

Principios de Bioestadística

Dra. Juliana Giménez

www.cii.org.ar

Nos permite

- Llegar a conclusiones correctas acerca de procedimientos para el diagnóstico
- Valorar protocolos de estudio e informes

Se pretende

- Inferir o generalizar resultados de una muestra a una población
- Esta inferencia se efectúa por medio de métodos estadísticos basados en la probabilidad

Los cinco pilares de la estadística

- Hipótesis
- Probabilidad
- Variables
- Población y muestra
- Inferencia

Hipótesis

Hipótesis nula (H_0):

- La relación entre variables se debe exclusivamente al azar
- Las diferencias observadas se deben al azar (No hay relación causa-efecto)

Hipótesis alternativa (H_1):

- La relación entre variables no es debida al azar
- Las diferencias observadas no son debidas al azar (relación causa-efecto)

Hipótesis y error

Error de tipo I o error α .

(falso positivo)

Puede rechazarse la hipótesis nula cuando en realidad es verdadera.

Error de tipo II o error β .

(falso negativo)

Probabilidad de dar por cierta H_0 cuando es cierta H_1 .

Aceptamos	H0 No relación Causa-efecto	H1 Relación Causa-efecto	Probabilidad total del acierto
H0 No relación Causa-efecto	Acierto 1- β (0.80)	Error β Falso negativo (0.20)	1
H1 Relación Causa -efecto	Error α Falso positivo (0.05)	Acierto 1- α Verdadero positivo (0.95)	1

Probabilidad

Ocurrencia de un determinado suceso que puede definirse como la proporción de veces que ocurriría dicho suceso si se repitiese un experimento o una observación en un número grande de ocasiones bajo condiciones similares.

Poblaciones y muestras

La población representa el conjunto de individuos que deseamos estudiar



La muestra es el subconjunto de individuos de la población sobre la cual realizamos las mediciones o experimento

El individuo es cada uno de los componentes de la población y la muestra

¿Porque trabajar con muestras?

Generalmente es inaccesible trabajar con toda la población

Por esto la muestra debe ser representativa de la población



¿Que estudiamos en la muestra?

Las variables

Cuantitativas

- Continuas (edad, peso, talla)
- Discretas (número de hijos)

Cualitativas

- Enfermo/sano
- Grupo sanguíneo

Por su asociación

- Dependientes
- Independientes



DATOS

Estadística descriptiva

Medidas de tendencia central

- Media , mediana , moda.

Medidas de dispersión

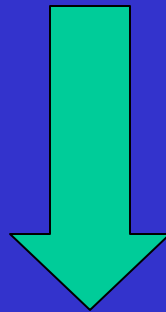
- Varianza, desvió estándar, coeficiente de variación.

Estadística de inferencia

Contraste de hipótesis

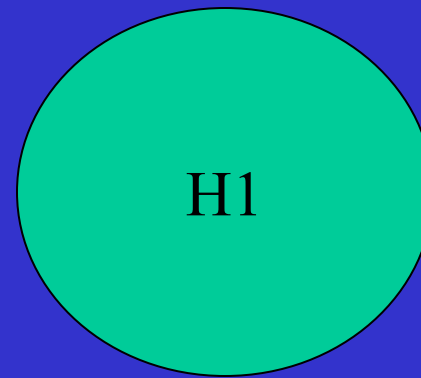
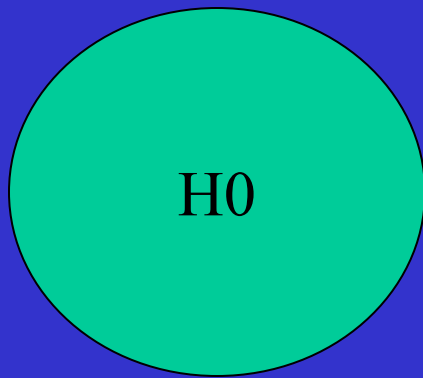
- El análisis de estudios clínico-epidemiológicos con frecuencia exige la comparación entre varios tratamientos o entre diferentes grupos de sujetos con respecto a una respuesta de interés.
- La hipótesis que se contrasta es la de que no existen diferencias entre ambos grupos o tratamientos (*hipótesis nula*), frente a la *hipótesis alternativa* de que estos sean realmente diferentes.

La probabilidad de cometer un error de este tipo es lo que mide precisamente el *grado de significación p*.

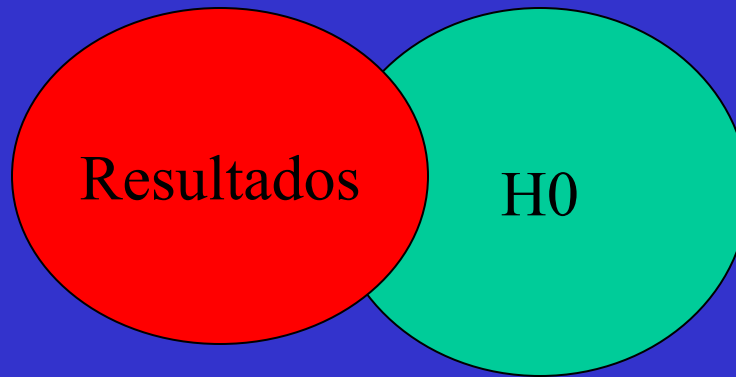


¡¡¡¡¡Si!!!!!! Esta es la famosa P

Prueba de hipótesis



Según los datos que tenga, pondré a prueba la hipótesis con un test de hipótesis conveniente



$P > 0.05$



$P < 0.05$

Rechazo H_0

¡Ejemplo!

