Intervalles de Confiance et de Prevision en R(3b)

October 17, 2018

Greta Laage, Luc Adjengue

Contents

1 Estimation par intervalle de confiance à partir d'un seul échantillon

On utilise les données du fichier Notes.csv pour illustrer les calculs des sections suivantes.

```
In [4]: Notes <- read.csv("Notes.csv", header = TRUE, sep = ";",dec = ".")
```

```
In [16]: d1 = Notes$ECONO
         n = length(d1)
         sigma = 2.4
```

```
In [17]: n
         sigma
```

80
2.4

1.1 Loi normale de variance connue

1.1.1 Moyenne

Résultat théorique avec :

- σ l'écart-type
- n le nombre d'observations dans l'échantillon
- α le seuil de confiance
- $\bar{X}$ la moyenne échantillonale.

L'intervalle de confiance à $100(1-\alpha)\%$ pour la moyenne théorique μ est donné par:

$$\bar{X} \pm z_{\alpha/2}\sigma/\sqrt{n}$$

```
In [18]: qnorm(1-0.025) # calcul de z_alpha/2
         L = mean(d1) - qnorm(1-0.025) * sigma / sqrt(n)
         U = mean(d1) + qnorm(1-0.025) * sigma / sqrt(n)
         L
         U
```

1.95996398454005
11.8078364756541
12.8596635243459

1.1.2 Taille de l'échantillon

Si on veut avoir confiance à $100(1-\alpha)\%$ que l'erreur commise en estimant la moyenne soit inférieure à ϵ, alors :

$$n \geq (\frac{z_{\alpha/2}\sigma}{\epsilon})^2$$

```
In [19]: # on regarde pour plusieurs valeurs de l'erreur
         epsilon = 0.1
         (qnorm(1-0.025) * sigma/epsilon)^2

         epsilon2 = 2
         (qnorm(1-0.025) * sigma/epsilon2)^2
```

2212.68028071981
5.53170070179954

1.2 Moyenne d'une loi normale de variance inconnue

Résultat théorique avec :

- σ l'écart-type
- n le nombre d'observations dans l'échantillon
- α le seuil de confiance
- $\bar{X}$ la moyenne échantillonale.

L'intervalle de confiance à $100(1-\alpha)\%$ pour la moyenne théorique μ est donné par:

$$\bar{X} \pm t_{\alpha/2;n-1} S/\sqrt{n}$$

```
In [23]: S = sd(d1)
```

```
In [24]: qt(1-0.025,n-1) # calcul de t_alpha/2;n-1
         L = mean(d1) - qt(1-0.025,n-1) * S / sqrt(n)
         U = mean(d1) + qt(1-0.025,n-1) * S / sqrt(n)
         L
         U
```

1.99045021023013
11.7570698364427
12.9015975881295

1.3 Variance d'une loi normale

Résultat théorique: L'intervalle de confiance à $100(1-\alpha)\%$ pour σ^2 lorsque les données proviennent d'une loi normale est:

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}$$

```
In [28]: L = (n-1)*S^2/qchisq(1-0.025,n-1)
         U = (n-1)*S^2/qchisq(0.025,n-1)
         L
         U
```

5.02972450142753
9.42122556657164

1.4 La proportion dans une population

Proportion Résultat théorique: Un intervalle de confiance à $100(1-\alpha)\%$ pour p lorsque n est grand est :

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

In [31]:
```
# Exemple 9.14 p261 du livre
pchap = 0.16
n = 75
L = pchap - qnorm(1-0.025) * sqrt(pchap*(1-pchap)/n)
U = pchap + qnorm(1-0.025) * sqrt(pchap*(1-pcahp)/n)
L
U
```

