Contrastos d'hipòtesis: Introducció

Decisions

Moltes situacions requereixen prendre una decisió sobre si s'ha d'acceptar o rebutjar alguna afirmació relativa al valor d'un paràmetre a partir de mostres

Decisions

Moltes situacions requereixen prendre una decisió sobre si s'ha d'acceptar o rebutjar alguna afirmació relativa al valor d'un paràmetre a partir de mostres

Exemple: Els responsables sanitaris del govern han determinat que el nombre mitjà de bacteris per cc a les aigües on es practiqui la recollida de cloïsses per al consum humà ha de ser $\leqslant 70$

Prenem una sèrie de mostres d'aigua d'una zona, i hem de decidir si s'hi poden recollir cloïsses

Decisions

Moltes situacions requereixen prendre una decisió sobre si s'ha d'acceptar o rebutjar alguna afirmació relativa al valor d'un paràmetre a partir de mostres

Exemple: Un hospital rep una partida de productes farmacèutics. L'encarregat té l'ordre de rebutjar les partides que continguin més d'un 5% d'unitats defectuoses

L'encarregat pren la decisió d'acceptar o rebutjar la partida basant-se en l'anàlisi d'una mostra aleatòria de productes d'aquesta

Decisions

Moltes situacions requereixen prendre una decisió sobre si s'ha d'acceptar o rebutjar alguna afirmació relativa al valor d'un paràmetre

Aquesta afirmació rep el nom d'hipòtesi i el mètode estadístic de presa d'una decisió sobre una hipòtesi rep el nom de contrast d'hipòtesis

Decisions

En un contrast d'hipòtesis, es contrasten dues hipòtesis alternatives: la hipòtesi nul·la H_0 i la hipòtesi alternativa H_1

- La hipòtesi alternativa H_1 és de la que cercam evidència
- La hipòtesi nul·la H_0 és la que rebutjarem si obtenim evidència de la hipòtesi alternativa
- Si no obtenim evidència a favor de H_1, no podem rebutjar H_0 ($\approx$ acceptam H_0, però és un abús de llenguatge)

Exemples

Els responsables sanitaris del govern han determinat que el nombre de bacteris per cc a les aigües on es practiqui la recollida de cloïsses per al consum humà ha de ser $\leqslant 70$. Hem de decidir si en una zona determinada s'hi poden recollir cloïsses

$\mu =$ nombre mitjà de bacteris per cc d'aigua

Contrast:

$$\begin{cases} H_0 : \mu \leqslant 70 \\ H_1 : \mu > 70 \end{cases}$$

Decisió: Prendrem algunes mostres i calcularem la mitjana mostral del nombre de bacteris per cc. Si és prou gran, ho considerarem evidència de H_1, i si no, acceptarem H_0.

Exemples

Un hospital rep una partida de productes farmacèutics. L'encarregat té l'ordre de rebutjar les partides que continguin més d'un 5% d'unitats defectuoses.

$p =$ proporció d'unitats defectuoses

Contrast:

$$\begin{cases} H_0 : p \leqslant 0.05 \\ H_1 : p > 0.05 \end{cases}$$

Decisió: L'encarregat comprova algunes unitats i calcula la proporció mostral de defectuoses. Si és prou gran, ho considera evidència de H_1, i si no, accepta H_0.

Decisions

Un contrast d'hipòtesis

$$\begin{cases} H_0 : \text{hipòtesi nul·la} \\ H_1 : \text{hipòtesi alternativa} \end{cases}$$

consisteix a plantejar:

- Hipòtesi nul·la H_0: És la hipòtesi que "per defecte" acceptam com a vertadera, i que rebutjarem si hi ha proves en contra
- Hipòtesi alternativa H_1: És la hipòtesi contra la qual contrastam la hipòtesi nul·la i que acceptarem en rebutjar la nul·la

i generar un regla de decisió per rebutjar o no la hipòtesi nul·la a partir de la informació continguda en una mostra

Exemple

En un judici, s'ha de declarar un acusat innocent o culpable

Contrast:

$$\left\{\begin{array}{l} H_0 : \text{L'acusat és innocent} \\ H_1 : \text{L'acusat és culpable} \end{array}\right.$$

Les proves són els elements de la mostra

Si el jurat troba prou incriminatòries les proves, declara culpable l'acusat (rebutja H_0 en favor de H_1)

Si no les troba prou incriminatòries, el declara no culpable (no rebutja H_0)

Considerar no culpable $\neq$ declarar innocent

Com triar H_0 i H_1?

Les proves han de poder donar evidència de H_1, la qual cosa permetrà rebutjar H_0

És impossible trobar evidència que $\mu =$ quelcom, en canvi sí que es pot demostrar que $\mu >$, o que $\mu <$, o que $\mu \neq$ quelcom

Com triar H_0 i H_1?

Les proves han de poder donar evidència de H_1, la qual cosa permetrà rebutjar H_0

És impossible trobar evidència que $\mu =$ quelcom, en canvi sí que es pot demostrar que $\mu >$, o que $\mu <$, o que $\mu \neq$ quelcom

Exemple: És $e^{\pi\sqrt{163}} = 262537412640768744$?

