

Metodología de la Investigación en Medicina Crítica (II parte) Bioestadística básica

Marco Muñoz¹, Daniel Morales¹, Cristian Grenett¹, Emiliano Bórquez¹, Evelyn Valdivia².

RESUMEN

En este segundo capítulo abarcaremos los fundamentos de la bioestadística de manera ágil para que los lectores puedan manejar con soltura las herramientas fundamentales que les permitirán recoger, clasificar, resumir, hallar regularidades y analizar los datos de índole biológica o clínica, siempre y cuando la variabilidad e incertidumbre sea una causa intrínseca de los mismos; así como de realizar inferencias a partir de ellos, con la finalidad de ayudar a la toma de decisiones y en su caso formular predicciones.

De lo anterior se desprende que la estadística no da fundamento a los hallazgos de un estudio sino que es un método que sistematiza la información de manera útil. También queda de manifiesto que el método estadístico a utilizar queda definido por las condiciones del diseño del estudio antes de su implementación.

Cabe advertir que la estadística es fundamentalmente un método matemático muy complejo de resumir por lo que el texto se focalizará en lo prioritario dejando al lector la labor específica de ahondar en aquellos apartados que le son de su interés.

El temario que abarcará esta serie será el siguiente:

1. Conceptos iniciales
2. Estadística descriptiva
3. Estadística inferencial
4. Estadística multivariada
5. Sugerencias para la estadística inferencial

SUMMARY

The second chapter will encompass and address the fundamentals of biostatistics from a practical approach so as to enable readers to manage some basic tools for collecting, classifying, summarizing, spotting regularities, and analyzing biological or clinical data, provided that variability and uncertainty be an intrinsic cause of such data. Readers will also be capable of drawing inferences from data in order to get assistance in decision-making processes, and in making predictions, when appropriate.

From the preceding statements it can be derived that Statistics does not support study findings but it constitutes a method for systematizing information and make it easily and usefully available. It is also stated that the statistical method to be applied depends largely on the study design features and it is established before its implementation.

It must also be said that Statistics is basically a complex and hard-to-summarize mathematical method, therefore our work will be focused on essential material, letting the reader the possibility to search for further information.

The Series will address the following topics:

1. Basic Concepts
2. Descriptive Statistics
3. Inferential Statistics
4. Multivariate Statistics
5. Suggestions for Inferential Statistics

1 Médicos Residentes UPC Clínica INDISA

2 Encargada Técnica, Enfermería UPC Clínica Indisa

1. CONCEPTOS INICIALES

Niveles de medición

Los científicos han desarrollado un sistema para clasificar los diferentes tipos de medición. La importancia de este sistema se refleja en que las operaciones analíticas que pueden llevarse a cabo a partir de un conjunto de datos dependen del nivel de medición utilizado, siendo los cuatro principales los siguientes: nominal, ordinal, intervalar y racional o proporcional.

a. Medición nominal

Nivel más básico de medición que consiste en la asignación de nombres o números con el simple propósito de clasificar determinadas características en categorías. Son ejemplos de variables susceptibles de medición nominal el género, el tipo sanguíneo y la profesión. La codificación con números no esconde un trasfondo matemático sino literal o de "etiqueta" por lo que las operaciones no tienen sentido salvo la suma de iguales (frecuencias) y su contraste. La medición nominal suministra sólo información de equivalencia: Juan es igual a Juan, 1 es igual a 1, hombre es igual a hombre y distinto de mujer (no equivalentes). El requisito básico para medir un atributo en escala nominal es que las categorías sean mutuamente excluyentes y exhaustivas en conjunto (una característica a cada medición y los resultados posibles abarcan todas las posibilidades).

b. Medición ordinal

Permite clasificar los objetos en función de un atributo conforme a su posición relativa respecto a otros. Así se mide equivalencias y posición relativa a un mismo nivel. Si se ordenaran sujetos por peso desde el más ligero al más pesado se estaría utilizando un nivel ordinal de medición. Los números (si se utilizan como etiquetas) no son arbitrarios, ya que significan incrementos o decrementos además de equivalencias según un criterio establecido (Capacidad Funcional I, II, III, IV. AHA). Sin embargo, la medición ordinal no informa acerca de cuánto mayor es un atributo respecto de otro nivel (la medición ordinal sólo indica la distribución relativa de los niveles de un atributo).

Como en el caso de las escalas nominales, los tipos de operaciones matemáticas viables con datos ordinales están restringidas. En general, los promedios carecen de sentido en las mediciones de orden jerárquico, en cambio, el recuento de frecuencias, porcentajes y otros procedimientos estadísticos sí resultan apropiados para analizarlos.

c. Medición intervalar

La medición intervalar ocurre cuando el investigador puede especificar tanto el orden jerárquico de los objetos en

función de un atributo como la distancia que media entre ellos. La distancia entre los valores numéricos de una escala intervalar representa distancias o diferencias equivalentes en el atributo que se mide. Así pues, las mediciones intercalares son más informativas que las ordinales, pero no suministran información acerca de la magnitud absoluta del atributo y el punto cero de esta escala es arbitrario, por ejemplo: un puntaje SOFA 10 puntos es el doble que uno de 5 puntos, pero un puntaje cero no existe (no se preestableció), o una temperatura de cero grados no tiene sentido de una escala a otra (es preestablecido arbitrariamente).

Las escalas intercalares amplían mucho las posibilidades analíticas del investigador. Los intervalos entre los números pueden sumarse o sustraerse y tienen un significado preciso, así estos datos pueden promediarse y otras operaciones matemáticas complejas son posibles.

d. Medición racional o proporcional

Estas suponen un cero significativo no arbitrario. Así la suma de características de una escala intervalar se enriquece con la particularidad de medición absoluta de un atributo para cada objeto de observación. Muchas medidas físicas proporcionan datos de nivel racional: el peso de una persona, la altura de un objeto.

El cero real de las escalas racionales permite realizar toda clase de operaciones matemáticas y así todos los procedimientos estadísticos son válidos en su aplicación a datos de este nivel de medición. Esta es la medición idónea para un científico, pero resulta inasequible para la mayor parte de los objetos de naturaleza cualitativa.

