ANOVA Simple

Resumen

El procedimiento **ANOVA Simple** (o de un criterio de clasificación) está diseñado para construir un modelo estadístico que describa el impacto de un solo factor categórico X sobre una variable dependiente Y. Se realizan pruebas para determinar si hay o no diferencias significativas entre las medias, varianzas y/o medianas de Y en los diferentes niveles de X. Además, los datos se pueden presentar gráficamente de varias formas, incluyendo un gráfico múltiple de dispersión, un gráfico de medias, un gráfico ANOM, y un gráfico de medianas.

En este procedimiento, se supone que los datos se colocarán en dos columnas, una para la variable dependiente Y y una segunda identificando los niveles de X. También se puede realizar un análisis de varianza simple empleando el procedimiento *Comparación de Varias Muestras*, que se debe usar si los datos se han puesto en columnas separadas para cada nivel de X.

StatFolio de Ejemplo: *oneway.sgp*

Datos de Ejemplo:

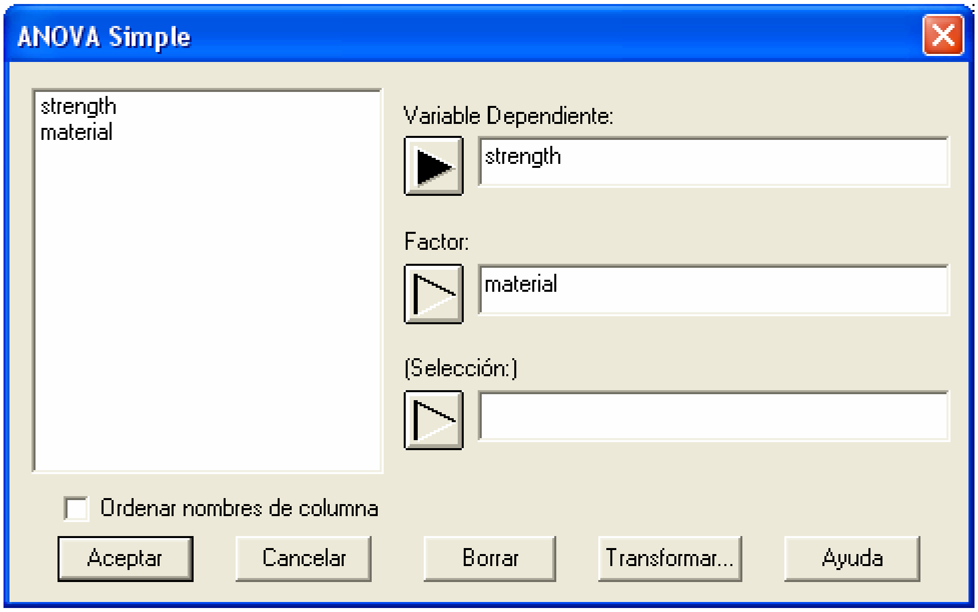
El archivo *breaking.sf6* contiene los resultados de un experimento en el cual se comparó la fuerza (*Strength*) de ruptura de artilugios de 4 diferentes materiales. A continuación se muestra una parte de los datos:

Strength	*Material*
39	A
57	A
42	A
32	A
43	A
50	A
31	A
51	A
27	B
43	B
25	B
11	B
39	B
…	…

El conjunto completo de los datos consiste de n = 32 artilugios, de los cuales 8 fueron hechos con cada uno de los q = 4 diferentes materiales.

Ingreso Datos

La caja de diálogo de ingreso de datos solicita los nombres de las columnas que contienen las mediciones Y y los niveles del factor X:



- **Variable Dependiente:** columna numérica que contiene las *n* observaciones de Y.

- **Factor:** columna numérica o no numérica que contiene un identificador para los niveles del factor X.

- **Selección:** selección de un subgrupo de datos.

Resumen del Análisis

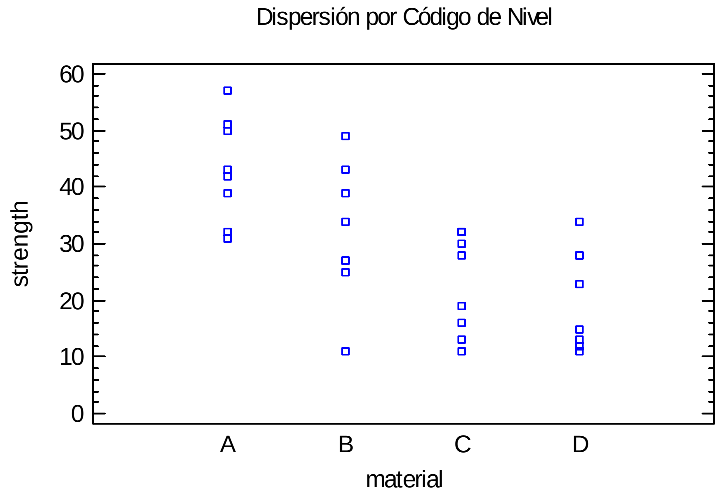
El *Resumen del Análisis* muestra el número de niveles de X y el número total de observaciones *n*.

ANOVA Simple - strength por material
Variable dependiente: strength
Factor: material

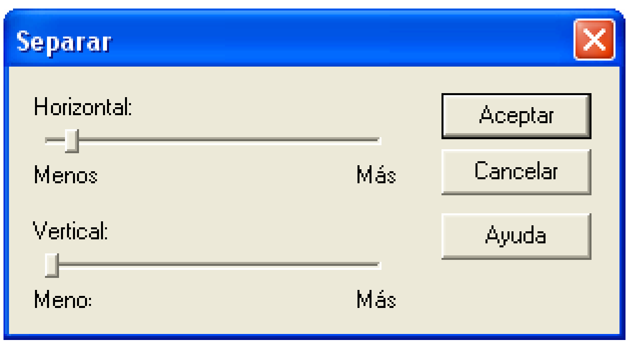
Número de observaciones: 32
Número de niveles: 4

Gráfico de Dispersión

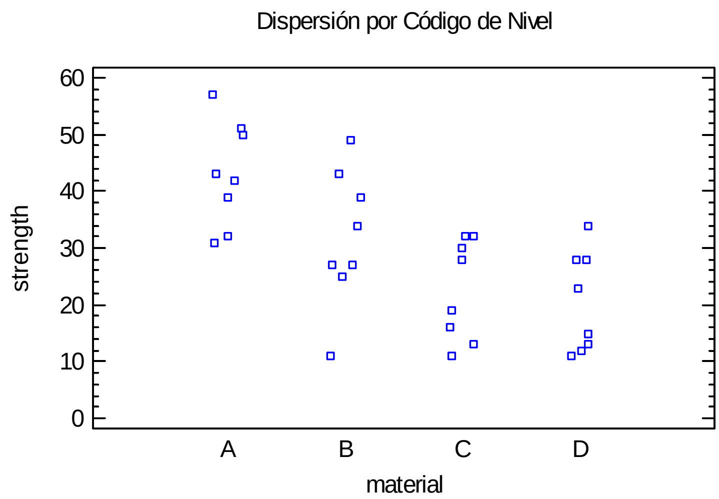
La ventana *Gráfico de Dispersión* grafica los datos por nivel del factor X.



