# Introduction aux tests statistiques

**Chargés de cours**
V. Léger & F. Leblanc (resp. UE)

- Problème : On a une pièce de monnaie dans notre poche, et on voudrait savoir si c'est une fausse pièce ou non.
- On ne peut pas distinguer visuellement une fausse pièce d'une vraie, mais on sait qu'une fausse pièce a $p_1 = 60\%$ de chances de tomber sur Pile (contre $p_0 = 50\%$ pour une vraie pièce).
- **Question : Comment savoir si la pièce qu'on a est une fausse ?**
- *Stratégie : lancer la pièce un grand nombre de fois et prendre la décision de la garder ou non en fonction de la fréquence de Pile.*

- On lance la pièce $n = 30$ fois. On note les différents tirages

$$x_1, \ldots, x_n \quad \text{où} \quad x = 1 \text{ si Pile et } x = 0 \text{ si Face}$$

- Calcul de la fréquence empirique obtenue : $\hat{p} = \frac{1}{n} \sum_{i=1}^{n} x_i$
- On veut prendre une décision entre les deux hypothèses (soit on garde la pièce, soit on s'en débarrasse)

$\rightarrow$ Quelle règle de décision devrait-on choisir ?

- 1er choix : On jette la pièce si $\hat{p} > p_0$
$\rightarrow$ Si la pièce n'est pas une fausse, on va quand même la jeter dans 50% de cas.
- 2ème choix : On jette la pièce si $\hat{p} > \frac{p_0+p_1}{2}$
$\rightarrow$ Le risque de rejeter à tort est égal au risque de garder à tort.
- Avec ces deux choix, on risque trop souvent de se débarrasser de la pièce à tort.
$\rightarrow$ **Les deux risques ne se valent pas. On préfère garder la pièce à tort plutôt que de la jeter à tort.**
- 3ème choix : On jette la pièce si $\hat{p} > \frac{1}{5}p_0 + \frac{4}{5}p_1$
$\rightarrow$ On favorise la décision de garder la pièce.

- **Modèle :** $X$ de loi $\mathcal{B}(p)$ avec $p \in \{0.5, 0.6\} = \{p_0, p_1\}$
- **Estimateur de p :** $\bar{X}_n$ sans biais de variance min. et de loi $\mathcal{N}(p, p(1-p)/n)$
- **Hypothèses à tester :**

$$\mathcal{H}_0 : p = p_0 \qquad \mathcal{H}_1 : p = p_1$$

- Sous $\mathcal{H}_0$ : $\bar{X}_n$ suit $\mathcal{N}(p_0, p_0(1-p_0)/n)$ et

$$\begin{aligned} P(\bar{X}_n > 0.55) &= P\left(\frac{\bar{X}_n - p_0}{\sqrt{p_0(1-p_0)/n}} > \frac{0.55 - p_0}{\sqrt{p_0(1-p_0)/n}}\right) \\ &= 1 - \Phi\left(\frac{0.55 - p_0}{\sqrt{p_0(1-p_0)/n}}\right) \\ &= P(accepter \mathcal{H}_1 | \mathcal{H}_0 vrai) = 0.31 \end{aligned}$$

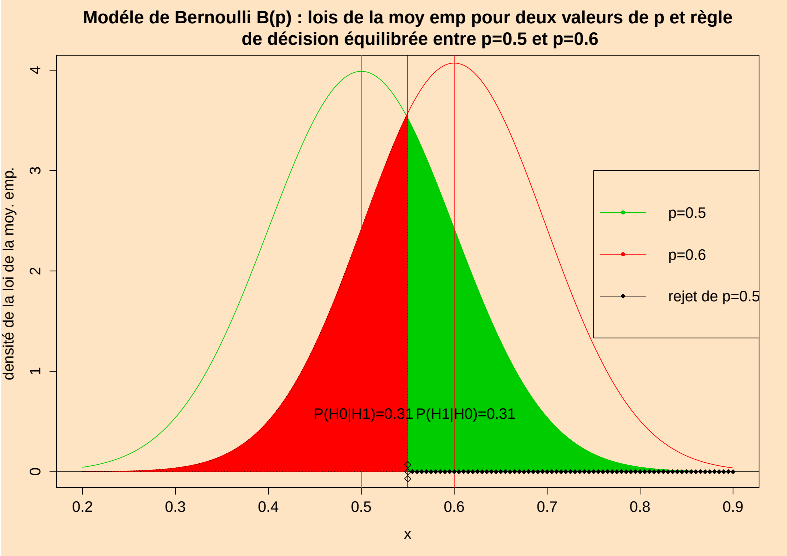
De même sous $\mathcal{H}_1$ : $\bar{X}_n$ suit une $\mathcal{N}(p_1, p_1(1-p_1)/n)$ et

