11. PRUEBAS NO PARAMÉTRICAS

Edgar Acuña
http://math.uprm/edu/~edgar

UNIVERSIDAD DE PUERTO RICO
RECINTO UNIVERSITARIO DE MAYAGUEZ

PRUEBAS NO PARAMÉTRICAS

Se estudiarán las pruebas noparamétricas, las cuales no requieren asumir normalidad de la población y que en su mayoría se basan en el ordenamiento de los datos. Todas las pruebas vistas en este capítulo requieren que la población sea continua. El parámetro que se usa para hacer las pruebas estadísticas es la Mediana y no la Media.

En **MINITAB**, para las pruebas noparamétricas se elige la secuencia **STAT** ▸ **Noparametrics.**

Pruebas No paramétricas para una sola muestra

1 Prueba de los Signos

Se usa para hacer pruebas de hipótesis acerca de la mediana de una población de una variable continua.

Ho: La Mediana poblacional es igual a un valor dado.

Ha: La mediana es menor (mayor ó distinta) del valor dado.

La prueba estadística está basada en la distribución Binomial con probabilidad de éxito p=½, puesto que la probabilidad de que un dato sea mayor o menor que la mediana es ½.

Para calcularla se determinan las diferencias de los datos con respecto al valor dado de la mediana y se cuentan los signos positivos y negativos.

Pruebas Noparamétricas para una sola muestra

Cuando la hipótesis alterna es "mayor que" y el número de diferencias positivas es mayor que las diferencias negativas entonces, el "p-value" se calcula por

$$P_1 = \sum_{i=0}^{c} \binom{n}{i} (\frac{1}{2})^n$$

donde c es el número de diferencias positivas y, *n* es igual al número de datos pero, si hay datos de valor igual a la mediana que se asume en la hipótesis nula entonces, *n* es igual al número de datos menos la cantidad de datos iguales a la mediana asumida,

cuando el número de diferencias positivas es menor que el número de diferencias negativas entonces el "*p-value*" es igual a

$$P_2 = \sum_{i=c}^{n} \binom{n}{i} (\frac{1}{2})^n$$

Pruebas Noparamétricas para una sola muestra

Si la hipótesis alterna es "menor que" y el número de diferencias positivas es mayor que el número de diferencias negativas entonces "*p-value*" = P_2 en caso contrario "*p-value*" = P_1. Cuando la hipótesis alterna es de dos lados y el número de diferencias positivas son mayores que el número de diferencias negativas entonces el "*p-value*" = $2P_2$, si hay menor número de diferencias positivas entonces "*p-value*"=$2P_1$ y si hay igual número de diferencias positivas y negativas entonces, "*p-value*"=1.
Si $n>20$ se puede usar aproximación Normal a una Binomial con $p = q = 0.5$, para calcular los "*p-values*". Es decir,

$$Z = \frac{X - .5n.}{.5\sqrt{n}}$$

Ejemplo

Probar si los datos del tiempo de vida después del transplante del ejemplo 7.5 sugieren que la mediana sea distinta de 5.

Ho, es que la mediana del tiempo de sobrevivencia es igual a 5 años

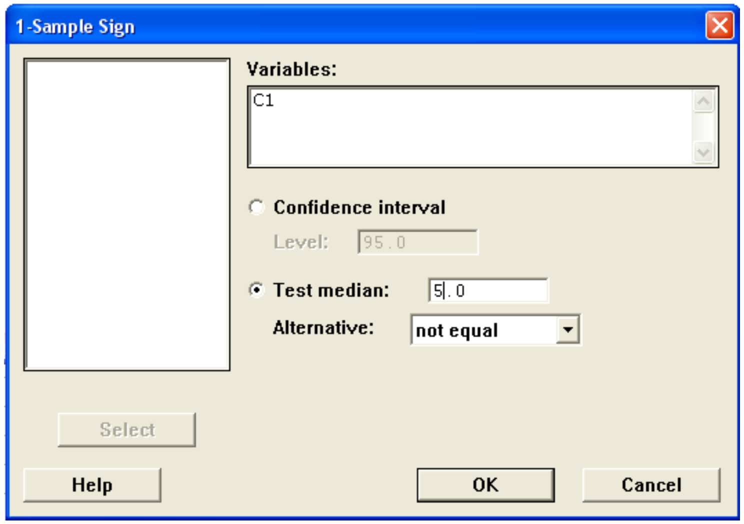
Ha, es que la mediana del tiempo de sobrevivencia es distinta de 5 años.

