CAPÍTULO 11

PRUEBAS NOPARAMÉTRICAS

En las pruebas estadísticas que se han discutido hasta ahora se hacen suposiciones acerca de la forma como se distribuye la población, la que por lo general se asume que se distribuye normalmente. De no haber normalidad las pruebas estadísticas no son válidas. Como se ha visto en el capítulo 5 la normalidad de la población se puede cotejar en base a la muestra tomada. Frecuentemente se arriva a la conclusión de que no hay normalidad y en consecuencia las pruebas que se hacen no son muy confiables, pero a pesar de todo se usan.

En este capítulo 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 contínua. El parámetro que se usa para hacer las pruebas estadísticas es la Mediana y no la Media. Existen una serie de pruebas noparámetricas, nosotros sólo veremos las más usadas.

En **MINITAB**, las pruebas noparamétricas aparecen cuando se elige la secuencia **STAT** • **Noparametrics**.

11.1 Pruebas Noparamétricas para una sola muestra

11.1.1 Prueba de los Signos

Se usa para hacer pruebas de hipótesis acerca de la mediana de una población de una variable continua. Es una alternativa a la prueba de Z o de t para la media poblacional. La hipótesis nula es H_o: La Mediana poblacional es igual a un valor dado y la Hipótesis alterna H_a: La mediana es menor (mayor ó distinta) del valor dado.

La prueba estadística está basada en la distribución Binomial con probabilidad de éxito p=1/2, 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 cuenta los signos positivos y negativos.

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

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"=1.0.

Si n>20 se puede usar aproximación Normal a una Binomial con p=q=.5, para calcular los "p-values". Es decir,

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

La aproximación mejora si incluímos el factor de corrección por continuidad igual a 1/2.

En MINITAB, para hacer la prueba de los signos, se sigue la secuencia STAT Noparametrics 1-sample Sign.

Ejemplo 11.1 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 H_o , es que la mediana del tiempo de sobrevivencia es igual a 5 años y, la hipótesis alterna H_a , es que la mediana de los tiempos de sobrevivencia es distinta de 5 años.

