Grado en Ingeniería Informática, curso 2015-16
Probabilidad y Estadística — Hoja 7

Contrastes paramétricos de hipótesis

1. La estatura X de los individuos varones de una cierta región sigue una distribución normal $\mathcal{N}(\mu, \sigma)$, con $\sigma = 6$ cm. Deseamos contrastar la hipótesis de que el valor medio de X es 180 cm, con el nivel de significación de 5 %. En una muestra al azar de 9 personas se han obtenido los siguientes resultados: 168, 180, 170, 175, 171, 173, 169, 184, 176. ¿Qué podemos concluir? ¿Y si no conociéramos σ?

2. Los lagartos del desierto se esconden del calor en verano para evitar que su temperatura corporal interna llegue al nivel letal de 45°C. Para estudiar el tiempo X (en minutos) requerido para que la temperatura de un lagarto alcance los 45°C partiendo de su temperatura a la sombra, se ha tomado una muestra, con los siguientes resultados: 10.1, 12.5, 12.2, 10.2, 12.8, 12.1, 11.2, 11.4, 10.7, 14.9, 13.9, 13.3.

Suponiendo que la variable X sigue una distribución $N(\mu, \sigma)$, y basándose en los datos obtenidos,

a) ¿puede concluirse (al nivel de significación 2.5 %) que el tiempo medio requerido para alcanzar la temperatura letal es menor que 13 minutos?;

b) ¿Puede concluirse (al mismo nivel) que la desviación típica de X es inferior a un minuto y medio?

3. Queremos contrastar la hipótesis de que una cierta moneda es equilibrada. Para ello, la lanzamos un número n de veces y anotamos el número de caras obtenidas.

a) Supongamos que $n = 200$ y pongamos un nivel de significación $\alpha = 5\,\%$. ¿Qué debe ocurrir en el experimento (los 200 lanzamientos de moneda) para que rechacemos la hipótesis?

b) Tomamos n general (pero grande). Fijamos el nivel de significación $\alpha = 5\,\%$. En el experimento sale un 57 % de caras. ¿Para qué valores de n rechazaremos la hipótesis, y para cuáles la aceptaremos?

c) Calcula explícitamente el p-valor de una muestra de tamaño n (grande) y media muestral $\overline{x}$.

4. Un grupo de investigadores afirma haber descubierto un tipo de alimento que no aumenta el colesterol en las personas que lo consumen. Para comprobarlo, se seleccionaron al azar 10 personas a las que se les midió su nivel de colesterol antes (X) y después (Y) de ser sometidos a una dieta a base de ese alimento. Suponiendo normalidad, contrastar la hipótesis nula de que el nivel de colesterol es el mismo antes y después de la dieta (al nivel 5 %), si los datos obtenidos son los siguientes:

X	120	312	243	161	314	234	143	287	423	155
Y	130	306	255	168	310	250	158	290	440	140

5. La siguiente tabla recoge los resultados de la medición de una cierta propiedad mediante dos técnicas experimentales diferentes.

Lote	1	2	3	4	5	6	7	8	9	10
Técnica A	84.63	84.38	84.08	84.41	83.82	83.55	83.92	83.69	84.06	84.03
Técnica B	83.15	83.72	83.84	84.20	83.92	84.16	84.02	83.60	84.13	84.24

Basándose en estos datos, al nivel de significación $\alpha = 5\,\%$, ¿hay diferencias entre ambas técnicas?

6. Se pretende estudiar la proporción de personas que apoyan una cierta iniciativa en dos regiones distintas. En la región 1 se toma una muestra de 100 personas, de las que un 60 % se decanta por el "sí". En la región 2, de las 250 personas encuestadas, 175 votarían "sí".

a) Contrasta, con nivel de significación $\alpha = 5\,\%$, la hipótesis de que las proporciones coinciden en la dos regiones.

b) Contrasta, con nivel de significación $\alpha = 5\,\%$, la hipótesis de que la proporción en la región 1 es menor que la proporción en la región 2. Calcula el p-valor en ambos apartados.

7. La siguiente tabla proporciona la concentración observada de tiol en el lisado sanguíneo de dos grupos de voluntarios. El primer grupo es "normal", mientras que el segundo padece artritis reumatoide.

Normal:	1.84	1.92	1.94	1.92	1.85	1.91	2.07
Artritis:	2.81	4.06	3.62	3.27	3.27	3.76	

¿Se puede considerar que la concentración media de tiol es igual en ambos grupos?

8. Se toman muestras del peso X en gallinas de dos granjas. En la primera granja se han pesado 140 gallinas, obteniéndose un peso medio de 3.5 kg, con una cuasidesviación típica de 0.5 kg. En la otra granja se pesaron 100 animales, con un peso medio de 3.2 kg y cuasidesviación típica de 1 kg. Fijamos un nivel de significación $\alpha = 5\,\%$. Suponiendo normalidad,

a) ¿se puede concluir que el peso medio de las gallinas es el mismo en las dos granjas?;

b) ¿se puede concluir que la variabilidad del peso de las gallinas es menor en la primera granja?

Contrastes χ^2

9. 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?

10. 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	30	80	140	110	60	30
Rango	≤ -2	$(-2, -1]$	$(-1, 0]$	$(0, 1]$	$(1, 2]$	> 2

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

11. 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 como se muestra a la derecha.

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.

12. Se desea estudiar el número de accidentes por día que se producen en cierto regimiento. El análisis de los partes de 200 días escogidos al azar dentro de los últimos 5 años ofrece 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 Poisson(λ), ¿hay suficiente evidencia estadística (tomar nivel de significación $\alpha = 5\,\%$) de que el verdadero valor medio λ del número de accidentes por día es menor que 1.35? ¿El p-valor es mayor o es menor que 5 %?

13. 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 como muestra la tabla de la derecha, 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?

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

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