Tests para igualdad de varianzas

En los problemas de comparaciones de dos o más tratamientos que hemos tratado hemos supuesto que, si existe una diferencia entre tratamientos, ésta se traduce en un corrimiento. Obviamente hay otras alternativas posibles, como por ejemplo que la diferencia entre tratamientos se traduzca en un cambio en la varianza. El primer test que veremos trabaja con los rangos tal como los definimos en el problema de posición.

Test de Mood para igualdad de varianzas: Sean $X_1, \dots, X_n$ e $Y_1, \dots, Y_m$ muestras aleatorias de distribuciones $F((x-\mu)/\sigma)$ y $F((x-\mu)/\tau)$ respectivamente, con μ común y conocida. Supongamos que la hipótesis nula de interés es $H_0: \sigma = \tau$.

Sea $R(Y_i)$ el rango de Y_i en la muestra agrupada de tamaño $N = n+m$, y definamos

$$M = \sum_{i=1}^m \left(R(Y_i) - \frac{N+1}{2} \right)^2$$

Valores grandes de M indicarían que $\tau > \sigma$ y valores pequeños de M indicarían que $\tau < \sigma$. Por lo tanto, se pueden construir tests para las diferentes alternativas basados en M .

Tablas exactas de la distribución de M se hallan en Laubscher, Steffens y DeLange (1968). Si n y m son suficientemente grandes se puede usar la aproximación normal, basada en

$$E(M) = m \frac{N^2 - 1}{12}$$

$$Var(M) = nm \frac{(N+1)(N^2 - 4)}{180}$$

Hipótesis a testear y región de rechazo:

A. $H_0: \sigma = \tau$ vs $H_1: \sigma > \tau$

Rechazamos H_0 si $M < m \frac{N^2 - 1}{12} - z_\alpha \sqrt{Var(M)}$

B. $H_0: \sigma = \tau$ vs $H_1: \sigma < \tau$

Rechazamos H_0 si $M > m \frac{N^2 - 1}{12} + z_\alpha \sqrt{Var(M)}$

C. $H_0: \sigma = \tau$ vs $H_1: \sigma \neq \tau$

Rechazamos H_0 si $\left| M - m \frac{N^2 - 1}{12} \right| > z_{\alpha/2} \sqrt{Var(M)}$

Notar que se supone que los parámetros de posición de ambas muestras son conocidos e idénticos. Si no se puede suponer que son idénticos, debe realizarse previamente algún tipo de ajuste. Si los valores que se utilizan en ese ajuste son estimaciones alejadas de los verdaderos valores de los parámetros, es posible que el rechazo de H_0 no implique necesariamente que las varianzas son distintas.

Eficiencia: La eficiencia el test de Mood fue estudiada por Hollander (1863). Respecto del test F y en el caso Normal es 0.76.

Ejemplo: Se desea testear si el orden en que las preguntas son formuladas influye en los resultados en un examen multiple choice. 21 alumnos que tomaron un curso de Inferencia Estadística fueron divididos en dos grupos aleatoriamente y se les tomó un examen que constaba de 20 preguntas. A 11 alumnos (Grupo 1) las preguntas se les presentaron en el orden en que se habían desarrollado los temas en clase y al resto (Grupo 2) en un orden aleatorio. Los resultados (número de respuestas correctas) se presentan en la siguiente tabla, juntamente con los rangos:

Grupo 1 (X)	Grupo 2 (Y)	Rango X	Rango Y
	3		1
	6		2
7		3	
	8		4
9		6	
9		6	
	9		6
10		8	
11		10.5	
11		10.5	
	11		10.5
	11		10.5
12		13	
	13		15
	13		15
	13		15
14		17	
15		18.5	
	15		18.5
17		20	
19		21	

Supongamos que se puede suponer que ambas poblaciones tienen el mismo centro y que se desea testear $H_0: \sigma = \tau$ vs $H_1: \sigma \neq \tau$

$E(M) = 366.67$ y $Var(M) = 5875.22$. El valor del estadístico $M = 359.75$ y por lo tanto el p-valor aproximado será

$$2 \left[\phi \left(\frac{359.75 - 366.67}{\sqrt{76.65}} \right) \right] = 0.93$$

y no se rechaza H_0 .

