

# MAP 433 : Introduction aux méthodes statistiques. Cours 6

2 Octobre 2015

# Aujourd'hui

- 1 Tests statistiques
  - Notion de test et d'erreur de test
  - Lemme de Neyman-Pearson
- 2 Tests gaussiens
  - Tests sur la moyenne
- 3 Compléments :  $p$ -valeur et liens entre tests et régions de confiance

## Exemple introductif

- On observe 10 lancers d'une pièce de monnaie et on obtient le résultat suivant :

$(P, P, F, F, P, F, P, P, F, P).$

La pièce est-elle équilibrée ?

- Répondre à cette question revient à **construire une procédure de décision** :

$$\begin{aligned} \varphi &= \varphi(P, P, F, F, P, F, P, P, F, P) \\ &= \begin{cases} 0 & \text{on accepte l'hypothèse la pièce est équilibrée} \\ 1 & \text{on rejette l'hypothèse la pièce est équilibrée} \end{cases} \end{aligned}$$

# Résolution

- On associe l'expérience statistique (par exemple)

$$\mathcal{E}^{10} = (\{0, 1\}^{10}, \text{parties de}(\{0, 1\}^{10}), \{\mathbb{P}_\theta^{10}, \theta \in \Theta = [0, 1]\}),$$

avec  $(P = 0, F = 1)$

$$\mathbb{P}_\theta^{10} = (\theta\delta_0(dx) + (1 - \theta)\delta_1(dx))^{\otimes 10}.$$

- Hypothèse nulle : la pièce est équilibrée

$$H_0 : \theta \in \Theta_0 = \{1/2\}$$

- Hypothèse alternative : la pièce est truquée

$$H_1 : \theta \in \Theta_1 = \Theta \setminus \{1/2\}$$

# Résolution

- $\Theta_0$  = ensemble des paramètres sous laquelle l'hypothèse nulle est satisfaite
- $\Theta_1$  = ensemble des paramètres sous laquelle l'hypothèse nulle n'est pas satisfaite = **alternative**
- $\Theta = \Theta_0 \cup \Theta_1$ .

# Règle de décision

- On note  $Z$  l'observation.
- On **construit** une **règle de décision simple** :

$$\varphi(Z) = \mathbb{1}_{\mathcal{R}}(Z) = \begin{cases} 0 & \text{on accepte l'hypothèse} \\ 1 & \text{on rejette l'hypothèse.} \end{cases}$$

- $\mathcal{R} \subset \mathfrak{Z}$  (espace des observables) : **zone de rejet** ou **région critique**.
- Exemple<sup>1</sup>

$$\mathcal{R} = \{|\hat{\theta}(z) - 1/2| > t_0\}, \quad \hat{\theta}(Z) = \hat{\theta}_n^{\text{mv}} \left( \overset{\text{exemple}}{=} 0, 6 \right)$$

où  $t_0$  est un seuil à choisir... **Comment ?**

---

1. léger abus de notation...

# Terminologie

- Une **règle de décision** (non-randomisée) assigne à chaque réalisation  $z$  de l'observation  $Z$  une **décision**.
- Si  $z \in \mathcal{A}$ , l'hypothèse nulle est **acceptée** ; autrement, l'hypothèse est rejetée.
- **Terminologie** :  $\mathcal{A}$  = **zone d'acceptation** ;  $\mathcal{R}$  = **zone de rejet** ou **région critique**.

# Erreur de décision

- Lorsque l'on prend la décision  $\varphi(Z)$ , on peut se **tromper de deux manières** :

$$\text{Rejeter } H_0 \text{ } (\varphi(Z) = 1) \text{ alors que } \theta = \frac{1}{2}$$

ou encore

$$\text{Accepter } H_0 \text{ } (\varphi(Z) = 0) \text{ alors que } \theta \neq \frac{1}{2}.$$

