

Module 8.1 - Tests d'hypothèses à un échantillon

Département de mathématiques et de statistique



UNIVERSITÉ
LAVAL

Introduction

Des hypothèses de recherche

Une question de recherche est souvent liée à un paramètre d'une population d'intérêt :

- Le taux moyen de défaillance des circuits imprimés produits par une nouvelle ligne de soudure est inférieur à 2%.
- La durée moyenne d'assemblage d'un composant est réduite après l'introduction d'un nouvel outil automatisé.
- L'écart-type des vibrations mesurées en vol sur une aile équipée d'un nouvel aileron est inférieur à celui mesuré avec l'ancien design.

Hypothèse statistique

Hypothèse statistique

Énoncé sur la loi de probabilité que suit une variable aléatoire.

Exemples :

Une population

- $\mu \neq 10$
- $\mu > 3$
- $\sigma^2 = 4$

Plusieurs populations

- $\mu_1 > \mu_2$
- $\sigma_1^2 \neq \sigma_2^2$
- $\mu_1 = \mu_2 = \mu_3 = \mu_4$

Test d'hypothèses

Dans un **test d'hypothèses**, on confronte deux hypothèses.

Par exemple,

$$H_0 : \mu = 10 \quad \text{contre} \quad H_1 : \mu > 10.$$

H_0 : l'hypothèse nulle. Hypothèse supposée vraie en l'absence de preuve suffisante du contraire.

H_1 : l'hypothèse alternative. Hypothèse soutenue par les données si celles-ci fournissent suffisamment d'éléments pour rejeter H_0 . Elle incarne un changement ou un effet à détecter.

On suppose que H_0 est vraie jusqu'à ce qu'une preuve suffisante (significative) que H_0 n'est pas vraie ait été amassée. La preuve provient des données qui sont observées.

Remarques

- H_0 et H_1 sont toujours mutuellement exclusives.
- H_1 peut avoir une forme bilatérale ($\mu \neq 10$)
ou unilatérale ($\mu < 10$ ou $\mu > 10$)
- C'est l'hypothèse de recherche (ce qui doit être démontré à partir de données) qui dicte le sens de H_1 .

QUIZ!

Vous devez valider qu'une structure de béton a une résistance moyenne supérieure à 80 MPa. Vous prélèverez 10 carottes à analyser.

1. Quelles sont vos hypothèses statistiques?
2. Vous obtenez une moyenne échantillonnale de $\bar{x} = 81,6$ MPa. Pouvez-vous conclure quelque chose? [Illustration](#)

Étapes d'un test d'hypothèses

1. On détermine les hypothèses H_0 et H_1 à confronter.
2. On recueille un échantillon aléatoire.
3. On décide si l'échantillon recueilli à l'étape 2 contient suffisamment de preuves pour rejeter H_0 en faveur de H_1 ou pas.

Il nous faut donc une règle pour prendre la décision à l'étape 3.

Règle de décision

Il n'y a que deux résultats possibles à un test d'hypothèses : **rejeter** H_0 ou **ne pas rejeter** H_0 .

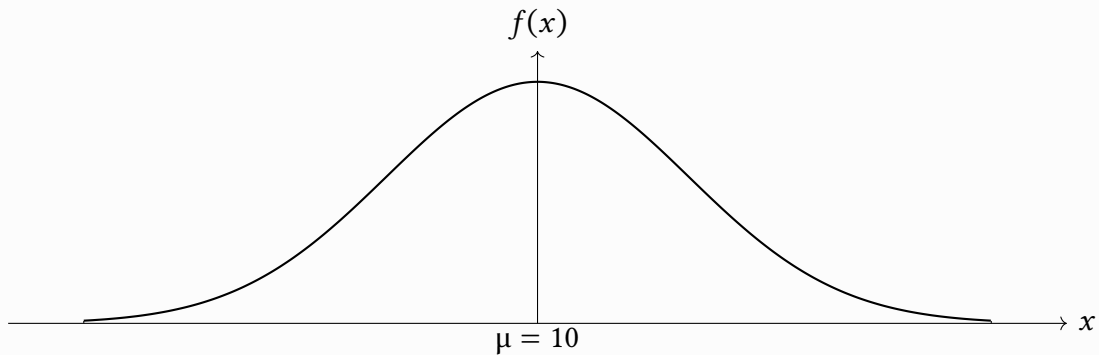
Une règle de décision est formulée avant d'observer les données. Elle prend la forme d'une **région critique** qui définit pour quels échantillons H_0 sera rejetée.

