Chapitre 8.
Tests d'hypothèses paramétriques

Introduction

Les tests d'hypothèses constituent un aspect important de la statistique inférentielle dont le principe général peut s'énoncer comme suit :

— Soit un caractère d'une population dont la valeur du **paramètre** (moyenne, proportion, variance, etc.) est **inconnue**. Une hypothèse dite **nulle ou H_0** est alors formulée sur ce paramètre inconnu, elle résulte de considérations théoriques, pratiques ou encore plus simplement basée sur un pressentiment.

— On veut **porter un jugement** sur cette hypothèse nulle H_0, sur la base des résultats d'un échantillon prélevé de cette population.

— Pour **décider** si l'hypothèse nulle formulée est supportée ou non par les observations, il faut une méthode qui permettra de conclure si l'écart observé entre **la valeur de la statistique** obtenue à partir de l'échantillon et **celle du paramètre spécifiée** dans l'hypothèse est trop important pour être uniquement imputable au fluctuations d'échantillonnage.

— La construction d'un test d'hypothèse consiste alors à déterminer entre quelles valeurs peut varier la statistique en **supposant vraie l'hypothèse nulle H_0** formulée.

— Les distributions d'échantillonnage des paramètres - **statistiques de test** que nous avons traitées dans le chapitre précédent vont être particulièrement utiles dans l'élaboration d'un test statistique.

— Concepts importants dans l'élaboration d'un test statistique :

1. **Hypothèse statistique** : est un énoncé (une affirmation) concernant les caractéristiques (valeurs des paramètres, forme de la distribution des observations) d'une population.

2. **Test d'hypothèse (ou test statistique)** : est une démarche qui à pour but de fournir une **règle de décision** permettant, sur la base de résultats d'échantillon, de faire un choix : **rejeter ou non-rejeter l'hypothèse nulle H_0**.

8.1 Notions générales sur les tests statistiques

1. **Hypothèse nulle** H_0 : hypothèse selon laquelle on fixe *a priori* **un paramètre inconnu** de la population à une valeur particulière.

2. **Hypothèse alternative** H_1 **(ou contre-hypothèse)** : n'importe quelle autre hypothèse qui diffère de l'hypothèse nulle H_0.

— **Remarque :** c'est toujours l'hypothèse nulle H_0 qui est soumise au test et toute la démarche du test s'effectue en considérant ou supposant cette hypothèse H_0 **comme vraie**.

— Conclusion du test : on doit soit **Rejeter** ou **Non-Rejeter** l'hypothèse nulle H_0, en aucun cas l'acceptée. Si H_0 est rejetée, on accepte l'hypothèse alternative H_1.

— **Seuil de signification** d'un test statistique : c'est le risque d'erreur α, consenti à l'avance de **rejeter à tort l'hypothèse nulle** H_0 alors qu'elle est vraie (favoriser alors l'hypothèse H_1) :

— Au seuil de signification correspond sur la distribution d'échantillonnage de la statistique de test, à la **région critique** de **Rejet** H_0 dont l'aire correspond à la probabilité α.

— Sur la distribution d'échantillonnage correspondra aussi une région complémentaire, dite **région d'acceptation** ou encore région de **Non-rejet de** H_0 dont la probabilité est égale au niveau de confiance $1-\alpha$.

— La valeur observée de la statistique de test **toujours calculée sous** H_0 appartient, soit à la région de Rejet H_0 (on favorisera alors l'hypothèse H_1), soit à la région de Non-rejet de H_0 (on favorisera alors l'hypothèse H_0).

— Formulation des hypothèses - Types de tests : à partir d'exemples, nous allons résumer les différents types de test **« Bilatéral » ou « Unilatéral »** qui peuvent se présenter et schématiser les régions de rejet et de non-rejet de l'hypothèse nulle H_0.

Remarques importantes :

1. Pour un test bilatéral, les **2 valeurs critiques** des tables statistiques sont des limites de la statistique qui conduisent au rejet de H_0, selon le seuil de signification α choisi.