Si hay muchos valores en común, tal vez quiera agregar algo de separación horizontal al gráfico presionando el botón *Separar* en la barra de herramientas del análisis:



Esto balancea cada punto aleatoriamente en dirección horizontal de tal manera que valores idénticos no se grafican encima uno de otro:



El gráfico anterior sugiere que hay diferencia en la fuerza de ruptura entre los cuatro materiales.

Resumen Estadístico

La ventana *Resumen Estadístico* calcula un número de diferentes estadísticas que se emplean comúnmente para resumir una muestra de datos de variables:

Resumen Estadístico para strength

material	*Recuento*	*Promedio*	*Desviación Estándar*	*Rango*	*Sesgo Estandarizado*	*Curtosis Estandarizada*
A	8	43.125	9.18753	26.0	0.0718672	-0.59635
B	8	31.875	11.9695	38.0	-0.369758	0.0369302
C	8	22.625	8.81456	21.0	-0.191636	-1.24894
D	8	20.5	8.86405	23.0	0.397909	-1.0005
Total	32	29.5313	13.0062	46.0	0.530301	-0.827139

La mayoría de las estadísticas caen en una de tres categorías:

1. Medidas de *tendencia central* – estadísticas que caracterizan el "centro" de los datos.
2. Medidas de *dispersión* – estadísticas que miden la dispersión de los datos.
3. Medidas de *forma* – estadísticas que miden la forma de los datos con respecto a una distribución normal.

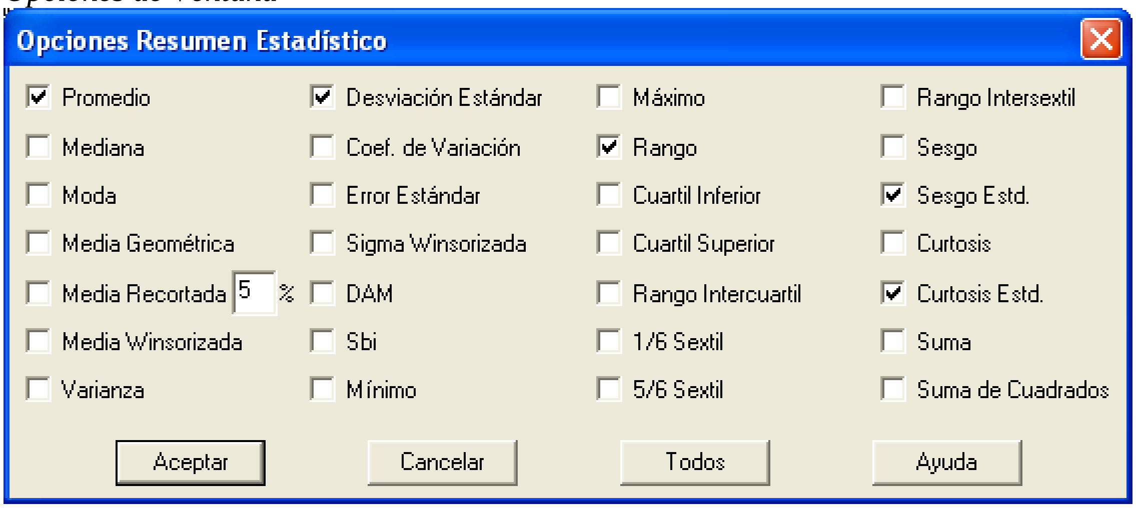
Las estadísticas incluidas en la tabla se controlan por las definiciones de configuración en la ventana *Estadísticas* de la caja de diálogo *Preferencias*. Estas selecciones pueden cambiarse usando la *Ventana de Opciones*. Para una descripción detallada de cada estadística, vea la documentación de *Análisis de Una Variable*.

De interés particular son:

1. *Medias muestrales* $\bar{Y}_j$: la fuerza de ruptura promedio para cada material.
2. *Desviaciones estándar muestrales* s_j: las desviaciones estándar para cada material.
3. *Sesgo y curtosis estandarizados*: Estas estadísticas deben estar entre –2 y +2 si los datos vienen de una distribución normal.

Para los artilugios, la mayor fuerza de ruptura promedio fue para el material A, seguida por la del material B. Ambas estadísticas sesgo y curtosis estandarizados están dentro del rango esperado para datos que provienen de una distribución normal.

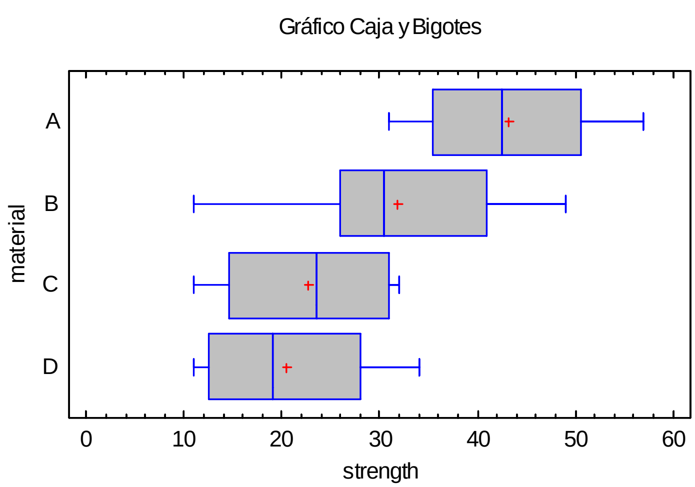
Opciones de Ventana



Seleccione las estadísticas deseadas.

Gráfico de Caja y Bigotes

Esta ventana presenta un gráfico de caja y bigotes para Y en cada nivel de X.



Los gráficos de caja y bigotes se construyen de la siguiente forma:

- Se dibuja una caja que se extienda desde el *cuartil inferior* de la muestra hasta el *cuartil superior*. Este es el intervalo cubierto por el 50% central de los valores de los datos cuando se ordenan de menor a mayor.

- Se dibuja una línea vertical en la *mediana* (el valor de en medio).

- Si se solicita, un signo de más se coloca en el lugar de la media muestral.

- Los bigotes se dibujan desde los extremos de la caja hasta los valores mínimo y máximo de los datos, a menos que haya valores inusualmente muy alejados de la caja (a los cuales Tukey llama *puntos extremos*). Los puntos extremos, que son puntos a más de 1.5 veces el rango intercuartílico (ancho de la caja) por arriba o por debajo de la caja, se indican por símbolos de señalamiento. Cualesquiera puntos a más de 3 veces el rango intercuartílico por arriba o por debajo de la caja se les llama *puntos extremos lejanos*, y se indican por símbolos de señalamiento con signos de más superpuestos por arriba de ellos. Si hay presentes puntos aberrantes (extremos o extremos lejanos), los bigotes se dibujan a los valores máximo y mínimo que no sean puntos extremos.