Par exemple, pour

$$H_0 : \mu = 10 \quad \text{contre} \quad H_1 : \mu > 10,$$

on devra trouver à partir de quelle valeur k on décide de rejeter H_0 .

Loi de $\bar{X}$ sous H_0



Deux erreurs possibles à l'issue d'un test d'hypothèses

Décision	Réalité (inconnue!)	
	H_0 est vraie	H_0 est fausse
On rejette H_0	Erreur de type I	
On ne rejette pas H_0		Erreur de type II

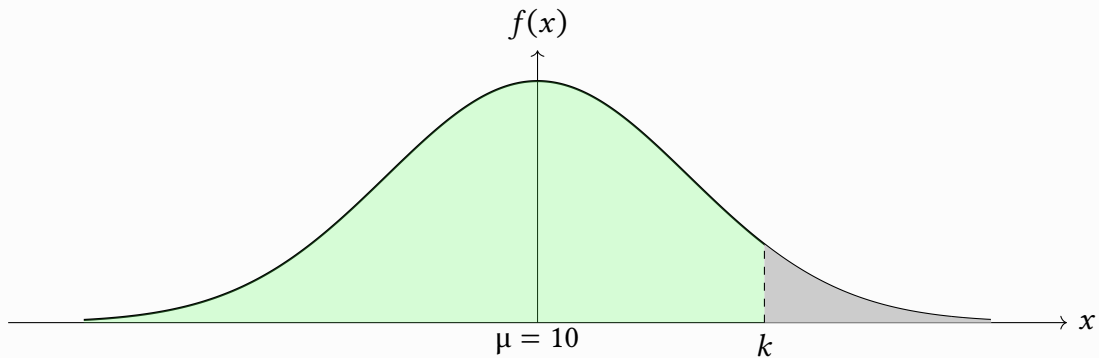
$$\begin{aligned}\alpha &= P(\text{erreur de type I}) \\ &= P(\text{rejeter } H_0 | H_0 \text{ vraie}) \\ &= \text{seuil du test}\end{aligned}$$

ET

$$\begin{aligned}\beta &= P(\text{erreur de type II}) \\ &= P(\text{ne pas rejeter } H_0 | H_0 \text{ fausse}) \\ &= P(\text{ne pas rejeter } H_0 | H_1 \text{ vraie})\end{aligned} \quad \Leftrightarrow$$

$$\begin{aligned}1 - \beta &= P(\text{rejeter } H_0 | H_0 \text{ fausse}) \\ &= P(\text{rejeter } H_0 | H_1 \text{ vraie}) \\ &= \text{puissance du test}\end{aligned}$$

Loi de $\bar{X}$ sous H_0

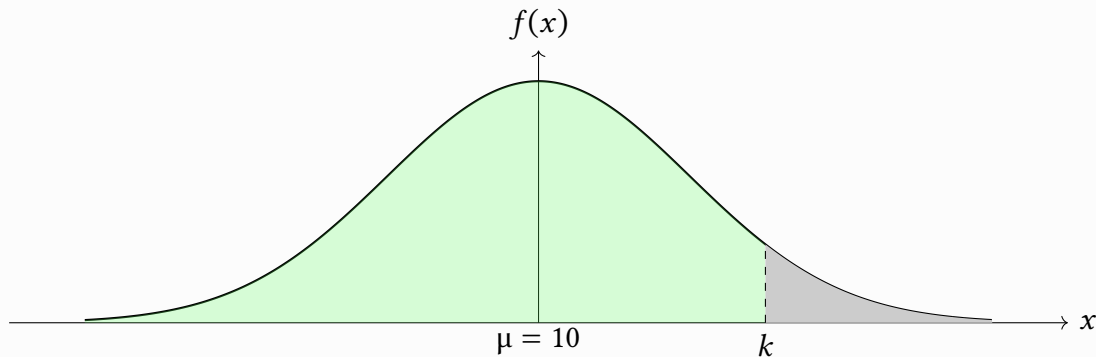


Quelques remarques

- **L'erreur de type I est la plus grave**, donc α est fixé par le chercheur à une faible valeur (souvent $\alpha = 0,05$ ou $0,01$).
- Plus α est petit, plus β est grand...
donc on fixe α et on contrôle souvent β avec la taille d'échantillon.

