LA DESCOMPOSICIÓN ORTOGONAL DEL TEST χ^2 DE PEARSON APLICADO AL TEST DE FRIEDMAN

Dindo Valdez

Introducción

Anderson[1] (1959) analizó las tablas de contingencia de doble entrada con los totales marginales constantes desarrollando la descomposición del test χ^2 de Pearson en componentes ortogonales y Rayner[2](1989) aplicó estas tablas y la prueba χ^2 a las pruebas de degustación de sabor de distintas variedades de un determinado producto. La descomposición del test χ^2 se puede aplicar como una prueba de homogeneidad entre filas (variedades), dicha descomposición ortogonal de este estadístico hace, esta prueba de hipótesis, más informativa y exhaustiva en este tipo de tablas donde su primer componente está relacionado con el estadístico Q_F del test de Friedman, y los subsecuentes componentes son extensiones a esta prueba.

Descomposición del test Chi cuadrado de Pearson

Primero, consideremos una tabla de contingencia de doble entrada de observaciones N_{ij} $i = 1,2,\ldots,r$; $j = 1,2,\ldots,c$, donde los totales de las filas $n_{i\bullet}$ y los totales de las columnas $n_{\bullet j}$ son valores conocidos (constantes), la clasificación de las columnas está ordenada y no así el de las filas. Para probar la hipótesis nula de homogeneidad entre las filas (tratamientos) es posible aplicar el conocido test χ^2 de Pearson, definido como:

$$\chi^2 = \sum_i \sum_j \frac{\left(N_{ij} - E_{ij}\right)^2}{E_{ij}}\ ,\ {}_{ij} = n_{i\bullet} n_{\bullet j} / n \text{ , siendo } n \text{ el total de datos} \qquad (1)$$

[1] Ver [1]

[2] Ver [2]

Sin embargo, es posible hacer mas potente a esta prueba descomponiéndola en c-1 componentes ortogonales para analizar efectos de escala, asimetría, curtosis, y momentos de mayor orden entre las filas de la tabla. Para definir tales componentes primero se otorga una puntuación a las categorías de las columnas (1,2,K , c), luego se definen polinomios ortogonales denotados por $h_r(j)$, lo cuales son ortogonales respecto de p_j (la probabilidad marginal de la columna j).

Tal que: $\sum h_r(j) \cdot h_s(j) \cdot p_j = \begin{cases} 0, & si \quad r = s \\ 1, & si \quad r \neq s \end{cases}$

Emerson[3] (1968) define estos polinomios $h_r(j)$ como:

$$h_r(j) = (jA_r + B_r)h_{r-1}(j) - C_r h_{r-2}(j), \quad j, r = 1,2, K\ c < n \tag{2}$$

donde A_r, B_r, C_r son constantes para un r dato tal que $h_r(j)$ llega a ser un polinomio de grado j, y los polinomios iniciales son: $h_{-1}(j)=0$ y $h_0(j)=1$. Tomando las probabilidades estimadas $p_j=n_{.j}/n$, $j=1,2,$K $,c$, proporcionales a las columnas marginales, se define al s-avo componente de la partición como:

$$V_{si} = \frac{1}{\sqrt{n_{i.}}} \sum_{j=1}^{c} N_{ij} \cdot h_s(j), \quad \text{para } s=1,K,c-1 \ e \ i=1,K,r \tag{3}$$

De esta manera el estadístico de Pearson resulta particionado en c-1 componentes[4],

$$\chi^2 = \sum_{s=1}^{c-1} \sum_{i=1}^{r} V_{si}^2 = \sum_{s=1}^{c-1} Q_s$$

Los componentes V_{si} satisfacen la condición lineal,

$$V_{s1}\sqrt{n_{1.}} + V_{s2}\sqrt{n_{2.}} + L + V_{sr}\sqrt{n_{r.}} = 0\,, s = 1, K, c-1$$

[3] P.L. Emerson ha realizado varios trabajos comparando su método con el de Gram Schimdt, llegando a la conclusión que su método es más rápido y preciso.

[4] Detalles de tal partición están dados en [3]

Bajo la hipótesis nula de homogeneidad entra las filas, cada componente V_{si} tiene distribución asintótica normal estándar, en consecuencia los componentes $Q_s = \sum V_{si}^2$ de La partición tienen distribución asintótica χ_{r-1}^2 . Tales componentes son llamados estadísticos score, y tiene similares propiedades a los estadísticos obtenidos por el método de la razón de verosimilitud[5].