Comparación de los niveles

Los cuatro niveles de medición constituyen una jerarquía desde el nivel superior proporcional al inferior nominal. La complejidad del análisis de un grupo de datos puede llevarnos a la necesidad de resumirlos en tablas de frecuencias ganando simplicidad visual para el investigador o para quien lee nuestra investigación. Pero el precio de esta ganancia es que al pasar de un nivel superior a uno más inferior se va perdiendo cantidad de información y capacidad de análisis matemático. Por ejemplo, respecto de la talla (cm.) de 15 pacientes adultos hombres en UCI (Tabla 1):

2. ESTADÍSTICA DESCRIPTIVA

Distribuciones de frecuencias

Los datos cuantitativos crudos (o en bruto) resultan angustiantes, no es posible discernir siquiera tendencias generales hasta que se los organiza. Como ejemplo consi-

TABLA 1. TALLA EN CENTÍMETROS DE 15 PACIENTES ADULTOS HOMBRES EN UCI

Proporcional	Intervalar	Ordinal	Nominal
185	185	Muy alto	Grande
175	175	Muy alto	Grande
170	175	Alto	Grande
170	170	Alto	Grande
165	165	Mediano	Grande
165	165	Mediano	Grande
165	165	Mediano	Grande
164	164	Bajo	Chico
163	163	Bajo	Chico
163	163	Bajo	Chico
163	163	Bajo	Chico
162	162	Bajo	Chico
160	160	Bajo	Chico
160	160	Bajo	Chico
160	160	Bajo	Chico

Nota: los datos de nivel proporcional e intervalar no tienen diferencia por la naturaleza proporcional intrínseca de la variable talla.

deremos el siguiente cuadro y observemos lo imposible de establecer tendencias en los datos de una variable apenas dicotómica y nominal, aun ordenando por trimestre:

Datos de 156 pacientes ingresados consecutivamente a UCI.
 Ordenados por trimestre A, B, C y D
 Variable: Ventilación Mecánica (VM)
 1 = si; 0 = no

```
101010101111010 010101010100001 101010101000010 010010101000001
010100111111111 001110100010010 111100000100100 001111110001
110000000 001001000100010 010010101010001
```

En el ejemplo siguiente usando una variable continua con un número menor de datos se hace igualmente complejo establecer tendencias.

Valores de glicemia capilar en pacientes de UTI-UCI control 07:00 AM

```
67 78 89 130 200 210 180 70 89 99 110 42 90
100 200 85 48 78 90 126 85 59 69 78 215 184
```

Sin embargo estos datos pueden organizarse de manera tal que podrán describirse en base a tres características: la forma de la distribución de valores, la tendencia central y la variabilidad de los datos.

Construcción de distribuciones de frecuencias

Las distribuciones de frecuencias nos ayudan a establecer orden en los datos numéricos. Este ordenamiento va desde el valor más bajo al más alto en frecuencia, haciendo un recuento del número de veces que se repite un dato. Así –en el ejemplo– se puede visualizar un puntaje del score X más frecuente en esta serie de pacientes, la tendencia de la agrupación y el número de datos obtenidos (N). El único requisito para que la construcción de una distribución de frecuencias sea válida es que los valores de la variable a medir sean mutuamente excluyentes y exhaustivos entre sí. Se puede aprovechar la tabla para anotar también los porcentajes de cada categoría en una columna adyacente. Los totales sumatorios de cada columna deben corresponder a N y 100% (Tabla 2).

Los datos así se pueden graficar en un histograma o en un polígono de frecuencias que tienen la ventaja de transmitir la información de forma casi instantánea y visual. El área bajo la curva de una de estas gráficas representa la frecuencia acumulada exacta de datos (Figura 1).

Formas de distribución

Una distribución de datos numéricos puede tomar infinitas formas visuales, sin embargo es importante tener algunas consideraciones acerca de ellas:

- Distribuciones según simetría: simétricas si constan de dos mitades similares o iguales, asimétricas si sus mitades son diferentes y éstas pueden dividirse en aquellas con sesgo positivo si la cola más larga está a la derecha y con sesgo negativo en caso contrario.

TABLA 2. SCORE X DE PACIENTES AL INGRESO UPC DURANTE 1° Y 2° TRIMESTRE

Puntaje	Recuento	Frecuencia (f)	Porcentaje (%)
1	111111	6	4,2
2	11111111	8	5,6
3	11111111	8	5,6
4	11111111	7	4,9
5	1111111111	10	7,0
6	1111111111	10	7,0
7	11111111111	11	7,7
8	1111111111111	13	9,2
9	1111111111111	13	9,2
10	11111111111111	15	10,6
11	1111111111111	13	9,2
12	11111111111	10	7,0
13	1111	4	3,8
14	111111	6	4,2
15	11111111	8	5,6

N = 142 = $\sum f$ $\sum \% = 100\%$

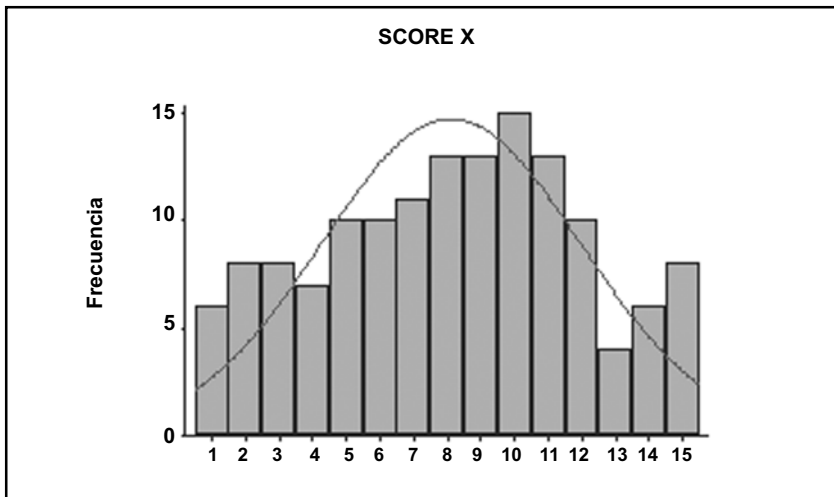


Figura 1.

- Distribuciones según modalidad: unimodales si presentan una cima o cresta, bimodal si presentan dos y multimodales si presentan más de dos.
- Distribuciones muy conocidas: normal o Gaussiana, Estándar (z), Student (t), Poisson.
- Distribuciones simétricas unimodal (1) y simétrica sin moda, multimodal (2), sesgada negativa (3) y positiva (4) (Figura 2).

Distribución normal y la regla empírica

La distribución normal tiene como media el cero y su área bajo la curva tiende a 1, es unimodal y simétrica no sesgada, algo achatada con colas largas; si se grafican las desviaciones estándar se puede establecer que el 68% de las observaciones está entre $\pm 1DS$, el 95% de las observaciones está entre $\pm 2DS$ y el 99,7% de las observaciones está entre $\pm 3DS$ (regla empírica) (Figura 3).

