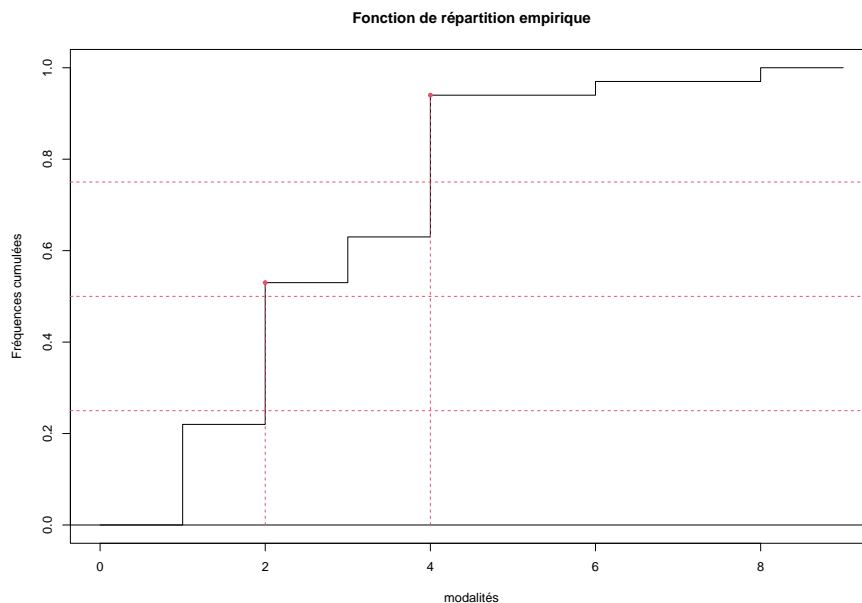


Solution Partiel STA401 Mars 2025 – 1h30

Exercice 1 (5 pts)

m_k	1	2	3	4	6	8
n_k	7	10	3	10	1	1
f_k	0.22	0.31	0.09	0.31	0.03	0.03
F_k	0.22	0.53	0.63	0.94	0.97	1



On lit sur le graphique précédent :

- La première modalité à partir de laquelle la fréquence cumulée dépasse 25% est 2. Donc le premier quartile vaut 2.
- La première modalité à partir de laquelle la fréquence cumulée dépasse 50% est 2. Donc le second quartile, appelé aussi médiane, vaut également 2.
- La première modalité à partir de laquelle la fréquence cumulée dépasse 75% est 4. Donc le troisième quartile vaut 4.

Exercice 2 (3 pts)

La population mondiale compte aujourd'hui 7,8 milliards d'habitants. Donc le groupe des 1% les plus riches contient 78 millions de personnes.

On sait que 12% des américains appartiennent au groupe des 1% les plus riches de la planète. Les américains sont $330 \cdot 10^6$, donc $0.12 \cdot 330 \cdot 10^6 = 39 \cdot 10^6$ font partie du groupe des 1% les plus riches. La proportion d'américains dans le groupe des 1% les plus riches est donc de $(78 \cdot 10^6)/(39 \cdot 10^6) \approx 51\%$.

Exercice 3 (7 pts) :

1. (1.5 pts) D'après le cours l'intervalle centré en μ_0 dans lequel la probabilité de trouver la variable \bar{X}_n vaut $1 - \alpha$ est l'intervalle de fluctuation de niveau de fluctuation $1 - \alpha$. L'expression de cet intervalle est :

$$IF = \left[\mu_0 - \frac{\sigma_0}{\sqrt{n}} u_{1-\frac{\alpha}{2}}, \mu_0 + \frac{\sigma_0}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right].$$

A.N.: pour $\mu_0 = 150, n = 25, \alpha = 5\%$ et $u_{0.975} = 1.96$ on obtient $IF_{calc} = [148.04, 151.96]$

2. (1.5 pts) La largeur de l'intervalle de fluctuation IF est donné par $2 \frac{\sigma_0}{\sqrt{n}} u_{1-\frac{\alpha}{2}}$. n tel que cette largeur soit inférieure à $2\% \mu_0$ satisfait :

$$2 \frac{\sigma_0}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \leq 2 \cdot 10^{-2} \mu_0 \iff n \geq \sigma_0^2 u_{1-\frac{\alpha}{2}}^2 \frac{10^2}{\mu_0} = 42.7.$$

Donc dès que la taille n de l'échantillon dépasse 42.7, soit $n \geq 43$ l'amplitude de l'intervalle de fluctuation est inférieur à $2\% \mu_0$.