- Erreur de première espèce (=rejeter à tort)

$$\mathbb{P}_{\frac{1}{2}}^{10} [\varphi(Z) = 1]$$

- Erreur de seconde espèce (=accepter à tort)

$$(\mathbb{P}_{\theta}^{10} [\varphi(Z) = 0], \theta \neq \frac{1}{2}).$$



## Retour à l'exemple

- Sous  $\mathbb{P}_\theta$ ,  $n\hat{\theta}(Z) = \sum_{i=1}^n Z_i$  suit une loi binomiale de paramètre de succès  $\theta$ . En notant  $\text{Bin}_{n,\theta}$  la fonction de répartition de la loi binomiale,

$$\begin{aligned}\mathbb{P}_\theta(|\hat{\theta}(Z) - 1/2| > t_0) \\ = 1 - \{\text{Bin}_{n,\theta}(n(1 + t_0)/2) - \text{Bin}_{n,\theta}(n(1 - t_0)/2)\}.\end{aligned}$$

- Erreur de première espèce :

$$\begin{aligned}\mathbb{P}_{1/2}(|\hat{\theta}(Z) - 1/2| > t_0) \\ = 1 - \{\text{Bin}_{n,1/2}(n(1 + t_0)/2) - \text{Bin}_{n,1/2}(n(1 - t_0)/2)\}\end{aligned}$$

# Erreur de première espèce

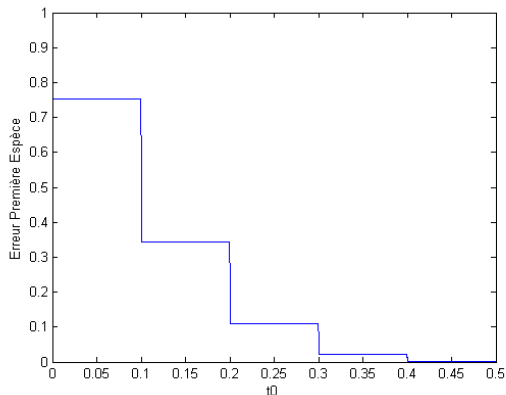


FIGURE – Erreur de Première Espèce en fonction de la valeur critique :  
fonction décroissante de  $t_0$  !

## Retour à l'exemple

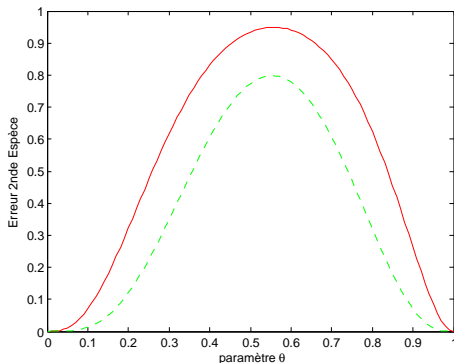
- Sous  $\mathbb{P}_\theta$ ,  $n\hat{\theta}(Z) = \sum_{i=1}^n Z_i$  suit une loi binomiale de paramètre de succès  $\theta$ . En notant  $\text{Bin}_{n,\theta}$  la fonction de répartition de la loi binomiale,

$$\begin{aligned}\mathbb{P}_\theta(|\hat{\theta}(Z) - 1/2| \leq t_0) \\ = \{\text{Bin}_{n,\theta}(n(1 + t_0)/2) - \text{Bin}_{n,\theta}(n(1 - t_0)/2)\}.\end{aligned}$$

- **Erreur de seconde espèce** : pour une valeur critique fixée :

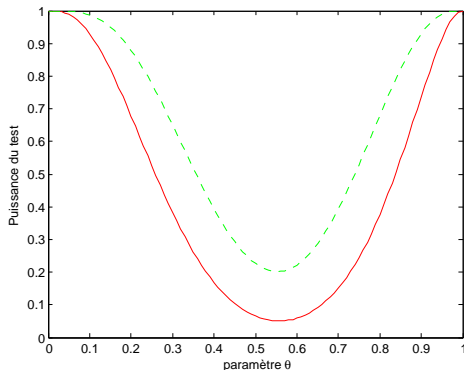
$$\begin{aligned}\theta \mapsto \mathbb{P}_\theta(|\hat{\theta}(Z) - 1/2| \leq t_0) \\ = \text{Bin}_{n,1/2}(n(1 + t_0)/2) - \text{Bin}_{n,1/2}(n(1 - t_0)/2)\end{aligned}$$

