



# **¿Qué dicen y qué no dicen las estadísticas? (Bioestadística sin dolor o como comprenderla sin morir en el intento)**

**Dr. Ricardo Bolaños**

**Dr. de la UBA. Área Farmacología**

**Médico Farmacólogo**

**Docente Autorizado de Farmacología. UBA**

**Profesor Asociado de Metodología de la Investigación Clínica. UAI**

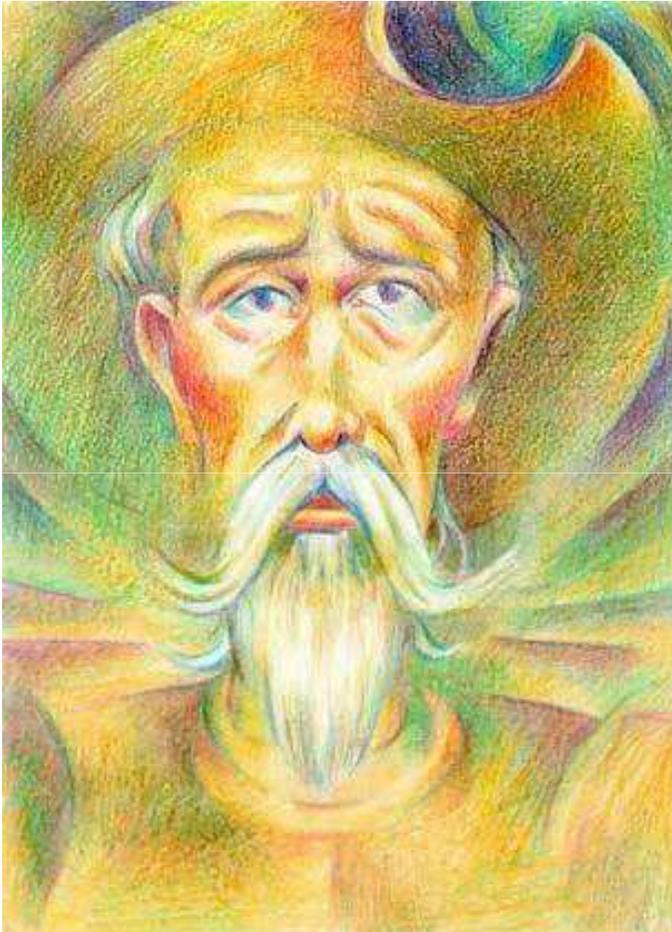
**Miembro Docente del Instituto de Investigaciones Epidemiológicas.**

**Academia Nacional de Medicina**

**Área de Ciencia Reguladora.**

**Dirección de Evaluación de Tecnologías Sanitarias. ANMAT**

**El disertante declara no poseer  
conflictos de interés académicos,  
éticos, empresariales o comerciales**



*"Entre los pecados mayores  
que los hombres cometen,  
aunque algunos dicen que es  
la soberbia, yo digo que es el  
desagradecimiento".*

*Don Quijote*

# Agradecimiento



**Al Profesor Emérito**

**Dr. Pablo M. Bazerque**

**Maestro de Farmacología y de  
Metodología de la Investigación, por sus  
permanentes, enseñanzas, formación,  
apoyo y confianza.**

**Agradecimiento  
(In memoriam)**

**Al Dr. Héctor Boffi Boggero (1921-2001).  
Maestro de Epidemiología.**

**(Aprendimos con él que lo más importante es “mirar” los problemas desde dos situaciones opuestas y que el sentido común es el valor más importante para acceder a la Epidemiología).**

**Al Prof. Dr. José Tessler,  
Maestro de Bioestadística.**

**(Aprendimos con él que pensar es un privilegio humano que va mucho más allá de lo que se piensa).**



**“El intelecto humano... se  
conmueve y se excita más por  
las afirmaciones que por las  
negaciones”.**

**Sir Francis Bacon, 1621**

**Sir Francis Bacon (UK): 1561-1626.  
Fundador del Empirismo en la Ciencia.**

# **Las preguntas que guiarán esta ponencia**

- 1.- ¿Qué tipo de estudio y qué variables se deben utilizar para cumplir el objetivo del estudio?**
- 2.- ¿Cuántas observaciones debemos realizar (tamaño muestral)?**
- 3.- ¿El test estadístico utilizado es el adecuado (cuando se lee un “paper”) o qué test seleccionar (cuando se debe hacer una investigación)?**
- 4.- ¿Siempre es necesario utilizar el análisis multivariado?**
- 5.- ¿La discusión siempre debe seguir el orden de los resultados?**

**Una breve Introducción.**

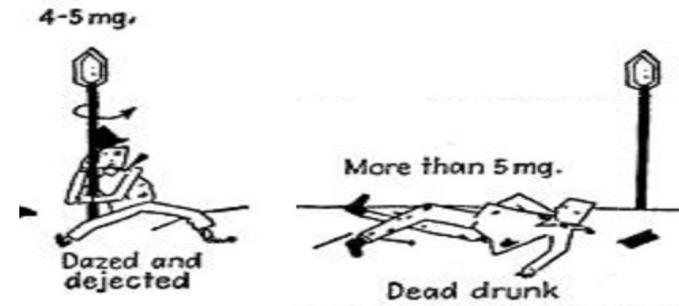
# Algunos conceptos a tener en cuenta

Algunos usan la estadística como los ebrios utilizan los faroles: más para apoyo que para iluminación. Andrew Lang (Escritor escocés. Escocia, 1844-1912)

Si Ud. tortura los datos lo suficiente, ellos confesarán.

Ronald Coase (Economista británico. Premio Nobel de Economía en 1991. 1910-2013)

Si los datos “hablan por sí mismos”, ¡No los interrumpa!!!  
(No haga interpretaciones más allá de lo que dicen).



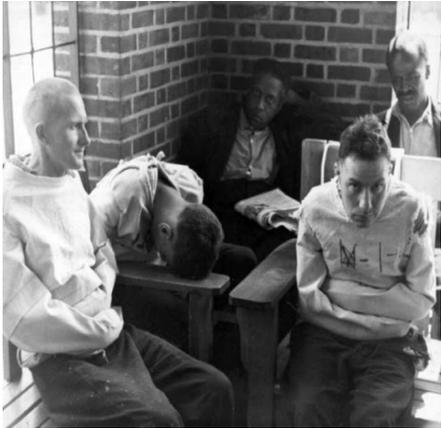
# Nuestra relación con “lo numérico”



<http://thales.cica.es/xvceam/?q=system/files/image001.png>  
<http://samuelsanchez.blogia.com/upload/20081101140758-matematica1.jpg>  
[http://4.bp.blogspot.com/-; UpQiiO9heSM/T2vJ5CKQJui/AAAAAAAAABU/\\_EwqriEiCt4/s1600/grito\\_desesperado.jpg](http://4.bp.blogspot.com/-; UpQiiO9heSM/T2vJ5CKQJui/AAAAAAAAABU/_EwqriEiCt4/s1600/grito_desesperado.jpg)  
[statics.cuantocabron.com/ccs/2014/01/CC\\_2274775\\_grumpy\\_cat\\_odio\\_a\\_primera\\_vista.jpg](http://statics.cuantocabron.com/ccs/2014/01/CC_2274775_grumpy_cat_odio_a_primera_vista.jpg)  
<http://web.educastur.princast.es/ies/rosarioa/archivos/matematicas/enrique.jpg>

# ¿Los números gobiernan el mundo?

Si la mayoría de los humanos fueran  
esquizofrénicos  
¿quiénes estarían internados?  
(Concepción frecuentista)



**El Investigador y sus Ídolos**

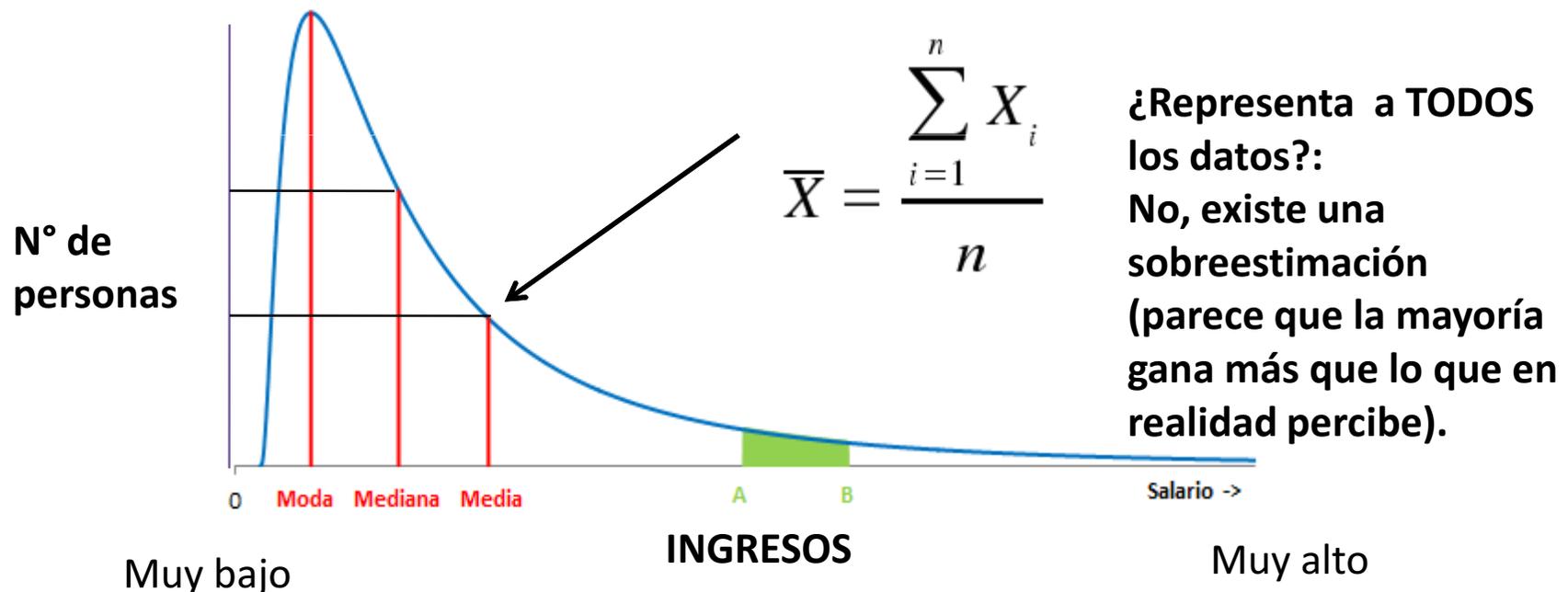
**Idolatría por desconocimiento**

**El misterio y la magia  
de los números son  
invasivos,  
embotan al sentido  
común.**

# “La veneración ciega de la Media Aritmética”

La Media Aritmética también mal llamada Promedio\* “suponemos” que nos “representa a un conjunto de datos”. Pues bien, no siempre es así.

\* (los promedios son una familia: Media Aritmética, Mediana, Media Geométrica, Media Armónica, Modo)



**En este caso la Mediana es el mejor “Estimador” (representa mejor la distribución de los datos).**

## Ejemplo de irracionalidad

Género
1
1
1
1
1
1
1
1
1
1
1
1
2
2
2
2
2
2
2
2
2
2
2

	Género
Tamanho da amostra =	20
Mínimo	1
Máximo	2
Amplitude Total	1
Mediana	1.5
Primeiro Quartil (25%)	1
Terceiro Quartil (75%)	2
Desvio Interquartílico	1
<b>Média Aritmética</b>	<b>1.5</b>
Variância	0.2632
Desvio Padrão	0.513
Erro Padrão	0.1147
Coefficiente de Variação	34.20%
Assimetria (g1)	0
Curtose (g2)	-2.2353
Média Harmônica =	1.3333
N (média harmônica) =	20
Média Geométrica =	1.4142
N (média geométrica) =	20
Variância (geom.) =	1.0564
Desvio Padrão (geom.) =	1.427

**Codificación:**

**1: Femenino**

**2: Maculino**

**Es una Media**

**Aritmética**

**perfectamente**

**calculada, pero**

**¿significa algo?**

**¿Qué es 1.5 de**

**género?**

# **Es fundamental entender\* para poder interpretar \*\***

**En cierta oportunidad un médico, extrajo del “Libro de Guardia” datos de todos los accidentados.**

**Observó lo siguiente:**

**De 100 accidentados, 25 (25%) se hallaban alcoholizados y 75 (75%) se encontraban sobrios.**

**Con esta “evidencia” hay más probabilidad de tener accidentes estando sobrio que estando alcoholizado (75% vs 25%).**

**Esta conclusión claramente va contra el sentido común. ¿Qué ocurrió?..... Una Falacia: Engaño, fraude o mentira con que se intenta dañar a alguien (DRAEL).**

**\*Entender: 1. tr. Tener idea clara de las cosas.**

**\*\*Interpretar: 3. tr. Explicar acciones, dichos o sucesos que pueden ser entendidos de diferentes modos.**

Otro colega, aplicando lo que había aprendido en un Curso de Epidemiología, construyó la siguiente Tabla de Contingencia (posibilidad que algo suceda o no suceda).