En los datos de la muestra, la variabilidad parece ser similar dentro de cada material, aunque la localización muestra algunas diferencias. No hay puntos aberrantes.

Ventana de Opciones

- **Dirección**: la orientación del gráfico, correspondiente a la dirección de los bigotes.

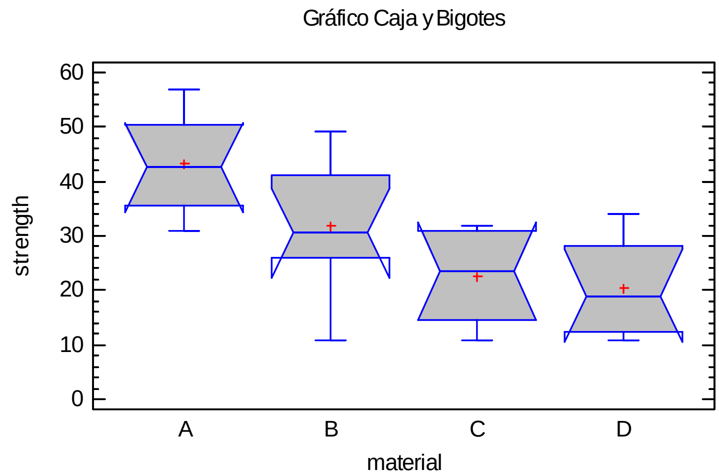
- **Muesca sobre la Mediana**: si se selecciona, se agregará una muesca al gráfico que muestra el error estimado asociado con cada mediana. Las muescas están escaladas de tal manera que, para muestras de igual tamaño, si no se traslapan, las medianas son significativamente diferentes al nivel de confianza por omisión del sistema(establecido en la pestaña *General* de la caja de diálogo de las *Preferencias* en el menú *Editar*).

- **Mostrar aberrantes**: si se selecciona, indica la localización de los puntos extremos.

- **Mostrar Media**: si se selecciona, muestra la localización de la media muestral así como la mediana.

<u>Ejemplo</u> – Gráfico de Caja y Bigote con Muescas

En el siguiente gráfico se agregaron muescas a las medianas a un nivel de confianza del 95%.



Cada muesca cubre el intervalo

$$\tilde{x}_j \pm \frac{z_{\alpha/2}}{2}\frac{1.25(RIC_j)}{1.35\sqrt{n_j}}\left(1+\frac{1}{\sqrt{2}}\right) \tag{1}$$

donde $\tilde{x}_j$ es la mediana de la *j*-ésima muestra, RIC_j es el rango intercuartílico muestral, n_j es el tamaño de la muestra, y $z_{\alpha/2}$ es el valor crítico (α/2)% superior de una distribución normal estándar. En casos donde el tamaño de la muestra es pequeño, puede que la muesca se extienda más allá de la caja, lo que resulta en una apariencia doblada.

En los datos de la muestra, la muesca para el material A no se traslapa con la de los materiales C ni D, indicando que la mediana es significativamente mayor para A que para C y D. Sin embargo, no es significativamente mayor que la del material B y que las muescas para A y B se traslapan.

Tabla de ANOVA

Para determinar si las medias de *q* grupos difieren o no significativamente unas de otras, se puede realizar un análisis de varianza simple (o de un criterio de clasificación). Los resultados se presentan en la *Tabla de ANOVA*:

Tabla ANOVA para strength por material

Fuente	*Suma de Cuadrados*	*Gl*	*Cuadrado Medio*	*Razón-F*	*Valor-P*
Entre grupos	2556.34	3	852.115	8.88	0.0003
Intra grupos	2687.63	28	95.9866		
Total (Corr.)	5243.97	31			

La tabla divide la variabilidad que existe en las *n* mediciones entre dos componentes:

1. Un componente "intra grupos", que mide la variabilidad de la fuerza de ruptura de los artilugios hechos del mismo material.

2. Un componente "entre grupos", que mide la variabilidad entre artilugios hechos de diferente material.

De particular importancia es la razón de F, que prueba la hipótesis de que la respuesta media para todos los niveles de X es la misma. Formalmente, prueba la hipótesis nula

$$H_0: \mu_1 = \mu_2 = ... = \mu_q$$

versus la hipótesis alterna

$$H_A: \text{no todas las } \mu_j \text{ iguales}$$

Si F es lo suficientemente grande, se rechaza la hipótesis nula.

La significancia estadística de la razón de F es mucho más fácil de juzgar por su valor de P. Si el valor de P es menor que 0.05, la hipótesis nula de medias iguales se rechaza al nivel de significancia del 5%, como en el presente ejemplo. Esto no implica que cada una de las medias es significativamente diferente de cada una de las otras. Simplemente implica que no todas las medias son iguales. Determinar qué medias son significativamente diferentes de cuáles otras requiere de pruebas adicionales, como se trata más adelante.

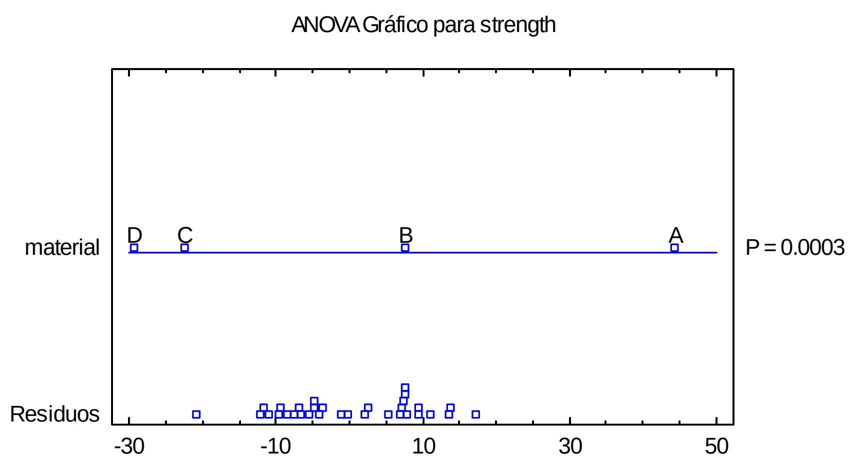
ANOVA Gráfico

El gráfico *ANOVA Gráfico*, desarrollado por Hunter (2005), es una técnica para exhibir gráficamente la importancia de las diferencias entre niveles del factor experimental. En un gráfico de los efectos escalados del factor, donde el "efecto" es igual a la diferencia entre la media para un nivel de ese factor y la gran media estimada. Cada uno de los efectos se multiplica por un factor de escala

$$\sqrt{\frac{\nu_R n_i}{\nu_T \bar{n}}} \qquad (2)$$

donde ν_R son los grados de libertad de los residuos, ν_T son los grados de libertad del factor, n_i es igual al número de observaciones en el *i-ésimo* nivel del factor, y $\bar{n}$ es el número promedio de observaciones en todos los niveles del factor. Esto escala los efectos de manera que la varianza natural de los puntos en el diagrama es comparable con la de los residuos, que se muestran en la parte de abajo del gráfico.