2. Un test unilatéral « risque à droite fg ou « risque à gauche fg ne comporte qu'**une seule valeur critique**.

3. Quelle que soit le type de test, l'hypothèse nulle H_0 doit **toujours** comporter **le signe égal** ($\leq$, $=$, $\geq$), elle **spécifie** la valeur du paramètre.
4. l'hypothèse alternative H_1 est formulée en choisissant l'une ou l'autre des trois formes ($<$, $\neq$, $>$). On choisira la plus **pertinente** à la situation pratique analysée.
5. Dans la plupart des tests d'hypothèses, l'inégalité dans l'hypothèse H_1 dénote dans quelle **direction** est localisée la **région de rejet (critique)** de l'hypothèse H_0.

Démarche d'un test statistique :

— Les principales étapes à suivre dans l'élaboration d'un test statistique :

1. Hypothèses statistiques,
2. Seuil de signification et conditions d'application du test,
3. La statistique de test qui convient,
4. Calcul de la statistique sous H_0
5. Règle de décision et conclusion.

8.2. Risques d'erreur, Niveau et Puissance d'un test statistique

— La règle de décision d'un test statistique comporte **2 risques** ou types d'erreur possibles. Ces deux risques varient en sens inverse : quand l'un augmente, l'autre diminue :

1. Risque de **première espèce** α : c'est le seuil de signification α ; risque d'erreur de rejeter à tort l'hypothèse nulle H_0 lorsque celle-ci est vraie :
$$\alpha = P(Rejet\, H_0\,/\,H_0\, vraie) = P(Choisir\, H_1/H_0 \text{ vraie})$$
2. **Niveau de confiance** $(1-\alpha)$: c'est la probabilité complémentaire du risque de $1^{ère}$ espèce α : la probabilité de ne pas rejeter l'hypothèse nulle H_0 alors que l'hypothèse H_0 est vraie :
$$1-\alpha = P(Non-Rejet\, H_0\,/\,H_0 \text{ vraie}) = \text{P(Choisir } H_0/H_0 \text{ vraie)}$$
3. Risque de **deuxième espèce** β : c'est le risque d'erreur de ne pas rejeter l'hypothèse nulle H_0 alors que l'hypothèse H_1 est vraie :
$$\beta = P(Non Rejet\, H_0\,/\,H_1\, vraie) = P(Choisir\, H_0/H_1 \text{ vraie})$$
4. **Puissance du test** $(1-\beta)$: c'est la probabilité complémentaire du risque de $2^{ème}$ espèce β : la probabilité de rejeter l'hypothèse nulle H_0 alors que l'hypothèse H_1 est vraie :
$$1-\beta = P(Rejet\, H_0\,/\,H_1\, vraie) = P(Choisir\, \mathrm{H}_1 \,/\, H_1 \text{ vraie})$$

- Le risque de première espèce α est choisi *a priori* par l'utilisateur.
- Le risque de deuxième espèce β dépend de l'hypothèse alternative H_1 et ne peut se calculer que si une valeur du paramètre est spécifiée dans l'hypothèse H_1, que l'on **suppose vraie**.
- Le graphique de β en fonction des diverses valeurs du paramètre posées en H_1 s'appelle la **courbe d'efficacité du test**.
- Plus le risque de deuxième espèce β **est petit**, plus le **test est puissant**.
- La puissance d'un test mesure sa capacité à séparer au mieux les deux hypothèses H_0 et H_1. Pour un risque de première espèce α donné, un bon test consiste à **minimiser** le risque de deuxième espèce β ou à **maximiser** la puissance $1-\beta$. Un test est satisfaisant si sa puissance est au moins égale à 80%.
- Pour une taille d'échantillon n et un risque α donnés, le **risque β diminue** lorsque **l'écart** entre la valeur du paramètre spécifiée sous H_0 et celle supposée vraie sous H_1 **augmente**.
- Réduire le risque de première espèce de α revient à élargir la zone de non-rejet de H_0 et donc à augmenter le risque de deuxième espèce β. **Lorsqu'un risque diminue, l'autre augmente.**
- Pour un risque α choisi et un écart-type σ déterminé, **augmenter la taille de l'échantillon n** revient à avoir une meilleure précision puisque $\frac{\sigma}{\sqrt{n}}$ diminue. La zone de non-rejet de H_0 sera ainsi plus restreinte, ce qui conduit à **diminuer le risque β** et à **augmenter la puissance** du test.
- Le graphique de $(1-\beta)$ en fonction des diverses valeurs du paramètre posées en H_1 s'appelle la **courbe de puissance du test**.
- Exemple : En contrôle industriel, **le risque α** correspond au risque pris par le **producteur** (ou le fournisseur) alors que le **risque β** correspond au risque pris par le **consommateur** (ou le client).

