UNIVERSIDAD POLITÉCNICA DE MADRID

conento

Aprendiendo idiomas V

María Jesús Vázquez Gallo

Disco de Festos. 2º milenio a. C.
Encontrado en Creta.
Sus jeroglíficos siguen sin ser descifrados.

Formación Conento: enero 2009

1. Introducción.

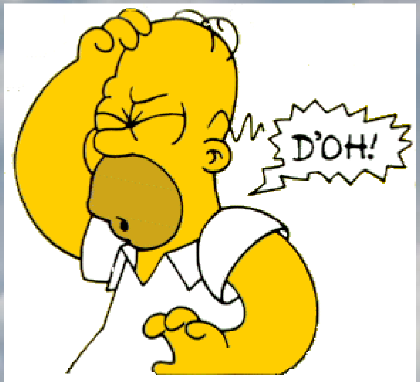
2. Contrastes de hipótesis y errores.

3. Potencia de un contraste.

4. Continuará…

La última vez…

Aprendiendo idiomas IV



estábamos hablando de errores…

…errores inherentes al contraste de hipótesis.

El **contraste de hipótesis** es una herramienta estadística para juzgar si cierta característica de una población es compatible con una muestra extraída de ella.

Ejemplo: Un fabricante de baterías de móvil afirma que, en promedio, duran 1000 minutos. Se selecciona una muestra aleatoria de 30 baterías. La media de la muestra es 1003 min. y la desviación típica de 5 min. ¿Hay motivos para desconfíar de la información aportada por el fabricante?

Si tomamos una decisión basada en un contraste **¿podemos equivocarnos?** En caso afirmativo, **¿con qué probabilidad?**

Hablaremos aquí sobre errores en los contrastes.

De los dos tipos de contrastes mencionados en **Aprendiendo idiomas IV**

- **Paramétricos** Hipótesis sobre un valor concreto o un intervalo para los parámetros del modelo.
  Ejemplo: Media igual a 1000.
- **No Paramétricos** Hipótesis sobre el tipo de distribución de probabilidad que ha generado los datos.
  Ejemplo: Distribución normal.

Hablaremos la próxima vez de los contrastes paramétricos.

En el contraste de hipótesis: el carácter probabilista del modelo hace que **todos los resultados** sean **estrictamente compatibles** con las **hipótesis**, o sea, cualquier conjunto de datos puede llegar a observarse, tanto si es cierta la hipótesis nula H_0 como si lo es la alternativa H_1.

No se trata de una verificación ideal de hipótesis.

Por tanto, en general:
No es posible descartar sin error ninguna de las dos hipótesis.

No sabemos con seguridad si la hipótesis que contrastamos con la información muestral –la hipótesis nula H_0- es cierta o es falsa.

¡La hipótesis que se contrasta nunca se considera probada!

Aceptamos o rechazamos dicha hipótesis con un nivel de significación α elegido previamente:
α es la probabilidad que hemos querido asignar al hecho de rechazar la hipótesis nula siendo cierta (error de tipo I).

El otro error posible es aceptar la hipótesis nula siendo falsa (error de tipo II) y su probabilidad se denota por β.

La probabilidad complementaria $1-\beta$, es decir, la de rechazar la hipótesis nula siendo falsa, se llama potencia del contraste.

No es posible reducir **simultáneamente** los dos errores en un contraste de hipótesis.

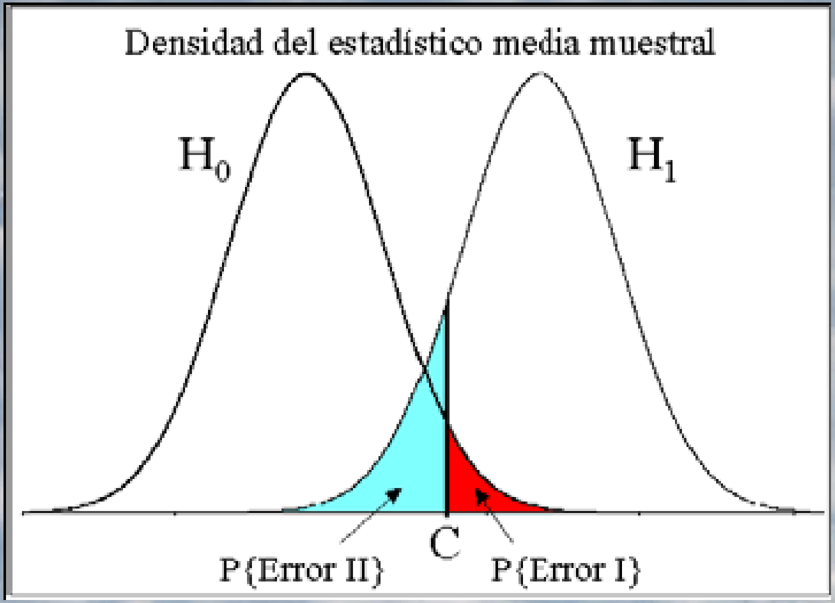
Supongamos tomamos el nivel de significación α igual a cero. Esto equivale a que la probabilidad de que rechacemos H_0, siendo cierta, sea cero.

