Chapitre II

Tests statistiques

Plan

Exemples introductifs

- **Exemple 1.**
  On s'intéresse à la longueur de pièces fabriquées par une machine. "En théorie" la longueur moyenne de ces pièces doit être de 150cm.On décide de mesurer 49 pièces choisies au hasard. La valeur moyenne des mesures est de 149.9.

Peut-on dire que la machine est toujours bien réglée ?

- **Exemple 2.**
  Un médicament couramment utilisé est connu pour guérir 30% des patients. Un nouveau traitement est expérimenté sur 10 patients. On observe 7 guérisons.

Doit-on remplacer l'ancien traitement ?

Problématique de test

- Ici, il ne s'agit plus d'estimer un paramètre à partir d'un échantillon mais de prendre une décision à l'aide de cet échantillon.
- Répondre aux questions posées revient à choisir une hypothèse parmi deux (on les notera H_0 et H_1).
- Un test statistique permet de réaliser un tel choix.

Plan

Exemple 16 (Exemple du traitement)

Un médicament couramment utilisé est connu pour guérir 30% *des patients. Un nouveau traitement est expérimenté sur 10 patients. On observe 7 guérisons. Doit-on remplacer l'ancien traitement ?*

$\theta_0 = 0.3$ probabilité de guérison avec l'ancien traitement
$\theta = ?$ probabilité de guérison avec le nouveau traitement
Modélisation :

	X	probabilité
patient guéri par le nouveau traitement	1	θ
patient non guéri	0	$1-\theta$

Modèle statistique : $\left(\{0,1\}^{10}, \mathcal{P}\left(\{0,1\}\right), (\mathcal{B}(\theta))_{\theta\in]0,1[}\right)$.

Exemple du traitement

Deux hypothèses sont en compétition à propos de la vraie valeur de θ :

- La première, notée H_0 et appelée hypothèse nulle ou fondamentale :

$$H_0 : \text{le nouveau traitement est comme l'ancien} : (\theta = \theta_0)$$

- On veut savoir si le nouveau traitement est meilleur, on oppose à H_0 une hypothèse appelée hypothèse alternative. On la note H_1 :

$$H_1 : \text{Le nouveau traitement est meilleur} : (\theta > \theta_0)$$

Comment départager ces deux hypothèses ?

Exemple du traitement

On constitue un échantillon aléatoire $(X_1, \dots, X_n)$ de n patients.
S_n : nombre de guérisons parmi les n patients testés.

$$S_n \rightsquigarrow \mathcal{B}(n, \theta).$$

On se fixe une règle de décision pour trancher entre les deux hypothèses :

- Si $S_n \geq k$, alors H_0 est rejetée au profit de H_1.
- Si $S_n < k$, alors H_0 n'est pas rejetée au profit de H_1.

k est appelé seuil critique.

- $\mathcal{R} = \{S_n \geq k\}$: zone de rejet de H_0 au profit de H_1= région critique.
- $\overline{\mathcal{R}} = \{S_n < k\}$: zone de non-rejet de H_0 au profit de H_1= région d'acceptation.

Deux types d'erreur peuvent apparaitre :

	Conclusion du test	
	rejet de H_0	non-rejet de H_0
$\theta = \theta_0$ (H_0 vraie)	erreur de 1ère espèce	pas d'erreur
$\theta > \theta_0$ (H_1 vraie)	pas d'erreur	erreur de 2ème espèce

- ☛ $\mathbb{P}_{\theta_0}(S_n \geq k)$: probabilité de commettre l'erreur de première espèce= risque de première espèce. C'est la probabilité de rejeter à tort H_0.
- ☛ $\mathbb{P}_{\theta}(S_n < k)$ avec $\theta > \theta_0$, θ inconnu : probabilité de commettre l'erreur de deuxième espèce= risque de deuxième espèce. C'est la probabilité de ne pas rejeter à tort H_0.

Le seuil k est choisi pour avoir une faible probabilité d'erreur de première espèce.

