

HERRAMIENTAS PARA LA INVESTIGACIÓN CIENTÍFICA**Microsoft Excel: una herramienta para la investigación.****Microsoft Excel: A tool for research.**

Lic. Luis Orlando Pérez González.

Licenciado en Matemática. Profesor Asistente. Facultad de Ciencias Médicas " Dr. Raúl Dorticós Torrado ". Cienfuegos.

RESUMEN

Con alguna frecuencia nuestros profesionales en su actividad investigativa carecen de algunos conocimientos y habilidades para manejar sistemas estadísticos potentes pero a la vez complejos como el SPSS; o bien, no disponen de estos y requieren entonces de la ayuda de otros colegas. Por otra parte se tiende a menospreciar infundadamente a Microsoft Excel como una herramienta para la investigación. Este software podría serles muy útil para este fin. Excel es básicamente una hoja de cálculo electrónica, pero es mucho más, es una opción muy buena para resolver la mayoría de las necesidades estadísticas de nuestras investigaciones.

Palabras clave: Programas de computación, Estadística, Investigación

ABSTRACT

Frequently, our professionals in their research activity, lack of some knowledge and skills to deal with strong as well as complex statistical systems like SPSS; even these last are not available occasionally, and they have to look for other partners help. Besides, sometimes Microsoft Excel is undervaluated as a tool for research. This software could be of much help for this aim. Excel is basically an electronic calculation sheet, but it is much more, it is a very good option to solve most statistical needs in our researches.

Key words: Software, Statistics, Investigation

Recibido: 6 de septiembre de 2006

Correspondencia:

Lic. Luis Orlando Pérez González.
Departamento de Informática
Facultad de Ciencias Médicas de Cienfuegos
Calle 51 A y Ave 5 de septiembre
Cienfuegos CP: 55100

INTRODUCCIÓN

El desarrollo de las investigaciones y sobre todo el manejo de los datos en la actualidad requiere de la utilización de herramientas potentes para facilitar el uso de los datos, agilizar la obtención de los resultados y para obtener una mejor precisión de los mismos. Una de esas potentes herramientas que pueden utilizarse en nuestro medio es el paquete de office de Windows y dentro de este al programa Microsoft Excel, el cual tiene una amplia gama de utilidades. En ese sentido, el objetivo de esta comunicación es exponer los elementos más importantes de este programa que son de utilidad para el desarrollo de las investigaciones científicas.

DESARROLLO

Microsoft Excel es, sin dudas, una buena opción para los investigadores. La captura de datos tiene múltiples facilidades así como la edición de estos y cuando el investigador logra disponer de una base de datos –los campos o las variables de interés para su investigación figurando en las columnas y los artículos, o sea, los casos observados o interrogados figurando como filas, teniendo así un arreglo rectangular “parejo”-, entonces puede con facilidad aplicarle a esos datos el tratamiento estadístico basado en las posibilidades existentes disponibles, tales como:

- Ordenar los artículos según el criterio de una o varias variables.
- Filtrar, que significa visualizar en un momento dado solo los artículos que cumplan ciertos requisitos.
- Aplicar diversos modos de subtotales para algunas variables, al grupo de artículos que tienen el mismo

Aprobado: 29 de septiembre de 2006

valor para alguna otra variable.

- Solicitar una tabla dinámica que resuma la distribución de frecuencias de alguna variable o que relacione varias variables.
- Rendir una estadística descriptiva de alguna(s) variable(s).
- Realizar pruebas de la inferencial, suministradas éstas en su análisis de datos.

Esto, entre las más comunes y más útiles, ya que hacer una exhaustiva mención de las posibilidades previstas en este software sería prácticamente imposible.

La barra de menú muestra entre sus opciones, el menú Datos y el menú Herramientas a los que nos referiremos en este trabajo.

El menú Datos agrupa, entre otras posibilidades, las facilidades de ordenamiento, filtrado y subtotales que son muy útiles en las etapas de captura de la información, edición y su puesta a punto; también a esta opción de menú Datos corresponde la emisión de tabla dinámica para una etapa posterior correspondiente a la recuperación de información y rendición de informes.

