

## ESTADÍSTICA II

### EJERCICIOS TEMA 3

---

1. Se quiere evaluar el efecto del ruido sobre la habilidad de las personas para realizar una determinada tarea. Para ello, se seleccionan 32 personas que son capaces de realizar la misma tarea en el mismo tiempo para realizar dicha tarea en un medio controlado y bajo 2 niveles distintos de ruido de fondo. De entre esas 32, se seleccionan aleatoriamente 16 para llevar a cabo la tarea sometidas a un ruido de fondo de nivel 2. Las restantes 16 las realizarán con poco ruido de fondo (nivel 1). Los tiempos (en minutos) que se registran son:

nivel 1	14	12	15	15	11	16	17	12	14	13	18	13	18	15	16	11
nivel 2	20	22	18	18	19	15	18	15	22	18	19	15	21	22	18	16

- (a) Plantea el contraste oportuno, especificando claramente la hipótesis nula y alternativa que consideras.
- (b) Calcula el  $p$ -valor del contraste, suponiendo normalidad y la misma varianza.
2. Se quiere determinar si el paro en dos grandes áreas urbanas del país, como son Madrid y Barcelona, es diferente. Para ello se toman muestras aleatorias en ambas ciudades, cada una de 500 personas, obteniéndose que en Madrid 35 estaban desempleadas y en Barcelona 25. ¿Existe alguna razón para creer que las frecuencias de paro en Madrid y Barcelona son diferentes?

- (a) Plantea el contraste oportuno con un nivel de significación del 5%, especificando claramente la hipótesis nula y alternativa que consideras, así como las hipótesis de trabajo.
- (b) Calcula el  $p$ -valor del contraste.

3. Construye un intervalo de confianza de nivel  $1 - \alpha$  para la diferencia de medias de dos poblaciones normales cuando podemos asumir que tienen la misma varianza,  $X \sim N(\mu_X, \sigma^2)$ ,  $Y \sim N(\mu_Y, \sigma^2)$ , y disponemos de dos muestras aleatorias simples  $X_1, \dots, X_{n_1}$ ,  $Y_1, \dots, Y_{n_2}$  independientes entre sí. ¿Qué amplitud tiene el intervalo?

Nota: emplea como estimador de  $\mu_X - \mu_Y$  la diferencia de las medias muestrales  $\bar{X} - \bar{Y}$ , y usa los resultados básicos vistos en clase para definir un estadístico que relacione  $\mu_X - \mu_Y$  con  $\bar{X} - \bar{Y}$  cuya distribución sea conocida.

4. El método MATWES<sup>1</sup> fue diseñado para medir las actitudes hacia las mujeres ejecutivas. Una puntuación alta indica actitudes negativas hacia las mujeres ejecutivas por parte de la persona que la otorga, puntuaciones bajas son indicativas de actitudes positivas. Se conjetura que la actitud hacia las mujeres ejecutivas cambia en función del sexo.

Para contrastar esta hipótesis se tomaron muestras aleatorias independientes de 151 hombres y de 108 mujeres estudiantes de M.B.A. En el grupo de los hombres se obtuvo una puntuación media de 85.8 con una desviación típica de 19.3. En el de mujeres se obtuvo una puntuación media de 71.5 con una desviación típica de 12.2.

- (a) ¿A qué conclusión llegas para un nivel de significación de  $\alpha = 0.01$ ? Plantea el contraste oportuno especificando claramente la hipótesis nula y alternativa que consideras, así como las hipótesis de trabajo.
- (b) A la vista de los resultados obtenidos, si construyeras un intervalo de confianza al 99% para la diferencia de puntuaciones medias, ¿contendría al 0?
5. Para comparar la velocidad de dos ordenadores, A y B, se mide el tiempo que invierten en realizar operaciones de una cierta clase definida. Se tomó una muestra de 5 operaciones de esta clase y cada operación fue realizada por ambos ordenadores. Se obtuvieron los siguientes tiempos (en milisegundos):

