

Utilidad de la hoja de cálculo Excel en el análisis de datos cuantitativos

Esperanza Bausela Herreras¹

esperanzabh@yahoo.es

Resumen: *En este artículo presentamos la hoja de cálculo Excel, que permite realizar distintos tipos de trabajos: Hojas de Cálculo, Bases de Datos y Gráficos y, también, Macros. Es una aplicación integrada en el entorno Windows. A estas capacidades se suma la posibilidad de utilizarla como base de datos. En este artículo exponemos la secuencia de diferentes tipos de análisis de tipo paramétrico. Previamente hemos contrastado la equivalencia de técnicas paramétricas y no paramétricas, para aquellas situaciones en las que no se cumplan los supuestos de normalidad.*

Palabras Clave: *Hoja de cálculo, Excel, análisis paramétricos, análisis no paramétricos.*

Summary: *In this paper we show the Excel paper, Excel, with which you can make several activities, such as: dates analysis, graphs, and, also macro. This is integrate in the Windows context. Also, it is add one new application as database. In this paper we analysing several parametrics analysis. Before, we have contrasted parametrics techniques with no-parametrics techniques, which can or must be used with the dates don't make the normal supposition.*

Words Keys: *Calculation paper, Excel, parametrics techniques and no-parametrics techniques*

¹ Doctora en Psicología y Ciencias de la Educación por la Universidad de León.

1. Introducción

La hoja de cálculo Excel de Microsoft es una aplicación integrada en el entorno Windows que permite la realización de cálculos, así como la representación de estos valores de forma gráfica. A estas capacidades se suma la posibilidad de utilizarla como base de datos.

Ofrece una gran ayuda, siendo relativamente fácil la comprensión de las operaciones básicas, y contando, además, con funciones estadísticas que facilitan los distintos cálculos. Dispone, además de un abundante repertorio de figuras y proporciona una hoja de trabajo estándar para registro de datos que permite una rápida transferencia a otros programas estadísticos con una dificultad mínima.

2. Comprobación de supuestos paramétricos: previos al análisis de datos

La violación de los supuestos paramétricos no invalida el análisis, ya que estas pruebas suelen ser lo suficientemente robustas para no verse seriamente afectas por ligeras violaciones de los supuestos paramétricos. Según Bisquerra (1989), existe clara evidencia empírica y matemática, de que las pruebas con una sola variable dependiente (por ejemplo ANOVA) son altamente robustas bajo la violación de normalidad y homoscedasticidad. La principal, excepción, según este profesor se halla cuando las muestras son muy pequeñas (menores de 10) y desiguales. La situación se complica cuando pasamos a los métodos multivariantes. Se observa, siguiendo al mismo autor, un cierto vacío en cuanto a recomendaciones concretas aceptadas de forma generalizada. La tendencia actual está en considerar que en muestras grandes ($n > 30$) los análisis multivariantes sin lo suficientemente robustos como para ser insensibles a ligeras desviaciones de los supuestos paramétricos, principalmente el normalidad multivariante y de la homoscedasticidad. Esto nos refuerza en la posibilidad de aplicación de los análisis multivariantes.

Siguiendo el procedimiento de varios estudios (Rincón, 1987; Goikoetxea, Aznar, Cortajarena, Ponte, Ugarte y Uriarte, 2001), en diversas situaciones se ha recurrido a la contrastación con la homóloga no paramétrica. Obteniéndose para el nivel de significación fijado ($p < .05$) los mismos resultados en comparación en todas las variables que con las técnicas paramétricas, siguiendo el diagrama de la tabla I.

Tabla I. Pruebas paramétricas y no paramétricas (elaboración propia)

TÉCNICAS PARAMÉTRICAS	TÉCNICAS NO PARAMÉTRICAS
I. Técnicas para el contraste de hipótesis entre dos medias	
Muestras Independientes	
Prueba t para muestras independientes	Prueba Mann – Whitney (U- test) Prueba de Kolmogorov – Smirnov Prueba de Rachas
Muestras relacionadas	
Prueba t de muestras relacionadas	Prueba de Wilcoxon
II. Técnicas para el contraste de hipótesis entre más de dos medias	
Análisis de la varianza	Prueba de Kruskal - Wallis
III. Coeficiente de correlación y regresión	
Correlación y regresión lineal simple	Spearman Correlación Biserial Correlación Tretacórica Coeficiente C

3. Desarrollo de la Hoja de cálculo Excel

0. Estructura general

0.1. Iniciar el programa → Inicio → Programas → Seleccionamos el programa Excel.

0.2. Pantalla → Excel trabaja con hojas de cálculo que están integradas en libros de trabajo. La pantalla esta compuesta:

- **Barra de Título.**
- **Barra de menú,** desde la que aparecen los menú con todas las órdenes fundamentales.
- **Barra Estándar** nos permite realizar rápidamente algunas de las operaciones más

comunes como abrir y guardar ficheros, imprimirlos o copiar y pegar lo copiado en otro lado.

