

INFERÈNCIA ESTADÍSTICA. 10-11

Pràctica 13 d'Estadística II.

Un enginyer informàtic vol comparar dos tipus de sistemes quant a la velocitat de processament de dades per a un determinat tipus de programes. Processa 70 treballs semblants, amb unes determinades quantitats d'inputs i outputs a cadascun, amb el sistema A i després processa uns altres 70 treballs, amb les mateixes quantitats d'inputs i outputs que cadascun dels primers treballs, amb el sistema B. Els resultats dels temps de processament en cadascun dels dos sistemes són

$$\begin{array}{ll}\bar{x}_A = 1.53 \text{ mseg} & \bar{x}_B = 1.34 \text{ mseg} \\ \hat{s}_A = 0.49 \text{ mseg} & \hat{s}_B = 0.53 \text{ mseg}\end{array}$$

Genereu unes col·leccions de valors que representin aquestes mostres de la variable amb la distribució de probabilitat que considereu adequada (justifiqueu-la).

Construiu un interval de confiança al 99% per a la diferència de mitjanes dels temps de processament amb els dos sistemes.

Amb un risc de 10%, podeu rebutjar la hipòtesi que la diferència de mitjanes és de 0.2, o no podeu fer-ho? I que la diferència de mitjanes és més gran que 0.1?

Oblidant ara tots els resultats obtinguts anteriorment, podeu assegurar en un 90% que no vénen de la mateixa distribució de probabilitat?

Genereu una tercera mostra que representi una variable aleatòria amb distribució normal de mitjana i desviació les mitjanes de les altres dues mostres. Oblidant-vos ara de com s'ha generat la mostra i dels resultats anteriors, considereu que la tercera mostra correspon als temps de processament en un tercer sistema i que les tres poblacions tenen una distribució normal amb la mateixa desviació. Contrastau si hi ha diferència significativa entre les mitjanes dels temps, és a dir, que el temps de processament depèn del sistema.

En el primer apartat ens demana generar la col·lecció de valors que representin les dues mostres escollin la probabilitat que vulguem nosaltres. Jo he escollit la distribució normal perquè es tracta de variables contínues perquè la velocitat es continua.

Per fer-ho creem les variables, anem al menú transformar i al submenú calcular variable, allí, seleccionem les variables i busquem l'expressió numèrica de la funció normal i hi afegim les dades.

En el segon apartat hem de construir l'interval de confiança al 99% per a la diferència de mitjanes de les dues variables.

Per fer-ho unim les dues variables i un cop unides creem una altra variable per definir els dos sistemes. Un cop fet anem al menú analitzar amb el submenú comparar mitges i aplicant la opció prova T de mitjanes independents, posem la variable de contrast i la

variable d'agrupació on toca i posar un interval de confiança al 99%.
El resultat és (-0.00434, 048739) ja que no assumim que les variàncies no son iguals.

En el tercer apartat ens demanen si podem rebutjar la hipòtesi que la diferencia de mitjanes és de 0.2. Primer fem la hipòtesi, la hipòtesi $H_0: \mu_1 - \mu_2 - 0.2 = 0$ i la hipòtesi $H_1: \mu_1 - \mu_2 - 0.2 \neq 0$.

El primer que fem es restar 0.2 a la nostra primera variable i tornem a realitzar la prova T de mitjanes independents però amb un $\alpha=0.1$, al fer el resultat obtenim el valor sig.(bilateral) que es 0.66 aquest valor es molt més gran que α , per tant, no rebutgem H_0 , per tant, podem suposar que la diferencia de mitjanes es igual a 0.2.

Per a la segona part de l'apartat el que hem de fer es comprovar que la diferencia de mitjanes sigui més gran que 0.1. Per tant las nostres hipòtesis ara serien $H_0: \mu_1 - \mu_2 - 0.1 \leq 0$ i $H_1: \mu_1 - \mu_2 - 0.1 > 0$, llavors fem el mateix que abans restem 0.1 a la primera variable i tornem a comparar els valors. En aquest cas hem de dividir la sig.(bilateral) entre 2, obtenint que sig./2=0.0675 que com es més petit que 0.1 rebutgem la hipòtesi H_0 , per tant podem suposar que la diferència de mitjanes és més gran que 0.1.

En el quart apartat ens demanen demostrar si no vénen de la mateixa distribució de probabilitat, hem de fer un test d'homogeneïtat. Les nostres hipòtesis són la H_0 : segueixen la mateixa distribució i H_1 : no segueixen la mateixa distribució. Tornem a separar les dues variables en dos columnes i anem al menú analitzar i al submenú de proves no paramètriques per agafar la opció de 2 mostres relacionades.

Un cop obtinguts els resultats comprovem que el sig. és igual a 0.027, la tornem a dividir entre dos i ens queda que $\alpha=0.1$ i que sig./2=0.0135, com que es més petit rebutjem H_0 , per tant, podem suposar que no segueixen una mateixa distribució.

Per al cinquè i últim apartat hem de generar una tercera mostra per fer un test ANOVA, la tercera mostra ha de ser de distribució normal de mitjana i desviació les mitjanes de les altres dues mostres i he suposat que havia de tenir 70 treballs com les altres dues i posem que $\alpha=0.1$.

Primer creem la variable aleatòria, després les posem totes en una sola variable amb una variable per definir els tres sistemes i les nostres hipòtesis son $H_0: \mu_1 = \mu_2 = \mu_3$ i H_1 : hi ha diferencia significativa. Un cop fet anem al menú analitzar, al submenú comparar mitjanes i al test d'ANOVA d'un factor.

Els resultats que obtenim que sig.=0.016, com que sig. es menor que α rebutgem H_0 , per tant podem suposar que hi ha diferència significativa entre les tres mostres.