Corrigé 9

Problème 1

Afin de déterminer si le taux de participation des Suisses aux votations est en augmentation, le test suivant a été réalisé sur la base d'un échantillon représentatif de données :

H_0 : le taux de participation est constant sur les 3 dernières années
H_1 : le taux de participation a significativement augmenté au cours des 3 dernières années

La p-valeur du test vaut 0.03.
Conclure en prenant un risque de première espèce égal à 5%.
Comme la p-valeur est inférieure au risque de première espèce, nous rejetons H_0

Problème 2 Interprétation

Le logiciel R fournit le résultat suivant concernant un test d'hypothèses :

```
x <- c(498, 502, 501, 499, 503, 503, 503, 499, 500, 502)
t.test(x, mu=500, alternative="two.sided", conf.level=0.99)

        One Sample t-test

data:  x
t = 1.6771, df = 9, p-value = 0.1278
alternative hypothesis: true mean is not equal to 500
99 percent confidence interval:
 499.0622 502.9378
sample estimates:
mean of x
      501
```

Interprétez ce résultat en indiquant
– les hypothèses nulle et alternatives

$$H_0 : \quad \mu = 500$$
$$H_1 : \quad \mu \neq 500$$

– la taille de l'échantillon
  $n = 10$
– le degré de confiance de l'intervalle de confiance
  $p = 0.99$

– la valeur de la statistique de test
  $\bar{x} = 501$
– le résultat du test au seuil $\alpha = 0.01$
  hypothèse nulle non rejetée

Problème 3

En 2002 et 2003, sur la base de deux échantillons de taille 20, une société d'assurances a signé chaque jour en moyenne respectivement 115 et 108 nouveaux contrats. Un test unilatéral à gauche (α=5%) dont l'hypothèse nulle était : "les nombres moyens de nouveaux contrats signés par jour sont identiques en 2002 et 2003" a conduit à une p-valeur de 0.45.
Selon vous, peut-on dès lors raisonnablement admettre que cette société a signé journellement le même nombre de contrats en 2002 et 2003 ?

On ne peut pas dire cela, mais simplement qu'il n'y a pas suffisamment d'évidence pour pouvoir rejeter cette hypothèse. Ainsi, un tel résultat peut être dû au hasard (45% de chance pour cela)

Problème 4

Le zoo des Marécottes est un zoo alpin situé en Valais. Des améliorations du site ont été faites cette année à la buvette du zoo, et le manager désire savoir si le nombre moyen de minutes passées par chaque visiteur à la buvette a augmenté, par rapport aux 36 minutes des précédentes années. Sur un échantillon aléatoire de 200 visites cette année, une moyenne de 36.8 minutes a été calculée, avec un écart type de 11 minutes.

En utilisant une erreur de type I de 0.05, le manager est-il en droit de croire que le nombre moyen de minutes passées à la buvette du zoo a augmenté cette année ? Utiliser la méthode de la p-valeur.

a) Le paramètre est le temps moyen passé à la buvette $\mu = 36$
b)
$$\begin{array}{ll} H_0: & \mu \leq 36 \\ H_1: & \mu > 36 \end{array}$$
c) Le niveau de signification est $\alpha = 0.05$
d) La statistique de test est celle d'un test de la moyenne
$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{36.8 - 36}{\frac{11}{\sqrt{200}}} = 1.0285$$
Il est aussi possible de résoudre en utilisant la distribution de Student.
e) La p-valeur vaut $P(z \geq 1.0285) \approx 0.1515$
f) Comme la p-valeur est supérieure au niveau de signification, l'hypothèse nulle n'est pas rejetée
g) La différence entre la valeur moyenne des précédentes années et celle de cette année n'est pas suffisamment grande. Ainsi, cette différence est probablement due simplement à une erreur d'échantillonnage.

Problème 5

En tant que responsable qualité d'une entreprise d'apiculture, l'une de vos tâches est de vérifier que le remplissage des pots de miel de 500g se fasse par une machine bien calibrée. La variable exprimant le remplissage suit une loi normale. Le niveau de signification choisi est de 0.01. Vous prélevez au hasard 10 pots et vérifiez précisément leur remplissage.
498 502 501 499 503 503 501 499 500 502

Utilisez la méthode de la p-valeur pour vérifier si la machine de remplissage est bien calibrée.

```
x <- c(498, 502, 501, 499, 503, 503, 501, 499, 500, 502)
t.test(x, mu=500, alternative="two.sided", conf.level=0.99)
```

a) Le paramètre est le remplissage moyen $\mu = 500$

b)
$$\begin{aligned} H_0 : &\quad \mu = 500 \\ H_1 : &\quad \mu \neq 500 \end{aligned}$$

c) Le niveau de signification est $\alpha = 0.01$

d) La statistique de test est celle d'un test de la moyenne avec écart type inconnu

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{500.8 - 500}{\frac{1.75119}{\sqrt{10}}} = 1.4446$$

e) $P(t \geq 1.4446) \approx 0.09123414$
Or, il s'agit d'un test bilatéral.
Donc la p-valeur vaut $2 * P(t \geq 1.4446) \approx 0.1825$

Avec les tables
$2 * P(t > 1.4446) = 2 * (1 - P(t \leq 1.4446)) \in [2 * (1 - (1 - 0.1)); 2 * (1 - (1 - 0.05))]$
et donc la p-valeur est comprise entre 0.1 et 0.2

f) Comme la p-valeur est supérieure au niveau de signification, l'hypothèse nulle n'est pas rejetée

g) La machine semble donc bien calibrée.