

L2Mass Statistique - Corrigé de l'interrogation du 3 avril 2019

Q.1 Soit (X_1, \dots, X_n) un échantillon d'une loi uniforme $\mathcal{U}([0, a])$.

Proposer deux estimateurs convergents du paramètre a . Sont ils biaisés ? (Argumenter sans forcément prouver.)

Réponse. Voici trois estimateurs de a ; les deux premiers sont construits avec la méthode des moments :

La moyenne empirique $\frac{1}{n}(X_1 + \dots + X_n)$ tend en probabilité (quand n tend vers $+\infty$) vers l'espérance de la loi $\mathcal{U}([0, a])$ qui vaut $\frac{a}{2}$, et est de même espérance que cette loi, donc $\frac{2}{n}(X_1 + \dots + X_n)$ est un estimateur convergent et sans biais de a .

Le deuxième moment empirique $\frac{1}{n}(X_1^2 + \dots + X_n^2)$ tend de même en probabilité vers le deuxième moment de $\mathcal{U}([0, a])$ qui vaut $\frac{1}{a} \int_0^a x^2 dx = \frac{a^2}{3}$ et ceci sans biais donc $\sqrt{\frac{3}{n}(X_1^2 + \dots + X_n^2)}$ tend en probabilité vers a , mais cette fois avec un biais parce que l'espérance d'une racine carré est plus petite que la racine carré de l'espérance.

$\max(X_1, \dots, X_n)$ tend en probabilité vers a , avec un biais car $\max(X_1, \dots, X_n) < a$ avec probabilité 1 donc l'espérance aussi.

Q.2 On teste une hypothèse \mathcal{H} par une "statistique de test" T en calculant la valeur de T en un échantillon (x_1, \dots, x_n) . A quelle condition sur cette valeur rejettera-t-on \mathcal{H} au seuil 0.02 ?

Réponse. On rejette \mathcal{H} au seuil 0.02 si la $F_T(t)$ est inférieure à 0.02 ou supérieure à $1 - 0.02$ (autrement dit si la p-valeur $\min(F_T(t), 1 - F_T(t))$ est inférieure à 0.02), où F_T est la fonction de répartition de $T(X_1, \dots, X_n)$ et t la valeur de T en l'échantillon observé (x_1, \dots, x_n) .

Q.3 Comment se compare le 3ème quartile de la loi gaussienne $\mathcal{N}(-1, 2^2)$ aux quantiles de la loi gaussienne $\mathcal{N}(0, 1)$?

Réponse. Le troisième quartile de $\mathcal{N}(-1, 2^2)$ est le nombre réel q_3 tel que $P(X \leq q_3) = \frac{3}{4}$, où X suit la loi $\mathcal{N}(-1, 2^2)$.

Or $P(X \leq q_3) = P\left(\frac{X - E(X)}{\sigma(X)} \leq \frac{q_3 - E(X)}{\sigma(X)}\right)$ et la variable $\frac{X - E(X)}{\sigma(X)}$ suit la loi $\mathcal{N}(0, 1)$.

Notons Q_3 le troisième quartile de $\mathcal{N}(0, 1)$ (sur une table on lirait $Q_3 \approx 0.67$); alors $\frac{q_3 - E(X)}{\sigma(X)} = Q_3$ donc $q_3 = \sigma(X)Q_3 + E(X) = 2Q_3 - 1$.

Ex.4 On veut estimer la proportion p de Oui à un référendum à venir en interrogeant N personnes choisies au hasard dans la population. 54% des personnes interrogées affirment qu'elles voteront Oui. On teste l'hypothèse $p = \frac{1}{2}$ par la statistique de test

$$\hat{p} = \text{proportion de Oui dans l'échantillon}$$

Que peut on dire de la loi de \hat{p} ?

Quelle doit être la taille N de l'échantillon pour pouvoir rejeter l'hypothèse $p = \frac{1}{2}$ au seuil 1% (à supposer que la proportion d'opinions favorables dans l'échantillon soit toujours 54%) ? On exprimera la réponse en terme des quantiles de la loi $\mathcal{N}(0, 1)$ (à défaut d'une table de valeurs).

Réponse. Voir le premier exercice corrigé de la feuille d'exercices type pour la seconde interrogation.

La variable aléatoire $N\hat{p}$ suit la loi binomiale $\mathcal{B}(N, p)$ qu'on peut approximer par la loi gaussienne de même espérance et de même variance $\mathcal{N}(Np, \sqrt{Np(1-p)^2})$ si Np et $N(1-p)$ sont suffisamment grands. Donc \hat{p} suit approximativement la loi gaussienne $\mathcal{N}(p, \sqrt{\frac{p(1-p)}{N}})$.

La valeur observée de \hat{p} est supposée être 0.54. On peut rejeter l'hypothèse $p = 0.5$ au seuil 1% si sous l'hypothèse $p = 0.5$ on a $P(\hat{p} \leq 0.54) > 1 - 1\%$ ce qui équivaut à $P(2\sqrt{N}(\hat{p} - 0.5) \leq 2\sqrt{N}(0.54 - 0.5)) > 1 - 1\%$ où $2\sqrt{N}(\hat{p} - 0.5)$ suit approximativement la loi $\mathcal{N}(0, 1)$.

Notons Q_{99} le 99-ième centile de $\mathcal{N}(0, 1)$ (sur une table on lit $Q_{99} \approx 2.33$); on veut donc $2\sqrt{N}(0.54 - 0.5) > Q_{99}$ donc $N > \left(\frac{Q_{99}}{0.08}\right)^2$.