Corrigé de la Série d'exercices :

Sol. Exo 1 :

1°/ Calcul de la Moyenne et de la Variance modifiée de la Variable d'intérêt :

* La moyenne ($\overline{Y}$) :

$$\overline{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i = \frac{1}{4}(8+3+11+4) = \underline{\underline{6,5}}$$

* La variance modifiée S_Y^2 =

$$S_Y^2 = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \overline{Y})^2.$$

$$= \frac{1}{3}\left[(8-6,5)^2 + (3-6,5)^2 + (11-6,5)^2 + (4-6,5)^2\right]$$

$$S_Y^2 = \frac{41}{3} = \underline{13,66}$$

2°/ La liste d'échantillons possibles : (tirage sans remise).

$$C_4^2 = \frac{4!}{2!\,2!} = 6. \rightarrow \text{tirage sans remise.}$$

$$\begin{cases} n = 2 : \text{taille de l'échantillon.} \\ N = 4 : \text{taille de la population} \end{cases}$$

Ech (e_i)	$\bar{y}_{ei}$	s^2_{ei}	s_{ei}
{8 - 3}	5,5	12,5	3,53
{8 , 11}	9,5	4,5	2,12
{8 , 4}	6	8	2,82
{3 , 11}	7	32	5,65
{3 , 4}	3,5	0,5	0,70
{11 - 4}	7,5	24,5	4,94

3°/ le taux de sondage : $t_x = \frac{n}{N} = \frac{2}{4} = \frac{1}{2}$

4°/ $\bar{y}_{ei} = \frac{1}{2} \sum_{i=1}^{2} y_i$ → voir le tableau ci-dessus.

$s^2_{ei} = \frac{1}{n-1} \sum_{i=1}^{2} (y_i - \bar{y}_{ei})^2$ → voir le tableau ci-dessus.

5°/ $\bar{y}$: moyenne-éch.

$\bar{y}$ est une variable aléatoire.

$P(\bar{y} = \bar{y}_{ei})$?

Les valeurs que peut prendre $\bar{y}$:

$\bar{y} = \{3,5 ; 5,5 ; 6 ; 7 ; 7,5 ; 9,5\}$.

Loi de probabilité de $\bar{y}$.

$\bar{y}$	3,5	5,5	6	7	7,5	9,5	Σ
$P(\bar{y}=\bar{y}_{ei})$	1/6	1/6	1/6	1/6	1/6	1/6	1

- Espérance de $\bar{y}$.

$$E(\bar{y}) = \sum_{ei=1}^{6} \bar{y}_{ei} \, P(\bar{y} = \bar{y}_{ei})$$

$$= \frac{1}{6}\left[3,5+5,5+6+7+7,5+9,5\right].$$

$$E(\bar{y}) = \frac{39}{6} = 6,5 = \bar{Y}.$$

6/ Vérification :

$E(\bar{y}) = 6.5 = \bar{Y} \Rightarrow \bar{y}$ estime sans biais $\bar{Y}$.

7/ s^2 variance-échantillon

s^2 estime sans biais $S_Y^2 \Rightarrow E(s^2) = S_Y^2$

$$E(s^2) = \sum_{ei=1}^{6} s_{ei}^2 \, P(s^2 = s_{ei}^2).$$

Les valeurs que peut prendre s^2 :

$$s^2 = \{0{,}5 \, , \, 4{,}5 \, , \, 8 \, , \, 12{,}5 \, , \, 24{,}5 \, , \, 32\}.$$

$$E(s^2) = \frac{1}{6}\left[0{,}5 + 4{,}5 + 8 + 12{,}5 + 24{,}5 + 32\right]$$

$$E(s^2) = \frac{82}{6} = 13{,}66 = S_Y^2.$$

d'où

s^2 estime sans biais S_Y^2.

Sol. Ex n^o 2 :

Soit P : la proportion d'individus atteints par une maladie professionnelle dans une entreprise de 1500 salariés.

$$\hat{P} = \frac{3}{10} = 0{,}3 \, . \quad \alpha = 5\% \, ; \; Z_{1-\frac{\alpha}{2}} = 1{,}96$$

$$IC_P = \begin{cases} \left[\hat{P} \mp Z_{1-\frac{\alpha}{2}} \sqrt{\dfrac{\hat{P}(1-\hat{P})}{n-1}}\right] \longrightarrow \text{PEAR} \\ \left[\hat{P} \mp Z_{1-\frac{\alpha}{2}} \sqrt{(1-t_x)\dfrac{\hat{P}(1-\hat{P})}{n-1}}\right] \longrightarrow \text{PESR} \end{cases}$$

→ Calcul de la taille de l'éch. pour que la longueur totale d'un IC avec un niveau de confiance 0,95 soit < 0,01.

a/ Cas d'1 PEAR :

Soient L : longueur de l'intervalle de confiance

et

$\hat{p}$: proportion d'individus atteints par la maladie professionnelle.

Donc,

$$IC_p = \left[\hat{p} - Z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{p})} \, ; \, \hat{p} + Z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{p})}\right]$$

tq $V(\hat{p}) = \dfrac{\hat{p}(1-\hat{p})}{n-1}$

$$L = \left[\hat{p} + Z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{p})}\right] - \left[\hat{p} - Z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{p})}\right]$$

$$\boxed{L = 2\, Z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{p})}}$$

$$L < 0,01 \Rightarrow 2\, Z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{p})} < 0,01$$

$$\Rightarrow 4\, Z^2_{1-\frac{\alpha}{2}} \left[\frac{\hat{p}(1-\hat{p})}{n-1}\right] < (0,01)^2$$

$$\Rightarrow (n-1) \geq \frac{4(1{,}86)^2 (0{,}3)(0{,}7)}{(10^{-4})}$$

$$\Rightarrow \boxed{n \geq 32256+1}$$

$$(n \geq 32257)$$

Sol. Exo n°3 :

1/ taux de sondage $t_x = \frac{n}{N} = \frac{7}{100}$

n : taille de l'éch.

