Contenidos

- 5.1: Diagnóstico: Análisis de los residuos
- 5.2: La descomposición ANOVA (ANalysis Of VAriance)
- 5.3: Relaciones no lineales y transformaciones para linealización
- 5.4: El modelo de regresión lineal en forma matricial
- 5.5: Introducción a la regresión lineal múltiple

Bibliografía

- Newbold, P. Estadística para los Negocios y la Economía (1997).
  - Capítulos 12, 13 y 14
- Peña, D. Regresión y Diseño de Experimentos (2002).
  - Capítulos 5 y 6

5.1. Diagnóstico en regresión

- Supuestos teóricos del modelo de regresión lineal simple de una var. respuesta y sobre una var. explicativa x:
  - Linealidad: $y_i = \beta_0 + \beta_1 x_i + u_i$, para $i = 1, \ldots, n$
  - Homogeneidad: $\mathrm{E}[u_i] = 0$, para $i = 1, \ldots, n$
  - Homocedasticidad: $\mathrm{Var}[u_i] = \sigma^2$, para $i = 1, \ldots, n$
  - Independencia: u_i y u_j son independientes para $i \neq j$
  - Normalidad: $u_i \sim \mathrm{Normal}(0, \sigma^2)$, para $i = 1, \ldots, n$
- Los métodos de diagnóstico se utilizan para contrastar si tales supuestos son adecuados para los datos disponibles (x_i, y_i); se basan en el análisis de los residuos $e_i = y_i - \widehat{y}_i$

5.1. Diagnóstico: diagrama de puntos

- El método más sencillo consiste en la observación visual del diagrama de puntos (x_i, y_i)
- A menudo, este sencillo pero potente método revela pautas que sugieren si el modelo teórico es o no adecuado
- Ilustraremos su uso con un ejemplo clásico. Consideremos los cuatro conjuntos de datos siguientes

5.1. Diagnóstico: diagrama de puntos

TABLE 3-10
Four Data Sets

DATA SET 1 X	DATA SET 1 Y	DATA SET 2 X	DATA SET 2 Y
10.0	8.04	10.0	9.14
8.0	6.95	8.0	8.14
13.0	7.58	13.0	8.74
9.0	8.81	9.0	8.77
11.0	8.33	11.0	9.26
14.0	9.96	14.0	8.10
6.0	7.24	6.0	6.13
4.0	4.26	4.0	3.10
12.0	10.84	12.0	9.13
7.0	4.82	7.0	7.26
5.0	5.68	5.0	4.74

DATA SET 3 X	DATA SET 3 Y	DATA SET 4 X	DATA SET 4 Y
10.0	7.46	8.0	6.58
8.0	6.77	8.0	5.76
13.0	12.74	8.0	7.71
9.0	7.11	8.0	8.84
11.0	7.81	8.0	8.47
14.0	8.84	8.0	7.04
6.0	6.08	8.0	5.25
4.0	5.39	19.0	12.50
12.0	8.15	8.0	5.56
7.0	6.42	8.0	7.91
5.0	5.73	8.0	6.89

SOURCE: F. J. Anscombe, *op. cit.*

5.1. Diagnóstico: diagrama de puntos

- Para cada uno de los cuatro conjuntos de datos anteriores, se obtiene el mismo modelo estimado de regresión lineal:
- $\hat{y}_i = 3.0 + 0.5x_i$
- $n = 11$, $\bar{x} = 9.0$, $\bar{y} = 7.5$, $r_{x,y} = 0.817$
- El error estándar estimado del estimador $\hat{\beta}_1$,

$$\sqrt{\frac{s_R^2}{(n-1)s_x^2}},$$

  toma el valor 0.118. El estadístico T correspondiente toma el valor $T = 0.5/0.118 = 4.237$
- Sin embargo, los diagramas de puntos correspondientes revelan que los cuatro conjuntos de datos son cualitativamente muy diferentes: ¿Qué conclusiones podemos extraer de estos diagramas?