	Accidente	No Accidente	Total de Filas
Sobrios	75	675	750
Ebrios	25	75	100
Total de Columna	100	750	850

De 750 sobrios hubo 75 accidentados: 10%

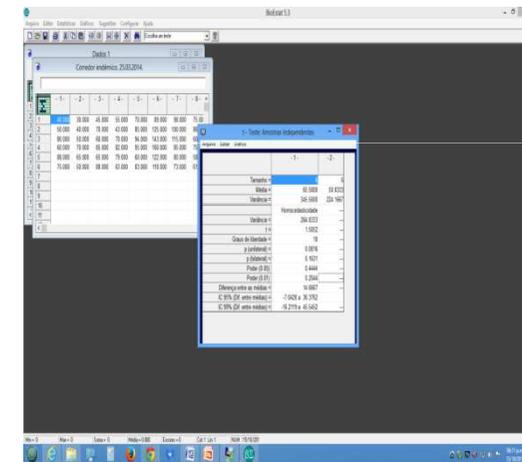
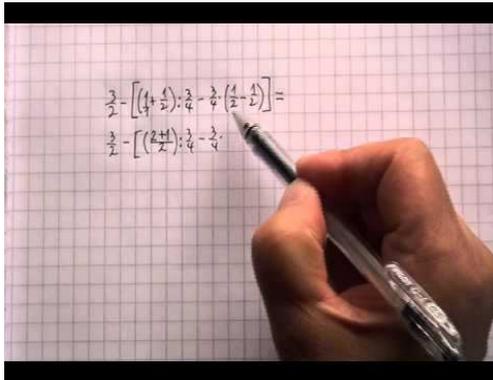
De 100 ebrios hubo: 25 accidentados: 25%

Test Binomial para diferencia de Proporciones:  $p < 0.0001$ .

Y esto está acorde con el sentido común.

Simplemente se utilizó un “Grupo Control”.

# El cambio conceptual



**De la “Estadística Resolutiva”**  
(Nos quedaban pocas ganas para la interpretación)

**A la “Estadística Interpretativa”**  
¿Cómo asociar los valores “p” o IC 95% al estudio que hemos realizado?.

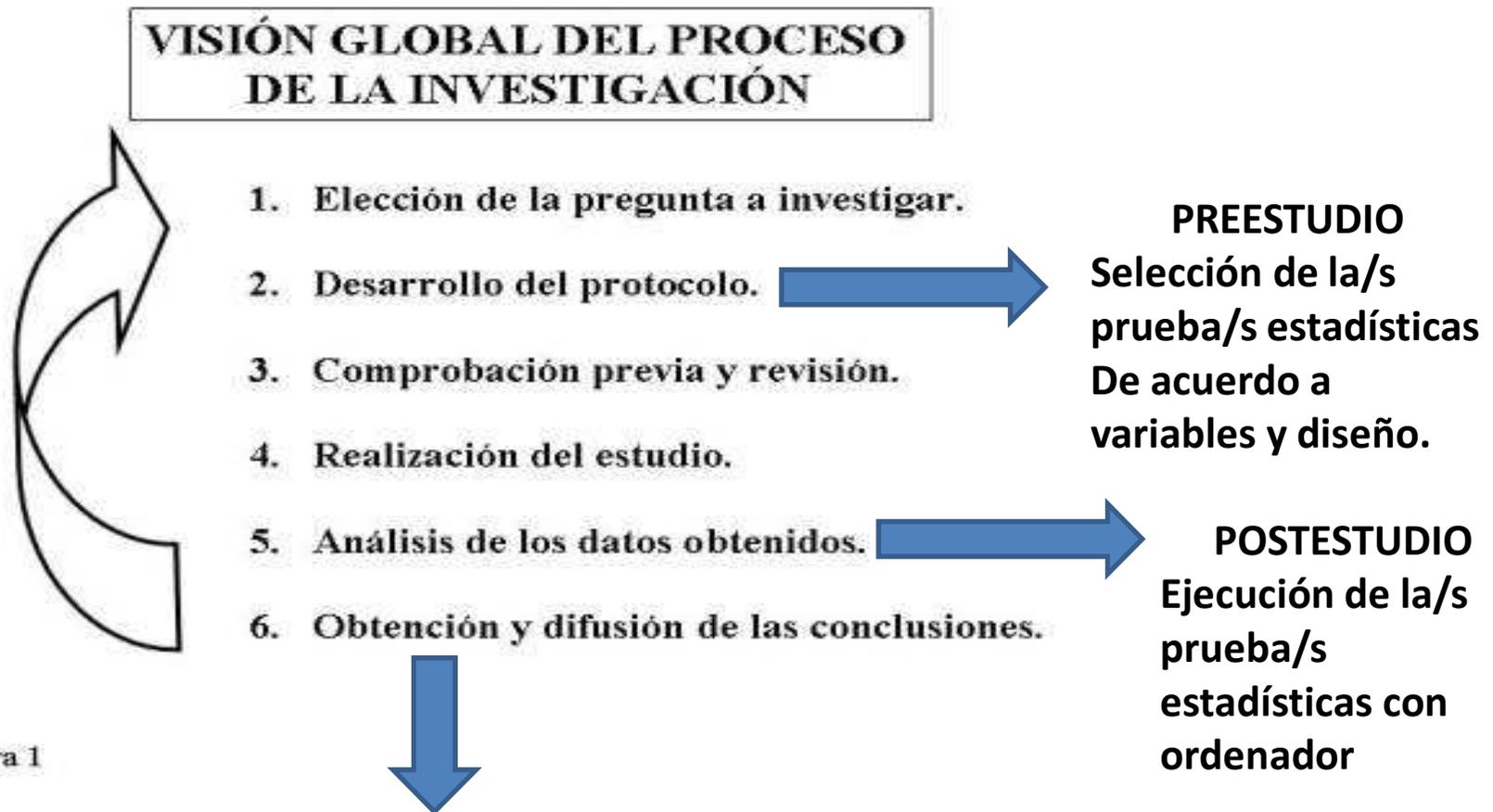


Figura 1

**“Traducción” de “lo numérico” a conceptos médicos.  
La tarea más gratificante del/los Investigador/es.**

**Sir Ronald Fisher\* uno de los padres de la estadística moderna (UK, 1890-1962) dijo:**



Indian Statistical Congress,  
Sankhya, c. 1938

**“Consultar al Estadístico una vez finalizado el trabajo de investigación, es como llevar un cadáver a la morgue, el profesional solamente podrá decir de qué murió ese estudio”.**



**\* Entre otros tantos aportes suyos figuran el ANOVA y el concepto de Hipótesis Nula.**



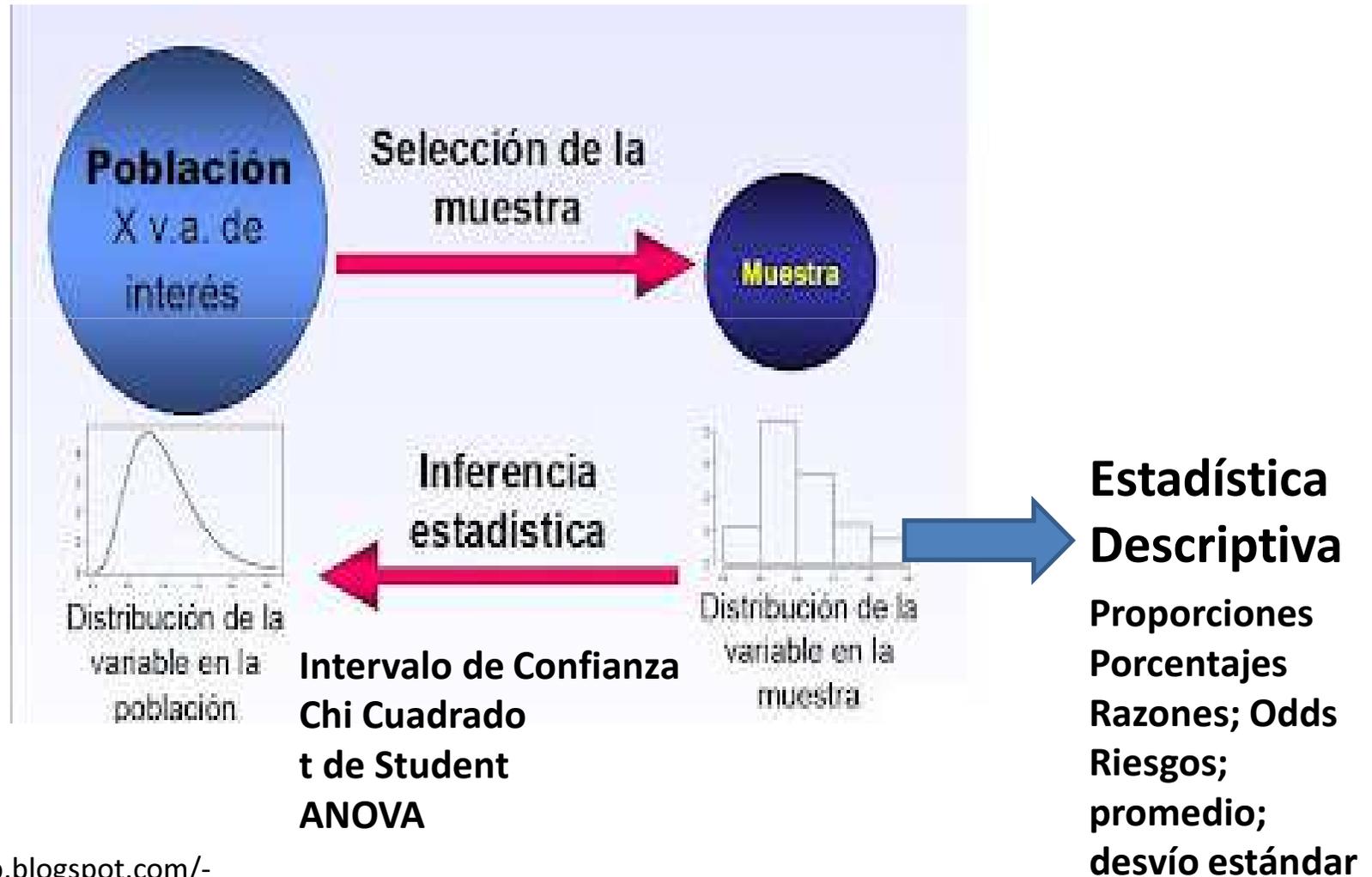
**Nuestros pacientes, ante todo, son personas que además, tienen diferentes aspectos representados por VALORES NUMÉRICOS (Peso, Estatura, Glucemia, Equilibrio ácido-base, etc).**

**Por estas últimas razones es que necesitamos una herramienta que, amigablemente, nos ayude a transitar por esa “jungla numérica”.**