Test de rangos al cuadrado (Conover): Sean $X_1, \dots, X_n$ e $Y_1, \dots, Y_m$ muestras aleatorias de distribuciones $F((x-\mu_1)/\sigma)$ y $F((x-\mu_2)/\tau)$ respectivamente. Supongamos que la hipótesis nula de interés es

$$H_0: \sigma = \tau$$

Transformamos los datos a desviaciones absolutas respecto de su media, es decir, definimos

$$U_i = |X_i - \mu_1|, \quad i = 1, \dots, n \quad \text{y} \quad V_j = |Y_j - \mu_2|, \quad j = 1, \dots, m$$

Sean $R(U_i)$ y $R(V_j)$ los rangos de U_i y V_j respectivamente en la muestra agrupada de tamaño $N=n+m$. Si no hay empates, el estadístico del test será

$$T = \sum_{i=1}^n [R(U_i)]^2$$

que, bajo H_0 , satisface

$$E(T) = n \frac{1}{N} \sum_{i=1}^n i^2 = \frac{n}{N} \frac{N(N+1)(2N+1)}{6} = \frac{n(N+1)(2N+1)}{6}$$

$$Var(T) = \frac{nm}{(N-1)N} \left\{ \sum_{i=1}^N i^4 - N \frac{(N+1)^2(2N+1)^2}{36} \right\} = \frac{nm(N+1)(2N+1)(8N+11)}{180}$$

Si hay empates, se utiliza el estadístico

$$T_1 = \frac{T - n \bar{R}_2}{\left[\frac{nm}{N(N-1)} \sum_{i=1}^N R_i^4 - \frac{nm}{N-1} (\bar{R}_2)^2 \right]^{1/2}}$$

donde $\bar{R}_2 = \frac{1}{N} \left\{ \sum_{i=1}^n [R(U_i)]^2 + \sum_{j=1}^m [R(V_j)]^2 \right\}$ es el promedio de los rangos al cuadrado de ambas muestras combinadas y $\sum_{i=1}^N R_i^4 = \sum_{i=1}^n [R(U_i)]^4 + \sum_{j=1}^m [R(V_j)]^4$.

Los cuantiles de la distribución exacta de T se encuentran tabulados (Tabla A9 de Conover) para $n \leq 10$, $m \leq 10$ y en el caso en que no hay empates. Para tamaños mayores de muestra o si hay empates se utiliza la aproximación Normal.

Supongamos que deseamos testear las hipótesis:

$$\mathbf{A.} \quad H_0: \sigma = \tau \quad \text{vs} \quad H_1: \sigma > \tau$$

Si n y m son menores o iguales que 10 y no hay empates, se rechaza H_0 si $T > w_{1-\alpha}$, siendo $w_{1-\alpha}$ el cuantil $1-\alpha$ de la distribución exacta.

Si n y m son grandes y no hay empates, se rechaza H_0 si $\frac{T - E(T)}{\sqrt{Var(T)}} > z_\alpha$.

Si hay empates, se rechaza H_0 si

$$T_1 = \frac{T - n \bar{R}_2}{\left[\frac{nm}{N(N-1)} \sum_{i=1}^N R_i^4 - \frac{nm}{N-1} (\bar{R}_2)^2 \right]^{1/2}} > z_\alpha$$

De manera análoga se aplica el test a las otras dos alternativas.

Eficiencia: La eficiencia de este test respecto al test F es similar a la del test de Mood. Es $15/(2\pi^2)=0.76$ en el caso Normal, 1 en el caso uniforme y 1.08 en el caso doble exponencial.

En caso de que las medias sean desconocidas, deben reemplazarse por estimaciones. En este caso la distribución exacta deja de ser válida y debe usarse la aproximación normal.

Extensión a más de dos muestras: Este test se puede extender fácilmente a más de dos muestras. Sean

$$X_{ij} \sim F\left(\frac{x - \mu_i}{\sigma_i}\right) \quad 1 \leq j \leq n_i, 1 \leq i \leq k$$

Si μ_i es conocido para todo i , se considera $U_{ij} = |X_{ij} - \mu_i|$. Si no se conoce, se estima y se considera $U_{ij} = |X_{ij} - \hat{\mu}_i|$.

Una vez construidos los U_{ij} se procede como en el test de Kruskal-Wallis, definiendo

$$R_{ij} = R(U_{ij}) = \text{rango de } U_{ij} \text{ en la muestra total ordenada de tamaño } N = \sum_{i=1}^k n_i$$

Sea $S_i = \sum_{j=1}^{n_i} (R_{ij})^2$ (en el caso $k=2$, $S_1 = T$)

Nos interesa testear

$$H_0: \sigma_1 = \sigma_2 = \dots = \sigma_k \quad \text{vs} \quad H_1: \text{existe al menos un par } i \neq j \text{ tal que } \sigma_i \neq \sigma_j$$

Como en el test de Kruskal-Wallis, se construye un estadístico que, bajo H_0 tiene distribución asintótica χ^2_{k-1} .

