

Contrôle : Statistique inférentielle

Patrick Tardivel, Cristian Mendico, Université Bourgogne Europe

19 mai 2025

Remarques : Calculatrice et formulaire autorisés. Durée 3 heures.

Les exercices sont notés sur 16 et la clarté de la rédaction est un pourcentage d'augmentation compris entre 0% et 25%. La note sur 20 est donnée par la formule : (note sur 16) $\times$ (1 + pourcentage).

Exercice 1. (3 points) On considère $(X_1, \dots, X_n)$ un échantillon d'une loi Bernoulli $\mathcal{B}(p)$ où $p \in]0, 1[$ est un paramètre inconnu modélisant la proportion de citoyens pro-européen en Allemagne choisis aléatoirement dans la population.

1. Rappeler la démarche mathématique permettant de construire un intervalle de confiance asymptotique pour p au niveau $1 - \theta$.
2. Sur 258 individus interrogés, 172 affirment être pro-européen. Calculer l'intervalle de confiance expérimental pour p de niveau asymptotique 0.90.

Exercice 2. (2.5 points) Avant la réforme du baccalauréat, la proportion des élèves qui réussissaient cet examen était de 82%. Depuis la réforme du baccalauréat, on a observé que sur 1685 élèves, 1410 ont réussi cet examen. D'après ces données, est-ce que le taux de réussite des élèves au baccalauréat est significativement supérieur à 82% ? Pour cet exercice, on traitera uniquement les étapes suivantes :

- Formulation des hypothèses.
- Statistique de test et sa loi sous l'hypothèse nulle.
- Calcul approché de la p -valeur (à l'aide de la normalité asymptotique) et conclusion.

Exercice 3. (2.5 points) Soit $(X_1, \dots, X_n)$ un échantillon d'une loi ayant pour densité

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

avec $\theta \in]0.5, 1[$. Déterminer l'estimateur du maximum de vraisemblance de θ .

Exercice 4 (Évaluation stochastique du « rêve du deuxième année » ; 8 points). Le « rêve du deuxième année » est la limite élégante, obtenue par Jean Bernoulli, de la série suivante :

$$S_N = \sum_{k=1}^N \frac{1}{k^k} \xrightarrow{N \rightarrow +\infty} \int_0^1 \exp(-x \ln(x)) dx \quad (1)$$

On note $\theta = \int_0^1 \exp(-x \ln(x)) dx$ cette constante. L'objectif de cet exercice est d'estimer stochastiquement θ . On rappelle que lorsque U suit une loi uniforme sur $[0, 1]$ alors $\mathbb{E}(f(U)) = \int_0^1 f(x) dx$; on pourra utiliser librement ce résultat dans la suite de l'exercice.

Soient $U_1, \dots, U_n$ des variables indépendantes et de même loi uniforme sur $[0, 1]$. On pose

$$T_n = \frac{1}{n} [\exp(-U_1 \ln(U_1)) + \dots + \exp(-U_n \ln(U_n))].$$

1. Justifier que T_n est un estimateur sans biais pour θ .
2. Justifier que T_n converge en probabilité vers θ lorsque n tend vers $+\infty$ et exprimer mathématiquement cette convergence.
3. En utilisant la normalité asymptotique

$$\frac{T_n - \theta}{\sqrt{\text{var}(T_n)}} \xrightarrow{n \rightarrow +\infty} \mathcal{N}(0, 1),$$

expliquer comment construire un intervalle de confiance pour θ de niveau asymptotique 95%.

Il est difficile d'obtenir une valeur exacte pour la variance de T_n . Néanmoins, il est possible de déterminer une majoration très précise de cette variance. On peut déjà observer que $\text{var}(T_n) = \text{var}(\exp(-U \ln(U))) / n$ où U suit une loi uniforme sur $[0, 1]$. Nous allons à présent déterminer une valeur numérique qui majore $\text{var}(\exp(-U \ln(U)))$.

4. Montrer l'identité suivante

$$\text{var}(\exp(-U \ln(U))) = \int_1^0 \exp(-2x \ln(x)) dx - \left(\int_1^0 \exp(-x \ln(x)) dx \right)^2.$$

5. Montrer l'inégalité

$$\sum_{k=0}^{N+1} \frac{k}{2^{k-1}} \leq \frac{1}{2} \frac{(N+1)^{N+1}}{2^{N+1}}.$$

6. Pour cette question, on utilisera librement la limite suivante :

$$Q_N = \sum_{k=1}^N \frac{k}{2^{k-1}} \xrightarrow{N \rightarrow +\infty} \int_1^0 \exp(-2x \ln(x)) dx. \quad (2)$$

(a) En utilisant la question 5 et les expressions (1) et (2), montrer que

$$Q_N + \frac{N+1}{2^N} - S_N^2 \geq \text{var}(\exp(U \ln(U))).$$

(b) En utilisant les expressions (1) et (2), établir la limite suivante

$$\lim_{N \rightarrow +\infty} Q_N + \frac{N+1}{2^N} - S_N^2 = \text{var}(\exp(-U \ln(U))).$$

7. Expliquer le code R suivant :

```
N=10
A = (1:N)~(1:N)
S = sum(A)
B = (1:N)~((1:N)*2~(1:N)-1)
Q = sum(B)
V = Q + 2~N/((N+1)~N*(N-1)) - S~2
n = 10000
U = runif(n)
Tn = mean(exp(-U*log(U)))
a = Tn - 1.96*sqrt(V/n)
b = Tn + 1.96*sqrt(V/n)
```