Aplicación de la descomposición al test de Friedman

Supongamos que r (>2) tratamientos son rankeados (clasificados) cada uno por n personas (jueces), entonces es posible realizar un cuadro de contingencia de doble entrada con r filas (tratamientos)y r columnas (rangos) como se muestra en la tabla siguiente:

Tratamiento	**Rango**				$n_{i.}$
	1	**2**	**N**	**r**	
T1	N_{11}	N_{12}		N_{1r}	n
T2	N_{21}	N_{22}		N_{2r}	n
⋮	⋮	⋮		⋮	⋮
Tr	N_{r1}	N_{r2}		N_{rr}	n
n.j	n	n		n	n x r

N_{ij} es el número de veces que el tratamiento i recibe el rango j

r es el número de muestras aleatorias

n es el tamaño de las muestras

Lo anterior queda claro ya que cada tratamiento y cada rango es asignado *n* veces. Aplicando la definición de la ecuación (1), el estadístico χ^2 de Pearson para este tipo de tablas resulta:

$$\chi^2 = \frac{r}{n}\sum_{i=1}^{r}\sum_{j=1}^{r}\left(N_{ij} - \frac{n}{r}\right)^2 = \sum_{s=1}^{r-1}\sum_{j=1}^{r} V_{si}^2 = \sum_{s=1}^{r-1} Q_s$$

[5] Ver [2] , sección 3.4

Con $V_{si} = \frac{\sqrt{r-1}}{r}\left(\sum_{j=1}^{r} h_s(j)\frac{N_{ij}-n/r}{\sqrt{n/r}}\right), \quad s = 1, K, r-1;\ i = 1, K, r.$

El primer componente de esta descomposición resulta ser el estadístico del test de Friedman, tal que:

$$Q_1 = \sum_{j=1}^{r} V_{1i}^2 = \frac{12}{rn(r+1)}\sum_{j=1}^{r}\left(R_i - \frac{n(r+1)}{2}\right)^2 = Q_F$$

Donde R_i es la suma de los rangos para el i-ésimo tratamiento $R_i = \sum_{j=1}^{r} N_{ij}$.

Ya que los componentes V_{si} envuelven momentos de orden s, los componentes Q_s detectan los efectos del s-avo momento bajo la hipótesis nula y por esta razón, se puede considerar al test de Friedman como un test de posición ya que envuelve polinomios de primer grado en su cálculo, de esta manera los subsecuentes componentes Q_s para $s=2$,K $r-1$, se consideran las extensiones al test de Friedman, puesto que éstos miden adicionalmente los efectos en la escala, asimetría, curtosis, y momentos de mayor orden respectivamente. Para mostrar la utilidad de tal descomposición se realizó un estudio en un grupo de personas universitarias entre los 18 y 25 años de edad con el objeto de medir la preferencia sobre cinco sabores de yogurt bebible (coco, frutilla, uva, limón y chicle) donde cada persona clasificó a cada sabor del 1º al 5º en el orden de preferencia llegando a realizar 95 observaciones (secuencias de preferencia) de manera que cada sabor recibió un total de 95 valoraciones (entre 1s, 2s, etc.). Es así que con el conjunto de observaciones se elaboró una tabla de contingencia, donde las filas corresponden a los sabores (tratamientos) y las columnas a los rangos (bloques).

Tratamiento (sabor)	**Escalas de Preferencia**					**Total**
	1	**2**	**3**	**4**	**5**	
Coco	12	18	20	19	26	95
Frutilla	8	17	24	22	24	95
Uva	13	16	18	21	27	95
Limón	15	31	21	16	13	95
Chicle	47	13	12	17	6	95
Total	95	95	95	95	95	475
La escala 1 es la menos gustada y 5 es la mayor preferencia						

La hipótesis a docimar es que cada variedad de yogurt tiene el mismo nivel de aceptación (homogeneidad), al aplicar la prueba χ^2 de homogeneidad entre filas se obtuvo un valor de 87,47 con 16 grados de libertad y una significancia de 7.2E-12; por tanto podemos concluir que la distribución en las preferencias de los 5 sabores de yogurt no es homogénea al 5% de significación. En cambio si utilizamos los datos brutos y aplicamos el test de Friedman para probar la hipótesis que los 5 sabores tienen similares resultados se obtiene Q_F = 41.364 con una significancia igual a 0. Consecuentemente rechazamos la hipótesis nula y concluimos que los cinco sabores no produjeron idénticos resultados respecto a sus rangos asignados.