5.1. Diagnóstico: diagrama de puntos

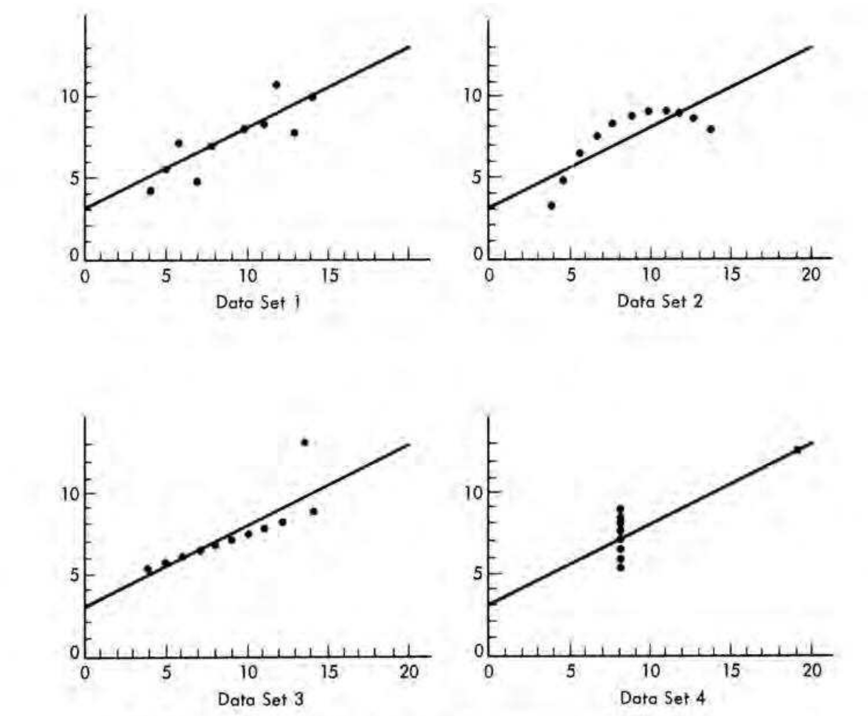


FIGURE 3-29 Scatterplots for the four data sets of Table 3-10
SOURCE: F. J. Anscombe, *op cit.*

5.1: análisis de los residuos

- Si la observación del diagrama de puntos no basta para descartar el modelo, se utilizan métodos de diagnóstico basados en el análisis de los residuos $e_i = y_i - \widehat{y}_i$
- El análisis comienza tipificando los residuos (dividiéndolos por la cuasi-desviación típica residual): Las cantidades resultantes se denominan residuos tipificados:

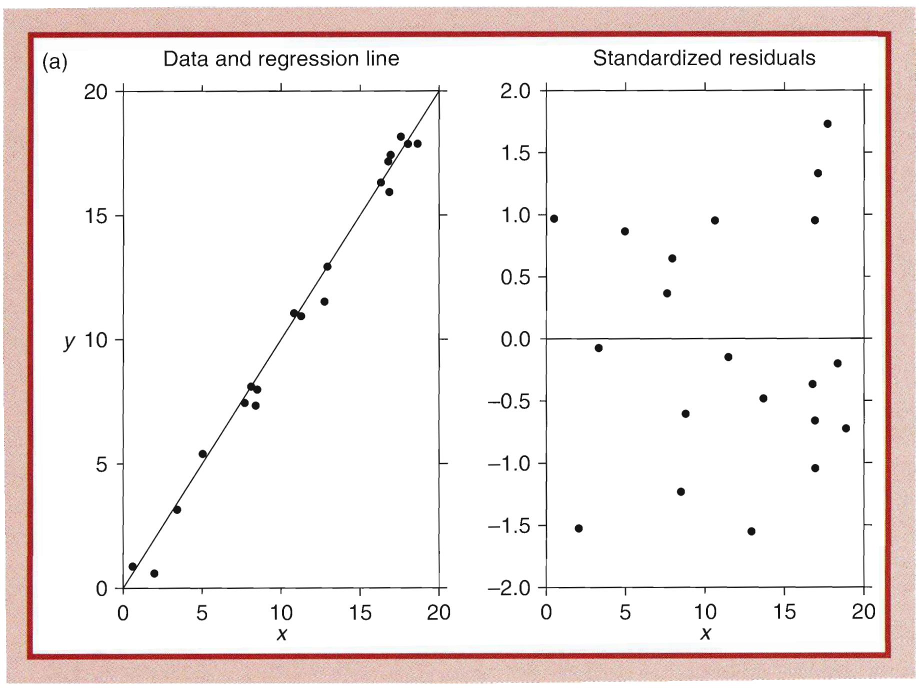
$$\frac{e_i}{s_R}$$

- Bajo los supuestos del modelo de regresión lineal, los residuos tipificados son aproximadamente variables aleatorias normales estándar independientes
- Un gráfico de los residuos tipificados no debería mostrar ninguna pauta clara

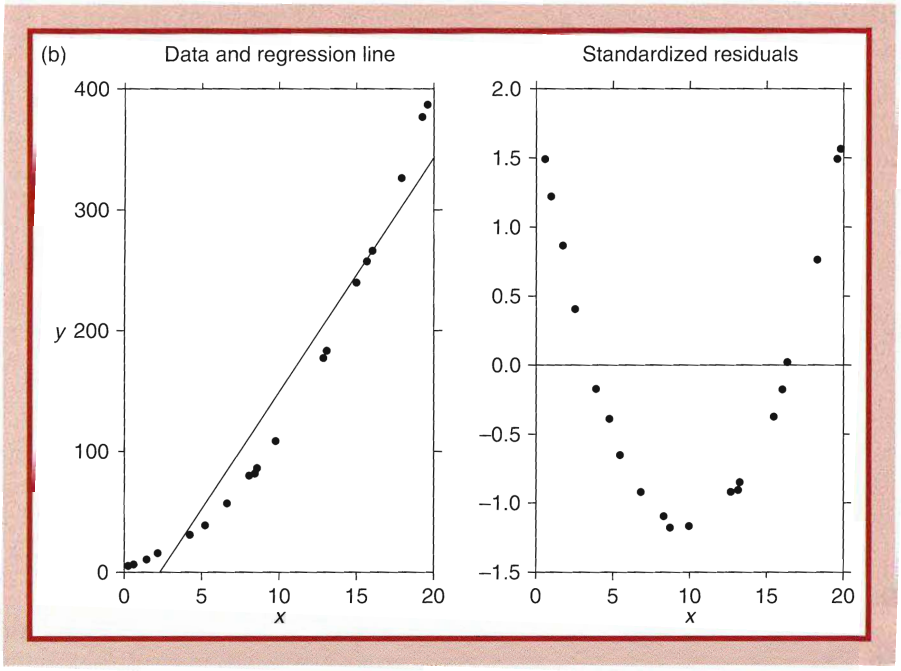
5.1: Diagramas de residuos

- Hay varios tipos de diagramas de residuos. Los más comunes son:
  - Diagrama de los residuos vs. x
  - Diagrama de los residuos vs. $\widehat{y}$
- Las desviaciones de los supuestos del modelo dan lugar a pautas, que se pueden identificar visualmente

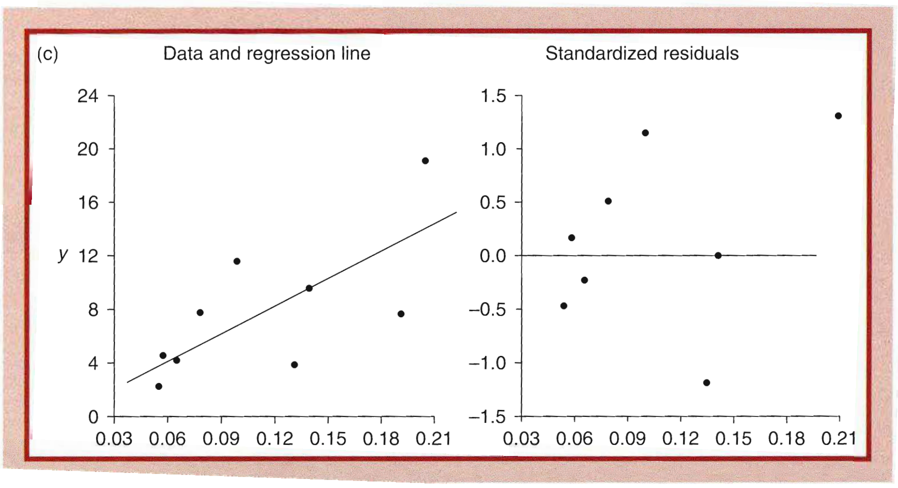
5.1: Ej: consistencia con el modelo teórico



