

LOI DE PROBABILITE DE LA PROPORTION D'ECHANTILLON

GASMI Abdelkader

November 4, 2024

POSITION DU PROBLEME

De même que pour la moyenne et pour la variance, si l'on considère une population infinie et que l'on y prélève un échantillon aléatoire et simple d'effectif n , on désigne par X_n le nombre d'individus possédant, dans chaque échantillon, le caractère étudié.

$$f_{X_n} = \frac{X_n}{n}$$

est la fréquence ou proportion des individus possédant, dans l'échantillon, le caractère étudié. On désigne par p la proportion des individus possédant, dans la population, le caractère étudié.

Les échantillons successifs possèdent les fréquences :

$$f_{X_n} = \frac{X_n}{n}, \quad f'_{X'_n} = \frac{X'_n}{n}, \quad f''_{X''_n} = \frac{X''_n}{n}$$

Ces fréquences peuvent être considérées comme des valeurs observées d'une même variable aléatoire :

$$f_{X_n} = \frac{X_n}{n}$$

La variable aléatoire F_n possède une distribution de probabilité, dite distribution d'échantillonnage de la proportion. On peut donc calculer l'espérance et la variance de cette distribution, en supposant que l'échantillon est aléatoire et simple.

LOI DE PROBABILITÉ SELON L'ÉCHANTILLON

Échantillon non exhaustif

$$E(F_n) = E\left(\frac{X_n}{n}\right) = \frac{1}{n}E(X_n) = \frac{1}{n}np = p$$

$$V(F_n) = V\left(\frac{X_n}{n}\right) = \frac{1}{n^2}V(X_n) = \frac{1}{n^2}npq = \frac{pq}{n}$$

L'écart type σ_{F_n} est appelé erreur standard de la fréquence d'un échantillon aléatoire et simple. En ce qui concerne la forme de cette distribution, on peut affirmer que la distribution de la proportion suit une loi normale de moyenne p et d'écart type

$$\sigma_{F_n} = \sqrt{\frac{pq}{n}}$$

à condition que la taille de l'échantillon soit supérieure ou égale à 30 ($n \geq 30$) et le produit $np \geq 5$.

Échantillon exhaustif

Dans le cas d'une population finie d'effectif N , au sein de laquelle est prélevé, sans remise, un échantillon aléatoire et simple d'effectif n , l'espérance mathématique et la variance de la fréquence sont :

$$E(F_n) = E\left(\frac{X_n}{n}\right) = \frac{1}{n}E(X_n) = \frac{1}{n}np = p$$

$$V(F_n) = V\left(\frac{X_n}{n}\right) = \frac{1}{n^2}V(X_n) = \frac{1}{n^2}\frac{N-1}{N}npq = \frac{N-1}{N}\frac{pq}{n}$$

L'erreur standard est alors :

$$\sigma_{F_n} = \sqrt{\frac{pq}{N}\frac{N-n}{N}}$$

En ce qui concerne la forme de cette distribution, on peut affirmer que la distribution de la proportion suit une loi normale de moyenne p et d'écart type

$$\sigma_{F_n} = \sqrt{\frac{pq}{N} \frac{N-n}{N}}$$

à condition que la taille de l'échantillon soit supérieure ou égale à 30 ($n \geq 30$) et le produit $np \geq 5$.

Exemple

Lors d'élections, les résultats ont montré qu'un des candidats a obtenu 46% des voix. Déterminer la probabilité pour que le vote donne une majorité de voix en faveur de ce candidat pour un échantillon de :

- 1. 200 personnes choisies au hasard parmi le corps électoral.*
- 2. 1000 personnes choisies au hasard parmi le corps électoral.*

Soit F la variable aléatoire qui désigne la fréquence des électeurs qui voteront pour ce candidat. La distribution d'échantillonnage de F suit une loi normale de moyenne p et d'écart type $\sigma_F = \sqrt{\frac{pq}{n}}$ à condition que la taille de l'échantillon soit supérieure ou égale à 30 ($n \geq 30$) et le produit $np \geq 5$.

- 200 personnes choisies au hasard parmi le corps électoral.

$$E(F) = 0,46$$

$$V(F) = \frac{200 \cdot 0,46 \cdot 0,54}{200} = 0,0012$$

$$\sigma_F = \sqrt{0,0012} = 0,035$$

Puisque $n = 200 > 30$ et le produit $np = 200 \cdot 0,46 = 92 > 5$, la distribution d'échantillonnage de F suit donc une loi normale de moyenne 0,46 et d'écart type 0,035.

Le vote donnera une majorité de voix en faveur de ce candidat si la fréquence des électeurs qui voteront pour ce candidat est supérieure ou égale à 50 %.

$$p(F \geq 0,5) = 1 - p(F < 0,5) = p\left(Z < \frac{0,5 - 0,46}{0,035}\right)$$

$$p(F \geq 0,5) = 1 - \Phi(1,14) = 1 - 0,8729 = 0,1271 = 12,71\%$$

► 1000 personnes choisies au hasard parmi le corps électoral

$$E(F) = 0,46$$

$$V(F) = \frac{1000 \cdot 0,46 \cdot 0,54}{1000} = 0,0002484$$

$$\sigma_F = \sqrt{0,0002484} = 0,016$$

Puisque $n = 1000 > 30$ et le produit $np = 1000 \cdot 0,46 = 460 > 5$, la distribution d'échantillonnage de F suit donc une loi normale de moyenne 0,46 et d'écart type 0,016.

Le vote donnera une majorité de voix en faveur de ce candidat si la fréquence des électeurs qui voteront pour ce candidat est supérieure ou égale à 50 %.