Corrigé Rattrapage STATS
ING 2 - GI

QCM 1° B 2° A C

Exo 2

1° $f_n = \frac{40}{200} = 0,2$ estimation ponctuelle de p.

2° $n = 200 \Rightarrow F_n \sim \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$

$\alpha = 5\% \Rightarrow$ I.D.C $= f_n \pm 1,96\sqrt{\frac{f_n(1-f_n)}{n}}$

$= 0,2 \pm 1,96\sqrt{\frac{0,2 \times 0,8}{200}} = 0,2 \pm 0,055.$

IDC $= [0,145 \,;\, 0,255]$

3° On voudrait avoir : $1,96 \times \sqrt{\frac{0,2 \times 0,8}{n}} = 0,01$

$\Leftrightarrow \sqrt{n} = \frac{1,96\sqrt{0,2 \times 0,8}}{0,01} = 196 \times 0,4$

$\Leftrightarrow n = (78,4)^2 = 6146,56.$

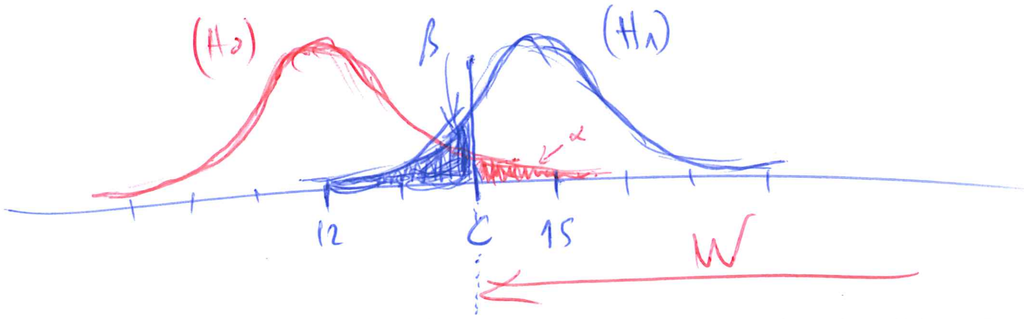
$\boxed{n \geq 6147}$

Exo 3

1° α = risque de 1ère espèce
= proba de rejeter (H_0) à tort.

β = risque de 2ème espèce = proba. de rejeter (H_1) à tort.

2°



4° $\alpha = P_{H_0}(W) = \underset{(H_0)}{P}(\bar{x} \geq 14)$
$\mathcal{N}(\mu_0, \sigma^2/n)$

Or sous (H_0), $\bar{X} \sim \mathcal{N}\left(12 ; \frac{5^2}{25}\right) = \mathcal{N}(12, 1)$.

$\rightarrow \alpha = \underset{(H_0)}{P}\left(\frac{\bar{X}-12}{1} \geq \frac{14-12}{1}\right) = P(Z \geq 2)$ avec $Z \sim \mathcal{N}(0,1)$

Table $\mathcal{N}(0,1) \rightarrow = 1 - 0{,}977 = 0{,}023$. $\alpha = 2{,}3\ \%$.

$\beta = P_{(H_1)}(\bar{X} \leq 14) = P_{(H_1)}\left(\frac{\bar{X}-15}{1} \leq \frac{14-15}{1}\right) = P(Z \leq -1)$

$= 1 - P(Z \leq 1) = 1 - 0{,}84 = 0{,}16 \Longrightarrow \beta = 16\%$.

5° Soit C le seuil correspondant ; on voudrait avoir :

$\alpha = 0{,}01 = P\left(Z \geq \frac{C-12}{5}\sqrt{n}\right) = 1 - F_Z\left(\frac{C-12}{5}\sqrt{n}\right)$

$\rightarrow F_Z\left(\frac{C-12}{5}\sqrt{n}\right) = 0{,}99 \Longrightarrow \frac{C-12}{5}\sqrt{n} = 2{,}33$

$\beta = 0{,}01 = P\left(Z \leq \frac{C-15}{5}\sqrt{n}\right) \Longrightarrow \frac{C-15}{5}\sqrt{n} = -2{,}33$

$$\begin{cases} (L_1) & \sqrt{n}\,(C-12) = 5 \times 2{,}33 \\ (L_2) & \sqrt{n}\,(C-15) = 5 \times -2{,}33 \end{cases}$$

$(L_1) - (L_2)$ $\sqrt{n}\,(15-12) = 5\,(2{,}33 + 2{,}33)$

$\sqrt{n} = \frac{23{,}3}{3} = 7{,}76$

$n = (7{,}76)^2 \simeq 60{,}3$

$\boxed{n \geq 61}$

Exo 4

1° Distrib. théorique faite sous hypothèse :

(H₀) : les 2 variables sont indépdtes.

2°
$$n_{\text{théo}} = \frac{\text{eff. ligne} \times \text{eff. colonne}}{\text{eff. total}}$$

$$22{,}70 = \frac{69 \times 51}{155}$$

3°
$$D = \sum \frac{(n_{obs} - n_{théo})^2}{n_{théo}} = \frac{(9 - 8{,}75)^2}{8{,}75} +$$

4° Sous (H_0) $D \sim \chi^2_{(p-1)(q-1)}$

p : nbre modalités var 1.
q : nbre modalités var 2.

$D \sim \chi^2_4$.

5° Avec $\alpha = 5\%$ et 4 degrés de liberté la table du khi-deux donne un seuil de : $c = 9{,}488$.

$$D = 18{,}47 > 9{,}488$$

On rejette (H_0). Il y a bien un lien entre les 2 caractères CSP et type d'éducation

---

Exo 5

1° $n = 200$ car 3 vars explicatives.

$n - 3 - 1 = 196$

2° Sales = 2,94 + 0,046 TV + 0,188 Radio − 0,001 Newspaper

3° $R^2 = 0{,}8972 \Rightarrow$ 89,72 % de la variabilité de Sales s'explique par la relation linéaire avec les 3 vars.

4° p-valeur < $2,2.10^{-16}$, on rejette l'hypothèse de nullité de tous les coef des 3 vars.
Il y a bien 1 relation linéaire entre Sales et au moins l'une des variables.

5° Si on note β_1 : le coef de TV dans le modèle :

(H_0) $\beta_1 = 0$
(H_1) $\beta_1 \neq 0$

$\hat{\beta}_1$ = estimateur de β_1.
$\hat{\sigma}(\hat{\beta}_1)$ = écart type de $\hat{\beta}_1$.

Sous (H_0), $\dfrac{\hat{\beta}_1}{\sigma(\hat{\beta}_1)} \sim T_{n-p-1}$ loi de STUDENT à 196 d.d.l

p-valeur < 2.10^{-16} → On rejette (H_0).
$\beta_1 \neq 0$, il y a bien un lien linéaire de Sales avec TV.

6° La p-valeur de la variable Newspaper = 0,86
On ne peut pas rejeter (H_0).
On doit considérer que $\beta_3 = 0$, et sortir la variable Newspaper du modèle.
On doit refaire une RLM pour expliquer la var : Sales par TV et Radio.