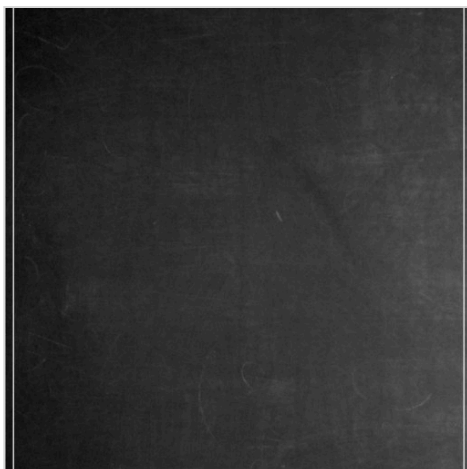


L U M B R E



Supuestos y diagnóstico

Grado: 4

Materia: Regresión y Modelado Estadístico · Probabilidad y Estadística

11/10/2026

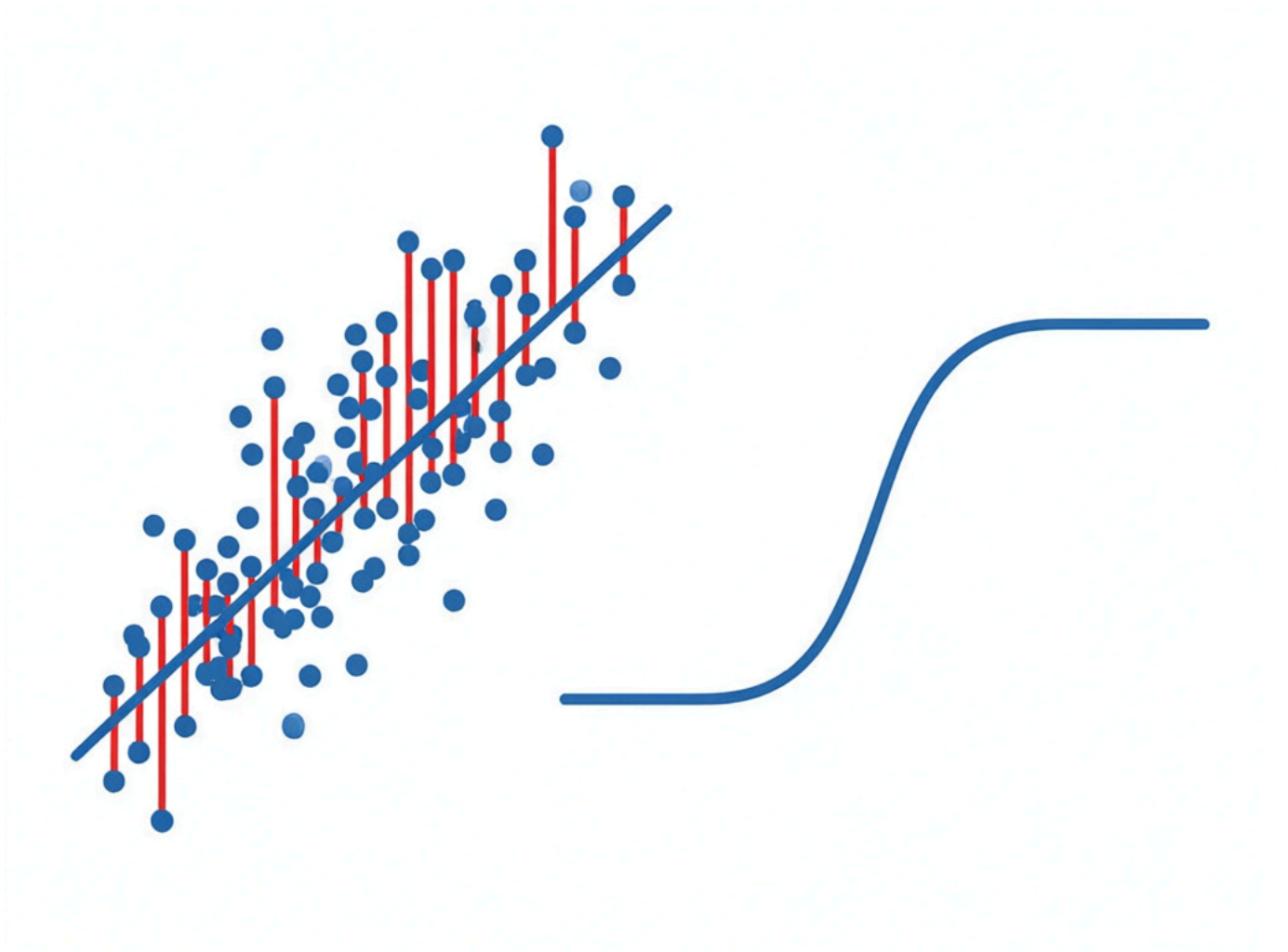
Antes de empezar

Una recta ajustada no vale nada si sus supuestos están rotos. El diagnóstico por residuos es la revisión médica del modelo: comprueba que la linealidad, la varianza constante y la independencia se sostienen antes de confiar en sus conclusiones.

QUÉ APRENDERÁS

- Verificar los supuestos del modelo lineal mediante gráficos y estadísticos de residuos, y detectar violaciones que invalidan la inferencia o la predicción.

Supuestos y diagnóstico



Regresión y modelado estadístico: lineal, múltiple y logístico.

Mapa conceptual

- **Regresión**
 - Lineal simple

- Beta 0 y 1
- R-cuadrado
- Residuos
- Supuestos
 - Normalidad
 - Homocedasticidad
 - Independencia
- Múltiple
 - Multicolinealidad
 - Ajuste R²
 - AIC/BIC
- Logística
 - Odds ratio
 - Sigmoide
 - Regularización

El mapa mental interactivo es interactivo en la versión online.



Antes de inferir, diagnosticar

El valor p de un coeficiente solo significa algo si el modelo está bien planteado. Sin diagnóstico, se corre el riesgo de hacer inferencia impecable sobre una estructura rota: matemáticas correctas, conclusiones falsas.

Los supuestos del modelo lineal

- Linealidad: la media de Y cambia de forma lineal con los predictores.
- Homocedasticidad: la varianza del error es constante en todo el rango.
- Independencia: los errores no están correlacionados entre sí.
- Normalidad de los errores: necesaria sobre todo para intervalos en muestras pequeñas.

