GM – PROBABILITÉS ET STATISTIQUE – CORRIGÉS 9

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

En salle

Exercice 1 Pour réaliser un test d'hypothèse :

1. Le paramètre d'intérêt est μ = la durée de vie moyenne des ampoules produites par l'usine
2. $H : \mu = 1600$
   $A : \mu < 1600$ (selon le problème)
3. $T = (\overline{X} - \mu_H)/(S/\sqrt{n})$, donc $t_{obs} = (1570 - 1600)/(85/\sqrt{31}) \approx -1.97$
4. (a) Sous H, $T \sim t_{30}$; $t_{30,0.05} = -t_{30,0.95} = -1.697$; (b) $t_{30,0.01} = -t_{30,0.99} = -2.457$
5. (a) En supposant un seuil (niveau) de signification $\alpha = 0.05$:
   $t_{obs} = -1.97$ (ou -1.95) $< -1.697 = t_{30,0.05}$, donc on REJETTE l'hypothèse NULLE H.

   (b) En supposant un seuil (niveau) de signification $\alpha = 0.01$:
   $t_{obs} = -1.97$ (ou -1.95) $> -2.457 = t_{30,0.01}$, donc on NE REJETTE PAS l'hypothèse NULLE H.

Exercice 2 (a) L'intervalle de confiance correspondant est $[\overline{Y} \pm t_{n-1,0.95}\, s/\sqrt{n}]$
(b) La longueur vaut $\overline{Y} + t_{n-1,0.95}\, s/\sqrt{n} - (\overline{Y} - t_{n-1,0.95}\, s/\sqrt{n}) = 2\, t_{n-1,0.95}\, s/\sqrt{n}$
(c) $t_{9,0.95} = 1.83$, donc pour un IC à 90% on a $[\overline{Y} \pm t_{9,0.95}\, s/\sqrt{n}] = [30 \pm 1.83 \times 1.7/\sqrt{10}]$ $\approx [29.02, 30.98]$

Exercice 3 Soit F = score examen final, T = score test bonus.
pente $= r s_F/s_M = 1.2$, ordonnée à l'origine $= \overline{F} - 1.2\,\overline{T} = -29$ points, donc
pred. Final $= 1.2 \times$ Test $- 29$ points; $RMSE\ (REQM) = s_F\sqrt{1-r^2} = 20\sqrt{1-0.6^2}$
$=$ 16 points.

Exercice 4 pred. taille $= 0.25$ pouces par an $\times$ ed. $+ 66.75$ pouces
Si $x = 12$, l'estimation de régression pour la taille est 0.25*12 + 66.75 = 69.75 pouces ;
si $x = 16$, l'estimation de régression pour la taille est 0.25*16 + 66.75 = 70.75 pouces.

Cependant, même s'il existe une association entre la taille et l'éducation, il est clair que le fait d'aller à l'université ne vous rend pas plus grand. La corrélation observée est plus facilement expliquée par l'association de la taille et de l'âge. En outre, il s'agit d'une étude d'observation, donc une corrélation entre la taille et l'éducation pourrait être due à d'autres facteurs, p. ex. des facteurs liés au milieu familial.

À domicile

Exercice 1 L'intervalle de confiance est $[\overline{Y} \pm t_{24,0.975}\, s/\sqrt{n}] = [1.6 \pm 2.064 \times 0.3/\sqrt{25}] \approx [1.48, 1.72]$

Exercice 2 (a) 1. Le paramètre d'intérêt est $\mu_x - \mu_y$ = la différence des moyennes des coûts des deux concepts

2. $H : \mu_x = \mu_y \Rightarrow \mu_x - \mu_y = 0$
   $A : \mu_x \neq \mu_y \Rightarrow \mu_x - \mu_y \neq 0$

3. $s_p^2 = ((n-1)s_x^2 + (m-1)s_y^2)/(n+m-2) = (11(37.00^2) + 5(36.40^2))/16 = 1355.237$, donc $s_p = \sqrt{1355.237} \approx 36.8$
   $T = (\overline{X} - \overline{Y})/(S_p\sqrt{(n+m)/nm})$, donc $t_{obs} = (400.00 - 327.00)/\left(36.8\sqrt{(12+6)/(12 \times 6)}\right)$ ≈ 3.97
4. Sous H (et en supposant que $\sigma_x^2 = \sigma_y^2$), $T \sim t_{16} t_{16,0.975} = 2.12 < 3.97$ $(= t_{obs})$, donc $p_{obs} \leq \alpha = 0.05$
5. Puisque $p_{obs} \leq \alpha = 0.05$, donc on REJETTE l'hypothèse NULLE H

(b) $X_1, \ldots, X_n \sim iid\, N(\mu_x, \sigma_x^2)$; $Y_1, \ldots, Y_m \sim iid\, N(\mu_y, \sigma_y^2)$; $\sigma_x^2 = \sigma_y^2$; X et Y indépendantes

Exercice 3 pred. longueur $= 0.05$ cm par kg $\times$ poids (kg) $+ 439.01$ cm
Si $x = 3$, pred. longueur = 0.05*3 + 439.01 = 439.16 cm ;
si $x = 5$, pred. longueur = 0.05*5 + 439.01 = 439.26 cm.

Oui, l'étude est une expérience. En supposant que la seule différence entre les unités expérimentaux est le poids, nous pouvons donc en déduire la causalité.