# Erreur de seconde espèce



**FIGURE** – Erreur de deuxième Espèce en fonction du paramètre  $\theta \in \Theta_1 = [0, 1] \setminus \{1/2\}$  pour deux valeurs du seuil critique :  $t_0 = 0.3$ , Erreur 1ere espèce 0.0654 et  $t_0 = 0.2$ , Erreur 1ère espèce 0.22

# Puissance du test



**FIGURE** – Puissance du test en fonction du paramètre  $\theta \in \Theta_1 = [0, 1] \setminus \{1/2\}$  pour deux valeurs du seuil critique :  $t_0 = 0.3$ .  
 Erreur 1ère espèce : 0.0654 et  $t_0 = 0.2$ , Erreur 1ère espèce : 0.22

# Erreur 1ère espèce / Puissance

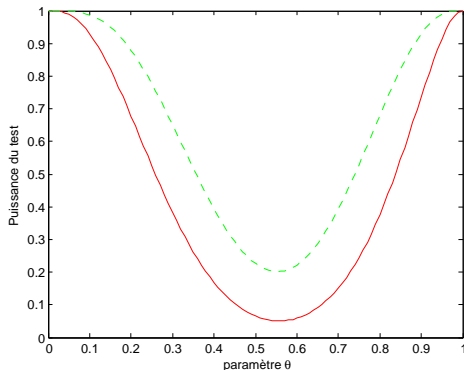


FIGURE – Erreur de première espèce / puissance pour une contre-alternative fixée  $\theta_1 = 0.75$  pour différentes valeurs du seuil critique  $t_0$ .

# Conclusion provisoire

- Un bon test  $\varphi$  devrait garantir **simultanément** des erreurs de première et seconde espèce **petites**.
- Mais il faut réaliser un compromis entre erreur de 1ère espèce et erreur de 2nde espèce (ou de façon équivalente entre erreur de 1ère espèce et puissance).
- **Question** : comment aborder la notion d'optimalité et comment construire une procédure de test satisfaisante ?

# Définition formelle

- Situation :  $\mathcal{E} = (\mathcal{Z}, \mathfrak{Z}, \{\mathbb{P}_\theta, \theta \in \Theta\})$  engendrée par l'observation  $Z$ .
- **Hypothèse nulle et alternative** :  $\Theta_0 \subset \Theta$  et  $\Theta_1 = \Theta \setminus \Theta_0$

## Definition (Test simple)

*Un test (simple) de l'hypothèse nulle  $H_0 : \theta \in \Theta_0$  contre l'alternative  $H_1 : \theta \in \Theta_1$  est une statistique  $\varphi = \varphi(Z) \in \{0, 1\}$ . (Fonction d') **erreur de première espèce** :*

$$\theta \in \Theta_0 \rightsquigarrow \mathbb{P}_\theta [\varphi(Z) = 1]$$

*(Fonction d') **erreur de seconde espèce***

$$\theta \in \Theta_1 \rightsquigarrow \mathbb{P}_\theta [\varphi(Z) = 0] = 1 - \text{puissance}_\varphi(\theta).$$



# Principe de Neyman

- On **disymétrise** les hypothèses  $H_0$  et  $H_1$  :  $H_0$  est plus importante que  $H_1$  dans le sens suivant : on **impose** une **erreur de première espèce prescrite**.

## Definition

*Pour  $\alpha \in [0, 1]$ , un test  $\varphi = \varphi_\alpha$  de l'hypothèse nulle  $H_0 : \theta \in \Theta_0$  contre une alternative  $H_1$  est de niveau  $\alpha$  si*

$$\sup_{\theta \in \Theta_0} \mathbb{P}_\theta [\varphi_\alpha = 1] \leq \alpha.$$

- Un test de niveau  $\alpha$  ne dit **rien** sur l'erreur de seconde espèce (comportement sur l'alternative).

# Principe de Neyman (cont.)

- Choix de la disymétrisation = choix de modélisation.
- Principe de Neyman :  $\alpha \in (0, 1)$ , parmi les test de niveau  $\alpha$ , chercher celui (ou ceux) ayant une **erreur de seconde espèce minimale**.