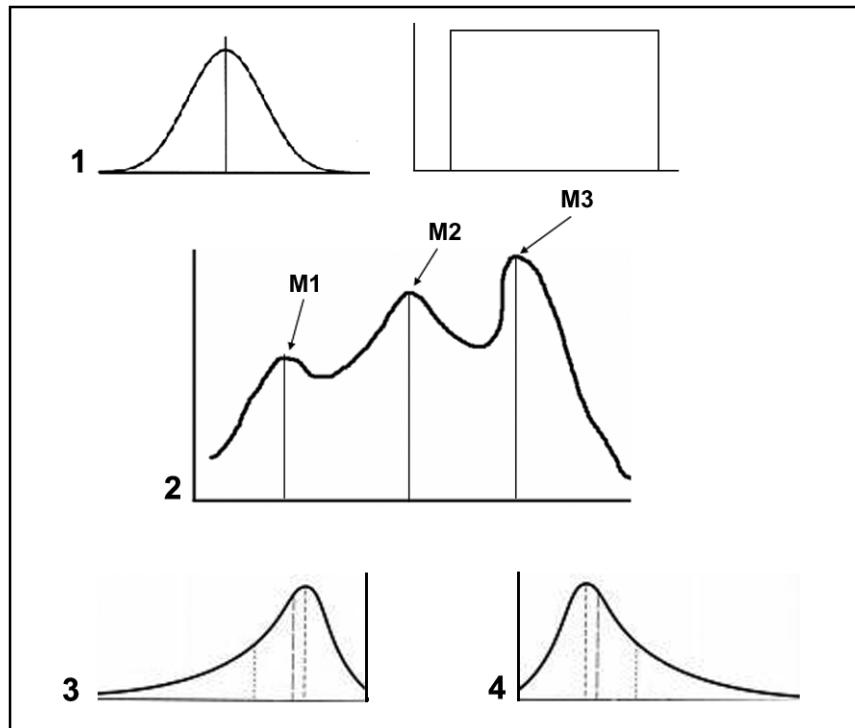


Figura 2.

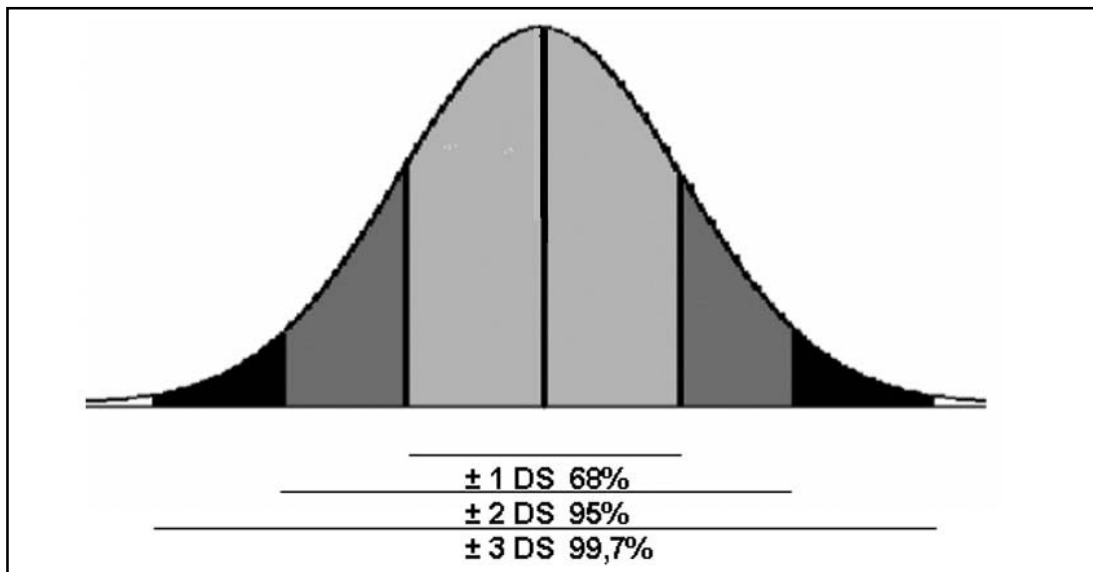


Figura 3. Regla empírica.

Tendencia central

Para el ordenamiento de datos crudos y aclaración de los patrones de grupo una distribución es muy útil. Sin embargo, el investigador necesita frecuentemente más que un patrón u orden, un número o dato que represente

al grupo y éste generalmente se localiza en el centro de la distribución. Existen tres medidas de tendencia central o promedios: moda, mediana y media.

- Moda: poco utilizada, es inestable en datos numéricos incluso de muestras de una misma población, es fácil de calcular. Se usa para datos nominales: el patrón

usual radiológico (modal) de neumonía neumocócica es con patrón radiológico lobar de condensación. (Variable nominal: patrón radiológico de neumonía neumocócica, categorías: intersticial, lobar, mixta, etc.)

- Mediana: punto de una escala numérica que separa el 50% superior e inferior de los datos, es entonces una medida de posición (no toma en cuenta el valor del dato, sólo su posición), es muy estable a pesar del sesgo que pueda tener una distribución. En nivel ordinal es la más usada.
- Media: punto en una escala numérica que se calcula como la suma de todos los datos dividida entre el número total de ellos, por lo que el valor de cada dato como la frecuencia de este valor hace variar a la media. En nivel intervalar y proporcional es la más usada.

Comparación de moda, mediana y media

Cuando se requiere representar las edades de los pacientes en UTI y compararles con otra UTI la media es la medida de tendencia central adecuada, pero si quisiera saber cuál es la edad más típica de los pacientes de mi UTI y compararla, entonces la mediana es más adecuada; si quisiera representar la edad más frecuente de los pacientes en UTI la moda es la adecuada.

Las tres medidas siguen un patrón determinado en una distribución sesgada siendo la moda la más alta, cayendo la mediana y luego la media hacia la cola más larga independientemente del tipo de sesgo (Figura 4).

Variabilidad

Si dos distribuciones tienen medidas de tendencia central semejantes no significa que los datos se distribu-

yen de igual manera, al estar diferentemente sesgadas la variabilidad de los datos puede ser diferente, por ejemplo dos distribuciones de valores de Presión Arterial Media en dos UTI pueden tener promedio de 70 mmHg, sin embargo la primera tiene una amplia variación de los valores y la segunda una muy pequeña. Puede inferirse así que los pacientes de la primera distribución están más graves al presentar valores muy por debajo y muy por alto de los considerados para una PAM normal, en cambio la segunda distribución presenta casi a todos los pacientes con PAM cercanas al valor normal. Ambas UTI en cuanto a esta variable son muy diferentes (Figura 5).

Las medidas de desviación de los datos entre sí pueden ser: rango, rango semicuartil y desviación estándar.