N : taille de la pop.

2/ Estimation ponctuelle de la moyenne des employés des 100 entreprises :

$$\hat{Y} = \bar{y} = \frac{1}{n} \sum_{i=1}^{7} y_i =$$

$$= \frac{1}{7}\left[10+5+15+20+12+8+10\right]$$

$$= \frac{80}{7} = 11{,}42 \simeq 11 \text{ employés / Etp.}$$

3/ ~~[illegible]~~

3/ L'ecart-type corrigé de $\hat{\bar{y}}$.

$$\hat{\sigma}_{\hat{\bar{y}}} = \sqrt{\hat{V}(\hat{\bar{y}})} = \begin{cases} \sqrt{\dfrac{s^2}{n}} & \text{PEAR} \\ \sqrt{(1-t_x)\dfrac{s^2}{n}} & \text{PESR} \end{cases}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{7} (y_i - \bar{y})^2$$

$$= \frac{1}{6} \left[(10-11{,}42)^2 + (5-11{,}42)^2 + \dots + (10-11{,}42)^2 \right]$$

$$s^2 = \frac{143{,}712}{6} = \underline{23{,}952}$$

d'où

$$\hat{\sigma}_{\hat{\bar{y}}} = \begin{cases} \sqrt{\dfrac{23{,}952}{7}} = 1{,}84 & \text{PEAR.} \\ \sqrt{(1-0{,}07)(3{,}42)} = 1{,}78 & \text{PESR.} \end{cases}$$

4/ $$IC_{\bar{Y}} = \left[\hat{\bar{Y}} \mp z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{\bar{Y}})} \right].$$

$$IC_{\bar{Y}} = \begin{cases} [11,42 \mp 1,96(1,84)] \longrightarrow PEAR. \\ [11,42 \mp 1,96(1,78)] \longrightarrow PESR. \end{cases}$$

5/ Soit e : l'incertitude de la moyenne de la pop.

$$IC_{\bar{Y}} = [\bar{y} \mp e].$$
$$= [\hat{\bar{Y}} \mp e].$$

Pour un PEAR : $e = \mp z_{1-\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}}$

Donc $|e| \leq 1 \Rightarrow \left| z_{1-\frac{\alpha}{2}} \sqrt{\frac{s^2}{n}} \right| \leq 1$

$$\Rightarrow z^2_{1-\frac{\alpha}{2}} \frac{s^2}{n} \leq 1$$

$$\Rightarrow n \geq z^2_{1-\frac{\alpha}{2}} s^2 \Rightarrow n \geq (1,96)^2 (23,9$$

$$n \geq 92.$$

Sol. EXϕ 04 :

$N = 50000$
$n = 250$
x : la Variable d'intérêt (distance parcourue).

$$\sum_{i=1}^{250} x_i = 15150 \,, \quad \sum_{i=1}^{250} x_i^2 = 1155400 .$$

1/ La quantité $\sum_{i=1}^{250} x_i = 15150$ signifie la distance totale parcourue par les individus de l'échantillon

2/ Variance Corrigée échantillon (s^2).

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2 \Rightarrow (n-1) s^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 \ldots (1)$$

tandis que la Variance échantillon σ_x^2 est donnée par :

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2$$

$$\Rightarrow n \sigma_x^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 \ldots (2)$$

de ① et ② :

$$(n-1)s^2 = n\,\sigma_x^2 \Rightarrow \boxed{s^2 = \frac{n}{n-1}\sigma_x^2}$$

on calcul σ_x^2 :

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$$

$$= \frac{1}{250}(1155400) - \left(\frac{15150}{250}\right)^2$$

$$= 4621{,}6 - (60{,}6)^2$$

$$= 4621{,}6 - (3672{,}36)$$

$$\boxed{\sigma_y^2 = 949{,}24}$$

Donc

$$s^2 = \frac{250}{250-1}(949{,}24)$$

$$\boxed{s^2 = 953{,}03}$$

3/

a/ $IC_w = \left[\hat{\bar{x}} \mp z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{\bar{x}})}\right]$.

$\alpha = 10\% \Rightarrow z_{1-\frac{\alpha}{2}} = 1,62$.

$\alpha = 5\% \Rightarrow z_{1-\frac{\alpha}{2}} = 1,96$.

$\hat{\bar{x}}$ et $\sqrt{V(\hat{\bar{x}})}$ étant fixes, alors l'intervalle le plus large est celui relatif à un niveau de confiance de 95%.

b/ pour $\alpha = 5\%$ (Niveau de confiance de 95%)

$$IC_w = \left[\hat{\bar{x}} \mp z_{1-\frac{\alpha}{2}} \sqrt{V(\hat{\bar{x}})}\right]$$

$$= \left[\bar{x} \mp z_{1-\frac{\alpha}{2}} \sqrt{\hat{V}(\bar{x})}\right].$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{250} x_i = \frac{1}{250} (15150) = 60,6$$

$$\hat{V}(\bar{x}) = \frac{s^2}{n} = \frac{953,03}{250} = 3,81$$

Donc

$$IC_{m} = \left[60,6 \mp 1,96 \left(\sqrt{3,81}\right) \right].$$

$$= \left[60,6 \mp 1,96(1,95) \right].$$

$$IC_{m} = \left[60,6 \mp 3,82 \right]$$

4/ la distance totale parcourue.

$$\hat{T} = N \cdot \bar{x} = 50000 (60,6) = 3030\,000 \text{ km})$$