3. (1pt) Dans le premier lot on a observé un poids moyen de $151mg$. Donc pour ce lot $\bar{x} = 152 \in IF$ et l'échantillon est conforme au niveau 95%. Dans le second $\bar{x} = 152mg \notin IF$ et l'échantillon n'est pas conforme au niveau 95%.
4. (3 pts) Dans le second lot de comprimés on note Y la variable poids de la molécule et on suppose que Y est de loi $\mathcal{N}(\mu, \sigma_0^2)$ avec μ inconnu et $\sigma_0 = 5mg$.

- (a) (1pt) L'estimateur de μ est \bar{Y}_n pour Y_1, \dots, Y_n échantillon de Y . Appliqué aux données il fournit l'estimation $\hat{\mu} = \bar{y} = 152$.
- (b) (1pt) L'expression formelle de l'intervalle de confiance symétrique pour le paramètre inconnu μ au niveau de confiance $1 - \alpha$ est dans le cas où la variance est connue :

$$IC = \left[\bar{Y}_n - \frac{\sigma_0}{\sqrt{n}} u_{1-\frac{\alpha}{2}}, \bar{Y}_n + \frac{\sigma_0}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right].$$

A.N.: pour $\hat{\mu} = 152, n = 25, \alpha = 5\%$ et $u_{0.975} = 1.96$ on obtient $IC_{calc} = [150.04, 153.96]$

- (c) (1pt) Par définition de l'intervalle IC : $P(\mu \in IC) = 1 - \alpha$. Donc sous l'hypothèse $\mu = \mu_0$: $P(\mu_0 \notin IC) = \alpha$. Autrement dit la probabilité de se tromper (soit refuser $\mu = \mu_0$ à tort) avec la règle de décision proposée vaut $\alpha = 5\%$.

Exercice 4 (6 pts) :

1. (1.5 pts) L'expression de l'intervalle de confiance de niveau approximatif $1 - \alpha$ pour p est donné par :

$$IC = \left[\bar{X}_n - \frac{\sqrt{\bar{X}_n(1 - \bar{X}_n)}}{\sqrt{n}} u_{1-\frac{\alpha}{2}}, \bar{X}_n + \frac{\sqrt{\bar{X}_n(1 - \bar{X}_n)}}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right].$$

A.N.: pour $\hat{p} = \bar{x}_n = 150, n = 100, \alpha = 10\%$ et $u_{0.95} = 1.645$ on obtient $IC_{calc} = [0.161, 0.299]$.

2. (2 pts) L'intervalle de confiance de niveau $1 - \alpha$ a la précision $\pm 1\%$ si sa largeur vaut 0.02 soit pour α tel que :

$$2 \frac{\sqrt{\bar{X}_n(1 - \bar{X}_n)}}{\sqrt{n}} u_{1-\frac{\alpha}{2}} = 0.02 \iff \alpha = 2 \left(1 - \Phi \left(\sqrt{\frac{n}{\bar{X}_n(1 - \bar{X}_n)}} 10^{-2} \right) \right) = 0.81$$

. Donc pour la précision demandée le niveau de confiance de l'intervalle est seulement de 19%.

3. (1 pt) L'intervalle (centré en p_0) de fluctuation de la moyenne empirique d'un échantillon de taille n de la loi de Bernoulli de paramètre p_0 au niveau approximatif $1 - \alpha$ est donné par (résultat de cours) :

$$IF = \left[p_0 - \frac{\sqrt{p_0(1 - p_0)}}{\sqrt{n}} u_{1-\frac{\alpha}{2}}, p_0 + \frac{\sqrt{p_0(1 - p_0)}}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \right].$$

A.N.: pour $p_0 = 0.15, n = 100, \alpha = 10\%$ et $u_{0.95} = 1.645$ on obtient $IF_{calc} = [0.091, 0.209]$.

4. (1 pt) Par définition de l'intervalle de fluctuation, la probabilité que $\bar{X}_n = \hat{p}$ soit dans l'intervalle vaut $1 - \alpha$. Donc $P(\hat{p} \notin IF | p = p_0) = \alpha$. Ici $\hat{p} = 0.23$ et $\hat{p} \notin IF$ donc on peut conclure $p \neq p_0$ avec un risque de se tromper de 10%.
5. (0.5 pt) On pourrait également conclure que $p > p_0$ puisque \hat{p} est supérieur à la borne supérieure de l'IF et donc préconiser à l'UGA d'installer de nouveaux U pour attacher les vélos.