Sign Test for Median: tiempo

Sign test of median = 5.000 versus not = 5.000

	N	Below	Equal	Above	P	Median
tiempo	12	7	0	5	0.7744	3.700

Interpretación: Como el "p-value" es mayor que .05 se aceptará la hipótesis nula. Es decir que la mediana del tiempo de vida después del transplante es 5. En este ejemplo el "p-value" es 2 veces la probabilidad de que una binomial con n=12 y p =.5 sea menor o igual que 5, ya que el número de diferencias positivas es menor que el de las negativas.



Si usamos aproximación normal a la binomial el *p-value* = $2P(X \leq 5) = .77728$,

Pruebas Noparamétricas para una sola muestra

2 La Prueba de Rangos con signos de Wilcoxon

Es usada para hacer pruebas de hipótesis acerca de la mediana.

La prueba estadística se basa en el estadístico de Wilcoxon (1945), el cual se calcula de la siguiente manera:

- Se resta de cada dato el valor de la mediana que se considera en la hipótesis nula.
- Se calcula los rangos de las diferencias sin tomar en cuenta el signo de las mismas (o sea en valor absoluto). En el caso de haber empate se asigna un rango promedio a todas las diferencias empatadas es decir; se les asigna el rango:
  (menor rango del grupo del empate + mayor rango del grupo del empate)/2.
- El estadístico W de Wilcoxon será la suma de los rangos correspondientes a las diferencias positivas.

Pruebas Noparamétricas para una sola muestra

Cuando la hipótesis alterna es "mayor que" y la suma de los rangos correspondientes a las diferencias positivas es mayor que el de las diferencias negativas, entonces el "*p-value*" se calcula por $P_1=P(W \geq W_c)$, donde W_c es el valor calculado de la prueba de Wilcoxon.

Cuando la suma de los rangos correspondientes a las diferencias positivas es menor que el de las diferencias negativas, entonces el "*p-value*" se calcula por $P_2=P(W \leq W_c)$

Pruebas Noparamétricas para una sola muestra

Si la hipótesis alterna es "menor que", y la suma de los rangos correspondientes a las diferencias positivas es mayor que el de las diferencias negativas, entonces "*p-value*"=P_2. En caso contrario "*p-value*"=P_1.

Cuando la hipótesis alterna es de dos lados y la suma de los rangos correspondientes a las diferencias positivas es mayor que el de las diferencias negativas, entonces el "*p-value*"=2P_2, si la suma de los rangos correspondientes a las diferencias positivas es la menor entonces "*p-value*"=2P_1 y si las sumas de los rangos correspondientes a las diferencias positivas y negativas son iguales entonces "*p-value*"=1.0.

Ejemplo

Probar si los datos del tiempo de vida después del transplante del ejemplo 7.5 sugieren que la mediana sea distinta de 5.

Solución:

La hipótesis nula Ho, es que la mediana del tiempo de sobrevivencia es igual a 5 años y la hipótesis alterna Ha, es que la mediana de los tiempos de sobrevivencia es distinta de 5 años.

Sign Test for Median: tiempo

Sign test of median = 5.000 versus not = 5.000

	N	Below	Equal	Above	P	Median
tiempo	12	7	0	5	0.7744	3.700

Interpretación: *Como el "P-value" es mayor que .05 se aceptará la hipótesis nula. Es decir que la mediana del tiempo de vida después del transplante es 5.0. En este ejemplo el "P-value" es 2 veces la probabilidad de que una binomial con n=12 y p=.5 sea menor o igual que 5, ya que el número de diferencias positivas es menor que el de las negativas.*

La Prueba de Rangos con signos de Wilcoxon

Al igual que la prueba de los signos, es usada para hacer pruebas de hipótesis acerca de la mediana. La prueba estadística se basa en el estadístico de Wilcoxon (1945), el cual se calcula de la siguiente manera:

- Se resta de cada dato el valor de la mediana que se considera en la hipótesis nula.
- Se calcula los rangos de las diferencias sin tomar en cuenta el signo de las mismas (o sea en valor absoluto). En el caso de haber empate se asigna un rango promedio a todas las diferencias empatadas es decir; se les asigna el rango: (menor rango del grupo del empate + mayor rango del grupo del empate)/2.
- Finalmente el estadístico W de Wilcoxon será la suma de los rangos correspondientes a las diferencias positivas.

