



## Aprendiendo idiomas I

María Jesús Vázquez Gallo



La Torre de Babel. Bruegel el Viejo. 1563.



1. **Introducción.**

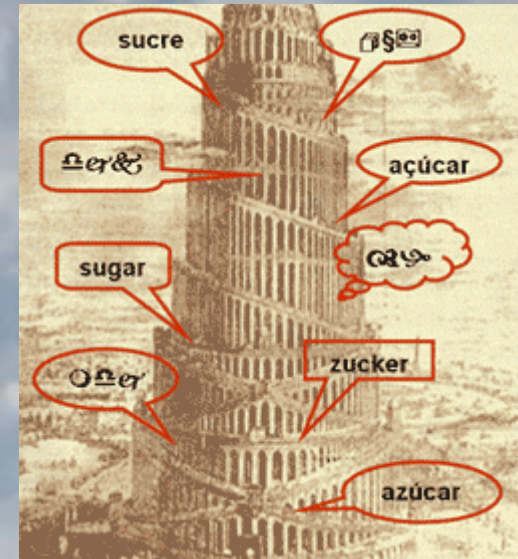
2. **Econometría.**

3. **Diccionario.**

**Continuará...**

En la Biblia, en el capítulo 11 del Génesis, se relata que los hombres pretendían alcanzar el cielo al construir la **torre de Babel**.. pero Yavé deseaba que la Humanidad se extendiera y multiplicara por toda la Tierra, así que hizo que los constructores de la torre hablasen en diferentes lenguas con lo cual se produjo tal confusión que la construcción de la torre no llegó a su fin...

Según el diccionario de la RAE, **babel** significa “lugar en que hay gran desorden y confusión o donde hablan muchos sin entenderse”.



Si alguien dice que utiliza un modelo **ARIMA** para construir un modelo lineal de una serie temporal en Economía, otra persona cuenta que maneja dos **series no estacionarias I(1) que están cointegradas**, y otra que, para modelizar una serie financiera, utiliza un modelo **ARCH**...

...quizá entendamos perfectamente lo que quiere decir...

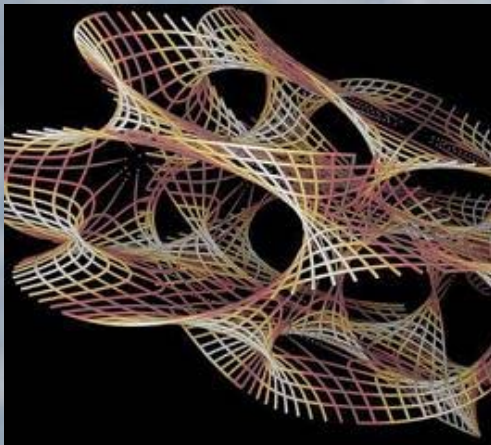


... o quizá no entendamos nada...



¿por qué se usa todo este lenguaje técnico –matemático- en lo relacionado con: la Economía, el Marketing, las finanzas...?

Está bien que se utilice en la Física... A mediados del S. XVII, Isaac Newton revolucionó las ciencias con sus leyes de Mecánica, Gravitación y Óptica, desarrollando el cálculo diferencial e integral (¡las famosas... derivadas e integrales!)