Com triar H_0 i H_1?

Les proves han de poder donar evidència de H_1, la qual cosa permetrà rebutjar H_0

És impossible trobar evidència que $\mu =$ quelcom, en canvi sí que es pot demostrar que $\mu >$, o que $\mu <$, o que $\mu \neq$ quelcom

Exemple: És $e^{\pi\sqrt{163}} = 262537412640768744$?

Si ho calculau

262537412640768743.99999999

Com triar H_0 i H_1?

Les proves han de poder donar evidència de H_1, la qual cosa permetrà rebutjar H_0

És impossible trobar evidència que $\mu =$ quelcom, en canvi sí que es pot demostrar que $\mu >$, o que $\mu <$, o que $\mu \neq$ quelcom

Exemple: És $e^{\pi\sqrt{163}} = 262537412640768744$?

Si ho calculau

262537412640768743.999999999999

Com triar H_0 i H_1?

Les proves han de poder donar evidència de H_1, la qual cosa permetrà rebutjar H_0

És impossible trobar evidència que $\mu =$ quelcom, en canvi sí que es pot demostrar que $\mu >$, o que $\mu <$, o que $\mu \neq$ quelcom

Exemple: És $e^{\pi\sqrt{163}} = 262537412640768744$?

Si ho calculau

$$262537412640768743.99999999999250072597198\ldots$$

Hem demostrat que no; emperò, per demostrar així que sí, hauríem d'haver calculat infinites xifres decimals

Com triar H_0 i H_1?

Les proves han de poder donar evidència de H_1, la qual cosa permetrà rebutjar H_0

És impossible trobar evidència que $\mu =$ quelcom, en canvi sí que es pot demostrar que $\mu >$, o que $\mu <$, o que $\mu \neq$ quelcom

- H_1 ha de ser definida per $>$, $<$, o $\neq$
- H_0 ha de ser definida per $=$, $\leqslant$, o $\geqslant$
- H_1 és la hipòtesi de la que podem trobar-ne proves, H_0 la que estam disposats a acceptar si no hi ha proves en contra

Com triar H_0?

Exemples

- Volem decidir si la mitjana és més petita que 2 o no

$$\begin{cases} H_0 : \mu = 2 \text{ (o } \mu \geqslant 2) \\ H_1 : \mu < 2 \end{cases}$$

Com triar H_0?

Exemples

- Volem decidir si la mitjana és més petita que 2 o no

$$\left\{\begin{array}{l} H_0 : \mu = 2 \text{ (o } \mu \geqslant 2) \\ H_1 : \mu < 2 \end{array}\right.$$

- Volem decidir si la mitjana és igual o diferent de 5

$$\left\{\begin{array}{l} H_0 : \mu = 5 \\ H_1 : \mu \neq 5 \end{array}\right. \text{ o } \left\{\begin{array}{l} H_0 : \mu \neq 5 \\ H_1 : \mu = 5 \end{array}\right. ?$$

Com triar H_0?

Exemples

- Volem decidir si la mitjana és més petita que 2 o no

$$\begin{cases} H_0 : \mu = 2 \text{ (o } \mu \geqslant 2) \\ H_1 : \mu < 2 \end{cases}$$

- Volem decidir si la mitjana és igual o diferent de 5

$$\begin{cases} H_0 : \mu = 5 \\ H_1 : \mu \neq 5 \end{cases}$$

Tipus d'hipòtesis alternatives

- Hipòtesi unilateral (*one-sided*; també d'una cua, *one-tailed*): $H : \theta > \theta_0$, $H : \theta < \theta_0$
- Hipòtesi bilateral (*two-sided*; també de dues cues, *two-tailed*): $H : \theta \neq \theta_0$

Els tests solen prendre el nom de la hipòtesi alternativa: test unilateral, test de dues cues,...

Tipus d'errors

Decisió	Realitat	
	H_0 certa	H_0 falsa
Acceptar H_0	Dec. correcta Prob$=1-\alpha$	Error Tipus II Prob$=\beta$
Rebutjar H_0	Error Tipus I Prob$=\alpha$	Dec. correcta Prob$=1-\beta$

- Error de Tipus I: Rebutjar H_0 quan és certa

  $P(\text{Error Tipus I}) = P(\text{Rebutjar } H_0 \mid H_0 \text{ certa}) = \alpha$

  α és el nivell de significació del contrast

Tipus d'errors

Decisió	Realitat	
	H_0 certa	H_0 falsa
Acceptar H_0	Dec. correcta Prob$=1-\alpha$	Error Tipus II Prob$=\beta$
Rebutjar H_0	Error Tipus I Prob$=\alpha$	Dec. correcta Prob$=1-\beta$

- Error de Tipus II: Acceptar H_0 quan és falsa

  $P(\text{Error Tipus II}) = P(\text{Acceptar } H_0 | H_0 \text{ falsa}) = \beta$

  $1-\beta = P(\text{Rebutjar } H_0 | H_0 \text{ falsa})$ és la potència del contrast