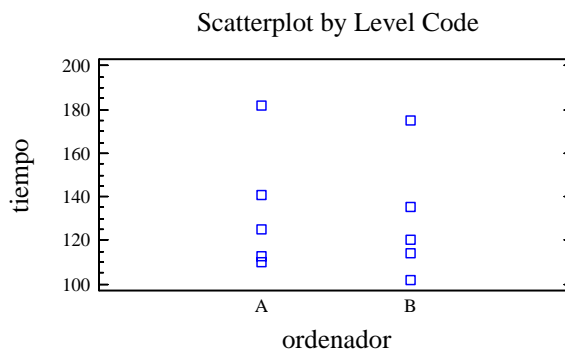
---

<sup>1</sup>P. Dubno, "Attitudes toward woman executives: A longitudinal approach", *Academy of Management Journal*, 28 (1985), 135-39, ver Newbold

A	110	125	141	113	182
B	102	120	135	114	175

Analizar si hay diferencias:

- Teniendo en cuenta que los datos están apareados (la misma operación es realizada por los 2 ordenadores). Obtén el  $p$ -valor del contraste.
- Considerando muestras independientes.
- En el segundo caso, al considerar muestras independientes, ¿qué crees que puede estar pasando? el siguiente gráfico, en el que se representan simultáneamente los tiempos obtenidos en las dos muestras, te puede servir de ayuda:

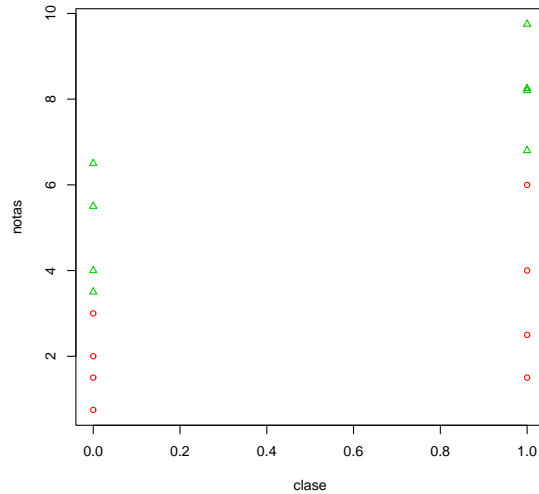


- Se conjetura que las acciones de una compañía sufrirían más variación en una industria con competencia en precios que en una en la que existiera un duopolio y colusión tácita. En un estudio sobre la industria de generadores mediante turbinas de vapor<sup>2</sup>, se halló que en 4 años de competencia en precios la variación de las acciones de la General Electric fue de 114.09. En los siguientes 7 años, en los cuales hubo un duopolio y colusión tácita, esta varianza fue de 16.08. Asumir que los datos pueden considerarse muestras aleatorias independientes de dos poblaciones normales y contrastar al 5% la conjetura anterior.
- Un usuario de grandes cantidades de componentes eléctricos adquiere éstos principalmente a dos proveedores, A y B. Debido a una mejor estructura en precios, el usuario hará negocios únicamente con el proveedor B siempre que la proporción de artículos defectuosos para B no exceda la del proveedor A (sea menos fiable). De 2 grandes lotes, el usuario selecciona al azar 125 unidades de A y 100 de B; inspecciona las unidades y encuentra 7 unidades defectuosas en cada una de las muestras (7 en la primera muestra y otras 7 en la segunda). Bajo las hipótesis adecuadas y con base a esta información, ¿existe alguna razón para no adquirir todas las componentes al proveedor B?
- Una profesora de Ciencias Políticas cree que el interés de los alumnos por la asignatura que imparte puede estar relacionado con su grado de participación en los procesos electorales del país. Con el fin de corroborar su hipótesis quiere comparar las notas medias de los alumnos que votaron en las últimas elecciones con la de aquellos alumnos que no lo hicieron. Para ello toma dos muestras aleatorias simples independientes de 114 alumnos que afirman haber votado y de 123 alumnos que manifiestan no haberlo hecho. En el primer grupo se obtuvo una calificación media de 2.71, con una desviación típica de 0.64. En el segundo grupo la nota media fue de 2.79, con una desviación típica de 0.56. Plantea el contraste oportuno especificando claramente la hipótesis nula y alternativa que consideras, así como las hipótesis de trabajo. Define la región de rechazo del contraste para un nivel de significación  $\alpha$  y obtén el  $p$ -valor del contraste.
- Se quiere analizar si la asistencia a clase tiene algún efecto sobre la calificación final que se obtiene en Estadística I. Para ello se tomaron muestras aleatorias independientes de 8 alumnos que asistieron regularmente a clase y de otros 8 que, por el contrario, rara vez habían asistido a clase. Se obtuvieron los siguientes resultados:

Regularmente (1)	2.5	1.5	7	4	8.2	6.8	9.5	8
Ocasionalmente (0)	0.75	1.5	3	1	4	4.5	6.5	5

<sup>2</sup>B.T.Allen, "Tacit collusion and market sharing: The case of steam turbine generators", *Industrial Organization Review* 4 (1976), 48-57. Ver Newbold.

