

Universidad Nacional Autónoma de México

Facultad de Estudios Superiores Iztacala

Carrera de Biología

**GUÍA PARA EL EXAMEN
EXTRAORDINARIO DE BIOESTADÍSTICA
II**

Por

Dr. Jorge Ricardo Gersenowies Rodríguez

Biol. Antonio Edmundo Cisneros Cisneros

2026

Presentación

Introducción

La Bioestadística no es una colección de fórmulas: es una forma de pensar la evidencia biológica.

En el tránsito por la carrera de Biología —y particularmente en una asignatura como Bioestadística II— se hace evidente que los datos no hablan por sí solos. Requieren organización, análisis, contraste y, sobre todo, interpretación crítica. Esta guía no pretende sustituir el estudio profundo de cada tema, sino acompañar la preparación para el examen extraordinario desde una perspectiva reflexiva: comprender qué estamos haciendo cuando aplicamos un método estadístico y por qué lo hacemos.

El **objetivo general** del curso establece que el alumno aplicará métodos estadísticos descriptivos e inferenciales para analizar datos obtenidos en investigaciones biológicas. En esa frase se condensa una responsabilidad metodológica: no basta con calcular, es necesario comprender.

La estadística como herramienta para construir conocimiento

En Biología, los fenómenos rara vez son deterministas y lineales. Trabajamos con variabilidad: organismos distintos, ambientes cambiantes, respuestas fisiológicas complejas. La estadística permite distinguir entre:

- Variación biológica real
- Variación experimental
- Error aleatorio
- Patrones con significado ecológico, evolutivo o fisiológico

Por ello, cada técnica incluida en el programa responde a una pregunta científica distinta.

1. Diseño y análisis de experimentos (ANOVA)

El diseño experimental no es un requisito burocrático: es el fundamento de la inferencia válida.

Cuando analizamos un **diseño factorial $A \times B$** , no solo comparamos medias; exploramos interacciones biológicas.

- ¿El efecto de un tratamiento depende del nivel de otro factor?
- ¿Existe sinergia o antagonismo?

El **ANOVA** permite separar fuentes de variación. La prueba **LSD** ayuda a identificar diferencias específicas entre medias. La **gráfica de interacción** convierte números en interpretación biológica.

Reflexión clave:

Un experimento mal diseñado no se corrige con un buen análisis estadístico.

2. Regresión lineal

La regresión responde a una pregunta central en Biología:

¿Cómo cambia una variable cuando otra varía?

En la regresión lineal:

- β_0 representa el punto de partida.
- β_1 representa la tasa de cambio.
- R^2 indica cuánto del fenómeno observamos está explicado por el modelo.

Pero el verdadero ejercicio crítico es preguntarse:

- ¿La relación es causal o solo asociativa?
- ¿El modelo lineal es biológicamente plausible?
- ¿El ajuste por mínimos cuadrados describe adecuadamente los datos?

La prueba de hipótesis para β_1 no es solo un cálculo: es la evaluación formal de si existe evidencia de relación entre variables.

3. Regresión no lineal

Muchos procesos biológicos no son lineales.

- La **hipérbola rectangular** modela cinética enzimática y saturación.
- La **regresión potencial** describe relaciones alométricas.
- La **regresión exponencial** modela crecimiento poblacional.
- La **regresión logística** integra límites ambientales.

Aquí la reflexión es profunda:

Ajustar un modelo no lineal implica reconocer que la naturaleza impone límites.

La elección del modelo no es matemática solamente: es biológica.

4. Correlación

El coeficiente R mide asociación, no causalidad.

La prueba de hipótesis sobre ρ permite evaluar si la correlación observada podría deberse al azar.

Pero interpretar correctamente significa reconocer:

- La magnitud del efecto.
- El contexto experimental.
- El tamaño de muestra.
- La plausibilidad biológica.

Una correlación alta puede ser espuria; una correlación moderada puede ser biológicamente relevante.

5. Métodos no paramétricos

Cuando los datos no cumplen supuestos de normalidad o homogeneidad de varianzas, la estadística ofrece alternativas:

- J_i^2 para variables categóricas.
- Wilcoxon para datos pareados.
- Mann–Whitney para muestras independientes.
- Kruskal–Wallis para múltiples grupos.

La decisión de usar métodos no paramétricos no es “un recurso de emergencia”; es una elección metodológica válida cuando la estructura de los datos lo exige.

Interpretar: el acto final y más importante

Los objetivos específicos del curso culminan en la interpretación.

Interpretar implica:

- Conectar resultados estadísticos con hipótesis biológicas.
- Reconocer limitaciones.
- Evitar conclusiones exageradas.
- Formular nuevas preguntas.

El examen extraordinario no debe verse únicamente como una evaluación cuantitativa, sino como una oportunidad para demostrar:

- Comprensión conceptual.
- Razonamiento estadístico.
- Capacidad crítica.
- Integración biológica.

Invitación al estudio consciente

Para prepararse adecuadamente:

1. Repase los conceptos antes que las fórmulas.
2. Practique ejercicios completos, desde el planteamiento hasta la interpretación.
3. Pregúntese siempre:
 - ¿Qué hipótesis estoy contrastando?
 - ¿Qué significan estos resultados en términos biológicos?
 - ¿Qué supuestos estoy aceptando?

La Bioestadística II no es un obstáculo curricular: es una herramienta que permitirá evaluar artículos científicos, diseñar investigaciones y defender resultados con rigor.

Comentario final

Comprender la estadística es comprender cómo se valida el conocimiento científico.

Que esta guía no sea solo preparación para un examen, sino consolidación de una forma rigurosa de pensar la Biología.

Los Reyes Iztacala, febrero 2026

Dr. Jorge Ricardo Gersenowies Rodríguez

Biol. Antonio Edmundo Cisneros Cisneros

Contenido Temático – Bioestadística II

1. ANOVA – Diseño y Análisis de Experimentos

1.1 Diseño Factorial A x B

A	B1	B2
1	2	3
2	4	5

1.2 Prueba LSD

1.3 Gráfica de Interacción



3. Regresión No Lineal

3.1 Hipérbola Rectangular



3.2 Potencial

$$y = a \cdot x^b$$



3.3 Exponencial

$$y = a \cdot e^{bx}$$



3.4 Logística



5. Métodos No Paramétricos

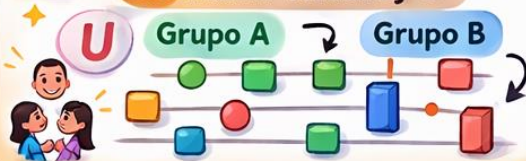
5.1 Ji² (Chi-cuadrado)



5.2 Wilcoxon (rangos)



5.3 Mann-Whitney U



2. Regresión Lineal

2.1 Introducción

$$y = f(X)$$

2.2 Regresión Lineal

$$Y = \beta_0 + \beta_1 X$$

2.3 Gráfica de Dispersión



2.4 Mínimos Cuadrados



2.5 R² Coef. de Determinación



2.6 Prueba de Hipótesis (β₁)

$$H_0: \beta_1 = 0$$

$$p < 0.05$$

4. Correlación

4.1 Introducción



4.2 Obtención de R

$$R = \frac{cov}{Sx \cdot Sy}$$



4.3 Prueba de Hipótesis (ρ)

$$H_0: \rho = 0$$

4.3.3 Interpretación



5.4 Kruskal-Wallis (H)



Comparación de rangos



Unidad 1.- Diseño y Análisis de Experimentos – ANOVA

El **Diseño y Análisis de Experimentos (DOE)** es una herramienta estadística que permite estudiar cómo influyen uno o más factores en una variable de respuesta.

Uno de los métodos más utilizados es el **ANOVA (Análisis de Varianza)**, que compara medias y determina si existen diferencias significativas entre tratamientos.

Diseño Factorial $A \times B$

¿Qué es un diseño factorial?

Un **diseño factorial** estudia **dos o más factores simultáneamente**, evaluando:

- Efecto del **Factor A**
- Efecto del **Factor B**
- **Interacción $A \times B$**

En un diseño **$A \times B$** , cada nivel del factor A se combina con cada nivel del factor B.

Ejemplo práctico

Una empresa quiere estudiar el efecto de:

- **Factor A:** Tipo de fertilizante (A1, A2)
- **Factor B:** Tipo de riego (B1, B2)

Variable respuesta: **Producción (kg)**

Diseño 2×2 con 2 repeticiones por tratamiento:

Fertilizante	Riego	Producción
A1	B1	20
A1	B1	22
A1	B2	25
A1	B2	27
A2	B1	30
A2	B1	32
A2	B2	28
A2	B2	29

Paso 1: Calcular medias

Medias por tratamiento:

- $A1B1 = (20+22)/2 = 21$
- $A1B2 = (25+27)/2 = 26$
- $A2B1 = (30+32)/2 = 31$
- $A2B2 = (28+29)/2 = 28.5$

Medias marginales:

- $A1 = (21+26)/2 = 23.5$
- $A2 = (31+28.5)/2 = 29.75$
- $B1 = (21+31)/2 = 26$
- $B2 = (26+28.5)/2 = 27.25$

Media general = 26.63

Paso 2: Tabla ANOVA (resumen conceptual)

Se descompone la variabilidad total en:

- SSA (Factor A)
- SSB (Factor B)
- SSAB (Interacción)
- SSE (Error)

Luego:

$$F = \frac{CM \text{ del factor}}{CM \text{ del error}}$$

Si $F \text{ calculado} > F \text{ tabla} \rightarrow$ efecto significativo.

Interpretación

- Si **A es significativo** $\rightarrow$ los fertilizantes producen diferencias.
- Si **B es significativo** $\rightarrow$ el tipo de riego influye.
- Si **A×B es significativo** $\rightarrow$ el efecto del fertilizante depende del tipo de riego.

Prueba de Comparación de Medias LSD

¿Qué es LSD?

LSD = **Least Significant Difference** (Diferencia Mínima Significativa).

Se usa después del ANOVA cuando el factor es significativo.

Permite comparar medias **dos a dos**.

Fórmula

$$LSD = t_{\alpha/2, gl\ error} \sqrt{\frac{2CM_e}{r}}$$

Donde:

- (CM_e) = Cuadrado medio del error
- r = número de repeticiones
- t = valor crítico de t de Student

Ejemplo (continuación)

Supongamos:

- CM error = 2.5
- r = 2
- t = 2.776 ($\alpha=0.05$)

$$LSD = 2.776 \sqrt{\frac{2(2.5)}{2}}$$

$$LSD = 2.776 \sqrt{2.5}$$

$$LSD = 2.776(1.58) = 4.39$$

Comparaciones

Comparación	Diferencia	¿Mayor que LSD?
A1B1 vs A2B1	10	✓ Sí
A1B2 vs A2B2	2.5	✗ No

Interpretación:

Existe diferencia significativa entre A1B1 y A2B1.

