SONDAGE ALEATOIRE SIMPLE ET SONDAGES ALEATOIRES STRATIFIES

1. Échantillon aléatoire simple

On tire un échantillon aléatoire simple sans remise de taille 100 dans une population de 3000 salariés d'une entreprise. On observe sur l'échantillon un salaire moyen de 1000 euros et un écart-type de 600 euros. Estimer par intervalle de confiance à 95% le salaire moyen des salariés de l'entreprise.

Solution

Soit μ et σ^2 le salaire moyen et la variance sur l'ensemble des 3000 salariés.
Soit $\bar{y} = 1000$ et s = 600 le salaire moyen et l'écart-type sur l'échantillon de taille 100.
Le taux de sondage $\frac{100}{3000}$ est négligeable. On déduit l'estimation de μ par intervalle de confiance à 95% :

$$\bar{y} \pm 1.96 \frac{s}{\sqrt{n}} \qquad 1000 \pm 1.96 \frac{600}{\sqrt{100}} \qquad 1000 \pm 118 \qquad [882\ ;\ 1118]$$

2. Décomposition de la moyenne et de la variance

Les 3000 salariés sont composés de 300 cadres et 2700 opérateurs. Dans chacun des deux groupes, la moyenne et l'écart-type du salaire sont donnés dans le tableau suivant :

	Effectif	Salaire moyen	Écart-type
Cadres	300	2400	1200
Opérateurs	2 700	900	100
Ensemble	3 000	?	?

Calculer la moyenne et l'écart-type du salaire de l'ensemble des salariés.

Solution

Soit μ et σ^2 la moyenne et la variance du salaire sur l'ensemble des salariés.

On a :

$$\mu = \frac{1}{10} \times 2400 + \frac{9}{10} \times 900 = 1050$$

$$\sigma^2_{\text{intra}} = \frac{1}{10} \times (1200)^2 + \frac{9}{10} \times (100)^2 = 153000$$

$$\sigma^2_{\text{inter}} = \frac{1}{10} \times (2400 - 1050)^2 + \frac{9}{10} \times (900 - 1050)^2 = 202500$$

$$\sigma^2 = \sigma^2_{\text{inter}} + \sigma^2_{\text{intra}} = 355500 = (596)^2$$

$$\boxed{\mu = 1050 \text{ et } \sigma = 596}$$

Remarque :

La moyenne des écart-types est $\overline{\sigma} = \frac{1}{10} 1200 + \frac{9}{10} 100 = 210$ et on a toujours : $\left(\overline{\sigma}\right)^2 \leq \sigma^2_{\text{int ra}}$.

3. Comparaison de l'échantillon aléatoire simple, stratifié proportionnel et stratifié optimal

Si l'on réalise un *sondage aléatoire simple* sans remise de taille 100, l'estimateur $\overline{Y}$ de μ a pour variance :

$$V\left(\overline{Y}\right) = \frac{\sigma^2}{n} = \frac{355500}{100} = (59.6)^2$$

Un intervalle de confiance déduit du sondage aléatoire simple est de la forme :

$$\overline{y} \pm 1.96 \times 59.6$$
$$\overline{y} \pm 116.8$$

Si l'on réalise un *sondage aléatoire stratifié proportionnel* sans remise de taille globale 100, on interroge 10 cadres et 90 opérateurs et l'estimateur $\overline{Y}_S$ de μ a pour variance :

$$V\left(\overline{Y}_S\right) = \frac{\sigma^2_{\text{intra}}}{n} = \frac{153000}{100} = (39.1)^2$$

Un intervalle de confiance déduit du sondage aléatoire stratifié proportionnel est de la forme :

$$\bar{y}_S \pm 1.96 \times 39.1$$
$$\bar{y}_S \pm 76.6$$

Si l'on réalise un *sondage aléatoire stratifié optimal* sans remise de taille globale 100, on interroge 57 cadres (120/210) et 43 opérateurs (90/210) et l'estimateur $\bar{Y}_S$ de μ a pour variance :

$$\mathrm{V}\left(\bar{Y}_S\right) = \frac{\bar{\sigma}^2}{n} = \frac{(210)^2}{100} = (21)^2$$

Un intervalle de confiance construit à partir de l'échantillon stratifié optimal est de la forme :

$$\bar{y}_S \pm 1.96 \times 21$$
$$\bar{y}_S \pm 41.2$$

L'effet de sondage est de 0.656 (= 39.1 / 59.6) pour le sondage stratifié proportionnel et de 0.352 (= 21 / 59.6) pour le sondage stratifié optimal.

C'est le coefficient de réduction appliqué à la longueur de l'intervalle de confiance du sondage aléatoire simple pour obtenir la longueur de l'intervalle de confiance du sondage considéré.