Puesto que existen diferencias entre los tratamientos, es importante analizar la influencia de los componentes ortogonales del test χ^2 para aclarar dónde radican estas diferencias exactamente, para lo cual calculamos los cuatro componentes. Aplicando la ecuación (2) los polinomios ortogonales de Emerson resultan:

J	1	2	3	4	5
$h_1(j)$	1,00	1,00	1,00	1,00	1,00
$h_2(j)$	-1,41	-0,70	0,00	0,70	1,41
$h_3(j)$	1,19	-0,59	-1,19	-0,59	1,19
$h_4(j)$	-0,70	1,41	0,00	-1,41	0,70
$h_5(j)$	0,26	-1,06	1,60	-1,06	0,26

Tales polinomios son ortonormales bajo la Ho con $p_j = 0.2$ y el valor esperado de la variable N_{ij} es 19 en todos los casos.

Efectos de posición y escala. Basados en los anteriores resultados, la media o efecto lineal para el i-ésimo sabor está relacionado con los elementos V_{1i} quedando definida como:

$$V_{1i} = \frac{2}{5}\sum_{j=1}^{5} h_1(j)\frac{N_{ij} - 19}{\sqrt{19}}, \qquad i = 1, \ldots\ldots, 5$$

De igual manera la variación llamada también efecto de dispersión o escala queda definido por:

$$V_{2i} = \frac{2}{5}\sum_{j=1}^{5} h_2(j)\frac{N_{ij} - 19}{\sqrt{19}}, \qquad i = 1, \ldots\ldots, 5$$

Los subsecuentes componentes son calculados aplicando la ecuación (3) y se presentan en la tabla:

Variedad	Efecto			
	Lineal	Cuadrático	Asimetría	Curtosis
	V_{1i}	V_{2i}	V_{3i}	V_{4i}
Coco	1,88	-0,05	0,77	0,24
Frutilla	2,40	-1,26	0,38	0,49
Uva	2,14	0,38	0,25	0,00
Limón	-1,36	-1,91	1,75	-0,85
Chicle	-5,06	2,85	-3,17	0,12

Fuente	Grados de Libertad	Suma de Cuadrados	Significancia
Posición Q_1	4,00	41,36	0,00
Escala Q_2	4,00	13,55	0,01
Residuo $Q_3 + Q_4$	8,00	15,05	0,06
Total χ^2	16,00	69,96	

Esta tabla tiene un estilo parecido al clásico análisis de varianza, basado en los componentes ortogonales del test χ^2 de Pearson. La tabla muestra que la suma de cuadrados de los efectos de ubicación es el estadístico Q_F del test de Friedman para datos ordenados por rangos. El componente Q_2 se considera como un nuevo estadístico, el cuál mide los efectos de escala. Esta tabla confirma una vez más que existen diferencias en cuanto a la posición entre las distintas variables de yogurt también muestra algo adicional y revelador, la existencia de diferencias significativas en la escala, tal diferencia no podría haber sido revelada si no se utilizaba la descomposición del test

Chi cuadrado. El residuo mide los efectos de asimetría y curtosis simultáneamente y éste no es significativo al 5%.

En estos estudios de comparación, tales datos usualmente son analizados mediante el tradicional análisis de varianza, pero de una manera distinta al análisis de rangos, ya que esta prueba puede verse afectada por valores atípicos que llevarían a tener una heterogeneidad en la varianza y por reasumir la normalidad de los datos. El análisis realizado es más robusto que el ANOVA, en el sentido que ninguno de los dos supuestos: homogeneidad en la varianza y normalidad de los datos, han sido asumidos. También este estudio ha demostrado ser más exhaustivo e informativo que la simple aplicación del test de Friedman.

Bibliografía

[1] P. Anderson, Revista Electrónica Biometrics.

[2] J.C. Rayner, Smooth Test of Goodness of Fit.

[3] Dindo Valdez, Pruebas Suavizadas de Bondad de Ajuste con Aplicación en Pruebas no Paramétricas.

----- O -----

Si amas algo

Según un Estadístico: “Si amas a alguien, déjalo ir; si el te quiere, las probabilidades que vuelva son de un 86.5 por ciento; si no te quiere, tus relaciones con él caen en el campo de lo improbable, con un margen de error de un 3 por ciento”.

Índice