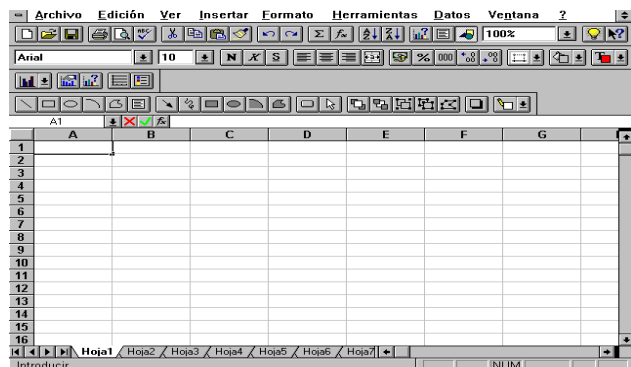


Imagen 1 Ventana de la hoja de cálculo de Excell

1. Empezar a trabajar con Excel

1.1. Hacer un nuevo documento

⇒ Archivo → Nuevo. También se puede crear un nuevo libro pulsando el botón de página en blanco que esta dentro de la barra estándar.

1.2. Guardar un documento

⇒ Es aconsejable guardar el documento por primera vez antes de empezar. → Archivo|Guardar como. La opción Archivo|Guardar, se utiliza para almacenar en el documento los cambios realizados desde la última vez que se guardó.

1.3. Abrir un documento ya creado

⇒ Archivo|Abrir → localizar el documento → Abrir. También puede realizar esta operación haciendo clic sobre este botón de archivo de la barra estándar.

1.4. Introducción de datos en las celdas

Un dato es una celda se introduce seleccionando con el ratón o con el teclado donde se desea introducir el valor → Teclear el dato → Enter.

1.5. Realizar operaciones numéricas

⇒ **Sumar** → Seleccionamos una celda y comenzamos escribiendo el signo igual "=", para indicar al programa que se trata de una operación matemática → Sumamos utilizando el

signo + → Pulsamos Enter o confirmar. No olvide pulsar Enter o confirmar antes de cambiar de celda. También podemos marcar los números que deseamos sumar → Sumatorio (Barra Estándar).

⇒ **Multiplicar** → Podemos multiplicar utilizando el signo *. Por ejemplo para multiplicar dos por dos escribimos en una celda =2*2 → pulsamos Enter en el teclado.

⇒ **Dividir** → Dividimos con el símbolo /. Por ejemplo si queremos dividir 4 entre 2 escribimos en una celda =4/2 → pulsamos Enter.

1.6. Estadísticos

En el menú *Herramientas* → Elegir *Análisis de datos* → Seleccionar la que se desee calcular, algunas opciones son:

- Análisis de varianza de un factor
- Análisis de varianza de dos factores con varias muestras por grupo
- Análisis de varianza de dos factores con una sola muestra por grupo
- Coeficiente de correlación
- Covarianza
- Estadística descriptiva
- Suavización exponencial
- Prueba t para varianzas de dos muestras
- Análisis de Fourier
- Histograma.

II.1. Características descriptivas de la muestra

2. Estadísticos descriptivos

→ *Rango de entrada*: Introducir la referencia de celda correspondiente al rango de datos que desee analizar.

→ *Agrupado por*: Filas o Columnas para indicar si los datos del rango de entrada están organizados en filas o en columnas.

→ *Rótulos en la primera fila y rótulos en la primera columna*: Si la primera fila del rango de entrada contiene rótulos, activar la casilla de verificación Rótulos en la primera fila. Si los rótulos están en la primera columna del rango de entrada, activar la casilla de verificación Rótulos en la primera columna. Esta casilla de

verificación estará desactivada si el rango de entrada carece de rótulos.

