1. $\mu = \mathcal{U}\left(]a,b[\right)$
2. $\mu = \beta\left(a,b\right)$
3. $\mu = \Gamma\left(n,\theta\right)$

Exercice 40 Complément de la feuille de TD1

Soit X la var de loi μ. Calculer $\mathbb{E}(X)$ et $\mathbb{V}(X)$ pour les deux lois suivantes

1. $\mu = \mathcal{B}\left(n,p\right)$
2. $\mu = \mathcal{P}\left(\theta\right)$

N.B.
Pour calculer $\mathbb{V}(X)$, vous devrez d'abord calculer $\mathbb{E}[X(X-1)]$ puis exprimer $\mathbb{E}[X(X-1)]$ en fonction de $\mathbb{V}(X)$ et $\mathbb{E}(X)$ puis en déduire $\mathbb{V}(X)$

Exercice 41 Complément de la feuille de TD2

Soient $(X_n)_{n\geq 1}$ et $(Y_n)_{n\geq 1}$ deux suites de var telles que les var sont toutes indépendantes (au sein de la même suite mais aussi pour les suites entre elles). On suppose de plus que toutes les var suivent la même loi de Bernoulli de paramètre θ

Pour tout $m \geq 1,\ S_m = \sum_{k=1}^{m}(X_k - Y_k)$

1. Calculer $\mathbb{E}(S_m)$, $\mathbb{V}(S_m)$, $\mathbb{C}\mathrm{ov}\left(S_m, X_k\right)$ quand $k > m$. Puis $\mathbb{C}\mathrm{ov}\left(S_m, S_n\right)$ quand $n > m$
2. Montrer que $\frac{S_n}{n}$ converge en probabilité vers 0

Exercice 42 Complément de la feuille de TD2

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{P}(\theta, c)$ avec θ inconnu et $t = (x_1, \cdots, x_n)$ une observation de ce vecteur.

Déterminer si cela est possible une Statistique Exhaustive associée au modèle et l'Information de Fisher.

Exercice 43 Complément de la feuille de TD4

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi μ^θ avec θ inconnu et $t = (x_1, \cdots, x_n)$ une observation de ce vecteur.

Dans tous les cas suivants, déterminer, si il existe, l'estimateur du maximum de vraisemblance de θ

1. $\mu^\theta = \mathcal{N}(m, \theta)$
2. $\mu^\theta = \mathcal{G}\left(\frac{1}{\theta}\right)$
3. $\mu^\theta = \Gamma\left(\alpha, \frac{1}{\theta}\right)$
4. $\mu^\theta = \mathcal{P}(\theta, c)$

Exercice 44 Complément de la feuille de TD5

Application numérique pour l'exercice 17 :

Une usine fabrique des barres en principe toutes de même diamètre. L'usine veut cependant s'assurer de leur diamètre moyen. Le diamètre de plusieurs barres est mesuré. Ces mesures ont donné les résultats suivants en cm

$$t = (10.3, 10.5, 10.2, 9.9, 9.7, 10.2, 10.4, 10.3, 9.6, 9.7)$$

On admet que chaque mesure est indépendante des autres et est une var de loi $\mathcal{N}\left(m, \sigma^2\right)$

1. On suppose que $\sigma^2 = 0.1$

   Déterminer les intervalles de confiance (exacts, bilatéraux et symétriques) observés $I_1 = \mathrm{I}_n^{0.95}(m)$ et $I_2 = \mathrm{I}_n^{0.90}(m)$. Comparer I_1 et I_2

2. On suppose que σ^2 est inconnue

   Déterminer les intervalles de confiance (exacts, bilatéraux et symétriques) observés $I_3 = \mathrm{I}_n^{0.95}(m)$ puis $I_4 = \mathrm{I}_n^{0.90}(m)$. Comparer I_1 et I_3

Exercice 45 Complément de la feuille de TD5

Application numérique pour l'exercice 18 :

L'usine précédente décide de changer de procédé de fabrication pour la production de ses barres. Le nouveau procédé est moins onéreux mais l'usine désire s'assurer que le diamètre des barres issues du nouveau procédé est identique au précédent mais avant tout que la fiabilité du nouveau procédé est équivalente à celle de l'ancien procédé.

Pour cela, l'usine procède à une nouvelle série de mesures indépendantes et supposées également de loi normale. Voici les nouveaux résultats

$$v = (9.6, 10.7, 9.8, 9.5, 10.3, 10, 10.9, 9.9, 10.5)$$

Peut-on considérer que les deux procédés sont aussi fiables l'un que l'autre ?

Exercice 46 Complément de la feuille de TD5

Soit $T = (X_1, \cdots, X_n)$ un n-échantillon de loi $\mathcal{U}(]0, \theta[)$

1. Déterminer la loi de $\frac{1}{\theta} \max\{X_1, \cdots, X_n\}$
2. Déterminer pour θ, $\widehat{\mathrm{I}}_n^{1-\alpha}(\theta)$, un intervalle de confiance exact, bilatéral et symétrique, au niveau $1-\alpha$ basé sur un estimateur naturel de θ
3. Déterminer pour θ, $\widetilde{\mathrm{I}}_n^{1-\alpha}(\theta)$ un intervalle de confiance asymptotique, bilatéral et symétrique, au niveau $1-\alpha$
4. On observe T et on obtient, pour $n = 1000$, $t = (x_1, \cdots, x_n)$ tel que

$$\sum_{k=1}^{n} x_k = 516.11 \quad \text{et} \quad \max(t) = 0.9991$$

   Comparer les deux intervalles précédents pour $\alpha = 0.1$

Exercice 47 Complément de la feuille de TD6

On reprend l'exercice 18.

Proposer un test au seuil de 5% pour savoir si les deux procédés sont aussi fiables l'un que l'autre.

Exercice 48 Complément de la feuille de TD6

Pour contrôler biologiquement l'eau d'une rivière, on prélève successivement 3 litres de cette eau et on observe x_1, x_2 et x_3 le nombre de germes dans chacun des litres prélevés.

On admet que le nombre de germes présents dans un litre suit la loi $\mathcal{P}(\lambda)$

On considère que l'eau est potable si $\lambda \leq 5$ et non potable si $\lambda > 5$

On choisit de déclarer que l'eau n'est pas potable si l'un au moins des x_i est supérieur à un seuil K où K est un entier naturel.

1. On veut tester "l'eau n'est pas potable" contre "l'eau est potable" au seuil de 5% Formaliser le modèle statistique et le test via sa zone de rejet.
2. On teste maintenant "l'eau est potable" contre "l'eau n'est pas potable" et on fixe $K = 7$. Formaliser le test. Calculer son seuil. Calculer sa puissance pour $\lambda = 7.5$