5.1: Ej: No linealidad



5.1: Ej: Heterocedasticidad



5.1: Datos atípicos

- A partir del gráfico de la recta de regresión podemos observar datos atípicos, que presentan desviaciones sustanciales de la recta de regresión
- Los estimadores $\widehat{\beta}_0$ y $\widehat{\beta}_1$ de los parámetros de la recta de regresión son muy sensibles a tales datos atípicos
- Por ello, es importante identificar tales datos y comprobar si son válidos

5.1: Normalidad de los errores

- Recordemos que uno de los supuestos teóricos del modelo de regresión lineal es que los errores tienen una distribución normal
- Podemos comprobar este supuesto visualmente a partir de la observación y análisis de los residuos e_i, empleando varios métodos:
  - Observación del histograma de frecuencias de los residuos
  - Observación de un "*Normal Probability Plot*" para los residuos (desviaciones importantes de los datos de la línea recta en este gráfico indican desviaciones sustanciales del supuesto de normalidad)

5.2: La descomposición ANOVA

- ANOVA: *ANalysis Of VAriance*
- Al ajustar un modelo de regresión lineal $\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i$ a un conjunto de datos (x_i, y_i), para $i = 1, \ldots, n$, podemos distinguir tres fuentes de variación en las respuestas:
  - variación debida al modelo: $\mathrm{SCM} = \sum_{i=1}^{n} (\widehat{y}_i - \bar{y})^2$, donde las siglas "SC" se refieren a "suma de cuadrados", y la "M" se refiere al "Modelo"
  - variación residual: $\mathrm{SCR} = \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 = \sum_{i=1}^{n} e_i^2$
  - variación total: $\mathrm{SCT} = \sum_{i=1}^{n} (y_i - \bar{y})^2$
- La descomposición ANOVA indica que $\mathrm{SCT} = \mathrm{SCM} + \mathrm{SCR}$

5.2: El coeficiente de determinación R^2

- La descomposición ANOVA indica que $\mathrm{SCT} = \mathrm{SCM} + \mathrm{SCR}$
- Notemos que: $y_i - \bar{y} = (y_i - \widehat{y}_i) + (\widehat{y}_i - \bar{y})$
- $\mathrm{SCM} = \sum_{i=1}^{n} (\widehat{y}_i - \bar{y})^2$ mide la variación de las respuestas debida a la regresión (explicada por los valores predichos $\widehat{y}$)
- Por lo tanto, el cociente $\mathrm{SCR}/\mathrm{SCT}$ es la proporción de variación de la respuesta no explicada por la regresión
- El cociente $R^2 = \mathrm{SCM}/\mathrm{SCT} = 1 - \mathrm{SCR}/\mathrm{SCT}$ es la proporción de variación de las respuestas explicada por la regresión; se conoce como coeficiente de determinación
- Resultado: $R^2 = r_{xy}^2$ (coef. de correlación al cuadrado)
- Ej: si $R^2 = 0.85$, la variable x explica un 85% de la variación de la variable y

5.2: Tabla ANOVA

Fuente de variación	SC	G.L.	Media	Cociente F
Modelo	SCM	1	$\mathrm{SCM}/1$	SCM/s_R^2
Residuos/Errores	SCR	$n-2$	$\mathrm{SCR}/(n-2) = s_R^2$	
Total	SCT	$n-1$		