El gráfico para los datos de la muestra se muestra a continuación:



En el lado derecho del gráfico se muestra el valor de P para la diferencia entre grupos, tomado de la tabla de ANOVA.

Comparando la variabilidad de los efectos del factor el gráfico anterior con la de los residuos, es fácil ver que las diferencias son de mayor magnitud de lo que podría justificarse con tan solo el error experimental. Dependiendo de la localización relativa de los efectos, también puede ser posible en algunos casos identificar visualmente qué efectos son significativamente diferentes de cuáles otros, lo que se hace formalmente con las *Pruebas de Rangos Múltiples* descritas a continuación.

Pruebas de Rangos Múltiples

Para determinar qué medias muestrales son significativamente diferentes de cuáles otras, se pueden realizar las *Pruebas de Rangos Múltiples*:

Contraste Múltiple de Rango para strength por material

Método: 95.0 porcentaje LSD

material	*Casos*	*Media*	*Grupos Homogéneos*
D	8	20.5	X
C	8	22.625	XX
B	8	31.875	X
A	8	43.125	X

Contraste	*Sig.*	*Diferencia*	*+/- Límites*
A - B	*	11.25	10.0344
A - C	*	20.5	10.0344
A - D	*	22.625	10.0344
B - C		9.25	10.0344
B - D	*	11.375	10.0344
C - D		2.125	10.0344

* indica una diferencia significativa.

La mitad superior de la tabla presenta en orden creciente de magnitud cada una de las medias muestrales estimadas. Se muestran:

- **Casos –** el número de observaciones n_j.
- **Media –** la media muestral estimada $\bar{Y}_j$.
- **Grupos Homogéneos –** una ilustración gráfica de qué media es significativamente diferente de cuáles otras, basada en los contrastes exhibidos en la segunda mitad de la tabla. Cada columna de X’s indica un grupo de medias dentro de las cuales no hay diferencia estadísticamente significativa. Por ejemplo, la primer columna en la tabla anterior contiene una X para los materiales C y D, indicando que sus medias no son significativamente diferentes. De igual modo, los materiales B y C no muestran diferencias significativas. La media del material A, por otro lado, es significativamente mayor que la media de cualquier otro material.
- **Diferencia –** la diferencia entre las dos medias muestrales

$$\hat{\Delta}_{j_1 j_2} = \bar{Y}_{j_1} - \bar{Y}_{j_2} \tag{3}$$

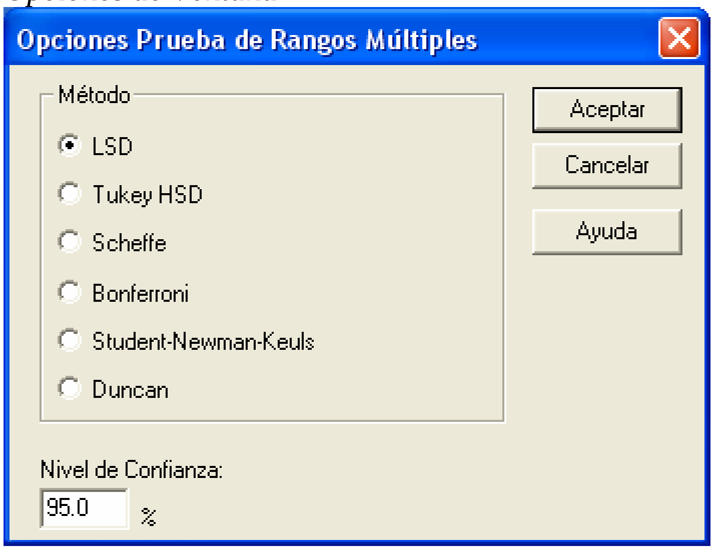
- **Límites –** una estimación por intervalo de tal diferencia, empleando el procedimiento de comparaciones múltiples actualmente seleccionado:

$$\hat{\Delta}_{j_1 j_2} \pm M \sqrt{CM_{intra}\left(\frac{1}{n_{j_1}} + \frac{1}{n_{j_2}}\right)} \tag{4}$$

donde M es la constante que depende del proceso elegido.

- **Sig**. – Se pone un asterisco junto a cualquier diferencia que estadísticamente sea significativamente diferente de 0 al nivel de significancia actualmente seleccionado, i.e., cualquier intervalo que no contenga al 0.

Opciones de Ventana



- **Método**: el método empleado para hacer las comparaciones múltiples.

- **Nivel de Confianza**: el nivel de confianza empleado por el procedimiento de comparaciones múltiples seleccionado.

Los métodos disponibles son:

- **LSD** – forma un intervalo de confianza para cada par de medias al nivel de confianza elegido usando:

$$M = t_{\alpha/2, n-q} \tag{5}$$

donde *t* representa el valor de la distribución t de Student con *n - q* grados de libertas que deja un área de $\alpha/2$ en la cola superior de la curva. Este procedimiento de debe a Fisher y es llamado el procedimiento de la *Mínima Diferencia Significativa* (*Least Significant Difference*), ya que la magnitud de los límites indica la diferencia mínima entre dos medias cualesquiera que puede ser declarada que representa una diferencia estadísticamente significativa. Sólo debe emplearse cuando la prueba de F en la tabla de ANOVA indica diferencias significativas entre las medias muestrales. La probabilidad α de cometer un error Tipo I aplica a cada par de medias por separado. Si se hace más de una comparación, la probabilidad global de declarar al menos un par de medias significativamente diferentes cuando no lo son puede ser considerablemente mayor que α.

- **Tukey HSD** – amplía los intervalos para permitir múltiples comparaciones entre todos los pares de medias, usando:

$$M = T_{\alpha/2,q,n-q} \tag{6}$$

que usa la T de Tukey en vez de la t de Student. La T de Tukey es igual a $\left(1/\sqrt{2}\right)$ veces la distribución de rangos estudentizados, la cuál se tabula en libros tales como en Neter et al. (1996). Tukey llamó a su procedimiento el procedimiento de la *Diferencia Honestamente Significativa* (*Honestly Significant Difference*) ya que la tasa de error por experimento en α. Si todas las media son iguales, la probabilidad de declarar *cualquiera* de los pares significativamente diferente en todo el experimento es igual a α. El procedimiento de Tukey es más conservador que el procedimiento de la LSD de Fisher, ya que hace más difícil declarar que cualquier par específico de medias sea significativamente diferente.