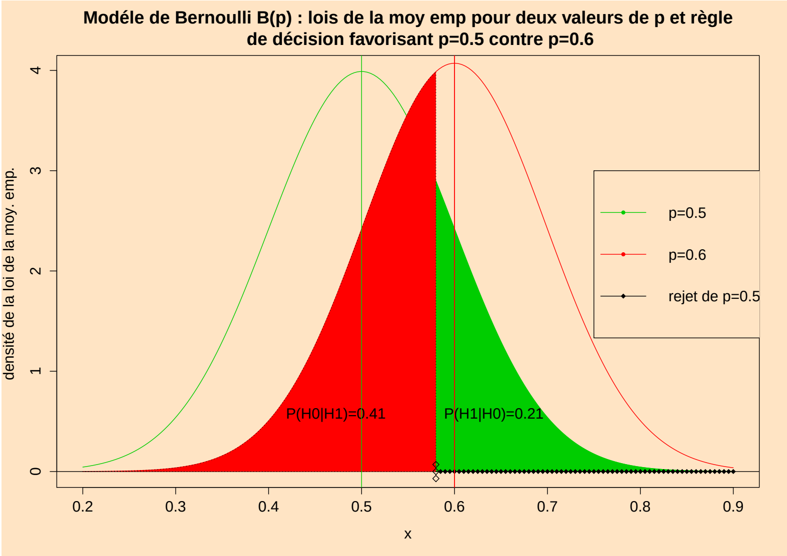
$$\begin{aligned} P(\bar{X}_n \leq 0.55) &= P\left(\frac{\bar{X}_n - p_1}{\sqrt{p_1(1-p_1)/n}} \leq \frac{0.55 - p_1}{\sqrt{p_1(1-p_1)/n}}\right) \\ &= \Phi\left(\frac{0.55 - p_1}{\sqrt{p_1(1-p_1)/n}}\right) \\ &= P(accepter \mathcal{H}_0 | \mathcal{H}_1 vrai) = 0.31 \end{aligned}$$

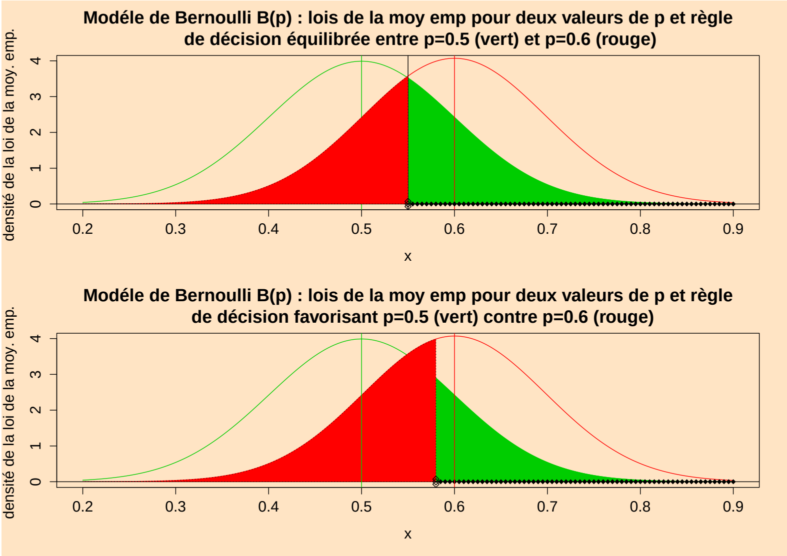
Ce qui caractérise un test en plus de sa région de rejet :

