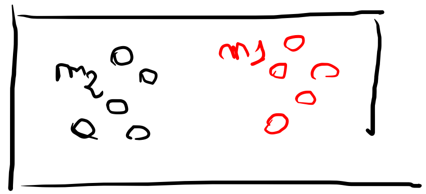
# Exercice 10




$$p = \frac{m_1}{m_1 + m_2}$$

Estimer p.

1) N tirages avec remise $\longrightarrow X_1, \dots, X_N$ i.i.d

$$X_i = \begin{cases} 1 & \text{ième tirage boule rouge} \\ 0 & \text{——— noire} \end{cases}$$

La prop de boules rouges $= \mathbb{P}(X_i = 1)$.

a) • Estimateur naturel de $p$ est $\bar{X}_N = \frac{1}{N} \sum_{i=1}^{N} X_i$

(car $X_i \sim$ Bernoulli$(p)$ $X_i$ iid $\mathbb{E}[X_1] = p$)
$\text{Var}(X_1) = p(1-p)$

• $\text{EQM}(\bar{X}_N, p) = \underbrace{\text{Var}(\bar{X}_N)}_{= \frac{1}{N} p(1-p)} + \underbrace{\text{Biais}^2(\bar{X}_N)}_{= 0}$

(b) · <u>Normalité asymptotique</u>

· $X_1, \dots, X_N$ iid

· $\mathbb{E}[X_1] = p$, $Var(X_1) = p(1-p)$

Par le TCL, $\sqrt{N}\,(\bar{X}_N - p) \xrightarrow[N \to +\infty]{\mathcal{L}} \mathcal{N}(0, p(1-p))$ (*)

· <u>Intervalle de confiance</u> : (*) implique :

$$\frac{\sqrt{N}\,(\bar{X}_N - p)}{\sqrt{p(1-p)}} \xrightarrow[N \to +\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

<u>Plug-in</u> On va s'intéresser à $\dfrac{\sqrt{N}\,(\bar{X}_N - p)}{\sqrt{\bar{X}_N(1-\bar{X}_N)}}$

$$\frac{\sqrt{N}\,(\bar{X}_N - p)}{\sqrt{\bar{X}_N(1-\bar{X}_N)}} = \underbrace{\frac{\sqrt{N}\,(\bar{X}_N - p)}{\sqrt{p(1-p)}}}_{\substack{\downarrow \mathcal{L} \\ \mathcal{N}(0,1)}} \; \underbrace{\frac{\sqrt{p(1-p)}}{\sqrt{\bar{X}_N\,(1-\bar{X}_N)}}}_{\substack{\downarrow \mathbb{P} \\ 1}}$$

car $\bar{X}_N \xrightarrow{\mathbb{P}} p$ (LGN) + théorème de cont.

Par Slutsky, $\dfrac{\sqrt{N}(\overline{X}_N - p)}{\sqrt{\overline{X}_N(1-\overline{X}_N)}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$.

ce qui implique

$$\lim_{N\to+\infty} \mathbb{P}\left(-q_{1-\alpha/2} \leq \frac{\sqrt{N}(\overline{X}_N - p)}{\sqrt{\overline{X}_N(1-\overline{X}_N)}} \leq q_{1-\alpha/2}\right) = 1-\alpha,$$

avec $q_{1-\alpha/2}$ le quantile d'ordre $1-\alpha/2$ de $\mathcal{N}(0,1)$.

$$\lim_{N\to+\infty} \mathbb{P}\left(\overline{X}_N - \frac{q_{1-\alpha/2}\sqrt{\overline{X}_N(1-\overline{X}_N)}}{\sqrt{N}} \leq p \leq \overline{X}_N + \frac{q_{1-\alpha/2}\sqrt{\overline{X}_N(1-\overline{X}_N)}}{\sqrt{N}}\right)$$

$= 1-\alpha$

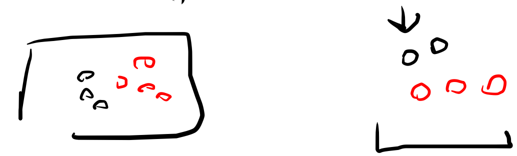
$$IC_{1-\alpha} = \left[\overline{X}_N \mp \frac{q_{1-\alpha/2}\sqrt{\overline{X}_N(1-\overline{X}_N)}}{\sqrt{N}}\right].$$

2) N tirages sans remise $\longrightarrow$ 1 exp / 5000 exp | $\rightarrow$ ~~i.i.d~~ exp.