- **Scheffe –** diseñado para permitir la estimación de todos los posibles contrastes entre las medias muestrales (no solo comparaciones por pares). Emplea un múltiplo relacionado con la distribución F:

$$M = \sqrt{(q-1)F_{\alpha,q-1,n-q}} \tag{7}$$

En este caso, este procedimiento es muy probable que sea muy conservador, ya que sólo se están estimando pares.

- **Bonferroni –** diseñado para permitir la estimación de cualquier número de contrastes previamente seleccionado. En este caso, usa un múltiplo igual a

$$M = t_{\alpha/(q(q-1)),n-q} \tag{8}$$

ya que se están estimando *q(q-1)/2* diferencias por pares. Estos límites generalmente son más amplios que los límites de Tukey cuando se hacen todas las comparaciones por pares.

- **Student-Newman-Keuls –** A diferencia de todos los métodos anteriores, este procedimiento no crea intervalos para las diferencias por pares. En su lugar, ordena las medias en orden creciente y luego comienza a separarlas en grupos de acuerdo con valores de la distribución de rangos estudentizados. Al final, las medias se separan en grupos homogéneos dentro de los cuales no hay diferencias significativas.

- **Duncan** - similar al procedimiento de Student-Newman-Keuls, excepto que usa un valor crítico diferente de la distribución de rangos estudentizados cuando define los grupos homogéneos. Un tratamiento detallado de los procedimientos de Duncan y Student-Newman-Keuls se encuentra en Milliken y Johnson (1992).

La elección entre el procedimiento LSD y un procedimiento de comparaciones múltiples tal como la HSD de Tukey deberá depender del costo relativo de cometer el error Tipo I (declarar diferentes un par de medias cuando realmente no lo son) versus el costo de cometer el error Tipo II (no declarar diferentes un par de medias cuando realmente sí lo son). En etapas tempranas de una investigación, uno puede no querer ser tan conservador como cuando se están haciendo verificaciones finales.

Tabla de Medias

Esta tabla presenta la media de cada nivel junto con un intervalo de incertidumbre:

Tabla de Medias para strength por material con intervalos de confianza del 95.0%

			Error Est.		
material	*Casos*	*Media*	*(s agrupada)*	*Límite Inferior*	*Límite Superior*
A	8	43.125	3.46386	38.1078	48.1422
B	8	31.875	3.46386	26.8578	36.8922
C	8	22.625	3.46386	17.6078	27.6422
D	8	20.5	3.46386	15.4828	25.5172
Total	32	29.5313			

El tipo de intervalo exhibido depende de las *Opciones de Ventana*.

Opciones de Ventana

- **Intervalos**: el método empleado para construir los intervalos.
- **Nivel de Confianza:** el nivel de confianza asociado con cada intervalo.

Los tipos de intervalos que se pueden seleccionar son:

- **Ninguno** – ningún intervalo es exhibido.
- **Errores Estándar (S Agrupada)** – muestra los errores estándar empleando la desviación estándar agrupada intra muestras:

$$\bar{Y}_j \pm \sqrt{\frac{CM_{intra}}{n_j}} \tag{9}$$

- **Errores Estándar (S individual)** - muestra los errores estándar usando la desviación estándar de cada muestra por separado:

$$\bar{Y}_j \pm \sqrt{\frac{s_j^{\ 2}}{n_j}} \tag{10}$$

- **Intervalos de Confianza (S Agrupada) –** presenta intervalos de confianza para las medias de los grupos empleando la desviación estándar agrupada intra grupos:

$$\bar{Y}_j \pm t_{\alpha/2,n-q} \sqrt{\frac{CM_{intra}}{n_j}} \tag{11}$$

- **Intervalos de confianza (s individual)** - presenta intervalos de confianza para las medias muestrales usando la desviación estándar de cada grupo por separado:

$$\bar{Y}_j \pm t_{\alpha/2,n_j-1} \sqrt{\frac{s_j^{\ 2}}{n_j}} \tag{12}$$

- **Intervalos LSD** - diseñados para comparar cualquier par de medias con el nivel de confianza establecido. Los intervalos están dados por

$$\bar{Y}_j \pm \frac{\sqrt{2}M}{2} \sqrt{\frac{CM_{intra}}{n_j}} \tag{13}$$

donde M se define como en las *Pruebas de Rangos Múltiples*. Esta fórmula aplica también a las tres secciones siguientes.

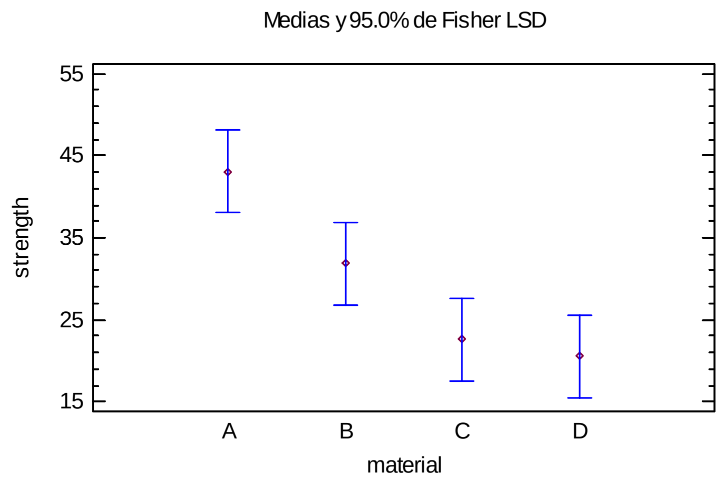
- **Intervalos Tukey HSD**- diseñados para comparar todos los pares de medias. El nivel de confianza establecido aplica a la familia completa de comparaciones por pares.

- **Intervalos Scheffe** - diseñado para comparar todos los contrastes. Generalmente no relevante aquí.

- **Intervalos Bonferroni** - diseñados para comparar un selecto número de contrastes. Los intervalos de Tukey generalmente son más estrechos.

Gráfico de Medias

Las medias de los niveles pueden ser graficadas junto con los intervalos de incertidumbre:



Los tipos de intervalos que se pueden emplear son los mismos que para la *Tabla de Medias* anterior.

Siempre que todos los tamaños muestrales sean los mismos (o casi iguales), el analista puede determinar qué media es significativamente diferente de cuáles otras empleando el procedimiento LSD, de Tukey, de Scheffe, o de Bonferroni simplemente viendo si un par de intervalos se traslapan o no en dirección vertical. Un par de intervalos que no se traslapan indican una diferencia estadísticamente significativa entre las medias al nivel de confianza elegido. En este caso, advierta que el intervalo para el material A no se traslapa con los intervalos de los otros materiales.

