

Statistiques mathématiques : Cours 6

Guillaume Lécué

21 septembre 2016

Aujourd'hui

Introduction aux tests

Hypothèse simple contre alternative simple

Lemme de Neyman-Pearson

Tests gaussiens

Tests sur la moyenne

Tests sur la variance

Exemple introductif

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

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

Question : La pièce est-elle équilibrée ?

- ▶ Répondre à cette question revient à prendre une décision :

$\varphi = \varphi(P, P, F, F, P, F, P, P, F, P)$

$= \begin{cases} \text{on accepte l'hypothèse « la pièce est équilibrée »} \\ \text{on rejette l'hypothèse « la pièce est équilibrée »} \end{cases}$

Exemple introductif : modélisation et définition des hypothèses

- ▶ Modélisation : On modélise ces observations par l'expérience statistique

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

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

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

- ▶ et on "traduit" la question en termes mathématiques : résoudre le problème de test suivant

$$H_0 : \theta = \frac{1}{2} \quad \text{contre} \quad H_1 : \theta \neq \frac{1}{2}$$

Définition

L'hypothèse H_0 est appelée *hypothèse nulle*.

L'hypothèse H_1 est appelée *hypothèse alternative*.

Définitions

Définition

1. Soit Z une observation de l'expérience statistique $\mathcal{E} = (\mathfrak{Z}, \mathcal{Z}, \{\mathbb{P}_\theta : \theta \in \Theta\})$.
2. Soit $\Theta = \Theta_0 \cup \Theta_1$ une partition de l'espace des paramètres.
3. Soit le problème de test

$$H_0 : \theta \in \Theta_0 \quad \text{contre} \quad H_1 : \theta \in \Theta_1$$

Un *test* ou *régle de décision* est une statistique de la forme

$$\varphi(Z) = I(Z \in \mathcal{R}) = \begin{cases} H_0 & \text{"on accepte } H_0\text{"} \\ H_1 & \text{"on rejette } H_0\text{"} \end{cases}$$

$\mathcal{R} \subset \mathfrak{Z}$ est appelée *zone de rejet* ou *région critique*.

Exemple introductif : construction d'un test (1/2)

- ▶ $Z = (X_1, \dots, X_{10})$: observation dans le modèle d'échantillonnage de Bernoulli $\{\mathbb{P}_\theta : 0 < \theta < 1\}$ où $\mathbb{P}_\theta = \theta\delta_1 + (1 - \theta)\delta_0$
- ▶ on regarde le problème de test suivant :

$$H_0 : \theta = \frac{1}{2} \quad \text{contre} \quad H_1 : \theta \neq \frac{1}{2}$$

- ▶ on construit un **test** à partir de l'EMV $\hat{\theta}_n^{\text{mv}} = \bar{X}_n = \bar{X}_{10}$ ($n = 10$) et d'un seuil t_0 donné :

$$\varphi(Z) = I(Z \in \mathcal{R}) = \begin{cases} H_0 & \text{quand } \left| \hat{\theta}_n^{\text{mv}} - \frac{1}{2} \right| \leq t_0 \\ H_1 & \text{sinon} \end{cases}$$

- ▶ la **zone de rejet** est ici

$$\mathcal{R} = \left\{ z \in \{0, 1\}^{10} : \left| \hat{\theta}_n^{\text{mv}}(z) - \frac{1}{2} \right| > t_0 \right\}$$

Exemple introductif : construction d'un test (2/2)

- ▶ Dans le test précédent

$$\varphi(Z) = I(Z \in \mathcal{R}) = \begin{cases} H_0 & \text{quand } \left| \hat{\theta}_n^{\text{mv}} - \frac{1}{2} \right| \leq t_0 \\ H_1 & \text{sinon} \end{cases}$$

l'EMV $\hat{\theta}_n^{\text{mv}}$ a été utilisé pour construire le test. C'est une approche classique : ici l'EMV joue le rôle de **statistique de test**

- ▶ On peut calculer l'EMV sur les données : $\hat{\theta}_n^{\text{mv}} \stackrel{\text{exemple}}{=} 0,6$ et donc prendre une décision.

Question : comment choisir le seuil t_0 ?