- Rango: o intervalo es la diferencia entre el valor más alto y el valor más bajo de una distribución, a menudo se representa por los valores referidos sin restarlos. Aunque es fácil de calcular, es muy inestable (de la misma forma que la moda) por lo que de una muestra a otra tiende a variar muchísimo. Además ignora por completo variaciones entre valores al interior de la distribución. Por lo tanto el rango se informa como un indicador grosero acompañando a otras medidas de variabilidad más certeras.
- Rango semicuartil: un cuartil se refiere a aquel punto que separa el 25% de los datos de una distribución, por lo que hay un cuartil superior (Q3) uno medio (mediana Q2) y uno inferior (Q1). El rango semicuartil corresponde a la mitad de la distancia entre Q1 y Q3. O sea el valor medio del 50% de los valores más frecuentes. Así es muy difícil que un valor extremo cambie el SQR (semiquartile range).
- Desviación estándar: esta medida toma en cuenta todos y cada uno de los valores de la distribución. Es

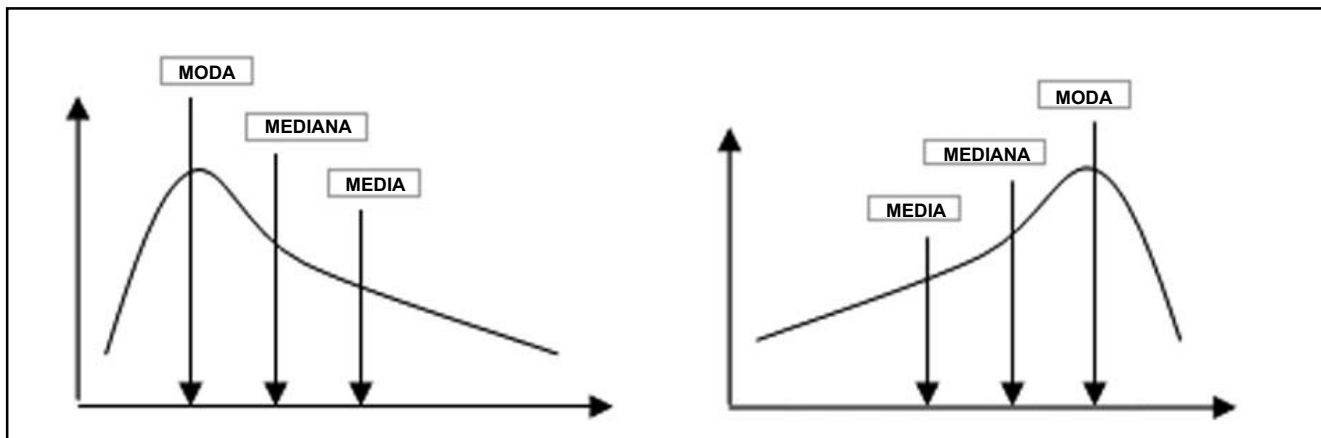


Figura 4.

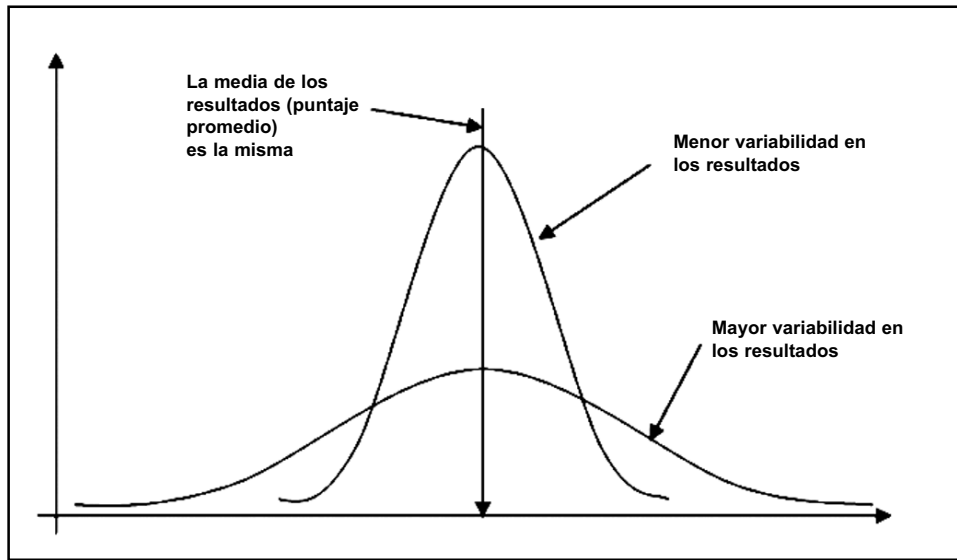


Figura 5. Distribuciones de valores de PAM en dos UTI adultos.

la magnitud promedio en que los valores individuales se desvían de la media. Si un paciente tiene natremia de 145 y la media es de 140 podríamos decir que la desviación individual es 5 si sumamos todas las diferencias y las dividimos por n tendríamos una desviación promedio, sin embargo este valor es cero puesto que se compensan diferencias positivas con negativas como propiedad de la media. Así, un artilugio es elevar al cuadrado cada desviación individual, sumarlas, dividir las por n (varianza) y luego al resultado extraer la raíz cuadrada: ésta es la desviación estándar.

Estadísticas descriptivas bivariadas: tablas de contingencia y correlación

La investigación se aboca generalmente al estudio de la relación entre variables y por lo tanto implica un método de descripción de esta relación.

Tablas de contingencia

Es una distribución bidimensional de las frecuencias de dos variables sobre la misma muestra en estudio. Por ejemplo género versus gravedad de la enfermedad (dos variables nominales) (Tabla 3):

En el mismo ejemplo pero ahora agregando a la variable gravedad las categorías de leve, grave, muy grave se establece una variable ordinal (Tabla 4):

Si las dos variables contrastadas son a lo menos ordinales también se puede tabular en una tabla 2x2. Pero si las variables son discretas o continuas (nivel intercalar o proporcional) se deben graficar en diagramas de dispersión donde cada punto une el valor de ambas variables, en el ejemplo edad versus APACHE II en 10 pacientes (Figura 6):

Correlación

Se entiende por correlación a la correspondencia entre los cambios de magnitud de una variable respecto

TABLA 3

	No grave	Grave	Totales
Femenino	2	3	5
Masculino	1	4	5
Totales	3	7	10

TABLA 4

	Leve	Grave	Muy grave	Totales
Femenino	2	1	2	5
Masculino	1	1	3	5
Totales	3	2	5	10

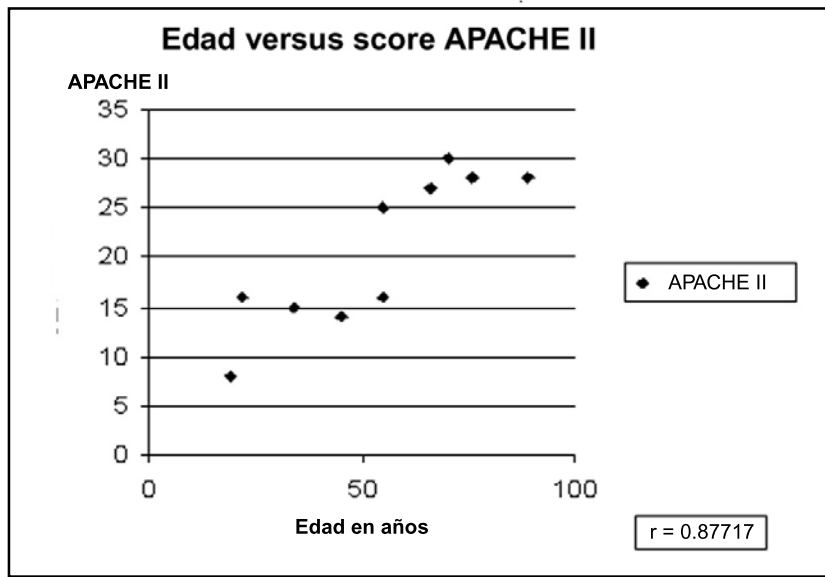


Figura 6.

de la otra. Su cálculo rara vez se hace a mano pudiendo obtenerse por una calculadora científica o en Excel o en cualquier programa de estadística. El aquí calculado es el coeficiente de correlación de Pearson (para dos variables al menos intervalares), existiendo el coeficiente rho de Spearman (para variables ordinales). Su valor puede fluctuar entre una correlación perfecta positiva (+1) que implica una pendiente positiva y la posibilidad de anticipar perfectamente el valor correspondiente a cada medida con ésta, y la correlación perfectamente negativa (-1) con una pendiente negativa y la misma capacidad anticipativa. El valor cero implica que una variable no se correlaciona en absoluto con la otra dando una gráfica recta con pendiente en 90 o 180 grados (Figura 7).