Tipus d'errors

En un judici, s'ha de declarar un acusat innocent o culpable

Contrast:

$$\begin{cases} H_0 : \text{L'acusat és innocent} \\ H_1 : \text{L'acusat és culpable} \end{cases}$$

Error de Tipus I: Declarar culpable un innocent

Error de Tipus II: Declarar no culpable un culpable

És pitjor l'error de Tipus I, convé minimitzar-lo

Tipus d'errors

L'ideal és trobar una regla de rebuig de H_0 que tingui poca probabilitat d'Error de Tipus I, α

Però també voldríem minimitzar la probabilitat d'Error de Tipus II, β

Problema: quan fem disminuir α, sol augmentar β

Què se sol fer?

1. Donar una regla de decisió per a un α màxim fixat
2. Després, si és possible, controlar la mida n de la mostra per minimitzar β

Tipus d'errors

L'ideal és trobar una regla de rebuig de H_0 que tingui poca probabilitat d'Error de Tipus I, α

Però també voldríem minimitzar la probabilitat d'Error de Tipus II, β

Problema: quan fem disminuir α, sol augmentar β

Què se sol fer?

1. Donar una regla de decisió per a un α màxim fixat
2. Després, si és possible, controlar la mida n de la mostra per minimitzar β

Aquesta segona part no la considerarem en aquest curs

Exemple: C.H. per a μ de normal amb σ coneguda

Sigui X una v.a. $N(\mu, \sigma)$ amb μ desconeguda i σ coneguda

Sigui $X_1, \ldots, X_n$ una m.a.s. de X de mida n

Considerem el contrast

$$\left\{ \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{array} \right.$$

Idea: Rebutjarem H_0 en favor de H_1 si $\overline{X}$ és prou més gran que μ_0

Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$

Si H_0 és vertadera,

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

Llavors, la regla consistirà a rebutjar H_0 si l'estadístic de contrast Z és major que un cert llindar, que determinarem amb α

Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$

Volem

$$\alpha = P(\text{Rebutjar } H_0 | H_0 \text{ certa}) = P(Z > \text{llindar})$$
$$\Longrightarrow 1 - \alpha = P(Z \leqslant \text{llindar}) \Longrightarrow \text{llindar} = z_{1-\alpha}$$

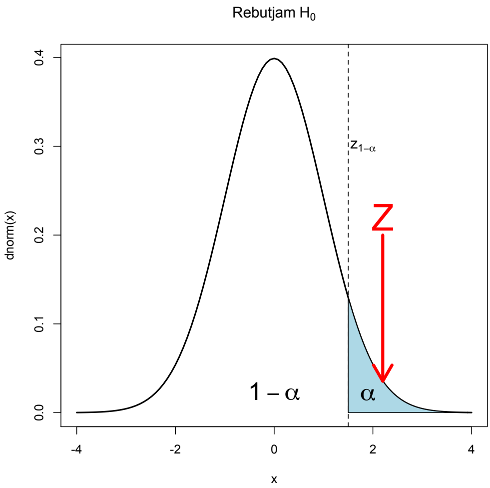
Per tant, perquè el nivell de significació del contrast sigui α, la regla de rebuig ha de ser

$$Z > z_{1-\alpha}$$

Rebutjam H_0 si $\displaystyle \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}$

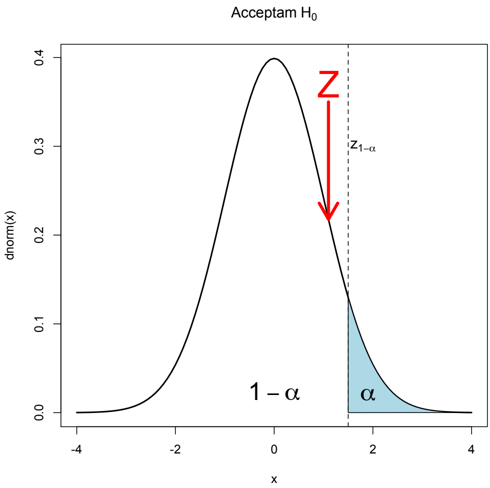
Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$



Exemple: C.H. per a μ de normal amb σ coneguda

$$\left\{\begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{array}\right.$$



Terminologia

Donat un contrast:

- Estadístic de contrast: el que ens permet definir una regla de rebuig de H_0
- Nivell de significació α: la probabilitat (màxima) d'error de Tipus I
- Regió crítica o de rebuig: el rang de valors tals que si l'estadístic de contrast hi cau, aleshores rebutjam H_0
- Regió d'acceptació: el complementari de la regió crítica

Terminologia

Donat un contrast:

- Interval de confiança del $(1-\alpha)\cdot 100\%$: un interval on el paràmetre poblacional té probabilitat $1-\alpha$ de pertànyer-hi (en el sentit dels intervals de confiança del tema anterior)