# ¿Son muy diferentes?

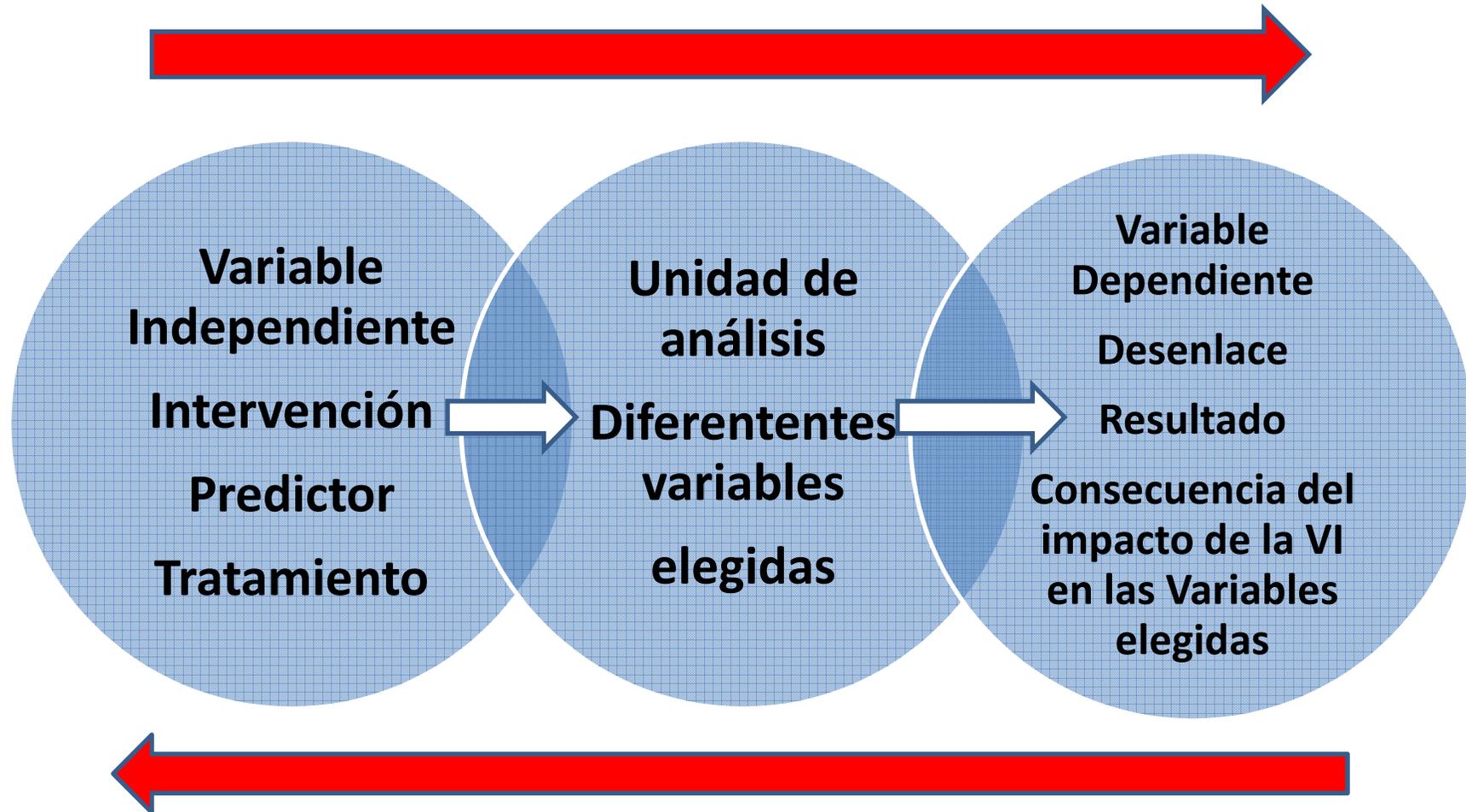
<b>Método clínico</b>	<b>Método estadístico Sus pilares</b>
<b>1.- Paciente</b>	<b>1.- Población y Muestra</b>
<b>2.- Signos y Síntomas</b>	<b>2.- Variables</b>
<b>3. Diagnóstico diferencial</b>	<b>3.- Probabilidad</b>
<b>4.-Diagnóstico presuntivo Estudios-Resultados</b>	<b>4.- Hipótesis (H0; H1)</b>
<b>5.- Comparación con patrones de enfermedad</b>	<b>5.- Inferencia o Confrontación (Comparación) de hipótesis</b>

# Modelo General



**1.- Definir el tipo de estudio y las variables a tener en cuenta para cumplir el objetivo del estudio.**

# Anátomo fisiología de toda Hipótesis



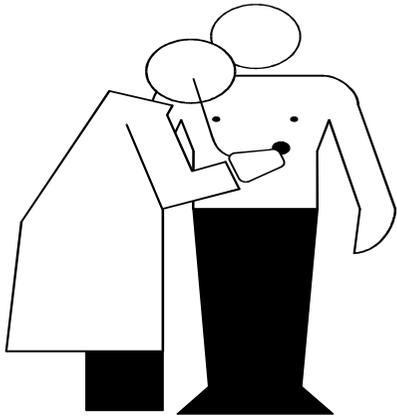
Las flechas rojas indican el sentido de lo que se busca.

## **1.a. Tipo de estudios clínicos**

**Estudio clínico es toda investigación que tiene como unidades de análisis a seres humanos.**

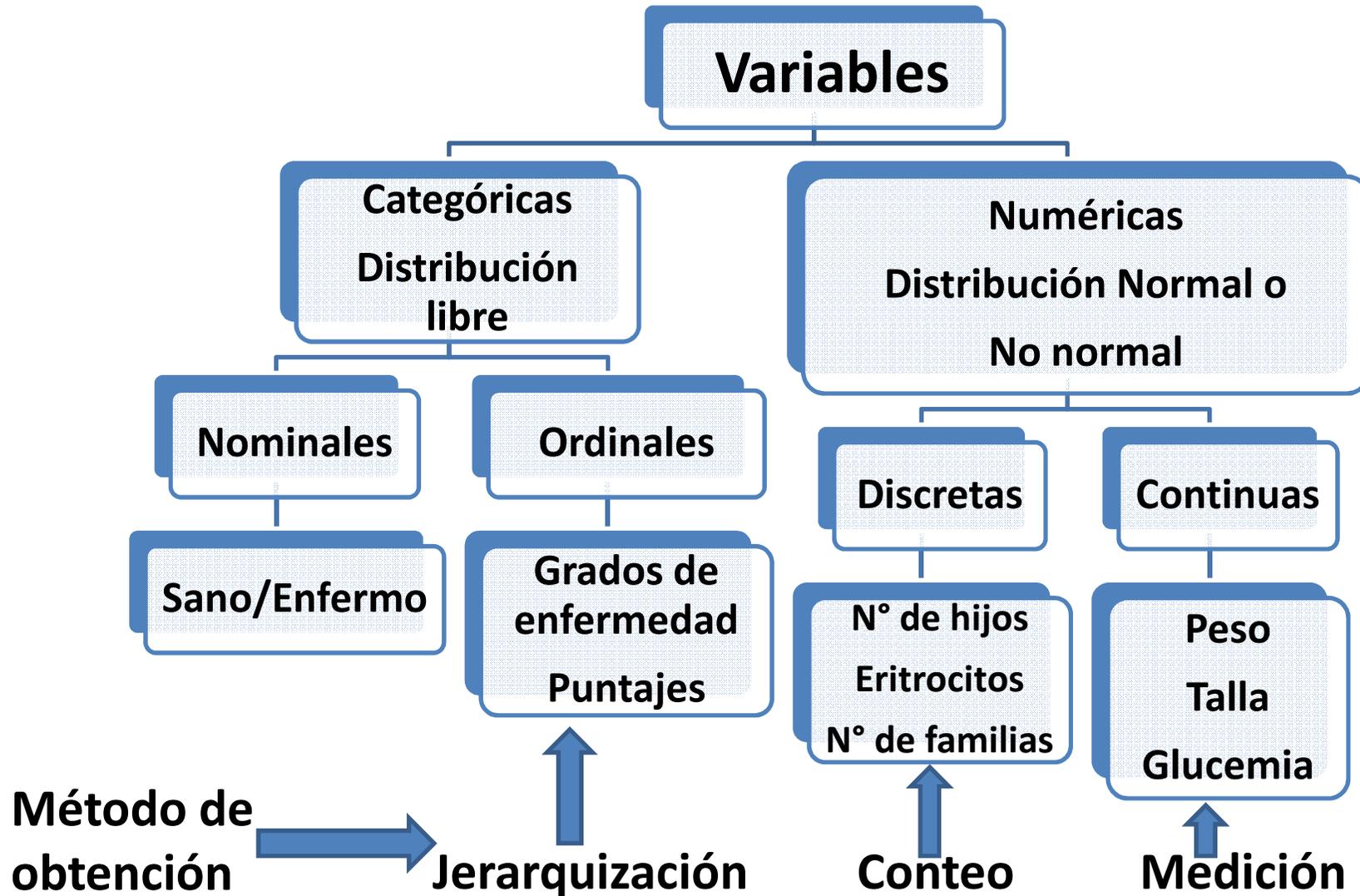
**No debe confundirse con Ensayo Clínico que es un tipo particular de Estudio Clínico.**

# Mapa conceptual de los Estudios Clínicos



# **1. b. En cuanto a las Variables**

# Mapa conceptual de las Variables aleatorias (V.a.)



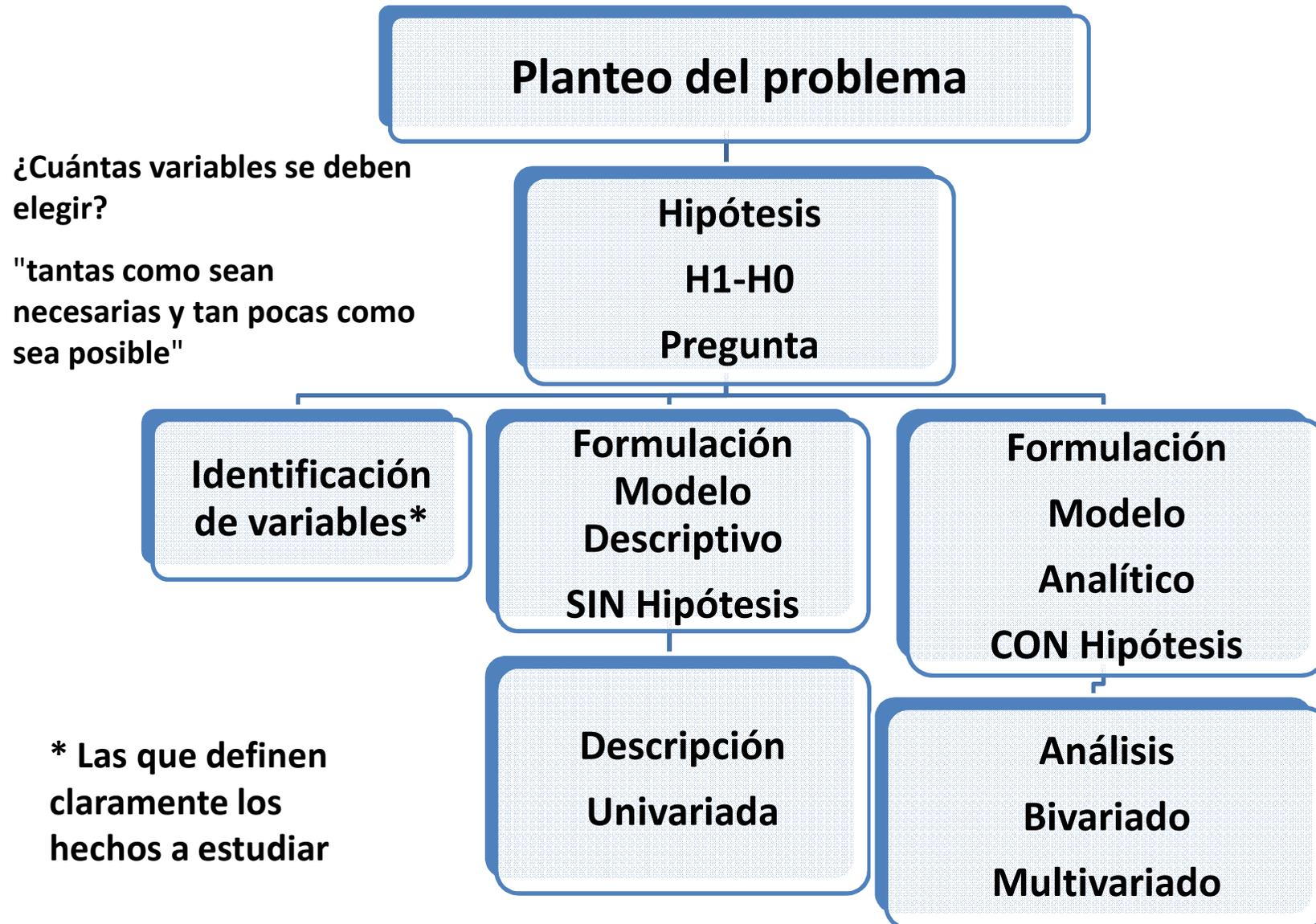
**Para abordar a las variables, primero  
debemos establecer que son “actores”  
en determinado “escenario”**



**Escenario  
Descriptivo**

**Escenario  
Analítico**

# Etapas metodológicas a partir del planteo del problema



**¿Cómo expresar sintéticamente  
los resultados descriptivos?**

# Resumen y graficación de Variables

Estimador	Variables			
	Categorías		Numéricas	
	Nominal	Ordinal	Discreta	Continua
n	SÍ	SÍ	SÍ	SÍ
Razón/proporción; %	SÍ	SÍ	En el caso de "Categorizar" a las variables numéricas	
Mediana; Rango: Cuantiles	NO	X	X En caso de No Normalidad	X En caso de No Normalidad
Media aritmética Variancia; Desvío estándar	NO	NO	SÍ Sólo si hay Normalidad	SÍ Sólo si hay Normalidad
Tablas	SÍ	SÍ	SÍ	SÍ
Gráfico Barras separadas	SÍ	SÍ	SÍ	NO
Gráfico Sectorial	SÍ	SÍ	En el caso de "Categorizar" a las variables numéricas	
Gráfico Caja y bigotes	NO	Sí, con: Mediana y Cuantiles	Sí, con: Media aritmética y Desvío estándar	Sí, con: Media aritmética y Desvío estándar
Histograma	NO	NO	NO	SÍ
Gráfico Cartesiano ortogonal	Toda vez que se desee presentar una supuesta Asociación entre una variable Independiente (Abscisa) y una Variable Dependiente (ordenada)			

## Datos resumen (BioEstat 5.3)

Variable Cuantitativa	
Tamaño muestral	15
Mínimo	355
Máximo	15404
	5440.06
Média Aritmética	67
	1.48E+0
Variância	7
	3841.32
Desvio Estándar	39
	991.825
Error Estándar	6
Coeficiente de Variación	70.61%

Variable Cualitativa Discreta	
Tamaño muestral	15
Mínimo	355
Máximo	15404
Mediana	4560
Primer Cuartil (25%)	3123.5
Tercer Cuartil (75%)	6745
Desvio Intercuartílico	3621.5
Média Geométrica =	4042.40
Variância (geom.) =	1.4508
Desvio Estándar (geom.) =	2.5234

**Precaución: los paquetes estadísticos  
brindan TODOS los datos resumen, el  
Investigador debe seleccionarlos**

**2.- ¿Cuántas observaciones  
debemos realizar (tamaño  
muestral) para que nuestros  
resultados tengan validez?**

# El tamaño muestral no surge de una compleja y oculta operación alquímica...



<https://www.puntmagicbarcelona.com/sites/default/files/puntmagicbarcelona/El%20Oro%20de%20los%20Alquimistas.jpg>

# Mitología del tamaño muestral



“Cuanto más grande la muestra, más confiables son los resultados”.

