Université de Bordeaux **PIN 401**
Mathématiques 2015–2016

FEUILLE D'EXERCICES n° 7 [*Correction*]

Exercice 1 –

1) loi binomiale de paramètres $(n = 1000, p = 10^{-3})$.

2) On approche X par Y suivant une loi de Poisson de paramètre $np = 1$.

a) $P(X = 2) = \binom{1000}{2} p^2(1-p)^{998} \approx 0.1840$. On trouve $P(Y = 2) = e^{-1}/2 \approx 0.1839$.

b) $P(Y = 0) + P(Y = 1) = e^{-1}(1 + 1) \approx 0.7358$. Donc $P(Y \geqslant 2) \approx 0.2642$.

Exercice 2 – Soit $n = 12$.

1) $\sum_{k=4}^{7} \binom{n}{k} \frac{1}{2^n} \approx 0.733$.

2) On note $X_k = 1$ si le i-ème lancer est une `Face` et 0 sinon. Les X_k sont des Bernoulli iid de paramètre $1/2$, donc de moyenne $1/2$, écart-type $1/2$. On pose $Z := 2\sqrt{n}(\frac{1}{n}\sum_{k=1}^{n} X_i - 1/2)$, qu'on approche par N normale centrée réduite. On pose $\alpha = \frac{-4}{\sqrt{12}}$, $\beta = \frac{2}{\sqrt{12}}$, soit

$$P(4 \leqslant \sum_{k=1}^{n} X_i \leqslant 7) = P\left(\frac{-2}{12} \leqslant \frac{1}{12}\sum_{k=1}^{12} X_i - \frac{1}{2} \leqslant \frac{1}{12}\right) = P(\alpha \leqslant Z \leqslant \beta).$$

On approche cette quantité par

$$P(\alpha \leqslant N \leqslant \beta) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-t^2/2}\, dt \approx 0.594.$$

Exercice 3 – On pose $q = 1 - p$.

1) La loi conditionnelle de Y sachant $X = n$ est une binomiale de paramètres (n, p), soit $P(Y = k \mid X = n) = \binom{n}{k} p^k q^{n-k}$, pour $k \in \{0, 1, \ldots, n\}$ et 0 sinon.

2) $P(Y = k, X = n) = P(Y = k \mid X = n)P(X = n) = \binom{n}{k} p^k q^{n-k} e^{-\lambda} \frac{\lambda^n}{n!}$, pour $k \in \{0, 1, \ldots, n\}$.

Exercice 4 – On pose $q = 1 - p$.

1) Soit $X_i = 1$ s'il y a rupture de stock en fin de semaine i et 0 sinon. Les X_i sont iid de Bernoulli, et $F_n = \frac{1}{n}(X_1 + \cdots + X_n)$. Donc nF_n est une binomiale de paramètres n, p. Soit $\mathbb{E}(F_n) = p$, $V(F_n) = \frac{1}{n}pq$ (par additivité de la variance pour des variables indépendantes).

2) Par Bienaymé-Tchebichev, on a

$$P(|F_n - \mathbb{E}(F_n)| > \lambda) \leqslant V(F_n)/\lambda^2 = \frac{pq}{n\lambda^2}.$$

Pour $\lambda = 0.01$, le majorant est < 0.25 pour $n \geqslant \frac{pq}{0.25\,\lambda^2} = 1900$.

Une deuxième solution utilise heuristiquement le théorème de la limite centrale pour approcher $Z := (F_n - \mathbb{E}(F_n))/\sqrt{pq/n}$ par une variable N normale centrée réduite :

$$P(|F_n - \mathbb{E}(F_n)| > \lambda) = P(|Z| > \delta), \quad \text{où } \delta = \lambda\sqrt{n/pq}.$$

On approche cette dernière probabilité par

$$P(|N| > \delta) = 2\int_{\delta}^{+\infty} e^{-t^2/2}\,\frac{dt}{\sqrt{2\pi}} = 2(1 - \mathcal{F}_N(\delta)),$$

qui est < 0.25 pour $\delta > 1.16$ (table), soit $n > pq(1.16/\lambda)^2 \approx 639$.

Une dernière solution utilise l'expression explicite de la loi de F_n et un logiciel de calcul formel pour déterminer le n minimal tel que

$$\sum_{k=\lceil n(p-\lambda)\rceil}^{\lfloor n(p+\lambda)\rfloor} \binom{n}{k} p^k q^{n-k} > 0.75,$$

en testant $n = 1, 2, \ldots$ (ou en faisant une dichotomie), on trouve $n = 701$.

