M1 finance: Statistiques non paramétriques
Fiche 1: Statistiques d'ordre et quantiles

Emeline Schmisser, `emeline.schmisser@math.univ-lille1.fr`, bureau 314 (bâtiment M3).

Exercice 1 : Loi discrètes et lois continues

1. Dire si la fonction de répartition est (à priori) discrète ou continue
   - on tire un ou plusieurs dés
   - la taille des enfants d'une classe de maternelle
   - le nombre de fautes par tranche de 1000 mots
   - le temps mis par des athlètes pour courire un 100 mètre
   - les salaires dans une entreprise
2. Est-ce que la fonction de répartition des lois suivantes est discrète ou continue ? Donner la formule de la densité/de la loi de probabilité suivant les cas.
   - loi de Cauchy
   - loi exponentielle de paramètre λ.
   - loi binomiale
   - loi multinomiale
   - loi normale
   - loi de Poisson

Exercice 2 : Statistiques d'ordre (notations)

On observe des variables aléatoires $X_1 = 5$, $X_2 = 0.3$, $X_3 = 2$, $X_4 = 4.1$, $X_5 = 2.6$.

1. Donner les valeurs des statistiques d'ordre $X_{(1)}, \ldots, X_{(5)}$.
2. Exprimer $X_{(1)}, \ldots, X_{(5)}$ en fonction de $X_1, \ldots, X_5$.

Exercice 3 : Simulation de variables aléatoires

Matlab ne sait simuler que des variables aléatoires suivant une loi uniforme ou une loi gaussienne. Expliquer comment on peut faire pour simuler :

1. Une variable aléatoire de loi β de paramètres $a = 2$ et $b = 1$. La densité de cette loi est :

$$f(x) = 2x\mathbf{1}_{x\in[0,1]}$$

2. Une exponentielle de densité $f(x) = \lambda \exp(-\lambda x)\mathbf{1}_{x\geq 0}$

Exercice 4 : Approximation par une loi normale, avec R

1. Soit $X \sim \mathcal{B}(n, p)$. Quelle est la moyenne de X ? Sa variance ? D'après le TCL, vers quoi tend la loi de $X - np$ quand n est grand ? On veut illustrer cette convergence et voir à partir de quel n on peut l'appliquer.
2. Simuler un vecteur aléatoire X de taille $m = 1000$ variables binomiales de paramètres $p = 0.5$ et $n = 50$. Créer un vecteur Y en recentrant et en renormalisant le vecteur le vecteur X (on veut que la moyenne théorique de Y soit égale à 0 et sa variance théorique à 1).
3. Faire un diagramme qqplot en utilisant la fonction `qqnorm` de `R`. Les diagrammes qqplot sont des diagrammes quantiles-quantiles : si X et N suivent la même loi, les points doivent être à peu près alignés sur la première bissectrice.
4. Est-ce que Y suit une loi normale ?

5. Même question si $n = 20$, puis $n = 10$.

On pose maintenant $p = 0,1$.

6. On pose maintenant $p = 0.1$.
7. À partir de quel n peut-on raisonnablement approcher Y par une loi normale ?

Savoir si la loi d'une variable aléatoire est proche ou non d'une loi normale est important : si c'est le cas, on pourra réaliser beaucoup de tests paramétriques : le calcul de la moyenne et de la variance sera assez simple. Si ce n'est pas le cas, on sera obligés de réaliser des tests non paramétriques : les intervalles de confiance donnés par les tests paramétriques ne seront pas bons.

Exercice 5 : Statistique d'ordre

1. On tire 5 fois un dé. Quelle est la loi de probabilité de $X_{(1)}$?

Exercice 6 : Loi uniforme, estimation d'un paramètre

On observe n variables indépendantes $X_1, X_2, \ldots, X_n$ suivant une loi $\mathcal{U}[0, \theta]$. On veut estimer θ.

1. Montrer que $X_{(n)}$ est l'estimateur du maximum de vraisemblance.
2. Donner un intervalle de confiance à 95% pour $X_{(n)}$.
3. En déduire un intervalle de confiance à 95% pour θ.
4. Est-ce que cet estimateur est paramétrique ou non paramétrique ?

Exercice 7 : Loi exponentielle et statistiques d'ordre

On observe 3 points suivant une loi exponentielle de paramètre θ.

1. Quelle est l'espérance de $X_{(3)}$?
2. Trouver c tel que $\mathbb{P}(X_{(1)} \geq c) = 0.5$.

Exercice 8 : Loi exponentielle

On observe n variables aléatoires iid de loi exponentielle de paramètre θ.

1. Soit $f_{X,Y}$ la densité jointe du couple (X, Y). Montrer que s'il existe deux fonctions g et h telles que, pour tout (x, y), $f_{X,Y}(x, y) = g(x)h(y)$, alors X et Y sont indépendantes.
2. Calculer la densité jointe de $X_{(r)}$ et de $X_{(s)}$ pour $r < s$.
3. En déduire la densité jointe de $X_{(r)}$ et de $X_{(s)} - X_{(r)}$.
4. Montrer que pour tous $(r < s)$, $X_{(r)}$ et $X_{(r)} - X_{(s)}$ sont indépendants.

Exercice 9 : Loi exponentielle et dispersion

On observe n variables aléatoires iid de loi exponentielle de paramètre 2.

1. Quelle est la densité de $R = X_{(n)} - X_{(1)}$?