Daena: International Journal of Good Conscience. 8(3)51-61. Octubre 2013. ISSN 1870-557X

Modelos Estadísticos No Paramétricos en Escala Nominal, Caso de una Muestra

(Nonparametric Statistical Models Nominal Scale, Case of a Sample)

Guillen, A., M.H. Badii, H. [1]Rodríguez & J. L. Abreu*

Resumen. Se discuta de forma breve la noción de métodos estadísticos tanto paramétricos como no paramétricos. Se describe por medio de ejemplos varias pruebas no paramétricas en escala nominal para el caso de una muestra. Se mencionan las virtudes y desventajas de algunas de estas pruebas y presentan pruebas alternativas viables en para estos casos.

Palabras claves. Función de distribución, probabilidad, prueba estadística no paramétrica.

Abstract. Both parametric and nonparametric tests are discussed. Nonparametric tests in nominal sclae for the case of one simple are described with examples, covering their advantages and their disadvantages. Some viable tests as alternative to some of the limited tests are presented and described.

Keywords. Distribution function, nonparametric test, probability.

Introducción

Todas las pruebas estadísticas derivan de la base de algunos supuestos sobre los datos que se van a usar, y la mayoría de las pruebas clásicas de significancia, como la prueba de t de Student, el análisis de varianza, las pruebas de regresión, asumen que los datos se distribuyen según algunas distribuciones clásicas pero mayormente, la distribución normal (Dynkin, 1961; Eberhardt and Flinger. 1977; Freedman et al., 1991; Gleser and Healy, 1976).

Debido a que todas las funciones de las distribuciones clásicas están escritas en expresiones matemáticas las cuales incluyen parámetros (como la media y la desviación estándar), se les denominan funciones de distribución paramétricas, y por tanto, las pruebas que asumen que los datos concuerdan a las funciones de distribución paramétricas se les denominan *pruebas paramétricas*. Ahora bien, debido a que la distribución normal es la distribución estadística más común, el término de la *prueba paramétrica* se utiliza mayormente para indicar y referir a aquella prueba que asuma que los datos tiene una distribución normal (Hettmansperger, and McKean. 199; Huber, 1981; Hohnson, and Kotz. 1972; Johnson et al., 1995; Waller, 1995; Waller et al, 1995).

A veces se puede transformar los datos para hacerlos que tengan una distribución normal y por ende poder usar pruebas paramétricas. Sin embargo, ocurre que los datos son tan anormales que no se puede usar ninguna transformación y tampoco pruebas no paramétricas. Afortunadamente, los estadísticos han desarrollado pruebas especiales que no requieren datos con distribución normal, y a éstas pruebas se les denomina *pruebas no paramétricas*. Como se puede esperar, las pruebas no paramétricas más conocidas y más comúnmente utilizadas son aquellas que corresponden a las pruebas paramétricas más conocidas y más utilizadas.

Una variable es una característica de un elemento de la población. Debido a que existen diferentes tipos de rasgos, también hay distintas escalas para expresar las variables que se definen de forma breve a continuación.

Escala Nominal: en donde se mide los atributos y/o las cualidades del elemento, los cuales pueden ser numéricos o no numéricos. Esta escala proporciona información muy general sobre los elementos y se trata de igualdad y desigualdad de los elementos.

Escala Ordinal: Aquí se ordenen o se ranquean los diferentes niveles (seleccionadas de forma subjetiva en una escala) de los atributos de los elementos los cuales de nuevo pueden ser numérico o no numérico. Esta escala contiene los rasgos de la escala anterior y por ende nos proporciona mayor cantidad de información que la escala nominal. La ordenación tiene sentido en esta escala. Estas dos escalas se utilizan en las investigaciones de tipo cualitativo.

Escala Intervalo: Se trata de medir la magnitud del elemento, es de tipo solamente numérico y tiene dos rasgos siguientes: a) hay distancias constantes entre los valores adyacentes de la variable, y b) tiene un cero no real, es decir convencional.