## Definition

*Un test de niveau  $\alpha$  est dit **Uniformément Plus Puissant (UPP)** si son erreur de seconde espèce est minimale parmi celles des tests de niveau  $\alpha$ .*

- Pour le cas d'une **hypothèse simple** contre une **alternative simple**, un test UPP existe.

## Règle de décision randomisée

- Pour toute valeur de l'observation  $z$ , la règle de décision choisit alternative avec une probabilité  $\varphi(z)$  et l'hypothèse nulle avec la probabilité  $1 - \varphi(z)$ .
- Une procédure de test randomisée est **entièrement spécifiée** par la donnée de la **fonction critique** du test  $\varphi : z \rightarrow \varphi(z) \in [0, 1]$ . Si  $\varphi$  prend simplement les valeurs 0 et 1, on obtient un test non randomisé.
- La **probabilité de rejet** est donnée, pour tout  $\theta \in \Theta$ , par  $\mathbb{E}_\theta[\varphi(Z)]$ .

# Principe de Neyman-Pearson

Le problème revient donc à maximiser la **puissance du test**

$$\pi(\theta) = \mathbb{E}_{\theta}[\varphi(Z)], \theta \in \Theta_1$$

sous la contrainte que le niveau du test soit inférieure à  $\alpha$

$$\mathbb{E}_{\theta}[\varphi(Z)] \leq \alpha$$

## Un cas élémentaire

- Supposons que  $\Theta_0 = \{\theta_0\}$  et  $\Theta_1 = \{\theta_1\}$ .
- On note  $p_0(z) = f(\theta_0, z)$  et  $p_1(z) = f(\theta_1, z)$  les densités des lois  $\mathbb{P}_{\theta_0}$  et  $\mathbb{P}_{\theta_1}$  par rapport à une mesure de domination  $\mu$  (existe toujours)

### Theorem (Existence d'un test de niveau $\alpha$ )

*Pour tester  $H_0 : \{\theta = \theta_0\}$  contre l'alternative  $H_1 : \{\theta = \theta_1\}$ , il existe un test  $\varphi$  et une constante  $c_\alpha$  telle que*

$$\mathbb{E}_{\theta_0}[\phi(Z)] = \alpha$$

*et*

$$\varphi(z) = \begin{cases} 1 & \text{quand } p_1(z) > c_\alpha p_0(z) \\ 0 & \text{quand } p_1(z) < c_\alpha p_0(z) \end{cases}$$

# Preuve

- Soit  $\alpha(c) = \mathbb{P}_0(p_1(Z) > cp_0(Z))$ , qui est la probabilité que la variable aléatoire  $p_1(Z)/p_0(Z)$  soit strictement plus grande que  $c$  (il suffit de se placer sur l'événement  $\{p_0(Z) > 0\}$ ).
- $c \mapsto \alpha(c)$  est donc décroissante, continue à droite et admet des limites à gauche ( $c \mapsto 1 - \alpha(c)$  est une fonction de répartition !)

$$\alpha(c^-) - \alpha(c) = \mathbb{P}_0(p_1(Z)/p_0(Z) = c), \alpha(-\infty) = 1, \alpha(+\infty) = 0$$

# Preuve

- Pour tout  $\alpha \in (0, 1)$ , il existe  $c_\alpha$  tel que  $\alpha(c_\alpha) \leq \alpha \leq \alpha(c_\alpha^-)$ .
- On considère le test  $\varphi_\alpha$  défini par

$$\varphi_\alpha(z) = \begin{cases} 1 & \text{quand } p_1(z) > c_\alpha p_0(z) \\ \frac{\alpha - \alpha(c_\alpha)}{\alpha(c_\alpha^-) - \alpha(c_\alpha)} & \text{quand } p_1(z) = c_\alpha p_0(z) \\ 0 & \text{quand } p_1(z) < c_\alpha p_0(z) \end{cases}$$

La seconde ligne a un sens seulement si  $\alpha(c_\alpha^-) > \alpha(c_\alpha)$ .

