Tests statistiques.

7.A.3 1/ On prend une région de rejet de la forme :
$W_n = \{\overline{X}_n \leq a_n\}$ où a_n est un réel que l'on va choisir.

- Sous H_0 ie sous $\mathbb{P}_{p_0}$, on a par limite centrale :

$$Y_n = \sqrt{\frac{n}{p_0(1-p_0)}}(\overline{X}_n - p_0) \xrightarrow[n\infty]{loi} \mathcal{N}(0,1)$$

Ainsi $\forall \beta \in \mathbb{R}$,

$$\mathbb{P}_{p_0}(Y_n \leq \beta) \xrightarrow[n\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta} e^{-\frac{x^2}{2}} dx \quad \begin{pmatrix} \text{lm} \\ 3.3.2 \end{pmatrix}$$

Donc avec $\beta = \Phi_\alpha$, on a $\mathbb{P}_{p_0}(Y_n \leq \beta) \xrightarrow[n\infty]{} \alpha$

Or $\{Y_n \leq \Phi_\alpha\} = \left\{\sqrt{\frac{n}{p_0(1-p_0)}}(\overline{X}_n - p_0) \leq \Phi_\alpha\right\}$

$$= \left\{\overline{X}_n \leq \sqrt{\frac{p_0(1-p_0)}{n}}\Phi_\alpha + p_0\right\}.$$

Bilan : Avec $a_n = \sqrt{\frac{p_0(1-p_0)}{n}}\Phi_\alpha + p_0$, on a

$$\mathbb{P}_{p_0}(W_n) \xrightarrow[n\infty]{} \alpha.$$

D'où : $\boxed{\alpha = \sup_{\substack{p\in H_0 \\ (\text{ie } p=p_0)}} \lim_{n\infty} \sup_p \mathbb{P}_p(W_n)}$

Le test est donc asymptotique de niveau α.

- Par loi forte, $\overline{X}_n \xrightarrow[n\infty]{ps} p < p_0$ sous H_1 (ie $p > p_0$)

Ainsi : $\mathbb{1}_{\{\overline{X}_n \leq a_n\}} \xrightarrow[n\infty]{ps} \mathbb{1}_{\{p \leq p_0\}} = 1$ sous H_1

(à noter que $a_n \xrightarrow[n\infty]{} p_0$).

D'où via cvg dominée (domination évidente par 1),

$$\mathbb{E}_p(\mathbb{1}_{\{\overline{X}_n \leq a_n\}}) \xrightarrow[n\infty]{} \mathbb{E}_p(1) = 1$$

ie : $\boxed{\mathbb{P}_p(\overline{X}_n \leq a_n) = \mathbb{P}_p(W_n) \xrightarrow[n\infty]{} 1 \quad \forall p \in H_1.}$

Le test est donc consistant.

2/ On calcule la p-valeur.

H_0 étant simple et comme $\omega_n = \{\overline{X}_n \leq a_n\}$, on a :

$$p\text{-valeur} = \mathbb{P}_{p_0}(\overline{X}_n \leq \overline{x}_n) \text{ avec } \overline{x}_n = 16\%.$$

On utilise l'approximation gaussienne (qu'on est légitime d'utiliser par (*)) :

$$p\text{-valeur} \simeq \mathbb{P}\left(\sqrt{\frac{p_0(1-p_0)}{n}}\,G + p_0 \leq \overline{x}_n\right) \text{ où } G \sim \mathcal{N}(0,1)$$

$$= \mathbb{P}\left(G \leq \sqrt{\frac{n}{p_0(1-p_0)}}\left(\overline{x}_n - p_0\right)\right)$$

$$\boxed{p\text{-valeur} = F_G\left(\sqrt{\frac{n}{p_0(1-p_0)}}\left(\overline{x}_n - p_0\right)\right)}$$

A.N $\sqrt{\frac{n}{p_0(1-p_0)}}\left(\overline{x}_n - p_0\right) \simeq -0{,}52$

$$\boxed{p\text{-valeur} \simeq 0{,}30 \quad > \quad \alpha = 5\%}$$

Ainsi : On ne rejette pas H_0. On ne peut pas conclure que le vaccin est efficace.

3/ $p = 15\%$ en vérité

On voit la p-valeur comme une v.a :

$$p\text{-valeur}(X_1, \dots, X_n) = F_G\left(\sqrt{\frac{n}{p_0(1-p_0)}}\left(\overline{X}_n - p_0\right)\right)$$

La proba recherchée est :

$$p_n = \mathbb{P}_p\left(p\text{-valeur}(X_1, \dots, X_n) \leq \alpha\right)$$

$$= \mathbb{P}_p\left(F_G\left(\sqrt{\frac{n}{p_0(1-p_0)}}\left(\overline{X}_n - p_0\right)\right) \leq \alpha\right)$$

$$= \mathbb{P}_p\left(\sqrt{\frac{n}{p_0(1-p_0)}}\left(\overline{X}_n - p_0\right) \leq F_G^{-1}(\alpha)\right)$$

$$= \mathbb{P}_p\left(\overline{X}_n \leq p_0 + \sqrt{\frac{p_0(1-p_0)}{n}}\,F_G^{-1}(\alpha)\right)$$

$$= \mathbb{P}\left(S_n \leq np_0 + \sqrt{np_0(1-p_0)}\,F_G^{-1}(\alpha)\right)$$

où $S_n \sim \mathcal{B}(n, p)$

Bilan : La proba recherchée est donnée par :

$$\boxed{p_n = F_{S_n}\left(np_0 + \sqrt{np_0(1-p_0)}\, F_G^{-1}(\alpha)\right)}$$

où F_{S_n} est la fonction de répartition de S_n.

A.N • $F_G^{-1}(\alpha) \simeq -1{,}64$

• $np_0 + \sqrt{np_0(1-p_0)}\, F_G^{-1}(\alpha) \simeq 11{,}7$

• $\boxed{p_n \simeq 0{,}16}$

4/ **Méthode** : En adaptant p en fonction de si le vaccin induit une diminution des infectés égale à $\frac{1}{6}$ ou $\frac{1}{3}$, il s'agit simplement de calculer la quantité $n_{min} = \inf\{n \in \mathbb{N}^* \mid p_n \geq 80\%\}$

A.N • Cas où $p = \left(1 - \frac{1}{6}\right) p_0$

$\boxed{n_{min} = 942.}$

• Cas où $p = \left(1 - \frac{1}{3}\right) p_0$

$\boxed{n_{min} = 219.}$