

## Relación 6 de problemas

---

### CONTRASTES PARAMÉTRICOS

1. Se analiza un envío de botellas sobre las que se afirma que contienen 100 cl de agua. Examinada una muestra de 5 botellas se obtiene que la media es de 95 cl. y la cuasivarianza muestral es  $s^2 = 1.1$ . Al nivel de significación 5 %, ¿existe evidencia empírica para afirmar que la cantidad media de agua no es de 100 cl?

2. La concentración media de dióxido de carbono en el aire a cierta altura es habitualmente de unas 355 ppm. Se sospecha que esta concentración es mayor en la capa de aire más próxima a la superficie. Para contrastar esta hipótesis se analizó el aire en 20 puntos elegidos aleatoriamente a una misma altura cerca del suelo. Resultó una media muestral de 580 ppm y una cuasi-desviación típica muestral de 180. Suponiendo normalidad para las mediciones, ¿proporcionan estos datos suficiente evidencia estadística, al nivel 0.01, a favor de la hipótesis de que la concentración es mayor cerca del suelo? Indicar razonadamente si el  $p$ -valor es mayor o menor que 0.01.

3. Una compañía petrolera afirma que más del 10 % de los hogares de una ciudad tienen calefacción de gasóleo. En una muestra aleatoria de 100 hogares de la ciudad se encontraron 14 con este tipo de calefacción.

¿Aportan estos datos evidencia empírica (a nivel  $\alpha = 5\%$ ) de que la afirmación de la compañía es cierta? Determina razonadamente si el  $p$ -valor del contraste utilizado es mayor o menor que 0.01.

4. Un fabricante de materiales para insonorización produce dos tipos A y B. De los 1000 primeros lotes vendidos, 560 fueron del tipo A. ¿Proporcionan estos datos suficiente evidencia estadística (al nivel de significación 1 %) para concluir que los consumidores prefieren mayoritariamente el tipo A?

5. Se tienen dos métodos, A y B, para determinar el calor latente de fusión del hielo. La siguiente tabla da los resultados obtenidos (en calorías por gramo de masa para pasar de  $-0.72^\circ\text{C}$  a  $0^\circ\text{C}$ ) utilizando ambos métodos independientemente:

Método A	79.98	80.04	80.02	80.04	80.03	80.03	80.04	79.97	80.05	80.03	80.02	80.00	80.02
Método B	80.02	79.94	79.98	79.97	79.97	80.03	79.95	79.97					

Con un nivel de significación del 10 %, ¿existen diferencias significativas entre los resultados medios proporcionados por los dos métodos?

6. Se recomienda que una persona mayor de 50 años consuma 15 mg de zinc al día en su dieta. En un informe sobre los hábitos alimentarios de una muestra de 100 individuos mayores de 50 años se señala que éstos consumieron una media de 11.3 mg de zinc diarios.

La cuasidesviación típica muestral correspondiente a estos datos fue de 6.43. Se supone que los datos siguen una distribución normal.

¿Permiten estos datos concluir (al nivel de significación del 5 %) que la ingesta media de zinc en la población de esa edad es inferior a la recomendada?

7. En un estudio sobre el efecto del ozono se mantuvo a un grupo de 23 ratas en un entorno sin ozono (grupo 1) y a otro grupo de 22 ratas en un entorno rico en ozono (grupo 2). Tras 7 días se midió el incremento en el peso de cada rata resultando que los incrementos medios fueron 22.40 g en el grupo 1 y 11.01 g en el grupo 2, mientras que las cuasidesviaciones típicas muestrales fueron 10.77 g y 19.02 g, respectivamente.

- Con los datos disponibles, ¿podemos afirmar a nivel  $\alpha = 0.01$  que el ozono influye en la variación de peso?
- Determina razonadamente si el  $p$ -valor del contraste es mayor o menor que 0.05.