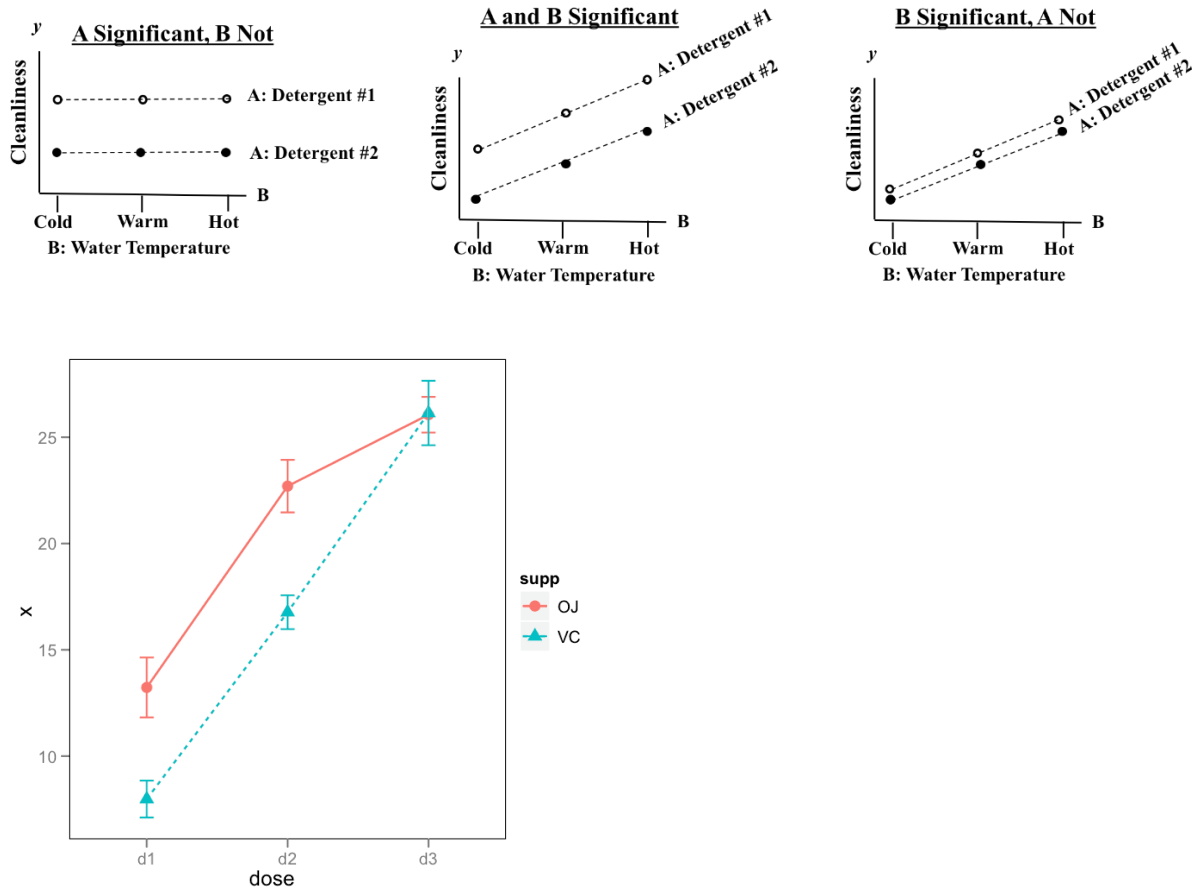
Gráfica de Interacción

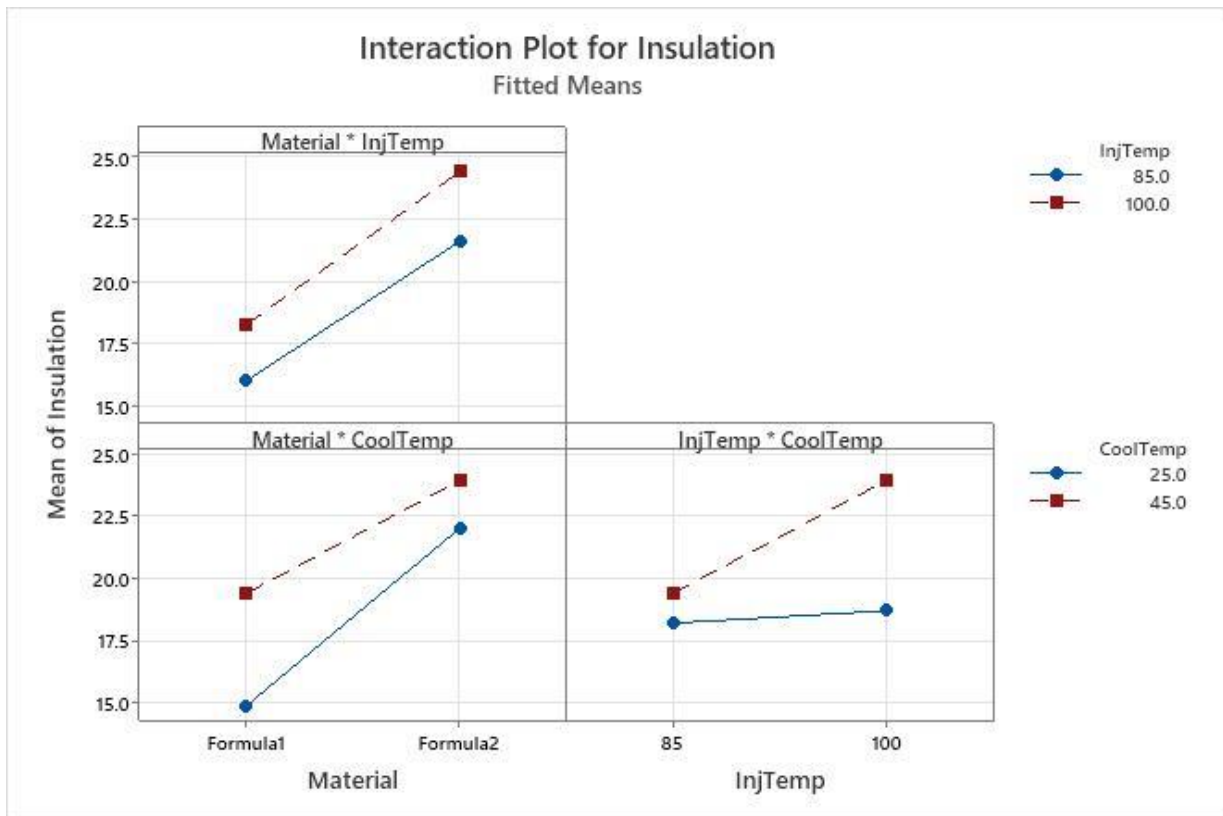
Una gráfica de interacción muestra si las líneas son:

- Paralelas → No hay interacción

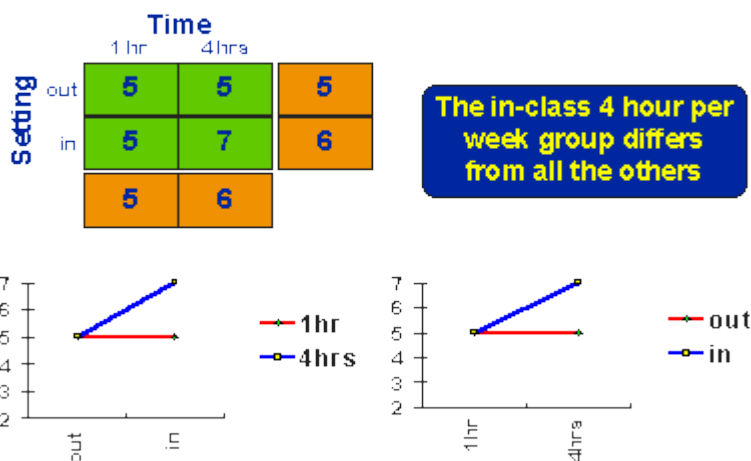
- Se cruzan → Sí hay interacción

Ejemplo visual conceptual





Interaction Effects



Interpretación del gráfico

En nuestro ejemplo:

- Para A1: 21 → 26
- Para A2: 31 → 28.5

Las líneas **no son paralelas**, por lo tanto existe indicio de interacción.

Ejercicios Resueltos

Ejercicio 1

Una empresa prueba:

- Factor A: Método de enseñanza (Tradicional, Virtual)
- Factor B: Horario (Mañana, Tarde)

Resultados promedio:

	Mañana	Tarde
Tradicional	70	75
Virtual	80	78

Preguntas:

1. ¿Existe interacción?
2. ¿Cuál método es mejor?

✓ Solución:

1. Las diferencias cambian de 5 a -2 → posible interacción.
2. Virtual tiene promedio mayor (79 vs 72.5).

Ejercicio 2

Se obtiene CM error = 4, r=3, t=2.12.

Calcular LSD:

$$LSD = 2.12 \sqrt{\frac{2(4)}{3}}$$

$$LSD = 2.12 \sqrt{2.67}$$

$$LSD = 2.12(1.63)$$

$$LSD = 3.45$$

Si diferencia entre medias = 5 → significativa.

Resumen Final

Tema	Objetivo
Diseño factorial	Evaluar efectos simultáneos
ANOVA	Detectar diferencias significativas
LSD	Comparar medias específicas
Gráfica de interacción	Visualizar interacción

Conclusión

El diseño factorial $A \times B$ permite:

- Ahorrar experimentos
- Detectar interacciones
- Obtener conclusiones más completas

El ANOVA determina si existen diferencias, y la prueba LSD indica **dónde están esas diferencias**.

Unidad 2. Análisis de Regresión Lineal

El **análisis de regresión lineal** es una técnica estadística que permite estudiar la relación entre una **variable independiente (X)** y una **variable dependiente (Y)**, con el objetivo de:

- Explicar la relación entre variables
- Predecir valores
- Evaluar la fuerza del modelo

2.1 Introducción a la regresión

La regresión responde preguntas como:

- ¿Aumentan las ventas cuando aumenta la publicidad?
- ¿Mejora la producción al aumentar la temperatura?
- ¿Influye el tiempo de estudio en la nota obtenida?

Se diferencia de la correlación porque:

Correlación	Regresión
Mide asociación	Modela y predice
No distingue causa	Sí distingue variable explicativa

2.2 Regresión Lineal

El modelo de regresión lineal simple es:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Donde:

- β_0 = intercepto
- β_1 = pendiente
- ε = error aleatorio

Interpretación:

- β_1 indica cuánto cambia Y cuando X aumenta una unidad.

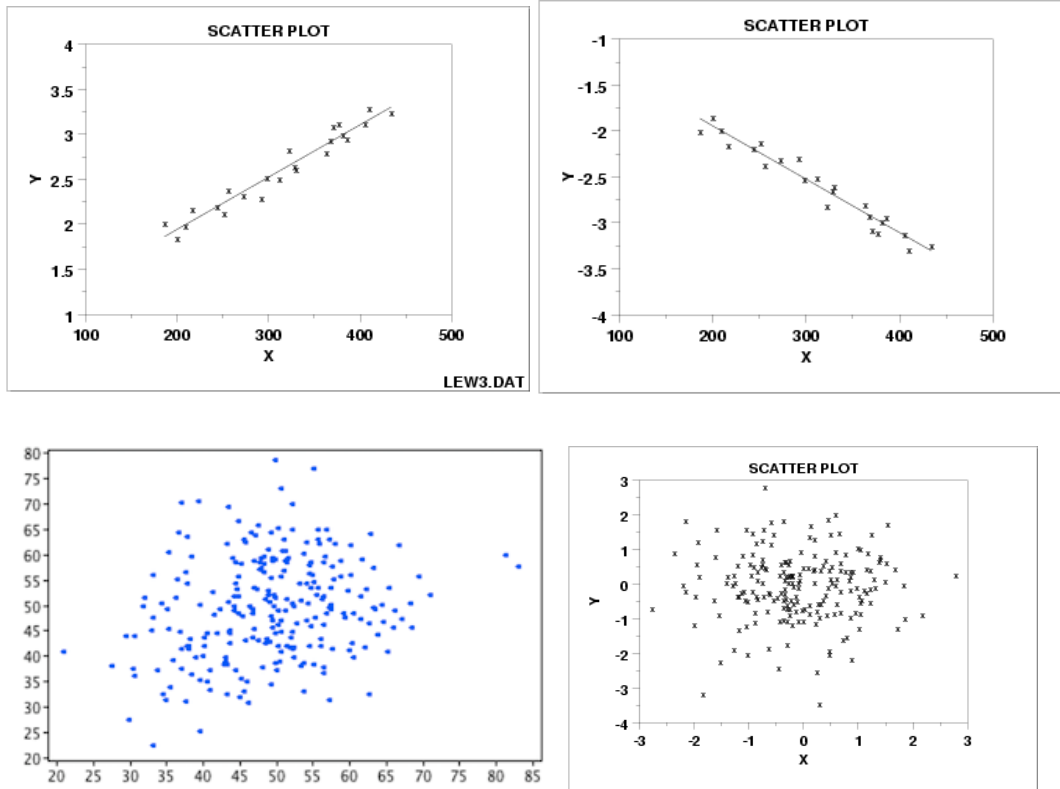
2.3 Gráfica de Dispersión

Antes de ajustar el modelo, se debe construir una **gráfica de dispersión**.