Compromis

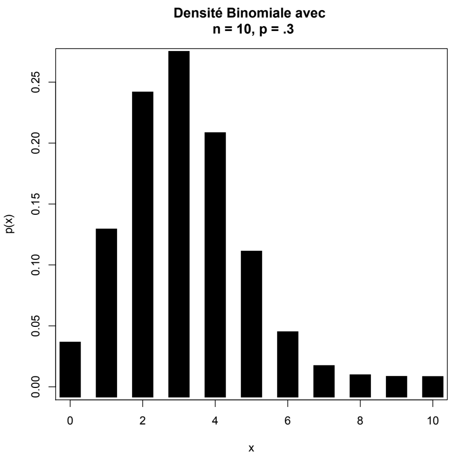
On fixe une probabilité maximale acceptable pour l'erreur de première espèce : $\alpha = 5\%, 10\%, \cdots$ est appelé niveau du test. On impose

$$\mathbb{P}_{\theta_0}(S_n \geq k) \leq \alpha$$

Parmi les k pour lesquels $\mathbb{P}_{\theta_0}(S_n \geq k) \leq \alpha$, on choisit le k (dépendant de α) qui rend la plus petite possible la probabilité d'erreur de 2ème espèce.

k	0	1	2	3	4	5	6	7	8
$\mathbb{P}(S_n = k)$	0.0282	0.1211	0.2335	0.2668	0.2001	0.1029	0.0368	0.0090	0.0[illegible]

$\mathbb{P}_{0.3}(S_{10} \geq 6) \approx 0.0473$ et on prend $k = 6$.



La zone de rejet du test sera l'événement

$$\mathcal{R} = \{(X_1, \ldots, X_{10}),\ S_{10} \geq 6\}$$

On observe $S_{10}(\omega) = s_{10} = 7$, $\mathcal{R}$ s'est produit, on décide de dire que ce traitement est meilleur.

Le nouveau traitement peut être pire que l'ancien ! ! ! ! on teste :

$$\begin{cases} H_0 : \text{le nouveau traitement n'est pas meilleur} : \theta \leq \theta_0 \\ H_1 : \text{le nouveau traitement est meilleur} : \theta > \theta_0 \end{cases}$$

Le seuil k est choisi pour contrôler l'erreur de première espèce :

$$\forall \theta \leq \theta_0, \quad \mathbb{P}_\theta(S_n \geq k) \leq \alpha = 5\%$$

$\theta \mapsto \mathbb{P}_\theta(S_n \geq k)$ est croissante lorsque $S_n \rightsquigarrow \mathcal{B}(n, \theta)$, donc cette condition se réduit à

$$\mathbb{P}_{\theta_0}(S_n \geq k) \leq \alpha$$

Donc le deuxième test a la même zone de rejet que celui avec l'hypothèse nulle plus simple :

$$k = 6, \quad \mathcal{R} = \{S_n \geq 6\}$$

On considère un n-échantillon $X_1, \dots, X_n$ associé au modèle $(\Omega, \mathcal{F}, (\mathbb{P}_\theta)_{\theta\in\Theta})$ et on se donne deux parties disjointes Θ_0 et Θ_1 telle que $\Theta_0 \cap \Theta_1 = \emptyset$ et $\Theta = \Theta_0 \cup \Theta_1$. On souhaite décider si $\theta \in \Theta_1$ en testant

- Hypothèse nulle $H_0 : \theta \in \Theta_0$ contre Hypothèse alternative $H_1 : \theta \in \Theta_1$

Exemple 17 (Traitement)

☛ $\Theta_0 = \{0.3\}$ *et* $\Theta_1 =]0.3, 1]$ *correspond au test*

$$H_0 : \theta = 0.3 \text{ contre } H_1 : \theta > 0.3$$

☛ $\Theta_0 =]0, 0.3]$ *et* $\Theta_1 =]0.3, 1[$ *correspond au test*

$$H_0 : \theta \leq 0.3 \text{ contre } H_1 : \theta > 0.3$$

Définition 12

Dans le modèle statistique $(\Omega, \mathcal{F}, (\mathbb{P}_\theta)_{\theta \in \Theta})$ *muni de l'échantillon* $\mathbf{X} = (X_1, \ldots, X_n)$*, construire un test de* $H_0 : \theta \in \Theta_0$ *contre* $H_1 : \theta \in \Theta_1$ *au niveau* α *revient à construire une zone de rejet* $\mathcal{R}$ *fonction de* $X_1, \ldots, X_n$ *de telle sorte que*

$$\sup_{\theta \in \Theta_0} \mathbb{P}_\theta(\mathcal{R}) \leq \alpha.$$

Règle de décision

- Si $\omega \in \mathcal{R}$, on rejette H_0 et on conclut H_1.
- Si $\omega \notin \mathcal{R}$, on ne rejette pas H_0.