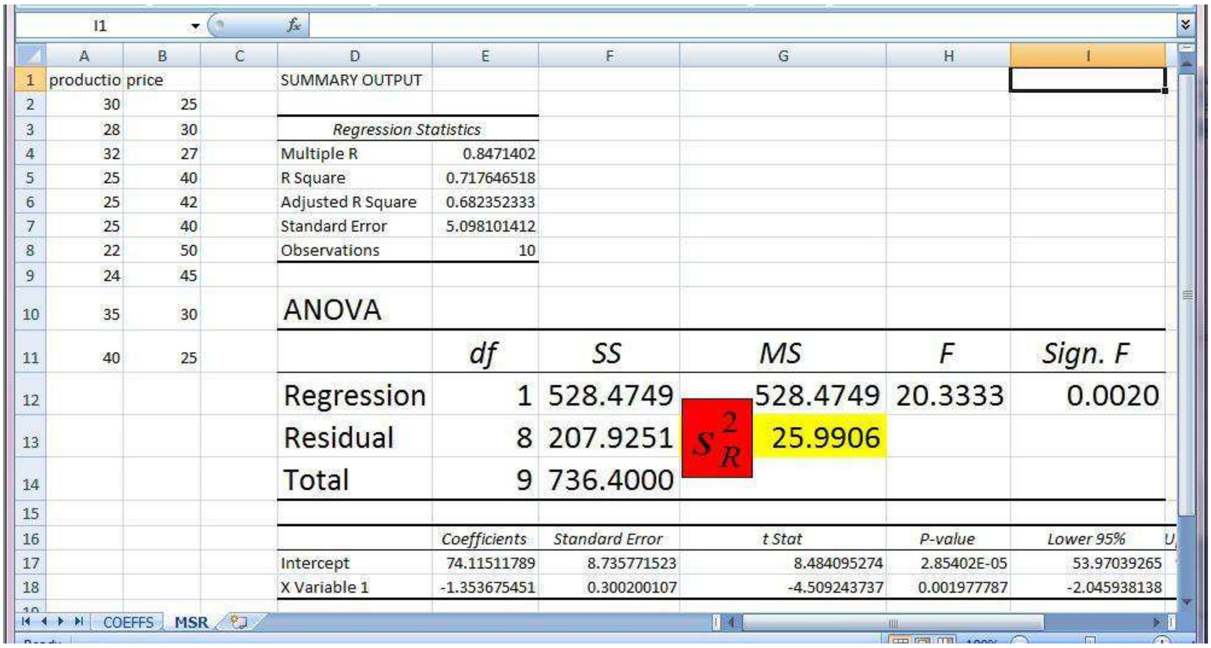
5.2: Contraste de hipótesis ANOVA

- Contraste de hipótesis $\mathrm{H}_0: \beta_1 = 0$ vs. $\mathrm{H}_1: \beta_1 \neq 0$
- Consideremos el cociente

$$F = \frac{\mathrm{SCM}/1}{\mathrm{SCR}/(n-2)} = \frac{\mathrm{SCM}}{s_R^2}$$

- Bajo H_0, F sigue una distribución $F_{1,n-2}$
- Contraste a nivel α: rechazar H_0 si $F > F_{1,n-2;\alpha}$
- ¿Cuál es la relación con el contraste basado en la t de Student que vimos en el Tema 4? Son equivalentes

5.2: Ej. ANOVA



	A	B	C	D	E	F	G	H	I
1	productio	price		SUMMARY OUTPUT					
2	30	25							
3	28	30		*Regression Statistics*					
4	32	27		Multiple R	0.8471402				
5	25	40		R Square	0.717646518				
6	25	42		Adjusted R Square	0.682352333				
7	25	40		Standard Error	5.098101412				
8	22	50		Observations	10				
9	24	45							
10	35	30		ANOVA					
11	40	25			*df*	*SS*	*MS*	*F*	*Sign. F*
12				Regression	1	528.4749	528.4749	20.3333	0.0020
13				Residual	8	207.9251	s_R^2 25.9906		
14				Total	9	736.4000			
15									
16					*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
17				Intercept	74.11511789	8.735771523	8.484095274	2.85402E-05	53.97039265
18				X Variable 1	-1.353675451	0.300200107	-4.509243737	0.001977787	-2.045938138