Con este reduccionismo se corre el **peligro** de incrementar los Falsos Positivos (Sobredimensionar los resultados)

“El hecho es tan común que bastan unos pocos casos para demostrarlo”

Con este reduccionismo se corre el **peligro** de incrementar los Falsos Negativos (Rechazar Hipótesis verdaderas)

**Se deben tener presentes  
3 aspectos  
que conoce el Investigador**

- a) Variabilidad de la/s variable/s en estudio.**
- b) Cantidad de Falsos Positivos ( $\alpha$ ) y Falsos Negativos ( $\beta$ ) que se está dispuesto a aceptar.**
- c)Cuál es la mínima diferencia a detectar.**

# Modelo general para la estimación del tamaño muestral

Inversa de Errores  $\alpha$  y  $\beta$

Variabilidad de la/s variable/s en estudio

$$n = \frac{(1/\alpha + \beta) * S^2}{\Delta^2}$$

Diferencia a detectar

$\Delta^2$

Esto significa que el n muestral es:  
Directamente proporcional a la Variabilidad e  
Indirectamente proporcional a la diferencia a detectar y a los Falsos + y Falsos – esperables.

# Ejemplos de obtención del tamaño muestral

**Variable Cuantitativa; Diferencia**

**De medias de 2 tratamientos**

**Variabilidad: Desvío Estándar en cada grupo es de 0.9**

**Error  $\alpha = 0.05$  ; Error  $\beta = 0.20$ , por lo que  $1 - \beta = 0.80$**

**La Diferencia a detectar es de 0.1 (10%)**

**Difference of means to be detected: 0.1**

**Expected standard deviation within groups: 0.9**

**Sample size: 624 for each sample**

**624 voluntarios recibirán A y otros 624 voluntarios recibirán B**

**Variable Cualitativa; Para diferencia de Proporciones (%)**

**Diferencia de un 10% en la eficacia de dos analgésicos por ejemplo (80% vs 70%)**

**Error  $\alpha = 0.05$ ; Error  $\beta = 0.20$ , por lo que  $1 - \beta = 0.80$**

**Sample size: 313 for each sample:**

**Para detectar una diferencia del 10%, se necesitarán 313 voluntarios en cada tratamiento.**

## En síntesis: relación entre diferencia a detectar y tamaño muestral

	Variable Cuantitativa		Variable Cualitativa	
<b>Diferencia a detectar</b>	<b>10%</b>	<b>20%</b>	<b>10%</b>	<b>5%</b>
<b>Desvío Estándar</b>	0.9	0.9	-----	-----
<b><math>\alpha</math></b>	0.05	0.05	0.05	0.05
<b><math>1 - \beta</math></b>	0.80	0.80	0.80	0.80
<b>n</b>	<b>624</b>	<b>157</b>	<b>313</b>	<b>1134</b>

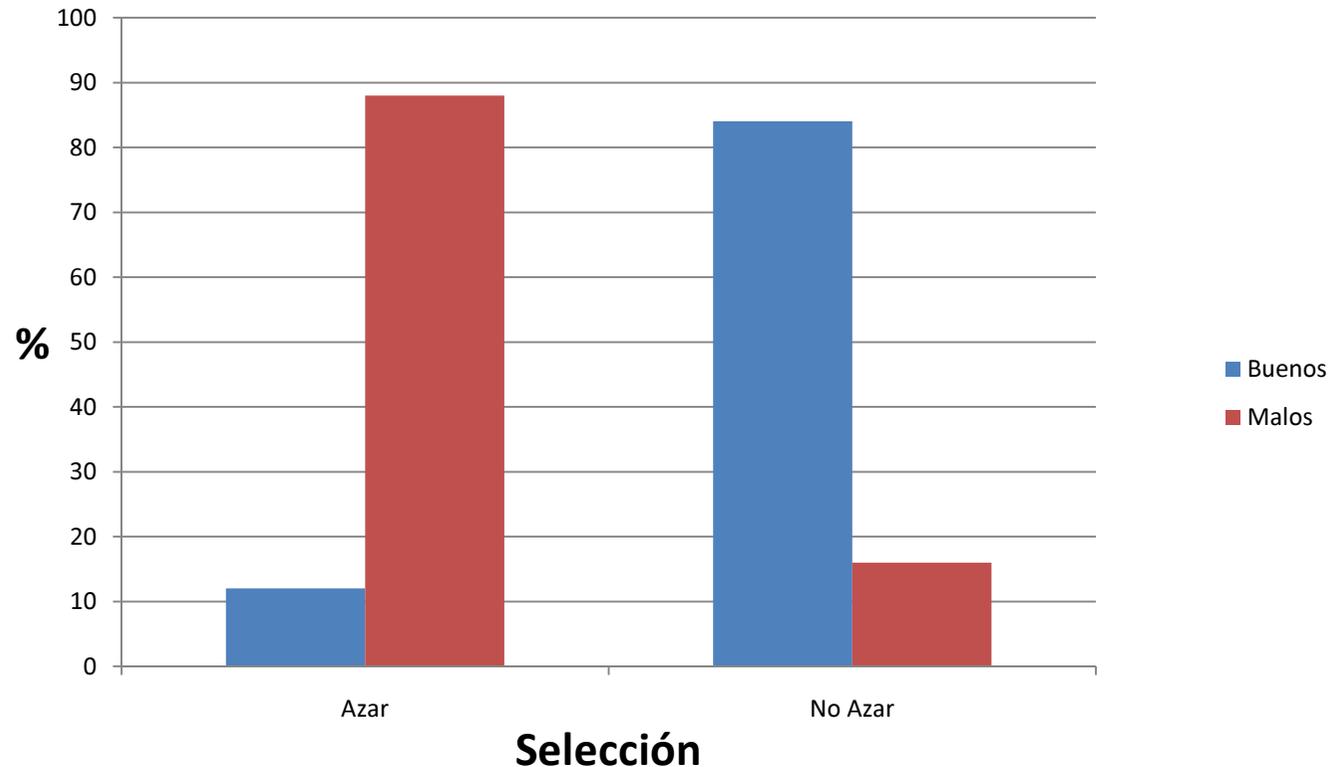
**Paradigma:**

**A menor diferencia a detectar, MAYOR n.**

# Muestra Apropiada

## El fundamento de la “Randomización”

Diferencias en los resultados  
de acuerdo a la selección



**Conclusión: las muestras no tomadas al azar tienen mayor probabilidad de sobreestimar los buenos resultados (Falsos Positivos).**

Modelo básico para la interpretación de resultados.

$p > 0.05$

No existen diferencias entre grupos

Diseño Paralelo

ANTES DEL ESTUDIO

OBSERVACIÓN (MEDICIÓN DE DATOS)

OBSERVACIÓN (MEDICIÓN DE DATOS)

$p > 0.05$ : no hay diferencia Antes Después  
 $p < 0.05$ : Existe diferencia entre Antes y Después

TRATAMIENTO CONTROL

TRATAMIENTO EXPERIMENTAL

$p > 0.05$ : no hay diferencia Antes-Después  
 $p < 0.05$ : Existe diferencia entre Antes y Después

OBSERVACIÓN (MEDICIÓN DE DATOS)

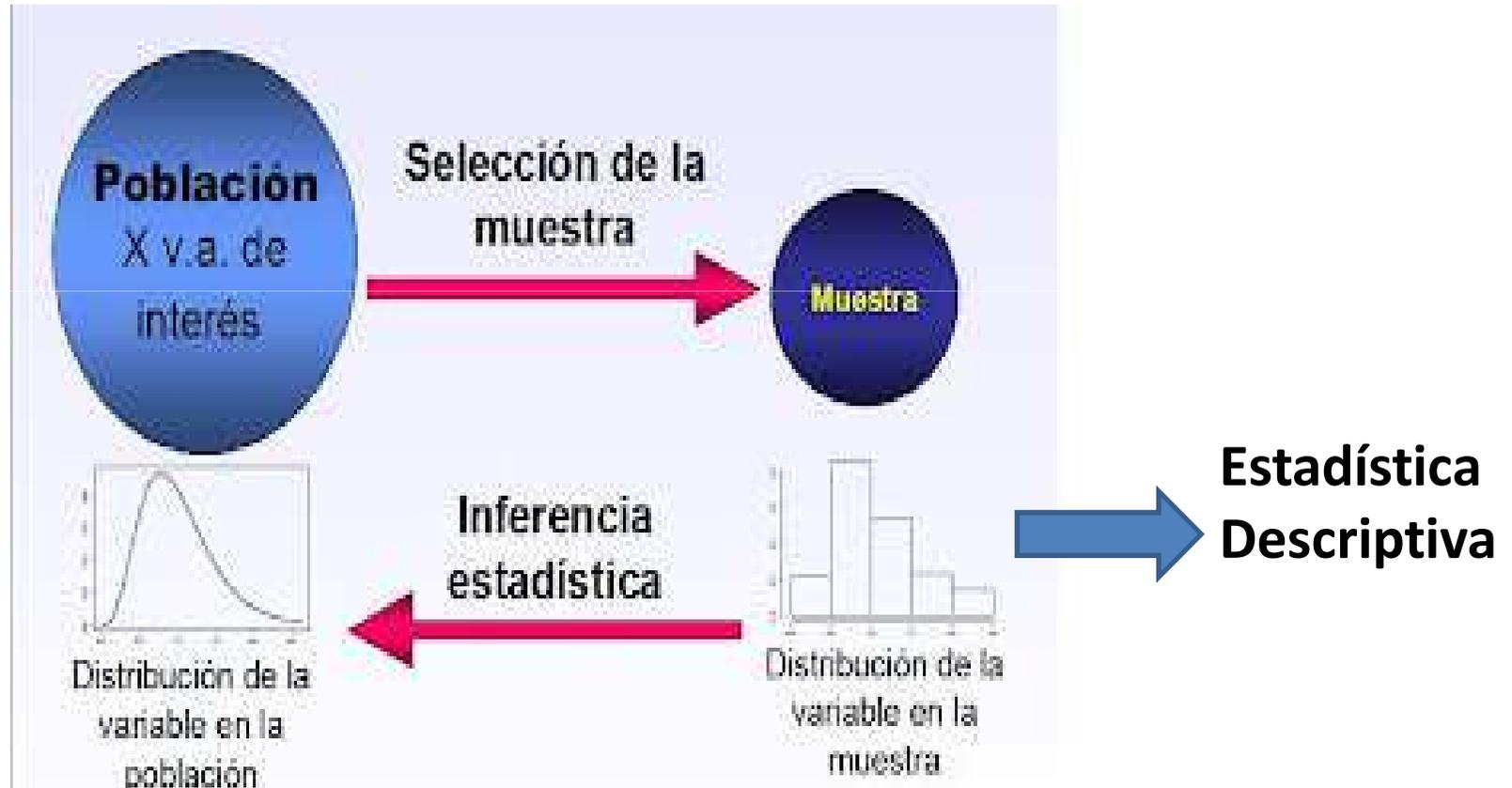
OBSERVACIÓN (MEDICIÓN DE DATOS)