  Se sol obtenir imposant que l'estadístic pertanyi a la regió d'acceptació per al nivell de significació α i aïllant el paràmetre

Exemple: C.H. per a μ de normal amb σ coneguda

Si la població és normal amb σ coneguda, un contrast al nivell de significació α de

$$\left\{\begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{array}\right.$$

té

- Estadístic de contrast: $Z = \dfrac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
- Regió crítica: $]z_{1-\alpha}, \infty[$
- Regió d'acceptació: $]-\infty, z_{1-\alpha}]$
- Regla de decisió: Rebutjar H_0 si $Z > z_{1-\alpha}$

Exemple: C.H. per a μ de normal amb σ coneguda

- Interval de confiança:

$$Z \leqslant z_{1-\alpha} \iff \frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \leqslant z_{1-\alpha}$$

$$\iff \mu_0 \geqslant \overline{X} - z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}$$

$$\iff \mu_0 \in \left[\overline{X} - z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}, \infty\right[$$

  És

$$\left[\overline{X} - z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}, \infty\right[$$

- Regla de decisió II: Rebutjar H_0 si el μ_0 contrastat no pertany a l'interval de confiança

Exemple

Sigui X una població normal amb $\sigma = 1.8$. Volem fer el contrast

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

amb un nivell de significació de 0.05

Prenem una m.a.s. de $n = 25$ observacions i obtenim $\overline{x} = 20.25$

Què decidim?

Exemple

$$\left\{\begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{array}\right.$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.25$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.25$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.25 - 20}{\frac{1.8}{\sqrt{25}}} = 0.694$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.25$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.25 - 20}{\frac{1.8}{\sqrt{25}}} = 0.694$
- Regió crítica: $]z_{1-0.05}, \infty[=]1.64, \infty[$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.25$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.25 - 20}{\frac{1.8}{\sqrt{25}}} = 0.694$
- Regió crítica: $]z_{1-0.05}, \infty[=]1.64, \infty[$
- Decisió: Com que $0.694 < 1.64$, no podem rebutjar H_0

Exemple

$$\left\{\begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{array}\right.$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.25$

- Interval de confiança:

$$\left[\overline{X} - z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \infty\right[= [19.66, \infty[$$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.25$

- Interval de confiança:

$$\left[\overline{X} - z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \infty\right[= [19.66, \infty[$$

- Decisió: Com que 20 pertany a l'interval de confiança, no podem rebutjar H_0

Exemple

Sigui X una població normal amb $\sigma = 1.8$. Volem fer el contrast

$$\left\{ \begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{array} \right.$$

amb un nivell de significació de 0.05

Prenem una m.a.s. de $n = 25$ observacions i obtenim $\overline{x} = 20.75$

Què decidim?

Exemple

$$\left\{\begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{array}\right.$$

$\alpha = 0.05$, $\sigma = 1.8$, $n = 25$, $\overline{x} = 20.75$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.75$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.75 - 20}{\frac{1.8}{\sqrt{25}}} = 2.083$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.75$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.75 - 20}{\frac{1.8}{\sqrt{25}}} = 2.083$
- Regió crítica: $]z_{1-0.05}, \infty[=]1.64, \infty[$
- Interval de confiança:
  $[\overline{X} - z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \infty[= [20.16, \infty[$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.75$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.75 - 20}{\frac{1.8}{\sqrt{25}}} = 2.083$
- Regió crítica: $]z_{1-0.05}, \infty[=]1.64, \infty[$
- Interval de confiança: $[\overline{X} - z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \infty[= [20.16, \infty[$
- Decisió: Rebutjam H_0: Concloem que $\mu > 20$

Exemple

Sigui X una població normal amb $\sigma = 1.8$. Volem fer el contrast

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

amb un nivell de significació de 0.05

Prenem una m.a.s. de $n = 25$ observacions i obtenim $\overline{x} = 19.75$

Què decidim?

Exemple

Sigui X una població normal amb $\sigma = 1.8$. Volem fer el contrast

$$\left\{ \begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{array} \right.$$

amb un nivell de significació de 0.05

Prenem una m.a.s. de $n = 25$ observacions i obtenim $\overline{x} = 19.75$

Què decidim?

Com voleu que rebutgem H_0 contra H_1, si $\overline{x} < 20$? (Exercici: Feu el càlcul, si no em creieu)

Exemple: C.H. per a μ de normal amb σ coneguda

Sigui X una v.a. $N(\mu, \sigma)$ amb μ desconeguda i σ coneguda

Sigui $X_1, \ldots, X_n$ una m.a.s. de X de mida n

Considerem el contrast

$$\left\{ \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{array} \right.$$

Rebutjarem H_0 si $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ és inferior a un cert llindar, que determinarem amb α

Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$

Volem

$$\begin{aligned} \alpha &= P(\text{Rebutjar } H_0 | H_0 \text{ certa}) \\ &= P(Z < \text{llindar}) \Longrightarrow \text{llindar} = z_\alpha \end{aligned}$$

