Test Multiples

Graciela Boente[1]

[1]Universidad de Buenos Aires and CONICET, Argentina

Ejemplo

En muchos ejemplos, se realizan múltiples tests y se desea combinar los resultados de estos tests independientes que se realizan sobre la misma hipótesis nula. Ejemplos de tales situaciones son

- Combinar los resultados obtenidos de efectividad de un medicamento en distintos hospitales o centros de salud, donde en general se toma un número grande de hospitales.
- Combinar resultados en microarreglos.

Ejemplo

Los microarreglos de ADN permiten medir los niveles de expresión de distintos genes.

Cuanto mayor sea el valor observado, más activo es el gen.

La tabla siguiente muestra el nivel de expresión de genes en 10 pacientes con dos tipos de cancer de hígado.

Hay 2638 genes pero solo mostramos los primeros dos. Los datos son el logaritmo del cociente de la intensidad de dos colorantes usados en los arreglos.

	Tipo I					Tipo II				
Paciente	1	2	3	4	5	6	7	8	9	10
Gen 1	230	-1350	-1580	-400	-760	970	110	-50	-190	-200
Gen 2	470	-850	-.8	-280	120	390	-1730	-1360	-1	-330

Interesa saber si el nivel medio (o mediano) del gen es diferente entre los dos grupos.

Ejemplo

Si testearamos diferencias sobre cada gen, estaríamos haciendo 2638 tests separados.

Si cada test tiene nivel α para cada gen, la probabilidad de falso rechazo de la igualdad entre ambos grupos es α.

Pero la probabilidad de **al menos un falso rechazo** es mucho mayor.

Este es el **problema de test múltiples**.

Ocurre en muchas situaciones y veremos algunos métodos para tratar este problema.

Supongamos que hemos realizado k tests independientes para testear

$$H_{0,i} \quad \text{versus} \quad H_{1,i} \quad 1 \leq i \leq k$$

obteniendo, respectivamente, los $p-$valores

$$p_1, \ldots, p_k$$

Supongamos que hemos realizado k tests independientes para testear

$$H_{0,i} \quad \text{versus} \quad H_{1,i} \quad 1 \leq i \leq k$$

obteniendo, respectivamente, los $p-$valores

$$p_1, \ldots, p_k$$

- La hipótesis global o hipótesis nula del meta–analisis es que todas las hipótesis nulas son ciertas ($H_0 = \cap_{i=1}^{k} H_{0,i}$)

- La hipótesis alternativa es que al menos una de las alternativas $H_{1,i}$ es cierta.

Cuando estudiamos estas hipótesis nos interesa responder dos preguntas:

1. Podemos rechazar H_0?
2. Si H_0 es rechazada, cual de las $H_{0,i}$ debe ser rechazada?

Cuando estudiamos estas hipótesis nos interesa responder dos preguntas:

1. Podemos rechazar H_0?
2. Si H_0 es rechazada, cual de las $H_{0,i}$ debe ser rechazada?

El procedimiento estadístico para responder 1) se llama *test global*, mientras que el utilizado para responder 2) es un *procedimiento de test múltiples*.

El método de Fisher es un *test global* y se aplica típicamente cuando testeamos la misma hipótesis en estudios independientes (como el caso de los hospitales).

Lema. *Si* $X \sim U(0,1)$ *entonces* $-2\log(X) \sim \chi^2_2$

Para los tests que hemos visto se cumple que:

- Si $H_{0,i}$ es cierta, $p_i \sim U(0,1)$.

Lema. *Si* $X \sim U(0,1)$ *entonces* $-2\log(X) \sim \chi^2_2$

Para los tests que hemos visto se cumple que:

- Si $H_{0,i}$ es cierta, $p_i \sim U(0,1)$.

Fisher sugiere combinar los $p-$valores usando el Lema y propone el estadístico

$$V = -2\sum_{i=1}^{k} \log(p_i)$$

Bajo H_0, $V \sim \chi^2_{2k}$, luego rechazamos H_0 si

$$V > \chi^2_{2k,\alpha}$$

donde $\mathbb{P}(\chi^2_{2k} > \chi^2_{2k,\alpha}) = \alpha$

Supongamos que hemos realizado k tests independientes para testear

$$H_{0,i} \quad \text{versus} \quad H_{1,i} \quad 1 \leq i \leq k$$

obteniendo, respectivamente, los $p-$valores

$$p_1, \ldots, p_k$$

- La hipótesis global o hipótesis nula del meta–analisis es que todas las hipótesis nulas son ciertas $(H_0 = \cap_{i=1}^{k} H_{0,i})$
- La hipótesis alternativa es que al menos una de las alternativas $H_{0,i}$ es cierta.