- ▶ y-a-t'il un meilleur choix de test ? une meilleure statistique de test ?

Les deux types d'erreurs de décision

Définition

Soit φ un test de zone de rejet \mathcal{R} (càd $\varphi(z) = I(z \in \mathcal{R})$)

- ▶ L'erreur de première espèce (*rejeter à tort*)

$$\sup_{\theta \in \Theta_0} \mathbb{P}_\theta[Z \in \mathcal{R}]$$

- ▶ La fonction d'erreur de seconde espèce (*accepter à tort*)

$$\theta \in \Theta_1 \mapsto \mathbb{P}_\theta[Z \notin \mathcal{R}] = 1 - \pi_\varphi(\theta)$$

où $\pi_\varphi(\theta) = \mathbb{P}_\theta[Z \in \mathcal{R}]$ est la *fonction puissance* du test

Note : "rejeter" = rejeter H_0 ; "accepter" = accepter H_0

Exemple introductif : les deux types d'erreurs

Le test $\varphi(z) = I(|\hat{\theta}_n^{\text{mv}}(z) - 0.5| > t_0)$ peut faire deux types d'erreurs :

- ▶ rejeter à tort :

Rejeter ($\varphi(Z) = H_1$) **alors que** $\theta = \frac{1}{2} \in \Theta_0 = \{\frac{1}{2}\}$

dans ce cas, l'**erreur de première espèce** est

$$\mathbb{P}_{0.5}[|\hat{\theta}_n^{\text{mv}}(Z) - 0.5| > t_0]$$

- ▶ accepter à tort :

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

dans ce cas, la fonction d'**erreur de seconde espèce** est

$$\theta \neq 1/2 \mapsto \mathbb{P}_\theta[|\hat{\theta}_n^{\text{mv}}(Z) - 0.5| \leq t_0]$$

Niveau asymptotique d'un test

Définition

Dans le modèle d'échantillonnage on construit une suite de test (φ_n) (où n est le nombre d'observations). Soit $\alpha \in (0, 1)$. On dit que (φ_n) est de **niveau asymptotique α** quand

$$\forall \theta \in \Theta_0, \quad \limsup_{n \rightarrow \infty} \mathbb{P}_\theta[\varphi_n = H_1] \leq \alpha$$

1. on veut s'assurer qu'asymptotiquement la probabilité de rejeter à tort est moins que α
2. idéalement, on aimerait pouvoir fixer un niveau pour les deux types d'erreurs (1ère et 2-ième) mais c'est en général pas possible. On va donc privilégier un contrôle de l'erreur de 1ère espèce en construisant des tests de niveau asymptotique donné : **introduction d'une dissymétrie**

Exemple introductif : étude asymptotique du test

Sous H_0 :

$$\sqrt{n}(\hat{\theta}_n^{\text{mv}} - 0.5) \xrightarrow{d} \mathcal{N}(0, 1/4)$$

en particulier, pour $g \sim \mathcal{N}(0, 1)$,

$$\mathbb{P}_{0.5}[|\hat{\theta}_n^{\text{mv}} - 0.5| > t_{n,\alpha}] \longrightarrow \mathbb{P}[|g| > q_{1-\alpha/2}] = \alpha$$

où, pour le quantile $q_{1-\alpha/2}$ d'ordre $1 - \alpha/2$ de g

$$t_{n,\alpha} = \frac{q_{1-\alpha/2}}{2\sqrt{n}}$$

Sous H_1 : pour tout $\theta \in \Theta_0^c = (0, 1) - \{0.5\}$,

$$\sqrt{n}|\hat{\theta}_n^{\text{mv}} - 0.5| \xrightarrow{p.s.} +\infty$$

Exemple introductif : prise de décision

Les données d'origines étaient :

$$z = (P, P, F, F, P, F, P, P, F, P)$$

l'EMV vaut donc $\hat{\theta}_{10}^{\text{mv}}(z) = 0.6$ et, pour un $\alpha \in (0, 1)$ donné le test est

$$\varphi(Z) = \begin{cases} H_0 & \text{quand } \left| \hat{\theta}_{10}^{\text{mv}}(Z) - \frac{1}{2} \right| \leq t_{10, \alpha} = \frac{q_{1-\alpha/2}}{2\sqrt{10}} \\ H_1 & \text{sinon} \end{cases}$$