Permite observar:

- Tendencia positiva
- Tendencia negativa
- Relación lineal o no lineal
- Valores atípicos

Ejemplo visual



Interpretación:

- Si los puntos siguen una línea ascendente → relación positiva.
- Si descienden → relación negativa.
- Si están dispersos sin patrón → no hay relación lineal.

2.4 Ajuste por Mínimos Cuadrados (β_0 y β_1)

El método de **mínimos cuadrados** minimiza la suma de los errores cuadrados:

$$\sum (Y_i - \hat{Y}_i)^2$$

Fórmulas:

$$\beta_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

Ejercicio Resuelto 1

Datos:

X (Horas estudio)	Y (Nota)
2	65
4	70
6	75
8	85

Paso 1: Calcular medias

$$\bar{X} = 5$$

$$\bar{Y} = 73.75$$

Paso 2: Calcular β_1

$$[\beta_1 = 3.25]$$

Paso 3: Calcular β_0

$$\beta_0 = 57.5$$

Modelo final:

$$\hat{Y} = 57.5 + 3.25X$$

Interpretación:

Por cada hora adicional de estudio, la nota aumenta 3.25 puntos.

2.5 Coeficiente de Determinación (R^2)

El coeficiente de determinación mide qué proporción de la variabilidad de Y es explicada por X.

$$R^2 = \frac{SS_{reg}}{SS_{total}}$$

Valores:

- $0 \rightarrow$ no explica nada
- $1 \rightarrow$ explica todo

Ejemplo

Supongamos:

- $SS_{reg} = 350$
- $SS_{total} = 400$

$$R^2 = 350/400$$

$$R^2 = 0.875$$

Interpretación:

El 87.5% de la variabilidad de las notas se explica por las horas de estudio.

2.6 Prueba de hipótesis para β_1

Se quiere probar si existe relación lineal significativa.

Hipótesis:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

Estadístico:

$$t = \frac{\beta_1}{SE(\beta_1)}$$

Si:

$$|t_{calc}| > t_{tabla}$$

$\rightarrow$ Se rechaza H_0

Ejercicio Resuelto 2

Supongamos:

- $\beta_1 = 3.25$
- $SE(\beta_1) = 0.8$
- $gl = 2$
- $t_{\text{tabla}} = 4.303$

$$t = 3.25/0.8$$

$$t = 4.06$$

Comparación:

$$4.06 < 4.303$$

Conclusión:

No se rechaza H_0 al 5%.

No hay evidencia suficiente para afirmar que la pendiente es distinta de cero.

Resumen General

Tema	Fórmula clave
Modelo	$Y = \beta_0 + \beta_1 X$
Pendiente	$\beta_1 = \text{Cov}(X, Y) / \text{Var}(X)$
Intercepto	$\beta_0 = \bar{Y} - \beta_1 \bar{X}$
R^2	$SS_{\text{reg}} / SS_{\text{total}}$
Prueba t	$\beta_1 / SE(\beta_1)$

Conclusión

La regresión lineal permite:

Explicar relaciones entre variables

Predecir valores futuros

Medir qué tan bueno es el modelo (R^2)

Verificar si la relación es significativa

3.- Análisis de Regresión No Lineal

La **regresión no lineal** se utiliza cuando la relación entre las variables no puede representarse adecuadamente con una línea recta. Es común en economía, biología, física, ingeniería y ciencias sociales.

A diferencia de la regresión lineal, estos modelos pueden representar crecimiento acelerado, saturación, rendimientos decrecientes o procesos de difusión.

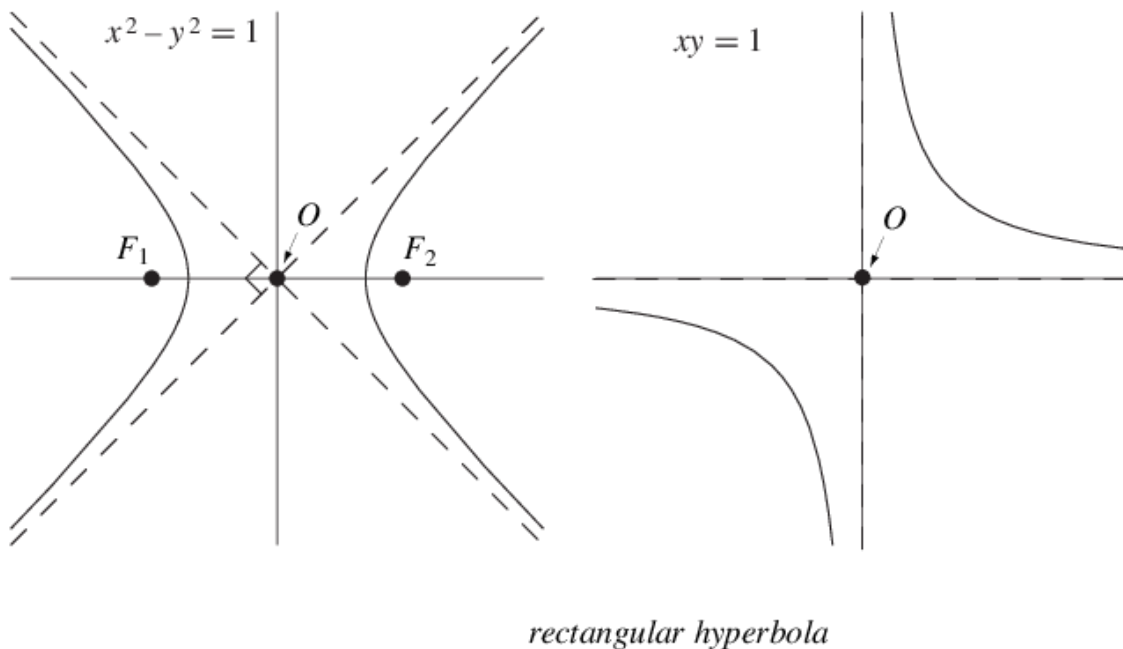
3.1 Hipérbola Rectangular

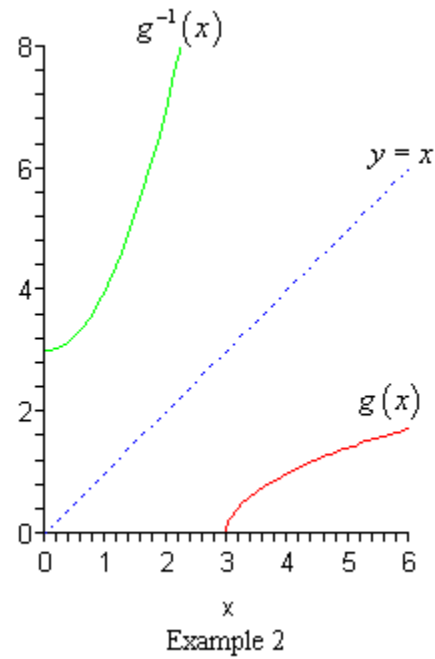
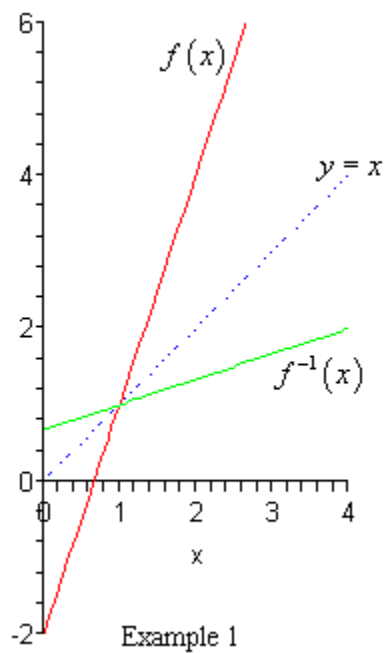
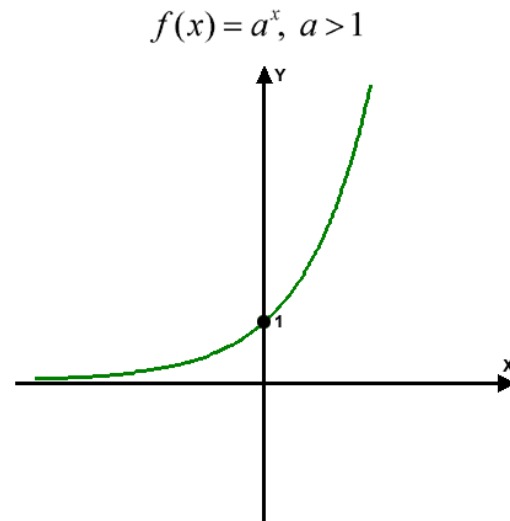
Modelo Matemático

$$y = \frac{a}{x} + b$$

Se utiliza cuando la variable dependiente disminuye a medida que la independiente aumenta, pero sin llegar a cero (relación inversa).

Forma de la Curva





4

Aplicaciones

- Curvas de demanda
- Rendimientos decrecientes
- Procesos físicos inversos

Ejercicio Resuelto

Se tiene la siguiente tabla:

x	y
1	12
2	7
3	5
4	4

Paso 1: Linealización

Si:

$$y = \frac{a}{x} + b$$

Sea:

$$z = \frac{1}{x}$$

Entonces:

$$y = az + b$$

Calculamos z:

x	y	z=1/x
1	12	1
2	7	0.5
3	5	0.33
4	4	0.25

Paso 2: Aplicar regresión lineal a (z,y)

Supongamos que obtenemos:

$$y = 8z + 3$$

Paso 3: Sustituimos

$$y = \frac{8}{x} + 3$$

Modelo ajustado.

3.1.1 Ecuación de Michaelis–Menten

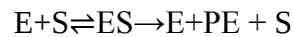
3.1.1.1 Introducción

La ecuación de Michaelis–Menten describe cómo varía la **velocidad de una reacción enzimática** en función de la concentración de sustrato.

Fue propuesta por Leonor Michaelis y Maud Menten en 1913, y es uno de los modelos fundamentales de la cinética enzimática.

3.1.1.2 Esquema básico de reacción

Una enzima (E) se une reversiblemente a su sustrato (S) para formar un complejo enzima-sustrato (ES), que luego produce el producto (P):



3.1.1.3 Supuestos del modelo

1. Se mide la **velocidad inicial** (v_0).
2. La concentración de sustrato es mucho mayor que la de enzima.
3. Se cumple el **estado estacionario**:

$$\frac{d[ES]}{dt} = 0$$

4. No hay acumulación significativa de producto.

3.1.1.4 La ecuación de Michaelis–Menten

$$v_0 = \frac{V_{max}[S]}{K_m + [S]}$$

Donde:

- v_0 = velocidad inicial
- V_{max} = velocidad máxima
- $[S]$ = concentración de sustrato
- K_m = constante de Michaelis

3.1.1.5 Significado biológico de K_m

$$K_m = \frac{k_{-1} + k_2}{k_1}$$

Interpretación práctica:

- Cuando $[S]=K_m$, entonces:

$$v_0 = \frac{V_{max}}{2}$$

K_m es la concentración de sustrato a la cual la velocidad es la mitad de V_{max} .

- K_m bajo $\rightarrow$ alta afinidad enzima-sustrato
- K_m alto $\rightarrow$ baja afinidad

3.1.1.6 Comportamiento de la curva

La gráfica de v_0 vs $[S]$ es una **hipérbola rectangular**:

- A bajas concentraciones $\rightarrow$ cinética de primer orden
- A altas concentraciones $\rightarrow$ cinética de orden cero

Ejercicios Resueltos

Ejercicio 1: Calcular velocidad inicial

Una enzima tiene:

- $V_{max}=120 \mu\text{mol}/\text{min}$
- $K_m=5 \text{ m}$
- $[S]=5 \text{ mM}$

Calcular v_0 .

Solución:

$$\begin{aligned} v_0 &= \frac{120(5)}{5 + 5} \\ v_0 &= \frac{600}{10} \\ v_0 &= 60 \mu\text{mol}/\text{min} \end{aligned}$$

Confirmamos que cuando $[S]=K_m$, entonces $v_0=V_{\max}/2$.