Escala Razón: Tiene todos los rasgos de la escala intervalo con la excepción de que en la escala razón, sí se utiliza un cero verdadero y por tanto en esta escala la noción de relacionar tiene sentido. Estos dos últimas escalas se usan para investigaciones de tipo cuantitativo.

El propósito de este trabajo es investigar y conocer las diferentes pruebas no paramétricas que se utilizan para el caso de una muestra en escala nominal.

Modelo Binomial

1. **Descripción:** Se utiliza para **1 muestra** en **escala nominal**

2. **Ecuación:**

$$P_x = {}_nC_r\,(p^r\, q^{(n - r)}) \text{ o } P_x = \{n! / [r!\,(n - r)!]\}\,(p^r\, q^{(n - r)})$$

Donde,

Px = Probabilidad del evento esperado
C = # de combinaciones
n = # de eventos
r = # de arreglo esperados del evento
! = Signo de factorial
p = Prevalencia del evento
q = 1 – p

Ejemplo: Un partido político acude a un asesor y plantea su deseo de tener 4 candidatos ganadores: 2 hombres y 2 mujeres, piensan que este es difícil, ya que antecedentes políticos indican inclinación fuerte hacia las mujeres y de hecho 3 de las últimas han sido solamente mujeres.

Así surgen los arreglos posibles:

1. Solo mujeres ($p_x = 4$)
2. 2 y 2 ($p_x = 2$)
3. Solo hombres ($p_x = 0$)

3. **Plantear hipótesis:**

Ho: Para este partido tener varones y mujeres como ganadores es un proceso aleatorio

Ha: Tener **solo varones** ganadores es una probabilidad muy pequeña por los antecedentes.

4. **Nivele de significancia:**

$\alpha = 0.05$ para probar la hipótesis

5. Procedimiento:

Probabilidad de tener los siguientes arreglos de 0, 2, o 4 mujeres se dan de manera siguiente: Primero probabilidad de tener 1 mujer es 0.75 ya que en base a los antecedentes del partido3 de los ganadores son mujeres y por tanto p = ¾ = .75, y consecuentemente, q = 1 – p = .25

Por tanto, probabilidad de tener 4 varones, 2 y 2 y solamente varones ($_4C_4$), 2 y 2 ($_4C_2$), y 0 hijas ($_4C_0$) son:

$P_x = {}_nC_r\,(p^r\, q^{(n - r)})$ o $P_x = \{n! / [r!\,(n – r)!]\}\,(p^r\, q^{(n – r)})$

$p(4) = {}_4C_4\,(.75^4\ .25^{(4 - 4)})$ o $P_x = \{4! / [4!\,(4 – 4)!]\}\,(.75^4\ .25^{(4 - 4)}) = 81/256 = .316$

$p(2) = {}_4C_2\,(.75^2\ .25^{(4 - 2)})$ o $P_x = \{4! / [2!\,(4 – 2)!]\}\,(.75^2\ .25^{(4 - 2)}) = 54/256 = .210$

$p(0) = {}_4C_0\,(.75^0\ .25^{(4 - 0)})$ o $P_x = \{4! / [0!\,(4 – 0)!]\}\,(.75^0\ .25^{(4 - 0)}) = 1/256 = .0039$

6. Inferencia o toma de decisión:

La probabilidad de tener 4 **mujeres ganadores** es igual a .316 (31.6%), tener 2 varones es igual a .21 (21%), y tener 4 **varones** es igual a .0039 (.39%). El último valor de p (.0039) que indica tener varones ganadores es menor que $\alpha = 0.05$ y por tanto: Ho se rechaza, es decir, Concebir hijos o hijas no es aleatorio.

7. Interpretación:

La probabilidad de tener SOLO ganadores del género masculino es muy baja (.39%), en contrario, la probabilidad de tener ganadores del sexo femenino es alta (31.6%) lo cual sería lo más probable, y finalmente el deseo del partido político de tener 2 varones y 2 mujeres ganadores es 21%.