Y : le nbre de boules rouges tirées.

$(N << m_1 + m_2)$.

N tirages $(=5)$



J'ai 3 boules rouges

a) $\mathbb{E}[Y] = Np$

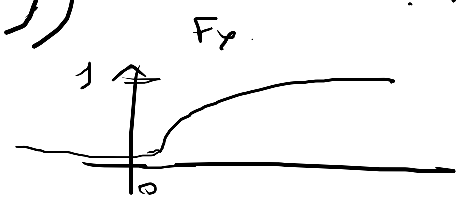
$$Var(Y) = Np(1-p)\frac{m_1+m_2-N}{m_1+m_2-1}.$$

↳ Loi hyper-géométrique (pas à savoir)

Y

continue

$\nearrow$ . $F_Y = \mathbb{P}(Y \leq x)$




↳ "Fonction de répartition" $k \in \mathbb{N}$.

$\mathbb{P}(Y = k) =$

Y discret

. $\mathbb{P}(Y = k)$, $\forall k \in \mathbb{N}$

$$\mathbb{P}(Y=k) = \frac{\left[\begin{array}{l}\text{nombre de façon de choisir}\\ k \text{ boules rouges parmi } m_1\\ \text{boules rouges}\end{array}\right]\left[\begin{array}{l}N-k \text{ (nombres de boules noires)}\\ \text{parmi } m_2 \text{ boules noires}\end{array}\right]}{\left[\begin{array}{l}\text{nombre de façons de choisir } N \text{ boules}\\ \text{parmi } m_1+m_2 \text{ boules}\end{array}\right]}$$

$$= \frac{\binom{m_1}{k}\binom{m_2}{N-k}}{\binom{m_1+m_2}{N}}.$$

(b) • $p$ : proportion de boules rouges.

Comment estimer $p$ à partir de $Y$ (nombre de boules rouges) ?

• Un estimateur <u>naturel</u> : $\dfrac{Y}{N}$ ← nbre de boules rouges, ← nbre de boules totales

• $\underline{\mathbb{E}[Y] = Np}$. Un estimateur <u>non biaisé</u> est $\dfrac{Y}{N} = \hat{p}_N$.

$$\mathbb{E}[\hat{p}_N - p] = 0.$$
$$= \mathbb{E}[\hat{p}_N] - p$$

$$\mathbb{E}[\hat{p}_N] = \mathbb{E}\left[\frac{Y}{N}\right] = \frac{1}{N}\mathbb{E}[Y] = p.$$

$$EQM(\hat{p}_N, p) = \text{Biais}^2(\hat{p}_N) + Var(\hat{p}_N).$$

$$= Var(\hat{p}_N)$$

$$= Var\left(\frac{Y}{N}\right)$$

$$= \frac{1}{N^2} Var(Y) = \frac{1}{N} p(1-p) \frac{m_1 + m_2 - N}{m_1 + m_2 - 1}$$

Comparaison des estimateurs $\overline{X}_N$ (question 1) et de $\hat{p}_N$ :

$$EQM(\hat{p}_N, p) < EQM(\overline{X}_N, p) \quad \text{car} \quad \frac{EQM(\hat{p}_N, p)}{EQM(\overline{X}_N, p)}$$

$$= \frac{m_1 + m_2 - N}{m_1 + m_2 - 1} < 1.$$

On choisit donc $\hat{p}_N$.

(c) IC (pas classique)

$p$

$A_m(\hat{p}_N)$ —— $B_m(\hat{p}_N)$

Quel écart $|p - \hat{p}_N|$ ??

$$\mathbb{P}\left(p \in [A_m, B_m]\right) = 1-\alpha.$$

$A_m, B_m$ quantités connues (dépendent pas de $p$. mais dépendent de $\hat{p}_N$).