La Prueba de Rangos con signos de Wilcoxon

- **Cuando la hipótesis alterna es "mayor que"** y la suma de los rangos correspondientes a las diferencias positivas es mayor que el de las diferencias negativas, entonces el "*p-value*" se calcula por $P_1=P(W \geq W_c)$,

  Cuando la suma de los rangos correspondientes a las diferencias positivas es menor que el de las diferencias negativas, entonces el "*p-value*" se calcula por $P_2=P(W \leq W_c)$.
- **Si la hipótesis alterna es "menor que"**, y la suma de los rangos correspondientes a las diferencias positivas es mayor que el de las diferencias negativas, entonces "p-value"=P_2. En caso contrario "p-value"=P_1.
- **Cuando la hipótesis alterna es de dos lados** y la suma de los rangos correspondientes a las diferencias positivas es mayor que el de las diferencias negativas, entonces el "p-value"=$2P_2$, si la suma de los rangos correspondientes a las diferencias positivas es la menor entonces "p-value"=$2P_1$ y si las sumas de los rangos correspondientes a las diferencias positivas y negativas son iguales entonces "p-value"=1.0.

Cuando n es mayor que 16, se usa aproximación Normal para hallar el "p-value" de la prueba pues, se puede mostrar que el estadístico de Wilcoxon se aproxima a una normal con media igual a $n(n+1)/4$, y varianza $n(n+1)(2n+1)/24$, cuando no hay empates.

$$z = \frac{W - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \sim N(0,1)$$

Si hubiera empates entonces, la varianza sufre una ligera modificación

$$z = \frac{W - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \sum_{i=1}^{g} \frac{t_i^3 - t_i}{2}}} \sim N(0,1)$$

g es el número de grupos empatados y t_i es el tamaño del i-ésimo grupo empatado.

En **MINITAB**, para hacer la prueba de Wilcoxon se sigue la secuencia

STAT ▸ Noparametrics ▸ 1-Sample Wilcoxon.

Ejemplo. Aplicar la prueba de Wilcoxon a los datos del ejemplo anterior.

Solución: La ventana de diálogo se completará como se muestra en la figura.

Los resultados en la ventana **session** serán

Wilcoxon Signed Rank CI: tiempo

				Confidence	
		Estimated	Achieved	Interval	
	N	Median	Confidence	Lower	Upper
tiempo	12	4.63	94.5	1.85	7.30

Interpretación: *Como el "p-value"=.906 es mayor que .05 no se rechaza la hipótesis nula. Es decir, hay suficiente evidencia estadística para concluir que la mediana de los tiempos de vida es 5.0.*

Pruebas Noparamétricas para muestras pareadas

La prueba de los signos y la prueba de Wilcoxon se pueden usar también como una prueba alterna a la prueba de *t* para comparaciones pareadas. En este caso se aplica la prueba noparamétrica a las diferencias entre los dos grupos.

Ejemplo 11.3. Se desea probar si el rendimiento en la prueba de aprovechamiento matemático es mejor que en la prueba de aptitud matemática. Para ello se toma una muestra de los resultados de 40 estudiantes:

Wilcoxon Signed Rank Test: diferenc

```
Test of median = 0.000000 versus median > 0.000000
              N
            for   Wilcoxon         Estimated
         N  Test  Statistic      P    Median
diferenc  40   40      591.0  0.008     27.75
```

Interpretación: Como el “p-value” es menor que .05, se rechaza la hipótesis nula y se concluye que hay evidencia estadística de que el rendimiento en aprovechamiento es mejor que en aptitud.

La prueba de Mann-Withney para dos muestras independientes

Se usa cuando se quiere comparar dos poblaciones usando muestras independientes, es decir; es una prueba alterna a la prueba de *t* para comparar dos medias usando muestras independientes. También es conocida como la ***prueba de suma de rangos de Wilcoxon***.

La hipótesis nula es que la mediana de las dos poblaciones son iguales y la hipótesis alterna puede ser que la mediana de la población 1 sea mayor (menor ó distinta) de la mediana de la población 2.

Cuando tanto n_1 como n_2 sean mayores que 10, se puede demostrar que si **No hay empates**, entonces W se distribuye aproximadamente como una normal con media $n1(n_1+n_2+1)/2$ y varianza $n_1n_2(n_1+n_2+1)/12$.

$$z = \frac{W - \frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}} \sim N(0,1)$$

La prueba de Mann-Withney para dos muestras independientes

Cuando **hay empates** entonces, la varianza es modificada y se obtiene:

$$z = \frac{W - \frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2}{12}[n_1+n_2+1-\sum_{i=1}^{g}\frac{t_i^3 - t_i}{(n_1+n_2)(n_1+n_2-1)}}} \sim N(0,1)$$

donde, g y t_i tienen el mismo significado dado anteriormente.