On doit maintenant trouver la valeur de k qui nous donne la probabilité d'erreur de type I (α) désirée.

Loi de $\bar{X}$ sous H_0



Si nous connaissons la loi de $\bar{X}$ sous H_0 , nous sommes capables d'en déduire la valeur de k qui nous donne la région critique désirée.

Exemple 1

Soit $X_1, \dots, X_n$ qui sont i.i.d. $\mathcal{N}(\mu, \sigma^2)$. On désire confronter les hypothèses :

$$H_0 : \mu = 10$$

$$H_1 : \mu > 10.$$

À partir de quelle valeur de $\bar{X}$ rejettera-t-on H_0 pour avoir un seuil α ?

Exemple 1

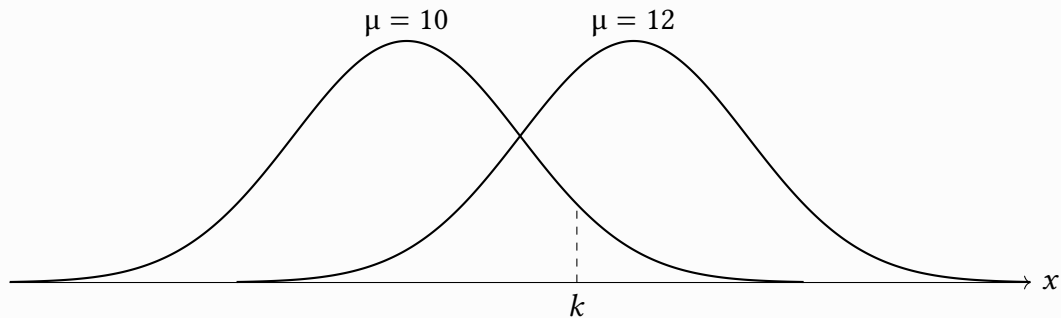
Soit $X_1, \dots, X_n$ qui sont i.i.d. $\mathcal{N}(\mu, \sigma^2)$. On désire confronter les hypothèses :

$$H_0 : \mu = 10$$

$$H_1 : \mu > 10.$$

Si en réalité $\mu = 12$, quelle est la puissance du test ?

Graphiquement



Seuil observé / valeur-p

Le seuil observé d'un test (valeur-p ou p -value)

Seuil observé, ou valeur-p

Le seuil observé d'un test (aussi appelé *valeur-p* ou p -value), noté α^* , est la **plus petite valeur de α** pour laquelle cet échantillon nous permettrait de rejeter H_0 .

α^* peut être vu comme la **probabilité d'observer un autre échantillon "au moins aussi extrême" si H_0 est vraie.**

$$\text{On rejette } H_0 \text{ au seuil } \alpha \iff \alpha^* \leq \alpha.$$

Illustration

Interprétation du seuil observé

α^* peut s'interpréter comme **une mesure de la compatibilité entre H_0 et les observations**, c'est la **significativité statistique** du test.

Plus le seuil observé est **faible, moins** notre échantillon semble plausible sous H_0 .

ATTENTION! Bien qu'on rejette H_0 quand α^* est petit, sa valeur **n'est PAS la probabilité que H_0 soit vraie!**

Exemple 2 : Un exemple pour clarifier les choses

On dispose d'un sac contenant **19 dés équilibrés** et **1 dé pipé** (ce dernier donne un « 6 » avec une probabilité de 80%). Un dé est tiré au hasard dans le sac, mais *on ne sait pas lequel*.