- (a) Plantea un contraste bilateral para ver si hay diferencias significativas en las calificaciones medias de los alumnos debidas a la asistencia a clase. ¿A qué conclusión llegas para un nivel de significación de  $\alpha = 0.05$ ? Da una cota para el  $p$ -valor.
- (b) A la vista de los resultados obtenidos en el apartado anterior, ¿cambiaría la conclusión a la que llegas si te replanteas el contraste y optas por uno unilateral?, ¿cómo cambiaría el  $p$ -valor?
- (c) En el siguiente gráfico se representan los datos obtenidos teniendo además en cuenta las horas que cada alumno dedica al estudio de la asignatura fuera de las horas de clase (triángulos: más de 1 horas semanal de estudio, círculos: menos de 1 hora).

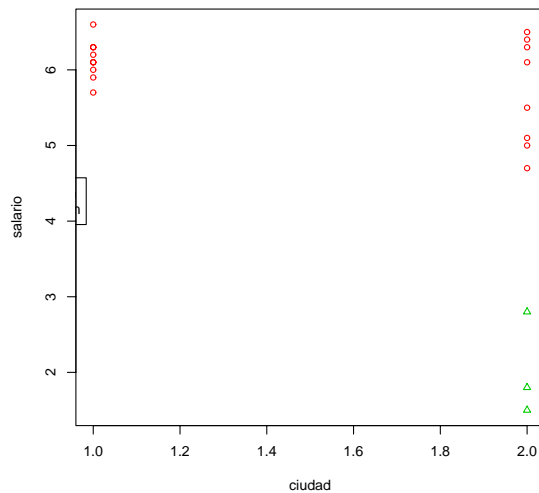


A la vista del gráfico, crees que aumentaría la potencia del contraste si se replantea el estudio para trabajar con muestras apareadas. En caso de que la respuesta sea afirmativa, ¿con qué criterio establecerías las parejas?

10. Estamos interesados en saber si los salarios medios en dos grandes ciudades de un país son homogéneos o si, por el contrario, se puede evidenciar que hay diferencias significativas en los salarios que perciben los trabajadores de estas 2 ciudades. Para ello se tomaron muestras aleatorias independientes de trabajadores (de actividades similares y niveles similares) en las 2 ciudades y se obtuvieron los siguientes datos:

Ciudad 1	5.9	6.1	6.3	6.1	6	6.2	5.7	6.3	6.6		
Ciudad 2	6.4	6.3	6.5	6.1	5	5.5	4.7	5.1	2.8	1.8	1.5

- (a) Plantea un contraste bilateral para ver si hay diferencias significativas en los salarios que perciben los trabajadores de estas 2 ciudades. ¿A qué conclusión llegas para un nivel de significación de  $\alpha = 0.1$ ? Da una cota para el  $p$ -valor.
- (b) En el siguiente gráfico se representan los datos obtenidos teniendo además en cuenta el sexo de los trabajadores (los triángulos corresponden a mujeres)



A la vista del gráfico, ¿te replantearías el estudio?

11. Para juzgar si la introducción de un nuevo plan de estudios ha mejorado la formación de los estudiantes se toma una muestra de 30 estudiantes del plan antiguo y 30 del nuevo. Los estudiantes se eligen de manera que tengan una nota similar de entrada en la universidad y la misma edad, para que las comparaciones sean homogéneas. Los 60 estudiantes realizan una misma prueba de conocimientos, obteniendo los estudiantes del plan antiguo una media de 7.2 y una cuasidesviación típica de 2,1. Para los estudiantes del nuevo plan la media es 7.4 con cuasidesviación típica 1.9. Para un nivel de significación del 1% se pide contrastar si esta introducción de un nuevo plan ha mejorado la formación de los estudiantes.