On cherche $\varepsilon$ tq : $\mathbb{P}(|\hat{p}_N - p| \geq \varepsilon) = \alpha$. $\left(\mathbb{P}(|\hat{p}_N - p| \leq \varepsilon) = 1-\alpha\right)$.

$$\mathbb{P}(|\hat{p}_N - p| \geq \varepsilon) \leq \frac{\mathrm{Var}(\hat{p}_N)}{\varepsilon^2} = \frac{1}{N\varepsilon^2} \frac{(m_1+m_2-N)\,p(1-p)}{(m_1+m_2-1)} = \alpha.$$

Inégalité de Markov

On choisit $\varepsilon$ qui minimise la borne sup de $\mathbb{P}(|\hat{p}_N - p| \geq \varepsilon)$.

$$\varepsilon = \sqrt{\frac{(m_1+m_2-N)(p(1-p))}{N\alpha\,(m_1+m_2-1)}}$$

$$IC_{1-\alpha} = \left[\hat{p}_N \pm \varepsilon\right]$$

# Exercice 4

$X$ densité $f(x) = \frac{2}{\theta^2} x \mathbb{1}_{[0,\theta]}(x)$ avec $\theta > 0$.

1) $\mathbb{E}[X] = \int_{\mathbb{R}} x f_\theta(x)\, dx$.

$$= \int_0^\theta \frac{2x^2}{\theta^2}\, dx = \left[\frac{2}{3\theta^2} x^3\right]_0^\theta = \frac{2}{3}\theta$$

$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

$$\mathbb{E}[X^2] = \int_{\mathbb{R}} x^2 f_\theta(x)\, dx = \int_0^\theta \frac{2x^3}{\theta^2}\, dx = \left[\frac{2}{4\theta^2} x^4\right]_0^\theta = \frac{1}{2}\theta^2$$

$$\mathrm{Var}(X) = \frac{1}{2}\theta^2 - \left(\frac{2}{3}\theta\right)^2 = \frac{1}{18}\theta^2.$$

2) Un estimateur de $\mathbb{E}[X] = \frac{2}{3}\theta$ est $\overline{X}_n$.

Donc par la méthode des moments, l'estimateur $\hat{\theta}_n$ est $\frac{3}{2}\overline{X}_n$.

3) <u>consistant</u> $\overline{X}_n \xrightarrow[n \to +\infty]{\mathbb{P}} \frac{2}{3}\theta$ (LGN)

$\hat{\theta}_n \xrightarrow[n \to +\infty]{\mathbb{P}} \theta$ (Th de continuité)

<u>asymptotiquement normal</u> $\sqrt{n}\left(\overline{X}_n - \frac{2}{3}\theta\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{18}\theta^2\right)$ (TCL)

$g(x) = \frac{3}{2}x$. $g'(x) = \frac{3}{2}$. (continue)

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \underbrace{\tfrac{1}{8}\theta^2}_{= g'(\theta)^2 \frac{1}{18}\theta^2}\Big) \quad (\Delta \text{ méthode})$$

4) IC asymptotique pour $\theta$.

- On a : $\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{8}\theta^2\right)$

- Cela implique $\dfrac{\overbrace{2\sqrt{2n}}^{=\sqrt{8}=2\sqrt{2}}\,(\hat{\theta}_n - \theta)}{\theta} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$. (Exo 3)

$\downarrow$

2 stratégies

1ère stratégie plug-in    On s'intéresse à $\dfrac{2\sqrt{2n}\,(\hat{\theta}_n-\theta)}{\hat{\theta}_n}$.

$$\frac{2\sqrt{2n}\,(\hat{\theta}_n-\theta)}{\hat{\theta}_n} = \underbrace{\frac{2\sqrt{2n}\,(\hat{\theta}_n-\theta)}{\theta}}_{\downarrow \mathcal{L}\ \mathcal{N}(0,1)} \underbrace{\frac{\theta}{\hat{\theta}_n}}_{\downarrow \mathbb{P}\ 1}$$

$\downarrow \mathbb{P}$ car $\hat{\theta}_n$ consistant. (+ théorème de continuité).