On lance ensuite le dé **cinq fois** de manière indépendante, et on observe **quatre résultats** « 6 ». On souhaite confronter les deux hypothèses suivantes :

H_0 : le dé lancé est équilibré

H_1 : le dé lancé est pipé

Si x désigne le nombre de « 6 » obtenus en cinq lancers indépendants, on peut définir une **région critique** de la forme $x \geq m$ où m est choisi de manière à obtenir le **niveau de signification** souhaité.

Exercice : Calculez α^* . Comparez-le à la probabilité que H_0 soit vraie (que vous pouvez exceptionnellement calculer ici, grâce au théorème de Bayes).

En simulation

4 cas étudiés

La règle de rejet de H_0 doit être basée sur la distribution échantillonnale de notre estimateur ($\bar{X}$ ou S^2). Nous considérerons 4 situations pour les tests d'hypothèses à un échantillon (les mêmes que pour les IC) :

Cas	Hyp. nulle	Situation
1	$H_0 : \mu = \mu_0$	$X_1, \dots, X_n \sim N(\mu, \sigma^2)$ avec σ^2 connue
2		$X_1, \dots, X_n \sim N(\mu, \sigma^2)$ avec σ^2 inconnue
3		$X_1, \dots, X_n \sim$ loi inconnue, mais n est grand
4	$H_0 : \sigma^2 = \sigma_0^2$	$X_1, \dots, X_n \sim N(\mu, \sigma^2)$

Cas 1 : Tests d'hypothèses sur la moyenne d'une distribution normale de variance connue

Cas 1 : Test sur μ , loi normale, σ^2 connue

Supposons un échantillon aléatoire $X_1, \dots, X_n$ issu d'une population $N(\mu, \sigma^2)$ où μ est inconnue mais σ^2 est connue.

Objectif : Tester si l'échantillon contient suffisamment de "preuves" pour rejeter $H_0 : \mu = \mu_0$, où μ_0 est une valeur donnée.

Trois hypothèses alternatives possibles :

$H_1 : \mu \neq \mu_0$ (bilatérale)

$H_1 : \mu < \mu_0$ (unilatérale à gauche)

$H_1 : \mu > \mu_0$ (unilatérale à droite)

Hypothèse alternative unilatérale à droite

Tester $H_0 : \mu = \mu_0$ contre $H_1 : \mu > \mu_0$.

On se doute que si $\bar{x}$ est trop élevée par rapport à μ_0 , c'est que H_0 est probablement fausse.

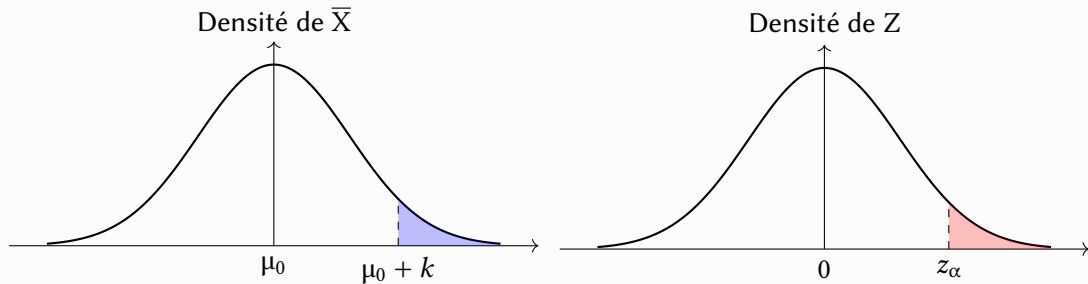
Pour déterminer à partir de quelles valeurs $\bar{x}$ "est trop élevée", on fixe le seuil α à une petite valeur, comme 0,05 ou 0,01.

Comme nous savons comment calculer des probabilités avec la loi normale, il est possible d'utiliser la standardisation pour travailler avec Z .

Hypothèse alternative unilatérale à droite

Autrement dit, on rejette H_0 au seuil α quand

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha.$$



Hypothèses alternatives unilatérales

H_1 unilatérale à droite :

Si on veut tester $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$ alors on rejette H_0 si $\bar{x}$ est suffisamment plus élevée que μ_0 .