El residuo como lupa

El residuo es la diferencia entre el valor observado y el predicho: lo que el modelo no pudo explicar. Es la materia prima del diagnóstico. Graficar residuos contra valores

predichos revela patrones que los resúmenes globales ocultan.

Linealidad: residuos con forma

Si al dibujar residuos frente a predichos aparece una curva (una U, una ola), el modelo lineal está dejando estructura sistemática sin capturar. La remedida es transformar variables o añadir términos no lineales hasta que el residuo quede sin forma.

Homocedasticidad: el embudo

Si la nube de residuos se abre como un embudo —poca dispersión para predichos pequeños, mucha para grandes— la varianza no es constante. La inferencia se resiente; la solución pasa por transformar la respuesta o usar errores estándar robustos.

Independencia y autocorrelación

En datos ordenados (series temporales, mediciones repetidas), los errores suelen correlacionarse: un residuo positivo tiende a seguir otro positivo. La autocorrelación infla la confianza del modelo. Un gráfico de residuos en el orden de recogida o el estadístico de Durbin-Watson lo delatan.

Normalidad y su alcance

Los supuestos de normalidad importan para intervalos de confianza y valores p en muestras pequeñas; con muestras grandes, el teorema central del límite protege las medias. Un Q-Q plot de residuos muestra desviaciones de la campana.

El Q-Q plot

El gráfico cuantil-cuantil enfrenta los cuantiles de los residuos con los de una normal teórica. Si los puntos siguen la diagonal, la normalidad se sostiene; colas que se despegan delatan asimetría o valores extremos.

