GM – Probabilités et Statistique – Corrigés 8

À noter : les **raisonnements/justifications** des étapes de résolution sont également importantes (pas seulement le résultat final).

En salle

Exercice 1 Le paramètre d'intérêt est μ = la dette moyenne de la population des diplômés récents d'université a (1) une *valeur inconnue*. On a (2) un *échantillon aléatoire*, et on suppose que $n = 25$ est (3) *suffisament grand que le TCL s'applique*. Alors, l'IC de 95% est :

$$\overline{x} \pm 1,96 \times s/\sqrt{n} \quad \Rightarrow \quad \boxed{2430 \pm 1,96\ \times\ 2500/\sqrt{25}}$$

Exercice 2 Pour réaliser un test d'hypothèses :

1. Le paramètre d'intérêt est la proportion p d'adultes de la population possédant un téléphone portable
2. $H : p = 0,92$ (affirmation des spécialistes du marketing)
   $A : p < 0,92$ (réclamation des fabricants)
3. $T = \dfrac{\hat{p} - p_H}{\sqrt{\frac{p_H(1-p_H)}{n}}}$; $t_{obs} = \dfrac{174/200 - 0,92}{\sqrt{(0,92)(0,08)/200}} = \dfrac{-0,05}{0,19} \approx -2,61$
4. Sous $H, T \sim N(0,\ 1)$, et donc $p_{obs} = P(Z < -2.61) = 1 - P(Z < 2,61) = 1 - 0,9955 = 0,0045$
5. $p_{obs} = 0,0045 \leq 0,05 = \alpha$, donc on REJETTE l?hypothèse NULLE et en conclut que la proportion d'adultes dans la population possédant un téléphone portable est statistiquement significativement (au niveau 5%) PLUS PETIT que 92%.

• Erreur de type I : Conclure que moins de 92% des adultes dans la population possède un téléphone portable quand, en fait, (au moins) 92% des adultes de la population possèdent un téléphone portable (*c.-à-d.*, on REJETTE l'hypothèse NULLE lorsque l'hypothèse NULLE est VRAIE).

• Erreur de type II : Conclure que 92% des adultes de la population possèdent un téléphone portable alors que, en fait, moins de 92% des adultes dans la population possède un téléphone portable (*c.-à-d.*, NE REJETTE PAS l'hypothèse NULLE lorsque l'hypothèse NULLE est FAUX).

Puisque nous avons REJETÉ l'hypothèse NULLE, s'il s'agit d'une erreur ce serait une Erreur de Type I (rejetant une hypothèse VRAIE NULL).

Exercice 3 Pour réaliser un test d'hypothèses :

1. Le paramètre d'intérêt est p = la vraie probabilité de fleurs bleues
2. $H : p = 0.75$
   $A : p \neq 0.75$ (le sens du biais possible est *inconnu* à l'avance)
3. $T = \dfrac{\hat{p} - p_H}{\sqrt{\frac{p_H(1-p_H)}{n}}}$; $t_{obs} = \dfrac{142/200 - 0.75}{\sqrt{(0.75)(0.25)/200}} \approx -1.31$
4. Sous $H, T \sim N(0,\ 1)$, et donc $p_{obs} = P(Z < -1.31 \textbf{ OU } Z > 1.31) = 2\,P(Z < -1.31)$ $= 2 \times 0.0951 \approx 0.19$
5. En supposant un seuil (niveau) de signification α (p.ex. 0.05 ou 0.01), on NE REJETTE PAS l'hypothèse NULLE ; c.-à-d., il n'y a PAS de preuve statistiquement significative que les données sont incompatibles avec le modèle.

Exercice 4 Pour réaliser un test d'hypothèses :

1. Le paramètre d'intérêt est $\mu_x - \mu_y$ = la vraie différence des moyennes (nouvelle − ancienne)
2. $H : \mu_x - \mu_y = \mathbf{7}$
   $A : \mu_x - \mu_y > \mathbf{7}$ (la durée de la nouvelle batterie est plus de 7 minutes de plus)
3. $t_{obs} = [(200 - 190) - \mathbf{7}]/\sqrt{40^2/100 + 20^2/100} \approx 0,67$
4. (En supposant le TCL) Sous H, $p_{obs} = P(Z > 0,67) = 1 - 0,7486 \approx 0,25$
5. $p_{obs} = 0,25 > \alpha = 0,05$, donc on NE REJETTE PAS l'hypothèse NULLE ; la différence n'est pas significative $\Longrightarrow$ la preuve en faveur de l'affirmation de l'ingénieur (l'hypothèse ALTERNATIVE) n'est pas suffisante.

À domicile

Exercice 1 Pour réaliser un test d'hypothèses :

1. Le paramètre d'intérêt est μ = le QI moyen de la population des étudiants du programme spécialisé
2. $H : \mu = 100$
   $A : \mu > 100$ (la direction qui nous intéresse est : les étudiants du programme spécialisé ont un QI *supérieur*)
3. $T = \dfrac{\overline{X} - \mu_H}{s/\sqrt{n}}$; $t_{obs} = \dfrac{107 - 100}{15/\sqrt{25}} = 2,33$
4. Sous $H, T \sim N(0,\,1)$, et donc $p_{obs} = P(Z > 2.33) = 1 - P(Z \leq 2,33) = 1 - 0,9901 \approx 0,01$
5. Pour α = 0,05, $p_{obs} \leq \alpha$, et donc on REJETTE l'hypothèse NULLE ; il y a un résultat hautement statistiquement significatif (*c'est-à-dire*, au niveau 1%) que les étudiant.e.s du programme ont, en moyenne, un QI plus élevé que la population générale.