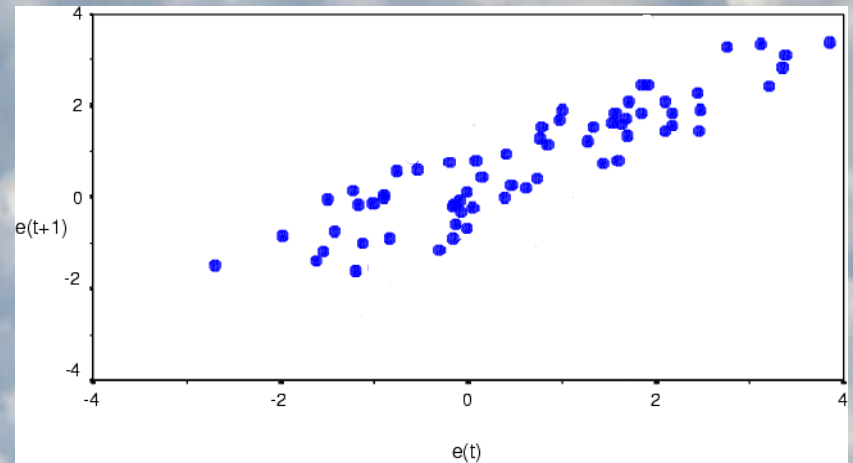
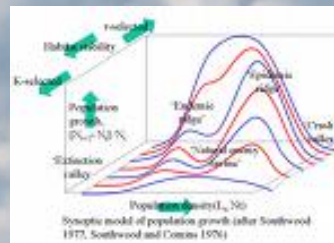
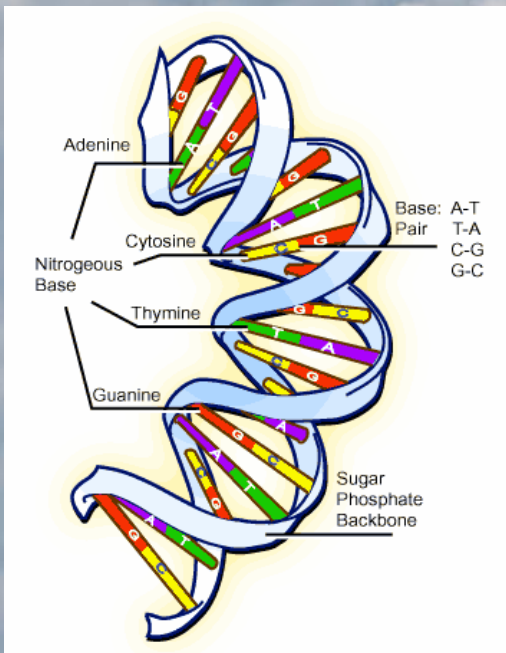
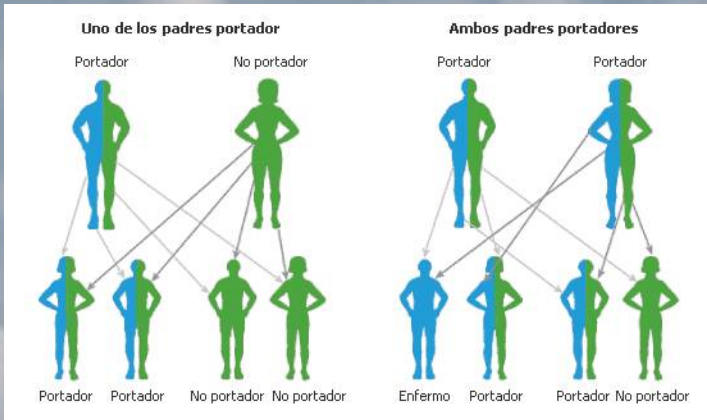
Usando este lenguaje, Albert Einstein explicó a la vez el las fuerzas del electromagnetismo y la gravedad, en su Relatividad General.

Sin este lenguaje, las modernas teorías físicas de las supercuerdas no se pueden entender...

**Formación Conento:** septiembre 2008

# No está mal que se use en la Biología...

A finales del S. XIX, Francis Galton (primo de Darwin) observó el fenómeno de **regresión a la media** al estudiar la herencia genética y la relación entre la estatura de los padres y sus hijos.