Pour obtenir un test de seuil α , on rejette H_0 si $z_0 > z_\alpha$.

H_1 unilatérale à gauche :

De même, si on veut tester $H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$ alors on rejette H_0 pour de faibles valeurs de $\bar{x}$, soit lorsque $z_0 < -z_\alpha$.

On rejette donc quand on se trouve dans le $100 * \alpha\%$ moins probable à droite ou à gauche.

Hypothèse alternative bilatérale

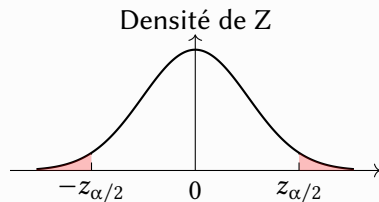
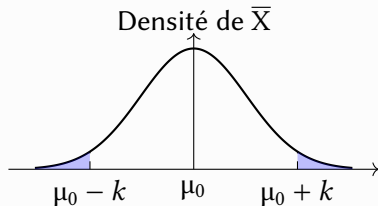
Dans le cas d'un test bilatéral, on voudra rejeter si ce qu'on observe est trop éloigné de μ_0 . C'est-à-dire si ce qu'on observe est trop grand ou trop petit. Autrement dit, on rejette H_0 au seuil α quand

$$Z_0 > z_{\alpha/2}$$

ou

$$Z_0 < -z_{\alpha/2}.$$

On peut résumer cette règle par $|Z_0| > z_{\alpha/2}$



Exemple 3

Une compagnie affirme que le remboursement d'impôt moyen aux étudiants qui ont utilisé son service est d'au moins 1100 \$.

Un échantillon aléatoire de 20 étudiants ayant utilisé le service a obtenu un remboursement moyen de 1060 \$.

En supposant une distribution normale d'écart-type 100 \$ pour le remboursement d'un étudiant, peut-on conclure que la compagnie fait de la fausse publicité, au seuil $\alpha = 0,05$?

1. Hypothèses :

2. Règle de décision :

Exemple 2 (suite)

3. Calcul de la statistique observée :

4. Décision :

L'hypothèse nulle...

5. Conclusion : (dans les termes de l'étude)

Au seuil de 5%, le remboursement moyen ...

QUIZ!

Vous faites un test sur 20 observations pour confronter les hypothèses $H_0 : \mu = 5$ et $H_1 : \mu < 5$.

Vous obtenez un seuil observé (valeur P) de 0,09.

Vrai, faux ou manque d'information ?

1. La moyenne échantillonnale est inférieure à 5.
2. Vous rejetez H_0 au seuil de 5%.
3. Vous rejetez H_0 au seuil de 10%.
4. Vous rejetez $H_0 : \mu = 4$ contre $H_1 : \mu < 4$ au seuil de 5%.
5. Vous rejetez $H_0 : \mu = 6$ contre $H_1 : \mu < 6$ au seuil de 5%.
6. Vous refaites l'expérience avec 40 observations. Vous obtenez la même moyenne échantillonnale. Votre seuil observé est alors plus élevé que 0,09 pour le test $H_0 : \mu = 5$ contre $H_1 : \mu < 5$.

Formules pour le seuil observé

De la définition du seuil observé α^* , on peut facilement déduire les formules générales suivantes (pour le cas 1 : loi normale, σ^2 connue) :

- Pour $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$:
 $\alpha^* = P(|Z| > |z_0|) = 2 P(Z > |z_0|).$
- Pour $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$:
 $\alpha^* = P(Z > z_0).$
- Pour $H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$:
 $\alpha^* = P(Z < z_0).$

Exemple 2 (suite)

Quel est le seuil observé du test réalisé sur la moyenne des remboursements d'impôts ?

Calcul de puissance

- Rappel :

Puissance du test = $P(\text{rejeter } H_0 \mid H_1 \text{ est vraie})$.

- En général, H_1 ne spécifie pas une valeur unique de μ , mais un intervalle de valeurs.
- On calcule la puissance pour une valeur plausible ou arbitraire de μ sous H_1 , que nous noterons μ_1 .