Recomendaciones en cuanto a los niveles de medición y la estadística descriptiva:

- Utilice el sistema de medición más alto posible para garantizar la utilización estadística más avanzada y eficiente de los datos.
- La excepción debe tenerse en cuenta cuando la situación clínica no amerite detalles engorrosos (rangos de edad, cuartiles de APACHE II, rangos de dosis de infusión de medicamentos) lo que facilita informar rápido y casi visual.
- La estadística descriptiva puede dar respuesta a preguntas de investigación de tipo específicamente descriptivas, pero lo más común es que sólo permita conocer o describir las variables de un estudio inferencial.

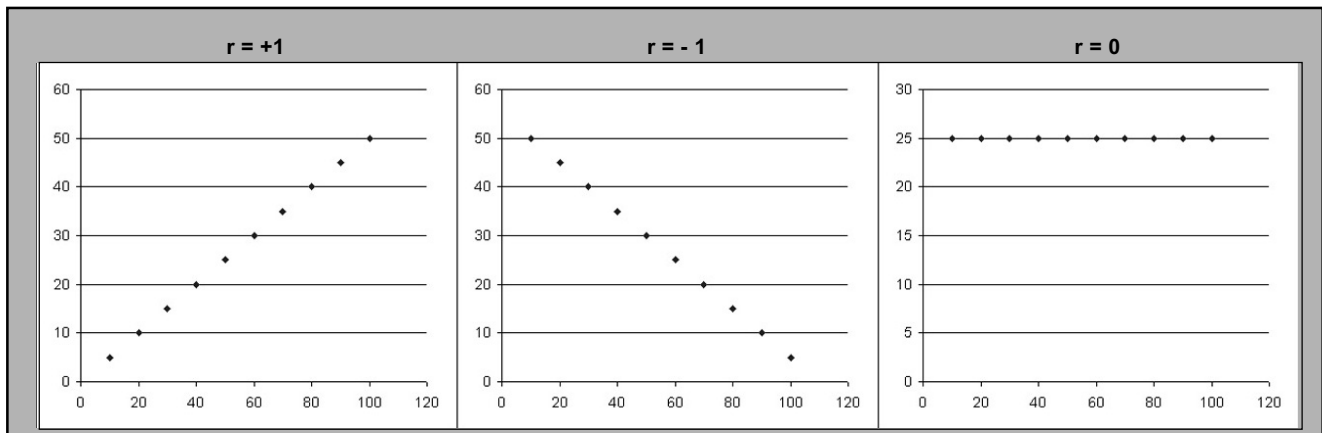


Figura 7. Tres ejemplos de correlaciones extremas entre dos variables.

- Si la distribución de los datos es sesgada (asimétrica) debe considerar que las tres medidas de tendencia central son pertinentes en información dependiendo de lo que se requiera informar.
- Las correlaciones no corresponden a relaciones fuertes de causalidad o asociación, pueden inducir falacia ecológica, por lo que corresponde muchas veces realizar estudios inferenciales para comprobarlas o descartarlas.

3. ESTADÍSTICA INFERENCIAL

Con base en las leyes de la probabilidad y admitiendo entonces siempre un grado de error en un marco de procesos sistemáticos y objetivos, la estadística inferencial suministra los medios necesarios para formular conclusiones acerca de una población a partir de los datos obtenidos de una muestra. Con su ayuda los investigadores pueden formular juicios o generalizar los resultados a grandes conjuntos de individuos con base en la información de un número limitado de sujetos.

Distribuciones de muestreo (DDM)

Si se utiliza una muestra para las estimaciones de los resultados es conveniente que ésta sea lo más representativa posible. Así las muestras aleatorias son ideales para este cometido. Aun siendo aleatorio no puede esperarse que las características de muestra y población fuente sean idénticas. La fluctuación de resultados entre una muestra y otra se conoce como error de muestreo. El desafío es demostrar que los valores obtenidos de una muestra son una buena estimación de los parámetros en la población.

Para empezar debemos explicar lo que es una distribución de muestreo en el cual un parámetro de tendencia es tomado sistemáticamente de un grupo de muestras idénticas de una misma población fuente. Este procedimiento es más bien teórico y matemático no siendo necesaria la extracción real de este número de muestras.

Características de las distribuciones de muestreo

Cuando se extrae de una población infinita un número infinito de muestras se obtienen algunas características de la DDM:

- 1) La distribución de muestreo de las medias es normal.
- 2) La media de una distribución de muestreo es siempre igual a la media poblacional.
- 3) La exactitud de una media cualquiera de una DDM de las medias es calculable gracias a la regla empírica (porque la DDM de medias es normal).

Error estándar de la media

La desviación estándar de una DDM de medias se conoce como error estándar de la media (EEM). Este se calcula a partir de la desviación estándar de la muestra y del tamaño de la población.

$$EEM = DE / \sqrt{N}$$

Este estadístico constituye una estimación del grado de error de muestreo entre una media muestral y otra. La probabilidad de obtener una media que esté entre dos valores extremos se puede estimar gracias al error estándar y la regla empírica o tablas de conversión de los valores de la distribución normal y sus áreas bajo la curva respectivas.

Estimación de parámetros

Existen dos técnicas para la inferencia estadística: una es la prueba de hipótesis (que es la más frecuente de ver en informes científicos) y la otra es la estimación de parámetros que es igualmente importante. Esta última permite determinar el valor aproximado de un parámetro poblacional específico, como el valor promedio de algún atributo. Esta estimación puede ser puntual o de intervalos, esta última indica además de un valor específico de estimación una probabilidad del rango de valores dentro del cual podría localizarse el parámetro. Así se construyen los llamados Intervalos de confianza (IC) con los respectivos límites superior e inferior del intervalo. Lo que arbitrariamente se acepta es un grado de confianza de 95% o 99%. El cálculo se basa en el EEM y en los principios relacionados con la distribución normal. En el caso de un IC95 el 95% de los valores de una distribución normal está entre 1,96 desviaciones estándar (1,96 DE), luego el intervalo de confianza de una estimación puntual está entre los límites de $\pm 1,96DE \times EEM$.