$p > 0.05$ : no hay diferencias entre tratamientos  
 $p < 0.05$ : Existe diferencia no debida al azar entre tratamientos

AL FINALIZAR EL ESTUDIO

**3- ¿El test estadístico utilizado es el adecuado (cuando se lee un “paper”) o qué test seleccionar (cuando se debe hacer una investigación)?**

# Preguntas (y respuestas) sobre los Métodos estadísticos



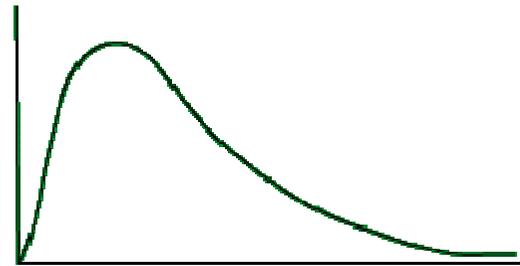
## NORMALIDAD BIOLÓGICA

Glucemia: 0.70-1.10 mg/dl

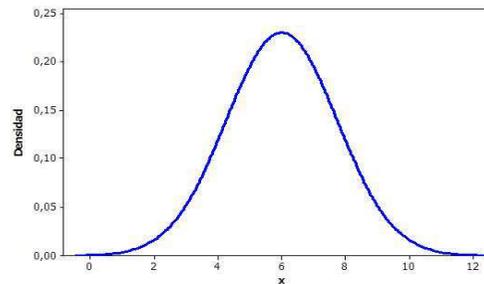
Eritrocitos: 4 millones

Urea: 10-40 mg/dL

Creatinina en sangre: 0,3-0,7 mg/dl



**ANORMALIDAD  
ESTADÍSTICA**



**NORMALIDAD  
ESTADÍSTICA**

## ANORMALIDAD BIOLÓGICA

Glucemia: 1.80 mg/dl

Eritrocitos: 3.5 millones

Urea: 60 mg/dL

Creatinina en sangre: 1 mg/dl



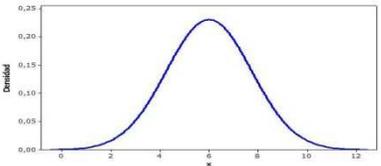
**ANORMALIDAD  
ESTADÍSTICA**

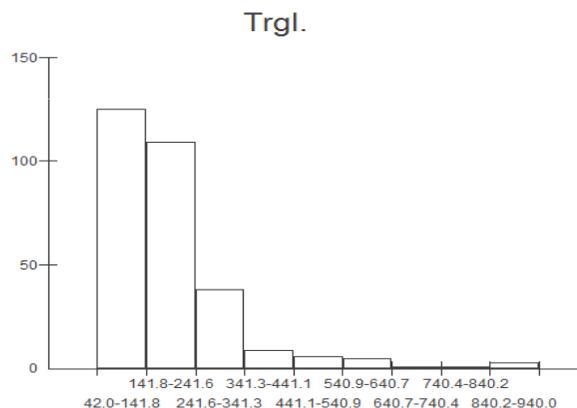
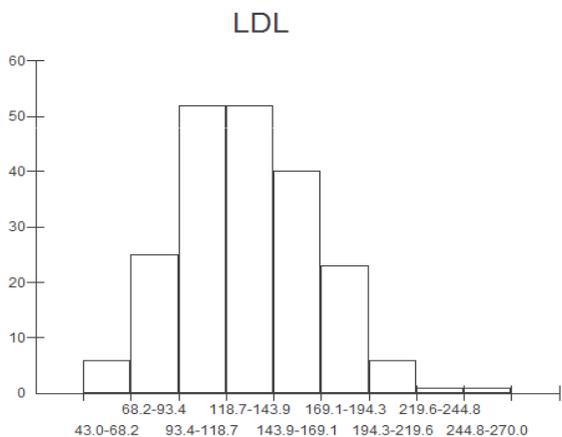
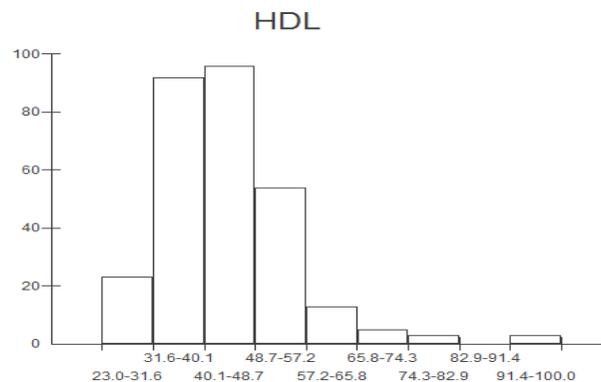
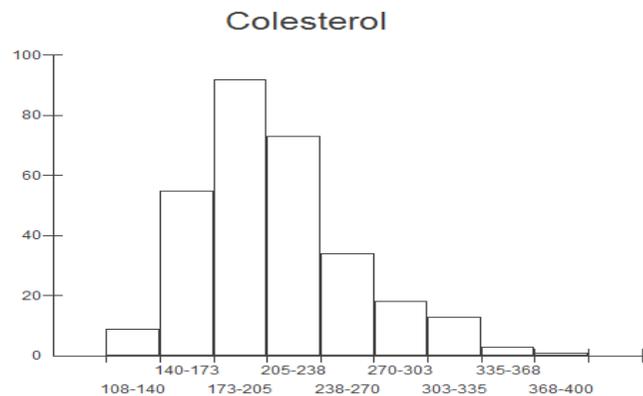
<http://www.ucv.cl/web/estadistica/imagenes/asimet1.gif>

<http://3.bp.blogspot.com/-wroZSsTHalo/UJecs8mEfxI/AAAAAAAAAAc/p1bYHUeAA4A/s1600/105542227.jpg>

[http://dc349.4shared.com/doc/aSVX4Z\\_1/preview\\_html\\_m57e836b3.png](http://dc349.4shared.com/doc/aSVX4Z_1/preview_html_m57e836b3.png)

# Las Normalidades no siempre son coincidentes

		Normalidad biológica 	
		SÍ	NO
Normalidad estadística 	SÍ	Normoglucemia	Hiperglucemia
	NO	Trigliceridemia	Enzimas hepáticas X 3; X4; X5, etc.



## Distribución de frecuencia de diversos lípidos sanguíneos en pacientes diabéticos

# Preguntas (y respuestas) sobre los Métodos (Pruebas) estadísticos

3.- ¿Qué nos dice un Método o Prueba o Test Estadístico?

Indica en qué magnitud ha  
intervenido el azar\* en el estudio.

**EVIDENCIA CIENTÍFICA:**

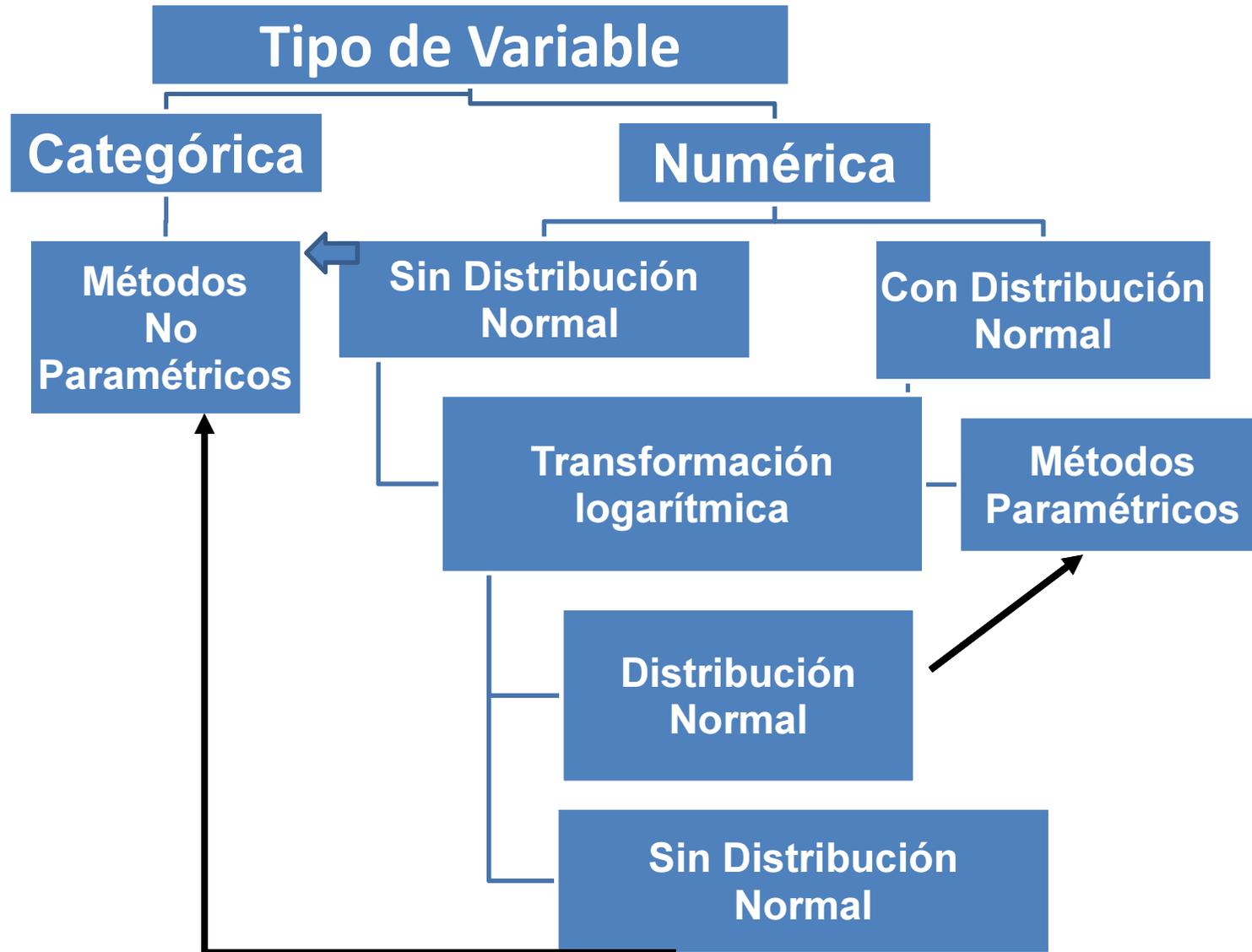
Probabilidad de intervención del azar  $< 5\%$



\* Probabilidad : cauntificación del azar.

**¿Qué se debe tener en cuenta para  
la selección de la Prueba  
Estadística?**

# Algoritmo para la elección del Método Estadístico



En la transparencia siguiente se brinda un listado con las pruebas estadísticas más frecuentemente utilizadas en Medicina

# Síntesis de las Pruebas Estadística más frecuentemente utilizadas en Medicina

## Variables Categóricas y cualquier Variable sin distribución normal

- 1.- Diferencia de proporciones (una muestra y dos muestras).
- 2.- Test Chi Cuadrado (Asociación). Diversas aplicaciones.
- 3.- Test de Wilcoxon (2 muestras apareadas o Antes-Después).
- 4.- Test de Mann Whitney (2 muestras independientes).
- 5.- Test de Kruskal-Wallis (más de 2 muestras independientes).
- 6.- Test de Friedman (más de dos mediciones repetidas).
- 7.- Correlación No paramétrica de Spearman.

## Variables Numéricas con Distribución Normal

- 1.- Test t de Student para 1 muestra.
- 2.- Test t de Student para dos muestras apareadas (Antes-Después).
- 3.- Test t de Student para dos muestras independientes.
- 4.- Aanálisis de Variancia de una vía o de una fuente de variación.
- 5.- Aanálisis de Variancia de más de una vía o fuente de variación.
- 6.- Análisis de Regresión, Correlación y Determinación

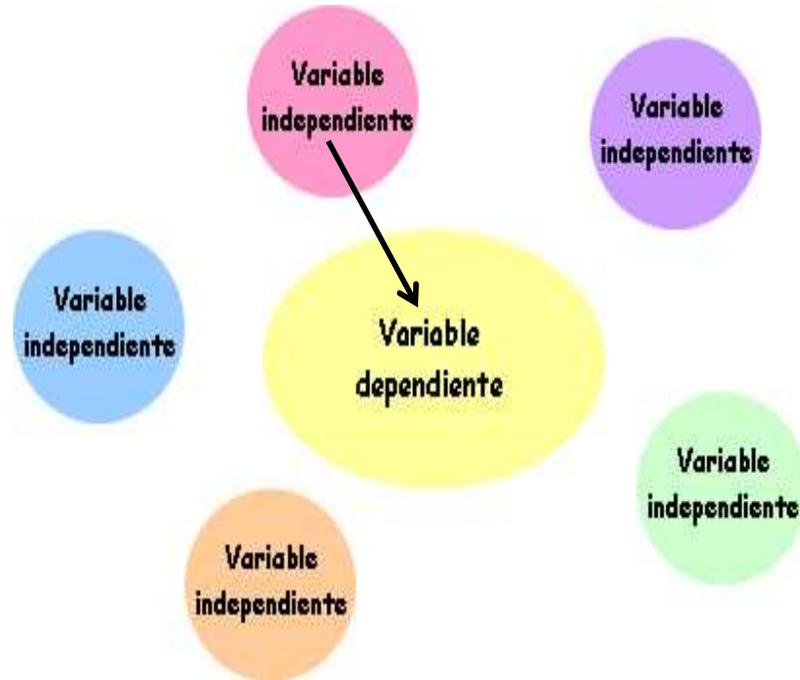
## 4.- ¿Siempre es necesario realizar el análisis multivariado?