5.3: Relaciones no lineales y linealización

- Supongamos que la parte determinista $f(x_i; a, b)$ de la respuesta en el modelo
$$y_i = f(x_i; a, b) + u_i, \quad i = 1, \ldots, n$$
es una función no lineal de x que depende de dos parámetros a y b (ej: $f(x; a, b) = ab^x$)
- En algunos casos podemos aplicar transformaciones a los datos para linearizarlos, y así poder aplicar los métodos de regresión lineal
- A partir de los datos (x_i, y_i) originales, obtenemos los datos transformados (x_i', y_i')
- Los parámetros β_0 y β_1 de la relación lineal entre las x_i' y las y_i' se obtienen como transformaciones de los parámetros a y b

5.3: Transformaciones para linealización

- Ejemplos de transformaciones para linealización:
  - Si $y = f(x; a, b) = ax^b$ entonces $\log y = \log a + b \log x$: tomamos $y' = \log y$, $x' = \log x$, $\beta_0 = \log a$, $\beta_1 = b$
  - Si $y = f(x; a, b) = ab^x$ entonces $\log y = \log a + (\log b)x$: tomamos $y' = \log y$, $x' = x$, $\beta_0 = \log a$, $\beta_1 = \log b$
  - Si $y = f(x; a, b) = 1/(a + bx)$ entonces $1/y = a + bx$: tomamos $y' = 1/y$, $x' = x$, $\beta_0 = a$, $\beta_1 = b$
  - Si $y = f(x; a, b) = \ln(ax^b)$ entonces $y = \ln a + b \ln x$: tomamos $y' = y$, $x' = \ln x$, $\beta_0 = \ln a$, $\beta_1 = b$

5.4: Regresión lineal en forma matricial

- Recordemos el modelo de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_i + u_i, \quad i = 1, \ldots, n$$

- Escribiendo una ecuación para cada observación obtenemos

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_1 + u_1 \\ y_2 &= \beta_0 + \beta_1 x_2 + u_2 \\ &\vdots \quad \vdots \quad \vdots \\ y_n &= \beta_0 + \beta_1 x_n + u_n \end{aligned}$$

5.4: Regresión lineal en forma matricial

- En forma matricial, podemos escribir

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix},$$

o, separando los parámetros β_j de las x_i,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix},$$

5.4: Regresión lineal en forma matricial

- Escribimos la relación matricial

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix},$$

como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

- $\mathbf{y}$: vector de respuestas; $\mathbf{X}$: matriz de variables explicativas (o del diseño experimental); $\boldsymbol{\beta}$: vector de parámetros; $\mathbf{u}$: vector de errores

5.4: La matriz de covarianzas de los errores

- Denotamos por $\mathrm{Cov}(\mathbf{u})$ la matriz $n \times n$ de covarianzas de los errores; su elemento (i, j) es

$$\mathrm{cov}(u_i, u_j) = \begin{cases} 0 & \text{si } i \neq j \\ \mathrm{Var}[u_i] = \sigma^2 & \text{si } i = j \end{cases}$$

- $\mathrm{Cov}(\mathbf{u})$ es la matriz identidad $\mathbf{I}_{n\times n}$ multiplicada por σ^2:

$$\mathrm{Cov}(\mathbf{u}) = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 \mathbf{I}$$

5.4: Estimación de mínimos cuadrados

- El vector estimado $\widehat{\beta}$ de mínimos cuadrados es la solución única de la ecuación matricial 2×2 (comprueba las dimensiones)

$$\left(\mathbf{X}^T\mathbf{X}\right)\widehat{\beta} = \mathbf{X}^T\mathbf{y},$$

es decir,

$$\widehat{\beta} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

- El vector $\widehat{\mathbf{y}} = (\widehat{y}_i)$ de respuestas estimadas es

$$\widehat{\mathbf{y}} = \mathbf{X}\widehat{\beta}$$

y el vector de residuos es $\mathbf{e} = \mathbf{y} - \widehat{\mathbf{y}}$

5.5: El modelo de regresión lineal múltiple

- Modelo de regresión lineal simple: predecir una respuesta y a partir de una variable explicativa x
- En numerosas aplicaciones, buscamos predecir la respuesta y a partir de múltiples variables explicativas $x_1, \ldots, x_k$
- Ej: predecir el precio de una casa en función de su superficie, localización, planta, y número de baños
- Ej: predecir el tamaño de un parlamento en función de la población, su tasa de crecimiento, el número de partidos políticos con representación, etc.