Ejemplo: 100 muestras de glicemias matinales en pacientes adultos de UCI:

$$EEM = DE / \sqrt{N}$$

$$X = 155 \text{ mg/dl}$$

$$DE = 2$$

$$N = 100$$

Estimación puntual:

$$155 \pm \text{límites de los IC95} = \pm 1,96 \times 2 \sqrt{100} = \pm 39$$

$$\text{IC95} = 116 \text{ a } 194 \text{ mg/dl}$$

Prueba de hipótesis

La prueba estadística de hipótesis es un proceso de toma de decisiones. Un investigador que plantea una hipótesis obtiene resultados numéricos que deben compararse mediante un método que limite la posibilidad de

		La situación real es que la hipótesis nula es	
		Verdadera	Falsa
El investigador calcula un estadístico de prueba y concluye que la hipótesis nula es:	Verdadera (se acepta la hipótesis nula)	Decisión Correcta	Error Tipo II
	Falsa (se rechaza la hipótesis nula)	Error tipo I	Decisión Correcta

Figura 8.

que las diferencias encontradas se deban al azar y representen verdaderas diferencias. Así frente a la pregunta de investigación, la maniobra y los resultados numéricos ya obtenidos, cabe explicar lo acaecido con sendas hipótesis; éstas deben escogerse estratégicamente de modo exhaustivo y excluyente.

¿El captopril reduce la Presión Arterial Sistólica en hipertensos no tratados?

Grupo A (Tratado), Diferencia de PAS media:
 30 ± 8 mmHg

Grupo B (No Tratado), Diferencia de PAS media:
 40 ± 12 mmHg

Ha (Hipótesis Alternativa): las diferencias entre grupos son debido al tratamiento con captopril.

Ho (Hipótesis nula): no hay diferencias significativas entre grupos tratados con captopril.

Así rechazando la hipótesis más fácil de descartar (la nula) se apoya a la hipótesis alternativa, siempre y cuando las reglas del experimento no hayan sido afectadas por sesgo.

Errores tipo I y II

Al tener muestras y no información sobre la población existen incertezas sobre la real dimensión de los resultados del estudio. Un investigador puede cometer dos tipos de errores en este juego de hipótesis: rechazar una hipótesis nula verdadera (I) o aceptar una hipótesis nula falsa (II). El riesgo de cometer un error tipo I recibe el nombre de nivel de significancia (error alfa) usualmente un 0,05% o 5% de probabilidad de rechazar una Ho verdadera o de encontrar diferencias cuando éstas sólo son por azar. Este error se preestablece por el investigador. El error tipo II o beta es el complementario a la potencia estadística que tiene que ver con: el tamaño muestral, el error alfa escogido y factores propios de la naturaleza de las diferencias entre grupos (Figura 8).

Regiones críticas

Se refiere a un área de la distribución teórica aplicable. Indica si un resultado es improbable para una hipótesis nula luego de definido un nivel de significancia, aceptar Ho en esta región nos haría cometer un error tipo I. Si la distribución de frecuencias fuera normal o casi normal ésta se podría graficar así (5% de nivel de significancia, 2,5% en cada cola) (Figura 9).

Pruebas estadísticas

En la práctica, no se construye una distribución de muestreo ni se calculan las regiones críticas de la distribución, sino que los resultados del estudio se emplean para calcular un estadístico de prueba por medio de una fórmula apropiada compatible con el nivel de medición de la variable. Luego se compara éste con los valores de las regiones críticas en una tabla específica para la distribución dada y según esto se acepta o rechaza Ho.

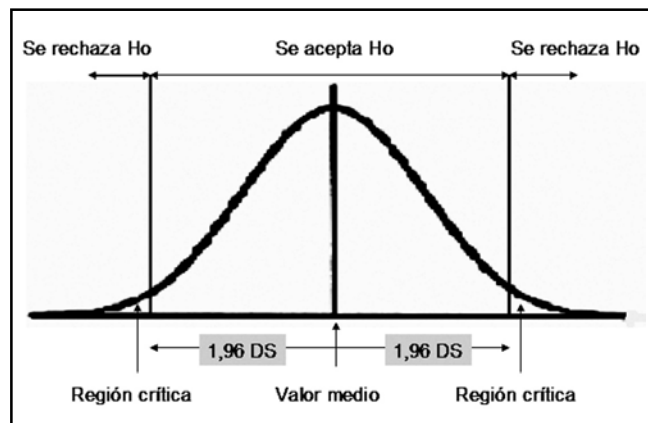


Figura 9.

Pruebas de una o dos colas

A veces H_0 sólo da lugar a un resultado unidireccional de rechazo, lo que lleva a usar la región crítica en totalidad a un extremo de la distribución (5% total hacia una cola). Es más fácil rechazar H_0 hacia una cola por lo que hay que justificar muy bien la decisión de usar una prueba unilateral (recurso poco honesto para dar positividad a un estudio)

Pruebas paramétricas y no paramétricas

Las pruebas paramétricas son las más usadas en investigación clínica, éstas reúnen tres requisitos esenciales:

1. Implican la estimación de cuando menos un parámetro.
2. Exigen a lo menos una escala intervalar de la variable.
3. Implican la suposición de que existe una distribución conocida de la variable (pe. Normal).

Las pruebas no paramétricas, no ocupan estimación de ningún parámetro y se usan cuando la variable está medida a nivel ordinal o nominal. Además son menos restrictivas respecto de la forma de la distribución de la variable lo que implica su utilidad en situaciones de esa índole.

Pruebas intersujeto e intrasujeto

Son pruebas intersujeto o de grupos independientes cuando se comparan dos muestras distintas respecto de una variable particular y son pruebas intrasujeto o dependientes cuando se mide una variable en forma repetitiva en un grupo único muestral. Los estadísticos de prueba a usar deben estar acordes con el tipo de muestra usada (una o varias).

Generalidades sobre prueba de hipótesis

Los pasos a seguir para una prueba de hipótesis correcta en general son:

1. Determinar la prueba estadística que conviene utilizar.
2. Establecer el nivel de significancia estadística previo al análisis.
3. Seleccionar una prueba de dos colas salvo excepciones justificadas.
4. Calcular una prueba estadística y calcular los grados de libertad (concepto que escapa a este artículo).
5. Obtener un valor tabulado para prueba estadística.
6. Comparar la estadística de prueba con el valor tabulado.
4, 5 y 6 se obtienen fácilmente con el programa estadístico computacional adecuado.

En general podemos distinguir pruebas estadísticas para comparar medias y proporciones además de otras que confirman la existencia de relaciones como se muestra en el cuadro siguiente. Ver anexo: Pruebas estadísticas.