Modelo de Ji^2 de Pearson

1. **Descripción:** Se utiliza para 1, 2 o más muestras independientes con varias variables en escala Nominal. Cuando el tamaño de muestra es mayor que 20, se reduce la probabilidad de cometer error tipo I y II. Sin embargo toma en cuenta las 2 siguientes condiciones:

(a) Cuando maneja Tabla de Contingencia de 2X2, el número de sujetos es ≤ 30, y hay 0 observaciones a algunas celdas, usar la prueba Exacta de Fisher & yates.
(b) Con múltiples grupos (muestras) pero con frecuencias menores que 5, usar Prueba $(Ji)^2$ de Proporciones.

Estas 2 opciones limitan error Tipo I e incrementan Potencia –eficacia de muestras pequeñas. Donde eficacia es reducir la probabilidad de rechazar un Ho cierto (error Tipo I) y la Potencia es el aumentar la probabilidad de rechazar un Ho falso.

2. **Ecuación: $X^2 = \sum [(O - E)^2 / E]$**
Donde,
X^2 = Valor calculado del modelo
∑ = Signo de sumatorio
O = Frecuencia observada
E = Frecuencia esperada o teórica
gl = K -1, donde, K = número de muestras o grupos

Ejemplo.

A un grupo o muestra de 36 alumnos divididos en 4 subgrupos de 9 cada uno, se aplican 4 métodos de enseñanza de una materia específica y se encontró datos diferentes de calificación de diferentes grupos. Se analizó los 36 casos para ver si hay diferencia significativa en función de método de enseñanza (Tabla 1). En la hilera superior o observada, cada dato en cada celda indica el número de los alumnos que pasaron con calificación aprobatoria. En caso de la hilera inferior o esperada, cada valor indica en número de los alumnos que teóricamente deben pasar la materia, en este caso ¼ (n) = ¼ (36) = 9. Las columnas indican el método de enseñanza del 1 hasta el 4.

Tabla 1. # de muertos bajo diferentes tratamiento o radiación.					**Total**
	Método 1	**Método 2**	**Método 3**	**Método 4**	
Observado (O)	12	7	8	9	36
Esperado (E)	9	9	9	9	36

3. **Plantear hipótesis:**

Ho: Las diferencias en la aprobación por tipo de método son efecto aleatorio
Ha: El método 1 donde hay mayor número de aprobación es mejor que otros métodos

4. Nivele de significancia:

$\alpha = 0.05$ para probar la hipótesis

5. Zona de rechazo:

Para todo valor menor que $p = .05$ ($p \leq .05$) se rechaza la Ho y viceversa.

6. Procedimiento:

Primero, hay que calcular los valores esperados. En este caso para una situación aleatorio del efecto del métod se maneja 36/4 = 9 y por tanto: para cuantificar el valor calculado del modelo de $X^2 = \sum [(O - E)^2 / E]$, se procede con $X^2 = [(12 - 9)^2 + (7 - 9)^2 + (8 - 9)^2 + (9 - 9)^2] / 9 = 1.55$
El gl es igual a k – 1 0 4 -1 = 3, y el valor tabulado para 3 gl es igual a 7. 815

7. Tomar decisión:

Valor calculado del modelo de 1.55 es menor que valor crítico de 7.815 y por tanto Ho se Acepta.

8. Interpretación:

Todos los tratamientos (métodos) tienen el mismo efecto de enseñanza, y por tanto, el hecho de que el método 1 aparentemente está causando mayor aprobación, se debe a un efecto aleatorio, es decir, no hay diferencia entre las 4 métodos en término de la enseñanza y por ende la calificación aprobatoria.

Alternativas a Ji^2

*Cuando la n es mayor que 30 la **Potencia-Eficacia** incrementa, es decir, la p de rechazar una Ho Correcto se reduce y la p de aceptar una Ho Falsa también se reduce.*
Pero cuando la n es menor que 30 hay 2 alternativas en lugar de usar la Ji^2.