→ *Opciones de salida: Rango de salida:* Esta herramienta genera dos columnas de información por cada conjunto de datos. La columna de la izquierda contiene los rótulos de estadística y la columna de la derecha contiene las estadísticas. Otras salidas: en una hoja nueva o en un libro nuevo.

→ *Resumen de estadísticas:* Nos muestra una tabla de resultados por cada una de las siguientes variables estadísticas:

- Media
- Error típico (de la media)
- Mediana
- Moda
- Desviación estándar
- Varianza
- Curtosis
- Coeficiente de asimetría
- Rango
- Mínimo
- Máximo
- Suma
- Cuenta
- Mayor(#)
- Menor (#)
- Nivel de confianza.

→ *K-ésimo mayor:* Activar esta casilla si se desea incluir una fila correspondiente al valor k-ésimo menor de cada rango de datos en la tabla de resultados. Si se escribe 1, esta fila contendrá el máximo conjunto de datos.

→ *K-ésimo menor:* Activar esta casilla si se desea incluir una fila correspondiente al valor k-ésimo mayor de cada rango de datos en la tabla de resultados. Si se escribe 1, esta fila contendrá el mínimo conjunto de datos.

3. Tablas de contingencia

⇒ *Datos* → *Informe de tablas y gráficos dinámicos* → ¿Dónde están los datos que desea analizar? Seleccionar en *Lista o base de datos de Microsoft Excel* → ¿Qué tipo de informe desea crear? Seleccionar *Tabla dinámica* →

Siguiente → ¿Dónde están los datos que desea usar? Seleccionar el rango → *Siguiente* → ¿Dónde desea situar el informe de la tabla dinámica? Seleccionar *Hoja de cálculo existente* → *Diseño:* Construir el informe de tabla dinámica arrastrando los botones de campo de la derecha sobre el diagrama de la izquierda → *Aceptar* → *Finalizar*.

⇒ *Prueba Estadísticas* → *PRUEBA.CHI* → *Rango actual* → *Rango esperado* → *Aceptar*.

4. Histograma

⇒ *Herramientas* → *Análisis de datos* → *Histograma* → *Rango de entrada* → *Rango de clases:* introducir el rango de celdas que definen los extremos de los intervalos de clase. Estos valores deberán estar en orden ascendente. Todos los valores por debajo del primer valor de clase formarán la clase inferior, y los valores por encima del último valor de clase formarán la clase superior. Si se omite el rango de clase, se creará un conjunto de clases distribuidas igualmente entre los valores mínimo y máximo de los datos → *Rótulos:* activar si la primera fila y la primera columna del rango de entrada contienen rangos → *Opciones de salida:* rango de salida, en una hoja nueva o en un libro nuevo → *Pareto* (histograma ordenado), permite presentar los datos en orden de frecuencia descendente en la tabla de resultados → *Porcentaje acumulado*, genera una columna de tabla de resultados con porcentajes acumulados e incluye una línea de porcentaje acumulado en el gráfico de histograma, si se desactiva omite los porcentajes acumulados → *Crear gráfico:* genera un gráfico de histograma incrustado con la tabla de resultados.

5. Frecuencias

⇒ *Estadísticas* → *FRECUENCIAS* → *Datos* → *Grupos* → *Aceptar*.

6. Gráficos estadísticos

⇒ *Asistente para gráficos* → tipo de gráfico: elegimos el tipo de gráfico entre los 14 que nos presenta. *Subtipo de gráfico:* seleccionamos el subtipo entre los que muestra (diagramas de barras y columnas- diagramas de líneas y dispersión- diagramas de sectores y áreas) y hacemos *clic* en el botón *Siguiente* → *datos de origen:* en la ficha *Rango de Datos*,

seleccionamos con el ratón el rango en el que están los datos. En la ficha *Serie* en Rótulos en el eje de categorías (X): elegimos también con el ratón el rango en el que están los rótulos y Siguiente → *opciones de gráfico*: rellenamos la ficha *Títulos*. En Eje de categorías (X): cumplimentamos el campo y hacemos lo mismo con el campo Eje de valores (Y). En la ficha *Leyenda*, desactivamos la opción *Mostrar leyenda* y Siguiente → *Ubicación del gráfico*: tenemos dos opciones: En una hoja nueva y Como objeto en: (podemos elegir la hoja, normalmente será la misma hoja) → Finalizar.