- Le niveau de  $\phi$  est donné par

$$\begin{aligned} \mathbb{E}_0[\varphi(Z)] &= \mathbb{P}_0 \left( \frac{p_1(Z)}{p_0(Z)} > c_\alpha \right) + \frac{\alpha - \alpha(c_\alpha)}{\alpha(c_\alpha^-) - \alpha(c_\alpha)} \mathbb{P}_0 \left( \frac{p_1(Z)}{p_0(Z)} = c_\alpha \right) \\ &= \alpha \end{aligned}$$

# Test le plus puissant

## Theorem

Un test  $\varphi$  vérifiant

$$\mathbb{E}_{\theta_0}[\phi(Z)] = \alpha$$

et

$$\varphi(z) = \begin{cases} 1 & \text{quand } p_1(z) > c_\alpha p_0(z) \\ 0 & \text{quand } p_1(z) < c_\alpha p_0(z) \end{cases}$$

est **uniformément le plus puissant** pour tester l'hypothèse nulle  $H_0 : \{\theta = \theta_0\}$  contre l'alternative  $H_1 : \{\theta = \theta_1\}$ .



# Preuve

Soit  $\varphi$  un test satisfaisant les conditions

$$\mathbb{E}_{\theta_0}[\phi(Z)] = \alpha$$

et

$$\varphi(z) = \begin{cases} 1 & \text{quand } p_1(z) > c_\alpha p_0(z) \\ 0 & \text{quand } p_1(z) < c_\alpha p_0(z) \end{cases}$$

et  $\varphi^*$  un test de niveau  $\mathbb{E}_0[\varphi^*(Z)] \leq \alpha$

# Preuve

- On note  $S^+ = \{z : \varphi(z) - \varphi^*(z) > 0\}$  et  $S^- = \{z : \varphi(z) - \varphi^*(z) < 0\}$ .
- Pour  $z \in S^+$ ,  $\phi(z) > 0$  et donc  $p_1(z) \geq c_\alpha p_0(z)$ .
- Pour  $z \in S^-$ ,  $\phi(z) < 1$  et donc  $p_1(z) \leq c_\alpha p_0(z)$ .
- Par conséquent :

$$\int (\varphi - \varphi^*)(p_1 - c_\alpha p_0) d\mu \geq 0$$

# Preuve

Comme

$$\int (\varphi - \varphi^*)(p_1 - c_\alpha p_0) d\mu \geq 0$$

Nous avons

$$\int (\varphi - \varphi^*)(p_1 - c_\alpha p_0) d\mu \geq c_\alpha \int (\varphi - \varphi^*)(p_1 - c_\alpha p_0) d\mu \geq 0.$$

# Puissance d'un test U.P.P

## Lemma

Notons  $\pi$  la puissance du test U.P.P de niveau  $\alpha$  du test  $H_0 = \{\theta = \theta_0\}$  contre l'alternative  $H_1 = \{\theta = \theta_1\}$ . Alors  $\alpha \leq \pi$  avec égalité si  $\mathbb{P}_{\theta_0} = \mathbb{P}_{\theta_1}$ .

## Démonstration.

Comme le test  $\varphi(z) \equiv \alpha$  a un niveau  $\alpha$ , nous avons  $\alpha \leq \pi$ . □

## Exemple de mise en oeuvre

- On observe

$$Z = (X_1, \dots, X_n) \sim_{\text{i.i.d.}} \mathcal{N}(\theta, 1).$$

- Construction du test de N-P. de  $H_0 : \theta = \theta_0$  contre  $H_1 : \theta = \theta_1$ , avec  $\theta_0 < \theta_1$ .
- Mesure dominante  $\mu^n =$  mesure de Lebesgue sur  $\mathbb{R}^n$  et

$$f(\theta, Z) = \frac{1}{(2\pi)^{n/2}} \exp \left( -\frac{1}{2} \sum_{i=1}^n X_i^2 + n\theta \bar{X}_n - \frac{n\theta^2}{2} \right).$$

- Rapport de vraisemblance

$$\frac{f(\theta_1, Z)}{f(\theta_0, Z)} = \exp \left( n(\theta_1 - \theta_0) \bar{X}_n - \frac{n}{2}(\theta_1^2 - \theta_0^2) \right).$$