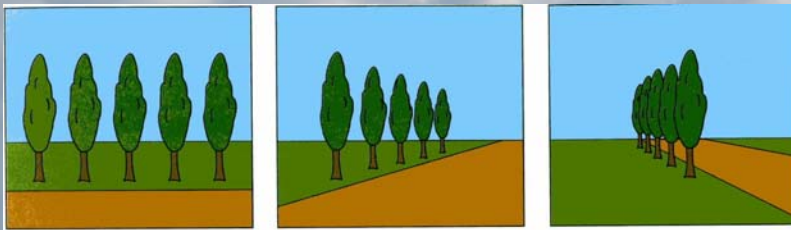
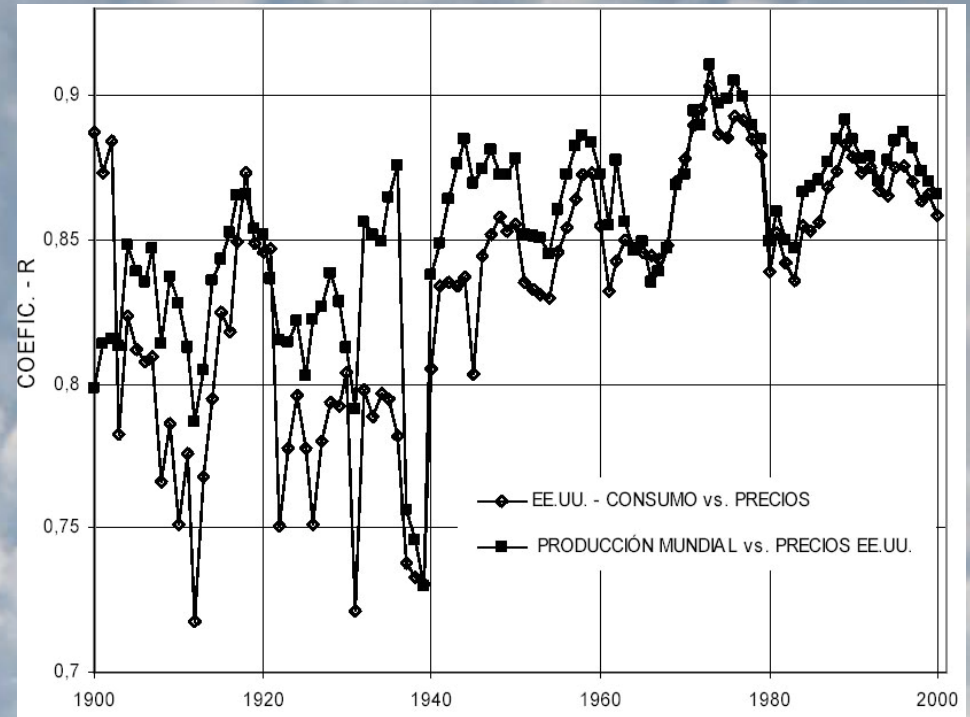


Cada vez más, se utilizan métodos matemáticos para analizar y predecir procesos biológicos.

## ¿Pero por qué en la Economía?

Porque sirven para:

- establecer reglas generales de comportamiento de muchos procesos, sabiendo qué confianza podemos depositar en ellas.
- generar predicciones sobre el comportamiento futuro de dichos procesos



Merece la pena aprender otro idioma, si nos permite entender otros puntos de vista.

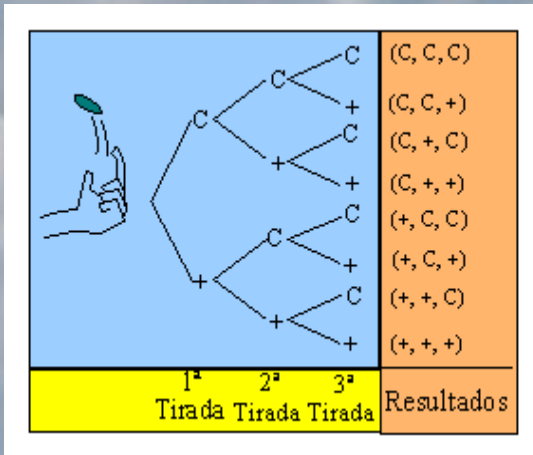


El vocabulario que vamos a intentar aprender es el de la **Econometría**.

La **Econometría** es uno de los campos de aplicación de la **Estadística**, en este caso, a la **Economía**.



La **Estadística** es un subcampo de las **Matemáticas** que se ocupa de sistemas que deben describirse usando **probabilidades** (porque los valores de las variables involucradas no se conocen con certeza, no son determinísticos).



