

TD n° 2 : Estimation d'autres paramètres ; PESR

Exercice 1. Une firme immobilière s'intéresse à la valeur totale des coûts pour la construction de 126 maisons. Pour cela, un échantillon de 15 maisons est sélectionné suivant un plan de sondage aléatoire de type PESR. Les coûts, en euros, pour chacune de ces 15 maisons sont :

35 505	30 200	28 700
36 450	29 800	35 300
32 600	26 400	38 000
38 100	33 200	28 500
31 600	23 400	38 200

1. Donner une estimation ponctuelle du coût total des 126 maisons.
2. Donner une estimation ponctuelle de l'écart-type de l'estimateur du coût total des 126 maisons.
3. Déterminer un intervalle de confiance au niveau 95% pour le coût total des 126 maisons.

Exercice 2. Un gérant d'hypermarché prélève un échantillon de 42 coupons de caisse parmi les 1550 de la journée suivant un plan de sondage aléatoire de type PESR. Il souhaite évaluer le montant total des ventes de produits de beauté. Il trouve que seulement 12 coupons parmi les 42 affichaient des achats de produits de beauté dont les montants sont :

12.55	30.10	28.55	14.00	13.20	24.10
48.90	14.00	35.70	26.00	24.20	11.00

(les 30 autres coupons ont des montants à 0 euro).

1. Donner une estimation ponctuelle du montant total des ventes en produit de beauté.
2. Donner une estimation ponctuelle de l'écart-type de l'estimateur du montant total des ventes en produit de beauté.
3. Déterminer un intervalle de confiance au niveau 95% pour le montant total des ventes en produit de beauté.

Exercice 3. Dans un premier temps, on considère le caractère $X = \text{"âge"}$ en années dans une population de 5 garçons : $U = \{\text{Bob, Malik, Charles, Alex, Seb}\} = \{u_1, \dots, u_5\}$. Pour tout $i \in \{1, \dots, 5\}$, soit x_i la valeur de X pour l'individu u_i . Les résultats, en années, sont :

x_1	x_2	x_3	x_4	x_5
12.5	14	17	15	16

On étudie le caractère

$$Y = \mathbb{1}_{\{X \geq 14.5\}}$$

qui vaut 1 si l'individu considéré a un âge supérieur ou égal à 14.5 ans et 0 sinon. Pour tout $i \in \{1, \dots, 5\}$, soit y_i la valeur de Y pour l'individu u_i . Les résultats sont :

y_1	y_2	y_3	y_4	y_5
0	0	1	1	1

On s'intéresse à la proportion des individus dont l'âge est supérieur ou égal à 14.5 ans.

1. Calculer la proportion-population p_U .
2. On prélève au hasard et simultanément 3 individus dans U formant ainsi un échantillon. Chaque individu a la même probabilité qu'un autre d'être sélectionné. On est donc dans le cadre d'un plan de sondage aléatoire de type PESR.
 - (a) Quel est le taux de sondage ? Combien d'échantillons peut-on former ? Expliciter les.
 - (b) Pour chaque échantillon ω , calculer la proportion-échantillon p_ω .
 - (c) Soit p_W la *var* égale à la proportion-échantillon, l'aléatoire étant dans l'échantillon considéré. Déterminer sa loi, puis calculer son espérance et sa variance.
 - (d) Retrouver les résultats de la question précédente avec les formules du cours.
3. Décrire brièvement l'enjeu des commandes R suivantes :

```

U = c("Bob", "Malik", "Charles", "Alex", "Seb")
x = c(12.5, 14, 17, 15, 16)
y = as.numeric(x >= 14.5)
library(sampling)
N = 5
n = 3
M = choose(N, n)
a = NULL
b = NULL
c = NULL
p_w = NULL
for (i in 1:M){
  t = srswor(n, N)
  a[i] = U[t != 0][1]
  b[i] = U[t != 0][2]
  c[i] = U[t != 0][3]
  p_w[i] = (1/n) * sum(y * t)
}
PESR = data.frame(a, b, c, p_w)
PESR

```

Cela renvoie :

```

      a      b      c      p_w
1     Bob Charles Seb 0.6666667
2     Bob Charles Alex 0.6666667
3     Bob   Malik Alex 0.3333333
4 Charles   Alex  Seb 1.0000000
5     Malik   Alex  Seb 0.6666667
6     Malik Charles Alex 0.6666667
7       Bob Charles Alex 0.6666667
8     Malik Charles Alex 0.6666667
9     Malik Charles Alex 0.6666667
10    Bob Charles Alex 0.6666667

```

Exercice 4. Un échantillon de 1768 étudiants d'une université est sélectionné suivant un plan de sondage aléatoire de type PESR. Dans cet échantillon, 865 étudiants fument. On suppose que cette université compte 10000 étudiants et on note p_U la proportion de fumeurs au sein des étudiants de cette université.

1. Calculer le taux de sondage.
2. Donner une estimation ponctuelle de p_U .
3. Donner une estimation ponctuelle de la variance de l'estimateur de p_U .
4. Déterminer un intervalle de confiance au niveau 95% pour p_U .
5. Déterminer la taille d'échantillon à choisir pour avoir une incertitude absolue sur p_U inférieure ou égale à 0.04 au niveau 95%.

Exercice 5. Un professeur désire évaluer le nombre d'étudiants ayant eut la moyenne à l'examen de son module. Dans le paquet contenant 30 copies au total, il prélève un échantillon de 8 copies suivant un plan de sondage aléatoire de type PESR. Parmi cet échantillon, 5 étudiants ont eut la moyenne.

1. Calculer le taux de sondage.
2. Donner une estimation ponctuelle de la proportion d'étudiants du module ayant eut la moyenne.
3. Donner une estimation ponctuelle de la variance de l'estimateur de la proportion d'étudiants du module ayant eut la moyenne.
4. Déterminer un intervalle de confiance au niveau 95% pour la proportion d'étudiants du module ayant eut la moyenne.

Exercice 6.

1. Que détermine la fonction R ci-dessous ?

```
icPESR = function(y, N) {  
  n = length(y)  
  p_w = mean(y)  
  var_p_w = (1 - n / N) * (p_w * (1 - p_w) / (n - 1))  
  a = p_w - 1.96 * sqrt(var_p_w)  
  b = p_w + 1.96 * sqrt(var_p_w)  
  print(c(a, b))  
}
```

2. Quelles sont les valeurs renvoyées par les commandes ci-dessous ?

```
icPESR(y = c(1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0), N = 85)
```

3. Que détermine la fonction R ci-dessous ?

```
n_ech = function(N, p_w, d0) {  
  n = N * p_w * (1 - p_w) * 1.96^2 / ((N * d0^2) + p_w * (1 - p_w) *  
  1.96^2)  
  print(ceiling(n))  
}
```

4. Quelles sont les valeurs renvoyées par les commandes ci-dessous ?

```
n_ech(N = 1000, p_w = 0.45, d0 = 0.2)
```