

Exercicis de comparació de 2 mitjanes

1. Un agricultor que es dedica a la producció de taronges té notícies de que podant els arbres de diferent forma a com ho fa tradicionalment s'aconsegueix una major collita de taronges. Per veure si això és així, decideix fer una prova i et demana consell sobre com portar-la a terme (amb quants arbres, ...). Indica quines són les preguntes que faries a l'agricultor per tenir la informació necessària per poder assessorar-lo correctament.

L'agricultor decideix finalment fer la prova amb 10 arbres que podarà amb el nou sistema, deixant com a grup de control altres 10 molt semblants (tant en desenvolupament com en localització i altres característiques que poden tenir influència sobre la producció). El pes, en kg, de taronges produïdes per cada arbre és:

Habitual: 41; 37; 46; 43; 48; 49; 47; 48; 47; 53

Nou: 50; 47; 44; 51; 40; 49; 48; 51; 40; 58

- a) Estan les dades en contra de la hipòtesi d'igualtat de variàncies poblacionals?
- b) Quin és l'estadístic de prova i la distribució de referència?
- c) Prenent un nivell de significació $\alpha = 0,05$. Pot rebutjar-se la hipòtesi nul·la?
- d) Quin és el p-valor de la prova?

2. Una empresa subministra tubs d'escapament a la indústria de l'automòbil. En el tub d'escapament es col·loca un sensor que comunica a l'ordinador del cotxe el contingut de CO en els gasos. L'empresa disposa de dos tipus de sensor, A i B, basats en principis de mesurament diferents, però tots dos es col·loquen en el silenciós del tub d'escapament. Si es desitja analitzar si els 2 tipus de sensor marquen igual, indica quin tipus de disseny (totalment aleatoritzat o amb dades aparellades) utilitzaries per a la recollida de les dades.

L'empresa que col·loca 2 sensors en cada cotxe, situats en 2 posicions del silenciós del tub d'escapament molt pròximes entre si. (Després d'analitzar el tema amb molt detall, els tècnics consideren que no hi ha cap raó per pensar que la posició en què es col·loquen afecti la lectura del sensor però, com a precaució addicional, en cada cotxe aleatoritzen la posició en què es col·loquen els sensors).

El contingut en parts per milió de CO de observat en l'experiment va ser:

Cotxe	1	2	3	4	5	6	7	8	9	10
Tipus A:	72,1	68,2	70,9	74,3	70,7	66,6	69,5	70,8	68,8	73,3
Tipus B:	74,0	68,8	71,2	74,2	71,8	66,4	69,8	71,3	69,3	73,6

- Calcula un interval de confiança del 95% per a la diferència de lectura entre sensors
- Es pot afirmar que tots dos sensors mesuren diferent?

- Per decidir quin de 2 vins és millor s'escullen 14 experts, que puntuaran els vins en una escala d'1 a 10. Com faries la prova? Segueix la resposta una llei normal? És això un inconvenient greu per a l'ús de la t-Student com a distribució de referència?

En la prova de tast de vins per a decidir com de 2 tipus és millor, cada un dels experts va tastar cada un dels vins i, seguint el protocol i els criteris usats en aquests casos, va assignar una puntuació (escala 0 a 10) a cada un d'ells. L'ordre de tast va ser aleatori i el nombre indicat com superíndex al costat de cada puntuació indica l'ordre en què es van provar els vins. Analitza els resultats per indicar quin vi és millor. Estigues atent a possibles sorpreses en les dades...

Tastador:	1	2	3	4	5	6	7	8	9	10	11	12*
Vi A:	5,4 ⁽²⁾	5,2 ⁽¹⁾	4,8 ⁽¹⁾	4,6 ⁽²⁾	6,3 ⁽¹⁾	5,4 ⁽¹⁾	5,0 ⁽¹⁾	8,0 ⁽²⁾	6,0 ⁽²⁾	7,9 ⁽¹⁾	4,2 ⁽¹⁾	5,0 ⁽¹⁾
Vi B:	4,9 ⁽¹⁾	8,8 ⁽²⁾	6,2 ⁽²⁾	3,4 ⁽¹⁾	8,7 ⁽²⁾	8,0 ⁽²⁾	8,4 ⁽²⁾	7,4 ⁽¹⁾	5,6 ⁽¹⁾	9,4 ⁽²⁾	5,5 ⁽²⁾	7,2 ⁽²⁾

*Estava previst comptar amb 14 tastadors, però al final 2 no van poder assistir

- Una empresa que es dedica a la venda per correu a través de catàleg decideix realitzar un estudi per comparar el poder de persuasió de 2 tipus de disseny del catàleg. Han decidit imprimir 5000 de cada tipus. Pot tenir sentit un disseny aparellat de recollida de dades en aquest cas?

L'empresa divideix la zona a què dirigeix la seva campanya en 10 àrees atenent a criteris de caràcter sociològic. A cada àrea s'envien 500 catàlegs de cada tipus. La taula següent conté el nombre de respostes que es van obtenir. Es pot dir que un tipus de catàleg és millor que un altre a l'efecte d'aconseguir respostes?

Àrea:	1	2	3	4	5	6	7	8	9	10
Tipus A:	28	15	43	16	54	12	34	67	28	46
Tipus B:	12	16	18	21	42	3	6	21	13	22

5. Un ramader que es dedica a la cria de vedelles és informat que afegint un cert "complex vitamínic" a l'alimentació dels animals aquests arriben a un major pes quan es porten a l'escorxador. La variabilitat del pes final de les vedelles es pot caracteritzar per $\sigma = 10$ kg.

El ramader decideix fer l'experiment alimentant a 20 animals amb la dieta enriquida, mantenint un altre grup de 20 com a grup de control. Recordant que se sap, perquè es tenen moltes dades històriques que així ho confirmen, que la desviació tipus del pes de les vedelles quan es porten a l'escorxador és de $\sigma = 10$ kg.

- a) Si la dieta enriquida aconsegueix engreixar les vedelles una mitjana de 8 kg i es pren la decisió amb un nivell de significació 0,05. Quina és la probabilitat que no canviï de la dieta normal a l'enriquida?
- b) Canviaria la resposta a la pregunta anterior si enlloc d'haver-se fet servir 20 animals a cada grup (dieta enriquida i control) s'haguessin fet servir 40?
6. Si es comet l'error d'analitzar com si fos un disseny totalment aleatoritzat unes dades que s'han obtingut de forma aparellada, com es poden veure afectades les conclusions de l'estudi?