	A	B	C	D
1				
2		Prueba de diferencia pareadas		
3		Automóvil	Neumático A	Neumático B
4		1	10.6	10.2
5		2	9.8	9.4
6		3	12.3	11.8
7		4	9.7	9.1
8		5	8.8	8.3
9		6	10	10.1
10		7	9.9	9.2
11		8	9	11.2
12		9	12.1	11
13		10	8.9	8.2
14		11	10.1	10.1
15		12	11	10
16		13	11.8	10.3
17		14	9.9	10.4
18		15	12.2	11.1
19		16	12.3	11.3
20		17	10.5	9.3
21		18	8.8	8.5
22		19	8.6	10.3
23		20	9.2	11

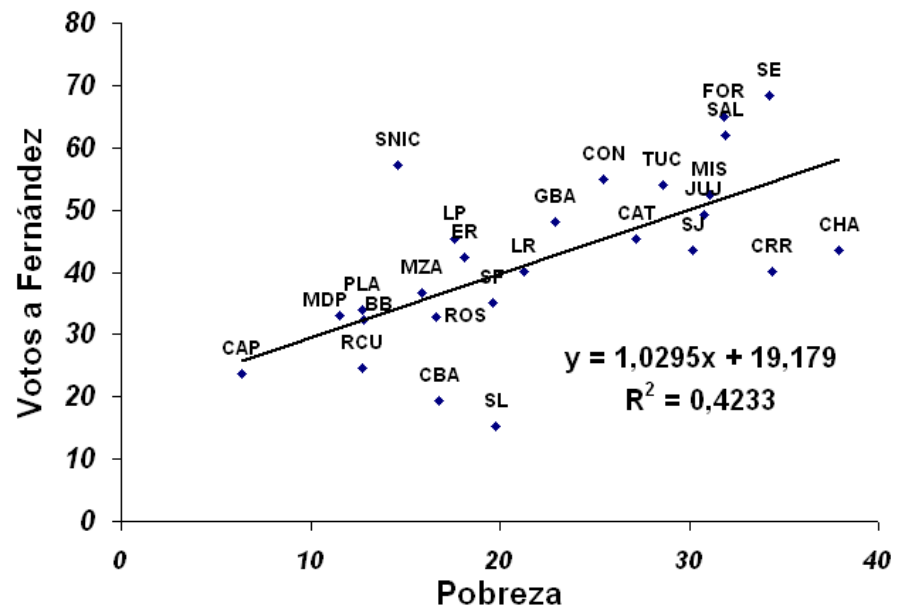
En **Estadística** es fundamental el análisis de datos.

Estadística + Economía + Matemáticas → Econometría

En palabras del economista **Ragnar Frisch** en los años 30:

“La experiencia ha mostrado que cada uno de estos tres puntos de vista, el de la estadística, la teoría económica y las matemáticas, es necesario, pero por sí mismo no suficiente para una comprensión real de las relaciones cuantitativas de la vida económica moderna. Es la *unión* de los tres aspectos lo que constituye una herramienta de análisis potente. Es la *unión* lo que constituye la econometría”

### La pobreza de Fernández





Pero ¿cómo y cuándo nace la Estadística?



