Université de Lausanne
Institut de Mathématiques Appliquées
Statistique I / non-psy
Cours du professeur André Berchtold

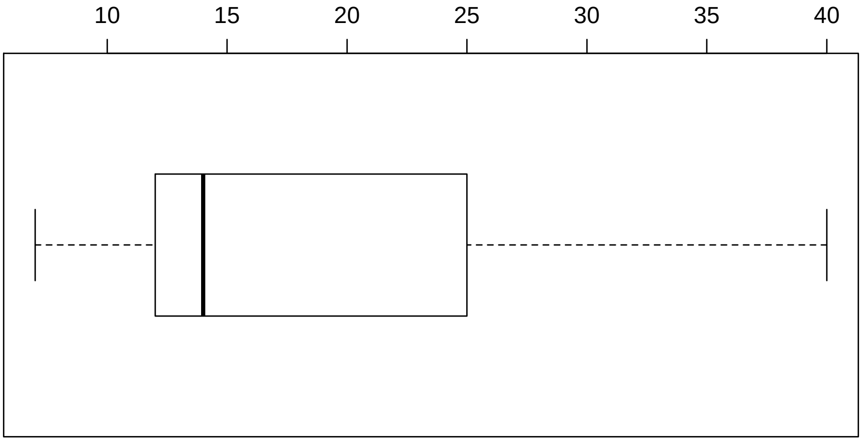
Test – 5 avril 2011 – Corrigé

Exercice 1

1. **Valeur : 9 points.**

0%	25%	50%	75%	100%
7	12	14	25	40

   Ecart interquartile : 25-12 = 13



2. **Valeur : 5 points.**
   Moyenne : $18.\overline{7}$ (18.78), variance : 93.28, écart-type : 9.66, coeff. de variation : 0.5144
3. **Valeur : 4 points.**
   – Faible nombre de données qui ne justifie pas réellement l'emploi d'un boxplot.
   – Le boxplot montre que la distribution est très fortement asymétrique avec un étalement sur la droite.
   – Les données sont particulièrement peu dispersées entre $q1$ et la médiane, et particulièrement fortement dispersées entre la médiane et le maximum.
   – La moyenne est nettement supérieure à la médiane, ce qui est aussi l'indice d'un étalement à droite.
   – La valeur 40 est très fortement supérieure aux autres et elle a donc une grande influence sur le calcul de la moyenne, de la variance et de l'écart-type.
   – L'écart auquel on peut s'attendre entre une observation prise au hasard et la moyenne est de 9.66.
   – En termes relatifs, ce même écart correspond à ± 0.5144 fois la moyenne.

Exercice 2

1. **Valeur : 8 points.**

   ```
   Expected Counts
            1        2
   1 492.7575 467.2425
   2 311.5664 295.4336
   3 122.6761 116.3239

   X-squared = 0.9544, df = 2, p-value = 0.6205
   ```

   $$V = \sqrt{\frac{0.9544}{1806}} = 0.0230$$

   Commentaire :
   – Les effectifs observés sont tous très proches des effectifs théoriques.
   – Le chi-2 montre qu'il n'y a pas formellement indépendance entre les 2 variables.
   – Le V de Cramer montre que l'association est extrêmement faible.

2. **Valeur : 2 points.**
   Etant donné que la p-valeur (0.6205) est strictement supérieure au risque de première espèce (0.05), l'hypothèse nulle est acceptée. On peut donc admettre qu'au niveau de l'ensemble de la population considérée, il y a indépendance entre le genre et le type d'études.

Exercice 3

Valeur : 3 points.
Caractéristiques de F (n=25, m=33) :

$$E(F) = \frac{33}{33-2} = 1.0645$$

$$Var(F) = \frac{2m^2(m+n-2)}{n(m-2)^2(m-4)} = \frac{121968}{696725} = 0.1751$$

Approximation :

$$\begin{aligned} P(F < 2.58) &\cong P\left(Z < \frac{2.58 - 1.0645}{\sqrt{0.1751}}\right) \\ &= P(Z < 3.62) \\ &= 0.9999 \end{aligned}$$

Exercice 4

Valeur : 4 points.
Intervalle pour la variance avec espérance inconnue.
$s^2 = 8.70^2 = 75.69$, $ns^2 = 20 \cdot 75.69 = 1513.8$

$$\left[\frac{ns^2}{q_{1-\alpha/2}^{(n-1)}} \quad ; \quad \frac{ns^2}{q_{\alpha/2}^{(n-1)}}\right]$$
$$\left[\frac{ns^2}{q_{0.975}^{(19)}} \quad ; \quad \frac{ns^2}{q_{0.025}^{(19)}}\right]$$
$$\left[\frac{1513.8}{32.85} \quad ; \quad \frac{1513.8}{8.91}\right]$$
$$[46.08 \quad ; \quad 169.90]$$

Cet intervalle de confiance représente l'ensemble des valeurs probables pour la variance de la population.

Exercice 5

1. **Valeur : 2 points.**
   Echantillon X : p=0.1639$>\alpha$=0.05 $\longrightarrow$ l'hypothèse nulle d'une moyenne égale à 5 est acceptée
   Echantillon Y : p=0.04981$<\alpha$=0.05 $\longrightarrow$ l'hypothèse nulle d'une moyenne égale à 5 est rejetée
2. **Valeur : 2 points.**
   Non, il n'y a pas de preuve, car les intervalles de confiance calculés sur les deux échantillons se chevauchent en grande partie, ce qui ne permet pas d'exclure l'hypothèse d'une égalité des moyennes.
3. **Valeur : 1 point.**

$$\bar{y} = \frac{3.600721 + 4.999279}{2} = 4.3$$