En **MINITAB**, para hacer la prueba de Mann-Withney, se sigue la secuencia **STAT** ▸ **Noparametrics** ▸ **Mann-Withney.**

Ejemplo

Usando los datos del ejemplo 7.11 probar si el rendimiento en la prueba de aprovechamiento matemático de los estudiantes de escuela pública y privada es el mismo. Los datos son como siguen:

privada	pública
642	580
767	638
641	704
721	694
625	615
689	617
	623
	689

Solución

Mann-Whitney Test and CI: privada, pública

	N	Median
privada	6	665.5
pública	8	630.5

Point estimate for ETA1-ETA2 is 26.5

95.5 Percent CI for ETA1-ETA2 is (-47.0,104.0)

W = 56.5

Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.1556

The test is significant at 0.1551 (adjusted for ties)

Interpretación: Como el "p-value" 0.1551 (ajustado por empates), es mayor que 0.05 se acepta hipótesis nula. Es decir; que hay evidencia estadística para concluir que el rendimiento en aprovechamiento matemático es el mismo para estudiantes de escuela pública y privada.

La prueba de Kruskal-Wallis para comparar más de dos grupos

La prueba de Kruskal-Wallis, es una alternativa a la prueba F del análisis de varianza para diseños de clasificación simple. En este caso se comparan varios grupos pero usando la mediana de cada uno de ellos, en lugar de las medias.

Ho: La mediana de las k poblaciones consideradas son iguales y

Ha: Al menos una de las poblaciones tiene mediana distinta a las otras.

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

donde, *n* es el total de datos.

La prueba de Kruskal-Wallis para comparar más de dos grupos

Si hay empates en los datos entonces, se aplica la siguiente modificación a *H*.

$$H' = \frac{H}{1 - \frac{\sum_{i=1}^{g} t_i^3 - t_i}{n^3 - n}}$$

Se puede mostrar que si los tamaños de cada grupo son mayores que 5 entonces, *H* se distribuye como una Ji-Cuadrado con, *k*-1 grados de libertad. Luego, la hipótesis nula se rechaza si $H > \chi^2_{k-1,1-\alpha}$.

Para hacer la prueba de Kruskal-Wallis en **MINITAB,** los datos de la variable cuantitativa deben ir en una columna y los niveles del factor en otra. No se permite en este caso entrar los grupos en columnas separadas.

Ejemplo

Usar la prueba de Kruskal-Wallis para comparar los métodos de enseñanza del ejemplo 10.1

Solución:

Ho: Las medianas de los tres métodos de enseñanza son iguales y

Ha: Al menos uno de los métodos de enseñanza tiene mediana distinta a los otros.

Kruskal-Wallis Test: notas versus método

Kruskal-Wallis Test on notas

método	N	Median	Ave Rank	Z
1	6	61.50	5.4	-2.29
2	7	85.00	13.8	2.72
3	5	74.00	8.4	-0.54
Overall	18		9.5	

H = 8.23 DF = 2 P = 0.016

H = 8.25 DF = 2 P = 0.016 (adjusted for ties)

Interpretación: Como el “p-value” es 0.016 menor que .05, se rechaza la hipótesis nula y se concluye que los métodos no son todos iguales. Es decir; al menos uno de los métodos tiene mediana distinta a los otros.

El Coeficiente de Correlación de Spearman

La correlación de Spearman mide el grado de asociación entre dos variables cuantitativas que siguen una tendencia siempre creciente o siempre decreciente. es más general que el Coeficiente de correlación de Pearson, la correlación de Spearman, en cambio se puede calcular para relaciones exponenciales o logarítmicas entre las variables.
Para hallar los ordenamientos, se usa la opción ***Rank*** del menú **Calc.** Los ordenamientos se guardan en otras columnas y luego se halla simplemente el coeficiente de correlación usual entre éstas dos columnas usando la opción **correlación** del submenú **Basic Statistics** del menú **STAT.**

MINITAB también incluye en el menú de *Pruebas Noparamétricas* a la ***Prueba de Friedman*** para análisis de diseños en bloques al azar y la ***prueba de Mood.***

Ejemplo

Calcular el coeficiente de Correlación de Spearman y compararlo con el coeficiente de correlación de Pearson para los siguientes datos:

Años como Realtor (X)	3	4	6	7	8	12	15	20	22	26
Casas Vendidas(Y)	9	12	16	19	23	119	34	37	40	45

Solución:

Ordenando los datos de cada variable se obtiene:

rankx	1	2	3	4	5	6	7	8	9	10
ranky	1	2	3	4	5	10	6	7	8	9

La correlación de Spearman de las variables X e Y será igual a la correlación de Pearson entre las variables rankx y ranky dando un valor de 0.879 lo que indica una alta asociación entre las variables. Sin embargo; la correlación de Pearson entre las variables X e Y da solamente 0.371, lo que indica una baja asociación lineal entre las variables.

Notar que el "outlier" 119 ha afectado grandemente al coeficiente de correlación de Pearson, pero no ha tenido efecto sobre la correlación de Spearman.