II.2. Procedimiento de análisis

7. Análisis de varianza de un factor

⇒ *Herramientas* → *Análisis de datos* → *Análisis de varianza de un factor* → *Rango de entrada* → *Agrupado Columnas* → *Rótulos en la primera fila* → *Alfa*: Introducir el nivel con el que desee evaluar los valores críticos de la función estadística F. El nivel alfa es un nivel de importancia relacionado con la probabilidad de que haya un error de tipo I (rechazar una hipótesis verdadera) (por ejemplo, 0,05) → *Opciones de salida*: *Rango de salida, en una hoja nueva o en un libro nuevo*.

8. Análisis de varianza de dos factores con una sola muestra por grupo

⇒ *Herramientas* → *Análisis de datos* → *Análisis de varianza de dos factores con una sola muestra por grupo* → *Rango de entrada* → *Rótulos*: Esta casilla de verificación estará desactivada si el rango de entrada carece de rótulos → *Alfa*: Introducir el nivel con el que desee evaluar los valores críticos de la función estadística F (valor 0,05) → *Opciones de salida*: *Rango de salida, en una hoja nueva o en un libro nuevo*.

9. Coeficiente de correlación

⇒ *Herramientas* → *Análisis de datos* → *Coeficiente de correlación* → *Rango de entrada* → *Agrupada por columnas* → *Rótulos en la primera fila* → *Opciones de salida*: *Rango de salida, en una hoja nueva o en un libro nuevo*.

10. Covarianza

⇒ *Herramientas* → *Análisis de datos* → *Rango de entrada* → *Agrupado por Filas o Columnas* →

Rótulos en la primera fila y rótulos en la primera columna → *Opciones de salida*: *Rango de salida, en una hoja nueva o en un libro nuevo*.

11. Prueba F para varianzas de dos muestras

⇒ *Herramientas* → *Análisis de datos* → *prueba F para varianzas de dos muestras* → *Rango para la variable 1*: Introducir la referencia correspondiente a la primera columna o fila de datos que se desean analizar → *Rango para la variable 2*: Introducir la referencia correspondiente a la segunda columna o fila de datos que desea analizar → *Rótulos*: Activar esta casilla si la primera fila y la primera columna del rango de entrada contienen rótulos → *Alfa*: Introducir el nivel de confianza para la prueba. Este valor deberá estar comprendido entre 0 y 1. El nivel Alfa es un nivel de importancia relacionado con la probabilidad de que haya un error de tipo I (rechazar una hipótesis verdadera) (valor, por ejemplo, 0,05) → *Opciones de salida*: *Rango de salida, en una hoja nueva o en un libro nuevo*.

12. Prueba t para dos muestras suponiendo varianzas iguales

⇒ *Herramientas* → *Análisis de datos* → *prueba F para varianzas de dos muestras* → *Rango para la variable 1* → *Rango para la variable 2* → *Rótulos* → Diferencia hipotética entre las medias (por ejemplo 0) (indica que, según la hipótesis, las medias de las muestras serán iguales) → *Alfa* (valor, por ejemplo, 0,05) → *Opciones de salida*: *Rango de salida, en una hoja nueva o en un libro nuevo*.

13. Prueba t para dos muestras suponiendo varianzas desiguales

⇒ *Herramientas* → *Análisis de datos* → *prueba F para varianzas de dos muestras* → *Rango para la variable 1* → *Rango para la variable 2* → *Rótulos* → Diferencia hipotética entre las medias (por ejemplo 0) (indica que, según la hipótesis, las medias de las muestras serán iguales) → *Alfa* (valor, por ejemplo, 0,05) → *Opciones de salida*: *Rango de salida, en una hoja nueva o en un libro nuevo*.

4. Bibliografía de Consulta

Bisquerra, R. (1989). *Introducción conceptual al análisis multivariable. Un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD*. Barcelona: PPU.

Goikoetxea, E., Aznar, S., Cortajarena, C., Ponte, J., Ugarte, K. y Urarte, J. (2001). El WAIS – III en pacientes esquizofrénicos. Hallazgos sobre su validez. *Revista de Psicología General y Aplicada*, 54 (2), 245 – 257.

Pérez, C. (2002). *Estadística aplicada a través del Excel*. Madrid: Prentice Hall.

Rincón, D. (1987). *Análisis de componentes motivacionales en el desarrollo de la personalidad ética*. Tesis Doctoral: Universidad Autónoma de Barcelona, Barcelona.