Contraste de Varianza

Uno de los supuestos subyacentes al análisis de varianza es que las varianzas de las poblaciones de donde provienen las muestras son la misma. La ventana *Contraste de Varianza* realiza cualquiera de varias pruebas para verificar este supuesto:

Contraste de Varianza

	Prueba	*Valor-P*
C de Cochran	0.373145	0.573459

El StatAdvisor
El estadístico mostrado en esta tabla evalúa la hipótesis de que la desviación estándar de strength dentro de cada uno de los 4 niveles de material es la misma. De particular interés es el valor-P. Puesto que el valor-P es mayor o igual que 0.05, no existe una diferencia estadísticamente significativa entre las desviaciones estándar, con un nivel del 95.0% de confianza.

Las hipótesis a contrastar en las prueba son:

Hipótesis Nula: todas las σ_j son iguales

Hipótesis Alt.: no todas las σ_j son iguales

Las cuatro pruebas son:

1. *Prueba de Cochran:* compara la máxima varianza intra-muestra con la varianza intra-muestra promedio. Un valor de P menor de 0.05 indica una diferencia significativa entre desviaciones estándar intra-muestra a un nivel de significancia del 5%. La prueba es apropiada sólo si todos los tamaños de grupo son iguales.

2. *Prueba de Bartlett:* compara un promedio ponderado de las varianzas intra-muestra con su media geométrica. Un valor de P menor de 0.05 indica una diferencia significativa entre desviaciones estándar intra-muestra a un nivel de significancia del 5%. La prueba es apropiada tanto para tamaños de grupo iguales como diferentes.

3. *Prueba de Hartley:* calcula la razón entre la mayor varianza muestral y la menor. Esta estadística debe ser comparada con una tabla de valores críticos, tal como la que se encuentra en Neter et al. (1996). Para 6 muestras y 62 grados de libertad para el error experimental, *H* tendría que exceder aproximadamente 2.1 para ser estadísticamente significativa a un nivel de significancia del 5%. Nota: esta prueba sólo es apropiada si el número de observaciones dentro de cada nivel de es el mismo.

4. *Pueba de Levene:* realiza un análisis de varianza simple sobre las variables

$$Z_{ij} = \left| y_{ij} - \bar{y}_j \right| \qquad (14)$$

   La estadística tabulada es la estadística F de la tabla de ANOVA.

Usando la prueba de Levene para los artilugios, no hay razón para rechazar el supuesto de que las desviaciones estándar son la misma para todos los materiales, dado que el valor de P es mayor que 0.05. Cualesquiera aparentes diferencias entre las desviaciones estándar muestrales no son estadísticamente significativas al nivel de significancia del 5%.

Gráficos de Residuos

Al igual que con todos los modelos estadísticos, es buena práctica examinar los residuos. En un análisis de varianza simple, los residuos están definidos por:

$$e_{ij} = y_{ij} - \bar{y}_j \qquad (15)$$

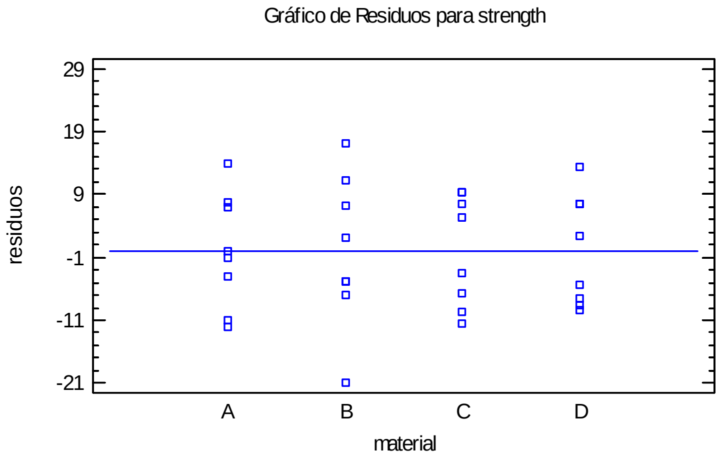
i.e., los residuos son las diferencia entre los valores de los datos observados y sus respectivas medias de nivel.

El procedimiento *Comparación de Varias Medias* crea 3 gráficos de residuos:

1. versus nivel del factor.
2. versus valor predicho.
3. versus número de fila.

Residuos versus Nivel del Factor

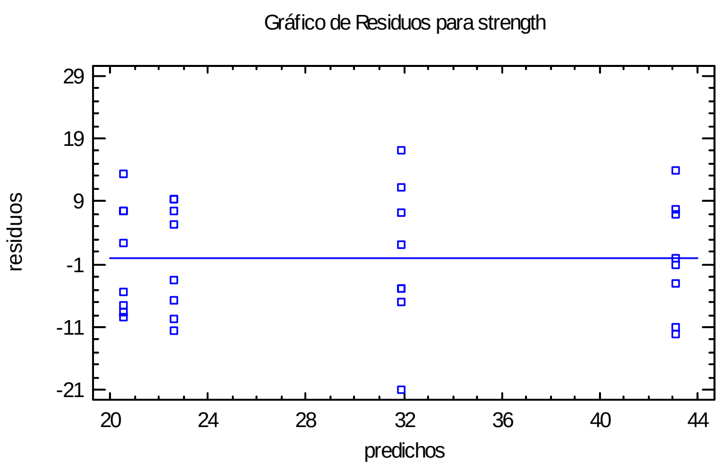
Este gráfico es útil para visualizar cualquier diferencia en variabilidad entre los niveles.



El residuo promedio en cada nivel es igual a 0.

Residuos versus Predichos

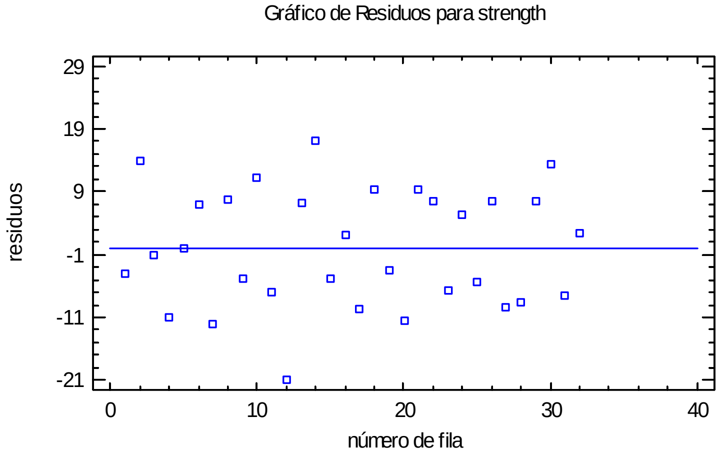
Este gráfico es útil para detectar cualquier heteroscedasticidad en los datos.



La heteroscedasticidad se presenta cuando la variabilidad de los datos cambia conforme cambia la media, y podría necesitar transformar los datos antes de realizar el ANOVA. Generalmente se evidencia por un patrón en forma de embudo en el gráfico de residuales.