Ejercicio 2: Determinar K_m

Una enzima alcanza:

- $V_{\max} = 80 \mu\text{mol}/\text{min}$
- Cuando $[S] = 2 \text{ mM}$, la velocidad es $40 \mu\text{mol}/\text{min}$

Determinar K_m .

Sabemos que 40 es la mitad de 80, entonces:

$$[S]=K_m$$

$$K_m = 2 \text{ mM}$$

Ejercicio 3: Determinar $V_{\max}$

Datos experimentales:

[S] (mM)	v_0 ($\mu\text{mol}/\text{min}$)
2	20
5	33.3
10	50
20	66.6

Observamos que la velocidad se aproxima a $80 \mu\text{mol}/\text{min}$.

$$V_{\max} \approx 80 \mu\text{mol}/\text{min}$$

Ejercicio 4: Cálculo de k_{cat}

Si:

- $V_{max} = 100 \mu mol/min$
- $[E]_T = 2 \mu mol$

$$k_{cat} = \frac{V_{max}}{[E]_T}$$

$$k_{cat} = \frac{100}{2}$$

$$k_{cat} = 50 min^{-1}$$

3.1.1.7 Eficiencia Catalítica

$$\frac{k_{cat}}{K_m}$$

Esta razón mide qué tan eficiente es una enzima.

El límite teórico es el **límite de difusión** ($\sim 10^8 - 10^9 M^{-1}s^{-1}$).

Forma Lineal: Lineweaver–Burk

Para linealizar la ecuación:

$$\frac{1}{v_0} = \frac{K_m}{V_{max}} \frac{1}{[S]} + \frac{1}{V_{max}}$$

Es una recta tipo:

$$y = mx + b$$

Donde:

- pendiente = K_m/V_{max}
- intercepto = $1/V_{max}$

Ejercicio 5 (Lineweaver–Burk)

Si la recta obtenida es:

$$\frac{1}{v} = 0.5 \frac{1}{[S]} + 0.02$$

Entonces:

$$\frac{1}{V_{max}} = 0.02$$

$$V_{max} = 50$$

$$\frac{K_m}{V_{max}} = 0.5$$

$$K_m = 0.5(50) = 25$$

$$V_{max} = 50$$

$$K_m = 25$$

Resumen Conceptual

Parámetro	Significado biológico
K _m	Afinidad enzima-sustrato
V _{max}	Capacidad máxima catalítica
k _{cat}	Número de recambio
k _{cat} /K _m	Eficiencia catalítica

Ejercicios Propuestos

1. Si K_m = 3 mM y V_{max} = 90 μmol/min, calcular v₀ cuando [S] = 1 mM.
2. Una enzima tiene K_m = 10 mM. ¿Qué ocurre si [S] = 100 mM?
3. Calcular k_{cat} si V_{max} = 200 μmol/min y [E]_T = 4 μmol.
4. Explicar por qué la cinética se vuelve de orden cero a altas concentraciones de sustrato.

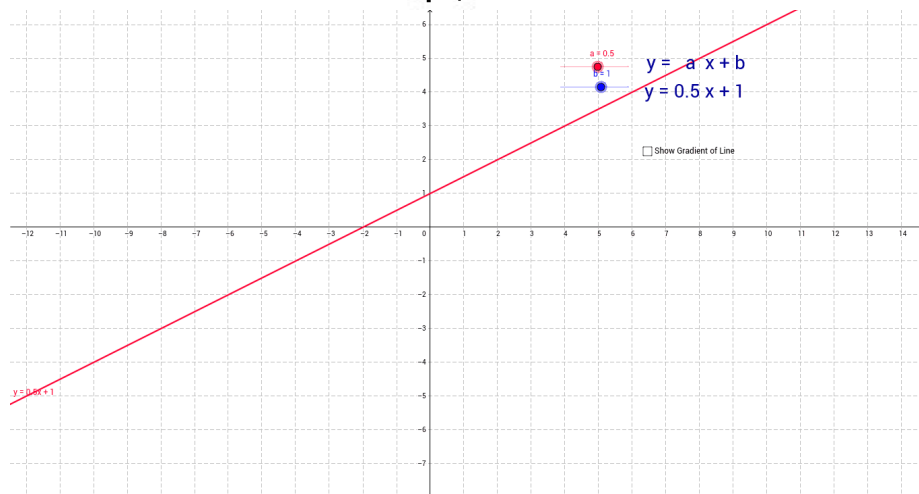
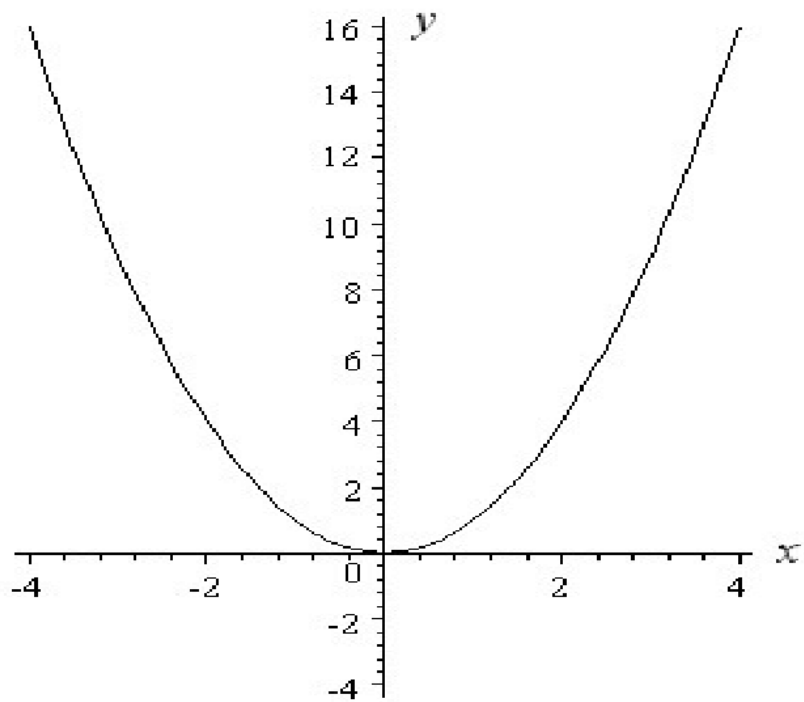
3.2 Regresión Potencial

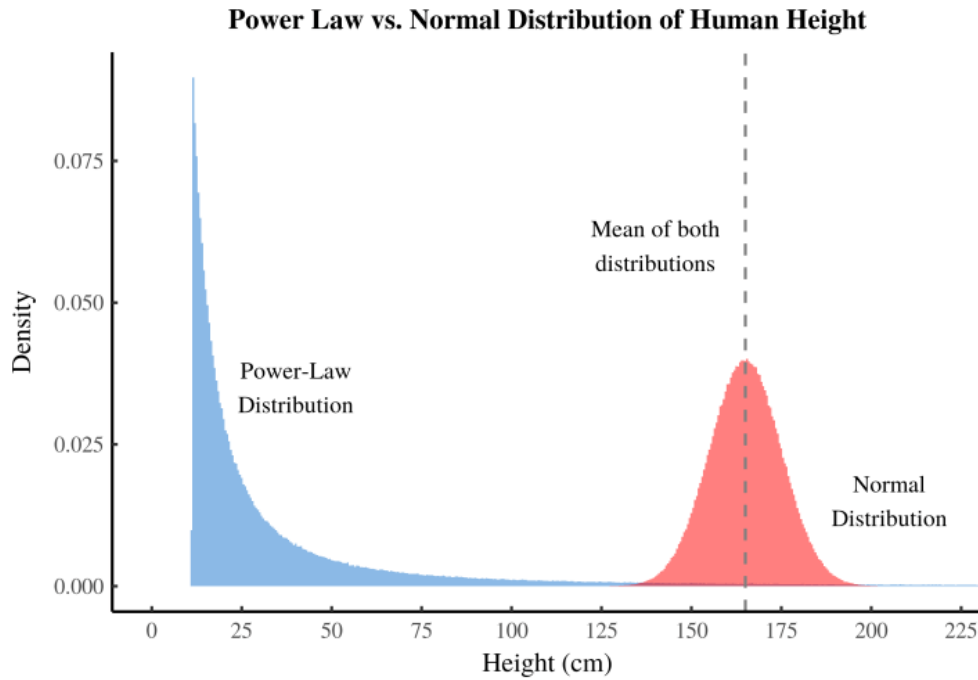
Modelo Matemático

$$y = ax^b$$

Se usa cuando el crecimiento es proporcional a una potencia de x.

Forma de la Curva





4

Aplicaciones

- Leyes físicas
- Elasticidad en economía
- Escalas biológicas

Ejercicio Resuelto

Datos:

x	y
1	2
2	8
3	18
4	32

Paso 1: Linealización

Tomamos logaritmos:

$$\ln y = \ln a + b \ln x$$

Sea:

$$Y = \ln y$$

$$X = \ln x$$

Paso 2: Tabla transformada

x	y	ln x	ln y
1	2	0	0.69
2	8	0.69	2.08
3	18	1.10	2.89
4	32	1.39	3.47

Aplicamos regresión lineal a (ln x, ln y).

Supongamos que obtenemos:

$$\ln y = 0.69 + 2 \ln x$$

Paso 3: Reexpresar modelo

$$a = e^{0.69} = 2$$

Modelo final:

$$y = 2x^2$$

Modelo ajustado.

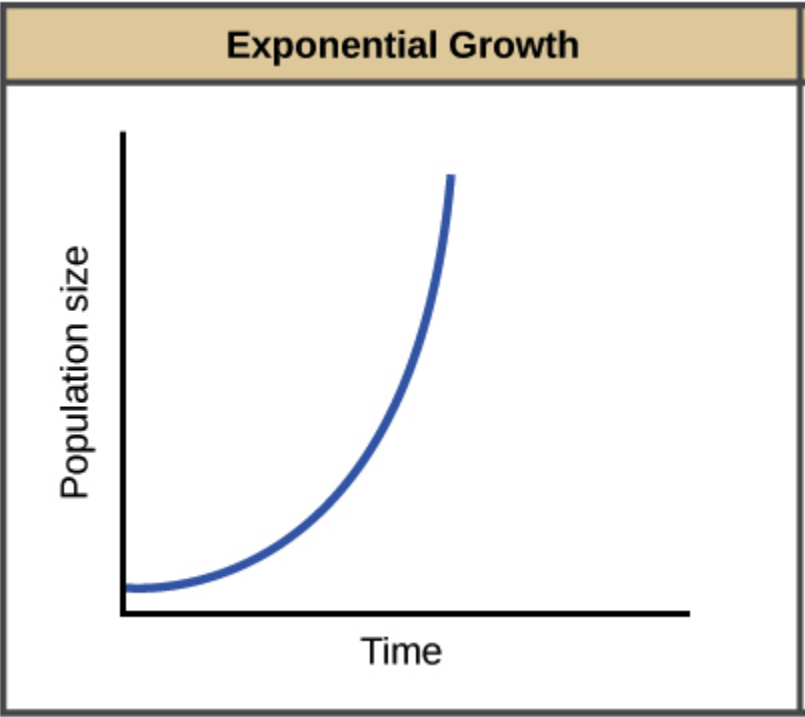
3.3 Regresión Exponencial

Modelo Matemático

$$y = ae^{bx}$$

Se utiliza cuando hay crecimiento o decrecimiento acelerado.