Exemple 4

On pige un échantillon $X_1, \dots, X_{16}$ d'une loi $N(\mu, 64)$ pour tester, au seuil $\alpha = 0,10$:

$$H_0 : \mu = 10$$

$$H_1 : \mu > 10$$

Quelle est la puissance du test si en vérité $\mu = 11$?

Exemple 4 (suite)

Déterminants de la puissance

Trois facteurs importants influencent la puissance d'un test :

1. la différence entre μ_0 et μ_1 : plus μ_1 s'éloigne de μ_0 , plus la puissance augmente.

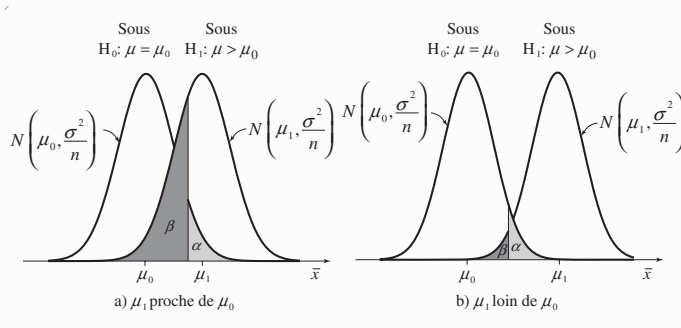


Figure 11.5

L'effet de la distance entre μ_1 et μ_0 sur la valeur de β : a) lorsque μ_1 est proche de μ_0 ; b) lorsque μ_1 est loin de μ_0 .

Déterminants de la puissance

2. le seuil du test : plus α augmente, plus la puissance augmente;

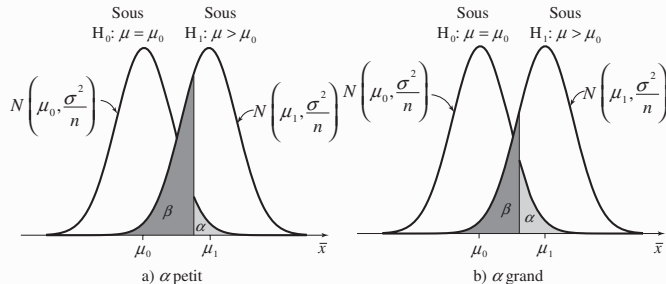


Figure 11.7

L'effet de la valeur de α sur la valeur de β : a) lorsque α est petit; b) lorsque α est grand.

Déterminants de la puissance

3. la taille d'échantillon : plus n augmente, plus la puissance augmente.

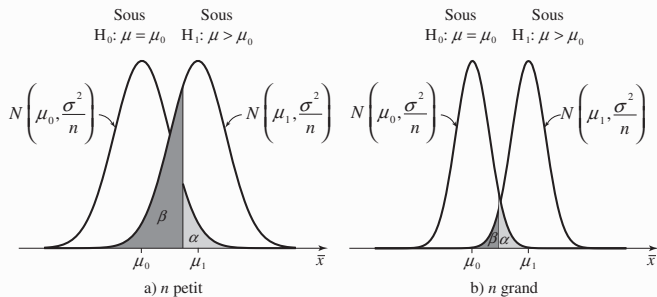


Figure 11.6

L'effet de la taille d'échantillon sur la valeur de β : a) lorsque la taille de l'échantillon (n) est petite; b) lorsque la taille de l'échantillon (n) est grande.

Déterminants de la puissance

1. Le statisticien n'a aucun contrôle sur μ , la vraie valeur de la moyenne. Il peut demander au chercheur la plus petite différence qu'il considère importante.
2. On peut augmenter α pour augmenter la puissance, mais on augmente les chances de faire une erreur de type I.
3. En pratique, la seule option est souvent d'augmenter la puissance en choisissant n suffisamment grand.

Exemple 4 (suite)

On reconsidère notre échantillonnage dans une population $N(\mu, 64)$ pour tester (avec $\alpha = 0,10$)

$$H_0 : \mu = 10$$

$$H_1 : \mu > 10$$

Quelle est la taille d'échantillon nécessaire pour que la puissance soit de 80% si $\mu = 11$?