Exercice 5 – On modélise la situation de la façon suivante : pour tout i inférieur au nombre total de votants, on pose $X_i = 1$ si la personne i vote Dupont, et 0 sinon. On suppose les X_i indépendants de même loi (Bernoulli de paramètre p inconnu). Soit $n = 100$; notre échantillon donne une réalisation de $X_1, \ldots, X_n$; on sait que

$$\hat{p} := \frac{1}{n}\sum_{i\leqslant n} X_i$$

est un estimateur de p (dont on connaît une réalisation : 58%). On cherche un intervalle de confiance $I = [\hat{p} - \delta, \hat{p} + \delta]$ aléatoire tel que $P(p \notin I) \leqslant 0.05$. Le TCL permet d'approcher

$$Z := \frac{\frac{1}{n}\sum_{i=1}^{n} X_i - p}{\sqrt{\frac{p(1-p)}{n}}}$$

par une variable N normale centrée réduite. Avec les notations précédentes et $\lambda := \delta/\sqrt{p(1-p)/n}$

$$P(p \notin I) = P\left(|Z| > \lambda\right) \approx P(|N| > \lambda),$$

qui est $\leqslant 0.05$ pour $\lambda > 1.96$ (table), soit

$$I = [\hat{p} - 1.96\sqrt{p(1-p)/n}, \hat{p} + 1.96\sqrt{p(1-p)/n}].$$

On a deux possibilités à ce stade : approcher $p(1-p)$ par notre réalisation de $\hat{p}(1-\hat{p})$, ou le majorer par $1/4$ (car $0 \leqslant p \leqslant 1$), ce qui remplace I par un intervalle légèrement plus grand.

On choisit par exemple la première solution et une réalisation des bornes de notre intervalle est

$$0.58 \pm 1.96\sqrt{\frac{0.58 \times 0.42}{100}},$$

soit $[0.48, 0.68]$. Comme la barre fatidique des 50% est à l'*intérieur* de cet intervalle, on ne peut pas conclure, au risque de 0.05 choisi.

Exercice 6 – Modélisation : pour l'individu i, le caractère étudié a la valeur X_i, de loi inconnue (en particulier, moyenne et variance inconnue). On suppose les X_i iid. Soit $n = 265$; on utilise les estimateurs standards de $\mathbb{E}(X_1)$ et $V(X_1)$:

$$\hat{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \hat{\sigma}^2 = \hat{V} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{X})^2.$$

On regroupe les valeurs prises par le caractère en classes et on arrondit de façon à identifier chaque valeur au centre de la classe. Donc une réalisation de $\hat{X}$ est

$$\mu := \frac{1}{265}(5 \times 1170 + 36 \times 1220 + 45 \times 1270 + 50 \times 1320 \\ + 61 \times 1370 + 49 \times 1420 + 19 \times 1470) \approx 1335.8,$$

et une réalisation de $\hat{V}$ est

$$\frac{1}{264}(5(1170-\mu)^2 + 36(1220-\mu)^2 + 45(1270-\mu)^2 + 50 \times (1320-\mu)^2 \\ + 61 \times (1370-\mu)^2 + 49 \times (1420-\mu)^2 + 19 \times (1470-\mu)^2) \approx 6016.8.$$

On calcule t_α tel que $P(|Y| > t_\alpha) \leqslant 0.01$ où Y suit une loi normale $N(0,1)$, par exemple $t_\alpha = 2.58$ (table). Un intervalle de confiance pour la moyenne est

$$I = \left[\hat{X} - \frac{\hat{\sigma} t_\alpha}{\sqrt{n}}, \hat{X} + \frac{\hat{\sigma} t_\alpha}{\sqrt{n}}\right],$$

dont une réalisation est $[1323, 1349]$.

Exercice 7 –

1) $\mu = 4.7$, $\sigma = 1.46$.

2) On utilise le test de Student (intervalle de confiance pour l'espérance d'une variable gaussienne d'espérance et variance inconnues) :

$$I = \left[\hat{X} - \frac{\hat{\sigma} t_\alpha}{\sqrt{n}}, \hat{X} + \frac{\hat{\sigma} t_\alpha}{\sqrt{n}}\right].$$

On veut $P(|Y| > t_\alpha) \leqslant \alpha = 0.05$, où Y suit une loi de Student à 9 degrés de libertés. La loi de Student étant symétrique, cette condition est équivalente à $2(1 - F_Y(t_\alpha)) \leqslant \alpha$ ou $F_Y(t_\alpha) \geqslant 1 - \alpha/2$, soit par exemple $t_\alpha = 2.27$ (table). Une réalisation de I est $[3.64, 5.75]$.