En la mayoría de situaciones aplicadas, esto produce una región de rechazo vacía, es decir, conlleva aceptar SIEMPRE H_0.

Pero esto es absurdo:
¡entonces podríamos prescindir de la muestra!

α=P(error I)=0 ⟶ rechazar siempre H_1 ⟷ β=P(error II)=1

β=0 ⟶ rechazar siempre H_0 ⟷ α=1



Los dos errores están relacionados: al disminuir uno aumenta el otro.

Un contraste debería buscar simultáneamente el nivel de significación α más bajo posible y la potencia 1-β más alta posible.

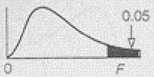
Pero al disminuir α, aumenta β y, por tanto, desciende la potencia.

Fijado el nivel de significación, se determina la región de rechazo cuya potencia es mayor entre todos los contrastes cuyo tamaño sea el fijado a priori.

¿Cómo?

Si las hipótesis son simples (es decir, bajo cada hipótesis sólo existe una distribución de probabilidad y, por tanto, existe una única probabilidad de error de tipo I y error de tipo II) entonces existe el contraste uniformemente de máxima potencia y se obtiene mediante el llamado Lema de Neyman-Pearson.

Pero ¿cómo elegir el nivel de significación?



Grados de libertad para el denominador	Grados de libertad para el numerador															
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39

Los valores 5 %, 1 % y 10 % fueron inicialmente elegidos como los más representativos en las colecciones de tablas que se utilizaban antes del la universalización del ordenador. Con el paso del tiempo estos valores se fueron convirtiendo en una convención.

Entonces

¿se elige uno de esos tres?, ¿cuál se prefiere?, ¿por qué?

La metodología más razonable es escoger un nivel α de acuerdo a la experiencia y después obtener el llamado *p-valor.*

El p-valor es el **nivel de significación más pequeño posible** que se puede escoger, para el que todavía se **rechazaría la hipótesis nula** con la muestra actual.

Escoger un nivel de significación menor que el p-valor conllevaría aceptar H_0.

El p-valor es una medida directa de lo creíble que resulta obtener una muestra como la actual si es cierta H_0.

En otras palabras, el p-valor es el menor α que permite aceptar H_1. Cuando el p-valor es pequeño (digamos, menor o igual que 0.05) se considera que hay una fuerte evidencia a favor de H_1.

¿para qué sirve el p-valor?

Conocer el p-valor tiene la ventaja de permitir que cualquiera decida qué hipótesis acepta basándose en su **propio nivel de significación α**. Esto no es posible cuando se indica sólo el resultado del contraste: si se acepta o se rechaza H_0 con un α fijo.

Al proporcionar el p-valor obtenido con la muestra actual, la decisión se hará de acuerdo a la regla siguiente:

$$\begin{cases} \text{si } \alpha >= \text{p-valor, rechazar } H_0 \\ \text{si } \alpha < \text{p-valor , aceptar } H_0 \end{cases}$$

El p-valor es un valor objetivo, cualquier experimentador dará el mismo valor una vez obtenida la muestra, mientras que el nivel de significación α es subjetivo, escogido por el experimentador.

En las situaciones prácticas más habituales, los contrastes óptimos ya están determinados, y el experimentador sólo ha de escoger el nivel de significación α que considere adecuado.

Escogido α quedan fijadas la región de rechazo y la potencia.

Para conseguir que un contraste mejore su potencia sin que ello implique un aumento excesivo de α la única posibilidad es **incrementar el tamaño de la muestra**.

Al aumentar el tamaño de la muestra, varía la ley de distribución del estadístico de contraste, y generalmente disminuye la varianza. En general, las propiedades del contraste mejoran.

¿Cuál es el tamaño necesario de la muestra?

Es una pregunta abierta.


Continuará...