La ventana de diálogo se completará como se muestra en la figura 11.1. En la ventana **session** aparecerán los siguientes resultados:

```
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.

Si usamos aproximación normal a la binomial el P-value= $2P(X \le 5)=2P(Z < \frac{5.5-6}{.5\sqrt{12}})=2P(Z < .2886)=2(.38864)=.77728$. El valor aproximado está bastante cerca al valor exacto a pesar de que el tamaño de muestra es n=12 menor que 20.

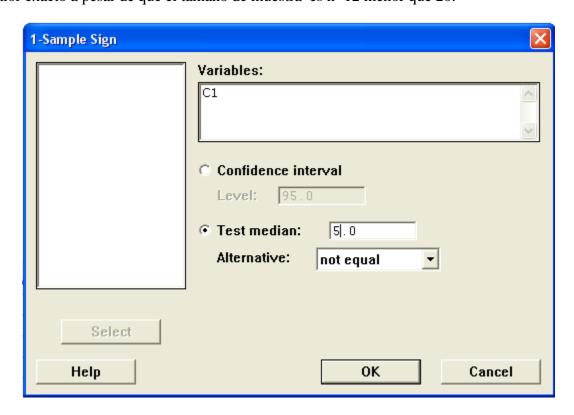


Figura 11.1. Ventana de diálogo para la prueba de signo del ejemplo 11.1

11.1.2 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:

- i) Se resta de cada dato el valor de la mediana que se considera en la hipótesis nula.
- ii) 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.
- iii) Finalmente el estadístico W de Wilcoxon será la suma de los rangos correspondientes a las diferencias positivas.

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 $P_2=P(W \le W_c)$.

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 correpondientes a las diferencias positivas y negativas son iguales entonces "p-value"=1.0.

Sea n, número de diferencias distintas de cero, es decir se está considerando que todos los valores de la muestra son distintos que el valor de la mediana que aparece en la hipótesis nula. Si $n \le 16$ entonces, los "p-values" se encuentran usando tablas de la distribución del estadístico de Wilcoxon.

Cuando n es mayor que 16, se usa aproximación Normal para hallar el "P-value" de la prueba pués, 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. Más especificamente, si no hay empates se tiene que:

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

No hay que olvidarse de aplicar un factor de corrección por continuidad igual a 1/2, pues se está aproximando una distribución discreta por una contínua. Si hubieran empates entonces, la varianza sufre una ligera moddificación y se aplica:

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

donde, 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 11.2. Aplicar la prueba de Wilcoxon a los datos del ejemplo anterior.

Solución: La ventana de diálogo se completará como se muestra en la figura 11.2

Los resultados en la ventana session serán:



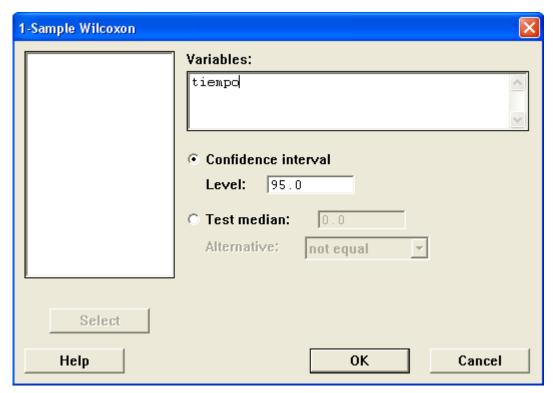


Figura 11.2. La ventana de diálogo para la prueba de Wilcoxon del ejemplo 11.2

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.

La media del estadístico de Wilcoxon es 39 y su varianza es 162.5. Como la hipótesis alterna es de dos lados entonces, el "P-value" es dos veces la probabilidad de que W≤37. Usando aproximación normal para calcular el "P-value", después de aplicar el factor de correción por continuidad y estandarizar, resulta que el "P-value"=2P(Z<-1.5/12.7475)=2(0.4532)=0.9064.

11.2 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. En el siguiente ejemplo se ilustra la prueba de Wilcoxon para comparar dos muestras pareadas.

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:

Row	aprovech	aptitud	diferenc
1 2 3 4 4 5 6 7 8 9 10 11 2 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 9 40	658 562 679 7310 6631 6654 6659 7100 7200 721 795 617 795 617 7641 7625 615 615 617 628 689 710 689 710 689 710 710 710 710 710 710 710 710 710 710	598 523 587 644 6616 6898 5694 674 674 674 674 674 674 674 674 674 67	60 -61 92 87 80 15 -19 56 -108 87 -25 63 46 91 13 122 -38 -77 -39 -13 -46 43 93 -19 16 -18 -86 -4 81 66 27 48 -34 64 65 112 116 117 117 117 118 119 119 119 119 119 119 119

```
Wilcoxon Signed Rank Test: diferenc
Test of median = 0.000000 versus median > 0.000000
                 Ν
               for
                     Wilcoxon
                                      Estimated
           Ν
              Test
                    Statistic
                                   Ρ
                                          Median
                                           27.75
          40
                40
                        591.0
                               0.008
diferenc
```

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.

11.3. 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.

Consideremos que se ha tomado una muestra de tamaño n_1 de la población 1 y de tamaño n_2 de la población 2. Para calcular la prueba estadística se combinan las dos muestras tomadas en una sola y se calculan los rangos en orden ascendente, en caso de datos empatados se asigna un rango promedio a ellos. Luego el estadístico W es igual a la suma de los rangos correspondientes a la muestra tomada de la población 1. Existen tablas para calcular los "p-values" de la prueba estadística.

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 $n_1(n_1+n_2+1)/2$ y varianza $n_1n_2(n_1+n_2+1)/12$. Es decir; cuando no hay empates:

$$z = \frac{W - \frac{n_1(n_1 + n_2 + 1)}{2}}{\sqrt{\frac{n_1n_2(n_1 + n_2 + 1)}{12}}} \sim N(0,1)$$

No hay que olvidarse de aplicar un factor de corrección por continuidad igual a 1/2, pués se está aproximando una distribución discreta por una continua. 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 11.4. 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: La ventana de diálogo se completará así:

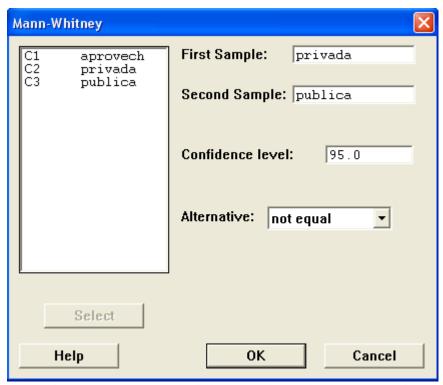


Figura 11.3. Ventana de diálogo para la prueba de Mann-Whitney del ejemplo 11.4

La ventana **session** mostrará los siguientes resultados:

```
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.