No, el análisis multivariado se aplica cuando se parte de la Hipótesis Alternativa que diversas variables independientes ( $X_1, X_2, X_3$ ) se relacionan con una o más variables Dependientes ( $Y$ ).

# Análisis Bivariado

Explora asociación entre:  
**UNA VARIABLE INDEPENDIENTE**  
Y  
**UNA VARIABLE DEPENDIENTE**



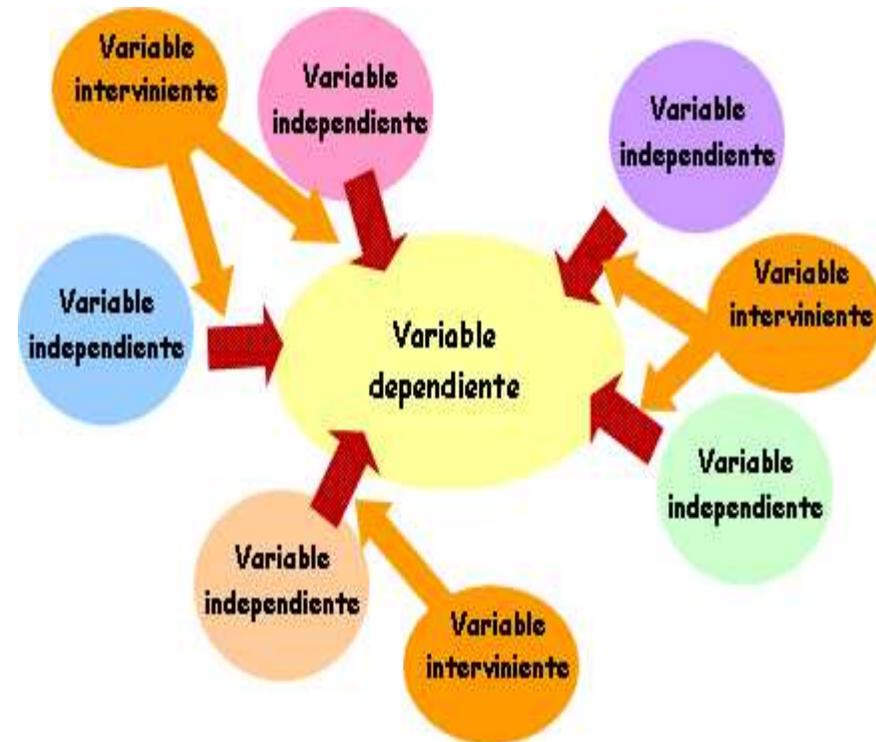
<http://www.rena.edu.ve/cuartaEtapa/metodologia/Imagenes/img1T6c.jpg>  
<http://www.rena.edu.ve/cuartaEtapa/metodologia/Imagenes/img3T6c.jpg>

# Análisis Multivariado

(AMV)

Explora asociaciones entre:

- Varias variables independiente y 1 variable dependiente.
- Varias variables independientes y varias variables dependientes.



El AMV, evalúa el “peso” de cada factor de riesgo, interviniendo todos los otros conjuntamente.

# Mapa conceptual de los más frecuentes métodos de Regresión utilizados en Medicina

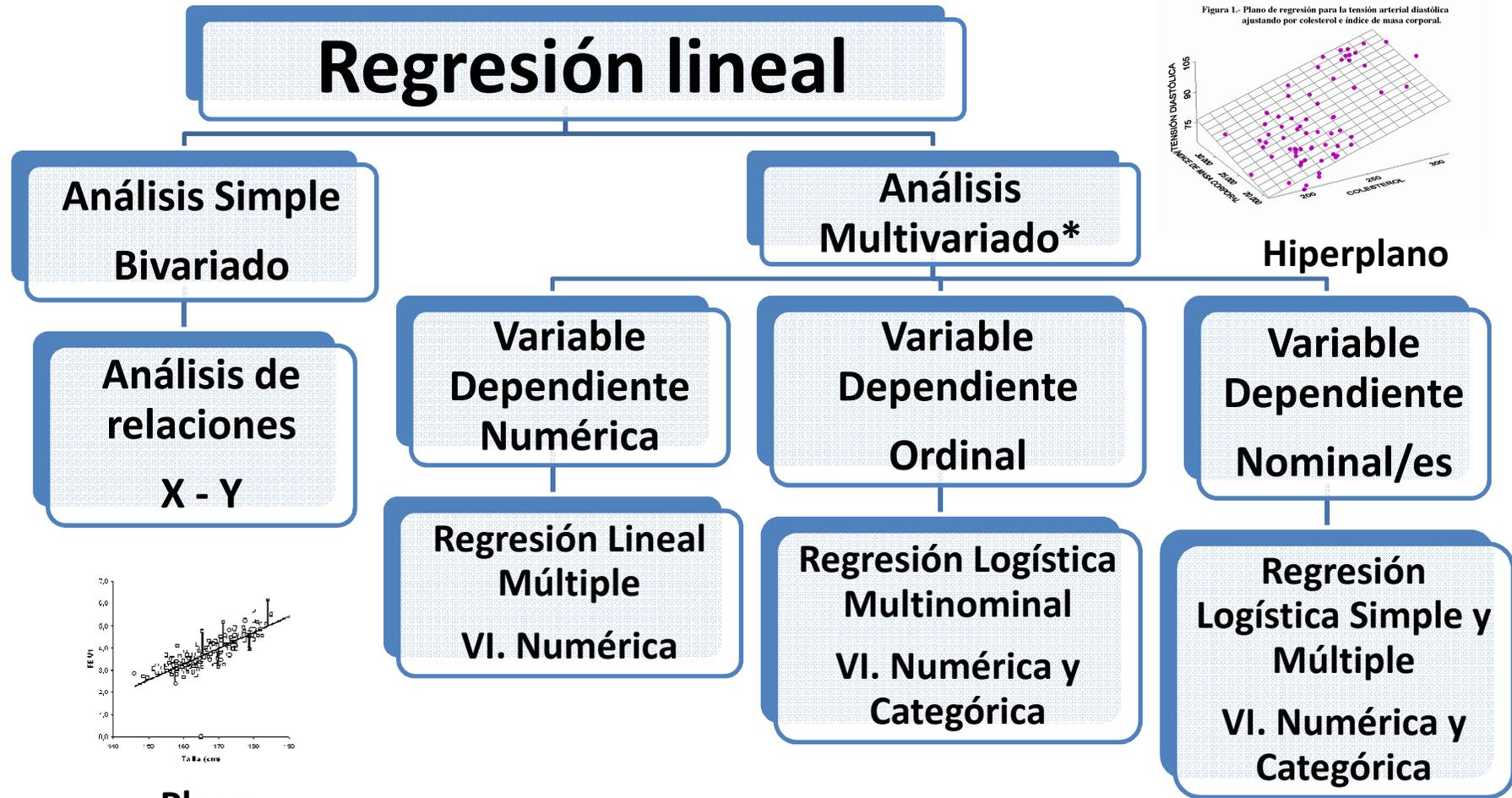
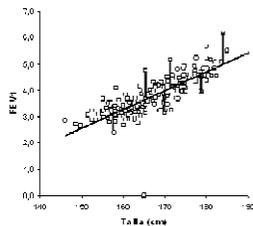
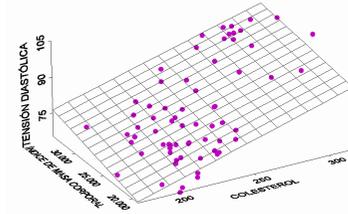


Figura 1- Plano de regresión para la tensión arterial diastólica ajustando por colesterol e índice de masa corporal.



\* El Análisis Multivariado evalúa más de 1 variable independiente y 1 o más variables dependientes.

<http://epidemiologiamolecular.com/wp-content/uploads/2009/05/clip-image008.gif>  
[https://www.fisterra.com/mbe/investiga/regre\\_lineal\\_multi/images/Image167.gif](https://www.fisterra.com/mbe/investiga/regre_lineal_multi/images/Image167.gif)

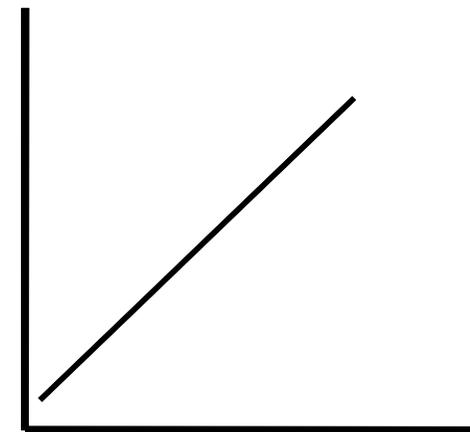
**En el análisis multivariado  
¿sólo se utilizan las variables que  
resulten significativas en el análisis  
univariado posterior? o todas?.**

# Ejemplo teórico de resultados de una regresión logística para evaluar la influencia de diferentes factores de riesgo en el infarto de miocardio

Variable Independiente	p
Exposición	
Edad	< 0001
IMC	<0.000001
Sedentarismo	<0.0001
Consumo de café	<0.05
Hábito de fumar	<0.00001
Nivel educacional	>0.05
Hipercolesterolemia	<0.00001
Hipetrigliceridemia	<0.001
Hiperglucemia	<0.05
Estatura	>0.05
Estado civil	>0.05
Consumo de agua	>0.05

Las variables resaltadas se evalúan individualmente mediante la Regresión Lineal Simple.

Riesgo



IMC  
Solamente se tienen en cuenta las que resultan significativas en el análisis univariado.

**5.- ¿La discusión siempre debe seguir el orden de los resultados?**

# Ordenamiento de la Discusión

1°

- En primer lugar comparar los resultados obtenidos con los de otros autores (coincidentes y no coincidentes).

2°

- Luego fundamentar los resultados obtenidos en una forma más literaria que numérica siguiendo un cierto orden establecido por los autores (que bien puede ser el de los resultados).

3°

- Finalmente, si es posible, dejar planteada una hipótesis (o al menos sus elementos) para futuros estudios.

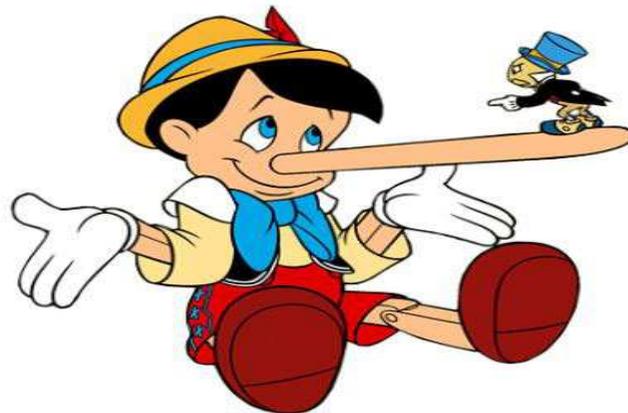
**\*Al final de la presentación se ofrece una guía para la redacción del “paper”.**

# Aspectos a tener en cuenta para no ser sorprendidos en nuestra buena fe.

Ideas tomadas de: Cómo mentir con estadísticas. Darrell Huff

<http://www.saladeprensa.org/art794.htm>

## CAPÍTULO 9: Cómo estadistiquear (Abuso de la Estadística)



<http://jumastorga.files.wordpress.com/2008/12/mentiras-seo1.jpg>



# Estadisticulación ¿qué es?

**Especulación estadística:**

**Informar**

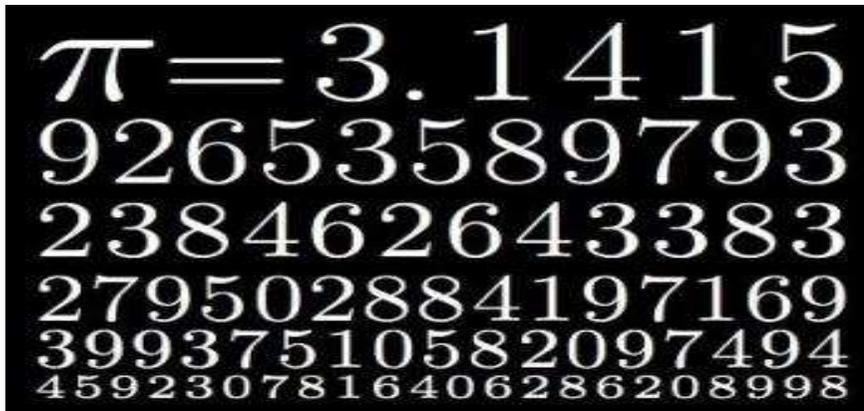
**engañosamente, manipuland**

**o\* los datos.**

**Manipular: 3. tr. Intervenir con medios hábiles y, a veces, arteros (mañoso, astuto), en la política, en el mercado, en la información, etc., con distorsión de la verdad o la justicia, y al servicio de intereses particulares.**

## 2 herramientas básicas de la “Estadisticulación”

1.- Decimales (cuantos más,  
“mejor”) pretende dar más  
“consistencia”\* a datos  
fraudulentos.