  Les risques liés aux tests d'hypothèses peuvent se résumer comme suit :

	Réalité	
Décision du test	H_0 vraie	H_1 vraie
Non-rejet H_0	bonne : $1-\alpha$	mauvaise : β
Rejet de H_0	mauvaise : α	bonne : $1-\beta$

Exemple d'application - Calcul des risques d'erreur

Un procédé de remplissage est ajusté de telle sorte que les contenants pèsent en moyenne 400g. Le poids des contenants est supposé normalement distribué avec un écart-type de 8g. On a prélevé un échantillon de 16 contenants pour vérifier si le procédé de remplissage se maintient en moyenne à 400g, on opte pour la règle de décision suivante :

$$\begin{cases} Si\ 396,08g \leq \overline{X}_{16} \leq 403,92g & \text{le processsus opère correctement} \\ Sinon & \text{arrêter le processus de remplissage} \end{cases}$$

1. Quelles sont les hypothèses statistiques que l'on souhaite tester avec cette méthode de contrôle ?
2. Déterminer la probabilité de commettre une erreur de première espèce.
3. Lors d'un contrôle, on a obtenu, pour un échantillon de 16 contenants, un poids moyen de 395g. Doit-on poursuivre ou arrêter la production ?
4. Quelle est la probabilité de commettre une erreur de deuxième espèce selon l'hypothèse alternative H_1 : m = 394 g ?
5. Avec ce plan de contrôle, quelle est la probabilité de rejeter l'hypothèse selon laquelle le procédé opère à 400g, alors qu'en réalité il opère à 394g ?
6. Faire de même pour les valeurs suivantes sous H_1 : m = 395g, 396g, 397g, 398g, 399g et 400g. Tracer la courbe d'efficacité du test.

Exemple d'application - Calcul des risques d'erreur - Solution

— *Hypothèses statistiques :*

$$\begin{cases} H_0 : m = m_0 = 400g & \textit{La machine fonctionne normalement} \\ H_1 : m \neq m_0 = 400g & \textit{La machine est déréglée} \end{cases}$$

— *Seuil de signification :* $\alpha = 5\%$.

— *Conditions d'application du test : test bilatéral symétrique de la moyenne (σ^2 connue). Petit échantillon de taille $n = 16$, prélevé d'une population normale.*

— *Statistique de test :* $\frac{\overline{X}_n - m}{\sigma/\sqrt{n}} \to N(0\,;\,1)$.

— *Risque de première espèce :*

$$\alpha = P(Rejet\ H_0/H_0\ vraie) = P(|\overline{X}_{16} - m| > 3.92\ /\ H_0\ vraie : m = 400\,g)$$
$$= 1 - P(|\overline{X}_{16} - 400| \leq 3.92) = 1 - P(|\frac{(\overline{X}_{16} - 400)}{\frac{\sigma}{\sqrt{n}}}| \leq \frac{3.92}{\frac{8}{\sqrt{16}}})$$
$$= 1 - P(|U| \leq 1.96) = 1 - P(-1.96 \leq U \leq 1.96) = 1 - (\Phi(1.96) - \Phi(-1.96))$$
$$= 1 - (2\Phi(1.96) - 1) = 2 - 2\Phi(1.96) = 2 - 2 \times 0.975 = 5\%.$$

