

L U M B R E



Regresión y correlación inferida

Grado: 3

Materia: Estadística Inferencial · Probabilidad y Estadística

11/10/2026

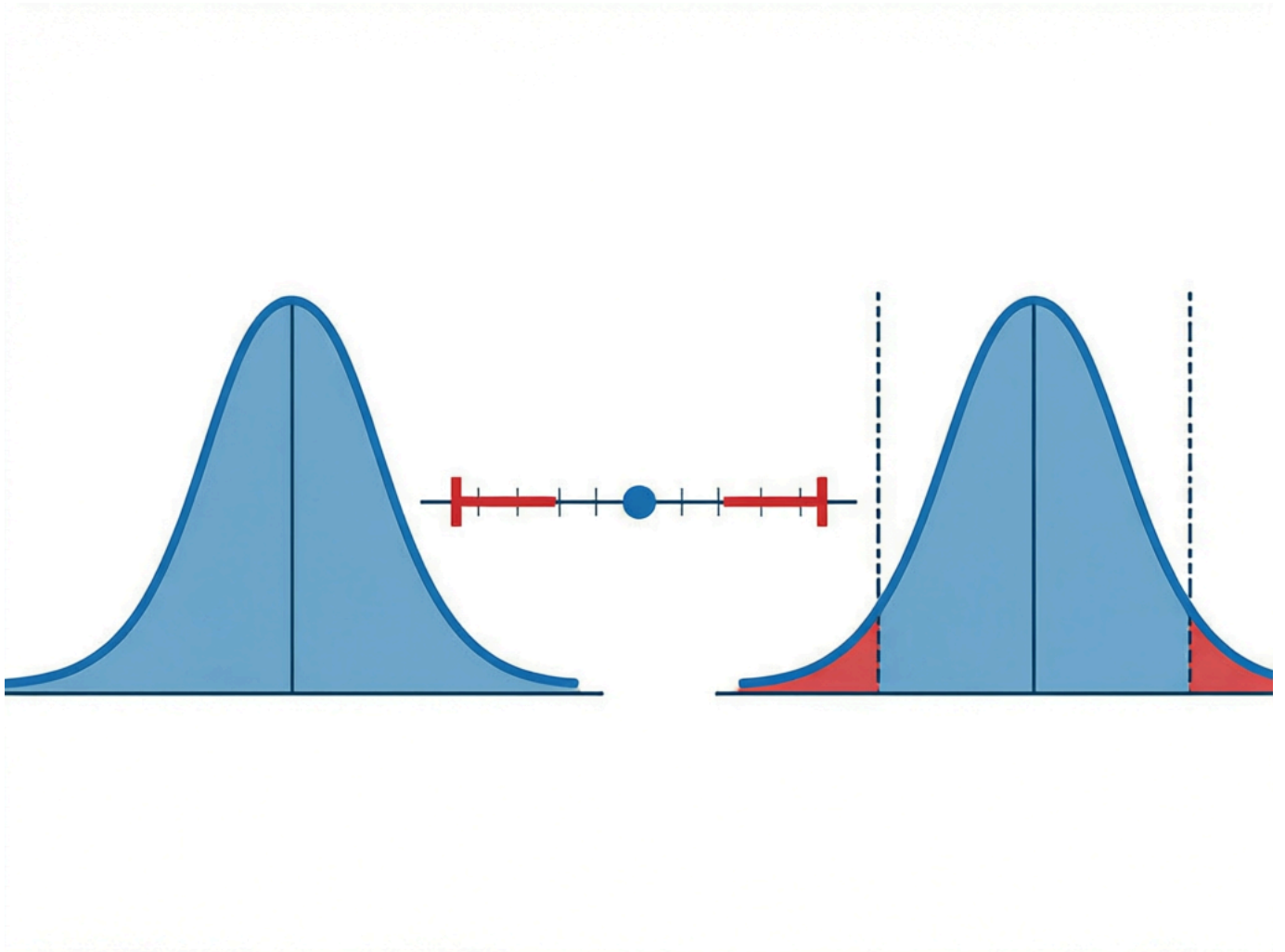
Antes de empezar

Ajustar una recta a los datos es fácil; generalizar más allá de la muestra es lo delicado. Esta lección une la regresión con la inferencia: qué podemos afirmar del coeficiente, con qué confianza y hasta dónde.

QUÉ APRENDERÁS

- Inferir a partir de un modelo de regresión: interpretar coeficientes y su incertidumbre, contrastar la relación y no extrapolar fuera del rango observado.

Regresión y correlación inferida



Estadística inferencial: estimación, pruebas de hipótesis y regresión.

Mapa conceptual

- **Estadística Inferencial**
 - Muestreo

- Aleatorio
- Estratificado
- Estimación
 - Puntual
 - Intervalos
- Hipótesis
 - H_0 y H_1
 - Valor p
 - Potencia
- Pruebas
 - t-Student
 - χ^2
 - ANOVA
- Regresión
 - Pendiente
 - Correlación

El mapa mental interactivo es interactivo en la versión online.



De la muestra a la relación subyacente

Una recta ajustada a tus datos es una foto de la muestra. La inferencia pregunta por la recta de la población: dónde está realmente su pendiente, cuán segura es, y si atreverse a predecir fuera de lo observado.

Del ajuste a la inferencia

La regresión de mínimos cuadrados estima la recta que mejor describe la muestra. Pero esa recta cambia de muestra a muestra. La inferencia trata los coeficientes como estimaciones de verdades poblacionales, con su error estándar, su intervalo y su prueba de hipótesis.

Interpretar la pendiente

La pendiente dice cuánto cambia, en promedio, la variable respuesta al aumentar en una unidad el predictor, manteniendo lo demás. Es un efecto condicional, no una promesa causal: describe asociación dentro de los datos observados.