Par exemple :

1. pour $\alpha = 5\%$, on a $q_{1-\alpha/2} \approx 1.96$ et $t_{10, \alpha} \approx 0.31$ alors comme $\left| \hat{\theta}_{10}^{\text{mv}}(z) - \frac{1}{2} \right| = 0.1 \leq t_{10, 5\%}$, on **accepte**
2. pour $\alpha = 10\%$, on a $q_{1-\alpha/2} \approx 1.64$, alors $t_{10, \alpha} \approx 0.26$ alors comme $\left| \hat{\theta}_{10}^{\text{mv}}(z) - \frac{1}{2} \right| = 0.1 \leq t_{10, 10\%}$, on **accepte**

Exemple introductif : introduction de la p -value

Le choix du niveau α est arbitraire. Dans l'exemple précédent, on va

- ▶ accepter tant que $q_{1-\alpha/2} \geq 2\sqrt{10} * \left| \hat{\theta}_{10}^{mv} - \frac{1}{2} \right|$
- ▶ rejeter dès que $q_{1-\alpha/2} \leq 2\sqrt{10} * \left| \hat{\theta}_{10}^{mv} - \frac{1}{2} \right|$

La **valeur limite de α pour laquelle la décision bascule** càd telle que

$$q_{1-\alpha/2} = 2\sqrt{10} * \left| \hat{\theta}_{10}^{mv} - \frac{1}{2} \right|$$

est appelée la **p-value**.

Ici la p-value est donnée par l'équation (en α)

$$q_{1-\alpha/2} = 2\sqrt{10} * 0.1 \approx 0.63$$

càd $\alpha = \text{p-value} \approx 0.525$.

p-value : introduction formelle

Définition

Soit φ_α un test de niveau asymptotique α et de zone de rejet \mathcal{R}_α . On appelle **p-value** du test la statistique

$$p\text{-value}(Z) = \inf(\alpha \in (0, 1) : Z \in \mathcal{R}_\alpha)$$

- ▶ c'est le seuil critique où la décision bascule :

$$\varphi_\alpha(Z) = \begin{cases} H_0 & \text{quand } \alpha \leq p\text{-value}(Z) \\ H_1 & \text{quand } \alpha > p\text{-value}(Z) \end{cases}$$

- ▶ la **p-value** quantifie le niveau de confiance sur l'acceptation (de H_0)

p-value : interprétation

p-value	niveau de confiance sur l'acceptation	décision
$p < 0.01$	très faible	<i>rejet</i> (avec confiance)
$0.01 \leq p < 0.05$	faible	<i>rejet</i>
$0.05 \leq p < 0.1$	fort	<i>acceptation</i>
$0.1 \leq p$	très fort	<i>acceptation</i> (avec confiance)

- ▶ forte p-value : le test ne permet pas de rejeter H_0
- ▶ petite p-value : même si on prend un niveau de test très petit, le test rejettera H_0 (alors qu'on a une forte aversion au risque de 1ère espèce, c-à-d de rejeter à tort)
- ▶ dans l'exemple, on a p-value ≈ 0.525 , on va donc accepter

Signification de l'acceptation

Accepter H_0 ne signifie pas que H_0 est vraie

1. par défaut, on accepte H_0 à moins qu'on apporte une "preuve" que H_0 n'est pas acceptable
2. *Accepter* signifie seulement qu'on n'a pas pu apporter une preuve que H_0 n'est pas acceptable : on préférera dire que le test ne permet pas de rejeter plutôt que "on accepte".
3. une *preuve* est l'observation d'un événement "rare" sous H_0 : "sous H_0 , la statistique de test prend une valeur qui peut être considérée comme rare" est une preuve de rejet
4. la "rareté" d'un événement est fixée par le niveau (asymptotique) α
5. si dans l'exemple, le vrai $\theta = 0.5 + 10^{-10}$, il est fort probable qu'on ne rejette pas alors que H_0 est fausse