## Exemple (cont.)

- **Zone de rejet** du test de N-P. :

$$\begin{aligned} & \{f(\theta_1, Z) > cf(\theta_0, Z)\} \\ &= \{n(\theta_1 - \theta_0)\bar{X}_n - \frac{n}{2}(\theta_1^2 - \theta_0^2) > \log c\} \\ &= \{\bar{X}_n > \frac{\theta_0 + \theta_1}{2} + \frac{\log c}{n(\theta_1 - \theta_0)}\}. \end{aligned}$$

- **Choix de  $c$** . On résout

$$\mathbb{P}_{\theta_0} \left[ \bar{X}_n > \frac{1}{2}(\theta_0 + \theta_1) + \frac{\log c}{n(\theta_1 - \theta_0)} \right] = \alpha.$$

- **Approche standard** : on raisonne sous  $\mathbb{P}_{\theta_0}$ . On a

$$\bar{X}_n = \theta_0 + \frac{1}{\sqrt{n}}\xi^{n,\theta_0}, \quad \xi^{n,\theta_0} \sim \mathcal{N}(0, 1)$$

sous  $\mathbb{P}_{\theta_0}$  **mais pas sous une autre probabilité  $\mathbb{P}_\theta$  si  $\theta \neq \theta_0$  !**

# Exemple (fin)

## ■ Résolution de

$$\mathbb{P}_{\theta_0} \left[ \theta_0 + \frac{1}{\sqrt{n}} \xi^{n, \theta_0} > \frac{1}{2}(\theta_0 + \theta_1) + \frac{\log c}{n(\theta_1 - \theta_0)} \right] = \alpha.$$

## ■ Equivalent à $\mathbb{P}_{\theta_0} \left[ \xi^{n, \theta_0} > \frac{\sqrt{n}}{2}(\theta_1 - \theta_0) + \frac{1}{\sqrt{n}} \frac{\log c}{\theta_1 - \theta_0} \right] = \alpha$ , soit

$$\frac{\sqrt{n}}{2}(\theta_1 - \theta_0) + \frac{1}{\sqrt{n}} \frac{\log c}{\theta_1 - \theta_0} = \Phi^{-1}(1 - \alpha),$$

$$\text{où } \Phi(x) = \int_{-\infty}^x e^{-u^2/2} \frac{du}{\sqrt{2\pi}}.$$

# p-valeur

- Supposons que sous  $\mathbb{P}_0$ , la distribution du rapport de vraisemblance  $p_1(Z)/p_0(Z)$  est continue. Le test U.P.P. de niveau  $\alpha$  est non-randomisé et rejette l'hypothèse nulle si  $p_1(Z)/p_0(Z) > k$ , où  $k$  est choisi de façon à ce que

$$\mathbb{P}_0(p_1(Z)/p_0(Z) \geq k) = \alpha.$$

- Lorsque l'on fait varier le niveau  $\alpha$ , on obtient ainsi une famille de régions de réjections,  $\{\mathcal{R}_\alpha\}_{\alpha \in [0,1]}$ , qui sont dans de nombreux cas d'intérêt emboîtées

$$\mathcal{R}_\alpha \subset \mathcal{R}_{\alpha'} \quad \text{si } \alpha < \alpha'.$$



## p-valeur

- Lorsque l'on fait varier le niveau  $\alpha$ , on obtient ainsi une famille de régions de réjections,  $\{\mathcal{R}_\alpha\}_{\alpha \in [0,1]}$ , qui sont dans de nombreux cas d'intérêt emboîtées

$$\mathcal{R}_\alpha \subset \mathcal{R}_{\alpha'} \quad \text{si } \alpha < \alpha'.$$

- Lorsque cette condition est satisfaite, il est intéressant de déterminer non seulement si le test est accepté ou rejeté à un niveau de signification donné, mais aussi de déterminer **le plus petit niveau de signification** auquel l'hypothèse serait rejetée.

# p-valeur

## Definition (p-valeur)

La  $p$ -valeur d'un test pour une observation  $Z$  donnée est le plus petit niveau de signification auquel le test serait rejeté

$$\hat{p}(Z) = \inf\{\alpha : Z \in S_\alpha\}.$$

- Une petite  $p$ -valeur suggère que l'observation contredit l'hypothèse.
- Une grande  $p$ -valeur s'interprète en faveur de ne pas vouloir rejeter l'hypothèse de base.

