DÉPARTEMENT DE MATHÉMATIQUES
ET DE GÉNIE INDUSTRIEL

MTH2302D - PROBABILITÉS ET STATISTIQUE

TD nº 9

Exercice 1

Le fichier distribution_moyenne.csv contient 1000 lignes. Chaque ligne peut être considérée comme :

- Un échantillon de taille 30 (les 30 premières colonnes) d'une loi exponentielle avec $\lambda = 1$.
- La moyenne échantillonnale, notée $\overline{X}$, est calculée à partir des n premières observations, où $n \in \{1, 3, 10, 30\}$. Cela revient à considérer successivement le cas de la moyenne pour un échantillon de taille 1, 3, 10 et 30.

Considérons le cas de la moyenne $\overline{X}$ avec $n = 3$. Sur chaque ligne se trouve un échantillon de taille 3 si on prend les 3 premières colonnes. On a :

ligne	X_1	X_2	X_3	$\overline{X}(n = 3)$
1	1,47	0,33	2,29	1,361
2	0,40	0,36	0,25	0,336
3	0,06	0,56	0,23	0,285
4	0,64	0,97	1,16	0,922
⋮	⋮	⋮	⋮	⋮
1000	0,06	1,02	0,37	0,483

On a ainsi un total de 1000 échantillons de taille $n = 3$. Cela donne 1000 valeurs (observations) de la moyenne $\overline{X}$. On peut donc ainsi étudier la distribution de la moyenne $\overline{X}$ pour différentes valeurs de n en prenant le nombre de colonnes correspondantes.

a) Analyser la distribution de $\overline{X}$: statistiques descriptives, histogramme, diagramme de normalité.

b) Vérifier que la formule suivante est vérifiée approximativement :

$$V\left(\overline{X}\right) = \frac{V(X)}{n}.$$

Exercice 2

L'intervalle de confiance pour la moyenne μ d'une population normale $\mathrm{N}(\mu, \sigma^2)$ avec σ inconnu et un coefficient de confiance $1-\alpha$ basé sur un échantillon $X_1, X_2, \ldots, X_n$ est

$$\mu = \overline{X} \pm t_{\alpha/2;n-1}\frac{S}{\sqrt{n}}.$$

On pose $n = 10$ et $1 - \alpha = 95\%$.

a) Déterminer $t_{\alpha/2;n-1}$.

b) Compléter le fichier intervalle_confiance.csv :

b.1) Simuler 1000 échantillons de taille $n = 10$ pour une population normale $\mathrm{N}(\mu = 2, \sigma^2 = 4)$ (ajouter 950 lignes au fichier).

b.2) Calculer l'intervalle de confiance pour chaque échantillon.

b.3) Calculer la proportion d'échantillons (parmi les 1000 intervalles simulés) pour lesquels l'intervalle de confiance estime correctement (contient) la moyenne de la population. Interpréter le résultat obtenu.

c) Choisir au hasard l'un des 1000 intervalles de confiance simulés. Cet intervalle est-il une bonne estimation pour la moyenne de la population ? Commenter.

Exercice 3

On considère l'intervalle de confiance approximatif avec niveau de confiance $1 - \alpha$ pour estimer la moyenne d'une population dont la variance σ^2 est inconnue, à partir d'un échantillon de taille n, où n est grand (≥ 50), exprimé par :

$$\mu = \overline{X} \pm z_{\alpha/2} \frac{S}{\sqrt{n}}.$$

On pose $1 - \alpha = 95\%$.

a) Déterminer $z_{\alpha/2}$.

b) Pour une population exponentielle avec $\lambda = 0{,}5$:

b.1) Simuler 1000 échantillons de taille $n = 50$.

b.2) Calculer l'intervalle de confiance pour chaque échantillon.

b.3) Calculer la proportion d'échantillons (parmi les 1000 intervalles simulés) pour lesquels l'intervalle de confiance estime correctement (contient) la moyenne de la population. Interpréter le résultat obtenu.

Exercices suggérés du livre

3 ème édition : 9.22, 9.27, 9.16, 9.29, 9.31, 9.33, 9.35, 9.36, 9.39.

2 ème édition : 10.22, 10.24, 10.29, 10.31, 10.32, 10.33, 10.35, 10.36, 10.40.