Cas 2 : Tests d'hypothèses sur la moyenne d'une distribution normale de variance inconnue

Cas 2 : Tests sur μ , loi normale, σ^2 inconnue

Comment faire un test sur μ si la variance est inconnue ?

Le test basé sur Z_0 est inapproprié puisque l'erreur de type I qui serait commise serait supérieure à α .

Si la variance est inconnue, on utilise la statistique

$$T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}.$$

Sous H_0 , on sait que $T_0 \sim t_{n-1}$.

Règles de décision

Au seuil α , on rejette $H_0 : \mu = \mu_0$...

- Test bilatéral ($H_1 : \mu \neq \mu_0$) :

on rejette H_0 si $|t_0| > t_{\alpha/2, n-1}$

$$\alpha^* = P(|t_{n-1}| > |t_0|) = 2 P(t_{n-1} > |t_0|)$$

- Test unilatéral à droite ($H_1 : \mu > \mu_0$) :

on rejette $H_0 : \mu = \mu_0$ si $t_0 > t_{\alpha, n-1}$

$$\alpha^* = P(t_{n-1} > t_0)$$

- Test unilatéral à gauche ($H_1 : \mu < \mu_0$) :

on rejette $H_0 : \mu = \mu_0$ si $t_0 < -t_{\alpha, n-1}$.

$$\alpha^* = P(t_{n-1} < t_0)$$

Seuils observés et puissance

Seuils observés :

Dans tous les cas, on peut énoncer la règle suivante :

On rejette H_0 si $\alpha^* \leq \alpha$.

C'est le calcul de la valeur-P (α^*) qui change selon le test.

Puissance :

Les calculs de puissance pour le cas 2 font appel à une loi *t décentrée* et ne seront donc pas couverts dans ce cours.

Exemple 5

Une fabrique achète régulièrement des transistors à un fournisseur. Les données passées montrent que le coefficient d'amplification de ces transistors avait une distribution normale de moyenne 155.

Lors d'un nouveau contrôle sur 20 transistors, on observe un coefficient d'amplification moyen de 152 et un écart-type de 10. Peut-on admettre, au seuil de 1%, que le coefficient moyen a changé ?

1. Hypothèses :

2. Règle de décision :

Exemple 5 (suite)

3. Calcul de la statistique observée :

4. Décision :

L'hypothèse nulle...

5. Conclusion : (dans les termes de l'étude)

Au seuil de 1%, le coefficient d'amplification moyen ...

Cas 3 : Tests d'hypothèses sur la moyenne d'une distribution inconnue (n grand)

Cas 3 : Tests sur μ , loi inconnue, n grand

Si la loi théorique des observations en cause est inconnue, comment faire un test sur la moyenne ?

Si n est grand (au moins 30), le test présenté au cas 1 peut être utilisé en remplaçant σ par l'écart-type échantillonnal S dans la définition de Z_0 .

Les probabilités d'erreur de type I et II (α et β) seront alors approximatives (car la loi normale est un modèle approximatif pour $\bar{X}$ lorsque n est grand).

Cas 4 : Tests d'hypothèses sur la variance d'une distribution normale

Cas 4 : Tests sur σ^2 , distribution normale

Si on veut contrôler la variabilité d'un procédé, on peut effectuer un test d'hypothèses portant sur la variance d'une distribution.

Ces tests sont relativement simples dans le cas où nous recueillons un échantillon aléatoire $X_1, \dots, X_n$ issu d'une population $N(\mu, \sigma^2)$.

Test bilatéral

On suppose $X_1, \dots, X_n$ un échantillon aléatoire d'une distribution $N(\mu, \sigma^2)$ pour laquelle on veut tester

$$H_0 : \sigma^2 = \sigma_0^2 \text{ contre } H_1 : \sigma^2 \neq \sigma_0^2.$$

On voudra rejeter H_0 lorsque S^2 sera éloignée de σ_0^2 , donc pour des valeurs éloignées de $n - 1$ pour la statistique

$$U_0 = \frac{(n - 1)S^2}{\sigma_0^2}.$$

Si H_0 est vraie, alors $U_0 \sim \chi_{n-1}^2$.