Forma de la Curva



a = 1

1

10

×

b = 1

-5

5

⊖

c = 1

-5

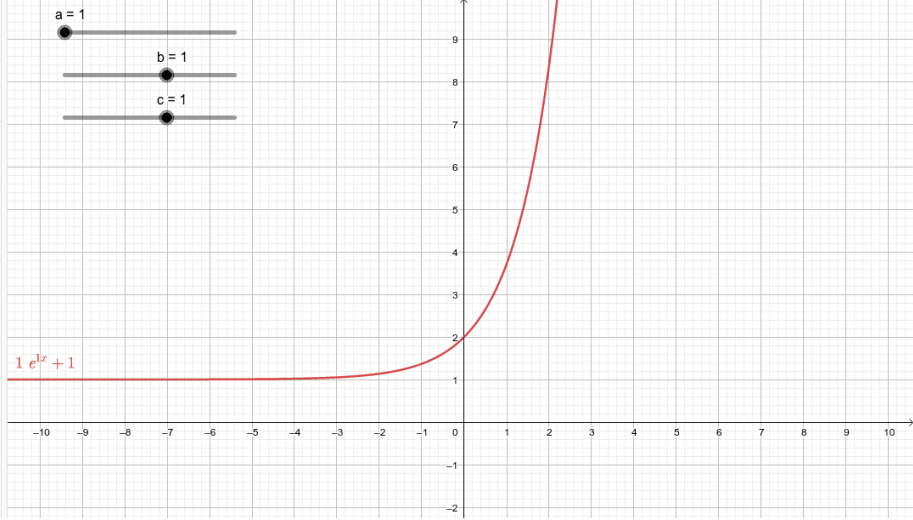
5

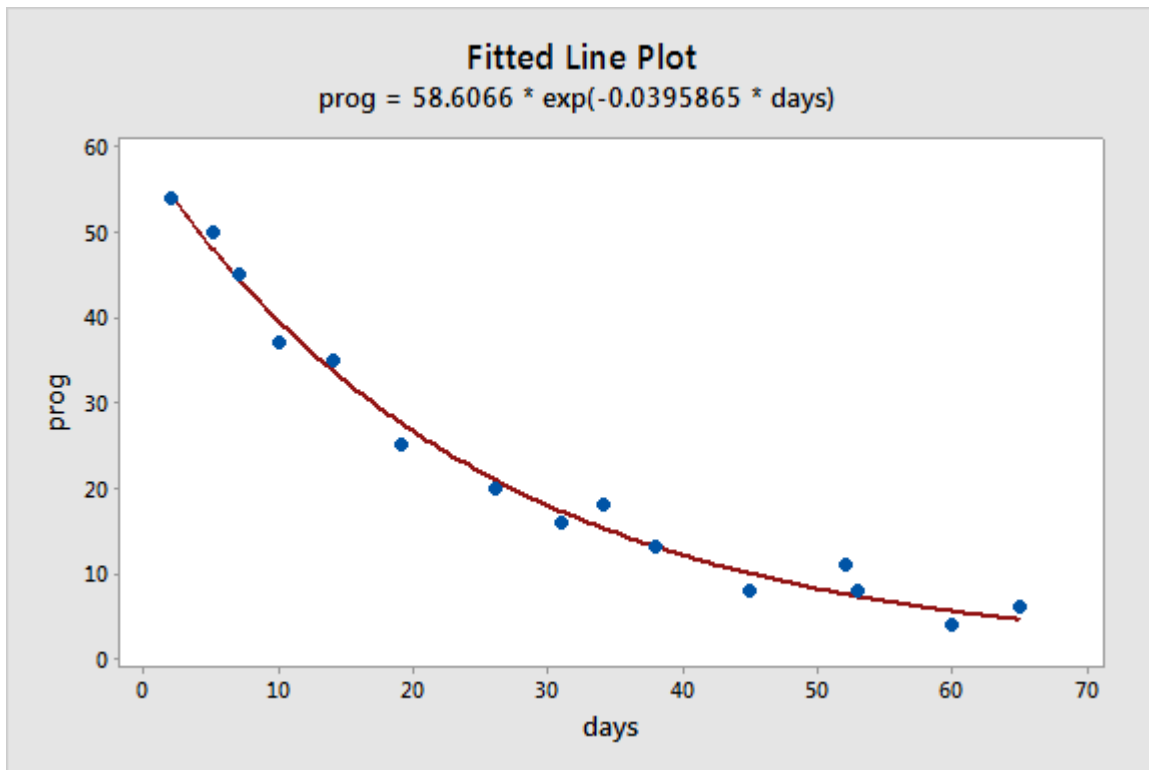
⊖

$f(x) = a e^{bx} + c$

→ $1 e^{1x} + 1$

×





Aplicaciones

- Crecimiento poblacional
- Interés compuesto
- Desintegración radiactiva

Ejercicio Resuelto

Datos:

x	y
1	3
2	6
3	12
4	24

Paso 1: Linealización

$$\ln y = \ln a + bx$$

Sea:

$$Y = \ln y$$

x	y	ln y
1	3	1.10
2	6	1.79
3	12	2.48
4	24	3.18

Aplicamos regresión lineal a (x, ln y).

Supongamos:

$$\ln y = 0 + 0.69x$$

Paso 2: Modelo final

$$a = e^0 = 1$$

$$y = e^{0.69x}$$

Modelo ajustado.

3.4 Regresión Logística

Modelo Matemático

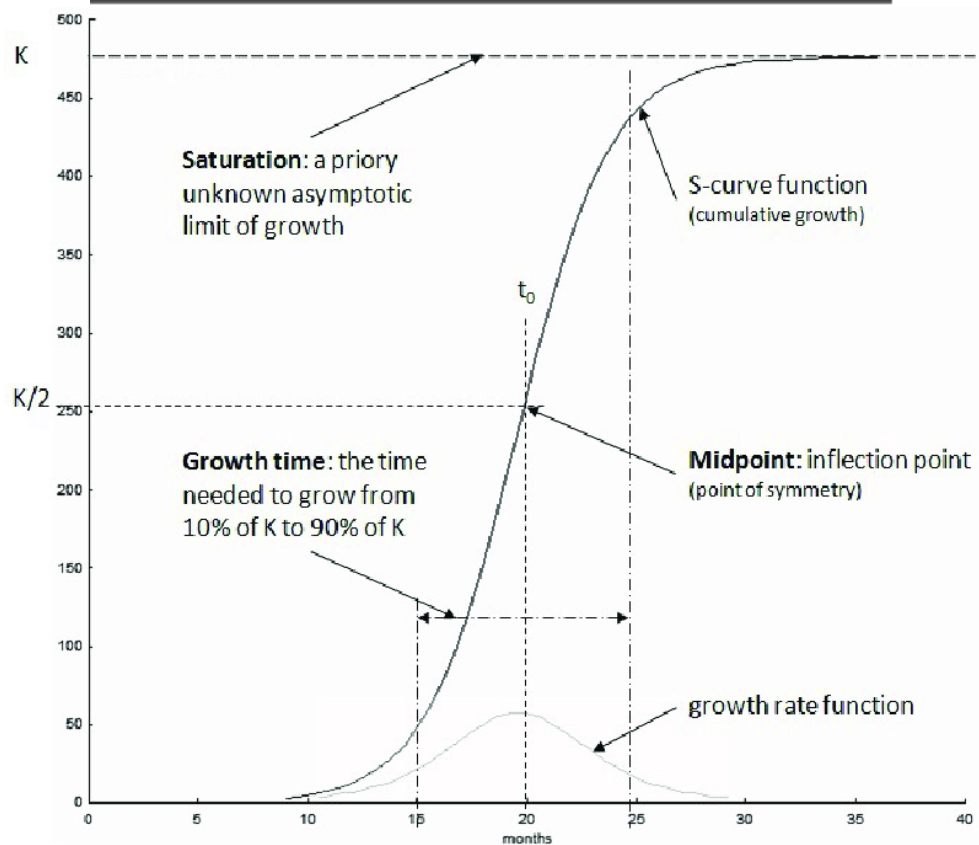
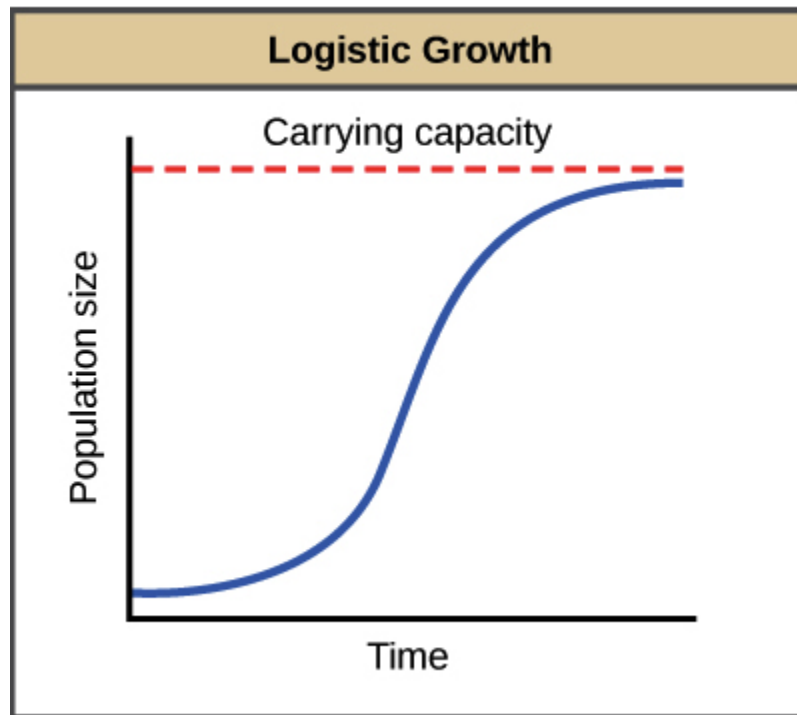
$$y = \frac{L}{1 + ae^{-bx}}$$

Donde:

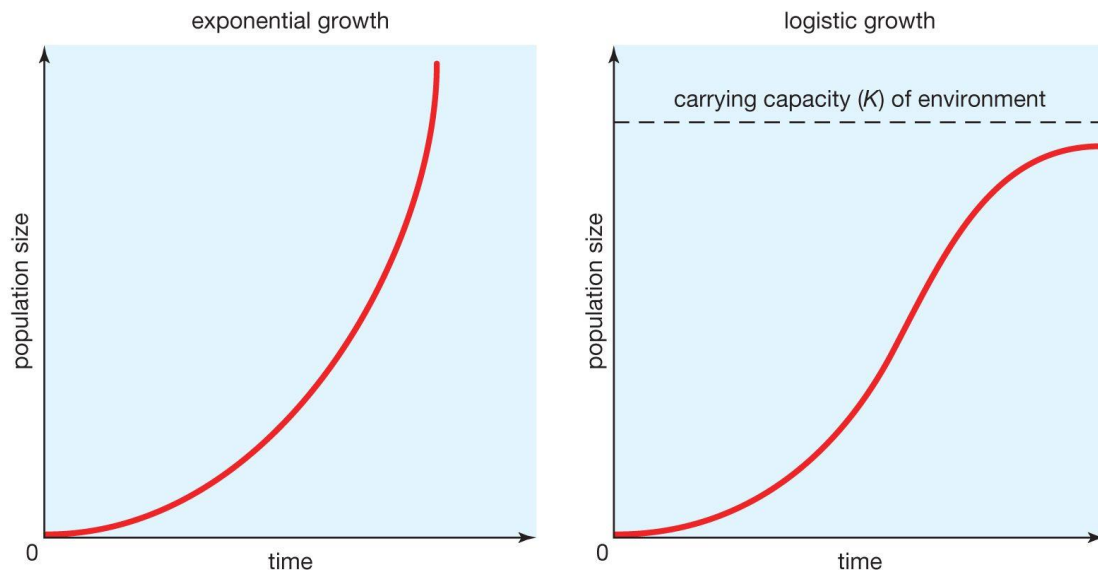
- L = valor máximo (capacidad límite)
- a y b = parámetros

Se usa cuando el crecimiento se desacelera al acercarse a un límite.

Forma de la Curva



Exponential versus logistic population growth



© 2012 Encyclopædia Britannica, Inc.

Aplicaciones

- Crecimiento poblacional con límite
- Difusión de tecnología
- Modelos epidemiológicos

Ejercicio Resuelto

Datos aproximados de adopción tecnológica:

x	y
1	5
2	12
3	25
4	40
5	48

Supongamos que el límite máximo es $L = 50$.

Paso 1: Transformación logística

$$\ln \left(\frac{L - y}{y} \right) = \ln a - bx$$

Calculamos:

$$Z = \ln \left(\frac{50 - y}{y} \right)$$

Se aplica regresión lineal entre (x, Z).