1. *Cuando tenemos una (a) en caso de una Tabla de Contingencia (T.C.) de 2 x 2, la muestra total es menor que 30 y (b) incluye 0 observaciones en alguna de las celdas, la prueba correcta sería la Probabilidad Exacta de Fisher y Yates.*

2. *Con hay grupos múltiples y con frecuencias menores que 5, se recomienda el uso de Ji^2 de proporciones.*

Probabilidad Exacta de Fisher & Yates

Esta prueba tiene mayor Potencia-Eficacia que Ji^2 normal Cuando a) se usa una T.C. de 2 x 2, la n es menor que 30 y b) hay alguna celda con valor de 0.

1. Ecuación:

P = [(A+B)! (C+D)! (A+C)! (B+D)!] / GT! (A! B! C! D!)

Donde,

P = Probabilidad exacta de Fisher & yates. Este es el nivel de probabilidad calculada del modelo.

A, B, C y D son las observaciones de las celdas de la Tabla 2.
GT = Gran total
! = Signo de factorial.

Tabla 2. Arreglo de la tabla para la Prueba Exacta de Fisher y Yates.

Grupos	**Variable**		**Total de Hilera**
	X	**Y**	
1	A	B	**A+B**
2	C	D	**C+D**
Total de Columna	**A+C**	**B+D**	**GT = A+B+C+D**

Ejemplo:

De un grupo de 15 alumnos, se selecciona 7 de ellos como grupo de tratamiento los cuales reciben un método nuevo de enseñanza y el resto (8) forman el grupo de control que recibe el método tradicional de enseñanza (Tabla 3). En ésta tabla, los datos de las celdas de cada columna indican el número de alumnos que obtuvieron calificaciones aprobatorias (6 y 2 en la columna aprobados) y no aprobatorios (1 y 6 en la columna no aprobados).

2. Hipótesis

Ho: Las frecuencias observadas para los 2 tipos de métodos son iguales y la diferencia es al azar
Ho: Las frecuencias observadas para los 2 tipos de métodos No son iguales y la diferencia es significativa, es decir, que hay mayor frecuencia de aprobación para el método nuevo versus el método tradicional

3. Nivel de significancia: Para todo valor de p menor o igual que 0.05, se rechaza Ho.

4. Zona de rechazo: Para todo valor mayor que 0.05, se acepta Ho.

5. Procedimiento:

Tabla 3. Los alumnos bajo el método tradicional (control) y el método nuevo (tratamiento).

Grupos	**Variable**		**Total de Hilera**
	Aprobados	**No aprobados**	
Método nuevo	6	1	**7**
Método tradicional	2	6	**8**
Total de Columna	**8**	**7**	**15**

P = [(A+B)! (C+D)! (A+C)! (B+D)!] / GT! (A! B! C! D!)
P = 7! 8! 8! 7! / 15! (6! 1! 2! 6!) = .031

6. Toma de decisión:

El valor calculado (0.031) es menor que 0.05, y por tanto, Ho se rechaza. Hay que acordar que este valor de 0.031 es el nivel de la probabilidad qye se debe comparar con el nivel de probabilidad de 0.05.

7. Interpretación:

A nivel de probabilidad de 0.031 (un nivel de confianza de casi 97%), el método nuevo es significativamente efectivo en la aprobación comparando con el método tradicional para el problema bajo el estudio.

Prueba de X_p^2 de proporción

La prueba de X_p^2 de proporción es para 2 o más muestras independiente (se menciona solo para ilustrar la otra alternativa a la prueba Ji^2), y cuando el tamaño de la muestra o el gran total es pequeño y hay cero en alguna celda. Esta prueba provee la misma potencia-eficacia que la prueba Ji^2 con muestras de tamaños muy grandes.

1. Ecuación:

$$X_p^2 = (1/m_p * m_q) \sum Ni\,(p - m_p)^2$$

Donde,
X_p^2 = Valor estadístico del modelo de X_p^2
m_p = Proporción promedio de que sucede, ocurre o acontezca el evento
m_q = Proporción promedio de que no ocurre el evento
p = Proporción de ocurrencia del evento
Ni = Tamaño de la muestra de subgrupo "i"

Ejemplo.