Test de seuil α :

$$\text{on rejette } H_0 \text{ si } u_0 < \chi_{1-\alpha/2; n-1}^2 \text{ ou si } u_0 > \chi_{\alpha/2; n-1}^2.$$

Tests unilatéraux

Au seuil α , on rejette $H_0 : \sigma^2 = \sigma_0^2 \dots$

- Test unilatéral à droite ($H_1 : \sigma^2 > \sigma_0^2$) :

on rejette $H_0 : \sigma^2 = \sigma_0^2$ si $u_0 > \chi_{\alpha;n-1}^2$

$$\alpha^* = P(\chi_{n-1}^2 > u_0)$$

- Test unilatéral à gauche ($H_1 : \sigma^2 < \sigma_0^2$) :

on rejette $H_0 : \sigma^2 = \sigma_0^2$ si $u_0 < \chi_{1-\alpha;n-1}^2$.

$$\alpha^* = P(\chi_{n-1}^2 < u_0)$$

Exemple 6

On évalue une nouvelle machine pour remplir des sacs de sucre. La compagnie est prête à tolérer une variance de 7 g^2 . Un test sur 31 sacs donne une variance échantillonnale de $8,2 \text{ g}^2$.

La variance est-elle trop grande au seuil $\alpha = 0,05$?

1. Hypothèses :

2. Règle de décision :

Exemple 6 (suite)

3. Calcul de la statistique observée :

4. Décision :

L'hypothèse nulle...

5. Conclusion : (dans les termes de l'étude)

Au seuil de 5%, la variance des poids ...

Lien entre intervalles de confiance et tests d'hypothèses bilatéraux

Intervalles de confiance et tests bilatéraux

Il existe une forte relation entre les intervalles de confiance et les tests bilatéraux.

L'examen des régions critiques des tests pour $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ montre que

On rejette $H_0 : \mu = \mu_0$ au seuil α	$\Leftrightarrow$	L'intervalle de confiance pour μ de niveau $1 - \alpha$ ne contient pas la valeur μ_0 .
---	-------------------	---

Idem pour σ^2 .

QUIZ!

À la suite d'une étude, on a calculé un intervalle de confiance à 99% pour μ : $[2, 26; 3, 41]$.

Vrai, faux, ou info insuffisante ?

1. Au seuil de 1%, on rejette $H_0 : \mu = 3$ pour $H_1 : \mu \neq 3$.
2. Au seuil de 1%, on rejette $H_0 : \mu = 4$ pour $H_1 : \mu \neq 4$.
3. Au seuil de 5%, on rejette $H_0 : \mu = 3$ pour $H_1 : \mu \neq 3$.
4. Au seuil de 5%, on rejette $H_0 : \mu = 4$ pour $H_1 : \mu \neq 4$.

Réponses aux exemples

1. $H_0 : \mu = 0,10$ contre $H_1 : \mu > 0,10$. $\alpha_1 > \alpha_2$; $\beta_1 < \beta_2$ $1 - \beta_1 > 1 - \beta_2$. ($\alpha^* = 0,041$)
2. $\alpha^* \approx 0,33\%$, alors que $P(H_0 \text{ est vraie} \mid X = 4) \approx 12,98\%$
3. $H_0 : \mu = 1100$ contre $H_1 : \mu < 1100$. Rejet de H_0 si $z_0 < -1,645$. $z_0 = -1,79$, donc le remboursement moyen est significativement inférieur à 1100 \$. ($\alpha^* = 0,037$)
4. $1 - \beta = 0,218$, $n = 291$
5. $H_0 : \mu = 155$ contre $H_1 : \mu \neq 155$. Rejet de H_0 si $|T_0| > 2,861$. $T_0 = -1,34$, donc le changement de moyenne n'est pas significatif au seuil de 1%. ($\alpha^* = 0,196$)
6. $H_0 : \sigma^2 = 7$ contre $H_1 : \sigma^2 > 7$. Rejet de H_0 si $U_0 > 43,77$. $U_0 = 35,14$, donc la variance n'est pas significativement trop grande au seuil de 5%. ($\alpha^* = 0,238$)