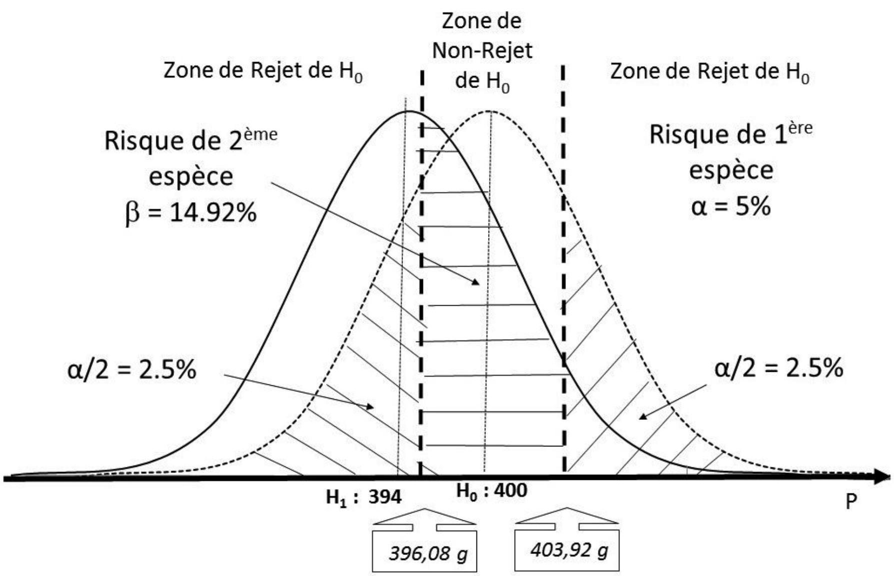
— *Doit-on poursuivre ou arrêter la production ?*
$\bar{x}_{16} = 395 \notin [396.08\,g\,;\,403.92\,g]$ *: intervalle de confiance de niveau* 95% *de m.* $\bar{x}_{16} = 395$ *g appartient à la zone de Rejet de l'hypothèse nulle* H_0*. On doit donc arrêter le processus de remplissage et le réajuster.*

— *Risque de deuxième espèce :*

$$\begin{aligned}\beta &= P(Non-rejet\,H_0/H_1\,vraie)\\ &= P(396.08 \leq \overline{X}_{16} \leq 403.92 \;/\; H_1\ vraie : m = 394g)\\ &= P[\tfrac{(396.08-394)}{\frac{8}{\sqrt{16}}} \leq U \leq \tfrac{(403.92-394)}{\frac{8}{\sqrt{16}}}] = P[1.04 \leq U \leq 4.96]\\ &= \Phi(4.96) - \Phi(1.04) \simeq 1 - 0.8508 = 14.92\% \quad \textit{cf. table N(0 , 1)}\end{aligned}$$

— *Puissance du test :*
Probabilité de rejeter l'hypothèse nulle H_0 *selon laquelle le procédé opère avec un poids moyen m = 400 g, alors qu'en réalité il opère avec un poids moyen m = 394 g :*

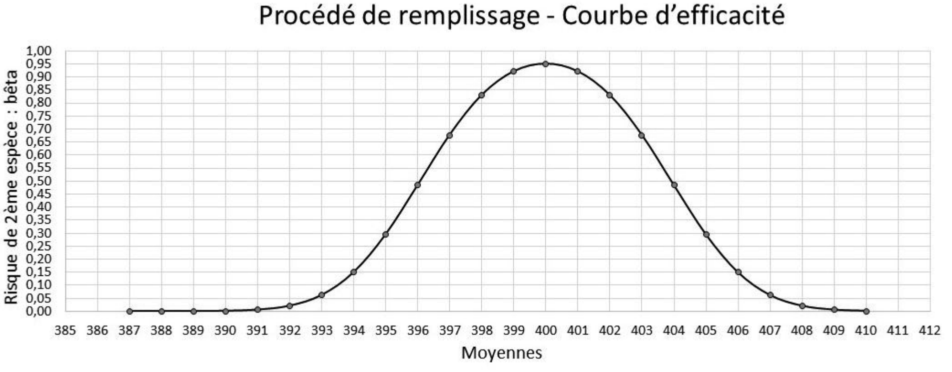
$$1-\beta = P(Rejet\,H_0 \;/\; H_1\ vraie : m = 394\,g) = 1 - 0.1492 = 85.08\%$$

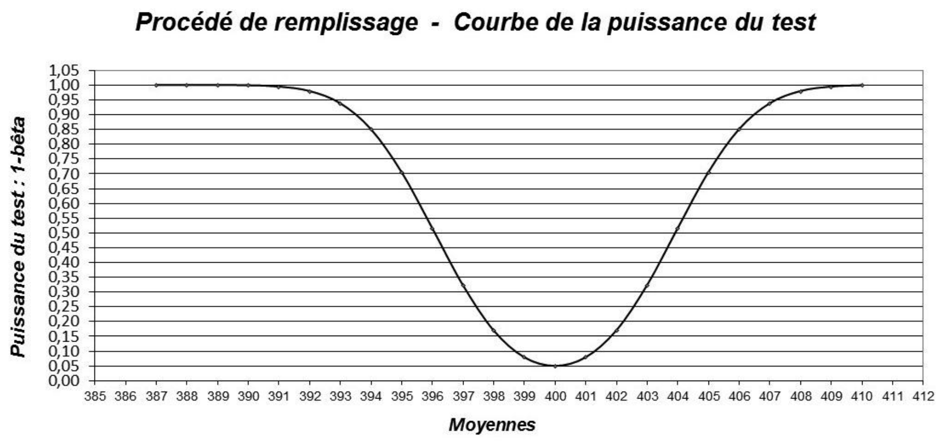


— *Courbe d'efficacité du test : l'erreur* β *de* $2^{\text{ème}}$ *espèce en fonction de la position moyenne du procédé sous* H_1*.*

