République algérienne démocratique et populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

Université Abou-bekr Belkaid –Tlemcen

Faculté des sciences humaines et sociales

Département des sciences sociales

COURS 6 et 7

Destiné aux étudiants en 2ème année (Semestre4)

Filière : Démographie

LES TECHNIQUES DE SONDAGE

Responsable du module : Mme .MORTAD Nedjlaà (M.C.A)

Année universitaire : 2019 - 2020

Cours 6

Estimation d'une proportion

Estimation ponctuelle d'une proportion

1- Estimation d'une proportion

1.2- Estimation ponctuelle d'une proportion

On considère un caractère (ou attribut) A sur une population dont la proportion p est inconnue. On suppose que l'on a prélevé un échantillon de taille n (tirage avec remise) sur lequel on a calculé la proportion p_e d'individus ayant le caractère A (Figure 9).

F est la variable aléatoire correspondant à la proportion du caractère A dans un échantillon de taille n pris au hasard. Alors F suit approximativement une loi normale :

$$F \backsim N\left(p\,;\,\sigma_p\right) \text{ où } \sigma_p = \sqrt{\frac{p(1-p)}{n}}$$

Une estimation ponctuelle $\hat{p}$ de la proportion p de l'attribut A dans la population est : $\hat{p} = p_e$

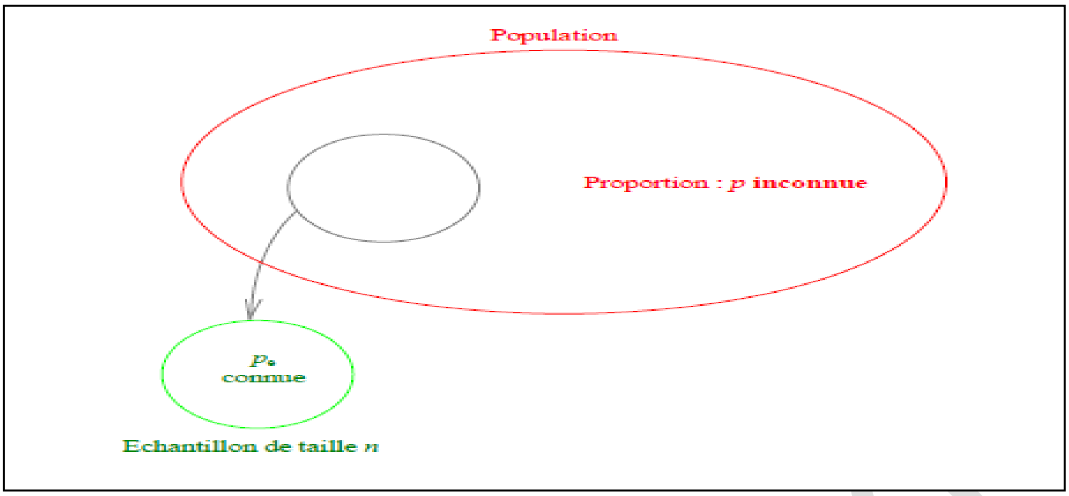

Figure 9 : Estimation ponctuelle de la proportion[1]

Une estimation ponctuelle $\widehat{\sigma_p}$ de l'écart type σ_p est selon le cas :

$$\widehat{\sigma_p} = \sqrt{\frac{p_e\,(1-P_e)}{n}}\sqrt{\frac{n}{n-1}} \quad \text{si } n \leq 30$$

$$\widehat{\sigma_p} = \sqrt{\frac{p_e\,(1-P_e)}{n}} \qquad \text{si } n > 30$$

Exemple 1 :

A quelques jours d'une élection, un candidat fait effectuer un sondage. Sur les 150 personnes interrogées, 45 se disent prêtes à voter pour lui aux prochaines élections.

1-Estimer ponctuellement la proportion d'individus prêts à voter pour ce candidat.

2-Estimer ponctuellement l'écart type de la proportion.

[1] - -COSTANTINI gille, Echantillonnage-estimation, statistiques inférentielles, BTS 2ème année , p5 https://docplayer.fr/13457270-Partie-a-echantillonnage.html

Solution :

$p_e = \frac{45}{150} = 0{,}3$

Donc : $\hat{p} = p_e = 0{,}3$

$\widehat{\sigma_p} = \sqrt{\frac{p_{e\,(1-P_e)}}{n}} = \sqrt{\frac{0{,}3\,.0{,}7}{150}} = 0{,}037$

Cours 7

Estimation d'une proportion

Estimation d'une proportion par intervalle de confiance

1- Estimation d'une proportion par intervalle de confiance

La proportion d'éléments (individus) possédant un caractère qualitatif dans la population, ayant un niveau de confiance 100(1-α) % de contenir la vraie valeur de p est en autant que $n\hat{p} \geq 5$ et $n(1-\hat{p}) \geq 5$, est donnée par l'intervalle ce confiance suivant :

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

Ou $\hat{p}$ représente la valeur observée de $\hat{p}$ (estimateur ponctuel de p) sur l'échantillon de taille n et $z_{\alpha/2}$ est la valeur de la variable normale centrée réduite.