Définition 13

Une statistique de test *est une fonction des v.a. $X_1, \ldots, X_n$ qui ne dépend pas du paramètre inconnu θ.*

Exemple 18 (traitement)

On a utilisé la statistique de test $S_n = \sum_{i=1}^{n} X_i$ pour construire la région de rejet $\mathcal{R} = \{S_n \geq 6\}$

Définition 14

- *Risque de première espèce : est la probabilité de rejeter à tort H_0.*

$$\alpha(\theta) = \mathbb{P}_\theta\left(\mathcal{R}\right) \text{ pour tout } \theta \in \Theta_0$$

- *Risque de deuxième espèce : est la probabilité de ne pas rejeter à tort H_0.*

$$\beta(\theta) = \mathbb{P}_\theta\left(\overline{\mathcal{R}}\right) \text{ pour tout } \theta \in \Theta_1.$$

- *Taille du test : est la probabilité maximale de rejeter à tort H_0.*

$$\alpha' = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta\left(\mathcal{R}\right)$$

- *Puissance du test ou fonction puissance est la probabilité de rejeter H_0 alors que H_0 est fausse (probabilité de tirer la bonne conclusion lorsque H_0 est fausse).*

$$\begin{aligned} \gamma(\theta) &= \mathbb{P}_\theta\left(\mathcal{R}\right) \text{ pour tout } \theta \in \Theta_1 \\ &= 1 - \beta(\theta) \end{aligned}$$

Définition 15

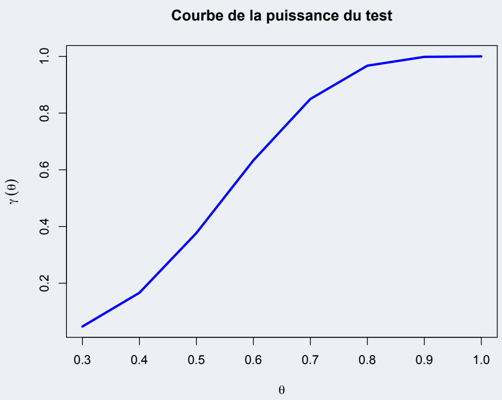
- *On dit qu'on effectue un test d'hypothèse simple contre une hypothèse simple lorsque $\Theta_0 = \{\theta_0\}$ et $\Theta_1 = \{\theta_1\}$ avec θ_0 et θ_1 connus.*
- *On dit qu'on effectue un test d'hypothèse multiple contre une hypothèse multiple lorsque $\Theta_0 =]-\infty, \theta_0]$ et $\Theta_1 =]\theta_0, +\infty[$.*
  - $H_0 : \theta \leq \theta_0$ *contre* $H_1 : \theta > \theta_0$ *Test unilatéral avec région de rejet à droite.*
  - $H_0 : \theta \geq \theta_0$ *contre* $H_1 : \theta < \theta_0$ *Test unilatéral avec région de rejet à gauche.*
  - $H_0 : \theta = \theta_0$ *contre* $H_1 : \theta \neq \theta_0$ *Test bilatéral.*

Exemple 19 (Traitement)

- *La taille du test est* $\alpha' = \mathbb{P}_{0.3}(S_{10} \geq 6) \approx 0.0473 \leq 0.05$ *(le niveau du test est de* 5%*).*
- *Le risque de deuxième espèce :* $\beta(\theta) = \mathbb{P}_\theta\,(S_{10} \leq 5)$ *pour tout* $\theta > 0.3$ $(< \mathbb{P}_{\theta_0}\,(S_{10} \leq 5) \approx 0.953)$

Exemple 20 (Traitement)

La fonction puissance vaut $\gamma(\theta) = \mathbb{P}_\theta(S_{10} \geq 6) = \sum_{k=6}^{10} \binom{10}{k} \theta^k (1-\theta)^{10-k}$ *avec* $\theta > 0.3$.



La puissance augmente lorsqu'on s'éloigne de l'hypothèse nulle !

Remarque 4

La zone de rejet de H_0 dépend du niveau du test choisi.

Exemple 21 (Traitement)

Niveau	**Seuil**	**Région de rejet**	**Décision**
$\alpha = 10\%$	$k = 6$	$\mathcal{R} = \{S_{10} \geq 6\}$	*On rejette H_0*
$\alpha = 5\%$	$k = 6$	$\mathcal{R} = \{S_{10} \geq 6\}$	*On rejette H_0*
$\alpha = 2\%$	$k = 7$	$\mathcal{R} = \{S_{10} \geq 7\}$	*On rejette H_0*
$\alpha = 1\%$	$k = 8$	$\mathcal{R} = \{S_{10} \geq 8\}$	*On ne rejette pas H_0*

On a observé $S_{10}(\omega) = 7$, on décide que le nouveau traitement est meilleur avec un test de niveau $\alpha = 10\%$ ou $\alpha = 5\%$ ou $\alpha = 2\%$. On décide qu'il n'est pas meilleur si on choisit $\alpha = 1\%$.