1. *Risque de première espèce* : $\alpha = P(accepter \mathcal{H}_1 | \mathcal{H}_0 vrai)$ que l'on souhaite contrôler :
   **celui de refuser $\mathcal{H}_0$ à tort**
2. *Risque de seconde espèce* : $\beta = P(accepter \mathcal{H}_0 | \mathcal{H}_1 vrai)$.
   Risque de refuser $\mathcal{H}_1$ à tort.
   Dépend aussi de l'écart entre les deux hypothèses testées. Si elle sont "collées" la somme de ces deux risques vaut 1.
3. *La puissance* : $\gamma = 1 - \beta = P(accepter \mathcal{H}_1 | \mathcal{H}_1 vrai)$

Sur la figure suivante : $\alpha$ : surface rouge (verte pour $\beta$)


Modéle de Bernoulli B(p) : lois de la moy emp pour deux valeurs de p et règle de décision équilibrée entre p=0.5 et p=0.6
densité de la loi de la moy. emp.
0
1
2
3
4
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
x
p=0.5
p=0.6
rejet de p=0.5
P(H0|H1)=0.31
P(H1|H0)=0.31


Modéle de Bernoulli B(p) : lois de la moy emp pour deux valeurs de p et règle de décision favorisant p=0.5 contre p=0.6
densité de la loi de la moy. emp.
0
1
2
3
4
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
x
P(H0|H1)=0.41
P(H1|H0)=0.21
p=0.5
p=0.6
rejet de p=0.5


Modéle de Bernoulli B(p) : lois de la moy emp pour deux valeurs de p et règle de décision équilibrée entre p=0.5 (vert) et p=0.6 (rouge)
densité de la loi de la moy. emp.
0
1
2
3
4
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
x
Modéle de Bernoulli B(p) : lois de la moy emp pour deux valeurs de p et règle de décision favorisant p=0.5 (vert) contre p=0.6 (rouge)
densité de la loi de la moy. emp.
0
1
2
3
4
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
x

A chaque test défini par sa région de rejet $W = \{\bar{X}_n > C\}$ correspond un risque de première espèce différent : pour $C = 0.55$ on a obtenu 31% et pour $C = 0.58$ on a obtenu 21%. On pourrait aussi chercher $C_\alpha$ pour que $W_\alpha = \{\bar{X}_n > C_\alpha\}$ soit la région de rejet d'un test de risque de première espèce donné $\alpha$.
Dans ce cas $C_\alpha$ est tel que

$$P\left(\frac{\bar{X}_n - p_0}{\sqrt{p_0(1-p_0)/n}} > \frac{C_\alpha - p_0}{\sqrt{p_0(1-p_0)/n}}\right) = \alpha$$

et après quelques manipulations algébriques on obtient :

$$C_\alpha = p_0 + \sqrt{\frac{p_0(1-p_0)}{n}}\, u_{1-\alpha}$$

Ex : calculer la puissance de ce test lorsqu'il est de niveau 5%.

Lorsque l'on dispose d'une seule estimation $\hat{p}$ (car un seul échantillon est observé) et que l'on souhaite faire le test à différents niveaux $\alpha$ on peut calculer les $C_\alpha$ correspondant et les décisions associées :

| $\alpha$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |
|---|---|---|---|---|---|
| $C_\alpha$ | 0.628 | 0.584 | 0.552 | 0.525 | 0.5 |

Ex : pour une estimation $\hat{p} = 0.57$ dans le test de niveau :

1. 20% : comme $0.57 \leq 0.584$ on ne rejette pas $p = 0.5$
2. 30% : comme $0.57 > 0.552$ on rejette $p = 0.5$ et on valide $p = 0.6$
3. 40% : comme $0.57 > 0.525$ on rejette $p = 0.5$ et on valide $p = 0.6$

Pour un échantillon la décision change selon le $\alpha$ choisi. **p-valeur :** la valeur seuil $\alpha^*$ la plus petite telle que pour tout risque $\alpha$ supérieur on rejette $\mathcal{H}_0$ et valide $\mathcal{H}_1$.