

Contrôle : Statistique inférentielle

Patrick Tardivel, Cristian Mendico, Université Bourgogne Europe

23 juin 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 (5 points). *On souhaite étudier l'impact d'un nouveau programme de formation sur les performances des employés d'une entreprise.*

Données : Avant puis après le programme de formation, les performances des employés sont notées $y_1^{avant}, y_2^{avant}, \dots, y_{150}^{avant}$ et $y_1^{après}, y_2^{après}, \dots, y_{150}^{après}$. On a obtenu les résultats suivants :

$$\sum_{i=1}^{150} y_i^{avant} = 7500, \quad \sum_{i=1}^{150} y_i^{après} = 7650 \quad \text{et} \quad \sum_{i=1}^{150} (y_i^{après} - y_i^{avant})^2 = 3900.$$

1. Faire un test statistique au niveau 5% pour conclure si le programme de formation a significativement amélioré les performances des employés. Rédiger soigneusement chaque étape.
2. Calculer la p-valeur.
3. Rappeler la démarche mathématique permettant de construire un intervalle de confiance pour une moyenne de niveau asymptotique $1 - \alpha$.
4. Sur les 150 employés, 95 ont vu leurs performances s'améliorer après le programme de formation. Donner un intervalle de confiance de niveau asymptotique 0,95 pour la proportion inconnue d'employés ayant amélioré leurs performances après le programme de formation.

Exercice 2 (3 points). *On souhaite savoir s'il existe une différence entre les proportions de bons hôtels (bien ou excellent) en banlieue et au centre-ville. Dans une ville, un échantillon d'hôtels évalués donne les résultats suivants.*

	Médiocre ou Moyen	Bien ou Excellent
En banlieue	28	47
Au centre-ville	19	53

1. Faire un test statistique de comparaison de proportions au niveau 5% pour conclure s'il existe une différence entre les proportions de bons hôtels en banlieue et au centre-ville. Rédiger soigneusement chaque étape.
2. Calculer la p-valeur.

Exercice 3 (8 points). *La densité d'une loi exponentielle de paramètre inconnu $\theta > 0$ est donnée ci-dessous :*

$$f(x) = \begin{cases} \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}.$$

Pour cet exercice, on considère un échantillon de taille n : $X_1, \dots, X_n$ de variables aléatoires de densité f .

1. Déterminer l'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ . En utilisant cet estimateur, donner une approximation de θ lorsque $x_1 = 11, x_2 = 11, x_3 = 6, x_4 = 14$ et $x_5 = 9$.
2. Calculer le biais et l'erreur quadratique moyenne de $\hat{\theta}$.
3. On considère le code R suivant :

```

n=1000
N=100000
theta=10
T=numeric(N)
for (i in (1:N))
{
  X=rexp(n,rate=1/theta)
  T[i]=mean(X)
}
Q=mean(T)
R=mean((T-theta)^2)

```

Que représentent les lettres n , N , Q et R ? Lorsque $n = 1000$, $\theta = 10$ et que N est infiniment grand que devraient valoir Q et R ?

4. Calculer les lois asymptotiques des statistiques suivantes :

(a) $\sqrt{n}(\hat{\theta} - \theta)$.

(b) $\sqrt{n}(\ln(\hat{\theta}) - \ln(\theta))$.

5. En utilisant la question 4-b), construire un intervalle de confiance asymptotique pour θ de niveau 0.95.