Análisis de potencia

Para realizar un análisis de potencia se requiere de cuatro variables de las cuales al menos tres deben ser conocidas:

- Nivel de significancia: alfa
- Tamaño muestral: N
- Potencia: (1-beta)
- Magnitud del efecto de la población: gama

Las opciones de pruebas estadísticas se resumen a cuatro situaciones comunes:

Situación	prueba necesaria
Comparación de dos medias:	prueba de t
Comparación de tres o más medias:	ANOVA
Comparación de proporciones:	prueba de χ^2
Probar la significancia estadística de una relación lineal bivariada:	r de Pearson

El tamaño muestral es la variable a despejar que para distintas situaciones de alfa, beta y gama están tabuladas previamente y se consultan en forma directa según la prueba estadística a utilizar. Los programas computacionales o paquetes estadísticos pueden rápidamente calcular las variables y obtener N desde los datos crudos originales o de las valores estimados de alfa, beta y gama. La situación general se puede resumir como: a menor alfa, a mayor beta a menor gama... mayor tamaño muestral.

Ejemplo del programa EPINFO 2000[®] donde se calcula el tamaño muestral para un estudio de cohortes donde alfa es 5% (95% de certeza), beta es 20% (80% de potencia), gama estimado de Riesgo Relativo de 2 (y frecuencias de no expuestos 10%) = 51000! (Figura 10).

4. ESTADÍSTICA MULTIVARIADA

Regresión

El concepto implica “volver a atrás”. La verdad es que la matemática se vale de la ecuación de la línea recta para obtener a través de puntos de correlación de dos variables, la línea recta que pasa más cercana a todos los puntos en general (mínimos cuadrados de las diferencias). Esta recta obtenida por regresión nos da la relación entre variables independientes (resultado) y variables dependientes (pronósticas o de riesgo).

C:\Epi_Info\STATCALC.EXE

Epinfo Version 6 Statcalc November 1993

Unmatched Cohort and Cross-Sectional Studies (Exposed and Nonexposed)
Sample Sizes for 0.10 % Disease in Unexposed Group

Conf.	Power	Unex:Exp	Disease in Exposed	Risk Ratio	Odds Ratio	Sample Size Unexp.	Sample Size Exposed	Total
95.00 %	80.00 %	1:1	0.20 %	2.00	2.00	25,471	25,471	50,942
90.00 %	"	"	"	"	"	20,470	20,470	40,940
95.00 %	"	"	"	"	"	25,471	25,471	50,942
99.00 %	"	"	"	"	"	36,956	36,956	73,912
99.90 %	"	"	"	"	"	53,127	53,127	106,254
95.00 %	80.00 %	"	"	"	"	25,471	25,471	50,942
"	90.00 %	"	"	"	"	33,443	33,443	66,886
"	95.00 %	"	"	"	"	40,898	40,898	81,796
"	99.00 %	"	"	"	"	57,012	57,012	114,024
"	80.00 %	4:1	"	"	"	58,484	14,621	73,105
"	"	3:1	"	"	"	47,535	15,845	63,380
"	"	2:1	"	"	"	36,550	18,275	54,825
"	"	1:2	"	"	"	19,039	39,670	59,517
"	"	1:3	"	"	"	17,931	53,792	71,723
"	"	1:4	"	"	"	16,967	67,868	84,835

Change values for inputs as desired, then press F4 to recalculate.

F1-Help F5-Print F6-Open File F10-Done

Figura 10.

La regresión puede aplicarse dada la distribución de los datos en curvas naturales más o menos lineales o curvas donde la correlación es más cercana a la curva logarítmica. Así se obtienen modelos de regresión (ecuaciones de curvas) que permiten predecir el riesgo o la probabilidad de un evento en función de las variables que se miden (Figura 11).

Debe considerarse que el modelamiento estadístico por regresión es un procedimiento altamente complejo donde deben considerarse la distribución de los datos, los sesgos y variables confusoras como también la interacción entre las distintas variables lo que lleva a un comportamiento distinto de éstas cuando se toman en cuenta en forma individual o combinada (análisis uni vs multivariado), además los modelos se pueden graficar en sinuosas curvas perfectas que parecen predecir con precisión un

proceso, sin embargo carecen de bondad de ajuste a los datos crudos y en el fondo a la realidad (Figura 12).

Como ya se esbozó se necesitan al menos 50 individuos (idealmente 100) por cada variable a analizar en el modelo y sus errores por insuficiente ajuste pueden ser garrafales (por ejemplo: modelo predictor de la calidad del aire en Santiago de Chile, tiene una bondad de ajuste inferior a lanzar una moneda al aire).

Por lo expuesto los autores creemos que no debe ahondarse más en el contexto de este capítulo y dejar sólo esta introducción para encarecer la necesidad de un apoyo estadístico y epidemiológico profesional en cada análisis de regresión, tablas de vida, curvas de sobrevivencia y análisis de covarianza (Figura 13).

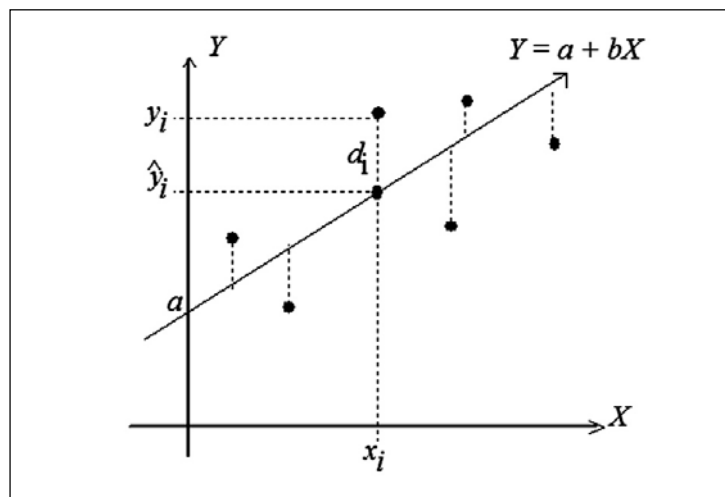


Figura 11. Mínimos cuadrados de las diferencias y regresión lineal.

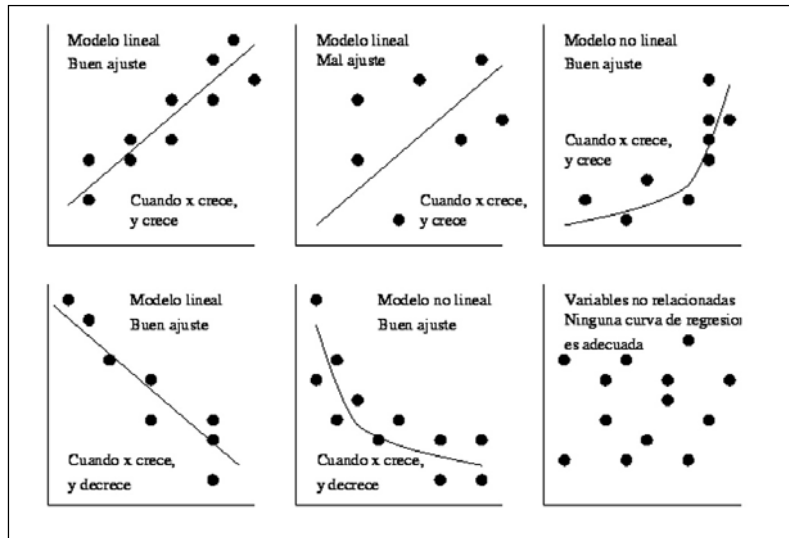


Figura 12. Modelos de regresión y bondad de ajuste con los datos originales.