Per tant, perquè el nivell de significació del contrast sigui α, la regla de rebuig ha de ser

$$Z < z_\alpha$$

La regió crítica és $]-\infty, z_\alpha[$

Exemple: C.H. per a μ de normal amb σ coneguda

Sigui X una v.a. $N(\mu, \sigma)$ amb μ desconeguda i σ coneguda

Sigui $X_1, \ldots, X_n$ una m.a.s. de X de mida n

Considerem el contrast

$$\left\{\begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array}\right.$$

Rebutjarem H_0 si $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ està a prou distància de 0, i la determinarem amb α

Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

Volem

$$\begin{aligned} \alpha &= P(\text{Rebutjar } H_0 | H_0 \text{ certa}) \\ &= P(Z < -\text{llindar o } Z > \text{llindar}) \\ &= P(Z < -\text{llindar}) + P(Z > \text{llindar}) \\ &= 2P(Z > \text{llindar}) = 2(1 - P(Z < \text{llindar})) \\ &\quad \Longrightarrow P(Z < \text{llindar}) = 1 - \frac{\alpha}{2} \\ &\quad \Longrightarrow \text{llindar} = z_{1-\frac{\alpha}{2}} \end{aligned}$$

Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

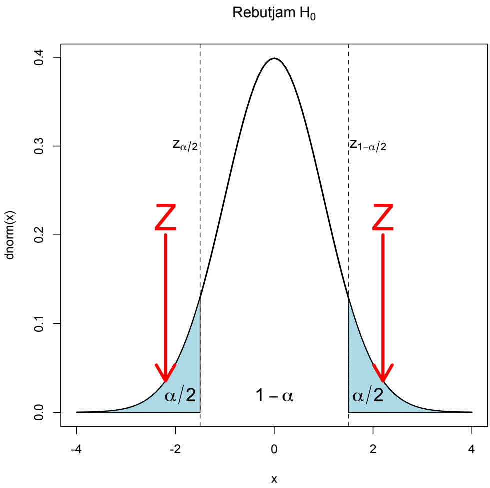
Per tant, perquè el nivell de significació del contrast sigui α, la regla de rebuig ha de ser

$$Z < -z_{1-\frac{\alpha}{2}} = z_{\frac{\alpha}{2}} \text{ o } Z > z_{1-\frac{\alpha}{2}}$$

La regió crítica és $]-\infty, z_{\frac{\alpha}{2}}[\cup]z_{1-\frac{\alpha}{2}}, \infty[$

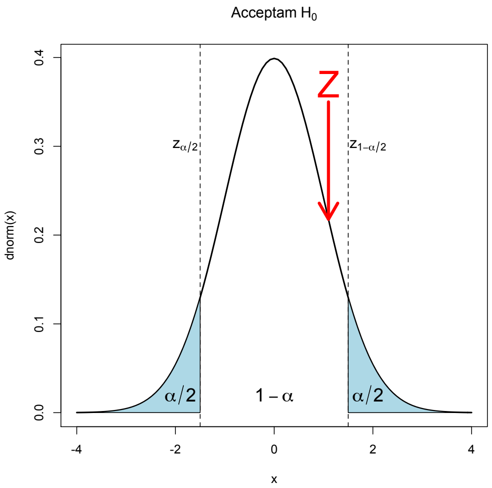
Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$



Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$



Exemple: C.H. per a μ de normal amb σ coneguda

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

Interval de confiança?

$$\begin{aligned} -z_{1-\frac{\alpha}{2}} \leqslant Z \leqslant z_{1-\frac{\alpha}{2}} &\Longleftrightarrow -z_{1-\frac{\alpha}{2}} \leqslant \frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \leqslant z_{1-\frac{\alpha}{2}} \\ &\Longleftrightarrow -z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leqslant \overline{X}-\mu_0 \leqslant z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \\ &\Longleftrightarrow \overline{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leqslant \mu_0 \leqslant \overline{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \\ &\Longleftrightarrow \mu_0 \in \left[\overline{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right] \end{aligned}$$

Us sona?

Exemple

Sigui X una població normal amb $\sigma = 1.8$. Volem fer el contrast

$$\left\{ \begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu \neq 20 \end{array} \right.$$

amb un nivell de significació de 0.05

Prenem una m.a.s. de $n = 25$ observacions i obtenim $\overline{x} = 20.5$

Què decidim?

