

CONTRASTES  $\chi^2$

1. Al lanzar un dado 300 veces, se han obtenido las siguientes frecuencias:

Resultado	1	2	3	4	5	6
Frecuencias	43	49	56	45	66	41

Al nivel de significación 5%, ¿se puede aceptar que el dado es regular?

2. Nos dicen que un programa de ordenador genera observaciones de una distribución  $N(0, 1)$ . Como no estamos seguros de ello, obtenemos una muestra aleatoria de 450 observaciones, mediante dicho programa, obteniéndose los siguientes resultados:

Frecuencia	Rango
30	menores que $-2$ ;
80	entre $-2$ y $-1$ ;
140	entre $-1$ y $0$ ;
110	entre $0$ y $1$ ;
60	entre $1$ y $2$ ;
30	mayores que $2$ .

¿Se puede aceptar, al nivel  $\alpha = 1\%$ , que el programa funciona correctamente?

3. 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.

4. En 1778, H. Cavendish realizó una serie de 29 experimentos con objeto de medir la densidad de la tierra. Sus resultados, tomando como unidad la densidad del agua, fueron:

5.50	5.61	4.88	5.07	5.26	5.55	5.36	5.29	5.58	5.65
5.57	5.53	5.62	5.29	5.44	5.34	5.79	5.10	5.27	5.39
5.42	5.47	5.63	5.34	5.46	5.30	5.75	5.68	5.85	

Al nivel de significación 5%, ¿se puede aceptar que la densidad de la tierra se ajusta a una distribución normal?

5. Se desea estudiar el número de accidentes por día que se producen en cierto regimiento. Para ello se toman al azar los partes de 200 días dentro de los últimos 5 años, encontrando los siguientes resultados:

Número de accidentes en el día	0	1	2	3	4	5	6
Número de días	58	75	44	18	3	1	1

- a) ¿Se puede aceptar, con nivel de confianza del 90%, que el número de accidentes por día sigue una distribución de Poisson?
- b) Asumiendo que el número de accidentes por día sigue una  $\text{Poisson}(\lambda)$ , ¿hay suficiente evidencia estadística (tomar nivel de significación  $\alpha = 5\%$ ) de que el verdadero valor medio  $\lambda$  del número de accidentes por día es menor que 1.35? ¿El  $p$ -valor es mayor o es menor que 5%?

6. 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:

$\frac{1}{2}p$	$\frac{1}{2}p^2 + pq$
$\frac{1}{2}q$	$\frac{1}{2}q^2$

donde  $q = 1 - p$  = proporción de genes defectuosos en la población. A partir de la muestra se ha estimado que  $q = 8.7\%$ . ¿Concuerdan los datos con el modelo?

7. Hemos desarrollado un modelo teórico para las diferentes clases de una variedad de moscas. El modelo depende de un parámetro  $p$ , con  $0 < p < 1$ , y dice que la mosca puede ser de tipo  $L$  con probabilidad  $p^2$ , de tipo  $M$  con probabilidad  $(1 - p)^2$  y de tipo  $N$  con probabilidad  $2p(1 - p)$ . Para confirmar el modelo experimentalmente tomamos una muestra de 100 moscas, obteniendo 10, 50 y 40, respectivamente.

- a) Hallar la estimación de máxima verosimilitud de  $p$  con los datos obtenidos.  
b) ¿Se ajustan los datos al modelo teórico, al nivel de significación 5%?

8. Un ayuntamiento decide poner 4 contenedores para reciclar papel en una zona de la ciudad, con la idea de que sean utilizados por la misma cantidad de personas (aproximadamente). Para ver si esto es cierto, hace una encuesta en la zona a 300 personas, preguntándoles qué contenedor utilizan. Los resultados obtenidos son los siguientes:

Contenedor	1	2	3	4
Personas	80	70	85	65

- a) Como consecuencia de estos resultados, ¿resulta aceptable que los 4 contenedores se usen con la misma frecuencia? Dar una respuesta razonada, con un nivel de significación de 10%.  
b) El  $p$ -valor del contraste anterior, ¿es inferior o superior a 10%? Dar una respuesta razonada.

9. Se quiere comparar la biodiversidad de dos montes cercanos. Para esto hacemos lo siguiente: en uno de los montes se eligen al azar 50 zonas, de  $4m^2$  cada una, y se hace el recuento del número de especies vegetales diferentes que hay en cada una, con los resultados de la tabla de la izquierda. En el otro monte se hace el mismo recuento en otras 40 zonas, obteniéndose los resultados de la tabla de la derecha:

Número de zonas	Número de especies	Número de zonas	Número de especies
20	menos de 6	12	menos de 6
17	entre 6 y 8	20	entre 6 y 8
13	más de 8	8	más de 8

¿Son similares los dos montes en lo que se refiere a su biodiversidad? Hacer el contraste correspondiente con un nivel de significación del 10%.

10. Se desea evaluar la efectividad de una nueva vacuna antigripal. Para ello se decide suministrar dicha vacuna, de manera voluntaria y gratuita, a una pequeña comunidad. La vacuna se administra en dos dosis, separadas por un período de dos semanas, de forma que algunas personas han recibido una sola dosis, otras han recibido las dos y otras personas no han recibido ninguna. La siguiente tabla indica los resultados que se registraron durante la siguiente primavera en 1000 habitantes de la comunidad elegidos al azar.

	No vacunados	Una dosis	Dos dosis
Gripe	24	9	13
No gripe	289	100	565

¿Proporcionan estos datos suficiente evidencia estadística (al nivel de significación 5%) para indicar una dependencia entre la clasificación respecto a la vacuna y la protección frente a la gripe?