$\pi = 3.1415$   
92653589793  
238462643383  
279502884197169  
39937510582097494  
4592307816406286208998

= 3,1416

Por alguna razón surgió  
el “redondeo”.

\*consistencia. (De *consistente*). 1. f. Duración, estabilidad, solidez.

## 2.- Herramientas básicas de la Estadística: Porcentajes (la preferida)



Utilización de %	Riesgos
Aplicado a pocas observaciones	Pueden resultar engañosos.
Expresados con más de un decimal	Absurdo y/o Fraudulento*
“Puntos” porcentuales	Interpretación ambigua, por lo que se puede utilizar la que más convenga**

**\*\*Si un valor varía de 2 a 4, se puede “interpretar”:** ¡Qué poco!  
**Aumentó sólo 2 puntos ó**  
**¡Cuánto! Se incrementó el 100%.**

\*absurdo, da. (Del lat. *absurdus*). 1. adj. Contrario y opuesto a la razón; que no tiene sentido.

\*fraudulento, ta. (Del lat. *fraudulentus*). 1. adj. Engañoso, falaz.



# Debemos estar prevenidos...



**Inmunización activa contra la “Estadisticulación”. Debe figurar en el calendario de todo investigador.**



# **Vacuna “Prevenestat” (inmunización activa frente a los engaños numéricos)**



**“Anti-  
Estadisticulación”**

# Componentes de la Vacuna Prevenestat (I)

Factor crítico	Tener en cuenta	Comentarios
Autores	Conflictos de interés	Recuerde que una estrategia utilizada es “convencer a los líderes de opinión”
Institución	En qué medida esa institución participó en el estudio	La sola mención de la Institución no es suficiente
Número de observaciones	¿Fueron suficientes para demostrar lo planteado en los objetivos?	Recuerde que no siempre más observaciones mejoran la confiabilidad: Riesgo de sobredimensionamiento
Comparaciones	¿Tuvo sentido médico esa/s comparación/es? ¿Se utilizó la prueba estadística apropiada? Recuerde que si una variable cuantitativa no tiene demostrada “Normalidad” la media aritmética y el desvío estándar no tienen sentido.	En muchas oportunidades el Médico Asistencial “huele” que algo no está bien en el “paper”, generalmente tiene razón el Médico Asistencial. Consulte con su Metodólogo de cabecera
Porcentajes	¿Están acompañados del número total que les dio origen? ¿Es correcta la interpretación de los mismos?	Al igual que con otros estimadores ¡sea escéptico!.

## Componentes de la Vacuna Prevenestat (II)

Factor crítico	Tener en cuenta	Comentarios
<b>Material y Métodos</b>	<p style="text-align: center;">¿La redacción es comprensible? ¿Con esa descripción Ud. podría repetir el estudio?.</p>	<p style="text-align: center;">Si la descripción no es clara, UD. tiene todo el derecho de pensar que los resultados tampoco lo serán.</p>
<b>Índices (agrupamiento de Indicadores)</b>	<p style="text-align: center;">Desconfíe si en Material y Método no está descrito cómo se contruyeron los índices e indicadores</p>	<p style="text-align: center;">Preste atención a la utilización de Índices e Indicadores no validados.</p>
<b>Asociación</b>	<p style="text-align: center;">Tener en cuenta no siempre es sinónimo de relación causal</p>	<p style="text-align: center;">Tenga en cuenta que en los '90 alguien asoció el crecimiento de las ventas de CD con el incremento de la infección HIV.</p>
<b>Resultados</b>	<p style="text-align: center;">Desconfíe si no está claramente descrita la metodología de obtención y tratamiento de los datos</p>	<p style="text-align: center;">No se deje impresionar por la "elegancia" y estética de tablas y figuras. Más que la forma busque el fondo de las mismas.</p>
<b>Conclusiones</b>	<p style="text-align: center;">¿Guardan relación con los objetivos planteados y los resultados obtenidos?</p>	<p style="text-align: center;">Recuerde que la disociación entre conclusiones y objetivos del trabajo es una de las causas más frecuentes de devolución de las publicaciones por parte de los árbitros.</p>

# **Adyuvantes de la Vacuna Prevenestat**

- a) Las “enfermedades inventadas” (“Mongering diseases”).**
- b) Filtrar con el sentido común el resultado de los consensos.**



## **Y, entonces, ¿Qué dicen y qué no dicen las estadísticas?**

- 1.- DICEN que es una forma de tomar decisiones con acotado nivel de incertidumbre (ver a continuación).**
- 2.- NO DICEN a qué atribuir biológica o médicamente los resultados.**

<http://d.mujer.hvimg.com/imagenes/salud/20110623121752.jpg>

[http://cl.selecciones.com/upload/contents/secondaryImage\\_639.jpg](http://cl.selecciones.com/upload/contents/secondaryImage_639.jpg)

# ¿Qué dicen las estadísticas?

Inquietud	Comentario
¿Con la estadística se resuelven todos los aspectos de la investigación?	<b>No, solamente los aspectos cuantitativos y probabilísticos. NUNCA LA INTERPRETACIÓN. La interpretación es patrimonio exclusivo del/los investigador/es.</b>
¿Debo conocer “todos” los métodos estadísticos?	<b>No, debemos tener a mano el Algoritmo de decisión y razonablemente tener en cuenta</b> a) <b>Tipo de Variables (Categóricas ó Numéricas)</b> b) <b>Diseño (1 muestra, 2 muestras, más de 2 muestras y si las mediciones que se harán serán repetidas o independientes).</b>
Para hacer investigación ¿necesito tener experiencia en estadística?	<b>No, solamente hay que tener inquietudes, sentido común y un GRAN PODER DE OBSERVACIÓN e ideas claras. “Quien no sabe lo que busca no entiende lo que encuentra” (L. Pasteur).</b>

# ¿Qué dicen las estadísticas?

Inquietud	Comentario
¿Por qué el número de observaciones es importante?	Porque una de las formas de “controlar” la variabilidad es el incremento del número de observaciones. Deben tenerse presentes la cantidad (%) de Falsos Positivos y Falsos Negativos que se está dispuesto a aceptar.
¿Debo conocer “todos” los métodos estadísticos?	No, debemos tener a mano el Algoritmo de decisión y razonablemente tener en cuenta: a) Tipo de Variables (Categóricas ó Numéricas) b) Diseño (1 muestra, 2 muestras, más de 2 muestras y si las mediciones que se harán serán repetidas o independientes).
¿Qué significan Riego Relativo (RR)y Odd Ratio (OR)?	Son 2 medidas de “Fuerza de Asociación” entre una variable independiente (predictora) y una variable de desenlace (evento). Tienen significación si son $> 1$ y el LI de su IC95% no es 1 o menos de 1.

# ¿Qué dicen las estadísticas?

Inquietud	Comentario
¿Por qué en algunos estudios se utiliza RR y en otros OR	<p>a) El RR es un estimador de Fuerza de Asociación en los estudios de Cohorte, <u>cuando se plantea la Hipótesis Alternativa que existe Asociación entre Variable Predictora (Factor de Riesgo) y Variable de Desenlace (Evento).</u></p> <p>b) El OR es también un estimador de Fuerza de Asociación, utilizado en estudios de Casos y Controles, cuando se plantea la Hipótesis Alternativa que existe <u>asociación entre Variable de Desenlace y Variable Predictora.</u> El OR se puede utilizar en cualquiera de los dos diseños</p>
¿Qué significa que un RR u OR dé menor a 1?	En primer lugar indica que NO ES un FACTOR de RIESGO y también que puede ser un FACTOR de PROTECCIÓN si el LS del IC 95% es menor que 1.
¿Por qué la importancia del 1?	Porque las medidas de los “Riesgos” (RR y OR) son COCIENTES entre Riesgos “Absolutos”. Al ser un cociente si el resultado es 1 significa que Numerador y Denominador son iguales (es lo mismo estar expuesto que no estarlo).

# ¿Qué dicen las estadísticas?

Inquietud	Comentario
¿Cómo se interpreta la “p”?	<p>Indica la <b>PROBABIOLIDAD</b> de intervención del azar en el estudio.</p> <p><math>p &gt; 0.05</math>: No se puede rechazar <math>H_0</math> (las diferencias obtenidas no “son significativas”).</p> <p><math>P &lt; 0.05</math>: Existe evidencia para rechazar <math>H_0</math> (las diferencias “son significativas”).</p> <p><b>ALERTA!!!!!!</b>: <math>p &lt; 0.05</math> es válida siempre y cuando la cantidad de Falsos Negativos sea menor del 20% (Poder <math>&gt; 80\%</math>). Si no, no es válida por más pequeña que sea.</p>
¿Por qué es importante el Intervalo de Confianza?	<p>El IC X%, es un espacio numérico en el cual se encuentra el <b>VERDADERO VALOR</b> de la variable en estudio, pero en la <b>POBLACIÓN</b>. Trasciende a la Muestra.</p>
Pasamos todo el tiempo leyendo o recogiendo información ¿para qué?	<p><b>¡Alerta!</b> Podemos sufrir dos padecimientos:</p> <p><b>a) Infoxicación</b></p> <p><b>b) Infobesidad</b> (se muestran en la diapositiva siguiente)</p>



**Para estos dos síndromes  
“rigurosos metaanálisis han  
demostrado que la terapéutica  
de elección es”:  
EL SENTIDO COMÚN Y LA  
GENEROSIDAD.**

# ¿Qué dicen las estadísticas?

## Tengamos siempre presente:

¿Cuándo con nuestro estudio ética, metodológica y estadísticamente intachable no encontramos “evidencia” es porque esa evidencia no existe?

La significancia ( $p < 0.05$ ) no es absoluta ya que:

**“Ausencia de evidencia no es evidencia de ausencia”.**

**(Bland y Altman, 1999).**

(“Nunca en 25 años de profesión ví un caso de Leishmaniasis”. ¿Esto significa que la Leishmaniasis no existe?)