Residuos versus Observación

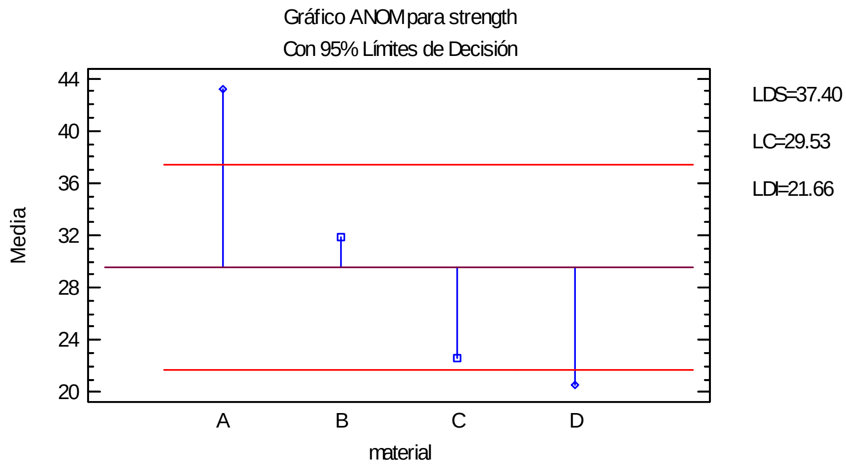
Este gráfico muestra los residuales versus el número de hilera en la hoja de datos:



Si los datos están arreglados en orden cronológico, cualquier patrón en los datos podría indicar una influencia externa. Ningún patrón tal es evidente en el gráfico anterior.

Gráfico Análisis de Medias (ANOM)

Si el número de muestras está entre 3 y 20, un enfoque ligeramente diferente a la comparación de medias de los nivel es presentado en el *Gráfico Análisis de Medias* o *ANOM* (del inglés ***An**alysis **of M**eans*):



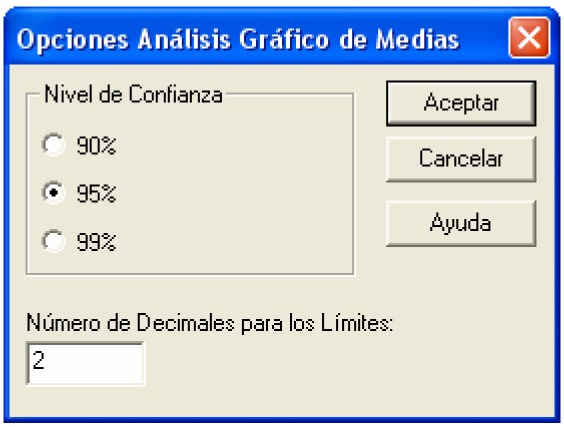
Este es similar a un gráfico de control estándar, donde se grafica cada media junto con una línea central y límites de decisión superior e inferior. La línea central se localiza en la gran media de todos las observaciones $\overline{Y}$. Los límites de decisión se localizan en

$$\overline{Y} \pm h_{n-q,1-\alpha} \sqrt{\frac{CM_{intra}}{n_j}\left(\frac{q-1}{q}\right)} \qquad (16)$$

donde *h* es un valor crítico obtenido de una tabla de la distribución t multivariada. El gráfico prueba la hipótesis nula de que todas las medias de los niveles son iguales a la gran media. Cualquiera media que caiga fuera de los límites de decisión indica que el nivel correspondiente difiere significativamente de esa media general.

La ventaja del gráfico ANOM es que muestra de un vistazo qué medias son significativamente diferentes del promedio de todos los niveles. Y lo hace usando un tipo de gráfico con el cual están muy familiarizados muchos ingenieros y operarios. Es fácil ver a partir de la gráfica anterior que el material A tiene una fuerza de ruptura media significativamente mayor que el promedio, mientras que el material D tiene una fuerza de ruptura media significativamente menor. El procedimiento es exacto si todos los tamaños muestrales son iguales y es aproximado si los tamaños de muestra no difieren mucho.

Opciones de Ventana



- **Nivel de Confianza:** nivel empleado para colocar los niveles de decisión.

- **Posiciones Decimales para los Límites:** número de cifras decimales mostradas en los límites de decisión.

Pruebas de Kruskal-Wallis y Friedman

Una alternativa al análisis de varianza estándar que compara las *medianas* de los niveles en vez de las medias es la *Prueba de Kruskal-Wallis*. Esta prueba es mucho menos sensible a la presencia de valores atípicos que un ANOVA simple estándar y debe usarse siempre que el supuesto de normalidad dentro de los niveles no es razonable. Prueba las hipótesis:

Hipótesis Nula: todas las medianas de los niveles son iguales

Hipótesis Alt.: no todas las medianas de los niveles son iguales

La prueba se realiza:

1. Ordenando todos los *n* valores del menor al mayor y asignándoles rango, 1 al menor y *n* al mayor. Si cualesquiera observaciones son exactamente iguales, entonces las observaciones empatadas se les da un rango igual al promedio de las posiciones en las que se encuentran los empates.

2. Calculando el rango promedio de las observaciones en cada nivel $\bar{R}_j$.

3. Calculando un estadístico de prueba para comparar las diferencias entre los rangos promedio.

4. Calculando un valor de P para probar las hipótesis.

La salida se muestra a continuación:

Prueba de Kruskal-Wallis para strength por material

material	*Tamaño Muestra*	*Rango Promedio*
A	8	26.125
B	8	17.6875
C	8	12.0625
D	8	10.125

Estadístico = 14.0787   Valor-P = 0.00279989

Valores pequeños de P (menores de 0.05 si se trabaja al nivel de significancia del 5%) indican que hay diferencias significativas entre las medianas de los niveles, como en el ejemplo anterior.

Prueba de Mood para la Mediana

La *Prueba de Mood para la Mediana* es otro método para determinar si las medianas de todos los q materiales son iguales o no. Es menos sensible a valores atípicos que la prueba de Kruskal-Wallis, pero también es menos potente cuando los datos provienen de distribuciones tales como la normal. La salida se muestra a continuación.