   [**Remarque :** nous **NE POUVONS PAS** conclure que les étudiant.e.s du programme sont plus intelligents, ou que nous avons « prouvé » que les étudiant.e.s du programme possèdent un QI plus élevé – c'est une interprétation sur le **VÉRITÉ** de la situation.

   Un test d'hypothèse ne peut pas prouver si les étudiant.e.s du programme ont un QI plus élevé ou non, seulement si l'échantillon est compatible (ou non) avec ayant le même QI que les individus du population générale, en moyenne.]

Exercice 2 Pour réaliser un test d'hypothèses :

1. Le paramètre d'intérêt est la vraie probabilité p d'un rouge
2. $H : p = 18/38$ (la roulette est équilibrée)
   $A : p \neq 18/38$ (la roulette n'est PAS équilibrée ; le sens du biais est *inconnu* à l'avance)
3. $T = \dfrac{\hat{p} - p_H}{\sqrt{\frac{p_H(1-p_H)}{n}}}$; $t_{obs} = \dfrac{(1890/3800) - (18/38)}{\sqrt{(18/38)(20/38)/3800}} \approx 2,93$
4. Sous $H, T \sim N(0,\,1)$, et donc $p_{obs} = P(Z > 2,93 \textbf{ OU } Z < -2,93) = 2\,P(Z > 2,93)$ $= 2 \times 0,0017 = 0,0034$
5. $p_{obs} = 0,0034 < 0,05 = \alpha$, donc on REJETTE l'hypothèse NULLE ; la différence est statistiquement significative et le résultat n'est pas compatible avec la variation aléatoire.

Exercice 3 (a) Pour réaliser un test d'hypothèses :

1. Le paramètre d'intérêt est $p_F - p_H$ = la différence (entre les populations des femmes et des hommes) des proportions de personnes qui ont attrapé un rhume.
2. $H : p_F - p_H = 0$
   $A : p_F - p_H \neq 0$ (les proportions sont différentes)
3. $\hat{p}_{ag} = [100(0,38) + 200(0,51)]/(100 + 200) = 0,467$,

   $$\text{donc } ES = \sqrt{\hat{p}_{ag}(1 - \hat{p}_{ag})[(1/n_1) + (1/n_2)]} = 0,061$$

   $$\implies t_{obs} = (0,38 - 0,51)/\sqrt{0,467(0,533)(1/100 + 1/200)} = -2,13$$

4. (En supposant le TCL) Sous H, $p_{obs} = 2\,P(Z >\mid t_{obs} \mid) = 2\,(1 - P(Z < 2,13)) = 2(1 - 0,9834) \approx 0,033$
5. $p_{obs} = 0,033 < 0,05 < \alpha$, donc on REJETTE l'hypothèse NULLE ; la différence entre les proportions pour les hommes et les femmes est statistiquement significative au niveau 5%. Il y a des preuves que l'efficacité diffère entre les hommes et les femmes.

(b) Pour réaliser un test d'hypothèses :

1. Le paramètre d'intérêt est $p_F - p_H$ = la différence (entre les populations des femmes et des hommes) de la proportion de personnes qui ont attrapé un rhume.
2. $H : p_F - p_H = 0$
   $A : p_F - p_H < 0$ (la proportions des femmes est plus petite)
3. $\hat{p}_{ag} = (100(0,38) + 200(0,51))/(100 + 200) = 0,467$,

   $$\text{donc } ES = \sqrt{\hat{p}_{ag}(1 - \hat{p}_{ag})[(1/n_1) + (1/n_2)]} = 0,061$$

   $$\implies t_{obs} = (0,38 - 0,51)/\sqrt{0,467(0,533)(1/100 + 1/200)} = -2,13$$

4. (En supposant le TCL) Sous H, $p_{obs} = P(Z < -2,13) = 1 - 0,9834 \approx 0,017$
5. $p_{obs} = 0,017 > 0,01 = \alpha$, donc on NE REJETTE PAS l'hypothèse NULLE ; la différence n'est pas significative. Au niveau 1%, il n'y a pas assez de preuves que le médicament est plus efficace pour les femmes que pour les hommes.

Exercice 4 Pour réaliser un test d'hypothèses :

1. Le paramètre d'intérêt est $\mu_x - \mu_y$ = la différence des moyennes des scores de la population des étudiant.e.s qui seront enseignés par Xavier et Yolande
2. $H : \mu_x - \mu_y = 0$
   $A : \mu_x - \mu_y \neq 0$ (les enseignant.e.s ne sont pas également efficaces)
3. $t_{obs} = [78 - 85]/\sqrt{(10^2/30) + (15^2/25)} \approx -1,99$
4. (En supposant le TCL) Sous H, $p_{obs} = 2\,P(Z >\mid -1,99 \mid) = 2\,P(Z > 1,99) = 2$ (1-0,9767) $\approx$ 0.046.
5. $p_{obs} = 0,046 < 0,10 = \alpha$, donc on REJETTE l'hypothèse NULLE ; la différence (de l'efficacité) est significative. (On suppose que le score du test fait une 'bonne' mesure de l'efficacité !)