Dans un pays tropical, un recensement a montré que 10% des plages étaient atteintes par des algues toxiques. Après avoir modifié le processus de rejets chimiques dans ce pays, on veut établir, puis mettre en œuvre, une procédure statistique permettant de décider, au seuil de 5% si le processus de rejet ainsi modifié a, ou non, un impact significatif sur la fréquence d'apparition de ces algues.

1. Détermination de la taille d'un échantillon

On fait l'hypothèse que le processus de rejet modifié n'a pas d'impact significatif, c'est à dire que la proportion de plages touchées dans le pays est $p = 0,1$.

On examine un échantillon de n plages au hasard (le pays possède suffisamment de plages pour considérer que cet échantillon résulte d'un tirage aléatoire avec remise). On désigne par X_n la variable aléatoire qui, à un échantillon de ce type, associe le nombre de plages atteinte par les algues et par $F_n = \frac{X_n}{n}$ la variable aléatoire associée à la " fréquence " de ce nombre de plages.

- Quelle est la loi de la variable X_n ? Donner ses paramètres
- Sous l'hypothèse $p = 0,1$, écrire l'intervalle de fluctuation asymptotique I_n à 95% de la variable aléatoire F_n et déterminer la valeur minimale de n permettant l'utilisation de cet intervalle.
- Déterminer la valeur minimale n_0 de n pour que le rayon de l'intervalle I_n soit inférieur à 0,05.

2. Construction d'une règle de décision

Énoncer la règle de décision permettant de rejeter, ou non, au seuil de décision de 5%, l'hypothèse selon laquelle 10% des plages du pays sont touchées par ce type d'algues sur un échantillon de taille n_0

3. Prise de décision

Sur un échantillon de n_0 plages, on constate que 11 plages seulement sont atteintes. Peut-on, au seuil de décision de 5%, rejeter l'hypothèse précédente ?

CORRECTION

1. Détermination de la taille d'un échantillon

- On a une succession de n expériences aléatoires identiques et indépendantes, chacune d'elles a deux issues :
réussite : la plage est atteinte par les algues ($p = 0,1$)
échec : la plage n'est pas atteinte par les algues ($q = 1 - p = 0,9$)
donc la variable aléatoire X_n suit une loi binomiale de paramètres $(n ; 0,1)$.

$$b. \quad \text{Sous l'hypothèse } p = 0,1 : I_n = \left[p - 1,96 \sqrt{\frac{p(1-p)}{n}} ; p + 1,96 \sqrt{\frac{p(1-p)}{n}} \right]$$

$$I_n = \left[0,1 - 1,96 \sqrt{\frac{0,1 \times 0,9}{n}} ; 0,1 + 1,96 \sqrt{\frac{0,1 \times 0,9}{n}} \right] \Leftrightarrow I_n = \left[0,1 - 1,96 \times \frac{0,3}{\sqrt{n}} ; 0,1 + 1,96 \times \frac{0,3}{\sqrt{n}} \right]$$

$$I_n = \left[0,1 - \frac{0,588}{\sqrt{n}} ; 0,1 + \frac{0,588}{\sqrt{n}} \right]$$

$$\text{On peut utiliser } I_n \text{ si } \begin{cases} n \geq 30 \\ n p \geq 5 \\ n(1-p) \geq 5 \end{cases} \text{ soit } \begin{cases} n \geq 30 \\ n \geq \frac{5}{0,1} \\ n \geq \frac{5}{0,9} \end{cases} \text{ donc } \begin{cases} n \geq 30 \\ n \geq 50 \\ n \geq \frac{50}{9} \end{cases} \text{ soit } n \geq 50 \text{ donc la valeur minimale de } n \text{ permettant l'utilisation de}$$

cet intervalle est 50

$$c. \quad \text{Le rayon de l'intervalle est } \frac{0,588}{\sqrt{n}} \text{ donc } \frac{0,588}{\sqrt{n}} < 0,05 \Leftrightarrow \sqrt{n} > \frac{0,588}{0,05} \Leftrightarrow \sqrt{n} \geq 11,76 \Leftrightarrow n \geq 138,2976 \Leftrightarrow n \geq 139$$

la valeur minimale n_0 de n pour que le rayon de l'intervalle I_n soit inférieur à 0,05 est $n_0 = 139$

2. Construction d'une règle de décision

Choisissons un échantillon de 139 plages et mesurons le nombre de plages du pays sont touchées par ce type d'algues.

$$I_{139} = [0,05 ; 0,15]$$

Si la fréquence f observée dans l'échantillon appartient à $[0,05 ; 0,15]$ au seuil de décision de 5%, on accepte l'hypothèse faite que 10% des plages du pays sont touchées par ce type d'algues.

Si la fréquence f observée dans l'échantillon n'appartient pas à $[0,05 ; 0,15]$ au seuil de 5 %, on rejette l'hypothèse faite que 10% des plages du pays sont touchées par ce type d'algues.

3. Prise de décision

Sur 139 plages, 11 sont atteintes donc $f = \frac{11}{139}$ soit $f \approx 0,079$

$f \in I_{139}$ donc au seuil de décision de 5%, on accepte l'hypothèse faite que 10% des plages du pays sont touchées par ce type d'algues.