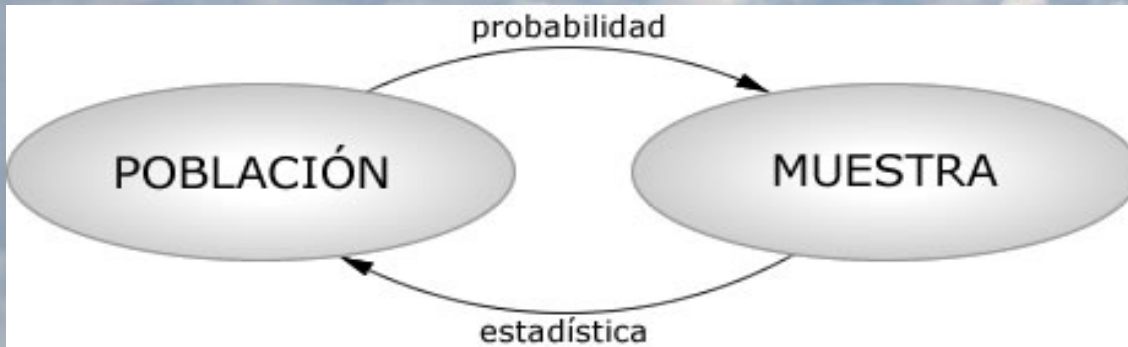
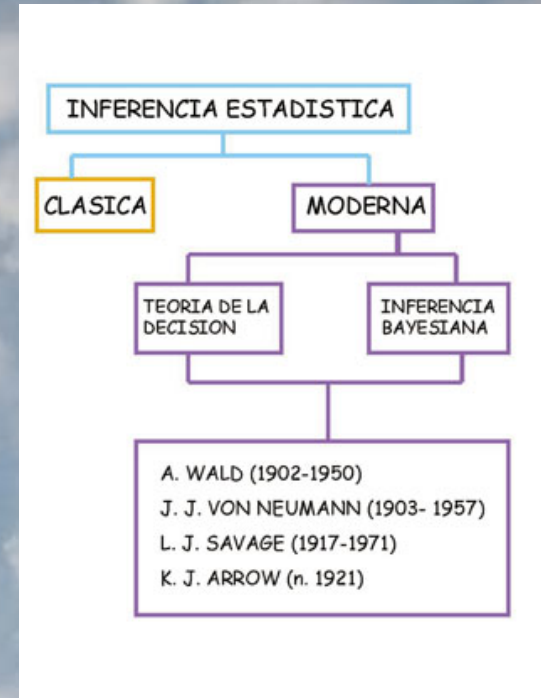
Dicen que el padre fue **Godofredo Achenwall** que realizó en el S. XVII los primeros análisis de datos sobre la economía de los Estados europeos y utilizó el término **Statistik**, relacionado con la palabra alemana **Staat**, que se refiere al estado como expresión de unidad política y se deriva del latín **status**, que significa estado también en el sentido político.

Pero el uso de herramientas cuantitativas para el tratamiento de datos, tiene su origen en épocas remotas. Hace más 5000 años, los egipcios empleaban censos para organizar el estado y la construcción de las pirámides.



Estadística

- Estadística DESCRIPTIVA  
Recolección, descripción, visualización y resumen de datos.
- INFERENCIA Estadística  
Generación de modelos y predicciones, teniendo en cuenta la aleatoriedad de los fenómenos observados.





Y empezamos por la A de **AZAR** pero en este diccionario no seguiremos el orden alfabético sino el de lo más fácil a lo menos fácil...

Ya sabemos, del curso sobre **Simulación Montecarlo** que:

**Azar:** del árabe *zahr*, dado, literalmente 'flores'.

**Aleatorio:** del latín *aleatorius*, propio del juego de dados.



Y también, del curso **Series temporales...** y **Fourier** que:

**Variable aleatoria:** variable que toma diferentes valores según el resultado de un experimento aleatorio. **Ejemplo:** **X** número de caras obtenidas al lanzar 2 monedas, puede tomar los valores:  $xx \rightarrow 0$ ,  $xc$  y  $cx \rightarrow 1$  y  $cc \rightarrow 2$ .





¿Y para qué sirven las **variables aleatorias**?

Una marca americana desea estudiar el consumo anual de latas de cerveza por habitante en Dakota del Norte.



Tomamos una **muestra** de la **población** de Dakota del Norte (unos 650.000 habitantes)...

¿**Cuántos**? Un número manejable.

¿**Quiénes**? La muestra debe ser representativa: cualquier individuo debe haber tenido la misma probabilidad de ser elegido. Hay técnicas para conseguir esto.

¿**Dónde está la variable aleatoria**?

Es el número de latas de cerveza que bebe anualmente cada habitante de Dakota del Norte.

Una vez recogidos los datos, se realiza el análisis descriptivo.

¿Alrededor de qué valor se agrupan los datos?

La respuesta está en las medidas de tendencia central. Todos recordamos algo así como **media, mediana y moda** ¿no?

Supuesto que se agrupan alrededor de un número, ¿cómo lo hacen?, ¿muy concentrados?

La respuesta está en las medidas de dispersión. Algunos recordamos algo así como **varianza y desviación típica** ¿no?