El método de Bonferroni es un *procedimiento de test múltiples* y propone rechazar H_0 si rechazamos alguna $H_{0,i}$

Método de Bonferroni

Para asegurar nivel menor o igual a α, el método de Bonferroni rechaza $H_{0,i}$ si

$$p_i \leq \frac{\alpha}{k}$$

Método de Bonferroni

Para asegurar nivel menor o igual a α, el método de Bonferroni rechaza $H_{0,i}$ si

$$p_i \leq \frac{\alpha}{k}$$

Teorema. *El método de Bonferroni produce un test para H_0 de nivel menor o igual a α.*

Método de Bonferroni

Para asegurar nivel menor o igual a α, el método de Bonferroni rechaza $H_{0,i}$ si

$$p_i \leq \frac{\alpha}{k}$$

Teorema. *El método de Bonferroni produce un test para H_0 de nivel menor o igual a α.*

En el ejemplo de los genes, si $\alpha = 0.05$, como $k = 2638$, $\alpha/k = 0.00001895375$. Luego, necesito mucha evidencia para rechazar.

El método de Bonferroni es muy conservador.

El método de Bonferroni no usa que los tests sean independientes.

False discovery rate

Supongamos que rechazamos todas las hipótesis nulas cuyos $p-$valores están por debajo de un valor dado C (por ejemplo, si usamos el método de Bonferroni $C = \alpha/k$)

False discovery rate

Supongamos que rechazamos todas las hipótesis nulas cuyos $p-$valores están por debajo de un valor dado C (por ejemplo, si usamos el método de Bonferroni $C = \alpha/k$) y consideremos la tabla

	No rechazo H_0	Rechazo H_0	Total
H_0 Cierta	U	V	k_0
H_0 Falsa	T	S	$k - k_0$
Total	$k - R$	R	k

False discovery rate

Supongamos que rechazamos todas las hipótesis nulas cuyos $p-$valores están por debajo de un valor dado C (por ejemplo, si usamos el método de Bonferroni $C = \alpha/k$) y consideremos la tabla

	No rechazo H_0	Rechazo H_0	Total
H_0 Cierta	U	V	k_0
H_0 Falsa	T	S	$k - k_0$
Total	$k - R$	R	k

- k_0 es el número de hipótesis nulas ciertas entre $H_{0,1}, \dots, H_{0,k}$, es fijo, pero desconocido
- V es el número de falsos positivos (o *false discoveries*)
- S es el número de verdaderos positivos (o *true discoveries*)
- R es el número de hipótesis nulas rechazadas (o *discoveries*)

False discovery rate

Cuando testeamos k hipótesis de las cuales k_0 son ciertas, R es una variable aleatoria observable, mientras que S, T, U y V no son observables.

False discovery rate

Cuando testeamos k hipótesis de las cuales k_0 son ciertas, R es una variable aleatoria observable, mientras que S, T, U y V no son observables.

Definamos la proporción de falsos positivos *False Discovery Proportion (FDP)* como

$$FDP = \begin{cases} \dfrac{V}{R} & \text{si} \quad R > 0 \\ 0 & \text{si} \quad R = 0 \end{cases}$$

FDP es la proporción de rechazos incorrectos.

False discovery rate

- El false discovery rate (FDR) se define como el valor esperado del cociente entre número de falsos rechazos V y el número de rechazos R.

False discovery rate

- El false discovery rate (FDR) se define como el valor esperado del cociente entre número de falsos rechazos V y el número de rechazos R.

- El false discovery rate (FDR) es igual a

$$FDR = \mathbb{E}(FDP)$$

False discovery rate

- El false discovery rate (FDR) se define como el valor esperado del cociente entre número de falsos rechazos V y el número de rechazos R.

- El false discovery rate (FDR) es igual a

$$FDR = \mathbb{E}(FDP)$$

- El FDR es la proporción esperada de rechazos incorrectos a lo largo de los k tests.

- Si $H_{0,i} = H_{0,1}$ para todo i, entonces $k_0 = k$). Por lo tanto, $FDR = \mathbb{P}(V \geq 1)$ y el FDR coincide con el nivel del procedimiento.

El método de Benjamini-Hochberg

1. Sean $p^{(1)}, \ldots, p^{(k)}$ los $p-$valores ordenados de menor a mayor. Indiquemos por $H_0^{(i)}$ la hipótesis nula asociada a $p^{(i)}$
2. Definamos
$$\ell_i = \frac{i\,\alpha}{k}$$
3. Si para todo i, $p^{(i)} > \ell_i$, ninguna $H_{0,i}$ se rechaza.
4. Si existe i tal que $p^{(i)} \leq \ell_i$. Sea
$$r = \max\{i : p^{(i)} < \ell_i\}$$
5. Sea $C = p^{(r)}$, C se llama el umbral de rechazo BH
6. Rechace todas las hipótesis nulas $H_{0,i}$ para las cuales $p_i \leq C$. Esto es equivalente a rechazar todas las $H_0^{(i)}$ para $i = 1, \ldots, r$.

