Modelo Lineal Generalizado

Test Exacto de Fisher

Cuando las muestras son pequeñas sabemos que el test de χ^2 y G^2 no son bien aproximados por la distribución χ^2 y en consecuencia las conslusiones a las que llegamos a partir de los p–valores calculados no son confiables.

En ese sentido, el Tests Exacto de Fisher es una solución a este problema en el caso de tablas de 2×2.

Consideremos el siguiente ejemplo: 13 individuos fueron operados de la rodilla. Los pacientes fueron clasificados según la dolencia en rodilla girada o rodilla directa y según el resultado de la operación en muy bueno o aceptable. La siguiente tabla muestra los resultados obtenidos.

	Resultado		n_{i+}
Rodilla	Muy Bueno	Aceptable	
Directa	3	2	5
Girada	7	1	8
n_{+j}	10	3	13

Table 1: Datos de Operación de Rodilla

Para estos datos, tenemos que el *valor observado del odds ratio* es

$$\theta = \frac{3 \times 1}{2 \times 7} = 0.2143$$

Si conociéramos los valores marginales, es claro que el valor de la primera casilla (podría ser cualquiera de ellas) determina los valores de los otros 3 casilleros:

	Resultado		n_{i+}
Rodilla	Muy Bueno	Aceptable	
Directa	3		5
Girada			8
n_{+j}	10	3	13

Table 2: Datos de Operación de Rodilla

Si en realidad $\theta = 1$, tendríamos que la probabilidad de observar un valor n_{11} en la casilla $(1,1)$ estará dada por la distribución multinomial:

$$P(n_{11}) = \frac{\binom{n_{1+}}{n_{11}}\binom{n_{2+}}{n_{+1} - n_{11}}}{\binom{n_{++}}{n_{+1}}}$$

En nuestro caso particular, si $\theta = 1$ tendríamos que la probabilidad de observar $n_{11} = 3$ sería

$$\frac{\binom{5}{3}\binom{8}{7}}{\binom{13}{10}} = 0.27972$$

Si quisiéramos realizar un test para las hipótesis:

$$H_o : \theta = 1 \quad vs. \quad H_1 : \theta < 1$$

deberíamos computar las probabilidades de todas las tablas que tiene θ menor que el observado. Recordemos que por la propiedad vista en clase θ es funciíon creciente de n_{11}, por ello las otras tablas favorables a H_1 serán aquellas con n_{11} menor al observado. En nuestro ejemplo hay sólo una posible:

Rodilla	Resultado Muy Bueno	Resultado Aceptable	n_{i+}
Directa	2	3	5
Girada	8	0	8
n_{+j}	10	3	13

Table 3: $\theta = 0$

con probabilidad

$$\frac{\binom{5}{2}\binom{8}{8}}{\binom{13}{10}} = 0.03497$$

por lo tanto el p–valor sería

$$0.27972 + 0.03497 = 0.31469$$

Si quisiéramos realizar un test para las hipótesis:

$$H_o : \theta = 1 \quad vs. \quad H_1 : \theta > 1$$

deberíamos computar las probabilidades de todas las tablas que tiene θ mayor al observado. Con el mismo crterio que antes consideraremos las tablas con n_{11} mayor al observado, que en nuestro caso son

con probabilidad

$$\frac{\binom{5}{4}\binom{8}{6}}{\binom{13}{10}} = 0.489510$$

	Resultado		n_{i+}
Rodilla	Muy Bueno	Aceptable	
Directa	4	1	5
Girada	6	2	8
n_{+j}	10	3	13

Table 4: $\theta = 1.33$

	Resultado		n_{i+}
Rodilla	Muy Bueno	Aceptable	
Directa	5	0	5
Girada	5	3	8
n_{+j}	10	3	13

Table 5: $\theta = \infty$

y con probabilidad

$$\frac{\binom{5}{5}\binom{8}{5}}{\binom{13}{10}} = 0.19580$$

por lo tanto el p–valor sería

$$0.27972 + 0.489510 + 0.19580 = 0.96503$$

Finalmente, si nos interesase testear

$$H_o : \theta = 1 \quad vs. \quad H_1 : \theta \neq 1$$

un criterio posible para calcular el p–valor es el de sumar la probabilidad de todas las tablas cuya probabilidad es **menor o igual** a la observada.

Las Tablas 3 y 5 son las tablas que tienen la propiedad de tener probabilidad menor o igual a la tabla obervada (Tabla 1) con una probabilidad asociada igual a 0.03497 y 0.19580, respectivamente.

Por lo tanto el p–valor para el test bilateral sería:

$$0.27972 + 0.03497 + 0.19580 = 0.51047$$