Supongamos que obtenemos:

$$Z = 2 - 0.8x$$

Entonces:

$$a = e^2$$

Modelo final:

$$y = \frac{50}{1 + e^2 e^{-0.8x}}$$

Modelo ajustado.

Comparación General

Modelo	Forma	Uso Principal
Hipérbola	Inversa	Rendimientos decrecientes
Potencial	Curva potencia	Elasticidades
Exponencial	Crecimiento rápido	Interés, población
Logística	Forma S	Crecimiento con límite

Conclusión

La regresión no lineal permite modelar fenómenos reales donde:

- No existe relación proporcional constante.
- Hay límites máximos.
- Hay crecimiento acelerado.
- Existen rendimientos decrecientes.

4.- Correlación

La **correlación** mide el grado de relación entre dos variables cuantitativas.

Indica **dirección** (positiva o negativa) e **intensidad** (débil, moderada o fuerte), pero **no implica causalidad**.

4.1 Introducción

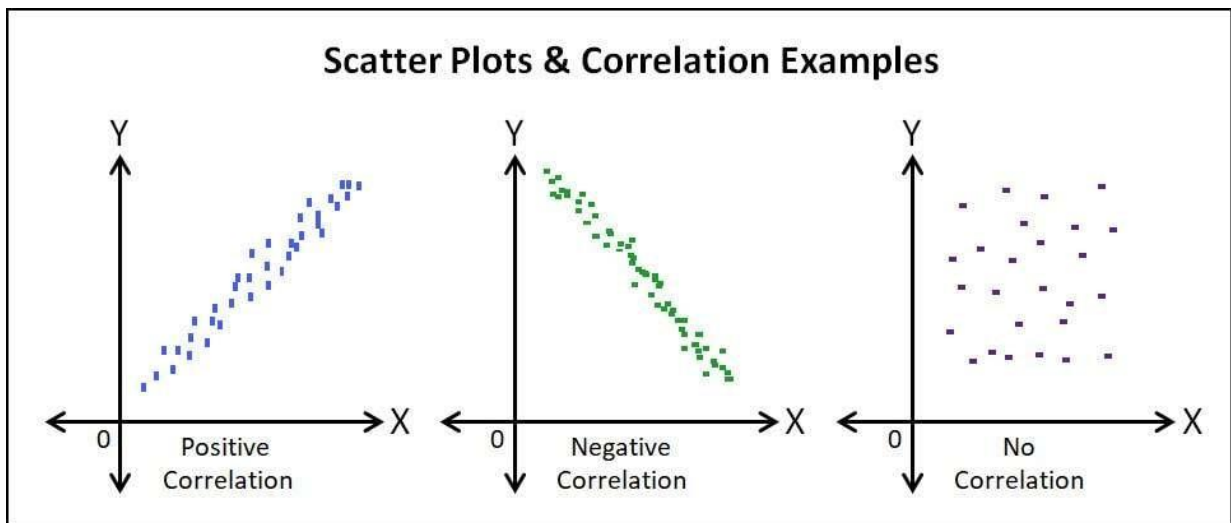
¿Qué es la correlación?

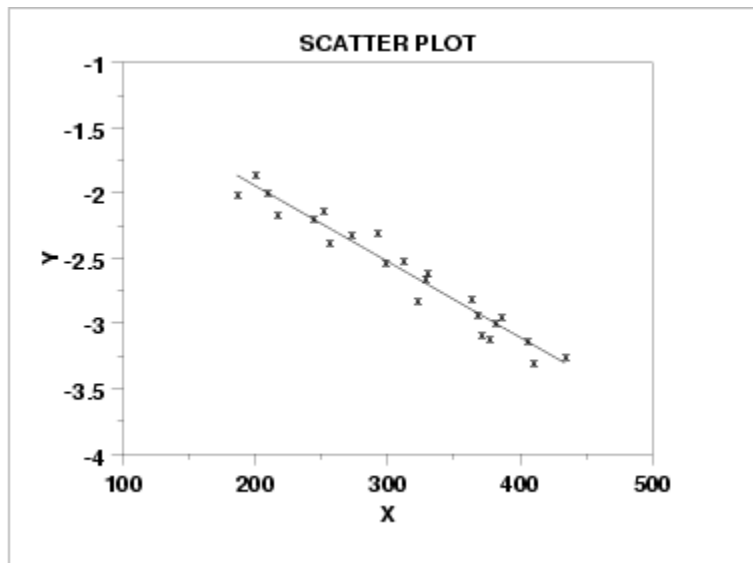
Es una medida estadística que cuantifica la relación lineal entre dos variables X y Y .

Se representa con:

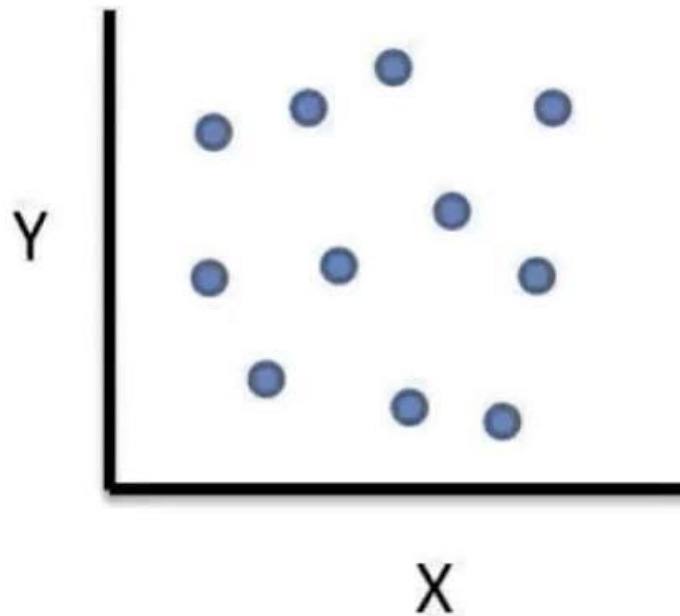
- r → coeficiente de correlación muestral (Pearson)
- ρ (**rho**) → coeficiente de correlación poblacional

Tipos de correlación





No Correlation



Tipo	Característica
Positiva	Cuando X aumenta, Y aumenta
Negativa	Cuando X aumenta, Y disminuye
Nula	No existe relación lineal

Rango del coeficiente r

$$-1 \leq r \leq 1$$

- $r = 1 \rightarrow$ correlación positiva perfecta
- $r = -1 \rightarrow$ correlación negativa perfecta
- $r = 0 \rightarrow$ no hay correlación lineal

4.2 Obtención de R (Coeficiente de Pearson)

✂ Fórmula

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

Donde:

- n = número de datos
- Σ = sumatoria

Ejercicio Resuelto

Se estudia la relación entre horas de estudio (X) y calificación (Y).

X	Y
2	65
4	70
6	75
8	85
10	95

Paso 1: Tabla auxiliar

X	Y	X ²	Y ²	XY
2	65	4	4225	130
4	70	16	4900	280
6	75	36	5625	450
8	85	64	7225	680
10	95	100	9025	950

Sumatorias:

$$\begin{aligned}\sum X &= 30 \\ \sum Y &= 390 \\ \sum X^2 &= 220\end{aligned}$$

$$\begin{aligned}\sum Y^2 &= 31000 \\ \sum XY &= 2490 \\ n &= 5\end{aligned}$$

Paso 2: Sustituir en la fórmula

$$\begin{aligned}r &= \frac{5(2490) - (30)(390)}{\sqrt{[5(220) - 30^2][5(31000) - 390^2]}} \\ &= \frac{12450 - 11700}{\sqrt{[1100 - 900][155000 - 152100]}} \\ &= \frac{750}{\sqrt{(200)(2900)}} \\ &= \frac{750}{\sqrt{580000}} \\ &= \frac{750}{761.58} \\ r &\approx 0.984\end{aligned}$$

Interpretación preliminar

Existe una **correlación positiva muy fuerte** entre horas de estudio y calificación.

4.3 Prueba de Hipótesis para ρ

Ahora verificamos si la correlación es estadísticamente significativa.

Hipótesis

$$\begin{aligned}H_0: \rho &= 0 \text{ (No hay correlación)} \\ H_1: \rho &\neq 0 \text{ (Sí hay correlación)}\end{aligned}$$

Estadístico de prueba

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Con:

- $n - 2$ grados de libertad

Continuamos el ejercicio anterior

$$r = 0.984$$

$$n = 5$$

Paso 1: Sustituir

$$t = \frac{0.984\sqrt{5-2}}{\sqrt{1-(0.984)^2}}$$

$$t = \frac{0.984\sqrt{3}}{\sqrt{1-0.968}}$$

$$t = \frac{0.984(1.732)}{\sqrt{0.032}}$$

$$t = \frac{1.704}{0.179}$$

$$t \approx 9.52$$

Paso 2: Valor crítico

Grados de libertad:

$$gl = 5 - 2 = 3$$

Para $\alpha = 0.05$ (bilateral):

$$t_{crítico} \approx 3.182$$

Paso 3: Decisión

Como:

$$9.52 > 3.182$$

Se **rechaza H_0** .

4.3.3 Interpretación

✓ Existe evidencia estadística suficiente para afirmar que la correlación es significativa.

Guía de interpretación de r

Valor de r	Interpretación
0.00 – 0.19	Muy débil
0.20 – 0.39	Débil
0.40 – 0.59	Moderada
0.60 – 0.79	Fuerte
0.80 – 1.00	Muy fuerte

(Mismos rangos para valores negativos)

Consideraciones importantes

1. Correlación $\neq$ causalidad.
2. Solo mide relación lineal.
3. Puede verse afectada por valores atípicos.
4. Se recomienda visualizar con diagrama de dispersión.

Conclusión General

La correlación permite:

- Medir intensidad y dirección.
- Evaluar relación lineal.
- Comprobar significancia estadística.
- Servir de base para regresión.

5.- Métodos No Paramétricos

Los **métodos no paramétricos** son técnicas estadísticas que **no requieren que los datos sigan una distribución normal** ni que se conozcan parámetros poblacionales como la media o la varianza.

Se utilizan cuando:

- La muestra es pequeña.
- Los datos son ordinales.
- No se cumple normalidad.
- Existen valores atípicos.

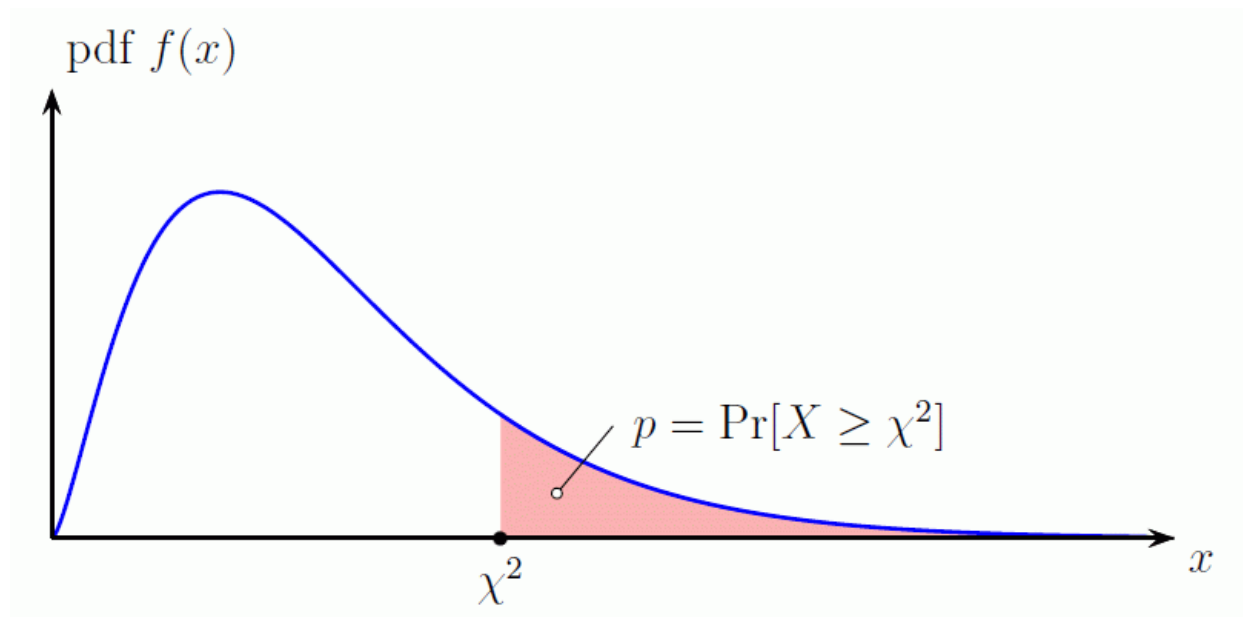
5.1 Prueba Ji-cuadrada (χ^2)