El método de Benjamini-Hochberg para tests independientes

Teorema. *Si se aplica el procedimiento de BH, entonces, independientemente de cuantas hipótesis nulas sean ciertas y de cual sea la distribución de los p−valores cuando la hipótesis nula es falsa,*

$$FDR = \mathbb{E}(FDP) \leq \frac{k_0}{k}\alpha \leq \alpha$$

El método de Benjamini-Hochberg para tests independientes

Teorema. *Si se aplica el procedimiento de BH, entonces, independientemente de cuantas hipótesis nulas sean ciertas y de cual sea la distribución de los $p-$valores cuando la hipótesis nula es falsa,*

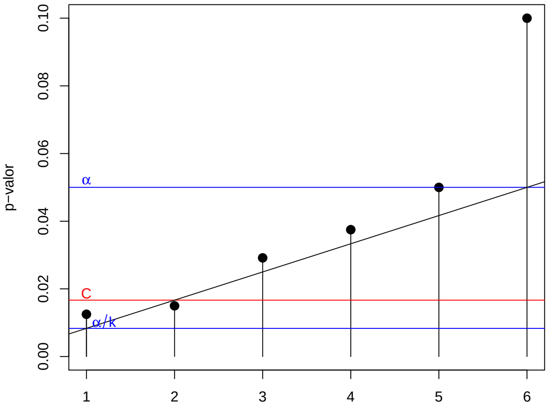
$$FDR = \mathbb{E}(FDP) \leq \frac{k_0}{k}\alpha \leq \alpha$$

Si $H_{0,i} = H_{0,1}$ **para todo** i, el $p-$valor del procedimiento de BH es

$$p_0 = k \min_{1 \leq i \leq k} \frac{p^{(i)}}{i}$$

Ejemplo

$$p^{(1)} = 0.0125, \quad p^{(2)} = 0.015, \quad p^{(3)} = 0.0292,$$
$$p^{(4)} = 0.0375, \quad p^{(5)} = 0.050, \quad p^{(6)} = 0.10$$



La recta oblicua es la recta de pendiente α/k

Ejemplo

- Si el test global de Fisher $V = -2\sum_{i=1}^{k}\log(p_i) = 41.39$ y $\chi^2_{2k,\alpha} = 21.03$. Luego rechazo $H_0 = \cap_{i=1}^{k} H_{0,i}$, o sea, **alguna hipótesis nula no es cierta**.

Ejemplo

- Si el test global de Fisher $V = -2\sum_{i=1}^{k}\log(p_i) = 41.39$ y $\chi^2_{2k,\alpha} = 21.03$. Luego rechazo $H_0 = \cap_{i=1}^{k} H_{0,i}$, o sea, **alguna hipótesis nula no es cierta**.

- Ahora hacemos test múltiples
  - Si testeamos separadamente con nivel $\alpha = 0.05$ rechazo si $p_i < \alpha$. En este caso 4 hipótesis se rechazan.

Ejemplo

- Si el test global de Fisher $V = -2\sum_{i=1}^{k}\log(p_i) = 41.39$ y $\chi^2_{2k,\alpha} = 21.03$. Luego rechazo $H_0 = \cap_{i=1}^{k} H_{0,i}$, o sea, **alguna hipótesis nula no es cierta**.

- Ahora hacemos test múltiples
  - Si testeamos separadamente con nivel $\alpha = 0.05$ rechazo si $p_i < \alpha$. En este caso 4 hipótesis se rechazan.
  - Si usamos el método de Bonferroni rechazo si $p_i < \alpha/k$. Ninguna hipótesis se rechaza.

Ejemplo

- Si el test global de Fisher $V = -2\sum_{i=1}^{k} \log(p_i) = 41.39$ y $\chi^2_{2k,\alpha} = 21.03$. Luego rechazo $H_0 = \cap_{i=1}^{k} H_{0,i}$, o sea, **alguna hipótesis nula no es cierta**.

- Ahora hacemos test múltiples
  - Si testeamos separadamente con nivel $\alpha = 0.05$ rechazo si $p_i < \alpha$. En este caso 4 hipótesis se rechazan.
  - Si usamos el método de Bonferroni rechazo si $p_i < \alpha/k$. Ninguna hipótesis se rechaza.
  - Si usamos el método de Benjamini-Hochberg, rechazo si $p_i \leq C$. En este caso, dos hipótesis se rechazan.