0.0770307634922204
0.24296923650778

1.4.1 Taille de l'échantillon si on a une idée de la valeur de p

$$n \geq (\frac{z_{\alpha/2}}{\epsilon})^2 p(1-p)$$

In [34]:
```
epsilon = 0.05
p = 0.16
(qnorm(1-0.025)/epsilon)^2 * p * (1-p)
```

206.516826200516

1.4.2 Taille de l'échantillon si on a aucune idée de la valeur de p

$$n \geq (\frac{z_{\alpha/2}}{\epsilon})^2 0.25$$

In [35]:
```
(qnorm(1-0.025)/epsilon)^2 * 0.25
```

384.145882069412

2 Estimation par intervalle de confiance à partir de deux échantillons

2.1 Différence entre les moyennes de deux lois normales de variances connues

Résultat théorique:

$$(\bar{X}_1 - \bar{X}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

In [39]:
```
# Exemple 9.16 du livre p264
xbar1 = 87.6
n1 = 10
sigma1 = 1.0
```

```
xbar2 = 74.5
n2 = 12
sigma2 = 1.5
L = xbar1 - xbar2 - qnorm(1-0.05) * sqrt(sigma1^2/n1 + sigma2^2/n2)
U = xbar1 - xbar2 + qnorm(1-0.05) * sqrt(sigma1^2/n1 + sigma2^2/n2)
L
U
```

12.2180454983093
13.9819545016907

2.2 Différence entre les moyennes de deux lois normales de variances inconnues mais égales

Résultat théorique :

$$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2;n_1+n_2-2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Avec

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

In [43]:
```
# Exemple 9.17 du livre p266
xbar1 = 24.6
xbar2 = 22.1
s1 = 0.85
s2 = 0.98
n1 = 12
n2 = 15

Sp2 = ( (n1-1)*s1^2 + (n2-1)*s2^2 )/(n1+n2-2)
Sp2

L = xbar1 - xbar2 - qt(1-0.025,n1+n2-2) * sqrt(Sp2) * sqrt(1/n1 + 1/n2)
U = xbar1 - xbar2 + qt(1-0.025,n1+n2-2) * sqrt(Sp2) * sqrt(1/n1 + 1/n2)
L
U
```

0.855724
1.76212565316654
3.23787434683346

2.3 Rapport des variances de deux lois normales

Résultat théorique: Un intervalle de confiance à $100(1-\alpha)\%$ pour σ_1^2/σ_2^2 sur deux échantillons indépendants de lois normales :

$$\frac{S_1^2}{S_2^2 F_{\alpha/2,n_1-1,n_2-1}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2 F_{1-\alpha/2,n_1-1,n_2-1}}$$

```
In [45]: # Exemple 9.19 p270 du livre
         xbar1 = 24.6
         xbar2 = 22.1
         s1 = 0.85
         s2 = 0.98
         n1 = 12
         n2 = 15

         L = s1^2 / (s2^2 * qf(1-0.05,11,14))
         U = s1^2 / (s2^2 * qf(0.05,11,14))
         L
         U
```

0.293233861870545
2.06025961574041

2.4 La différence entre deux proportions

Résultat théorique: Un intervalle de confiance à $100(1-\alpha)\%$ pour $p_1 - p_2$ lorsque n_1 et n_2 sont grands :

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

```
In [48]: # Exemple 9.20 p271 du livre
         n1 = 75
         phat1 = 0.16
         n2 = 0.85
         phat2 = 0.12
         L = phat1-phat2-qnorm(1-0.025)*sqrt((phat1*(1-phat1))/n1 + (phat2*(1-phat2))/n2)
         U = phat1-phat2+qnorm(1-0.025)*sqrt((phat1*(1-phat1))/n1 + (phat2*(1-phat2))/n2)
         L
         U
         qnorm(1-0.025)
```

-0.655793547423694
0.735793547423694
1.95996398454005

3 Intervalles de prévision

Résultat théorique:

Un intervalle de prévision à $100(1-\alpha)\%$ pour X_{n+1} basé sur un échantillon de taille n provenant d'une loi normale est donné par :

$$\bar{X} \pm t_{\alpha/2;n-1}\sqrt{S^2\left(1+\frac{1}{n}\right)}$$

In [6]:
```
# Exemple 9.21 du livre

xbar = 1.666
s = 0.273
n = 10
alpha = 0.05

L = xbar - qt(1-0.025,9)*sqrt(s^2*(1+1/n))
U = xbar + qt(1-0.025,9)*sqrt(s^2*(1+1/n))
L
U
```

1.01828826761567
2.31371173238433