Probabilidad y Estadística para Inteligencia Artificial
Examen 2

Fausto Juárez Yélamos

8 de Junio del 2024

Resolución del Problema 1

Se pretende estimar los valores de producción Y (en miles de toneladas) de cierto material, en función del tiempo transcurrido X (en meses) usando los valores de la tabla:

X	Y
2	2
5	53
13	135
16	199
22	509

Se plantea un modelo de la forma $Y = a + bX + cX^2$. Para encontrar los estimadores de mínimos cuadrados para a, b y c, utilizamos la siguiente notación matricial:

$$\mathbf{Y} = \mathbf{X}\beta$$

donde:

$$\mathbf{Y} = \begin{pmatrix} 2 \\ 53 \\ 135 \\ 199 \\ 509 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 5 & 25 \\ 1 & 13 & 169 \\ 1 & 16 & 256 \\ 1 & 22 & 484 \end{pmatrix}, \quad \beta = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

El estimador de mínimos cuadrados es:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$$

Calculamos primero $\mathbf{X}^T\mathbf{X}$ y $\mathbf{X}^T\mathbf{Y}$:

$$\mathbf{X}^T\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 5 & 13 & 16 & 22 \\ 4 & 25 & 169 & 256 & 484 \end{pmatrix} * \begin{pmatrix} 1 & 2 & 4 \\ 1 & 5 & 25 \\ 1 & 13 & 169 \\ 1 & 16 & 256 \\ 1 & 22 & 484 \end{pmatrix} = \begin{pmatrix} 5 & 58 & 938 \\ 58 & 938 & 15074 \\ 938 & 15074 & 243318 \end{pmatrix}$$

$$\mathbf{X}^T\mathbf{Y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 5 & 13 & 16 & 22 \\ 4 & 25 & 169 & 256 & 484 \end{pmatrix} * \begin{pmatrix} 2 \\ 53 \\ 135 \\ 199 \\ 509 \end{pmatrix} = \begin{pmatrix} 898 \\ 16406 \\ 322288 \end{pmatrix}$$

Ahora, encontramos $\hat{\beta}$:

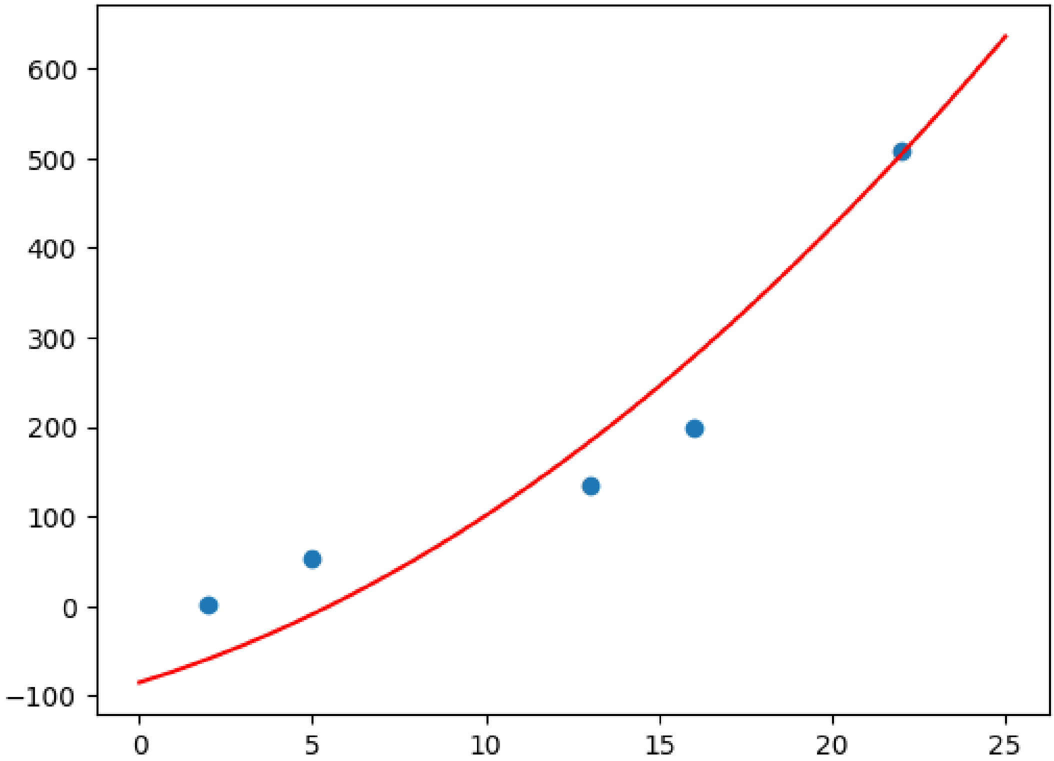
$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \approx \begin{pmatrix} -85.226 \\ 11.804 \\ 0.682 \end{pmatrix}$$

Por lo tanto, los estimadores de mínimos cuadrados son:

$$a \approx -85.226$$
$$b \approx 11.804$$
$$c \approx 0.682$$

El modelo estimado es:

$$\hat{Y} = -85.226 + 11.804X + 0.682X^2$$



Resolución del Problema 2

Una empresa aseguradora desea estimar el porcentaje de morosidad en una determinada cartera de gran dimensión que tiene emitidas 9 pólizas, de las cuales 1 está en mora. El investigador encargado, teniendo en cuenta toda la información disponible y su conocimiento sobre el funcionamiento interno de la empresa, considera que puede modelar el porcentaje p de morosidad, según una distribución Beta$(2, 4)$. Determinar la distribución a posteriori del parámetro p. Determinar su media y su varianza.

Paso 1: Especifico la distribución a priori

La distribución a priori de p es:

$$p \sim \text{Beta}(2, 4)$$

Paso 2: Especifico la verosimilitud

La verosimilitud, dado que se observa una variable binomial con parámetros n (número de pruebas) y p (probabilidad de éxito), es:

$$\text{Verosimilitud} \sim \text{Binomial}(n = 9, x = 1)$$

Paso 3: Calculo la distribución a posteriori

La distribución a posteriori sigue una distribución Beta con parámetros actualizados:

$$p \mid \text{datos} \sim \text{Beta}(\alpha + x, \beta + n - x)$$

donde:

- $\alpha = 2$ (parámetro de la distribución a priori)
- $\beta = 4$ (parámetro de la distribución a priori)
- $x = 1$ (número de éxitos observados)
- $n = 9$ (número de pruebas)

Sustituyendo estos valores en la fórmula:

$$p \mid \text{datos} \sim \text{Beta}(2 + 1, 4 + 9 - 1) = \text{Beta}(3, 12)$$

Paso 4: Determino la media y varianza de la distribución a posteriori

Para una distribución Beta $\text{Beta}(\alpha, \beta)$:

- La media es:
$$\mu = \frac{\alpha}{\alpha + \beta}$$
- La varianza es:
$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

Para nuestra distribución a posteriori $\text{Beta}(3, 12)$:

- Media:
$$\mu = \frac{3}{3 + 12} = \frac{3}{15} = 0.2$$
- Varianza:
$$\sigma^2 = \frac{3 \times 12}{(3 + 12)^2(3 + 12 + 1)} = \frac{36}{15^2 \times 16} = \frac{36}{225 \times 16} = \frac{36}{3600} = 0.01$$