UGA Grenoble
IUT II Département informatique

TDS 3 et 4 : Échantillons et simulations

1 Échantillon

En statistique, la population est l'ensemble sur lequel on étudie une série statistique. Un échantillon est une partie (un sous-ensemble) de la population. D'un point de vue pratique, on ne dispose jamais des réalisations d'une expérience sur toute la population mais seulement sur une partie de la population qui peut être naturellement comme un échantillon.
Le mot a un sens un peu différent en probabilités. Répéter n fois une même expérience dans des conditions identiques, c'est ce qu'on appelle prélever un échantillon de taille n. On obtient alors n réalisations de cette expérience.

2 Loi forte des grands nombres (Kolmogorov, 1929)

Une loi forte des grands nombres est une loi mathématique selon laquelle la moyenne des n premiers termes d'une suite de variables aléatoires converge presque sûrement vers une constante (non aléatoire), lorsque n tend vers l'infini. Lorsque ces variables ont même espérance, par exemple lorsqu'elles ont toutes même loi, cette limite constante est l'espérance commune à toutes les variables aléatoires de cette suite. La loi forte est vérifiée sous diverses conditions portant sur les variables aléatoires de la suite :

Théorème 2.1 *Si $(X_n)_{n>0}$ est une suite de variables aléatoires indépendantes et de même loi, on a l'équivalence entre :*
(i) $E(|X_1|) < +\infty$,
(ii) la suite $\frac{X_1+\cdots+X_n}{n}$ converge presque sûrement.
De plus, si l'une de ces deux conditions équivalentes est remplie, alors la suite $\frac{X_1+\cdots+X_n}{n}$ converge presque sûrement vers la constante $\mathbb{E}(X_1)$.

3 Théorème central limite

Le théorème central limite dit que, si un grand nombre de variables aléatoires indépendantes ayant la même loi sont ajoutées, leur somme suit approximativement une loi normale.

1. Pour des échantillons distribués suivant une loi binomiale, la loi binomiale $B(n, p)$ se comporte comme la loi normale $\mathbb{N}(np, np(1-p))$ pour n grand.
2. Si $\{X_i\}_{i=1}^{\infty}$ est une suite de variables aléatoires indépendantes de même loi et de moyenne μ et de variance σ^2, alors $\bar{X}_n = \dfrac{\sum_{i=1}^{n} X_i}{n}$ suit approximativement une loi normale $\mathbb{N}(\mu, \sigma^2/n)$ pour n grand
3. ou, $Z_n = \frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}$, variable centrée réduite issue de $\bar{X}_n$, suit approximativement une loi normale $\mathbb{N}(0,1)$ pour n grand.

Exercices

Exercice 1 *Théorème central limite*
On lance un dé à 6 faces et on note par x_i le chiffre sorti au i-ième lancer. On définit de plus la suite des moyennes m_i de ces lancers par

$$m_1 = x_1, \quad m_2 = (x_1 + x_2)/2, \quad m_3 = (x_1 + x_2 + x_3)/3, \ldots$$

1. Modéliser le problème (c'est-à-dire, donner les variables aléatoires qui nous intéressent ainsi que leurs lois).

2. Quelle est la loi de M_n, variable aléatoire de modalité m_n ? Quel résultat du cours utilisez vous ?
3. Combien de lancers faudra-t-il pour que $M_n \in [3, 4]$ avec une probabilité supérieure ou égale à 0.99 ?

Exercice 2 Une société fabrique des vis en grandes quantités. La probabilité qu'une vis soit défectueuse est $p = 0.04$. Au cours de la production, on prélève un échantillon aléatoire de 600 vis. On note X, la variable aléatoire représentant le nombre de vis défectueuses parmi ces 600.

1. Quelle est la loi de X ? Donner ses paramètres.
2. Montrer que l'on peut approcher cette loi par la loi normale.
3. Donner les deux paramètres de cette loi normale.
4. Calculer la probabilité d'avoir au moins 27 vis défectueuses parmi les 600. On donnera la commande R associée.

Exercice 3 Un agent commercial doit faire du démarchage par téléphone. En moyenne, un appel sur cinq conduit à une commande.

1. Soit X, la variable aléatoire représentant le nombre de commandes passées après 60 appels.
   1. Donner la loi de X.
   2. Justifier que cette loi peut être approximée par une loi normale. En donner les paramètres.
   3. Calculer $P(X > 15)$, $P(X < 10)$,
4. Quel est le nombre minimal d'appels que l'on doit passer pour avoir au moins 75% de chances d'obtenir au moins 15 commandes ?

Exercice 4 Dans une station de ski de la région grenobloise, on admet que le nombre de skieurs sur une journée, pendant la haute saison, suit une loi normale de moyenne 50000 et d'écart-type 8000.

1. La préfecture estime que ce tourisme est "gérable" (accueil, nuisances, environnement,...) lorsque la probabilité d'accueillir moins de 55000 skieurs dans une journée dépasse 70%. Est-ce le cas ?
2. Elle souhaite réfléchir sur la base d'échantillons de 10 journées de vacances.
   1. Quelle est la loi de $\bar{X}$: "nombre moyen de skieurs sur une journée" ?
   2. Quelle est la probabilité que, dans un tel échantillon, le nombre moyen de skieurs soit inférieur à 55000 ?

Exercice 5 Une population, contenant de nombreux individus, passe une test de Q.I. Les résultats forment une variable aléatoire X, que l'on suppose de loi normale d'espérance 102 et d'écart type 15.

1. Quelle est la probabilité pour qu'un individu pris au hasard ait un QI inférieur à 100 ?
2. On veut analyser les résultats de quelques échantillons de cette population. Pour cela, on forme des groupes de 20 individus et on calcule sur ces groupes le QI moyen de chaque groupe.
   1. Donner les paramètres de la loi normale des QI moyens des échantillons de taille 20.
   2. Quelle est la probabilité pour qu'un groupe ait un QI moyen inférieur à 100 ?
   3. Au lieu de 20, combien d'individus faudrait-il choisir au hasard pour avoir moins de 5% de risque que le QI moyen groupe d'un soit inférieur à 100 ?

Exercice 6 *Loi Normale*
La taille X d'un homme français tiré au sort dans la population est supposée suivre une loi normale de moyenne 172 cm et de variance 196 :

-1- quelle est la probabilité que X soit inférieure à 160cm ?

-2- quelle est la probabilité qu'un homme tiré au sort mesure plus de deux mètres ?

-3- Que vaut $P(165 < X < 185)$?

-4- La taille des femmes françaises est modélisée par une loi gaussienne de moyenne 162cm et de variance 144. Quelle est la probabilité qu'un homme tiré au sort soit plus grand qu'une femme choisie au hasard ?