5.5: El modelo de regresión lineal múltiple

- Modelo de regresión lineal múltiple: predecir una respuesta y a partir de múltiples variables explicativas $x_1, \ldots, x_k$
- Tenemos n observaciones: para $i = 1, \ldots, n$,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i$$

- Suponemos que las u_i son v.a. independientes con distribución $\mathrm{Normal}(0, \sigma^2)$

5.5: Ajuste de mínimos cuadrados

- Tenemos n observaciones: para $i = 1, \ldots, n$,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i$$

- Buscamos ajustar a los datos $(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i)$ un hiperplano de la forma

$$\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{i1} + \widehat{\beta}_2 x_{i2} + \cdots + \widehat{\beta}_k x_{ik}$$

- El residuo para la observación i es: $e_i = y_i - \widehat{y}_i$
- Utilizamos la estimación de los parámetros $\widehat{\beta}_j$ que minimiza la suma de los cuadrados de los residuos

5.5: Modelo en forma matricial

- Escribimos la relación matricial

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix},$$

como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

- $\mathbf{y}$: vector de respuestas; $\mathbf{X}$: matriz de variables explicativas (o del diseño experimental); $\boldsymbol{\beta}$: vector de parámetros; $\mathbf{u}$: vector de errores

5.5: Estimación de mínimos cuadrados de β

- El vector estimado $\widehat{\beta}$ de mínimos cuadrados es la solución única de la ecuación matricial $(k+1) \times (k+1)$ (comprueba las dimensiones)

$$\left(\mathbf{X}^T\mathbf{X}\right)\widehat{\beta} = \mathbf{X}^T\mathbf{y},$$

como en el caso $k = 1$ visto anteriormente, es decir,

$$\widehat{\beta} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

- El vector $\widehat{\mathbf{y}} = (\widehat{y}_i)$ de respuestas estimadas es

$$\widehat{\mathbf{y}} = \mathbf{X}\widehat{\beta}$$

y el vector de residuos es $\mathbf{e} = \mathbf{y} - \widehat{\mathbf{y}}$

5.5: Estimación de la varianza σ^2

- Para el modelo de regresión lineal múltiple, estimamos la varianza σ^2 mediante la cuasi-varianza residual,

$$s_R^2 = \frac{\sum_{i=1}^n e_i^2}{n-k-1},$$

que es un estimador insesgado (nótese que para regresión lineal simple el denominador vale $n-2$)

5.5: Distribución muestral de $\widehat{\boldsymbol{\beta}}$

- Bajo los supuestos del modelo, el estimador de mínimos cuadrados $\widehat{\boldsymbol{\beta}}$ del vector de parámetros $\boldsymbol{\beta}$ sigue una distribución normal multivariante
- $\mathrm{E}\left[\widehat{\boldsymbol{\beta}}\right] = \boldsymbol{\beta}$ (i.e., es un estimador insesgado)
- La matriz de covarianzas de $\widehat{\boldsymbol{\beta}}$ es $\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) = \sigma^2 \left(\mathbf{X}^T\mathbf{X}\right)^{-1}$
- Estimamos $\mathrm{Cov}(\widehat{\boldsymbol{\beta}})$ por $s_R^2 \left(\mathbf{X}^T\mathbf{X}\right)^{-1}$
- La estimación de $\mathrm{Cov}(\widehat{\boldsymbol{\beta}})$ nos da estimaciones $s^2(\widehat{\beta}_j)$ de la varianza $\mathrm{Var}(\widehat{\beta}_j)$; $s(\widehat{\beta}_j)$ es el error estándar del estimador $\widehat{\beta}_j$
- Al tipificar $\widehat{\beta}_j$ obtenemos: $\dfrac{\widehat{\beta}_j - \beta_j}{s(\widehat{\beta}_j)} \sim t_{n-k-1}$ (t de Student)