Signification du rejet

Seul le rejet est informatif

1. *rejeter* signifie qu'on a apporté la preuve que H_0 ne peut pas être acceptée
2. à un niveau α fixé, on ne peut donc pas accepter puisqu'on a observé un événement qui est rare sous H_0 (déclarer un événement rare dépend du niveau α)
3. on rejettera quand la statistique de test prend une valeur rare (= peu vraisemblable) sous H_0
4. en général, **sous H_0** , on connaît la loi asymptotique de la statistique de test (ex. : $\hat{\theta}_n^{\text{mv}} \sim \mathcal{N}(0.5, (4n)^{-1})$) donc si la valeur prise par cette statistique en les observations : $\hat{\theta}_n^{\text{mv}}(z)$ est peu vraisemblable pour sa loi limite (ex. : $\hat{\theta}_n^{\text{mv}} = 0.9$) alors on rejettera = on aura apporté une preuve que H_0 n'est pas acceptable

Choix des hypothèses : dissymétrie

le choix des hypothèses est important

1. Pour une partition $\Theta = A \cup B$, il n'y a pas équivalence entre les deux problèmes de test

$$H_0 : \theta \in A \text{ contre } H_1 : \theta \in B$$

et

$$H_0 : \theta \in B \text{ contre } H_1 : \theta \in A$$

2. on choisit les hypothèses en fonction de l'intérêt qu'on porte au problème : l'hypothèse H_0 est privilégiée
3. l'hypothèse H_0 est privilégiée car on a décidé de se couvrir contre le risque de 1ère espèce avant le risque de 2-ième espèce : càd, on souhaite éviter avant tout de rejeter à tort et par conséquent, on a tendance à "trop accepter"

Méthodologie pour les tests asymptotiques (1/2)

- a) trouver une **statistique de test** (souvent un estimateur) $\hat{\theta}_n$
- b) tel que sous H_0 , on a une normalité asymptotique du style

$$\sqrt{\frac{n}{v(\hat{\theta}_n)}} (\hat{\theta}_n - \theta_0) \xrightarrow{d} V$$

(ou $v(\theta_0)$ à la place de $v(\hat{\theta}_n)$)

- c) et tel que sous H_1 :

$$\sqrt{\frac{n}{v(\hat{\theta}_n)}} (\hat{\theta}_n - \theta_0) \xrightarrow{p.s.} +\infty$$

(avec ou sans valeurs absolues)

Méthodologie pour les tests asymptotiques (2/2)

- d) on utilise cette statistique pour construire un test de niveau asymptotique $\alpha \in (0, 1)$ en posant

$$\varphi(Z) = \begin{cases} H_0 & \text{quand } (\hat{\theta}_n - \theta_0) \leq \frac{q_{1-\alpha}^V \sqrt{v(\hat{\theta}_n)}}{\sqrt{n}} := t_{n,\alpha} \\ H_1 & \text{sinon} \end{cases}$$

- e) on a, sous H_0 ,

$$\mathbb{P}[\text{rejet}] = \mathbb{P}[\hat{\theta}_n - \theta_0 > t_{n,\alpha}] \longrightarrow \mathbb{P}[V > q_{1-\alpha}^V] = \alpha$$

C'est donc un test de niveau asymptotique α .

Choix de tests : notion d'optimalité pour les tests

Hypothèse simple contre alternative simple

- ▶ Cas où $\Theta = \{\theta_0, \theta_1\}$ avec $\theta_0 \neq \theta_1$ et

$$\boxed{\Theta_0 = \{\theta_0\} \text{ contre } \Theta_1 = \{\theta_1\}}$$

- ▶ Existe-t-il un test φ^* **optimal**, au sens où : $\forall \varphi$ test, on a **simultanément** un meilleur contrôle sur les deux erreurs (1ère et 2-ième espèce)

$$\begin{cases} \mathbb{P}_{\theta_0} [\varphi^* = H_1] \leq \mathbb{P}_{\theta_0} [\varphi = H_1] & \text{"proba de rejet à tort"} \\ \mathbb{P}_{\theta_1} [\varphi^* = H_0] \leq \mathbb{P}_{\theta_1} [\varphi = H_0] & \text{"proba d'accepter à tort"} \end{cases}$$

- ▶ Si \mathbb{P}_{θ_0} et \mathbb{P}_{θ_1} ne sont pas **étrangères** (cf. Cours 5) un tel test φ^* **ne peut pas exister**.

