

Probabilités et Approximation Normale d'un Sondage

Un éditeur réalise un sondage sur la présence de pages TICE dans ses futurs manuels : chaque sondé peut répondre par « ***pour*** » ou bien par « ***contre*** », et on note ***p*** la probabilité qu'un sondé réponde « ***pour*** ».

a. Justifier que la variable aléatoire ***R*** qui, à un sondé pris au hasard, associe **1** s'il réponde « ***pour*** », et **0** sinon, suit la loi de Bernoulli de paramètre ***p***.

On définit la variable aléatoire ***R*** par :

- $R = 1$ si le sondé répond « pour ».
- $R = 0$ si le sondé répond « contre ».

Chaque sondé choisit **indépendamment** sa réponse, et la probabilité qu'un sondé réponde « pour » est ***p***. Il y a donc deux issues possibles :

- $R = 1$ avec probabilité ***p***,
- $R = 0$ avec probabilité $1-p$.

Ainsi, ***R*** suit une loi de **Bernoulli** de paramètre ***p*** : $P(R = 1) = p$, $P(R = 0) = 1 - p$.

b. On note $(R_1, \dots, R_{800})$ un échantillon de taille 800 de la loi de la variable aléatoire ***R*** et on suppose qu'il modélise les réponses apportées par 800 personnes au sondage. Déterminer la loi de probabilité de la variable aléatoire $S = R_1 + \dots + R_{800}$.

La **loi de probabilité** de la variable aléatoire ***S***, qui est la somme des 800 variables de Bernoulli indépendantes $(R_1, \dots, R_{800})$, est une **loi binomiale**.

Loi de probabilité de *S*

- ***S*** suit une loi binomiale $B(800, p)$.
- La fonction de probabilité de cette loi est donnée par :

$$P(S = k) = \binom{800}{k} \times p^k \times (1 - p)^{800-k} \text{ avec où } \binom{800}{k} \text{ est le coefficient binomial : } \binom{800}{k} = \frac{800!}{k! (800 - k)!}$$

où :

- $\binom{800}{k}$ est le coefficient binomial qui compte le nombre de façons de choisir k succès parmi 800 essais.
- p^k est la probabilité d'avoir exactement k réponses « pour ».
- $(1 - p)^{800-k}$ est la probabilité d'avoir $800 - k$ réponses « contre ».

Cette loi décrit donc la répartition des réponses favorables parmi les 800 sondés.

c. Le sondage a conclu que 61% des sondés étaient favorables à la présence de pages TICE.

En supposant que $p = 61\%$, quelle est la probabilité que, sur un sondage de 800 personnes, la proportion de « pour » soit comprise entre 56% et 75% ?

On note :

- ***S*** : le **nombre de sondés** ayant répondu « ***pour*** », qui suit une **loi binomiale** : $S \sim B(800, 0,61)$.
- $\hat{p}$: la **proportion observée** de réponses « ***pour*** » dans l'échantillon, définie par : $\hat{p} = \frac{S}{800}$

$\hat{p}$ est une **variable aléatoire** qui donne une estimation de la proportion réelle ***p***.

Nous devons donc calculer : $P(0,56 \leq \hat{p} \leq 0,75)$.

En multipliant cette inégalité par 800, cela revient à calculer : $P(0,56 \times 800 \leq S \leq 0,75 \times 800) = P(448 \leq S \leq 600)$

Pour un grand effectif ***n***, la loi binomiale peut être approximée par une **loi normale** : $S \approx N(\mu, \sigma^2)$

avec :

- **Espérance** : $\mu = n \times p = 800 \times 0,61 = 488$
- **Écart-type** : $\sigma = \sqrt{n \times p(1 - p)} = \sqrt{800 \times 0,61(1 - 0,61)} \approx 13,98$

On reformule donc la probabilité : $P(448 \leq S \leq 600)$

On transforme S en une variable centrée réduite Z, définie par :

$$Z = \frac{S - \mu}{\sigma}$$

Les bornes deviennent :

$$Z_1 = \frac{448 - 488}{13,98} \approx -2,86 \text{ et } Z_2 = \frac{600 - 488}{13,98} \approx 8,01$$

Ainsi : $P(448 \leq S \leq 600) = P(-2,86 \leq Z \leq 8,01) = 1 - 0,0021 = 0,9981 = 99,81\%$

Le sondage a conclu que 61% des sondés étaient favorables à la présence de pages TICE.

En supposant que $p = 61\%$, quelle est la probabilité que, sur un sondage de 800 personnes, la proportion de « pour » soit supérieure à 85% ?

On note :

- S : le **nombre de sondés** ayant répondu « **pour** », qui suit une **loi binomiale** : $S \sim B(800, 0,61)$.
- $\hat{p}$: la **proportion observée** de réponses « **pour** » dans l'échantillon, définie par : $\hat{p} = \frac{S}{800}$

$\hat{p}$ est une **variable aléatoire** qui donne une estimation de la proportion réelle p.

On doit calculer $P(\hat{p} \geq 0,85)$

En multipliant cette inégalité par 800, cela revient à calculer :

$$P(S \geq 0,85 \times 800) = P(S \geq 680)$$

Pour un grand effectif n, la loi binomiale peut être approximée par une **loi normale** : $S \approx N(\mu, \sigma^2)$

avec :

- **Espérance** : $\mu = n \times p = 800 \times 0,61 = 488$
- **Écart-type** : $\sigma = \sqrt{n \times p(1 - p)} = \sqrt{800 \times 0,61(1 - 0,61)} \approx 13,98$

On reformule donc la probabilité : $P(S \geq 680)$

On transforme S en une variable centrée réduite Z, définie par :

$$Z = \frac{S - \mu}{\sigma} \sim N(0,1)$$

La borne devient :

$$Z = \frac{680 - 488}{13,98} \approx 13,75$$

Avec une table on trouve

$$P(Z \leq 13,65) \approx 1$$

Ainsi :

$$P(S \geq 680) = 1 - P(Z \leq 13,65) = 0$$

La **probabilité** que la proportion de réponses « pour » soit **supérieure à 85%** est **pratiquement nulle**.

Cela signifie qu'il est **quasiment impossible** qu'un sondage de 800 personnes donne une proportion aussi élevée de « pour ».