Se conduce un experimento en diseño factorial de 2 x 3, con 24 alumnos (12 con un método de enseñanza nueva y 12 con un método de enseñanza tradicional) para ver el efecto de prácticas de laboratorio medida vía el aumento o reducción de # de alumnos aprobados (Tabla 4). En esta tabla, la ni indica el número de los alumnos por cada hilera o subgrupo.

Tabla 4. Efecto de prácticas de laboratorio en aprobación (MN= Método Nuevo, MT= Método Tradicional).

Subgrupo	Reducción en aprobación	Aumento de aprobación	ni	P (reducción en Aprobación)
MN & PL	1	3	4	0.25
MN& Control	0	4	4	0.00
MN & Aislados	3	1	4	0.75
MT& PL	1	3	4	0.25
MT & Control	2	2	4	0.50
MT& Aislados	4	0	4	1.00
Total	**11**	**13**	**24**	Proporción de ocurrencia *al azar* del evento **(m_p) = 12/24 = 0.5**

2. Hipótesis:

Ho: La diferencia entre número de alumnos aprobados y no aprobados es al azar
Ha: La diferencia entre aprobaciones y reprobaciones es significativa y se debe a prácticas de laboratorio

3. Nivel de significancia. Para todo valor de probabilidad ≤ 0, Ho se rechaza.

4. Zona de rechazo. Para todo valor de probabilidad > 0, Ho se acepta.

5. Procedimiento.

$$\mathbf{X_p^2 = (1/m_p * m_q) \sum Ni\ (p - m_p)^2}$$

$X_p^2 = (1 /0.5*0.5)\ [4(.25 - 0.5)^2 + 4(0 - 0.5)^2 + \ldots 4(1 - 0.5)^2] = 11$
gl = Número de Hileras – 1 = 6 - 1 = 5

Xi^2 de Tabla con 5 gl es igual a 11.07.

6. Toma de decisión. El valor calculado es menor que el valor tabulado y por tanto Ho se acepta.

7. Interpretación. La diferencia entre el número de alumnos aprobados entre los individuos con y sin práctica de laboratorio se debe al azar.

Referencias

Dynkin, E.B. 1961. Necessary and sufficient statistics for a family of probability distributions. Mathematical Statistics and probability. 1: 23-41.

Eberhardt, K.R. and M.A. Flinger. 1977. A comparison of two tests for equality of two proportions. Amer. Statist. 31: 151-155.

Freedman, D., R. Pisani, R. Purves and A. Adhikari. 1991. Statistics. 2nd edition. New York: Norton.

Gleser, L.J. and J.D. Healy. 1976. Estimating the mean of a normal distribution with known coefficient of variation. J. Amer. Statis. Assoc. 71: 977-981.

Hettmansperger, T.P. and J.W. McKean. 1998. Robust Nonparametric Statistical methods. New York: Wiley.

Huber, P.J. 1981. Robust Statistics. New York: Wiley.

Hohnson, N.L. and S. Kotz. 1972. Distributions in statistics. New York: Wiley.

Johnson, N.L., S. Kotz and N. Balakrishnan. 1995. Continuous Univariate Distributions. Vol. 2. 2nd. Edition. New York: Wiley.

Johnson, N.L., S. Kotz and N. Balakrishnan. 1995. Univariate Discrete Distributions. Vol. 2. 2nd. Edition. New York: Wiley.

Waller, L.A. 1995. Dose the characteristic function numerically distinguish distributions? Amer Statis. 49: 150-151.

Waller, L.A., B.W. Turnball J.M. Hardin. 1995. Obtaining distribution functions by numerical inversion of characteristic functions with applications. Amer. Statist. 49. 346-350.

---

*UANL, San Nicolás, N.L., México, [1]U. Agraria de la Habana, Cuba