Exemple

$$\left\{ \begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu \neq 20 \end{array} \right.$$

$\alpha = 0.05$, $\sigma = 1.8$, $n = 25$, $\overline{x} = 20.5$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu \neq 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.5$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.5 - 20}{\frac{1.8}{\sqrt{25}}} = 1.39$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu \neq 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.5$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.5 - 20}{\frac{1.8}{\sqrt{25}}} = 1.39$
- Regió crítica: $]-\infty, z_{0.025}[\cup]z_{0.975}, \infty[=]-\infty, -1.96[\cup]1.96, \infty[$
- Interval de confiança: $[19.794, 21.206]$

Exemple

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu \neq 20 \end{cases}$$

$\alpha = 0.05,\ \sigma = 1.8,\ n = 25,\ \overline{x} = 20.5$

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$
- Pren el valor $z_0 = \dfrac{20.5 - 20}{\frac{1.8}{\sqrt{25}}} = 1.39$
- Regió crítica:
  $]-\infty, z_{0.025}[\cup]z_{0.975}, \infty[=]-\infty, -1.96[\cup]1.96, \infty[$
- Interval de confiança: $[19.794, 21.206]$
- Decisió: No podem rebutjar H_0

El p-valor

El p-valor o valor crític (p-*value*) d'un contrast és la probabilitat que, si H_0 és vertadera, l'estadístic de contrast prengui un valor tan extrem o més que el que ha pres

Exemple: En un contrast

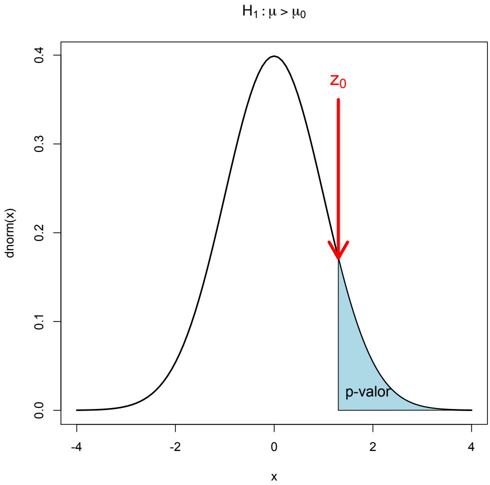
$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$

si l'estadístic Z ha valgut z_0,

$$p\text{-valor} = P(Z \geqslant z_0)$$

El p-valor

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$



El p-valor

El p-valor o valor crític (p-*value*) d'un contrast és la probabilitat que, si H_0 és vertadera, l'estadístic de contrast prengui un valor tan extrem o més que el que ha pres

Exemple: En un contrast

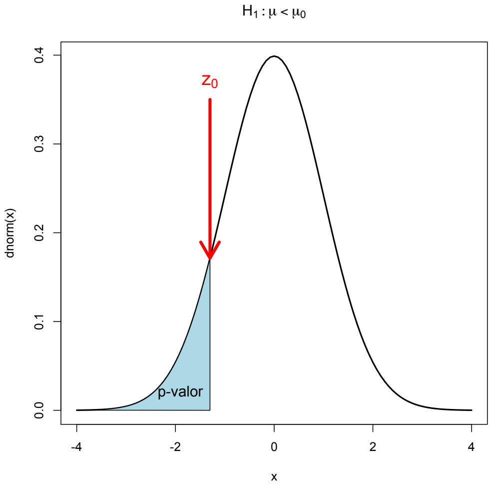
$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$

si l'estadístic Z ha valgut z_0,

$$p\text{-valor} = P(Z \leqslant z_0)$$

El p-valor

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$



El p-valor

El p-valor o valor crític (p-*value*) d'un contrast és la probabilitat que, si H_0 és vertadera, l'estadístic de contrast prengui un valor tan extrem o més que el que ha pres

Exemple: En un contrast

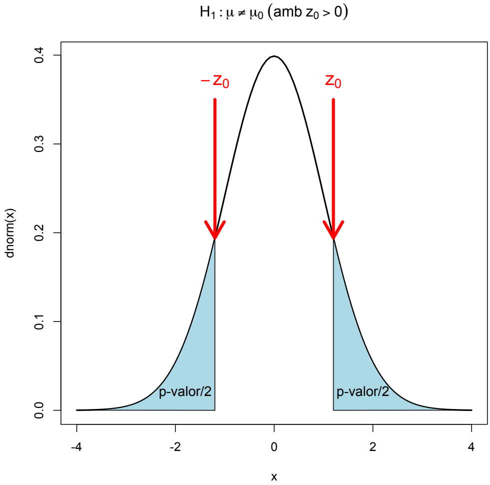
$$\left\{\begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array}\right.$$

si l'estadístic Z ha valgut z_0,

$$\begin{aligned} p\text{-valor} &= 2 \cdot \min\{P(Z \leqslant -|z_0|), P(Z \geqslant |z_0|)\} \\ &= 2 \cdot P(Z \geqslant |z_0|) \end{aligned}$$

El p-valor

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$



El p-valor

El p-valor o valor crític (p-*value*) d'un contrast és la probabilitat que, si H_0 és vertadera, l'estadístic de contrast prengui un valor tan extrem o més que el que ha pres

És una mesura inversa de la força de l'evidència de H_1: si H_0 és vertadera, com més petit sigui el p-valor, més improbable és observar el que hem observat.

Per tant, com més petit sigui el p-valor, amb més força podem rebutjar H_0.

El p-valor

Per exemple, p-valor $= 0.03$

- Significa que la probabilitat que, si H_0 és vertadera, l'estadístic de contrast prengui un valor tan extrem o més que el que ha pres és 0.03 (petita: evidència que H_0 és falsa)
- No significa:
  - La probabilitat que H_0 sigui vertadera és 0.03
  - H_0 és vertadera un 3% de les vegades

El p-valor

Important!