Si no hay empates, el estadístico es

$$T = \frac{180}{N(N+1)(2N+1)(8N+11)} \left[\sum_{j=1}^k \frac{S_j^2}{n_j} - \frac{N(N+1)^2(2N+1)^2}{6} \right]$$

Si hay empates, se debe corregir la esperanza y la varianza y se utiliza

$$T = \frac{1}{D^2} \sum_{i=1}^k \left(\frac{S_i^2}{n_i} - N \bar{S}^2 \right)$$

donde $\bar{S} = \frac{1}{N} \sum_{i=1}^k S_i$ y $D^2 = \frac{1}{N-1} \left(\sum_{i=1}^k \sum_{j=1}^{n_i} R_{ij}^4 - N \bar{S}^2 \right)$.

Se rechaza H_0 a nivel α si $T > \chi_{k-1, \alpha}^2$.

Comparaciones múltiples: Diremos que σ_i es significativamente distinta de σ_j si

$$\left| \frac{S_i}{n_i} - \frac{S_j}{n_j} \right| > t_{N-k, \alpha/2} \left(D^2 \frac{N-1-T}{N-k} \right)^{1/2} \left(\frac{1}{n_i} + \frac{1}{n_j} \right)^{1/2}$$

donde $t_{N-k, \alpha/2}$ es el valor crítico $\alpha/2$ de la distribución t con N-k grados de libertad.

Test de Siegel-Tukey: Sean $X_1, \dots, X_n$ e $Y_1, \dots, Y_m$ muestras aleatorias de distribuciones $F((x-\mu)/\sigma)$ y $F((x-\mu)/\tau)$ respectivamente. Supongamos que las hipótesis de interés son

$$H_0: \sigma = \tau \quad \text{vs} \quad H_1: \sigma > \tau$$

Se asignan rangos de la siguiente manera:

- rango 1 a la menor observación en la muestra agrupada
- rango 2 a la mayor observación
- rango 3 a la segunda mayor observación
- rango 4 a la segunda menor observación y así sucesivamente.

Si la muestra X tiene mayor dispersión que la muestra Y, los rangos correspondientes a la muestra Y deberían ser mayores que los de la muestra X. Llamemos $S_1, S_2, \dots, S_n$ a los rangos de la muestra Y en la muestra agrupada, el estadístico del test es

$$T_v = \sum_{i=1}^m S_i$$

y se rechazará H_0 a favor de H_1 si $T_V > c$. ¿Cómo se halla el valor c ?

Se puede probar que T_V tiene la misma distribución que el estadístico del test de Mann-Whitney, y por lo tanto si no hay empates y los tamaños de muestra son menores o iguales que 20, se puede utilizar la Tabla A7 de Conover. Si n y m son grandes se rechaza H_0 si

$$T_3 > m(N+1) + z_\alpha \sqrt{\frac{mn(N+1)}{12}}$$

Este test se basa en un supuesto importante: las medias de ambas muestras son iguales. Si ésto no ocurriese, podrían reemplazarse las medias por valores estimados y trabajar con $X_i - \hat{\mu}_X$ y $Y_j - \hat{\mu}_Y$. En este caso, debe usarse la distribución aproximada.

Eficiencia: Este test tiene menor eficiencia que los tests anteriores. Su eficiencia respecto al test F bajo normalidad es $6/\pi^2 = 0.608$ y bajo la distribución doble exponencial es 0.94. Otro problema adicional de este test es su falta de simetría, es decir que si se invierte el orden de ranqueo, el p-valor obtenido no será el mismo.

Test de Klotz: Es un test basado en scores normales, con alta eficiencia si la distribución es simétrica. Bajo normalidad su eficiencia respecto al test F es 1. Sean $X_1, \dots, X_n$ e $Y_1, \dots, Y_m$ muestras aleatorias de distribuciones $F((x-\mu)/\sigma)$ y $F((x-\mu)/\tau)$.

El estadístico del test de Klotz es

$$K = \sum_{i=1}^m \left[\Phi^{-1} \left(\frac{R(Y_i)}{N+1} \right) \right]^2$$

Klotz (1962) calculó la distribución exacta, pero si los tamaños de muestra son grandes puede utilizarse la aproximación Normal y

$$E(K) = \frac{m}{N} \sum_{i=1}^N \left[\Phi^{-1} \left(\frac{i}{N+1} \right) \right]^2$$

$$Var(K) = \frac{nm}{N(N-1)} \sum_{i=1}^N \left[\Phi^{-1} \left(\frac{i}{N+1} \right) \right]^4 - \frac{[E(K)]^2 n}{m(N-1)}$$