El análisis de datos, que pertenece al menú Herramientas encierra una mayoría importante de pruebas inferenciales, además de reportes descriptivos. De no presentarse la opción al desplegarse la opción de menú Herramientas, se elige por única vez la opción Complementos de esta misma opción de menú Herramientas y se marca la casilla correspondiente a análisis de datos. Hecho esto, debe mostrarse en Herramientas la opción análisis de datos.

Excel cuenta con un considerable número de funciones suministradas, entre las que figuran las "Estadísticas", junto a las "Matemáticas", las de "Texto", las "Financieras" y otras, capaces de ofrecer una increíble cantidad de resultados que podrían ser aprovechados en las investigaciones.

Para la recuperación de la información disponible en la base de datos a través de resúmenes son incuestionables las posibilidades de "Informe de tablas y gráficos dinámicos", también de la opción de menú Datos. Estas tablas se caracterizan por tener en su diseño cuatro secciones: la de página, la de filas, la de columnas y la de campos, que así nombradas definen las partes de una tabla tridimensional. ¿Por qué? Porque el campo que sea designado para la sección de filas se interrelacionará con el que sea designado para la sección de columnas definiendo las celdas o cuadros de la tabla, y en el campo datos o sección de datos se especifica la información que se desea ver en cada celda o cuadro, o intersección de fila con columna. Así estaríamos haciendo alusión a una tabla bidimensional, plana, pero que complementada con el campo designado para la sección de página, permitiría, por ejemplo, que en un momento pudiera estarse mostrando la tabla de datos de un departamento, o de otro, o de todos en conjunto, asumiendo que exista en la base de datos un campo o

variable que se refiera al departamento.

más formal.

Las tablas dinámicas dan múltiples facilidades como designar dos campos para filas (o columnas) y en ese caso, el segundo abre el detalle para cada una de los valores del primero. Puede diseñarse solo con filas -o bien con columnas-, en este caso la tabla sería de frecuencias, no de contingencia o relación, y por supuesto, también la sección de página es opcional. En el cuerpo de la tabla (sección de datos) pueden definirse funciones como contar los casos (lo más frecuente) o usar otras funciones como sumar, promediar, dar el valor mayor, o el menor de esa intersección de fila con columna, así como acompañar el dato con algún porcentaje con variantes diferentes para su cálculo. Se pueden ocultar categorías o valores de alguna variable a voluntad, se pueden ordenar las categorías de cualquier variable de fila o de columna. Se puede manejar la presencia o no de totales y subtotales, así como hacer fusiones o agrupamientos de algunos valores puntuales de la variable de fila o de columna, entre otras ventajas; y lo que es muy importante y le confiere el apellido a la tabla de dinámica: los cambios que se realicen sobre la base de datos subyacente que la alimenta se reflejan en la tabla con solo una simple orden de actualización. Son una herramienta poderosísima para la actividad investigativa con un único señalamiento desfavorable: su falta de estética, la cual requiere de alguna elaboración posterior para una presentación.

Pongamos como ejemplo el caso de una base de datos de una investigación que incluye entre otros campos: el diagnóstico clínico, el diagnóstico histológico, la operación realizada, el nombre del especialista, sexo del paciente, etc., y en un momento determinado quisiéramos tener conocimiento para cada uno de los diagnósticos clínicos (campo de página), cuántos casos operados (en el campo de datos) existen, relacionando el tipo de operación realizada (campo de columna), con cada uno de los diferentes diagnósticos histológicos (campo de fila). Obtenida la tabla con todos los diagnósticos clínicos diferentes recogidos en la base de datos resultaría:

D.C.	(Todas)		
Contar de Dr*.O.R.			
D.H.	Parotidectomia	Submaxilectomia	Total general
Adenoc.	2		2
A monomorfo	3		3
A pleomorfo	22	4	26
CAQ	3		3
C mucosopi	1		1
Linf no HDG		1	1
Meta CEBD-piel	3	2	5
Oncocitoma	1		1
T Warthin	10	1	11
Total general	45	8	53

* Campo sobre el que se ordenó contar para llenar la tabla y que puede editarse de manera personalizada.

