

## Tema 11.- Test de hipótesis

- Supongamos que  $\mu$  es la media de una distribución normal de varianza unidad  $\sigma^2 = 1$ . Queremos testear la hipótesis nula  $H_0 : \mu = 2$  frente a la hipótesis alternativa  $H_1 : \mu = 0$  para una muestra aleatoria de tamaño  $n = 4$  para una significancia  $\alpha = 0.05$  (*single-sided*).
  - Calcularla las regiones de aceptación y rechazo
  - Calcular la potencia del test  $1 - \beta$  del test.
  - Estimar el número de medidas necesarias para conseguir la misma potencia pero con una significancia del test de  $5\sigma$
- Supongamos que lanzamos 20 monedas y 17 de ellas son cara. ¿Cuál es el p-value de este resultado?
- Para comprobar la hipótesis de que la probabilidad de que al lanzar una moneda salga cruz es de  $p = 0.5$ , frente a la hipótesis de que no lo sea, se realiza un test que consiste en realizar diez lanzamientos independientes. La hipótesis  $p = 0.5$  se rechaza si se obtiene un resultado de diez caras o diez cruces.
  - ¿Cuál es el nivel de significancia del test?
  - Si resulta que la probabilidad de que salga cruz es  $p = 0.1$ , ¿Cuál es la potencia (*power*) del test?
- Supongamos un experimento de conteo de sucesos donde esperamos un fondo de 0.5 sucesos. Realizamos la medida y obtenemos un número total de 5 sucesos. ¿Podemos anunciar que hemos realizado un descubrimiento (*discovery*)? En caso contrario, ¿Cuántos sucesos deberíamos haber medido?
- Los últimos resultados de IceCube sobre búsqueda de neutrinos cósmicos con energías por encima de 100 TeV anuncian un total de 28 sucesos sobre un fondo esperado de neutrinos atmosféricos de 12.1 sucesos ¿Podemos decir que se trata de un descubrimiento (*discovery*)? En caso contrario, ¿cuántos sucesos deberían haber sido observados como mínimo para publicar un descubrimiento?
- Dos conjuntos diferentes de 22 medidas supuestamente gaussianas arrojan los siguientes valores medios:  $\bar{x}_1 = 181.87$  y  $\bar{x}_2 = 183.04$ . La estimación de la desviaciones estándares a partir de los datos proporciona los valores  $s_1 = 6.59$  y  $s_2 = 7.00$ . A la vista de estos resultados, ¿existe alguna razón para rechazar la hipótesis de que ambas distribuciones tienen el mismo valor medio?
- Diez estudiantes toman una muestra de medidas todas con un error de 0.2 g:

10.2	10.4	9.8	10.5	9.9	9.8	10.3	10.1	10.3	9.9
------	------	-----	------	-----	-----	------	------	------	-----

- Testear la hipótesis de que todas las medidas proceden de una muestra cuya verdadera masa es 10.1g.
  - Repetir el test suponiendo que no conocemos la resolución de las medidas.
  - Testear la hipótesis de que todas las medidas proceden de la misma muestra. Estimar el número de grados de libertad y el  $\chi^2$ .
- El valor medio de una muestra aleatoria de 16 medidas obtenidas de una distribución normal es de  $\bar{x} = 41.5$  y la estimación de la desviación estándar es de  $s = 2.795$ . A la vista de estos resultados, ¿existe alguna razón para rechazar la hipótesis de que la media verdadera es  $\mu = 43.0$  ?
  - Dos conjuntos diferentes de 16 y 7 medidas respectivamente arrojan los siguientes valores medios:  $\bar{x}_1 = 20.3$  y  $\bar{x}_2 = 19.0$ . La estimación de la desviaciones estándares a partir de los datos proporciona los valores  $s_1 = 3.38$  y  $s_2 = 1.83$ . A la vista de estos resultados, ¿existe alguna razón para rechazar la hipótesis de que ambas distribuciones tienen el mismo valor medio?

10. El 23 de febrero de 1987, el experimento IMB (Irvine-Michigan-Brookhaven) realizaba medidas de interacciones de neutrinos. El tiempo durante el cual el detector estaba funcionando fue dividido en intervalos de 10 segundos de duración, de manera que en cada intervalo se registraba el número de interacciones de neutrinos detectadas. El número de intervalos con “ $i$ ” sucesos se muestra en la siguiente tabla:

Número de sucesos	0	1	2	3	4	5	6	7	8	9
Número de intervalos	1042	860	307	78	15	3	0	0	0	1

Ese mismo día fue cuando los astrónomos observaron la supernova SN1987A. Testear la hipótesis de que los datos vienen descritos por una distribución de Poisson (cuyo valor medio es un parámetro libre), encontrar el valor de  $\chi^2$  y el p-value.

- Tomando todos los datos salvo el intervalo con 9 sucesos.
- Tomando todos los datos.

11. En un experimento que investiga interacciones protón-fotón, se hace incidir un haz de fotones de alta energía en una cámara de burbujas. Para medir la intensidad del haz, se cuentan las veces que se producen conversiones de pares electrón-positrón. La frecuencia de casos en los que se observan 0,1, 2,... pares simultáneamente, es decir, en la misma fotografía sigue una distribución de Poisson. Desviaciones del comportamiento poissoniano pueden proporcionar información sobre errores sistemáticos. Los resultados de observar  $n = 355$  fotografías se muestran en la siguiente tabla:

12.

Número de pares $e^+e^-$ por fotografía ( $k$ )	0	1	2	3	4	5	6	7	8
Número de fotos con $k$ pares $e^+e^-$	47	69	84	76	49	16	11	3	0

Testear si los datos vienen descritos por una distribución de Poisson mediante un test de  $\chi^2$  al 1% de CL.

13. Los métodos *binned* para detectar fuentes puntuales de neutrinos cósmicos se basan en dividir el cielo observable en celdas donde la probabilidad de observar sucesos de fondo es constante. Supongamos que dividimos el cielo en 4000 celdas y contamos los sucesos observados en cada una de ellas siendo los resultados los que se muestran en la tabla:

Número de sucesos	0	1	2	3	4
Número de celdas	2758	1023	186	31	2

- Calcular el número medio de sucesos observado por celda.
- Testear la hipótesis de que los datos vienen descritos por una distribución de Poisson mediante un test de  $\chi^2$  al 1% de CL.
- Si hubiéramos encontrado una celda con 6 sucesos ¿Con qué rotundidad podríamos decir que se trata de una fuente de neutrinos cósmica?

14. En una distribución plana donde el fondo promedio por bin es de 1.6 sucesos.

- Si observamos dos bins consecutivos con 6 y 5 sucesos ¿Cuál es el p-value correspondiente a dicho resultado?
- Calcular el valor del  $\chi^2$  que se obtendría mediante un test de Pearson.
- ¿Está justificado el uso de dicho test en este caso?

