Biostatistique

Tests d'hypothèses

Anicet Ebou, Institut National Polytechnique Félix Houphouët-Boigny,
ediman.ebou@inphb.ci

Plan

Introduction

Introduction

Il s'agit d'une méthode statistique permettant de vérifier, entre autres, la valeur d'un paramètre, la forme d'une distribution, etc. Pour cela, les hypothèses décrivant la situation doivent être formulées et un test statistique est ensuite exécuté.

Exemple 1: On étudie le rendement d'une exploitation laitière. Le cahier des charges de la coopérative exige que la vitesse moyenne de combustion doit de 40 l/j.

Supposons que l'écart-type de ce rendement soit d'environ $\sigma = 2$ l/j.

Le responsable veut vérifier si effectivement la moyenne est de 40 l/j, à partir d'un échantillon de taille $n = 25$ dont la vitesse moyenne de combustion est $\overline{x} = 41,25$ l/j.

Hypothèses et erreurs

Définitions

Une *hypothèse statistique*, noté H, est une affirmation concernant:

1. La valeur d'un paramètre (moyenne, variance, proportion, etc.)
2. L'égalité des paramètres de deux distributions (deux moyennes, deux variances, etc.)
3. La forme d'une distribution.

Remarques:

- Dans les deux premiers cas, on a une *hypothèse paramétrique*.
- Dans le troisième cas, on a une *hypothèse non paramétrique*.

Hypothèses

On suppose que l'on cherche à vérifier la valeur d'un paramètre inconnu θ de la distribution d'une population X.

Pour cela, on compare deux hypothèses portant sur la valeur de θ.

Définition:
On distingue deux types d'hypothèses:

1. L'*hypothèse nulle*: $H_0 = \theta = \theta_0$.
2. L'*hypothèse alternative* (ou *contre hypothèse*) H_1, qui peut prendre l'une des formes suivantes:
   - $H_1 : \theta \neq \theta_0$ (*bilatérale*).
   - $H_1 : \theta < \theta_0$ (*unilatérale à gauche*).
   - $H_1 : \theta > \theta_0$ (*unilatérale à droite*).

Hypothèses (ii)

Remarque: L'hypothèse nulle peut provenir:

- D'expériences antérieures. Dans ce cas on cherche à savoir si les conditions ont changé.
- D'une théorie ou d'un modèle du problème. Dans ce cas, on cherche à déterminer si le modèle est valide.
- De considérations extérieures, comme des spécifications techniques. Dans ce cas on cherche à savoir si l'objet ou le processus est conforme.

Erreurs

Définiion

1. L'*erreur de première espèce* (de type I) est

$$\alpha = P(\text{rejeter } H_0 | H_0 \text{ est vraie})$$

2. L'*erreur de deuxième espèce* (de type II) est

$$\beta = P(\text{accepeter } H_0 | H_0 \text{ est fausse})$$

		Réalité	
		H_0 est vraie	H_0 est fausse
Décision	accepter H_0	$1 - \alpha$	erreur de type II (β)
	rejeter H_0	erreur de type I (α)	$1 - \beta$ (puissance)

Erreurs (ii)

Exercice 1: Entre rejeter ce qui est vrai et accepter ce qui est faux, quel est selon vous l'erreur la plus grave? Justifiez.

Définition:

1. On appelle α le *seuile critique* ou le *seuil de signification* du test.
2. La *puissance* du test est la probabilité de rejeter H_0 si H_0 est effectivement fausse, c'est-à-dire $1 - \beta$.

Exécution du test: les étapes détaillées

1. Formuler H_0 et H_1.
2. Choisir α.
3. Considérer un échantillon de taille n.
4. Exécuter le test: par exemple vérifier si la statistique $|Z_0|$ est plus grande que $z_{\frac{\alpha}{2}}$.
5. Conclure en acceptant ou en rejetant H_0:
   - Rejeter: H_0: conclusion forte.
   - Ne pas rejeter H_0: conclusion faible.
6. Facultatif: calculter β et le risque de deuxième espèce.
7. Facultatif: si β est trop élevé, indiquer un nouveau α et/ou un nouveau n et recommencer.

Remarques

- En général, on veut que α et β soient petits.
- La probabilité α de l'erreur de première espèce est habituellement fixée à l'avance.
- La probabilité β de l'erreur de deuxième espèce dépend de
  - La valeur réelle du paramètre θ.
  - La taille n de l'échantillon.
- Pour α et n fixés, $\beta \equiv \beta(0)$ est une fonction du paramètre θ. Le graphe de cette fonction est appelé *courbe caractéristique* du test.
- **Idée**: On choisit H_0 et α de façon à minimiser le risque d'erreur de type I (la plus grave), et β (la probabilité d'erreur de type II) est une conséquence de ce choix. Souvent $\alpha < \beta$.
- Si on veut diminuer β, il faut augmenter α et/ou n.