**Todo lo dicho se sintetiza en dos frases:**

**No todo “lo numérico” es verdad revelada  
(debemos conocer esto y estar prevenidos).**



**Recuerde la frase de  
Almafuerte (Pedro B. Palacios,  
1854-1917):**

**“Quien no viene por donde  
debe venir, no viene a lo que  
dice”.**

En Medicina Asistencial: “LA CLÍNICA ES SOBERANA”  
En Investigación: “EL SENTIDO COMÚN ES SOBERANO”

**“Nuestro objetivo como científicos no es publicar tantos artículos como podemos, sino descubrir y diseminar la verdad”.**

**Detrás de nuestras verdades científicas están aguardando nuestros pacientes.**



[http://adriabl.files.wordpress.com/2010/11/inicios\\_microscopio.jpg](http://adriabl.files.wordpress.com/2010/11/inicios_microscopio.jpg)

[http://img.medicaexpo.es/images\\_me/photo-g/estetoscopio-doble-cabeza-acero-inoxidable-78888-163871.jpg](http://img.medicaexpo.es/images_me/photo-g/estetoscopio-doble-cabeza-acero-inoxidable-78888-163871.jpg)

**Tengamos presente siempre este  
monstruo de 3 cabezas:**

**Estadisticulación**

**Infoxicación**



**Infobesidad**

# **Agradecimiento:**

**A las Dras.:**

**Florencia Lución y “Mavi” Juárez por las inteligentes preguntas que motivaron esta presentación.**

**Después de 46 años de docencia, Dios me dio la gracia de poder seguir aprendiendo de mis alumnos jóvenes.**



**“Pueden obligarme a buscar la  
verdad pero no a encontrarla”**

**Denis Diderot. Pensador y  
enciclopedista francés (1713-1784)**

**Muchas gracias por su tiempo y  
paciencia**

**Para la reflexión....**

## Aspectos a tener en cuenta al momento de redactar un “paper”

Acápite	Sugerencias
<b>Título</b>	Breve y explicativo (quizás sea lo único que sea)
<b>Autores</b>	Sólo los que realmente trabajaron o hicieron aportes de algún tipo. Filiaciones (cargos). Evitar el “amiguismo”
<b>Resumen</b>	300 palabras. Estructuración: de acuerdo a la revista. Objetivo/s(15%)-Métodos(10%)-Resultados(60%)-Conclusiones(15%)
<b>Palabras clave</b>	Por lo menos 3
<b>Introducción</b>	Problema a resolver/estudiar. Estado actual del conocimiento (no debe ser una revisión bibliográfica). Último párrafo: Objetivos del estudio. Si el Protocolo fue bien redactado, este acápite ya está escrito.
<b>Métodos</b>	Unidades de análisis (personas, muestras, etc.). Diseño del estudio. Aspectos éticos Acciones realizadas (extracciones, diversas mediciones, etc.). Método/s estadísticos seleccionados (nivel $\alpha < 0.05$ y poder fijados 0.80). Recordar que de acuerdo al tipo de variable serán los estimadores que se utilicen. Ídem Introducción.
<b>Resultados</b>	“Mostrar más que decir” (tablas, figuras). No sacar conclusiones en este acápite
<b>Discusión</b>	a) Comparación con otros autores b) Fundamento de los hallazgos c) Posibles hipótesis posteriores
<b>Conclusiones</b>	Deben estar totalmente de acuerdo con los Objetivos planteados en la Introducción
<b>Referencias (Bibliografía)</b>	Actualmente “por orden de aparición en el texto” Siempre consultarlas instrucciones para los autores de la revista a la que se enviará el trabajo

# **Algunas preguntas que Ud. siempre se hizo y las respuestas nunca se entendieron bien**

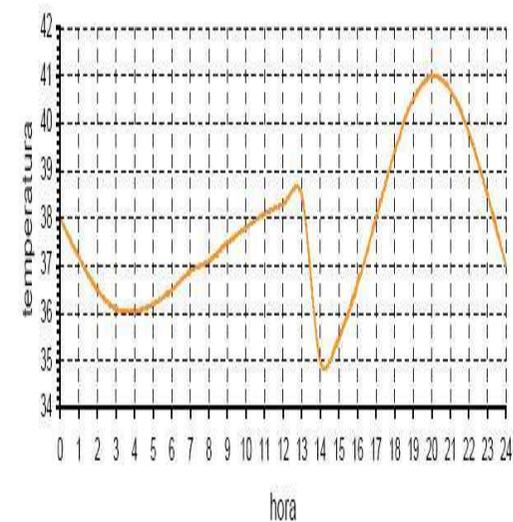
**1.- Si yo para ser buen Pediatra Asistencial tengo que estar bien entrenado en Clínica, Diagnóstico y Tratamiento ¿por qué necesito la Bioestadística?.**

- a) Por la Clínica: ya que para las enfermedades es necesario fijar “modelos” o “patrones”. Estos patrones son más sólidos cuanto más evidencia los respalde y la evidencia, entendemos hoy, es una muy baja probabilidad de intervención del azar. Para la interpretación de la información epidemiológica.**
- b) Por el Diagnóstico: los métodos diagnósticos deben estar “validados”, es decir ser Sensibles (que evalúen lo que se desea evaluar) y Específicos (que puedan distinguir entre salud y enfermedad).**
- c) Por el tratamiento: las medidas de intervención, deben demostrar que son notoriamente más útiles que no aplicarlas y esto surge, por ejemplo, de los Metaanálisis.**

# Algunas preguntas que Ud. siempre se hizo y las respuestas nunca se entendieron bien

2.- ¿En mi práctica asistencial de qué me sirve (además de los aspectos legales y administrativos) recolectar datos numéricos?

- a) Por una cuestión de ordenamiento: siempre se encuentra algo que está adecuadamente ordenado que si está “en cualquier parte”.
- b) Porque los datos numéricos, además, nos permiten “medir” diferentes aspectos (por ejemplo: Percentilos de estatura o peso).
- c) Porque si medimos podremos **COMPARAR** y si comparamos en diferentes tiempos, podremos establecer la evolución de un proceso ¿recordemos los clásicos cuadros de evolución de temperatura que se encontraban a los pies de las camas de nuestros pacientes internados).



Termograma diario

## **Algunas preguntas que Ud. siempre se hizo y las respuestas nunca se entendieron bien**

**3.- Supongo que con esto me alcanza ¿o hay más?**

- a) Sí, hay más. La propia esencia de nuestra Profesión es conocer y establecer relaciones entre diferentes manifestaciones que se llaman VARIABLES (no tienen siempre el mismo valor).**
- b) Porque esos diferentes valores (por ejemplo dentro del rango de normalidad) ocurren por CASUALIDAD o AZAR. Es más, ese rango de “valores normales” fueron confirmados y validados mediante métodos brindados por la Bioestadística.**
- c) Porque cuando esos valores están fuera del rango de normalidad, debemos saber si es por CASUALIDAD O por CAUSALIDAD.**

## Relaciones entre Observación, Intervención, Descripción y Análisis

<b>Tipo de Estudio</b>	<b>Descriptivos Sin Hipótesis</b>	<b>Analíticos Con Hipótesis</b>
<b>Observacionales</b>	Descripción de casos Prevalencia Incidencia	Casos y Controles Cohorte Caso-Cohorte Transversales
<b>De Intervención</b>	De haber algo aquí, sería una incongruencia	Ensayos Clínicos

# ¿Cuál es la diferencia entre Riesgo Relativo y Odds Ratio?

## Estudio de Cohortes: H1: Existe Asociación entre Exposición y Evento

	Evento	No Evento	Total filas
Expuestos	a	b	a+b
No exp.	c	d	c+d
Total columnas	a+c	b+d	a+b+c+d

Riesgo Absoluto en Expuestos (RAE):

$$a/a+b$$

Riesgo Absoluto en No expuestos (RANE):

$$c/c+d$$

Riesgo Relativo (RR) = RAE / RANE

Es una Razón de Proporciones (RA)

## Estudio de Casos y Controles: H1: Existe Asociación entre Evento y Exposición

	Casos	Controles	Total filas
Expuestos	a	b	a+b
No exp.	c	d	c+d
Total columnas	a+c	b+d	a+b+c+d

Odd de exposición en Casos (OEC):

$$a/c$$

Odd de exposición en Controles (OECt):

$$b/d$$

Odd Ratio:

$$OEC / OECt$$

Es un Razón de Razones (Odd)

El OR es una buena estimación de RR cuando este no puede utilizarse (no se puede sumar "en horizontal"). Como se observa Odd Ratio puede utilizarse en ambas situaciones

Para abordar estos aspectos, debemos recordar los componentes de toda Hipótesis

**VARIABLE**



**VARIABLE INDEPENDIENTE**

Que actúa sobre el/los



**SUJETO O INDIVIDUO**

Exposición  
Factor de Riesgo  
Tratamiento



**UNIDAD DE ANÁLISIS**



Y modifica las variables elegidas



**VARIABLE**



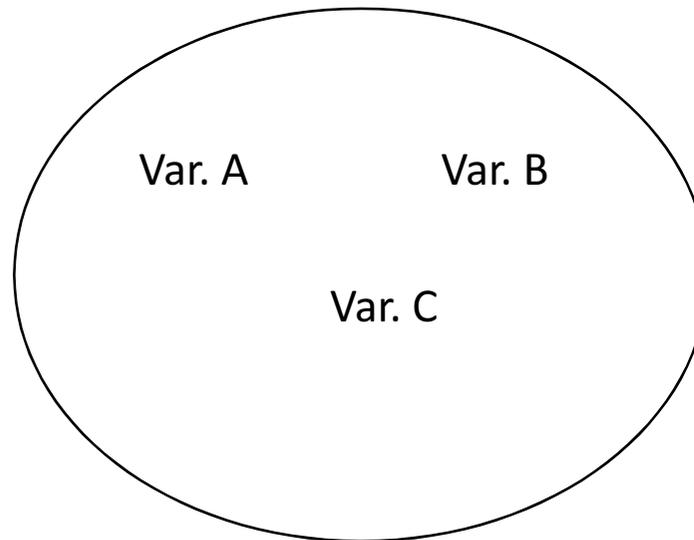
**VARIABLE DEPENDIENTE**



Desenlace; Evento; Respuesta

## Modelo General

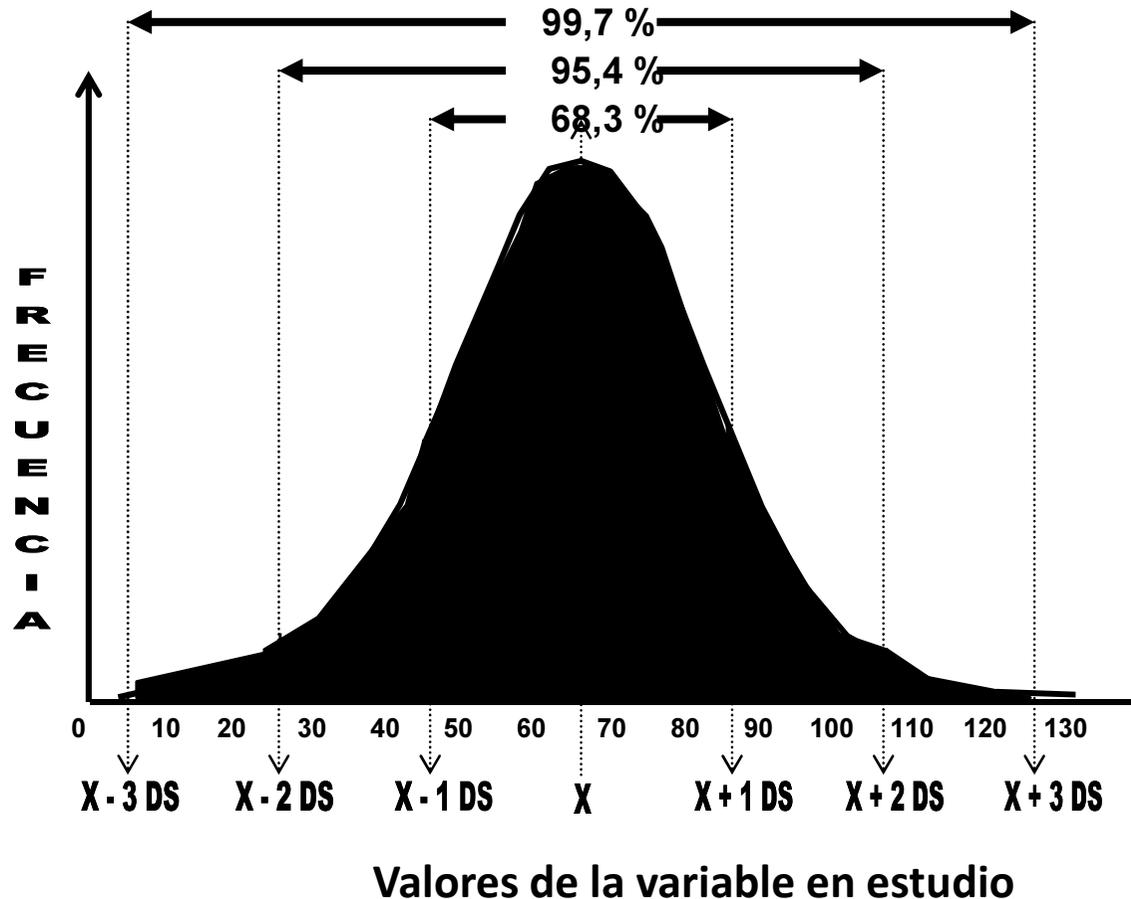
Cuando queremos estudiar un hecho, el mismo está dado por ciertas características, que son precisamente las Variables



Si Describimos de “a una por vez”: **Descripción Univariada** (no importa cuántas variables sean).

Si Analizamos las posibles **RELACIONES** entre ellas : **Análisis Bi o Multivariado.**

**Veamos un concepto cardinal para comprender muchos aspectos de la Bioestadística. La Distribución Normal o Gaussiana**



La razón por la cual es importante este concepto es porque si una serie de datos numéricos se aproxima a esta distribución, se pueden aplicar las técnicas “Paramétricas”:

- a) Descriptivas: Media, Desvío Estándar, Variancia, Coeficiente de Variación %.
- Inferenciales: tests t de Student, ANOVA, Regresión lineal, etc.