— *Courbe de puissance du test : la puissance du test* $1-\beta$ *en fonction de la position moyenne du procédé sous* H_1*.*

Moyenne sous H_1 : m	Risque β Bêta	1 - ß Puissance du test	Moyenne sous H_1 : m	Risque β Bêta	1 - ß Puissance du test
387	0,0000	1,0000	399	0,9209	0,0791
388	0,0000	1,0000	400	0,9500	0,0500
389	0,0002	0,9998	401	0,9209	0,0791
390	0,0012	0,9988	402	0,8299	0,1701
391	0,0055	0,9945	403	0,6770	0,3230
392	0,0207	0,9793	404	0,4840	0,5160
393	0,0618	0,9382	405	0,2946	0,7054
394	0,1492	0,8508	406	0,1492	0,8508
395	0,2946	0,7054	407	0,0618	0,9382
396	0,4840	0,5160	408	0,0207	0,9793
397	0,6770	0,3230	409	0,0055	0,9945
398	0,8299	0,1701	410	0,0012	0,9988





8.3. Application des tests d'hypothèses

8.3.1. Tests paramétriques usuels

Exemple d'application 1

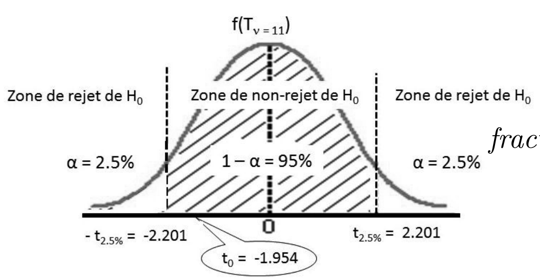
Une entreprise fournit à un client des tiges d'acier. Le client exige que les tiges aient en moyenne, une longueur de 29 mm. La longueur des tiges est supposée normalement distribuée. On veut vérifier si le procédé de fabrication opère bien à 29 mm. Un échantillon aléatoire de 12 tiges provenant de la fabrication donne une longueur moyenne de 27.25 mm et un écart-type empirique de 2.97 mm.

Doit-on conclure, au seuil $\alpha = 5\%$, que la machine est déréglée ?

Exemple d'application 1 - Solution

On s'intéresse à l'égalité ou la différence spécifiée sous H_1 par le signe « $\neq$ », on opte pour un test bilatéral symétrique.

1. *Hypothèses statistiques :*
$$\begin{cases} H_0 : m = m_0 = 29 \ \textit{mm}, & \textit{la machine n'est pas déréglée} \\ H_1 : m \neq m_0 = 29 \ \textit{mm}, & \textit{la machine est déréglée} \end{cases}$$
2. *Seuil de signification et conditions d'application du test : $\alpha = 5\%$, petit échantillon $n = 12$ provenant d'une population normale. Test bilatéral de la longueur moyenne des tiges (variance inconnue) à une moyenne donnée $m_0 = 29$.*
3. *Statistique de test :* $\frac{\overline{X}_n - m}{S_n^* / \sqrt{n}} \to T_{n-1=11 \ d.d.l.}$
4. *Calcul de la statistique de test sous l'hypothèse nulle $H_0 : m = m_0 = 29$*
$$t_0 = \frac{\bar{x}_{12} - m}{s_{12}^* / \sqrt{12}} = \frac{27.25 - 29}{3.10 / \sqrt{12}} = -1.954 \ \ \textit{avec} \ \ s_{12}^* = \sqrt{\frac{12}{11}} s_{12} = \sqrt{\frac{12}{11}} 2.97 = 3.10$$



fractiles de Sudent

$$T_{11} : t_{\frac{\alpha}{2} = 2.5\%} = \pm 2.201.$$