Riposte : principe de Neyman (1/2)

- ▶ On « **dissymétrise** » 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** :
on souhaite éviter avant tout de rejeter à tort

Définition

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 α** si

$$\sup_{\theta \in \Theta_0} \mathbb{P}_\theta [\varphi_\alpha = H_1] \leq \alpha$$

- ▶ Un test de niveau α ne dit **rien** sur la fonction erreur de seconde espèce (ou la puissance).

Riposte : principe de Neyman (2/2)

Définition

Soit φ un test de zone de rejet \mathcal{R} . La **puissance** du test φ est la fonction $\pi_\varphi : \theta \in \Theta_1 \mapsto \mathbb{P}_\theta[Z \in \mathcal{R}]$ (proba de rejeter à raison)

- ▶ **Principe de Neyman** : $\alpha \in (0, 1)$, parmi les tests de niveau α , chercher celui (ou ceux) qui sont **les plus puissants**

Définition

Un test de niveau α est dit **Uniformément Plus Puissant (UPP)** si sa puissance est maximale parmi celles des tests de niveau α :

1. $\sup_{\theta \in \Theta_0} \mathbb{P}_\theta [\varphi^* = H_1] \leq \alpha$
2. et si φ est tel que $\sup_{\theta \in \Theta_0} \mathbb{P}_\theta [\varphi = H_1] \leq \alpha$ alors $\forall \theta \in \Theta_1, \pi_{\varphi^*}(\theta) \geq \pi_\varphi(\theta)$

Test de Neyman-Pearson (test du rapport de vraisemblance)

Pour le cas d'une **hypothèse nulle simple** (càd $\Theta_0 = \{\theta_0\}$) contre une **hypothèse alternative simple** (càd $\Theta_1 = \{\theta_1\}$), un test UPP existe : c'est le test de Neyman-Pearson (ou test du rapport de vraisemblance) dont la construction est comme suit :

- ▶ $f(\theta, z) = \frac{d\mathbb{P}_\theta}{d\mu}(z)$, $z \in \mathfrak{Z}$, $\theta \in \{\theta_0, \theta_1\}$, μ mesure dominante
- ▶ On choisit une **région critique** de la forme

$$\mathcal{R}(c) = \{z \in \mathfrak{Z} : f(\theta_1, z) > cf(\theta_0, z)\}, \quad c > 0$$

et on **calibre** $c = t_{n,\alpha}$ de sorte que

$$\mathbb{P}_{\theta_0} [Z \in \mathcal{R}(t_{n,\alpha})] = \alpha$$

- ▶ Le test ainsi construit (si cette équation admet une solution) **est de niveau α** . On **montre** qu'il est UPP.

Lemme de Neyman-Pearson

Proposition

Soit $\alpha \in [0, 1]$. S'il existe $t_{n,\alpha}$ solution de

$$\mathbb{P}_{\theta_0} [f(\theta_1, Z) > t_{n,\alpha} f(\theta_0, Z)] = \alpha$$

alors le test de région critique

$$\mathcal{R}_\alpha = \{z : f(\theta_1, z) > t_{n,\alpha} f(\theta_0, z)\}$$

est de niveau α et **UPP** pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1$.

- ▶ Si $U = f(\theta_1, Z)/f(\theta_0, Z)$ est bien définie et $\mathbb{P}_U \ll \lambda$ (sous \mathbb{P}_{θ_0}), alors $\mathbb{P}_{\theta_0} [U > t_{n,\alpha}] = \alpha$ admet une solution.