Illustration du concept: cas 1

On juge une personne et on formule les hypothèses suivantes:

$$H_0 : \text{personne innocente et} \quad H_1 : \text{personne coupable}$$

		Réalité	
		personne innocente	personne coupable
Décision	innocenter la personne	$1-\alpha$	erreur de type II (β)
	condamner la personne	erreur de type I (α)	$1-\beta$ (puissance)

- $P(\text{condamner un innocent}) = \alpha$.
- $P(\text{libérer un coupable}) = \beta$.

Ne pas condamner un innocent est prioritaire par rapport à ne pas libérer un coupable.

Illustration du concept: cas 2

On juge une personne et on formule les hypothèses suivantes:

$$H_0 \text{ : personne coupable et } \; H_1 \text{ : personne innocente}$$

		Réalité	
		personne coupable	personne innocente
Décision	condamner la personne	$1 - \alpha$	erreur de type II (β)
	innocenter la personne	erreur de type I (α)	$1 - \beta$ (puissance)

- $P(\text{libérer un coupable}) = \alpha$.
- $P(\text{condamner un innocent}) = \beta$.

Ne pas libérer un coupable est prioritaire par rapport à ne pas condamner un innocent.

Région critique

Le principe général d'un test d'hypothèse repose sur la considération d'une statistique et d'une *région critique*.

Une région critique est une région où il est peu probable que la statistique prenne des valeurs lorsque l'hypothèse nulle est vraie.

Le test consiste alors à:

- Rejeter H_0 si la valeur calculée de cette statistique est dans la région critique: Conclusion forte.
- Ne pas rejeter ($\simeq$ accepter) H_0 si la valeur calculée est en dehors de la région critique: Conclusion faible.

Tests d'hypothèses sur un seul échantillon

Test d'hypothèse bilatéral sur la moyenne: cas où σ^2 est connue

Les trois étapes principales pour le test sont:

1. Formulation des hypothèses: $H_0 : \mu = \mu_0$ et $H_1 : \mu \neq \mu_0$.
2. Calcul de la statistique pertinente avec les valeurs de l'échantillon:

$$Z_0 = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\text{ . Si } H_0 \text{ est vraie, alors } Z_0 \sim N(0,1) \ (n \text{ grand}).$$

3. Acceptation ou rejet de H_0:
   - On calcule $z_{\frac{\alpha}{2}}$ tel que $\Phi\left(z_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2}$.
   - On rejette H_0 si $|Z_0| > Z_{\frac{\alpha}{2}}$ ou on accepte H_0 si $|Z_0| \leq Z_{\frac{\alpha}{2}}$

Test d'hypothèse bilatéral sur la moyenne: cas ou σ^2 est inconnue

A la place de Z_0, on utilise la statistique

$$T_0 = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

et

- On rejette H_0 si $|T_0| > t_{\frac{\alpha}{2};n-1}$.
- On accepte H_0 si $|T_0| \leq t_{\frac{\alpha}{2};n-1}$.

Tests d'hypothèse unilatéraux sur la moyenne

Les formulations pour l'hypothèse alternative H_1 sont:

1. $H_0 : \mu = \mu_0$ (ou $\mu > \mu_0$), $H_1 : \mu < \mu_0$ (unilatéral à gauche).
2. $H_0 : \mu = \mu_0$ (ou $\mu \leq \mu_0$), $H_1 : \mu > \mu_0$ (unilatéral à droite).

Les statistiques du test sont les mêmes:

- $Z_0 = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ (σ^2 connue)
- $T_0 = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}$ (σ^2 inconnue)

Les critères de rejet changent:

- $Z_0 < -z_\alpha$ ou $T_0 < -t_{\alpha;n-1}$ (unilatéral à gauche).
- $Z_0 > Z_\alpha$ ou $T_0 > t_{\alpha;n-1}$ (unilatéral à droite).

Erreur de deuxième espèce β

Considérons les hypothèses

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Si H_0 est fausse alors $\mu = \mu_1 = \mu_0 + \Delta$ avec $\Delta \neq 0$. Si σ^2 **est connue**, on peut montrer que

$$\beta = \beta(\Delta) = \Phi\left(z_{\frac{\alpha}{2}} - \frac{\Delta}{\sigma}\sqrt{n}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{\Delta}{\sigma}\sqrt{n}\right)$$

Exercice 2: Prouver cette formule en commençant par montrer que $Z_0 \sim N\left(\frac{\Delta\sqrt{n}}{\sigma}, 1\right)$.