Puntos influyentes

No todos los atípicos importan igual: uno con elevado apalancamiento, situado en un extremo del rango de X, puede arrastrar la recta casi él solo. La distancia de Cook combina lejanía y residual para medir cuánto cambiaría el ajuste sin ese punto.

Apalancamiento frente a influencia

El apalancamiento mide cuán lejos está un punto en el espacio de X, la palanca potencial que tiene. La influencia es cuánto cambia realmente el modelo. Un punto con

palanca pero que cae en la recta no influye; hay que distinguir potencial de daño real.

Escalado de residuos

Los residuos crudos varían en magnitud según la escala. Estandarizarlos o tipificarlos los hace comparables entre observaciones y permite juzgar cuáles son inusualmente grandes respecto a su propia varianza esperada.

Transformar la respuesta

Cuando el diagnóstico muestra heterocedasticidad o curvatura, una transformación de la respuesta (logaritmo, raíz, una Box-Cox) puede enderezar la relación y estabilizar la varianza de golpe. Es la primera palanca a probar.

Modelos que no son lineales

Si el diagnóstico resiste toda transformación, quizá la relación exija otro modelo: un generalizado para respuestas no normales, o un método no paramétrico. El diagnóstico no solo arregla: a veces anuncia que la familia de modelo es la equivocada.

El orden del diagnóstico

- Ajustar el modelo provisional y dibujar residuos frente a predichos.
- Buscar forma (falta de linealidad) y embudo (heterocedasticidad).
- Revisar puntos de alta influencia y decidir con criterio, no por umbral ciego.
- Corregir con transformación o variables, y volver a diagnosticar.

Errores frecuentes

- Reportar valores p sin haber mirado nunca un gráfico de residuos.
- Eliminar automáticamente todo punto con alta influencia, borrando datos válidos.
- Aplicar correcciones a ciegas sin revalidar que el diagnóstico mejoró.

1. Un gráfico de residuos frente a valores predichos que se abre en abanico señala principalmente:

- ☐ Falta de linealidad
- ☐ Heterocedasticidad: varianza del error no constante
- ☐ Autocorrelación
- ☐ Normalidad excesiva

2. Los supuestos del modelo lineal deben comprobarse antes de confiar en los valores p de los coeficientes.

☐ Verdadero

☐ Falso

No hay atajo visual

Un buen diagnóstico es visual y repetido: unos pocos gráficos bien leídos dicen más que una batería de test formales aplicados mecánicamente. El estadístico confirma; el ojo descubre lo que no se esperaba.

Iterar el ajuste

El diagnóstico no es un trámite final sino un bucle: cada corrección sugiere volver a mirar. Un modelo maduro es resultado de varias pasadas de ajustar, examinar residuos y refinar, no de un único ajuste ingenuo.

ACTIVIDAD INTERACTIVA DISPONIBLE EN LUMBRE

3. Diagnóstico



4. Ordena el bucle de diagnóstico de un modelo lineal:

Elementos:

- ajusta el modelo provisional
- dibuja residuos y busca patrones
- transforma o añade estructura según lo visto
- repite el diagnóstico sobre el modelo corregido

Ajustar	. . .
Graficar	. . .
Corregir	. . .
Revalidar	. . .

Residual diagnostics <https://www.r-bloggers.com/Gráficos de diagnóstico de residuos explicados con ejemplos>.

VIDEO

Video complementario

Recurso audiovisual para reforzar los conceptos.



Escanea para verlo en Lumbre

<https://www.youtube.com/watch?v=1wK1K2M7tXI>

5. Tus residuos frente a predichos dibujan una parábola clara. ¿Qué supuesto se viola y qué dos baterías de corrección considerarías, en qué orden?

NIST: Regression <https://www.itl.nist.gov/div898/handbook/pmd/section1/pmd16.htm>

Referencia completa de metodos de regresion.

Referencia curricular

Grado:	4
Materia:	Regresión y Modelado Estadístico · Probabilidad y Estadística

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



<https://univ.lumbre.lat/object/cdia-reg-002-supuestos-y-diagnostico>

Material educativo generado desde Lumbre · lumbre.lat