Exemple de test de Neyman-Pearson (1/4)

On observe $Z = (X_1, \dots, X_n) \stackrel{i.i.d.}{\sim} \mathcal{N}(\theta, 1)$ et pour $\theta_0 < \theta_1$, on considère le problème de test

$$H_0 : \theta = \theta_0 \text{ contre } H_1 : \theta = \theta_1$$

La vraisemblance en θ est

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

Le rapport de vraisemblance est

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

Exemple de test de Neyman-Pearson (2/4)

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

$$\begin{aligned}\mathcal{R}(c) &= \{z \in \mathbb{R}^n : f(\theta_1, z) > cf(\theta_0, z)\} \\ &= \{(x_1, \dots, x_n)^\top \in \mathbb{R}^n : n(\theta_1 - \theta_0)\bar{x}_n - \frac{n}{2}(\theta_1^2 - \theta_0^2) > \log c\} \\ &= \{(x_1, \dots, x_n)^\top \in \mathbb{R}^n : \bar{x}_n > \frac{\theta_0 + \theta_1}{2} + \frac{\log c}{n(\theta_1 - \theta_0)}\}\end{aligned}$$

- ▶ **Choix de c** : on résout

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

Sous \mathbb{P}_{θ_0} :

$$\bar{X}_n \sim \mathcal{N}\left(\theta_0, \frac{1}{n}\right)$$

Exemple de test de Neyman-Pearson (3/4)

Résoudre en c : pour $g \sim \mathcal{N}(0, 1)$,

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

càd $\mathbb{P} \left[g > \frac{\sqrt{n}}{2}(\theta_1 - \theta_0) + \frac{1}{\sqrt{n}} \frac{\log c}{\theta_1 - \theta_0} \right] = \alpha$, soit

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

où $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ d'une $\mathcal{N}(0, 1)$

- **Conclusion** : le test de NP de niveau α a pour zone de rejet $\mathcal{R}(c_\alpha)$
où

$$c_\alpha = \exp \left(\sqrt{n}(\theta_1 - \theta_0)q_{1-\alpha} - \frac{n(\theta_1 - \theta_0)^2}{2} \right)$$

Exemple de test de Neyman-Pearson (4/4)

On voit que le test de NP s'écrit sous la forme :

$$\varphi(Z) = \begin{cases} H_0 & \text{quand } \bar{X}_n \leq \theta_0 + t_{n,\alpha} \\ H_1 & \text{sinon} \end{cases} \quad \text{où } t_{n,\alpha} = \frac{q_{1-\alpha}}{\sqrt{n}}$$

rem. : la valeur θ_1 n'intervient pas dans le test de NP.

- ▶ la **puissance** du test est ici :

$$\pi_\varphi(\theta_1) = \mathbb{P}_{\theta_1}[\bar{X}_n > \theta_0 + t_{n,\alpha}] = \mathbb{P}[g > \sqrt{n}(\theta_0 - \theta_1) + q_{1-\alpha}]$$

car sous \mathbb{P}_{θ_1} , $\bar{X}_n \sim \mathcal{N}(\theta_1, 1/n)$.

rem. : La puissance augmente quand n augmente et quand $|\theta_0 - \theta_1|$ augmente. L'alternative n'intervient que dans la puissance.

Tests classiques dans le modèle d'échantillonnage gaussien

Test sur la moyenne à variance connue

On observe $Z = (X_1, \dots, X_n) \sim \mathcal{N}(\mu, \sigma^2 \text{Id}_n)$ où σ est connue. On considère le problème de test

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

Principe on estime μ et on rejette H_0 si l'estimateur est « plus grand » que μ_0 . On considère des tests de la forme

$$\varphi_\alpha(Z) = \begin{cases} H_0 & \text{si } \bar{X}_n \geq \mu_0 + t_{n,\alpha} \\ H_1 & \text{sinon} \end{cases}$$

On choisit le **seuil** $t_{n,\alpha}$ tel que

$$\sup_{\mu \leq \mu_0} \mathbb{P}_\mu [\varphi_\alpha(Z) = H_1] = \alpha$$

Détermination de $t_{n,\alpha}$

Majoration de l'erreur de première espèce. Soit $\mu \leq \mu_0$. Sous \mathbb{P}_μ , $\bar{X}_n \sim \mathcal{N}(\mu, \sigma^2/n)$, alors pour $g \sim \mathcal{N}(0, 1)$

$$\begin{aligned}\mathbb{P}_\mu [\bar{X}_n - \mu_0 \geq t_{n,\alpha}] &= \mathbb{P} \left[\left(\mu + \frac{\sigma}{\sqrt{n}} g \right) - \mu_0 \geq t_{n,\alpha} \right] \\ &= \mathbb{P} \left[\frac{\sigma}{\sqrt{n}} g \geq t_{n,\alpha} + (\mu_0 - \mu) \right] \\ &\leq \mathbb{P} \left[\frac{\sigma}{\sqrt{n}} g \geq t_{n,\alpha} \right] \text{ on veut } \alpha\end{aligned}$$

