Dans la suite, pour un échantillon i.i.d. $X = (X_1, X_2, \ldots, X_n)$, on désigne par $S(X)$ l'écart-type empirique défini par

$$S^2(X) := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

On a déjà montré que $S^2(X)$ est un estimateur sans biais pour la variance.

Exercice 1. Transformation de variables normales et loi du χ^2.
Soient $X = (X_1, X_2, \ldots, X_n)^T$ un vecteur aléatoire où les X_j sont i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$

(a) Montrer que si $Y = AX$ où A est une matrice orthogonale alors le vecteur Y a la même loi que le vecteur X.
Indication : Utiliser le théorème de transformation.

(b) Montrer que $\frac{n-1}{\sigma^2}S^2(X)$ suit une loi du χ^2 à $n-1$ degrés de liberté. Remarquer que la loi de ce quotient ne dépend ni de μ, ni de σ^2.
Indications : Commencer par centrer les variables X_i pour obtenir des variables $Y_i \sim \mathcal{N}(0, 1,)$, puis considérer la transformation $Z = AY$ où A est une matrice orthogonale ayant $(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$ comme première ligne.

(c) Vérifier graphiquement le résultat à l'aide de `R` (`rnorm`, `dchisq`).

Exercice 2. Intervalle de confiance pour la variance.
On considère un échantillon aléatoire $X_1, \ldots, X_n$ où les X_i suivent une loi $\mathcal{N}(\mu, \sigma^2)$.

(a) Construire un intervalle de confiance de niveau $1 - \alpha$ pour σ^2.
Indication : Utiliser l'exercice 1.

(b) Dans `R`, générer 100 échantillons de taille 6 de loi $\mathcal{N}(10, 0.8^2)$. Pour chaque échantillon, déterminer l'intervalle de confiance pour σ^2 avec $\alpha = 0.05$ et indiquer quelle proportion des intervalles contient le vrai paramètre $\sigma^2 = 0.8^2$.

Exercice 3. Intervalle de confiance pour la moyenne.
Soient $X_1, X_2, \ldots, X_n$ des variables i.i.d. de loi normale $\mathcal{N}(\mu, \sigma^2)$. Les paramètre μ et σ^2 sont inconnus mais on s'intéresse uniquement à la valeur de μ.
Soit $T(X) := \sqrt{n}\frac{\bar{X}_n - \mu}{S(X)}$.

(a) Montrer que T peut s'écrire comme

$$T \stackrel{d}{=} \frac{Y_1}{\sqrt{\frac{1}{n-1}\sum_{i=2}^{n} Y_i^2}} \stackrel{d}{=} \frac{Y_1}{\sqrt{\frac{1}{n-1}\chi^2_{n-1}}}$$

où $Y_1, Y_2, \ldots, Y_n$ sont des variables i.i.d. $\mathcal{N}(0, 1)$ et χ^2_{n-1} est indépendante de Y_1 et suit une loi du χ^2 à $n-1$ degrés de liberté.
Indication : Utiliser le résultat de l'exercice 1 ainsi que la transformation orthogonale utilisée.

Par l'exercice 2 de la série 10 du semestre d'automne, on sait que T suit une loi de Student à $n-1$ degrés de liberté.

(b) Dans `R`, les informations sur la distribution t de Student sont données en entrant `help(TDist)`. Comparer graphiquement les densités de Student pour différents degrés de liberté avec la densité normale standard.

(c) Dans une étude sur la qualité nutritionnelle des fast-food, on a mesuré la quantité de graisse pour un échantillon aléatoire de 35 hamburgers d'une certaine chaîne de restaurants. On suppose les données normales $\mathcal{N}(\mu, \sigma^2)$. La moyenne et l'écart-type empiriques de l'échantillon sont respectivement $\bar{x} = 30.2$ et $s = 3.8$ grammes de graisse.

  (i) A l'aide de cet échantillon, donner un intervalle de confiance à 95% pour le contenu en graisse moyen d'un hamburger servi dans ce restaurant.

  (ii) Peut-on affirmer que la vraie moyenne se trouve dans cet intervalle ?

  (iii) Cet intervalle serait-il plus grand ou plus petit si on avait choisi 90% ?

  (iv) Même question que (i) en approximant la loi de Student par une loi normale, i.e. en considérant que s est le vrai écart-type.

Facultatif Sur un échantillon normal $\mathcal{N}(\mu, \sigma^2)$ de taille 12, on calcule $(18.6, 26.2)$ comme intervalle de confiance à 95% pour μ. En déduire $\bar{x}$ et s les moyenne et écart-type de l'échantillon, et donner un intervalle de confiance pour μ à 98%.

À rendre jusqu'au vendredi 3 mai 2013, à 12h.