# Bilan provisoire

- Si l'on accepte **le principe de Neyman**, on sait résoudre le problème à deux points.
- Que faire si l'hypothèse nulle  $H_0$  ou l'alternative  $H_1$  sont **composites** ?
  - On peut proposer des extensions si l'on dispose de structures particulières sur la vraisemblance du modèle (Poly. Ch. 7.3, hors programme).
  - On sait dire **beaucoup de choses** dans le cas gaussien.
- **Critique méthodologique de l'approche de Neyman**  $\rightsquigarrow$  notion de  $p$ -valeur.

# Tests gaussiens incontournables

- On observe

$$\mathbf{Y} = (Y_1, \dots, Y_n) \sim \mathcal{N}(\mu, \sigma^2 \text{Id}_n).$$

- Test sur la moyenne, variance connue

$$H_0 : \mu \leq \mu_0 \quad \text{contre} \quad H_1 : \mu > \mu_0$$

- **Principe** on estime  $\mu$  et on rejette  $H_0$  si l'estimateur est plus grand que  $\mu_0$ .

$$\mathcal{R}(c_\alpha) = \{\bar{Y}_n - \mu_0 \geq c_\alpha\}, \quad c_\alpha \text{ à déterminer.}$$

- On choisit  $c_\alpha$  de sorte que

$$\sup_{\mu \leq \mu_0} \mathbb{P}_\mu [\mathcal{R}(c_\alpha)] \leq \alpha.$$

- Il y a **plusieurs choix possibles**. On fait le choix rendant  $\mathcal{R}(c_\alpha)$  maximale.

# Calcul de $c_\alpha$

- Majoration de l'erreur de première espèce. Si  $\mu \leq \mu_0$ , on a

$$\begin{aligned}\mathbb{P}_\mu [\overline{Y}_n - \mu_0 \geq c_\alpha] &= \mathbb{P}_\mu \left[ \left( \mu + \frac{\sigma}{\sqrt{n}} \xi^{n,\mu} \right) - \mu_0 \geq c_\alpha \right] \\ &= \mathbb{P}_\mu \left[ \frac{\sigma}{\sqrt{n}} \xi^{n,\mu} \geq c_\alpha + (\mu_0 - \mu) \right] \\ &\leq \mathbb{P}_\mu \left[ \frac{\sigma}{\sqrt{n}} \xi^{n,\mu} \geq c_\alpha \right].\end{aligned}$$

où  $\xi^{n,\mu}$  est, en loi sous  $\mathbb{P}_\mu$ , une gaussienne standard.

- Petit miracle : la loi de  $\xi^{n,\mu}$  sous  $\mathbb{P}_\mu$  ne dépend pas de  $\mu$  Donc

$$\mathbb{P}_\mu \left[ \frac{\sigma}{\sqrt{n}} \xi^{n,\mu} \geq c_\alpha \right] = 1 - \Phi \left( \frac{\sqrt{n}}{\sigma} c_\alpha \right)$$

on veut  
 $\leq \alpha$ .

- Le choix  $c_{\alpha,n} = \frac{\sigma}{\sqrt{n}} \Phi^{-1}(1 - \alpha)$  conduit à la zone de rejet  $\mathcal{R}(c_\alpha)$  maximale.

# Contrôle de l'erreur de seconde espèce

- On a construit un test de niveau  $\alpha$  parmi une classe **donnée a priori** de tests basés sur un estimateur raisonnable, de sorte que l'on ait une zone de rejet maximale. Désormais,  $c_{\alpha,n}$  est **fixé**.
- On **évalue à la main** l'erreur de seconde espèce ou la **fonction de puissance**

$$\begin{aligned}\mu \in (\mu_0, +\infty) &\rightsquigarrow \mathbb{P}_\mu [\overline{Y}_n - \mu_0 < c_{\alpha,n}] \\ &= 1 - \text{puissance du test au point } \mu\end{aligned}$$

- **Montrer que** pour tout  $\mu > \mu_0$ , on a  $\mathbb{P}_\mu [\overline{Y}_n - \mu_0 < c_{\alpha,n}] \rightarrow 0$  lorsque  $n \rightarrow \infty$ .
- Pour l'optimalité dans un sens plus fort, il faut **d'autres outils**.

