

Continguts: www.acclc.cat/ivv_docs.php?any=2015*In vitro veritas*Pàgina web de la revista: www.acclc.cat/ivv.php

Document docent

Contrast d'hipòtesis

Associació Catalana de Ciències de Laboratori Clínic
Secció d'Estadística i Metrologia ¹

Henar Valbuena Parralejo

Laboratoris Clínics, Hospital Universitari Vall d'Hebron, Barcelona

¹Membres de la Secció d'Estadística i Metrologia durant la preparació d'aquest document: B. Candás Estébanez, X. Fuentes Arderiu (coordinador), M. Martínez Casademont, S. Miró Cañís, J.M. Queralto Compañó, H. Valbuena Parralejo

2015 © Publicat per l'Associació catalana de ciències de Laboratori Clínic

1. Introducció

En el primer document docent publicat per la Secció d'Estadística i Metrologia (1) ja es van introduir els conceptes d'hipòtesis i d'inferència estadística com a eines útils en la confirmació o el rebuig de conclusions obtingudes a partir d'unes dades.

En un altre document docent s'ha aclarit què són i quines són les principals distribucions de probabilitats (2) que permeten dur a terme els contrastos d'hipòtesis, anomenats *paramètrics*, encara que també s'han descrit contrastos d'hipòtesis, anomenats *no paramètrics*, que no requereixen el coneixement de distribucions de probabilitats.

En aquest document docent es presenten els conceptes necessaris per comprendre els fonaments dels contrastos d'hipòtesis.

2. Contrast d'hipòtesis (2, 3)

Un contrast d'hipòtesis (també anomenat prova de contrast d'hipòtesis, prova de decisió estadística o prova de significació) és un procés que permet decidir probabilísticament entre una hipòtesi nul·la i una hipòtesi alternativa.

En aquest procés, el primer pas és plantejar les dues hipòtesis que s'han de contrastar. La hipòtesi nul·la, H_0 , és la que acceptarem, llevat que les dades indiquin la seva falsedat. Cal dir que encara que H_0 pot ser rebutjada per les dades, mai pot ser verificada: en el cas que les dades no indiquin el seu rebuig es dirà que «res s'oposa a afirmar que H_0 és certa».

La hipòtesi alternativa, H_1 , és la hipòtesi que enfrontem a H_0 . Si no es diu una altra cosa, H_1 indica simplement que és H_0 falsa.

Per exemple, en un estudi epidemiològic sobre la concentració de substància de colesterol en el plasma dels barcelonins de més

de 50 anys es podrien voler estudiar les diferències entre els homes i les dones. H_0 indicaria que no existeix diferència en la concentració de substància de colesterol en el plasma entre els homes i les dones de més de 50 anys, i H_1 indicaria que sí que hi ha diferència. Es plantejaria com un problema estadístic per decidir si la mitjana poblacional per aquesta magnitud en els homes i les dones és diferent:

$$H_0: \mu_{\text{homes}} = \mu_{\text{dones}}$$

$$H_1: \mu_{\text{homes}} \neq \mu_{\text{dones}}$$

A continuació es du a terme una prova estadística per contrastar les dues hipòtesis. Si el contrast és paramètric la selecció de la prova estadística concreta dependrà fonamentalment de la distribució de probabilitat de la població en estudi. Si el contrast és no paramètric la distribució de probabilitat esmentada no té cap interès, encara que generalment seran necessàries més dades per poder-lo aplicar.

En l'exemple anterior, un cop realitzat el contrast d'hipòtesis oportú, es pot decidir probabilísticament si es rebutja la H_0 .

3. Errors de tipus I i II. Riscs α i β (4, 5)

Després de realitzar el contrast, s'escull una de les dues hipòtesis, H_0 o H_1 . Hem de tenir en compte, però, que el fet d'escollir una d'aquestes dues hipòtesis comporta sempre un risc d'equivocar-se:

a) Existeix un risc (probabilitat) α que es rebutgi H_0 quan en realitat és certa. Aquesta decisió equivocada rep el nom d'*error de tipus I* o de *primera espècie*. El risc α és una probabilitat coneguda que es defineix *a priori*. És habitual fixar un $\alpha = 0,05$, encara que sempre es pot modificar per aconseguir condicions més relaxades o més estrictes.

El risc α escollit per construir intervals de probabilitat que serveixen com a criteri en els contrastos d'hipòtesis rep el

nom de *nivell de significació*. Així, per exemple, quan definim $\alpha = 0,05$, es parla de diferència significativa amb un nivell de significació de 0,05 o del 5 %.

Paral·lelament, es defineix com a *nivell de confiança* el valor $1-\alpha$, i en aquest cas es parlaria de diferència significativa amb un nivell de confiança de 0,95 o del 95 %.

- b) També es pot cometre una equivocació a l'acceptar H_0 quan en realitat és falsa. Aquest error es coneix com *error de tipus II* o *de segona espècie*; el risc de cometre un error de tipus II rep el nom de *risc β* . És un valor desconegut, i disminueix en augmentar el nombre d'individus estudiats en la mostra.

Es defineix el valor $1-\beta$ com a *potència* del contrast d'hipòtesis i és igual a la probabilitat de rebuig de la hipòtesi nul·la quan, en realitat, és falsa.

4. Bibliografia

- (1) Fuentes Arderiu X. Estadística i ciències del laboratori clínic. *In vitro veritas* 2013; 14:25-30. <<http://www.acclc.cat/continguts/ivv148.pdf>> (Accés: 2015-03-04).
- (2) Miró Cañís S, Fuentes Arderiu X. Distribucions de freqüències i distribucions de probabilitats. *In vitro veritas* 2015;16:35-39. <<http://www.acclc.cat/continguts/ivv178.pdf>> (Accés: 2015-03-04).
- (3) García Pérez A. Estadística aplicada: Conceptos básicos. Madrid: UNED; 2012.
- (4) Parés Mariné N, Pozo Montero F, Mazaira Font F, Vidal Seguí Y. Probabilitat i estadística matemàtica: Teoria i problemes resolts. Barcelona: Universitat Politècnica de Catalunya; 2010.
- (5) Doménech Massons J M. Bioestadística: Métodos estadísticos para investigadores. Barcelona: Herder; 1977.