En un contrast amb nivell de significació α,

- rebutjam H_0 si p-valor $< \alpha$
- acceptam H_0 si $\alpha \leqslant p$-valor

Exemple: En un contrast

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$

suposem que l'estadístic Z ha valgut z_0. El p-valor és $P(Z \geqslant z_0)$

$$\text{Rebutjam } H_0 \iff z_0 > z_{1-\alpha}$$
$$\iff P(Z \geqslant z_0) < P(Z \geqslant z_{1-\alpha}) = 1 - (1-\alpha) = \alpha$$

El p-valor

Important!

En un contrast amb nivell de significació α,

- rebutjam H_0 si p-valor $< \alpha$
- acceptam H_0 si $\alpha \leqslant p$-valor

Exemple: En un contrast

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

suposem que l'estadístic Z ha valgut $z_0 > 0$. El p-valor és $2P(Z \geqslant z_0)$

$$\text{Rebutjam } H_0 \iff z_0 > z_{1-\frac{\alpha}{2}}$$
$$\iff 2P(Z \geqslant z_0) < 2P(Z \geqslant z_{1-\frac{\alpha}{2}}) = 2(1-(1-\tfrac{\alpha}{2})) = \alpha$$

El p-valor

Important!

En un contrast amb nivell de significació α,

- rebutjam H_0 si p-valor $< \alpha$
- acceptam H_0 si $\alpha \leqslant p$-valor

El *p-valor* d'un contrast és

- El nivell de significació α més petit per al qual rebutjaríem la hipòtesi nul·la
- El nivell de significació α més gran per al qual acceptaríem la hipòtesi nul·la

El p-valor

Important!

En un contrast amb nivell de significació α,

- rebutjam H_0 si p-valor $< \alpha$
- acceptam H_0 si $\alpha \leqslant p$-valor

El *p-valor* d'un contrast és

- El nivell de significació α més petit per al qual rebutjaríem la hipòtesi nul·la
- El nivell de significació α més gran per al qual acceptaríem la hipòtesi nul·la
- La probabilitat mínima d'error de Tipus I que permetem si rebutjam la hipòtesi nul·la amb el valor de l'estadístic de contrast obtingut

El p-valor

Important!

En un contrast amb nivell de significació α,

- rebutjam H_0 si p-valor $< \alpha$
- acceptam H_0 si $\alpha \leqslant p$-valor

El *p-valor* d'un contrast és

- El nivell de significació α més petit per al qual rebutjaríem la hipòtesi nul·la
- El nivell de significació α més gran per al qual acceptaríem la hipòtesi nul·la
- La probabilitat màxima d'error de Tipus I que permetem si acceptam la hipòtesi nul·la amb el valor de l'estadístic de contrast obtingut

El p-valor

Important!

Si no establim un nivell de significació α, aleshores

- Acceptam H_0 si el p-valor és "gran" ($\geqslant 0.1$)
- Rebutjam H_0 si el p-valor és "petit" (< 0.05). En aquest cas, el p-valor és
  - Significatiu si és < 0.05
  - Fortament significatiu si és < 0.01
  - Molt significatiu si és < 0.001

```
Signif. codes: 0‘***’0.001‘**’0.01‘*’0.05‘.’0.1‘ ’1
```

Si el p-valor està entre 0.05 i 0.1 i no tenim nivell de significació, es requereixen estudis posteriors per prendre una decisió (la zona de crepuscle, *twilight zone*)

Exemple

Sigui X una població normal amb $\sigma = 1.8$. Volem fer el contrast

$$\left\{ \begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{array} \right.$$

Prenem una m.a.s. de $n = 25$ observacions i obtenim $\overline{x} = 20.25$

Què decidim?

Exemple

Sigui X una població normal amb $\sigma = 1.8$. Volem fer el contrast

$$\begin{cases} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{cases}$$

Prenem una m.a.s. de $n = 25$ observacions i obtenim $\overline{x} = 20.25$

Què decidim?

No tenim α: Mirem el p-valor

Exemple

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$

  Pren el valor $z_0 = \dfrac{20.25 - 20}{\frac{1.8}{\sqrt{25}}} = 0.694$

Exemple

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$

  Pren el valor $z_0 = \dfrac{20.25 - 20}{\frac{1.8}{\sqrt{25}}} = 0.694$

- p-valor$= P(Z \geqslant 0.694) = 0.2438 > 0.1$ gran

Exemple

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$

  Pren el valor $z_0 = \dfrac{20.25 - 20}{\frac{1.8}{\sqrt{25}}} = 0.694$

- p-valor$= P(Z \geqslant 0.694) = 0.2438 > 0.1$ gran

- Decisió: No podem rebutjar H_0

Exemple

Sigui X una població normal amb $\sigma = 1.8$. Volem fer el contrast

$$\left\{ \begin{array}{l} H_0 : \mu = 20 \\ H_1 : \mu > 20 \end{array} \right.$$

Prenem una m.a.s. de $n = 25$ observacions i obtenim $\overline{x} = 20.75$

Què decidim?

