

Statistique inférentielle

TP 3 : Tests d'hypothèses

Le fichier `sinistres.csv` représente les montants d'indemnisations qu'une société d'assurance a versés pour un produit donné. On suppose que ces montants sont distribués selon une loi normale et que la variance σ^2 est connue ($\sigma^2 = 3600$). On cherche à tester $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$.

1. Générer et stocker dans une matrice 1000 échantillons de taille $n = 30$ à partir de la population de données initiales.
2. Écrire une fonction permettant d'effectuer un test de nullité de la moyenne. Cette fonction aura en arguments d'entrée un échantillon de données, l'écart-type σ et le risque de première espèce α , et retournera 1 si H_0 est rejetée et 0 sinon.
3. Parmi les 1000 échantillons générés, combien de fois peut-on affirmer que le montant de l'indemnisation moyenne est différent de 650 euros (au risque $\alpha = 5\%$) ?
4. Retrouver la formule explicite de la p-valeur de ce test. Écrire une fonction qui à partir d'un échantillon de données et de l'écart-type σ renvoie la p-valeur du test.
5. Retrouver la formule explicite de la puissance de ce test. Écrire une fonction qui à partir de n , de σ , de α et de $\delta = \mu_1 - \mu_0$ (où μ_1 est la "vraie" valeur de μ sous H_1) renvoie la puissance du test.
6. En fixant $\alpha = 5\%$, tracer la puissance du test pour $n = 30, 50$ et 100 en fonction de $\delta \in [-50, 50]$.
7. Refaire les questions 5 et 6 en utilisant la fonction `pwr.norm.test` disponible dans le package `pwr`. On rappelle que la fonction pour charger un package déjà installé sur le système est `library` et que des nouveaux packages peuvent être téléchargés et installés à partir des dépôts CRAN à l'aide de la fonction `install.packages`.
8. En fixant $\alpha = 5\%$, calculer le nombre d'observations nécessaires pour que le risque de seconde espèce ne dépasse pas 10% pour $\delta = 20$.