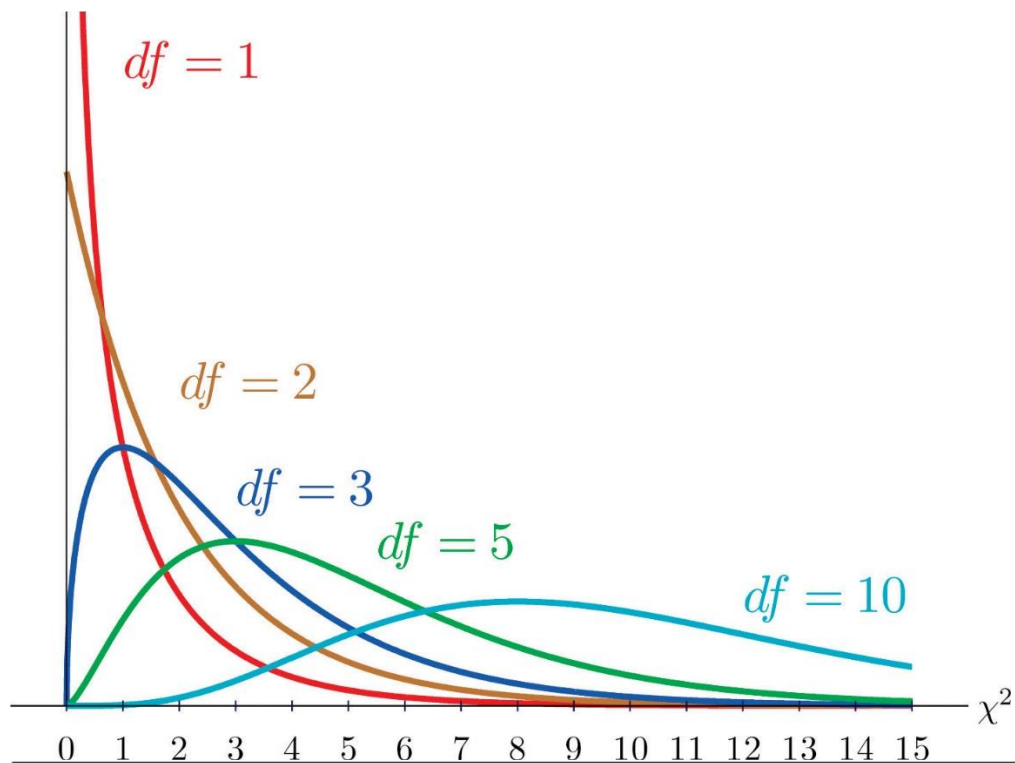
✂ ¿Para qué sirve?

Se utiliza para:

- Probar independencia entre variables cualitativas.
- Comparar frecuencias observadas vs. esperadas.
- Probar bondad de ajuste.

Distribución χ^2





SEXPREF				
	Males	Females	Both	
Observed	4	2	3	9
Expected	2.1	5.4	1.5	
(O-E) ² /E	1.72	2.14	1.5	
Observed	3	16	2	21
Expected	4.9	12.6	3.5	
(O-E) ² /E	.74	.92	.64	
	7	18	5	30

Fórmula

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Donde:

- O = frecuencia observada
- E = frecuencia esperada

Ejercicio Resuelto (Independencia)

Se estudia si el género influye en preferencia de producto.

	Producto A	Producto B	Total
Hombres	30	20	50
Mujeres	20	30	50
Total	50	50	100

Paso 1: Calcular frecuencias esperadas

$$E = \frac{(Total\ fila)(Total\ columna)}{Total\ general}$$

Para Hombres-A:

$$E = \frac{50 \times 50}{100} = 25$$

Tabla esperada:

	A	B
Hombres	25	25
Mujeres	25	25

Paso 2: Calcular χ^2

$$\chi^2 = \frac{(30-25)^2}{25} + \frac{(20-25)^2}{25} + \frac{(20-25)^2}{25} + \frac{(30-25)^2}{25}$$

$$\chi^2 = 1 + 1 + 1 + 1 = 4$$

Paso 3: Grados de libertad

$$gl = (r - 1)(c - 1) = (2 - 1)(2 - 1) = 1$$

Valor crítico ($\alpha=0.05$): 3.84

Como:

$$4 > 3.84$$

Se rechaza $H_0 \rightarrow$ Las variables no son independientes.

5.2 Prueba de Rangos con Signo de Wilcoxon

Alternativa no paramétrica a la prueba t para muestras relacionadas.

Se usa cuando:

- Datos pareados.
- No hay normalidad.

Esquema del procedimiento

Tested Pair ^a	Z	Asymp. Sig. (2-tailed)	Tested Pair ^a	Z	Asymp. Sig. (2-tailed)	Tested Pair ^a	Z	Asymp. Sig. (2-tailed)
1 - 2	-5.604 ^b	.000	2 - 9	-1.554 ^c	.120	5 - 6	-8.348 ^b	.000
1 - 3	-4.802 ^b	.000	2 - 10	-2.176 ^c	.030	5 - 7	-2.482 ^b	.013
1 - 4	-.703 ^b	.482	3 - 4	-4.879 ^c	.000	5 - 8	-8.103 ^b	.000
1 - 5	-5.635 ^c	.000	3 - 5	-7.986 ^c	.000	5 - 9	-7.354 ^b	.000
1 - 6	-5.144 ^b	.000	3 - 6	-.614 ^b	.540	5 - 10	-7.562 ^b	.000
1 - 7	-3.655 ^c	.000	3 - 7	-8.057 ^c	.000	6 - 7	-7.613 ^c	.000
1 - 8	-4.790 ^b	.000	3 - 8	-.622 ^b	.534	6 - 8	-.190 ^c	.849
1 - 9	-2.939 ^b	.003	3 - 9	-1.847 ^c	.065	6 - 9	-2.822 ^c	.005
1 - 10	-3.962 ^b	.000	3 - 10	-1.933 ^c	.053	6 - 10	-2.314 ^c	.021
2 - 3	-.426 ^c	.670	4 - 5	-5.949 ^c	.000	7 - 8	-7.935 ^b	.000
2 - 4	-5.077 ^c	.000	4 - 6	-4.361 ^b	.000	7 - 9	-6.962 ^b	.000
2 - 5	-8.593 ^c	.000	4 - 7	-4.342 ^c	.000	7 - 10	-6.451 ^b	.000
2 - 6	-.912 ^b	.362	4 - 8	-4.448 ^b	.000	8 - 9	-3.165 ^c	.002
2 - 7	-7.867 ^c	.000	4 - 9	-2.321 ^b	.020	8 - 10	-2.299 ^c	.021
2 - 8	-.578 ^b	.563	4 - 10	-3.202 ^b	.001	10 - 9	-.018 ^b	.985

Note: a - Post-hoc testing based on Wilcoxon signed ranks test, b - Based on negative ranks, c - Based on positive ranks

Source: Own processing on the basis of the respondents' answers

WILCOXON SIGNED RANK TEST - MINI EXAMPLE

	score_1	score_2	difference	absolute_difference	rank	signed_rank
1	8	10	-2	2	3.0	-3.0
2	5	4	1	1	1.5	1.5
3	7	6	1	1	1.5	1.5
4	9	5	4	4	4.0	4.0
5	8	8	0	.	.	.
6						

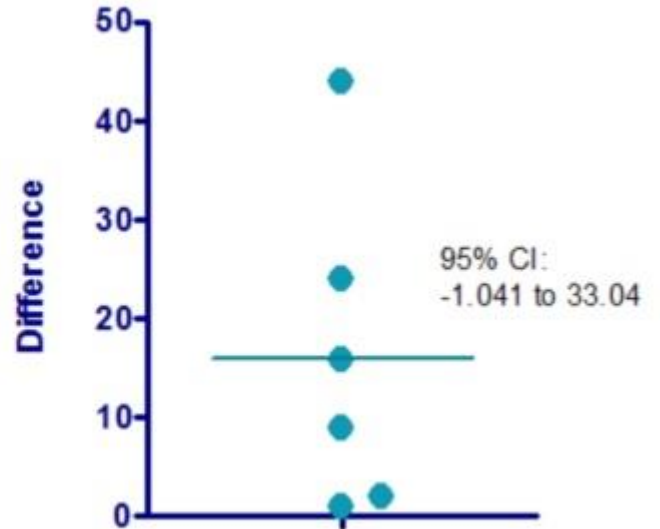
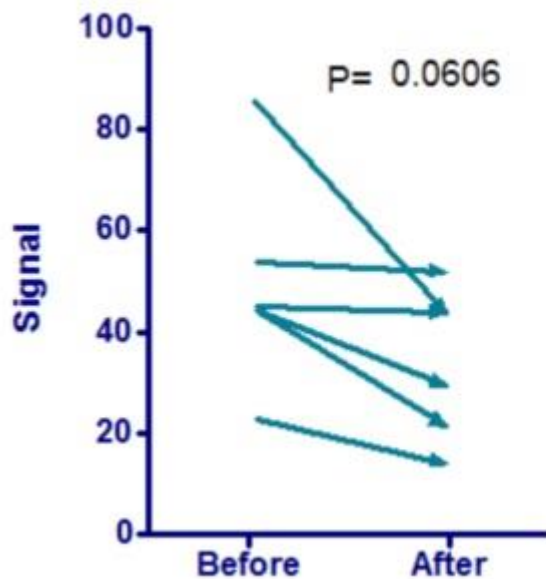
Ranks			
	N	Mean Rank	Sum of Ranks
Negative Ranks	1 ^a	3.00	3.00
Positive Ranks	3 ^b	2.33	7.00
Ties	1 ^c		
Total	5		

a. score_1 < score_2

b. score_1 > score_2

c. score_1 = score_2

Test statistic = **Wilcoxon W+** =
sum over positive signed ranks = **7**.



Ejercicio Resuelto

Medimos presión arterial antes y después de un tratamiento:

Antes	Después
150	140
160	150
145	142
155	150

Antes	Después
170	160

Paso 1: Diferencias

Diferencia
-10
-10
-3
-5
-10

Paso 2: Ordenar por valor absoluto

Dif	Dif	Rango
-3	3	1
-5	5	2
-10	10	3.5
-10	10	3.5
-10	10	3.5

(Se promedian rangos repetidos)

Paso 3: Sumar rangos negativos

Todos son negativos:

$$T = 1 + 2 + 3.5 + 3.5 + 3.5 = 13.5$$

Se compara con tabla crítica (n=5, $\alpha=0.05$).

Se concluye que hay diferencia significativa.

5.3 Prueba U de Mann-Whitney

¿Para qué sirve?

Alternativa no paramétrica a la prueba t para muestras independientes.