Se quiere estudiar la posible asociación entre madre gestante fumadora y bajo peso del bebe al nacer.

1) Se plantean las hipótesis

H_0 = la relación entre las variables se debe al azar.

No hay asociación entre las variables

H_1 = la relación entre las variables no es al azar, existe una relación causa - efecto

2) Variables

- Habito de fumar (fumadora- no fumadora)
- Peso del recién nacido (bajo peso- peso normal)

Las variables son independientes

3) Población y muestra

Población: Todas las mujeres gestantes.

Muestra: 2000 gestantes que participan del estudio.

4) Contraste de hipótesis

Elijo una prueba de contraste de hipótesis, que contemple variables independientes aleatorias cualitativas.

La prueba χ^2 permite determinar si dos variables cualitativas están o no asociadas.

Este estadístico asume:

$$\chi^2 = \sum \frac{(\text{observado} - \text{esperado})^2}{\text{esperado}}$$

Obtención del esperado:

$$A = [(a+c) \times (a+b)] / n$$

$$B = [(b+d) \times (a+b)] / n$$

$$C = [(a+c) \times (c+d)] / n$$

$$D = [(b+d) \times (c+d)] / n$$

Frecuencias
observadas

	Recién nacido de bajo peso		
Gestante	Sí	No	Total
Fumadora	43 (a)	207 (b)	250
No fumadora	105 (c)	1645 (d)	1750
Total	148	1852	2000

Gestante	Recién nacido de bajo peso		Total
	Sí	No	
Fumadora	43 (18.5)	207 (231.5)	250
No fumadora	105 (129.5)	1645 (1620.5)	1750

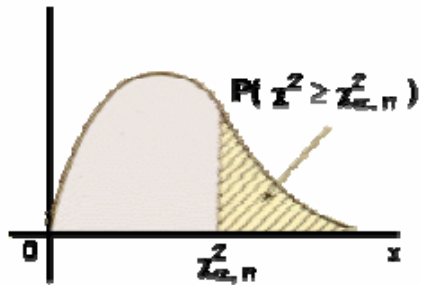
Frecuencias
observadas
y esperadas

$$\chi^2 = \sum \frac{(\text{observado} - \text{esperado})^2}{\text{esperado}}$$

$$\chi^2 = \frac{(43 - 18.5)^2}{18.5} + \dots = 40.04$$

Valor del estadístico
para este caso

Los valores del estadístico se distribuyen según una distribución Chi-cuadrado que depende de un parámetro g.l (grados de libertad) en una tabla r filas x k columnas. Los g.l $(r-1) \times (k-1) = 1$



De ser cierta la H_0 el valor del estadístico debería estar dentro del rango de la tabla.

Con una seguridad del 95% ($\alpha=0.05$) el valor teórico es 3,84.

El valor de nuestro estadístico es 40,04 supera el valor de tabla.

Grados de libertad	Probabilidad de un valor superior				
	0,1	0,05	0,025	0,01	0,005
1	2,71	3,84	5,02	6,63	7,88
2	4,61	5,99	7,38	9,21	10,60
3	6,25	7,81	9,35	11,34	12,84
4	7,78	9,49	11,14	13,28	14,86
5	9,24	11,07	12,83	15,09	16,75
6	10,64	12,59	14,45	16,81	18,55
7	12,02	14,07	16,01	18,48	20,28
8	13,36	15,51	17,53	20,09	21,95
9	14,68	16,92	19,02	21,67	23,59
10	15,99	18,31	20,48	23,21	25,19
11	17,28	19,68	21,92	24,73	26,76
12	18,55	21,03	23,34	26,22	28,30
13	19,81	22,36	24,74	27,69	29,82
14	21,06	23,68	26,12	29,14	31,32
15	22,31	25,00	27,49	30,58	32,80
16	23,54	26,30	28,85	32,00	34,27
17	24,77	27,59	30,19	33,41	35,72
18	25,99	28,87	31,53	34,81	37,16
19	27,20	30,14	32,85	36,19	38,58
20	28,41	31,41	34,17	37,57	40,00
21	29,62	32,67	35,48	38,93	41,40
22	30,81	33,92	36,78	40,29	42,80
23	32,01	35,17	38,08	41,64	44,18
24	33,20	36,42	39,36	42,98	45,56
25	34,38	37,65	40,65	44,31	46,93
26	35,56	38,89	41,92	45,64	48,29
27	36,74	40,11	43,19	46,96	49,65
28	37,92	41,34	44,46	48,28	50,99
29	39,09	42,56	45,72	49,59	52,34
30	40,26	43,77	46,98	50,89	53,67
40	51,81	55,76	59,34	63,69	66,77
50	63,17	67,50	71,42	76,15	79,49
60	74,40	79,08	83,30	88,38	91,95
70	85,53	90,53	95,02	100,43	104,21
80	96,58	101,88	106,63	112,33	116,32

En este caso:

Rechazamos la hipótesis nula (H0) y
aceptamos la hipótesis alternativa (H1)

$$\chi^2 = 40.04 > 3,84 \quad p = 0,05$$

Conclusión:

Existe asociación entre las variables, es
decir esta diferencia no se debe al azar.

El habito de fumar en la mujer embarazada
se relaciona significativamente con un
bajo peso del recién nacido

La estadística no aporta
soluciones mágicas por si sola



Solo nos dice que probabilidad
hay de certeza sobre la
hipótesis planteada

juliana.gimenez@pfizer.com

jg@cii.org.ar