11.4 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. Es decir, la hipótesis nula es Ho: La mediana de las k poblaciones consideradas son iguales, y la alterna Ha: Al menos una de las poblaciones tiene mediana distinta a las otras.

La prueba estadística, denotada por H, se calcula hallando primero los rangos de cada una de los k grupos pero, considerando que se ha combinado todos los grupos en uno sólo. En caso de haber datos empatados se asigna un rango promedio a cada dato del grupo empatado.

Sea R_k la suma de los rangos del grupo k, el estadístico de Kruskal-Wallis necesario para hacer la prueba estadística se calcula por.

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

donde, n es el total de datos. Si hay empates en los datos entonces, se aplica la siguiente modificación a H.

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

$$1 - \frac{1}{n^3 - n}$$

Se puede mostrar que si los tamaños de cada grupo son mayores que 5 entonces, H se distribuye como una Chi-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 11.5. Usar la prueba de Kruskal-Wallis para comparar los métodos de enseñanza del ejemplo 10.1

Solución:

La hipótesis nula es Ho: Las medianas de los tres métodos de enseñanza son iguales y la hipótesis alterna es Ha: Al menos uno de los métodos de enseñanza tiene mediana distinta a los otros.

La ventana de diálogo se completará así:

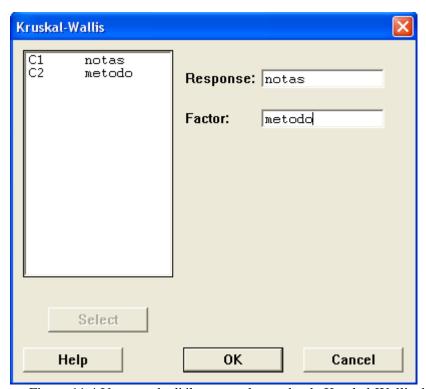


Figura 11.4 Ventana de diálogo para la prueba de Kruskal-Wallis del ejemplo 11.5

En la ventana **session** se obtendrá:

```
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.

11.5. 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 decir, es más general que el Coeficiente de correlación de Pearson, el cual asume que la relación entre las dos variables solamente es lineal, la correlación de Spearman, en cambio se puede calcular para relaciones exponenciales o logaritmicas entre las variables.

El coeficiente de correlación de Spearman es simplemente la correlación de Pearson entre los rangos del los valores de las dos 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**.

Ejemplo 11.6. Calcular el coeficiente de Correlación de Spearman y compararlo con el coeficiente de correlación de Pearson para los siguientes datos:

Años Realtor (X)	como	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.

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*.

EJERCICIOS

 En un hospital oncológico se llevan estadisticas acerca del tiempo de supervivencia de pacientes de cáncer Los resultados en una muestra aleatoria de 25 pacientes fueron los siguientes.

42 45 51 46 340 81 246 63 155 151 37 138 245 377 455 365 776 163 20 1234 201 2970 456 1235 1581

Usar una prueba noparamétrica para probar que la mediana del tiempo de supervivencia de pacientes de cáncer, es mayor de 300 dias. Escribir las hipótesis y comentar sus resultados.

2. Se eligen al azar 10 empleados de una empresa y se anotan sus sueldos mensuales (en dólares

1500 1250 900 800 1450 990 1200 1900 1300 1050

Usando un nivel de significación del 10% ¿se podrá concluir que el sueldo mediana excede a 1200?

3. El tiempo de producción (en horas) de un artículo de 15 máquinas elegidas al azar en una gran planta de fabricación son:

5.80 6.06 5.90 5.92 5.68 6.27 6.08 6.15 5.93 5.96 5.88 5.63 6.00 5.96 5.70 A un nivel de significación de 0.05:

- a) Probar si la mediana del tiempo requerido difiere de 5.8
- b) Probar si la mediana es mayor a 5.8
- 4. Las notas de una evaluación hecha a 40 estudiantes elegidos al azar son:

```
78 75 52 65 68 75 52 62 73 75 77 70 50 72 66 62 77 76 74 75 68 71 70 66 68 66 67 85 82 66 72 65 71 77 67 82 65 69 82 87
```

- a) Probar si la mediana de las notas difiere de 70. Usar $\alpha = 0.05$
- b) Probar si la mediana de las notas es menor de 70. Usar $\alpha = 0.05$
- 5. La efectividad de Bezendrine en acelerar el rítmo cardíaco (medido en pulsaciones por minuto), fue cotejado en 10 pacientes elegidos al azar. Cada paciente sirvió como su propio control con la mitad de los pacientes asignados al recibir Bezendrine durante el primer período de estudio y, la otra mitad a recibir un Placebo(solución alcalina). Todos los pacientes fueron examinados para determinar su rítmo cardiaco, 2 horas después de recibir el medicamento. Después de dos semanas donde no se les dio ninguna medicina a los pacientes que habían recibido el Placebo se les dio Bezendrine, y a la otra mitad el Placebo. Los resultados son como siguen:

Paciente Placebo Bezendrine 1 250 258

2	271	285
3	243	245
4	252	250
5	266	268
6	272	278
7	293	280
8	296	305
9	301	319
10	298	308

Usar una prueba no paramétrica para probar la efectividad del Bezendrine. Escribir las hipótesis correspondientes. Comparar su resultado con el de la prueba T

6. Se está estudiando la efectividad de un nuevo medicamento para reducir la presión arterial Sistólica, el medicamento fue suministrado a 20 pacientes . Se les ha medido la presión arterial, antes y dos horas después de suministrarles el medicamento. Los datos se presentan en la siguiente tabla:

Sujeto	pre. Inic	Pre. Post	sujeto	pre. Inic	pre. Post
1	102	103	11	118	114
2	142	140	12	144	139
3	185	182	13	136	137
4	110	108	14	130	126
5	143	140	15	121	125
6	131	129	16	151	150
7	115	111	17	137	135
8	124	126	18	142	136
9	150	145	19	120	117
10	108	108	20	153	149

A un nivel se significación del 1%, probar si hay evidencia suficiente para afirmar que el medicamento es efectivo.

7. Se desea comparar la eficacia de dos compuestos en la producción de glucosa en la sangre, para tratar a personas que padecen de Hipoglicemia, para tal propósito se seleccionaron al azar 7 ratones, los resultados del estudio se presentan en la siguiente tabla:

Ratón	Comp 1	Comp 2
1	4.6	5
2	5.3	5.2
3	3.8	3.5
4	7.2	6.3
5	8.4	8.6
6	4.8	4.2
7	3.5	4.4

¿A qué conclusión llegará usando un nivel de significación del 5 por ciento?

8. El presidente de una empresa piensa que el número de ausencias injustificadas para el personal gerencial es menor que la de los obreros. Para esto se eligen al azar 7 gerentes y 10 obreros, y se registran sus ausencias injustificadas durante un año.

Gerentes: 20 14 19 22 25 30 17

Obreros: 37 29 51 18 40 26 41 24 19 28

- a) Probar la hipótesis usando una prueba noparametrica. ¿A qué conclusión llega?
- b) Probar la hipótesis usando una prueba de t. ¿A qué conclusión llega? ¿Cómo se comparan los P-values?
- 9. Se seleccionaron al azar ministros de 3 religiones: 8 metodistas, 10 católicos y 9 pentecostales y se desea probar si poseen el mismo conocimiento sobre enfermedades mentales. Los resultados de un test para medir sus conocimientos son los siguientes

Católico	Pentecostal
32	28
32	21
26	15
26	15
22	14
20	14
14	09
16	11
14	08
15	
	32 32 26 26 22 20 14 16 14

Escribir la hipótesis correspondiente y usar una prueba noparamétrica para probarla. Analizar sus resultados y compararlo con la prueba F del análisis de varianza.

10. El peso (en libras), y estatura (en pulgadas) de 15 jóvenes se muestra en la siguiente tabla. Calcule el coeficiente de correlación de Spearman y compararlo con el coeficiente de correlación de Pearson.

estatura	Peso
4.8	115.3
4.9	124.9
5.1	123.8
5.2	137.2
5.3	138.3
4.8	113.1
5.2	137.9
4.8	101.2
5.4	131.9
4.8	102.7
4.9	115.0
5.3	130.5
5.2	108.0