Incertidumbre del coeficiente

Cada coeficiente lleva un error estándar que mide cuánto variaría de muestra a muestra. Un coeficiente grande pero con error enorme es poco fiable; uno pequeño pero preciso puede ser sólido. La magnitud sin su incertidumbre no dice nada.

Probar si el coeficiente difiere de cero

Se divide el coeficiente entre su error estándar para obtener un estadístico t , que se contrasta con su distribución bajo «pendiente cero». Un p pequeño indica que la relación lineal es improbable que sea inexistente en la población.

Intervalo de confianza del coeficiente

El coeficiente más o menos un crítico por su error estándar da el rango plausible de la pendiente poblacional. Si el intervalo incluye cero, la relación no está claramente establecida. Reportar el intervalo supera al simple «significativo sí o no».

La R cuadrada y su lectura

R cuadrada es la fracción de la varianza de la respuesta explicada por el modelo. Alta no implica causalidad ni buen pronóstico fuera de la muestra, y en regresión múltiple sube mecánicamente al añadir predictores. Hay que juzgarla con mesura y con la R cuadrada ajustada.

Predicción frente a explicación

Un modelo puede predecir bien sin explicar nada (correlaciones espurias útiles) o explicar sin predecir mucho. Saber qué buscas —pronóstico o comprensión— cambia qué métricas importan y cuánto confiar en el coeficiente.

El peligro de extrapolar

La recta ajustada en un rango de datos no vale fuera de él. La relación puede curvarse, saturar o invertirse donde no hay observaciones. Extrapolar es la forma más elegante de afirmar algo sin evidencia: la inferencia solo respalda predicciones dentro del dominio visto.

Interpolación con cautela

Incluso dentro del rango, predecir la media de Y para un X dado tiene un intervalo de confianza (para la media) más estrecho que el intervalo de predicción (para un caso individual), que añade el ruido propio de cada observación. Confundir ambos subestima la incertidumbre.

Supuestos del modelo lineal

- Linealidad en los parámetros y en la relación media.
- Residuos con varianza constante (homocedasticidad).
- Residuos independientes y aproximadamente normales para inferir.
- Sin colinealidad perfecta entre predictores.

Diagnóstico antes de inferir

Antes de confiar en el p de un coeficiente, hay que mirar los residuos: un patrón (curvatura, embudo) señala que el modelo lineal no es adecuado y que la inferencia asociada es frágil. Los supuestos no son un trámite, son la garantía del valor p .

Regresión y correlación se tocan

En una recta simple, probar si la pendiente difiere de cero equivale a probar si la correlación difiere de cero: son la misma pregunta de ángulos distintos. El coeficiente de correlación es una pendiente estandarizada.

Multicolinealidad

Cuando los predictores se parecen entre sí, sus coeficientes individuales se vuelven inestables y sus errores se inflan: el modelo acierta en conjunto pero no sabe atribuir el mérito a cada variable. La inferencia sobre coeficientes sufre, aunque la predicción no.

Inferir es generalizar

Toda inferencia en regresión descansa en que la muestra represente a la población (muestreo aleatorio) y en que el proceso sea estable. Romper esas condiciones invalida los intervalos: reflejarían la variabilidad, no el sesgo ni los cambios estructurales.

Predicción sin modelo

Para solo predecir, a veces conviene un método flexible que capture no linealidades, renunciando a la interpretación de coeficientes. La inferencia clásica privilegia la transparencia del coeficiente; el pronóstico puro, la precisión. Son objetivos legítimos pero distintos.

Errores frecuentes

- Extrapolar la recta fuera del rango de datos observado.
- Atribuir causalidad a un coeficiente significativo sin control de confusores.

- Ignorar residuos con patrón y reportar el valor p del coeficiente como si fuera válido.

1. Un coeficiente de correlación alto en la muestra garantiza el mismo valor al predecir sobre toda la población.

☐ Verdadero

☐ Falso

Del coeficiente a la decisión

Traducir una pendiente inferida a una decisión exige su intervalo, su contexto causal y su dominio de validez. La regresión entrega la relación estimada; el criterio humano decide cuánto peso darle fuera del gráfico.

ACTIVIDAD INTERACTIVA DISPONIBLE EN
LUMBRE

2. Inferir de una regresión



3. Ordena el uso responsable de una regresión inferencial:

Elementos:

- ajusta la recta por mínimos cuadrados
- revisa que los residuos no tengan patrón
- interpreta coeficientes con su error e intervalo
- predice solo dentro del rango válido

Ajustar	. . .
Diagnosticar	. . .
Inferir	. . .
Predecir	. . .

R-bloggers [https://www.r-bloggers.com/Ejemplos reproducibles de regresión inferencial y diagnósticos](https://www.r-bloggers.com/Ejemplos-reproducibles-de-regresión-inferencial-y-diagnósticos).

VIDEO

Video complementario

Recurso audiovisual para reforzar los conceptos.



Escanea para verlo en Lumbre

<https://www.youtube.com/watch?v=1v1ySMzTQzI>

4. Un modelo ajusta una recta entre X de 10 y 20 con fuerte R cuadrada. ¿Por qué es arriesgado predecir para X igual 100? ¿Qué mostrarían los residuos que debería advertirte?

NIST Engineering Statistics Handbook <https://www.itl.nist.gov/div898/handbook/>
Referencia completa de métodos estadísticos.

Referencia curricular

Grado:	3
Materia:	Estadística Inferencial · Probabilidad y Estadística

Versión interactiva completa (con actividades, videos y autoevaluación):



<https://univ.lumbre.lat/object/cdia-esti-007-regresion-y-correlacion-inferida>

Material educativo generado desde Lumbre · lumbre.lat