Exemple

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$

  Pren el valor $z_0 = \dfrac{20.75 - 20}{\frac{1.8}{\sqrt{25}}} = 2.083$

Exemple

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$

  Pren el valor $z_0 = \dfrac{20.75 - 20}{\frac{1.8}{\sqrt{25}}} = 2.083$

- p-valor$= P(Z \geqslant 2.083) = 0.0186$ petit

Exemple

- Estadístic de contrast: $Z = \dfrac{\overline{X} - 20}{\frac{1.8}{\sqrt{25}}}$

  Pren el valor $z_0 = \dfrac{20.75 - 20}{\frac{1.8}{\sqrt{25}}} = 2.083$

- p-valor$= P(Z \geqslant 2.083) = 0.0186$ petit

- Decisió: Rebutjam H_0 contra H_1

Decisions

Podem decidir un contrast:

- Amb la regió crítica: Si l'estadístic de contrast cau dins la regió crítica per al nivell de significació α, rebutjam H_0
- Amb l'interval de confiança: Si el paràmetre poblacional a contrastar cau dins l'interval de confiança per al nivell $(1-\alpha)\cdot 100\%$, acceptam H_0
- Amb el p-valor: Si el p-valor és més petit que el nivell de significació α, rebutjam H_0
- Amb el p-valor i sense α: Si el p-valor és petit, rebutjam H_0, i si és gran, l'acceptam

Aquí emprarem el p-valor

El mètode dels sis passos (amb α)

1) Establiu la hipòtesi nul·la H_0 i la hipòtesi alternativa H_1
2) Fixau un nivell de significació α
3) Seleccionau l'estadístic de contrast apropiat
4) Calculau el valor de l'estadístic de contrast a partir de les dades mostrals
5) Calculau el p-valor del contrast
6) Decisió: rebutjau H_0 en favor de H_1 si el p-valor és més petit que α; en cas contrari, acceptau H_0

El mètode dels cinc passos (sense α)

1) Establiu la hipòtesi nul·la H_0 i la hipòtesi alternativa H_1
2) Seleccionau l'estadístic de contrast apropiat
3) Calculau el valor de l'estadístic de contrast a partir de les dades mostrals
4) Calculau el p-valor del contrast
5) Decisió: rebutjau H_0 en favor de H_1 si el p-valor és petit (< 0.05), acceptau H_0 si el p-valor és gran ($\geqslant 0.1$), i ampliau l'estudi si el p-valor està entre 0.05 i 0.1

Exemple

Els anys que viuen els habitants d'un país segueixen aproximadament una llei normal amb $\sigma = 8.9$ anys

Una mostra aleatòria de 100 morts enguany ha donat una vida mitjana de 71.8 anys

Volem decidir si la vida mitjana en aquest país és superior als 70 anys

Emprarem el contrast

$$\left\{ \begin{array}{l} H_0 : \mu = 70 \\ H_1 : \mu > 70 \end{array} \right.$$

Exemple: amb α

Suposem que prenem un nivell de significació $\alpha = 0.05$

L'estadístic de contrast és

$$Z = \frac{\overline{X} - 70}{8.9/\sqrt{100}} = \frac{\overline{X} - 70}{0.89}$$

El seu valor en aquest contrast és $z_0 = \dfrac{71.8 - 70}{0.89} = 2.02$

El p-valor és

$$P(Z \geqslant 2.02) = 0.0217$$

Com que $0.0217 < \alpha$, rebutjam H_0: concloem que $\mu > 70$

Exemple: amb α

Suposem que prenem un nivell de significació $\alpha = 0.01$

L'estadístic de contrast és

$$Z = \frac{\overline{X} - 70}{6.8/\sqrt{100}} = \frac{\overline{X} - 70}{0.68}$$

El seu valor en aquest contrast és $z_0 = \dfrac{71.8 - 70}{0.68} = 2.022$

El p-valor és

$$P(Z \geqslant 2.022) = 0.0217$$

Com que $0.0217 > \alpha$, no podem rebutjar H_0: concloem que $\mu \leqslant 70$ amb aquest nivell de significació

Exemple: sense α

L'estadístic de contrast és

$$Z = \frac{\overline{X} - 70}{6.8/\sqrt{100}} = \frac{\overline{X} - 70}{0.68}$$

El seu valor en aquest contrast és $z_0 = \dfrac{71.8 - 70}{0.68} = 2.022$

El p-valor és

$$P(Z > 2.022) = 0.0216$$

Com que és petit (< 0.05), rebutjam H_0: concloem que $\mu > 70$

Un darrer consell

El que avui en dia es recomana als informes és:

- Si donen α: Donar el *p*-valor i l'interval de confiança del contrast per a l'α donat (nivell de confiança $(1-\alpha)\cdot 100\%$)
- Si no donen α: Donar el *p*-valor, i l'interval de confiança del contrast al nivell de confiança 95%