Prueba de la Mediana de Mood para strength por material
Total n = 32
Gran mediana = 29.0

material	*Tamaño de Muestra*	*n<=*	*n>*	*Mediana*	*LC inferior 95.0%*	*LC superior 95.0%*
A	8	0	8	42.5		
B	8	4	4	30.5		
C	8	5	3	23.5		
D	8	7	1	19.0		

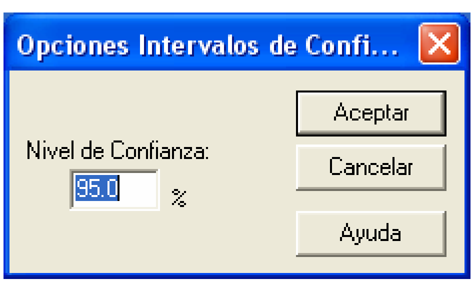
Estadístico = 13.0   Valor-P = 0.00463639

En la parte superior de la tabla se presentan el número total de observaciones n y la mediana general. Para cada nivel de X, la tabla muestra

1. *Tamaño de Muestra*: el número de observaciones n_j.
2. *n*<=: de las observaciones en ese nivel, cuántas son menores o iguales que la mediana general.
3. *n*>: de las observaciones en ese nivel, cuántas son mayores que la mediana general.
4. *Mediana*: la mediana del nivel.
5. *LC*: los límites superior e inferior de confianza para la mediana de la población de la cual provienen los datos. Si los tamaños de muestra son demasiado pequeños, como en el presente ejemplo, puede no ser posible obtener límites de confianza.

En la parte inferior de la pantalla se presentan un estadístico de prueba y el valor de P. Tratando las columnas *n*<= y *n*> como columnas de una tabla de contingencia de dos entradas, se calcula un estadístico de prueba chi-cuadrada. Valores pequeños de P (menores de 0.05 si se trabaja al nivel de significancia del 5%) llevan a la conclusión de que las medianas no son todas iguales, como en el presente ejemplo.

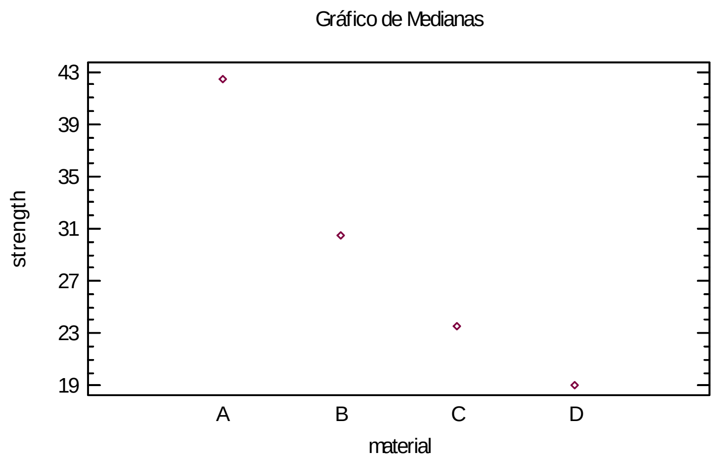
Opciones de Ventana



- **Nivel de Confianza:** nivel empleado para los límites de confianza.

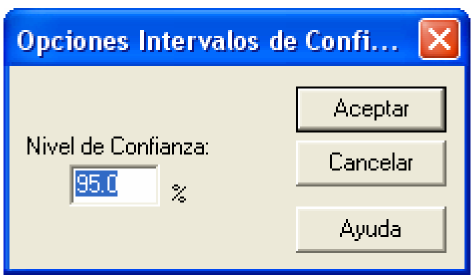
Gráfico de Medianas

El *Gráfico de Medianas* exhibe los intervalos de confianza para las medianas presentados por la ventana de la *Prueba de Mood para la Mediana*.



En el presente caso, los tamaños de muestra son demasiado pequeños para permitir la estimación de intervalos de confianza.

Opciones de Ventana



- **Nivel de Confianza:** nivel empleado para los límites de confianza.

Salvar Resultados

Se pueden salvar los siguientes resultados en la hoja de datos::

1. *Casos por Nivel* – los q tamaños de muestra n_j.
2. *Medias de Niveles* – las q medias de nivel.
3. *Medianas de Niveles* – las q medianas de nivel.
4. *Desviaciones Estándar de Niveles* – las q desviaciones estándar nivel.
5. *Errores Estándar de Niveles* – los errores estándar de cada media de nivel, $\sqrt{CM_{intra} / n_j}$.
6. *Etiquetas de Niveles* – q etiquetas, una para cada nivel.
7. *Indicadores de Nivel* – n indicadores del nivel, identificando cada residual.
8. *Residuales* – los n residuales.
9. *Rangos de Niveles* – los q rangos de nivel.

Cálculos

Análisis de Varianza

Fuente	Suma de Cuadrados	G.L.	Cuadrado Medio	Razón F
Entre grupos	$SC_{entre} = \sum_{j=1}^{q} n_j \left(\bar{Y}_j - \bar{Y}\right)^2$	$gl_{entre} = q - 1$	$CM_{entre} = \frac{SC_{entre}}{gl_{entre}}$	$F = \frac{CM_{entre}}{CM_{dentro}}$
Intra grupos	$SC_{intra} = \sum_{j=1}^{q} \sum_{i=1}^{n_j} \left(Y_{ij} - \bar{Y}_j\right)^2$	$gl_{intra} = \sum_{j=1}^{q} \left(n_j - 1\right)$	$CM_{intra} = \frac{SC_{intra}}{gl_{intra}}$	
Total	$SC_{total} = \sum_{j=1}^{q} \sum_{i=1}^{n_j} \left(Y_{ij} - \bar{Y}\right)^2$	n-1		

Prueba de Cochran

El estadístico presentado es calculado por

$$A = \frac{\max\left(s_j^2\right)}{\sum_{j=1}^{q} s_j^2} \tag{17}$$

Para probar significancia estadística,

$$C = \left(q - 1\right)\left(\frac{A}{1 - A}\right) \tag{18}$$

se compara con una distribución F con *(n/q - 1)* y *(n/q - 1)(q - 1)* grados de libertad.

Prueba de Bartlett

El estadístico presentado es calculado por

$$B = \frac{1}{C}\left[\left(gle\right)\ln\left(CME\right) - \sum_{j=1}^{q} (n_j - 1)\ln\left(s_j^2\right)\right] \tag{19}$$

donde

$$C = 1 + \frac{1}{3\left(q - 1\right)}\left[\left(\sum_{j=1}^{q} \left(n_j - 1\right)^{-1}\right) - \frac{1}{gle}\right] \tag{20}$$

$$CME = \frac{1}{gle}\sum_{j=1}^{q}\left(n_j - 1\right)s_j^2 \tag{21}$$

$$gle = \sum_{j=1}^{q}\left(n_j - 1\right) \tag{22}$$

B se compara con la distribución chi-cuadrada con *(q-1)* grados de libertad.

Prueba de Hartley

$$H = \frac{\max\left(s_j^2\right)}{\min\left(s_j^2\right)} \tag{23}$$

Límites de Confianza para la Mediana

Los límites presentados son una interpolación no lineal de los intervalos de confianza al nivel más cercano por arriba y por debajo del nivel solicitado. Después de ordenar las observaciones, el intervalo que se extiende del *d*-ésima observación más pequeña a la *d*-ésima observación más grande forma un intervalo de confianza para la mediana con nivel de confianza 1 – 2 P_B(d-1), donde P_B representa la distribución binomial acumulada con $p = 0.5$ y $n = n_j$.