La respuesta debe incluir:

- Las hipótesis nula y alternativa para este contraste.
  - La expresión del estadístico de contraste y la región crítica del contraste.
  - Los cálculos necesarios para evaluar si los datos pertenecen a la región crítica, o para obtener el p-valor. Tu conclusión sobre el contraste.
  - Especificar los supuestos en los que se basa el análisis en los apartados anteriores.
12. Disponemos del número de días que cada paciente permanece ingresado en un hospital para dos muestras independientes de 39 hombres y 35 mujeres.

Hombres (1)					Mujeres (2)				
1	2	5	9	14	1	3	5	9	15
1	2	6	10	15	1	4	6	9	18
1	3	6	10	17	1	4	6	10	21
1	3	6	11	18	1	4	7	12	
1	3	7	12	19	2	4	7	12	
1	3	7	12	23	2	5	7	14	
2	4	8	13	24	3	5	7	14	
2	4	9	13		3	5	7	15	

$$\bar{x}_1 = 7.90, \quad s_1 = 6.41, \quad \bar{x}_2 = 7.11, \quad s_2 = 5.16$$

- Calcula un intervalo de confianza al 95% para el número medio de días que un hombre permanece ingresado en el hospital. Interpreta dicho intervalo. ¿Necesitamos asumir normalidad? ¿Por qué?
- Especificando claramente la hipótesis nula y la alternativa, ¿podemos afirmar que el número medio de días que una mujer permanece ingresada es significativamente diferente que el número medio de días que un hombre permanece ingresado? Calcula el p-valor.

- c) También estamos interesados en estudiar cuantos pacientes permanecen ingresados más de una semana (8 días o más). Con un nivel de significación del 10 %, y especificando claramente la hipótesis nula y la alternativa, ¿podemos afirmar que la proporción de pacientes que permanecen ingresados más de una semana es mayor para los hombres que para las mujeres?
13. Un fabricante asegura que la utilización de un aditivo en la gasolina hace que disminuya el consumo de los coches en al menos 2 litros por cada 100 Kilómetros. Para contrastar esta afirmación se ha evaluado la variable aleatoria “número de litros consumidos en 100 Kilómetros” para dos muestras. La primera la formaban los nueve coches en los que se había utilizado el aditivo, y la segunda aquellos trece coches en los que no se había utilizado. Los datos de la primera y segunda muestra verifican:

$$\sum_{i=1}^9 x_i = 74, \quad \sum_{i=1}^9 x_i^2 = 645, \quad \sum_{i=1}^{13} y_i = 130, \quad \sum_{i=1}^{13} y_i^2 = 1370$$

- a) Realiza un contraste de igualdad de varianzas para un nivel de significación del 10 %.
- b) Supongamos varianzas desconocidas e iguales y muestreos independientes. ¿Podemos creer lo que afirma el fabricante?

Especifica las hipótesis empleadas para llevar a cabo cada contraste, así como el estadístico del contraste, la región de rechazo y la conclusión a la que llegas.

14. De acuerdo a la última encuesta de pernoctaciones en establecimientos hoteleros (Encuesta de ocupación hotelera 2009) disponemos de datos que nos indican que el grado de ocupación del 2008 fue del 59,75 % mientras que en 2009 fue del 58,28 %. Si suponemos que los tamaños muestrales sobre los que se ha hecho la encuesta para conocer el grado de ocupación eran de 2400 establecimientos en el año 2008 y 2500 en 2009:

- a) Calcula un intervalo de confianza al 95 % para la diferencia de la ocupación en 2009 frente a la del 2008.
- b) Valora a través de un contraste y con un nivel de significación del 5 % si el porcentaje del grado de pernoctación ha disminuido en el 2009 en relación al 2008.
- c) Razona si el hecho que el intervalo contenga el cero está relacionado con el hecho que en el segundo apartado no hayamos rechazado la hipótesis nula de igualdad, pese a que es un contraste unilateral.
- d) ¿Cambiaría la respuesta a este razonamiento en los dos apartados si la confianza fuese del 99 % (nivel de significación del 1 %) en lugar del 95 % (nivel de significación del 5%)?