Exercice 1:

On réalise un test de compréhension de lecture pour des enfants de sixième dans un collège. Un échantillon de 81 garçons a obtenu une note moyenne de 17,23 et un écart-type de 4,35. Un échantillon de 86 filles a obtenu une note moyenne de 19,95 et un écart-type de 4,78. En supposant une répartition normale des résultats, peut-on dire que les garçons sont moins bons que les filles dans cette épreuve ?

On pose $m_G = 17,23$, $m_F = 19,95$. Les écarts-types expérimentaux sont $S_G = 4,35$ et $S_F = 4,78$.
On se trouve dans des grands échantillons ($n_G, n_F \geqslant 30$). Les écarts-types théoriques σ_G et σ_F sont inconnus. Pour faire un test des moyennes, il faut donc tout d'abord effectuer un test des variances. Notons que $S_G < S_F$. On établit donc le test suivant :

Test des variances :

• Étape 1 (Formulation des Hypothèses :)

$\mathcal{H}_0$: "Les variances sont égales pour les garçons et les filles",
$\mathcal{H}_1$: "La variance est plus élevée pour les filles que pour les garçons.".
On testera ces hypothèses grâce aux écarts-types théoriques σ_G et σ_F des notes obtenues. Ainsi, sous $\mathcal{H}_0$, on a $\sigma_G = \sigma_F$. Sous $\mathcal{H}_1$, on a $\sigma_G < \sigma_F$.

• Étape 2 (Choix de la statistique :)

On note S_G et S_F la proportion aléatoire de des notes obtenues, sur des échantillons respectifs de taille $n_G = 81$, $n_F = 86$.
Alors, sous $\mathcal{H}_0$, on a $F_{n_F,n_G} = \frac{n_F(n_G-1)S_F^2}{n_G(n_F-1)S_G^2}$ suit une loi de Fisher-Snedecor à (85,80) ddl.

• Étape 3 (Calculs de région critique) :

On pose $\alpha = 5\%$. On cherche f_α tel que $P[F_{n_G,n_F} \geqslant f_\alpha] = \alpha = 5\%$. On trouve $f_\alpha \approx 1,44$. On en déduit que la région critique est $\boldsymbol{K_\alpha(F) = \{F_{n_F,n_G} \geqslant 1,44\}}$**.**

• Étape 4 (Prise de décision) :

$F^{exp} \approx 1,2 \notin K_\alpha$. On décide donc d'accepter l'hypothèse $\mathcal{H}_0$: les variances sont identiques.

Dans le test précédent, on a déterminé que les variances étaient identiques. On pose donc $s = \sqrt{\frac{n_G S_G^2 + n_F S_F^2}{n_G + n_F - 2}} \approx 4,6041$.
On note μ_G et μ_F les moyennes théoriques des notes obtenues.

Test des moyennes :

• Étape 1 (Formulation des Hypothèses :)

$\mathcal{H}_0$: "Les garçons et les filles sont aussi bons à cette épreuve.",
$\mathcal{H}_1$: "Les garçons sont moins moins à cette épreuve".
On testera ces hypothèses grâce aux moyennes théoriques μ_G et μ_F des notes obtenues. Ainsi, sous $\mathcal{H}_0$, on a $\mu_G = \mu_F$. Sous $\mathcal{H}_1$, on a $\mu_G < \mu_F$.

• Étape 2 (Choix de la statistique :)

On note M_G et M_F les moyennes aléatoires des notes obtenues, sur des échantillons respectifs de taille $n_G = 14$, $n_F = 15$.
Alors, sous $\mathcal{H}_0$, comme on a des grands échantillons et que les variances sont identiques, on a $Z = \frac{M_F - M_G}{s\sqrt{\frac{1}{n_G} + \frac{1}{n_F}}} \approx \frac{M_F - M_G}{0{,}7129} \rightsquigarrow \mathcal{N}(0;1)$

• Étape 3 (Calculs de région critique) :

On pose $\alpha = 5\%$. On cherche z_α tel que $P[Z \geqslant z_\alpha] = \alpha = 5\%$. On trouve $z_\alpha = 1,645$. On en déduit que la région critique est $\boldsymbol{K_\alpha(Z) = \{T \geqslant 1,645\}}$**.**

• Étape 4 (Prise de décision) :

$Z^{exp} \approx 3,81 \in K_\alpha$. On décide donc d'accepter l'hypothèse $\mathcal{H}_1$: Les filles sont meilleures que les garçons dans cette épreuve.