Corrigé 5

Problème 1 Importation

Les fabricants d'ordinateurs achètent des barrettes mémoire à Taiwan au meilleur prix. Toutefois, il arrive que certaines importations ne soient pas de qualité suffisante. Un lot vient d'arriver chez Dell d'un fournisseur de Taïwan. Afin de contrôler la qualité du lot, 130 barrettes sont sélectionnées aléatoirement et testées : 13 se sont révélées de qualité insuffisante.

a) Donnez un intervalle de confiance à 90% concernant la proportion de barrettes défectueuses dans le lot.
$np = 130 \times 0.1 \geq 5$ et $n(1-p) = 130 \times 0.9 \geq 5$.
Les hypothèses sont satisfaites pour utiliser la formule de l'IC pour une proportion
$0.1 \pm 1.645\sqrt{\frac{0.1(1-0.1)}{130}} = 0.1 \pm 0.0433$ i.e. $[0.0567; 0.1433]$

b) Supposez que Dell souhaite réduire la marge d'erreur pour l'estimation donnée précédemment. Quelles options a-t-il ?
Réduire le degré de confiance ou prendre un échantillon de taille plus grande

Problème 2 Passagers par voiture*

Dans le but d'estimer le nombre moyen de passagers (conducteur compris) par véhicule automobile circulant sur l'autoroute Genève-Lausanne, un observateur a recueilli les données suivantes :

Nombre de passagers	1	2	3	4	5	6	Total
Effectif	230	248	117	76	14	3	688

```
data <- c(rep(1,230),rep(2,248),rep(3,117),rep(4,76),rep(5,14),rep(6,3))
dataT <- table(data)
moyenne <- mean(data); ecarttype <- sd(data)^2/sqrt(margin.table(dataT))
```

À l'aide d'intervalles de confiance à 95%

a) Estimer la moyenne μ de la population
Intervalle de confiance pour μ avec σ^2 inconnu

$$\mu = \bar{x} \pm t_{\frac{\alpha}{2},n-1}\ \hat{\sigma}_{\bar{X}}$$

est valide car la taille de l'échantillon (688) est suffisamment grande (≥ 30), et le théorème central limite permet de justifier l'utilisation d'une distribution normale.
Dans le cas où les données sont groupées par classes, la moyenne s'écrit

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{c} n_i \cdot x_i$$

où n_i représente la fréquence de la classe i et c le nombre de classes.
Nous avons donc

$$\overline{x} = \frac{230 \cdot 1 + 248 \cdot 2 + \cdots + 3 \cdot 6}{688} = \frac{1469}{688} \approx 2.1352$$

L'écart type de l'estimateur $\bar{X}$ se calcule de la manière suivante

$$\hat{\sigma}_{\overline{x}} = \frac{\hat{\sigma}}{\sqrt{n}}$$

et l'estimation de l'écart type de la population se calcule ainsi

$$\hat{\sigma} = \frac{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}}{\sqrt{n-1}}$$

Dans le cas où les données sont groupées par classes, cette formule s'écrit

$$\hat{\sigma} = \frac{\sqrt{\sum_{i=1}^{c} n_i \cdot (x_i - \overline{x})^2}}{\sqrt{n-1}}$$

Nous avons donc

$$\hat{\sigma} = \frac{\sqrt{230 \cdot (1 - 2.1352)^2 + 248 \cdot (2 - 2.1352)^2 + \cdots + 3 \cdot (6 - 2.1352)^2}}{\sqrt{687}} \approx 1.0875$$

et par conséquent

$$\hat{\sigma}_{\overline{x}} = \frac{1.0875}{\sqrt{688}} \approx 0.0415$$

Pour $\alpha = 5\%$ et $n = 688$, nous avons

$$t_{\frac{\alpha}{2}}^{(n-1)} = t_{0.025}^{(687)} \approx z_{0.025} \approx 1.96$$

Nous obtenons alors

$$\mu = \bar{x} \pm t_{\frac{\alpha}{2},n-1}\, \hat{\sigma}_{\bar{X}} = 2.1352 \pm 1.96 \cdot 0.0415 = 2.1352 \pm 0.0813$$

En d'autres termes, l'intervalle de confiance pour $\alpha = 5\%$ est donné par

$$[2.0538; 2.2165]$$

```
t.test(data, conf.level=0.95)$conf.int
```

b) Estimer la variance σ^2 de la population
Intervalle de confiance pour σ^2 avec μ inconnu :

$$\left[\frac{\sum(x_i - \overline{x})^2}{q_{\frac{\alpha}{2},n-1}} \; ; \; \frac{\sum(x_i - \overline{x})^2}{q_{1-\frac{\alpha}{2},n-1}}\right]$$

$n = 688$, $n - 1 = 687$, $1 - \frac{\alpha}{2} = 0.975$ D'où $\frac{\alpha}{2} = 0.025$.

Avec la fonction EXCEL KHIDEUX.INVERSE nous pouvons trouver que

$$q_{0.975}^{(687)} = 616.26 \qquad q_{0.025}^{(687)} = 761.53$$

Notons que cette fonction d'EXCEL définit la probabilité comme le risque de première espèce.
Pour des données groupées par classes, la formule $\sum_{i=1}^{n}(x_i - \overline{x})^2$ s'écrit

$$\sum_{i=1}^{c} n_i \cdot (x_i - \overline{x})^2$$

Nous avons donc

$$\sum_{i=1}^{c} n_i \cdot (x_i - \overline{x})^2 = 230 \cdot (1 - 2.1352)^2 + 248 \cdot (2 - 2.1352)^2 + \cdots + 3 \cdot (6 - 2.1352)^2 \approx 812.43$$

L'intervalle de confiance pour la variance est donné par

$$\left[\frac{812.43}{761.53}\ ;\ \frac{812.43}{616.26}\right] = [1.07\ ;\ 1.32]$$