Remarque :

Cet intervalle de confiance n'est valable que sous les conditions d'application mentionnées précédemment soit : $n\hat{p} \geq 5$ et $n(1-\hat{p}) \geq 5$[2].

Exemple2[3] *:*

Dans une municipalité, on a effectué un sondage pour connaitre l'opinion des contribuables sur un nouveau règlement d'emprunt. D'une liste informatisée de 6000 payeurs de taxes, on a prélevé, par tirage au sort, 150 noms. Sur ces 150, 45 étaient en faveur du nouveau règlement d'emprunt. Déterminer un intervalle de confiance pour p, la proportion vraie de

[2] - BAILLARGEON gérald , Probabilités, statistique et techniques de régression, ed SMG, 1989, P255.
[3] - idem, p255.

contribuables de cette municipalité qui sont en faveur du nouveau règlement d'emprunt, avec les niveaux de confiance de 90%, de 95% et de 99%.

Solution :

L'estimation ponctuelle de p est $\hat{p} = \frac{45}{150} = 0{,}30$

L'écart type de la proportion d'échantillon est :

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0{,}30)(0{,}70)}{150}} = 0{,}0374$$

Les intervalles de confiances obtenus pour les différents niveaux de confiance sont :

$Ic_{90\%} = [0{,}30 - 1{,}645.0{,}0374; 0{,}30 + 1{,}645.0{,}0374\ ]$

$Ic_{90\%} = [0{,}2385; 0{,}3515\ ]$

$Ic_{95\%} = [0{,}30 - 1{,}96.0{,}0374; 0{,}30 + 1{,}96.0{,}0374\ ]$

$Ic_{95\%} = [0{,}2267; 0{,}3733\ ]$

$Ic_{99\%} = [0{,}30 - 2{,}58.0{,}0374; 0{,}30 + 2{,}58.0{,}0374\ ]$

$Ic_{99\%} = [0{,}2035; 0{,}3965\ ]$

Le calcul de ces intervalles de confiance est valable en autant que : $n\hat{p} \geq 5$ et $n(1-\hat{p}) \geq 5$. Ces conditions d'application sont largement vérifiées : puisque $n\hat{p} = (150)(0{,}30) = 45$ et $n(1-\hat{p}) = (150)(0{,}70) = 105$.

Remarques[4]

-Dans le calcul de tout intervalle de confiance, plus le niveau de confiance est élevé, plus l'amplitude de l'intervalle est grande.

[4] - BAILLARGEON gérald , Probabilités, statistique et techniques de régression, ed SMG, 1989, p 256.

-Si l'échantillonnage s'effectue sans remise, à partir d'une population finie de taille N, on doit apporter une correction à $\sigma(\hat{p})$, dans ce cas :

$$\sigma(\hat{p})=\sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}$$

Pour considérer que ce facteur de correction est négligeable : $\frac{n}{N} \leq 0{,}10$.

-Dans le cas d'un tirage sans remise, l'écart type de la statistique concernée est donc plus faible que celui obtenu lors d'un tirage avec remise (meilleur précision).

2- Marge d'erreur associée à l'estimation de p et taille d'échantillon requise

Comme la valeur observée $\hat{p}$ est prédisposée à des fluctuations d'échantillonnage, il y a toujours un écart entre la valeur observé $\hat{p}$ et la valeur réelle p. Cet écart constitue la marge d'erreur dans l'estimation de p.

On l'appelle aussi : la précision du sondage.

Lorsque $\hat{p}$ est utilisée comme estimation de p, alors, pour un niveau de confiance 100 (1-α)%, la marge d'erreur E sera au plus égale à:

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

Par la suite, on peut déterminer la taille minimale de l'échantillon requise pour une marge d'erreur fixée.

Ainsi, pour une marge d'erreur désirée E et un niveau de confiance 100(1-α)% ; la taille d'échantillon minimale requise est :

$$n = \frac{Z^2_{\alpha/2}\,\hat{p}(1-\hat{p})}{E^2}$$

Exemple

Un groupe d'étudiants effectuent un sondage auprès de la population étudiante pour estimer le pourcentage des fumeurs. La population étudiante compte 8000.

Déterminer la taille d'échantillon requise pour assurer une marge d'erreur d'excédant pas 5%, avec un niveau de confiance de 95%, si l'enquête indique que 32% des étudiants fument.

Solution :

Marge d'erreur = 5%

Niveau de confiance : 95% , $Z_{\alpha/2}$= 1,96

$\hat{p}$= 0,32

Alors, $n = \frac{(1{,}96)^2\,(0{,}32)(0{,}68)}{(0{,}05)^2} = 334{,}37 \approx 335$