¿Qué otras medidas nos permiten describir los datos?

La respuesta está en las medidas de posición. Los papás recientes seguro que conocen los **percentiles**, y existen también **cuartiles**, etc.

Supongamos que 315, 321, 332, 359, 360, 360, 361, 364, 371 y 387 son diez datos recogidos en la muestra de Dakota del Norte y ordenados de menor a mayor. (En la realidad serían más, claro)

La **media (aritmética)** sería la suma de todos los datos dividida por 10:

$$3530/10=353$$

→ Los habitantes de Dakota del Norte en promedio consumen 353 latas de cerveza al año.

La **mediana** sería... como hay diez datos que es un número par...

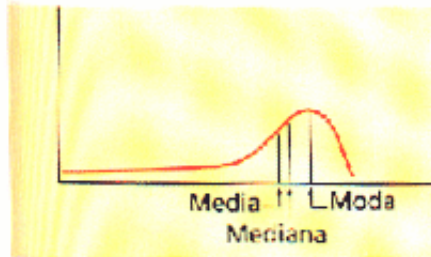
La media de los dos datos en el centro de la ordenación:

$$(360+360)/2=360$$

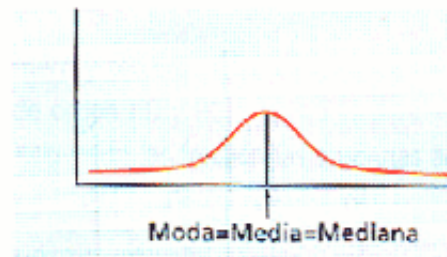
→ Hay tantos consumos por encima de 360 como por debajo de esa cantidad.

La **moda** es el dato que más se repite, en este caso, el 360.

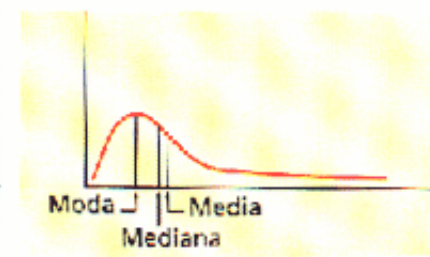
→ 360 es el número de latas de cerveza por año preferido por los de Dakota del Norte.



Sesgada a la izquierda



Simétrica o incesgada



Sesgada a la derecha

La media es la medida de centralización más utilizada pero

La **media** es muy sensible a la variación de los datos.

La **mediana** es menos sensible a dichos cambios.

Cuando una serie de datos contiene algunos valores muy grandes o muy pequeños en relación con el resto, la media no es representativa. La tendencia central se describe mejor en este caso usando la mediana.

En cuanto a las **medidas de dispersión** de los datos:

La **varianza** de los datos es la media de los cuadrados de las diferencias entre cada valor de la variable y la media.

$$[(353-315)^2 + (353-321)^2 + \dots + (353-387)^2] / 10 = 5228 / 10 = 522.8$$

→ Pero esos exponentes... ¿Esto es un número de latas al cuadrado?

La **desviación típica** es la raíz cuadrada de la varianza

22.86

→ En promedio un individuo se aleja 22.86 del número medio de latas.

La **desviación típica** es la desviación en promedio de los datos con respecto a la media. Si nuestra variable fuera el valor de un índice bursátil como el IBEX, por ejemplo; la desviación típica, es una medida de riesgo, que se conoce también como **volatilidad**.



Algunos datos curiosos que publicaba la revista **TIME** hace un año:

1. En media, el consumo anual de cerveza en Dakota del Norte es de 353 latas por habitante.
2. Duermen en promedio 8 horas 38 min.
3. 12% de los sueños de los hombres son sobre sexo comparado con 4% de las mujeres.
4. Están conectados a Internet un promedio de 2 horas 36 min. desde la oficina y 1 hora 7 minutos desde casa.
5. Los más felices son los “sacerdotes”, los menos felices “ayudantes de gasolineras”.
6. Hay más televisores que personas en casa. (2.7 televisores por casa, comparado con 2.6 personas por casa)
7. El 83% de la población no hace ejercicio.