# Autres tests classiques gaussiens

## ■ Ingrédient principal :

$$s_n^2 := \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2 = \frac{n}{n-1} (\hat{\sigma}_n^2)^{mv}$$

alors

$$(n-1) \frac{s_n^2}{\sigma^2} \sim \chi^2(n-1)$$

et

$$\frac{\sqrt{n}(\bar{Y}_n - \mu)}{s_n} \sim \text{Student}(n-1)$$

et ces variables sont **pivotales** : leur loi ne dépend pas de  $\mu, \sigma^2$  sous  $\mathbb{P}_{\mu, \sigma^2}$ .

- Les lois du  $\chi^2$  et de **Student** (à  $k$  degrés de liberté) sont classiques et s'étudient indépendamment.

## $p$ -valeurs

- Exemple : on observe

$$X_1, \dots, X_n \sim_{\text{i.i.d.}} \mathcal{N}(\mu, \sigma^2), \quad \sigma^2 \text{ connu.}$$

- Objectif : tester  $H_0 : \mu = 0$  contre  $H_1 : \mu \neq 0$ .
- Au niveau  $\alpha = 5\%$ , on rejette si

$$|\bar{X}_n| > \frac{\phi^{-1}(1 - \alpha/2)}{\sqrt{n}}$$

- Application numérique :  $n = 100$ ,  $\bar{X}_{100} = 0.307$ . On a  $\frac{\phi^{-1}(1 - 0.05/2)}{\sqrt{100}} \approx 0.196$ . on rejette l'hypothèse....



## $p$ -valeur (cont.)

- Et pour un autre choix de  $\alpha$ ? Pour  $\alpha = 0.01$ , on a  $\frac{\phi^{-1}(1-0.05/2)}{\sqrt{100}} \approx 0.256$ . On rejette toujours... Pour  $\alpha = 0.001$ , on a  $\frac{\phi^{-1}(1-0.05/2)}{\sqrt{100}} \approx 0.329$ . On accepte  $H_0$  !
- Que penser de cette petite expérience ?
  - En pratique, on a une observation une bonne fois pour toute (ici 0.307) et on choisit  $\alpha$ ... comment ?
  - On ne veut pas  $\alpha$  trop grand (trop de risque), mais en prenant  $\alpha$  de plus en plus petit... on va fatalement finir par accepter  $H_0$  !
- Défaut de méthodologie inhérent au principe de Neyman (contrôle de l'erreur de première espèce).

# $p$ -valeur

- Quantité **significative** : non par le niveau  $\alpha$ , mais le **seuil de basculement de décision** : c'est la  $p$ -valeur ( $p$ -value) du test.

## Definition

Soit  $\mathcal{R}_\alpha$  une famille de zones de rejet d'un test de niveau  $\alpha$  pour une hypothèse  $H_0$  contre une alternative  $H_1$ . Soit  $Z$  l'observation associée à l'expérience. On a  $Z \in \mathfrak{Z}$  et  $\mathcal{R}_0 = \mathfrak{Z}$ . On appelle  **$p$ -valeur du test** la quantité

$$p\text{-valeur}(Z) = \inf\{\alpha, Z \in \mathcal{R}_\alpha\}.$$

# Interprétation de la $p$ -valeur

- Une grande valeur de la  $p$ -valeur s'interprète en faveur de **ne pas vouloir rejeter l'hypothèse**.
- Ne pas vouloir rejeter l'hypothèse peut signifier deux choses :
  - L'hypothèse est vraie
  - L'hypothèse est fausse **mais** le test n'est pas **puissant** (erreur de seconde espèce **grande**).
- **Souvent** : la  $p$ -valeur est la probabilité (sous  $H_0$ ) que la statistique de test d'une expérience copie soit  $\geq$  à la statistique de test observée.
- **Exemple du test du  $\chi^2$  et de l'expérience de Mendel (à suivre)**