Par Slutsky, $\dfrac{2\sqrt{2n}\,(\hat{\theta}_n-\theta)}{\hat{\theta}_n} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$

Ce qui implique :

$$\lim_{n\to+\infty} \mathbb{P}\left( -q_{1-\alpha/2} \leq \frac{2\sqrt{2n}\,(\hat{\theta}_n - \theta)}{\hat{\theta}_n} \leq q_{1-\alpha/2} \right) = 1-\alpha .$$

$$\Leftrightarrow \lim_{n\to+\infty} \mathbb{P}\left( -\frac{\hat{\theta}_n q_{1-\alpha/2}}{2\sqrt{2n}} \leq \hat{\theta}_n - \theta \leq \frac{\hat{\theta}_n q_{1-\alpha/2}}{2\sqrt{2n}} \right) = 1-\alpha .$$

$$\Leftrightarrow \lim_{n\to+\infty} \mathbb{P}\left( -\hat{\theta}_n - \frac{\hat{\theta}_n q_{1-\alpha/2}}{2\sqrt{2n}} \leq -\theta \leq -\hat{\theta}_n + \frac{\hat{\theta}_n q_{1-\alpha/2}}{2\sqrt{2n}} \right) = 1-\alpha$$

$$\Leftrightarrow \lim_{n\to+\infty} \mathbb{P}\left( \hat{\theta}_n - \frac{q_{1-\alpha/2}\,\hat{\theta}_n}{2\sqrt{2n}} \leq \theta \leq \hat{\theta}_n + \frac{q_{1-\alpha/2}\,\hat{\theta}_n}{2\sqrt{2n}} \right) = 1-\alpha$$

$$IC_{1-\alpha} = \left[ \hat{\theta}_n \pm \frac{q_{1-\alpha/2}\,\hat{\theta}_n}{2\sqrt{2n}} \right].$$

2ème stratégie

$$2\sqrt{2n}\,\underbrace{\frac{(\hat{\theta}_n - \theta)}{\theta}}_{} \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

$$= 2\sqrt{2n}\left(\frac{\hat{\theta}_n}{\theta} - 1\right)$$

ce qui implique :

$$\lim_{n \to +\infty} \mathbb{P}\left(-q_{1-\alpha/2} \leq 2\sqrt{2n}\left(\frac{\hat{\theta}_n}{\theta} - 1\right) \leq q_{1-\alpha/2}\right) = 1-\alpha.$$

$$\Leftrightarrow \lim_{n \to +\infty} \mathbb{P}\left(-\frac{q_{1-\alpha/2}}{2\sqrt{2n}} + 1 \leq \frac{\hat{\theta}_n}{\theta} \leq \frac{q_{1-\alpha/2}}{2\sqrt{2n}} + 1\right) = 1-\alpha$$

$$\Leftrightarrow \lim_{n \to +\infty} \mathbb{P}\left(\frac{-q_{1-\alpha/2}}{2\sqrt{2n}\,\hat{\theta}_n} + \frac{1}{\hat{\theta}_n} \leq \frac{1}{\theta} \leq \frac{q_{1-\alpha/2}}{2\sqrt{2n}\,\hat{\theta}_n} + \frac{1}{\hat{\theta}_n}\right) = 1-\alpha$$

$$\Leftrightarrow \lim_{n \to +\infty} \mathbb{P}\left(\hat{\theta}_n\left(1 + \frac{q_{1-\alpha/2}}{2\sqrt{2n}}\right)^{-1} \leq \theta \leq \hat{\theta}_n\left(1 - \frac{q_{1-\alpha/2}}{2\sqrt{2n}}\right)^{-1}\right) = 1-\alpha$$

↳ Bornes $\frac{1}{\hat{\theta}_n} \pm \frac{q_{1-\alpha/2}}{2\sqrt{2n}\,\hat{\theta}_n} \geq 0$ p.s.

5) EMV

$$L_x(\theta) = \prod_{i=1}^{n} f_\theta(x_i)$$

$$= \prod_{i=1}^{n} \frac{2}{\theta^2} x_i \, \mathbb{1}_{[0,\theta]}(x_i)$$

Maximiser en $\theta$

→ $L_x$ prend 2 valeurs $L_x(\theta) = \begin{cases} 0 & \text{si } \exists i,\ x_i \notin [0,\theta] . \\ \prod_{i=1}^{n} \frac{2}{\theta^2} x_i , & \forall i,\ x_i \in [0,\theta] . \end{cases}$

→ $L_x$ sera maximale $\forall i,\ x_i \in [0,\theta] \iff \forall i,\ x_i \leq \theta$.

On pose : $\hat{\theta}^{MV} = \max_{i=1,\dots,n} x_i$.