On prend

$$t_{n,\alpha} = \frac{\sigma q_{1-\alpha}}{\sqrt{n}}$$

En particulier, on a :

$$\sup_{\mu \leq \mu_0} \mathbb{P}_\mu [\varphi_\alpha(Z) = H_1] = \mathbb{P}_{\mu_0} [\varphi_\alpha(Z) = H_1]$$

Calcul de la puissance du test

Soit $\mu > \mu_0$. Sous \mathbb{P}_μ , la loi de \bar{X}_n est $\mathcal{N}(\mu, \sigma^2/n)$ alors la **fonction de puissance** du test est

$$\begin{aligned}\mu \in (\mu_0, +\infty) &\mapsto \mathbb{P}_\mu [\bar{X}_n - \mu_0 \geq t_{\alpha, n}] \\ &= \mathbb{P} \left[g \geq \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} + q_{1-\alpha} \right]\end{aligned}$$

Rem. :

- ▶ la puissance tend vers 1 quand n tend vers $+\infty$,
- ▶ c'est un test UPP ; il faut **d'autres outils** pour le montrer.

Test sur la moyenne à variance inconnue

- ▶ **Ingrédient principal** :

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

alors

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

et

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

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

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

Tests sur la moyenne

- ▶ On teste $H_0 : \mu \leq \mu_0$ contre $H_1 : \mu > \mu_0$. Un test de niveau α : donné par la zone de rejet

$$\mathcal{R}_\alpha = \{z \in \mathbb{R}^n : T(z) > q_{1-\alpha, n-1}^{\mathfrak{T}}\}$$

où

$$T(Z) = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{S_n}$$

et $q_{1-\alpha, n-1}^{\mathfrak{T}}$ = quantile d'ordre $1 - \alpha$ de la loi de Student à $n - 1$ degrés de liberté :

$$\mathbb{P} [\text{Student}_{n-1} > q_{1-\alpha, n-1}^{\mathfrak{T}}] = \alpha$$

- ▶ On teste $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$. Un test de niveau α : donné par $\mathcal{R}_\alpha = \{z \in \mathbb{R}^n : |T(z)| > q_{1-\alpha/2, n-1}^{\mathfrak{T}}\}$.

Test sur la variance

- ▶ On teste $H_0 : \sigma^2 \leq \sigma_0^2$ contre $H_1 : \sigma^2 > \sigma_0^2$. Un test de niveau α : donné par la zone de rejet

$$\mathcal{R}_\alpha = \{z \in \mathbb{R}^n : V(z) > q_{1-\alpha, n-1}^{\chi^2}\},$$

où

$$V(Z) = \frac{1}{\sigma_0^2} \sum_{i=1}^n (X_i - \bar{X})^2$$

et

$$\mathbb{P} [\text{Chi-deux}_{n-1} > q_{1-\alpha, n-1}^{\chi^2}] = \alpha.$$

- ▶ **Mêmes remarques méthodologiques** sur l'optimalité de ces tests que précédemment.

Introduction de la p -value dans le cas Gaussien

Exemple : on observe

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

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 α ?**. 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 » α ... **comment ?**
 - ▶ On ne veut pas α trop grand (trop de risque), mais en prenant α 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).

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

Définition

Soit \mathcal{R}_α une famille de zones de rejet d'un test de niveau α pour une hypothèse H_0 contre une alternative H_1 . Soit Z l'observation associée à l'expérience. On a $Z \in \mathfrak{J}$ et $\mathcal{R}_0 = \mathfrak{J}$. 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**).
- ▶ la p -value mesure la rareté de ce qu'on a observé sous H_0 . Quand la p -value est petite on va rejeter H_0 avec confiance sinon, accepter H_0 dans ce cas, reviendrait à admettre qu'on a observé quelque chose de rare.