Y a una orden de solo ver el diagnóstico clínico de tumor submandibular obtenemos dinámicamente la tabla limitada a este diagnóstico que es subconjunto de la anterior, como se ve a continuación:

D.C.	T submand.		
Contar de Dr.*O.R.			
D.H.	Parotidectomia	Submaxilectomia	Total general
A pleomorfo	2	2	4
CAQ	1		1
C mucoepi	1		1
Linf no HDG		1	1
Meta CEED-piel	1	2	3
Total general	5	5	10

*Campo sobre el que se ordenó contar para llenar la tabla y que puede editarse de manera personalizada.

El Análisis de datos provee diferentes pruebas, entre ellas: Estadística descriptiva que rinde el siguiente informe para una o más variables seleccionadas, como se mostrará a continuación o el análisis comparativo de dos o mas poblaciones.

En ese sentido el análisis de datos provee, mediante la inferencial paramétrica, una serie de pruebas entre las que figuran aquellas más comúnmente usadas como: análisis de varianza de un factor, regresión, prueba f para la igualdad de dos varianzas, prueba t para dos muestras pareadas, prueba t para dos muestras independientes con varianzas iguales, prueba t para dos muestras independientes con varianzas desiguales y prueba z para las medias de dos muestras. Algunas de estas son presentadas a continuación:

Análisis descriptivo de los datos.

Columna1	
Media	5
Error típico	0,816496581
Mediana	5
Desviación estándar	2,160246899
Varianza de la muestra	4,666666667
Curtosis	-1,2
Coefficiente de asimetría	0
Ranqo	6
Mínimo	2
Máximo	8
Suma	35
Cuenta	7
Mayor (2)	7
Menor (2)	3
Nivel de confianza(95,0%)	1,997896621

*Respuesta del programa para señalar un valor no está disponible para una función o fórmula.

Análisis inferencial paramétrico de los datos

Prueba F para varianzas de dos muestras		
	Variable 1	Variable 2
Media	5	15,75
Varianza	4,666	5,071
Observaciones	7	8
Grados de libertad	6	7
F	0,920	
P(F<=f) una cola	0,467	
Valor crítico para F (una cola)	0,237	

De esa manera todas las pruebas antes mencionadas se pueden realizar utilizando la opción análisis de datos del menú herramientas.

Pero además de las opciones de tablas dinámicas y del análisis estadístico descriptivo e inferencial paramétrico de los datos, este programa cuenta con otras herramientas muy importantes para el análisis y presentación de los datos de una investigación. Una de ellas está relacionada con la posibilidad de construir gráficos a partir de los datos presentados en una hoja de cálculo o gráficos dinámicos, de gran utilidad para un mejor análisis y resumen de los datos.

También este programa cuenta con la posibilidad de desarrollar un enorme número de funciones matemáticas, económicas y estadísticas.

Los diferentes tipos de funciones pueden usarse situándose en la celda donde se desea el resultado de la función que se necesite, tecleando un signo "=" que prepara a la celda para que le sea pegada la función y seguidamente el usuario investigador deberá ir respondiendo las especificidades de la función ya que cada una tiene sus propios requerimientos y argumentos de acuerdo a su objetivo. La lista que se presenta – incompleta por demás y nombradas exactamente como las suministra Excel-, puede ilustrar el alcance del universo de funciones de este software para el trabajo estadístico.

COEF.DE.CORREL
COEFICIENTE ASIMETRÍA
COEFICIENTE .R2
CUARTIL
CURTOSIS
DESVEST
DISTR.BINOM.
DISTR.CHI
DISTR.NORMAL
DISTR..NORMAL ESTAND
DISTR.T
ESTIMACIÓN.LINEAL
FISHER
FRECUENCIA
INTERVALO.CONFIANZA
K.ESIMO MAYOR
K.ESIMO MENOR
MAX
MEDIANA
MIN
MODA
PERCENTIL
PROMEDIO
PRONÓSTICO
PRUEBA.CHI
PRUEBA.T
PRUEBA.Z
PRUEBA.F
TENDENCIA
VAR