- Cas bilatéral: $H_1 : \mu \neq \mu_0 (\Delta \neq 0)$:

Erreur de deuxième espèce β (ii)

$$\beta(\Delta) = \Phi\left(z_{\frac{\alpha}{2}} - \frac{\Delta}{\sigma}\sqrt{n}\right) - \Phi\left(z_{\frac{\alpha}{2}} - \frac{\Delta}{\sigma}\sqrt{n}\right)$$

- Cas unilatéral gauche: $H_1 : \mu < \mu_0 (\Delta < 0)$:

$$\beta(\Delta) = \Phi\left(z_{\alpha} + \frac{\Delta}{\sigma}\sqrt{n}\right)$$

- Cas unilatéral droite: $H_1 : \mu > \mu_0$ $(\Delta > 0)$:

$$\beta(\Delta) = \Phi\left(z_{\alpha} - \frac{\Delta}{\sigma}\sqrt{n}\right)$$

Exercice 3: Illustrer le cas unilatéral à droite.

Détermination de n si σ^2 est connue

- Cas bilatéral: $H_1 : \mu \neq \mu_0$ ($\Delta \neq 0$):

$$n \simeq \left(\frac{\sigma \left(z_{\frac{\alpha}{2}} + z_\beta \right)}{\Delta} \right)^2$$

- Cas unilatéral gauche ou droite: ($\Delta < 0$ ou $\Delta > 0$):

$$n = \left(\frac{\sigma \left(z_\alpha + z_\beta \right)}{\Delta} \right)^2$$

Exercice 3: Prouver le cas unilatéral.

Détermination de n si σ^2 est connue (ii)

Exercice 4:

On étudie la vitesse de croissance d'un arbre. Les études précédentes montrent que la vitesse de croissance de cet arbre est de 10 mm/jour.

On sait que l'écart-type de cette vitesse est d'environ $\sigma = 2$ mm/jour.

On se préocuppe de la probabilité β d'accepter $H_0 : \mu = 10$ mm/jour alors que la vitesse moyenne de croissance est en réalité $\mu = 11$ mm/jour.

1. Calculter la probailité β si $\alpha = 0,05$ et $n = 25$.
2. Si on veut que le test détecte avec probabilité 0,80 un écart de 1 mm/jour entre $\mu_0 = 40$ cm/s et la vitesse de croissance moyenne réelle, quelle taille d'échantillon faut-il utiliser ?

Niveau critique observé (P-value)

Défintion

Le *niveau critique observé* (ou P-value) *PV* est la valeur minimale de α telle que H_0 est toujours rejetée.

- Avantage: une fois que la P-value est connue, le décideur peut déterminer la décision du rejet ou non-rejet en utilisant n'importe quel seuil α.
- Si $\alpha > PV$, H_0 est rejetée.
- Inconvénient: calcul pas toujours facile.
- Lors d'un test, les logiciels donnent *PV*.
- Habituellement, si *PV* est grande, H_0 est acceptée.

Calcul de la P-value lors du test $H_0 : \mu = \mu_0$

Soit Z_0 la statistique employée pour un test d'hypothèse et z_0 (ou t_0) sa valeur calculée à partir d'un échantillon.

- Si σ^2 est connue:
  - cas unilatéral gauche ($H_1 : \mu < \mu_0$) $PV = \Phi(z_0)$
  - cas unilatéral droite ($H_1 : \mu > \mu_0$) $PV = 1 - \Phi(z_0)$
  - cas bilatéral ($H_1 : \mu \neq \mu_0$) $PV = 2(1 - \Phi(|z_0|))$
- Si σ^2 est inconnue:
  - cas unilatéral $PV = P(T > |t_0|)$ avec $T \sim T_{n-1}$
  - cas bilatéral $PV = 2P(T > |t_0|)$ avec $T \sim T_{n-1}$

Relation avec les intervalles de confiance

Soit $[L, U]$ un intervalle de confiance pour un paramètre θ, de niveau de confiance $100(1-\alpha)\%$.

Alors le test d'hypothèses au seuil critique α pour

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \neq \theta_0$$

mène au rejet de H_0 si et seulement si $\theta_0 \notin [L, U]$.

Tests d'hypothèses avec un échantillon: autres cas

Voir polycope.

Exerice 5:

Un fabricant de boissons gazeuses s'intéresse à l'unformité du remplissage des canettes par une machine.

Soit X le volume de remplissage d'une canette. On suppose que X suit une loi normale.

Si la variance du volume de remplissage excède 15 mL^2 alors un pourcentage inacceptable de canettes ne seront pas remplies correctement.

Le fabricant veut donc tester si la variance est inférieure à 15 mL^2.

Supposons que le niveau critique soit fixé à $\alpha = 0,05$.

Quelle est la probabilité d'accepter $H_0 : \sigma^2 = 15$ mL^2 alors qu'en réalité la variance est d'au moins 25 mL^2, avec un échantillon de 20 canettes ?

Tests d'hypothèses avec un échantillon: autres cas (ii)

Exercice 6:

On veut tester $H_0 : p = 0,2$ contre $H_1 : p > 0,2$. Un échantillon de taille $n = 50$ donne 11 "succès".

1. Que peut-on conclure au niveau $\alpha = 5\%$?
2. Quelle taille supplémentaire doit-on prélever afin que H_0 soit rejetée 9 fois sur 10 lorsque $p = 0,24$?