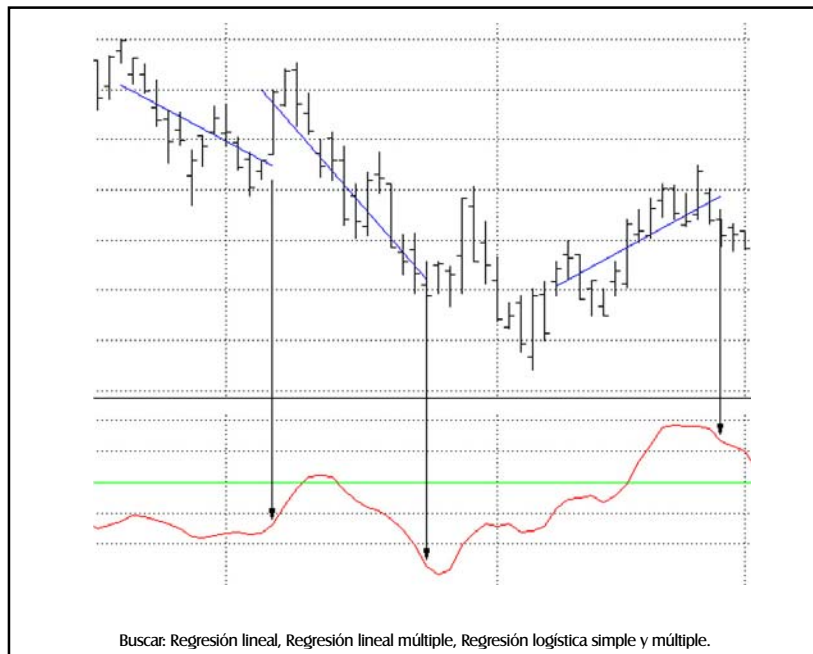


Figura 13. Modelo obtenido por regresión (abajo) y distribución real de los datos anticipados (parece un buen modelo pronóstico).

Computadoras y estadística multivariada

Los paquetes estadísticos disponen de múltiples recursos para el análisis y desarrollo de los datos y resultados de la estadística multivariada. Su uso también necesita de un gasto en tiempo y práctica además de conocimientos profundos acerca de lo que se ingresa y se obtiene en tablas de resúmenes o gráficos complejos.

5. SUGERENCIAS PARA ESTADÍSTICA INFERENCIAL Y MULTIVARIADA

La estadística inferencial es mucho más compleja que lo presentado en este resumen. La sugerencia esencial que se desprende de ello es que el clínico sin formal aprendizaje de metodología se abstenga de aplicar esta matemática avanzada y considere seriamente hacerse asesorar por

un profesional calificado tanto en el diseño de investigación como en las pruebas estadísticas a utilizar. De lo contrario existe serio riesgo de caer en errores garrafales de análisis e interpretación no necesariamente voluntarios, los que corregidos con antelación le darán renombre y valor agregado a nuestras investigaciones.

REFERENCIAS

1. Marcello Pagano, Kimberlee Gauvreau. Principios de Bioestadística. 2ª edición International Thomson Editores 2001.
2. Martin Bland. An Introduction to Medical Statistics. 3th edition Oxford University press 2002.
3. Stanton A. Glantz. Primer of biostatistics. The McGraw-Hill Companies 2005.
4. Robert G. Trapp. Basic & Clinical Biostatistics. 4th Edition Beth Dawson. The McGraw-Hill Companies 2007.

ANEXO. PRUEBAS ESTADÍSTICAS

Nombre de la prueba	Estadístico de prueba	Inter o Intragrupal	Objetivo	Nivel de medición Variable independiente	Nivel de medición Variable dependiente
PRUEBAS PARAMÉTRICAS					
Prueba de t para grupos independientes	t	Intergrupal	Probar la diferencia entre las medias de dos grupos independientes	Nominal	Intervalar, racional
Prueba de t pareada	t	Intragrupal	Probar la diferencia entre las medias de dos grupos relacionados	Nominal	Intervalar, racional
Análisis de varianza (ANOVA)	F	Intergrupal	Probar la diferencia entre las medias de tres o más grupos independientes o de dos o más variables independientes	Nominal	Intervalar, racional
ANOVA de Mediciones repetidas	F	Intragrupal	Probar la diferencia entre las medias de tres o más grupos relacionados	Nominal	Intervalar, racional
Correlación producto-momento de Pearson	r	Inter e intragrupal	Probar la existencia de una relación entre dos variables	Intervalar, racional	Intervalar, racional
PRUEBAS NO PARAMÉTRICAS					
Prueba U de Mann-Whitney	U	Intergrupal	Probar la diferencia en los rangos de valores de dos grupos independientes	Nominal	Ordinal
Prueba de la mediana	x ²	Intergrupal	Probar la diferencia entre las medianas de dos grupos independientes	Nominal	Ordinal
Prueba de Kruskal Wallis	H	Intergrupal	Probar la diferencia entre los rangos de valores de tres o más grupos independientes	Nominal	Ordinal
Prueba de los rangos y signos de Wilcoxon	Z	Intragrupal	Probar la diferencia en los rangos de valores de dos grupos relacionados	Nominal	Ordinal
Prueba de Friedman	x ²	Intragrupal	Probar la diferencia en los rangos de valores de tres o más grupos relacionados	Nominal	Ordinal
Prueba de Chi cuadrado	x ²	Intergrupal	Probar la diferencia entre dos proporciones de dos o más grupos independientes	Nominal	Nominal
Prueba de Mc Nemar	x ²	Intragrupal	Probar la diferencia entre las proporciones de muestras pareadas (2x2)	Nominal	Nominal
Prueba exacta de Fisher		Intergrupal	Probar la diferencia entre las proporciones en una tabla de contingencia 2x2 cuando N<30	Nominal	Nominal
Rho de Spearman		Inter e intragrupal	Probar que una correlación es distinta de cero (es decir, que hay una relación)	Ordinal	Ordinal
Tau de Kendall		Inter e Intragrupal	Probar que una correlación es distinta de cero (es decir, que hay una relación)	Ordinal	Ordinal
Coefficiente Phi		Intergrupal	Examinar la magnitud de una relación entre dos variables dicotómicas	Nominal	Nominal
V de Cramér	V	Intergrupal	Examinar la magnitud de una relación entre variables en una tabla de contingencia (no restringida a 2x2)	Nominal	Nominal