8. Se han utilizado dos procedimientos de centrifugado para extraer el agua intersticial de 12 muestras diferentes de areniscas. Los resultados obtenidos (en términos del porcentaje de agua recuperada respecto al total disponible) han sido los siguientes:

Método	%						$\sum_i x_i$	$\sum_i x_i^2$
1	33.3	33.3	35.7	38.1	31.0	33.3	204.7	7013.77
2	43.6	45.2	47.7	45.4	43.8	46.5	272.2	12361.14

Bajo hipótesis de normalidad,

- determina mediante un contraste de nivel  $\alpha = 0.1$  si es aceptable la hipótesis de igualdad de varianzas de los resultados obtenidos por ambos métodos;
- ¿Permiten estos datos afirmar a nivel  $\alpha = 0.05$  que el método 2 es más eficiente que el 1?

9. En Nigeria hay miles de industrias madereras, que producen ingentes cantidades de residuos de madera. Algunos científicos han estudiado la posibilidad de extraer potasas de estos desechos en lugar de simplemente quemarlos y arrojarlos a los vertederos. Para ello en una fábrica concreta se tomaron muestras de serrín proveniente de nueve árboles (A1-A9). En cada muestra de serrín se midió la cantidad de potasa (en kg) por unidad de volumen (en m<sup>3</sup>) y luego se repitió la medida como comprobación. A continuación aparecen ambas medidas:

	A1	A2	A3	A4	A5	A6	A7	A8	A9
Primera medida	0.36	1.53	0.43	0.35	0.30	0.33	0.21	0.44	0.25
Segunda medida	0.36	1.42	0.47	0.32	0.31	0.32	0.20	0.45	0.24

¿Hay diferencias significativas a nivel  $\alpha = 0.05$  entre las concentraciones medias de potasa correspondientes a las dos medidas? Escribe claramente la hipótesis nula, la hipótesis alternativa y todas las condiciones necesarias para que las conclusiones sean válidas.

10. En un estudio se informa de los resultados de un experimento diseñado para comparar dos tratamientos de cierto tipo de cáncer, uno de ellos a base de quimioterapia y el otro basado en una combinación de quimioterapia y radioterapia. De 154 individuos que recibieron únicamente quimioterapia, 76 sobrevivieron más de 15 años, mientras que 98 de los 164 individuos que recibieron la combinación de quimioterapia y radioterapia sobrevivieron más

de 15 años. ¿Se puede afirmar con un nivel del 5 % que la probabilidades de sobrevivir más de 15 años es superior para el tratamiento combinado?

**11.** El nivel ceráunico de un lugar es el número de días al cabo del año en los que hay tormenta (se considera día con tormenta a aquel en el que al menos se oye un trueno). En cierta comarca, se conoce por datos históricos que el nivel ceráunico sigue una distribución de Poisson con una media ( $\lambda$ ) de 20 días. Sin embargo, se piensa que últimamente su nivel ceráunico posiblemente haya aumentado, como consecuencia del cambio climático. En un seguimiento de los últimos 50 años, se ha obtenido que el número medio de días al año con tormenta ha sido de 22. ¿Se puede afirmar que efectivamente se ha producido un aumento del nivel ceráunico? Da una respuesta al nivel de significación del 5 %.

**12.** Queremos comparar dos métodos rápidos para estimar la concentración de una hormona en una solución. Tenemos 10 dosis preparadas en el laboratorio y vamos a medir la concentración de cada una con los dos métodos. Se obtienen los siguientes resultados:

Dosis	1	2	3	4	5	6	7	8	9	10
Método A	10.7	11.2	15.3	14.9	13.9	15.0	15.6	15.7	14.3	10.8
Método B	11.1	11.4	15.0	15.1	14.3	15.4	15.4	16.0	14.3	11.2

Contrasta para un nivel de significación del 10 % si los dos métodos proporcionan, en media, las mismas estimaciones.