5.5: Inferencia sobre los parámetros $\widehat{\beta}_j$

- Intervalo de confianza a nivel $1-\alpha$ para β_j:

$$\widehat{\beta}_j \pm t_{n-k-1;\alpha/2}\, s(\widehat{\beta}_j)$$

- Contraste de hipótesis a nivel α para $\mathrm{H}_0: \beta_j = 0$ vs. $\mathrm{H}_1: \beta_j \neq 0$
- Rechazar H_0 si $|T| > t_{n-k-1;\alpha/2}$, donde $T = \widehat{\beta}_j / s(\widehat{\beta}_j)$ es el estadístico de contraste

5.5: La descomposición ANOVA

- ANOVA: *ANalysis Of VAriance*
- Al ajustar un modelo de regresión lineal múltiple

$$\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{i1} + \cdots + \widehat{\beta}_k x_{ik}$$

  a un conjunto de datos $(x_{i1}, \ldots, x_{ik}, y_i)$, para $i = 1, \ldots, n$, podemos distinguir tres fuentes de variación en las respuestas:
  - variación debida a la regresión: $\mathrm{SCM} = \sum_{i=1}^{n} (\widehat{y}_i - \bar{y})^2$, donde las siglas "SC" se refieren a "suma de cuadrados", y la "M" se refiere al "Modelo"
  - variación residual: $\mathrm{SCR} = \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 = \sum_{i=1}^{n} e_i^2$
  - variación total: $\mathrm{SCT} = \sum_{i=1}^{n} (y_i - \bar{y})^2$
- La descomposición ANOVA indica que $\mathrm{SCT} = \mathrm{SCM} + \mathrm{SCR}$

5.5: El coeficiente de determinación R^2

- La descomposición ANOVA indica que $\mathrm{SCT} = \mathrm{SCM} + \mathrm{SCR}$
- Notemos que: $y_i - \bar{y} = (y_i - \widehat{y}_i) + (\widehat{y}_i - \bar{y})$
- $\mathrm{SCM} = \sum_{i=1}^{n} (\widehat{y}_i - \bar{y})^2$ mide la variación de las respuestas debida a la regresión (explicada por los valores predichos $\widehat{y}_i$)
- Por lo tanto, el cociente $\mathrm{SCR}/\mathrm{SCT}$ es la proporción de variación de la respuesta no explicada por la regresión
- El cociente $R^2 = \mathrm{SCM}/\mathrm{SCT} = 1 - \mathrm{SCR}/\mathrm{SCT}$ es la proporción de variación de las respuestas explicada por las variables explicativas; se conoce como coeficiente de determinación múltiple
- Resultado: $R^2 = r^2_{\widehat{y}y}$ (coef. de correlación al cuadrado)
- Ej: si $R^2 = 0.85$, las variables $x_1, \ldots, x_k$ explican un 85% de la variación de la variable y

5.5: Tabla ANOVA

Fuente de variación	SC	G.L.	Media	Cociente F
Modelo	SCM	k	SCM/k	$(\mathrm{SCM}/k)/s_R^2$
Residuos/Errores	SCR	$n-k-1$	$\frac{\mathrm{SCR}}{n-k-1} = s_R^2$	
Total	SCT	$n-1$		

5.5: Contraste de hipótesis ANOVA

- Consideremos el contraste de hipótesis $\mathrm{H}_0 : \beta_1 = \beta_2 = \cdots \beta_k = 0$ vs. $\mathrm{H}_1 : \beta_j \neq 0$ para algún $j = 1, \ldots, k$
- H_0 : la respuesta no depende de las x_j
- Consideremos el cociente

$$F = \frac{\mathrm{SCM}/k}{\mathrm{SCR}/(n-k-1)} = \frac{\mathrm{SCM}/k}{s_R^2}$$

- Bajo H_0, F sigue una distribución $F_{k,n-k-1}$
- Contraste a nivel α: rechazar H_0 si $F > F_{k,n-k-1;\alpha}$