

Statistiques Appliquées

TD 5

Calculs et simulations avec Scilab

5.1 Intelligence

Les tests d'intelligence (QI) sont conçus de façon à ce que les résultats suivent une distribution gaussienne $N(100, 10)$.

Calculer en utilisant Scilab (fonctions `cdfnor`) :

- Quelle proportion de la population obtient un résultat inférieur à 105.
- Quelle proportion de la population obtient un résultat entre 95 et 105.
- Quelles valeurs de QI, symétriques autour de la moyenne, couvrent 90% de la population.
- Que représente un résultat égal à 120 ?

Vérifier en utilisant les tables de la normale standard.

5.2 Un grand nombre d'échantillons

L'application java http://www.ruf.rice.edu/~lane/stat_sim/sampling_dist/ prélève un nombre d'échantillons de taille n à partir d'une population de distribution $p_X(x)$ donnée, calcule la valeur de la moyenne (v.a. \bar{X}) de chaque échantillon et trace un histogramme avec toutes les valeurs \bar{x} ainsi obtenues. Jouez avec cette applications pendant 5-10 minutes pour vous persuader du théorème de la limite centrale et vous familiariser avec la notion d'échantillonnage. Observez les valeurs des moyennes et des variances.

On propose ici de faire ces mêmes calculs avec Scilab.

- Générer une population de $N = 10^6$ individus suivant la loi normale (`rand`, paramètres de votre choix).
- Créer un vecteur vide pour stocker les moyennes et un autre pour les variances, les deux de longueur $k = 100$ (le nombre d'échantillons qu'on va prélever), avec les commandes :

```
x_bar_vec = zeros(k,1)
s2_vec = zeros(k,1)
```
- Dans une boucle `for` de k itérations, prélever un échantillon de taille $n = 10$ de la population.
- Calculer la moyenne \bar{x} et la variance s^2 de l'échantillon et stocker les résultats dans les vecteurs `x_bar_vec` et `s2_vec` déjà définis.
- Après la fin des itérations, générer un histogramme avec les valeurs recueillies dans `x_bar_vec` et un autre avec celles du `s2_vec`. (On peut choisir de montrer les *fréquences* (effectifs de chaque classe) ou les *densités* (fréquences relatives de chaque classe divisées par la largeur de la classe,

équivalentes à une densité de probabilité si le nombre d'observations est très élevé) avec l'option `normalization` de la fonction `histplot`.)

- f. Essayer d'autres valeurs pour N, n, k .
- g. Essayer avec une autre population, p.ex. uniforme.
- h. Lancer plusieurs simulations avec les mêmes paramètres et observer les résultats.
- i. Est-ce qu'on retrouve les mêmes phénomènes qu'avec l'applet Java ?
- j. Remarques, conclusions . . .

5.3 Écart-type et déciles de la normale

Les figures produites par le script de la v.a. uniforme contiennent plus d'informations (écart-type autour de l'espérance et déciles) que celles produites par le script de la v.a. normale.

Modifier le deuxième script de telle façon qu'il produise des figures équivalentes à celles du premier.

Script loi uniforme http://www.i3s.unice.fr/~deneire/va_uniforme.sce

Script loi normale http://www.i3s.unice.fr/~deneire/va_normale.sce