Concepto gráfico

Gender	Ranks
female	1
female	2.5
male	2.5
male	4
male	5
female	7
female	7
male	7
male	9
male	10
female	11

Calculation of the rank sums

$$T_{female} = 1 + 2.5 + 7 + 7 + 11 = 28.5$$

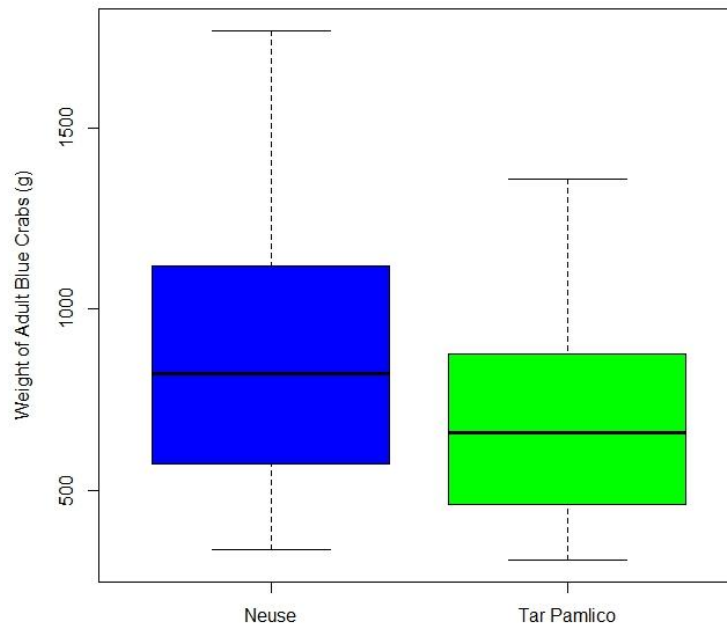
$$T_{male} = 2.5 + 4 + 5 + 7 + 9 + 10 = 37.5$$

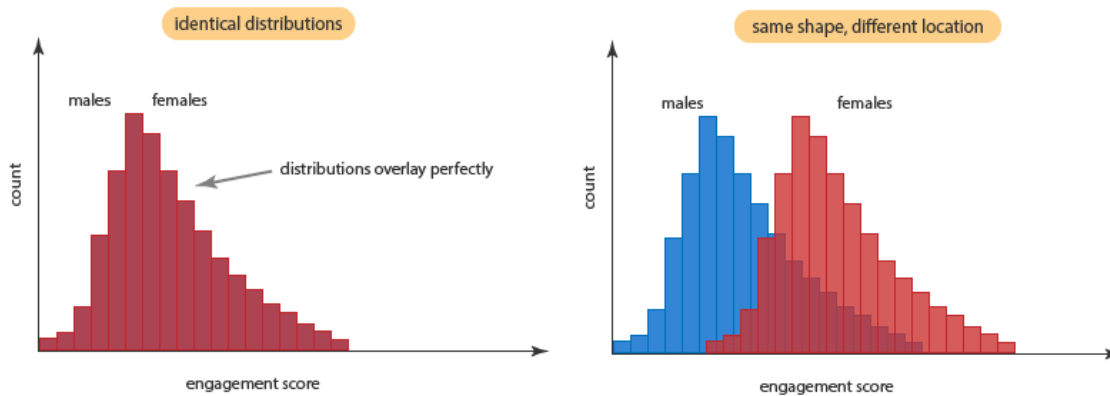
Number of tied
ranks

Number of people
sharing rank i

$$\sum_{i=1}^k \frac{t_i^3 - t_i}{12} = \frac{2^3 - 2}{12} + \frac{3^3 - 3}{12} = 2.5$$

Weight of Adult Blue Crabs in the Neuse and Tar Pamlico Basins





Ejercicio Resuelto

Grupo A: 5, 7, 8

Grupo B: 2, 4, 6

Paso 1: Ordenar y asignar rangos

Datos combinados:

2(1), 4(2), 5(3), 6(4), 7(5), 8(6)

Suma de rangos:

- Grupo A: $3 + 5 + 6 = 14$
- Grupo B: $1 + 2 + 4 = 7$

Paso 2: Fórmula

$$U_A = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_A$$

$$U_A = 3 \times 3 + \frac{3 \times 4}{2} - 14$$

$$U_A = 9 + 6 - 14 = 1$$

Se compara con tabla crítica.

Si $U_{\text{calculado}} < U_{\text{crítico}} \rightarrow$ diferencia significativa.

5.4 Prueba H de Kruskal-Wallis

¿Para qué sirve?

Alternativa no paramétrica al ANOVA.

Se usa cuando:

- Tres o más grupos independientes.
- No hay normalidad.

Concepto gráfico

Kruskal-Wallis Test

Ranks

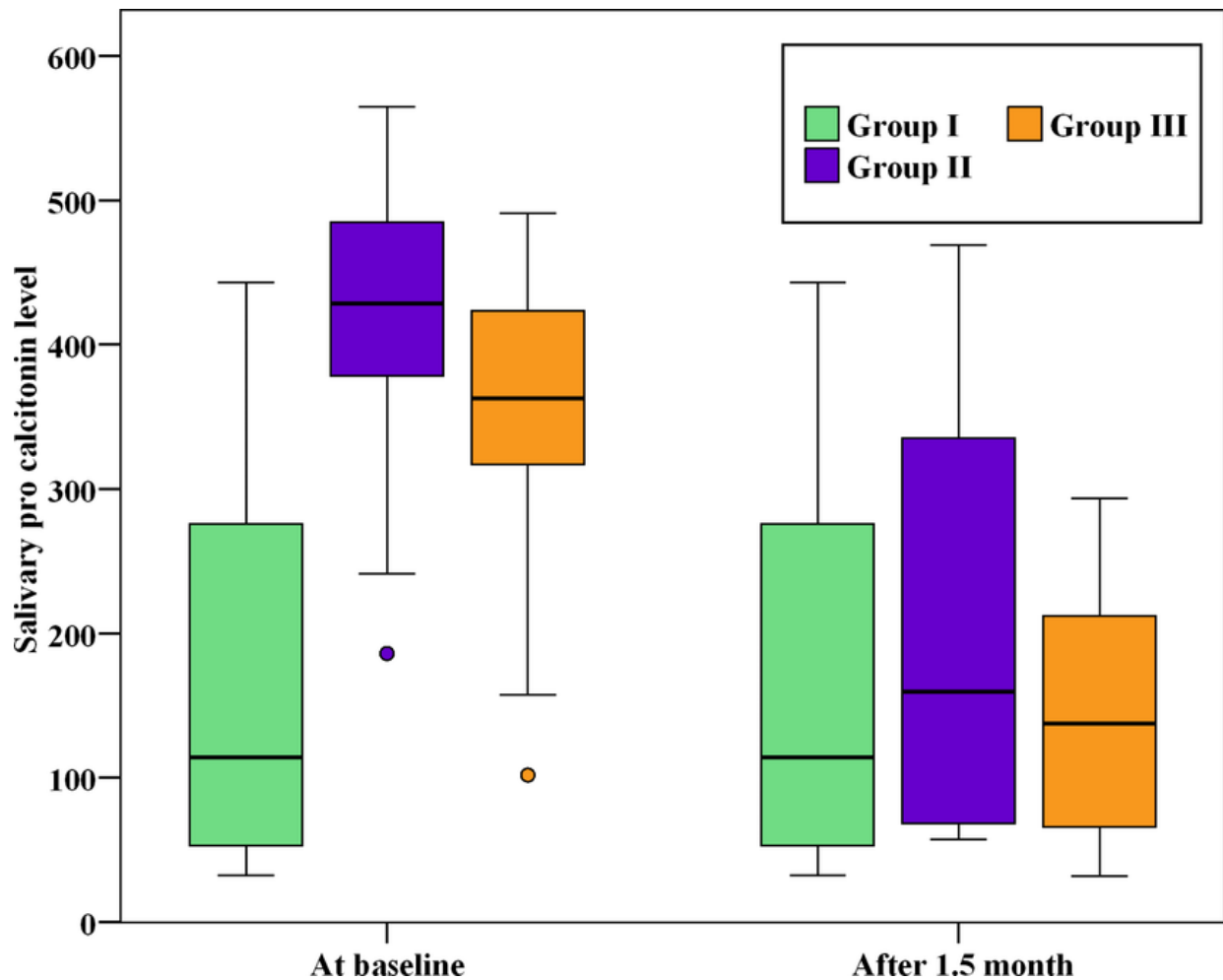
Drug Treatment Group		N	Mean Rank
Pain_Score	Drug A	20	35.33
	Drug B	20	34.83
	Drug C	20	21.35
	Total	60	

Test Statistics^{a,b}

	Pain_Score
Chi-square	8.520
df	2
Asymp. Sig.	.014

a. Kruskal Wallis Test

b. Grouping Variable:
Drug Treatment
Group



Feature	Parametric Tests	Non-Parametric Tests
Assumptions	Requires normal distribution and equal variance	No strict distribution assumptions
Examples	t-tests, ANOVA	Kruskal-Wallis, Chi-square
Data Type	Continuous (interval/ratio)	Ordinal, categorical, or skewed continuous
Power	More powerful if assumptions are met	Less powerful but more flexible
Handling of Outliers	Sensitive to outliers	More robust to outliers
Sample Size	Works best with larger samples	Suitable for small sample sizes
Best Used When	Data follows a normal distribution and variance is equal across groups	Data is skewed, ordinal, or categorical, or assumptions of parametric tests are violated

Fórmula

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1)$$

Ejercicio Resuelto

Tres métodos de enseñanza:

Grupo 1: 5,6

Grupo 2: 7,8

Grupo 3: 9,10

Paso 1: Rangos

5(1),6(2),7(3),8(4),9(5),10(6)

Sumas de rangos:

- $R_1 = 1+2=3$
- $R_2 = 3+4=7$
- $R_3 = 5+6=11$

Paso 2: Sustituir

$$\begin{aligned} N &= 6 \\ H &= \frac{12}{6(7)} \left(\frac{3^2}{2} + \frac{7^2}{2} + \frac{11^2}{2} \right) - 3(7) \\ H &= \frac{12}{42} (4.5 + 24.5 + 60.5) - 21 \\ H &= \frac{12}{42} (89.5) - 21 \\ H &= 25.57 - 21 \\ H &\approx 4.57 \end{aligned}$$

Se compara con χ^2 con gl=2.

Comparación General

Prueba	Uso
χ^2	Variables categóricas
Wilcoxon	Muestras relacionadas
Mann-Whitney	2 muestras independientes

Prueba	Uso
Kruskal-Wallis	3 o más muestras independientes

Conclusión

Los métodos no paramétricos:

No requieren normalidad

Funcionan con datos ordinales

Son robustos ante valores extremos

Son alternativa cuando fallan los métodos paramétricos

Bibliografía Básica

1. Box, G. E. P, Hunter, W. G., and Hunter J. S. Estadística para investigadores. Reverté, 1993.
2. Daniel, W. Bioestadística, bases para el análisis de las ciencias de la salud. México, Limusa, 2009.
3. Hawkins, D. Biomeasurement: A Student's Guide to Biostatistics. Oxford University Press, 2014.
4. Milton, J. S. Estadística para Biología y ciencias de la salud. 3ª ed. España, McGraw-Hill Interamericana, 2007.
5. Sokal, R. R. y Rohlf, J. F. Introducción a la Bioestadística. México, Reverté, 2000.
6. Triola, M. F. Estadística. México, Trillas, 2004.
7. Weimer, R. Estadística. México, Patria, 2009.
8. Zar, J. Biostatistical Analysis. 5ª ed. USA, Pearson Prentice Hall, 2010.

Diseño y ANOVA

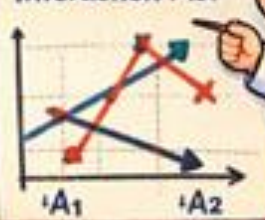
Diseño Factorial

Factorial $A \times B$

	A1	A2	B2
B1	$A \neq B$		
B1	$A \neq B$		

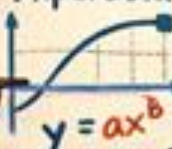
$A \times B$

Interacción Plot



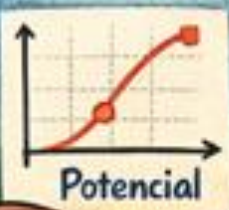
Regresión No Lineal

Hipérbola



$$y = ax^b$$

Potencial



$$y = \beta_0 + \beta_1 x$$



$$H_0: \beta_1 = 0?$$
$$H_1: \beta_1 = 0!$$

Correlación

$$r = 0.8$$

$$p < 0.05$$

Interpretación

Métodos No Paramétricos

$$\chi^2$$

Wilcoxon

Mann-Whitney
Kruskal-Wallis



Carrera de BIOLOGÍA