**13.** La existencia de trazas de metales en el agua afecta a su sabor y, si las concentraciones son altas, puede afectar a la salud. En un estudio se seleccionaron seis localizaciones en un río y, para cada localización, se determinó la concentración de zinc en el agua de la superficie y en el agua del fondo (en mg/l). Los resultados fueron los siguientes:

Localización	1	2	3	4	5	6
Concentración en el fondo ( $y$ )	0.43	0.27	0.57	0.53	0.71	0.72
Concentración en la superficie ( $x$ )	0.41	0.24	0.39	0.41	0.60	0.61

Asumiendo normalidad, ¿existe evidencia empírica para afirmar, con un nivel de significación  $\alpha = 0.05$ , que la concentración media de zinc en el fondo es diferente a la concentración media en la superficie?

---

## CONTRASTES NO PARAMÉTRICOS

**14.** Al lanzar un dado 300 veces se han obtenido las siguientes frecuencias:

valores	1	2	3	4	5	6
frecuencias	43	49	56	45	66	41

Al nivel de significación 0.05, ¿se puede afirmar que el dado está trucado?

**15.** Para estudiar el número de ejemplares de cierta especie en peligro de extinción que viven en un bosque, se divide el mapa del bosque en nueve zonas y se cuenta el número de ejemplares de cada zona. Se observa que 60 ejemplares viven en el bosque repartidos en las 9 zonas de la siguiente forma:

8	7	3	5	9	11	6	4	7
---	---	---	---	---	----	---	---	---

Mediante un contraste de hipótesis, analiza si estos datos aportan evidencia empírica de que los animales tienen tendencia a ocupar unas zonas del bosque más que otras.

**16.** Sea  $X$  la variable aleatoria que representa el número de defectos en una gran bobina de acero galvanizado. Se han inspeccionado 100 bobinas para las que se obtuvieron los siguientes valores de la variable  $X$ :

Valores	0	1	2	3
Frecuencias absolutas	20	25	30	25

- Si se supone que  $X$  tiene distribución de Poisson de parámetro  $\lambda$ , estima el valor de  $\lambda$  a partir de los datos utilizando el método de momentos. Calcula un intervalo de confianza para  $\lambda$  de nivel aproximado 0.95.
- A partir de los datos anteriores, ¿es aceptable el modelo de Poisson? Contesta a la pregunta mediante un contraste de bondad de ajuste de nivel  $\alpha = 0.05$ .

**17.** Se considera la hipótesis nula de que el tiempo de respuesta de un ordenador en un procedimiento sigue una distribución exponencial de media 1. Para contrastar esta hipótesis se desea aplicar un contraste de bondad de ajuste  $\chi^2$  basado en cinco intervalos equiprobables bajo  $H_0$ . Determina estos intervalos y aplica el contraste a la muestra:

0.10	0.99	1.14	1.26	3.24	0.12	0.26	0.80
0.79	1.16	1.76	0.41	0.59	0.27	2.22	0.66
0.71	2.21	0.68	0.43	0.11	0.46	0.69	0.38
0.91	0.55	0.81	2.51	2.77	0.16	1.11	0.02
2.13	0.19	1.21	1.13	2.93	2.14	0.34	0.44

**18.** Se clasificaron 1000 individuos de una población según el sexo y según fueran normales o daltónicos.

	Masculino	Femenino
Normal	442	514
Daltónicos	38	6

Según un modelo genético, las probabilidades deberían ser:

	Masculino	Femenino
Normal	$p/2$	$p^2/2 + pq$
Daltónicos	$q/2$	$q^2/2$

donde  $q = 1 - p =$  es la proporción de genes defectuosos en la población. A partir de la muestra se ha estimado que  $\hat{q} = 0.087$ . ¿Concuerdan los datos con el modelo?

**19.** Se ha desarrollado un modelo teórico para las diferentes clases de una variedad de moscas. Este modelo nos dice que la mosca puede ser de tipo L con probabilidad  $p^2$ , de tipo M con probabilidad  $q^2$  y de tipo N con probabilidad  $2pq$  ( $p + q = 1$ ). Para confirmar el modelo experimentalmente tomamos una muestra de 100 moscas, obteniendo 10, 50 y 40, respectivamente.

- Halla la estimación de máxima verosimilitud de  $p$  con los datos obtenidos.
- ¿Se ajustan los datos al modelo teórico, al nivel de significación 0.05?