Quelle est la valeur de α pour laquelle la décision est modifiée ?

Définition 16

On appelle p-valeur ou probabilité critique le plus petit niveau α^* pour lequel on rejette H_0 au vu des observations.

La $p-$valeur quantifie la marge de sécurité avec laquelle on accepte ou rejette H_0.

$$\begin{cases} \text{Si } \alpha^* \geq \alpha \text{ on ne rejette pas } H_0 \\ \text{Si } \alpha^* < \alpha \text{ on rejette } H_0 \end{cases}$$

Exemple : Pour $S_{10}(\omega) = 7$, la p-valeur vaut $\alpha^* = \mathbb{P}_{0.3}(S_{10} \geq 7) \approx 1.1\%$

Lorsque n est relativement grand, on peut utiliser l'approximation gaussienne de la moyenne empirique par exemple, ce qui donne une zone de rejet très proche de $\mathcal{R}$. Le test est alors asymptotique.

Définition 17

Soit le modèle statistique $(\Omega, \mathcal{F}, (\mathbb{P}_\theta)_{\theta\in\Theta})$ et $\mathcal{R}_n$ une zone de rejet pour le test H_0 contre H_1. Le test est de niveau asymptotique α si

$$\limsup_{n\to\infty} \sup_{\theta\in\Theta_0} \mathbb{P}_\theta(\mathcal{R}_n) = \alpha$$

Exemple de test de niveau asymptotique α

On effectue un sondage sur $n = 900$ personnes. Parmi elles, 486 personnes ont l'intention de voter pour A. Tester au niveau 5% si A aura la majorité absolue dans l'électorat. Soit

$$X_i = \begin{cases} 1 \text{ si le sondé } i \text{ vote pour } A \\ 0 \text{ sinon.} \end{cases}$$

$X_1, \dots, X_{900}$ i.i.d. de loi $\mathcal{B}(\theta)$, θ inconnu $\theta \in \Theta =]0, 1[$.
On effectue un test unilatéral à droite

$$H_0 : \theta \leq 0.5 \text{ (A ne gagne pas) contre } H_1 : \theta > 0.5 \text{ (A gagner).}$$

Statistique de test

$S_n = \sum_{i=1}^{n} X_i$ nombre de sondés qui votent pour A, $S_n \rightsquigarrow \mathcal{B}(n, \theta)$ avec θ inconnu et $n = 900$.

Exemple de test de niveau asymptotique α

Sous H_1, S_n a tendance à prendre des valeurs plus élevées
Sous H_0, $S_n \rightsquigarrow \mathcal{B}(n, \theta)$ avec $\theta \leq 0.5$.

Règle de décision et zone de rejet

- Si $S_n \geq k$ on rejette H_0.
- Si $S_n < k$, on ne rejette pas H_0.

On cherche k tel que

- $\mathbb{P}_\theta(S_n \geq k) \leq \alpha = 5\%$ pour tout $\theta \leq 0.5$.
- le plus petit possible

La zone de rejet est de la forme $\mathcal{R} = \{S_n \geq k\} = \{\overline{X}_n \geq k'\}$.

Principe des tests statistiques

Démarche d'un test statistique

- ☛ Choix de H_0 et de H_1
- ☛ Choix d'un risque α.
- ☛ Choix d'une statistique de test S_T et de sa loi sous H_0
- ☛ Détermination de la région critique ou région de rejet $\mathcal{R}$
- ☛ Conclusion : observation de la réalisation de S_T sur l'échantillon :
  - Si $\omega \in \mathcal{R}$ alors Rejet de H_0
  - Si $\omega \notin \mathcal{R}$ alors Non rejet de H_0

Types de tests :

- ☛ Tests paramétriques : comparaison de paramètres (moyennes, proportions, variances, ...)
- ☛ Tests semi et non-paramétriques : comparaison de distributions

Remarque 5

Les hypothèses nulle et alternative ne jouent pas des rôles symétriques. Il faut montrer que H_0 est peu probable pour décider H_1.

Remarque 6

Pour un α fixé, on peut construire différentes régions critiques et donc différents tests. Le meilleur des tests à α fixé est celui qui minimise le risque β (maximise la puissance) ! Tout test de région critique $\mathcal{R}' \subseteq \mathcal{R}$ sera de niveau α.