Generalized Linear Models – Part II

Johns Hopkins Engineering

625.461 Statistical Models and Regression

Module 11 – Lecture 11C

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Estimation of Parameters in Logistic Regression

$$y_i = E(y_i) + \varepsilon_i$$

The observed responses y_i are independently Bernoulli random variables with mean

$$E(y_i) = \pi_i = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i'\boldsymbol{\beta})}$$

Maximum Likelihood Estimation for β

$$f_i(y_i) = \pi_i^{y_i}(1-\pi_i)^{1-y_i}, \quad i = 1, 2, \ldots, n$$

$$L(y_1, y_2, \ldots, y_n, \boldsymbol{\beta}) = \prod_{i=1}^{n} f_i(y_i) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i}$$

$$\ln L(\mathbf{y}, \boldsymbol{\beta}) = \sum_{i=1}^{n} y_i \mathbf{x}_i' \boldsymbol{\beta} - \sum_{i=1}^{n} \ln[1 + \exp(\mathbf{x}_i' \boldsymbol{\beta})]$$

Maximum Likelihood Estimation for β

Numerical search methods could be used to compute the MLEs $\hat{\boldsymbol{\beta}}$ (no closed analytic form).

Often we have repeated observations at each level of the x variables. If the model is correct, it can be proved that asymptotically (i.e., sufficiently large n)

$$E\left(\hat{\beta}\right)=\beta \quad \text{and} \quad \mathrm{Var}\left(\hat{\beta}\right)=\left(\mathbf{X}'\mathbf{V}\mathbf{X}\right)^{-1}$$

$$V_{ii}=\boxed{n_i}\hat{\pi}_i\left(1-\hat{\pi}_i\right)$$

Prediction of y in Logistic Regression

$$\hat{y}_i = \hat{\pi}_i = \frac{\exp(\hat{\eta}_i)}{1+\exp(\hat{\eta}_i)} = \frac{\exp\left(\mathbf{x}_i'\hat{\boldsymbol{\beta}}\right)}{1+\exp\left(\mathbf{x}_i'\hat{\boldsymbol{\beta}}\right)} = \frac{1}{1+\exp\left(-\mathbf{x}_i'\hat{\boldsymbol{\beta}}\right)}$$

Pneumoconiosis Data (Ex. 13.1, page 426)

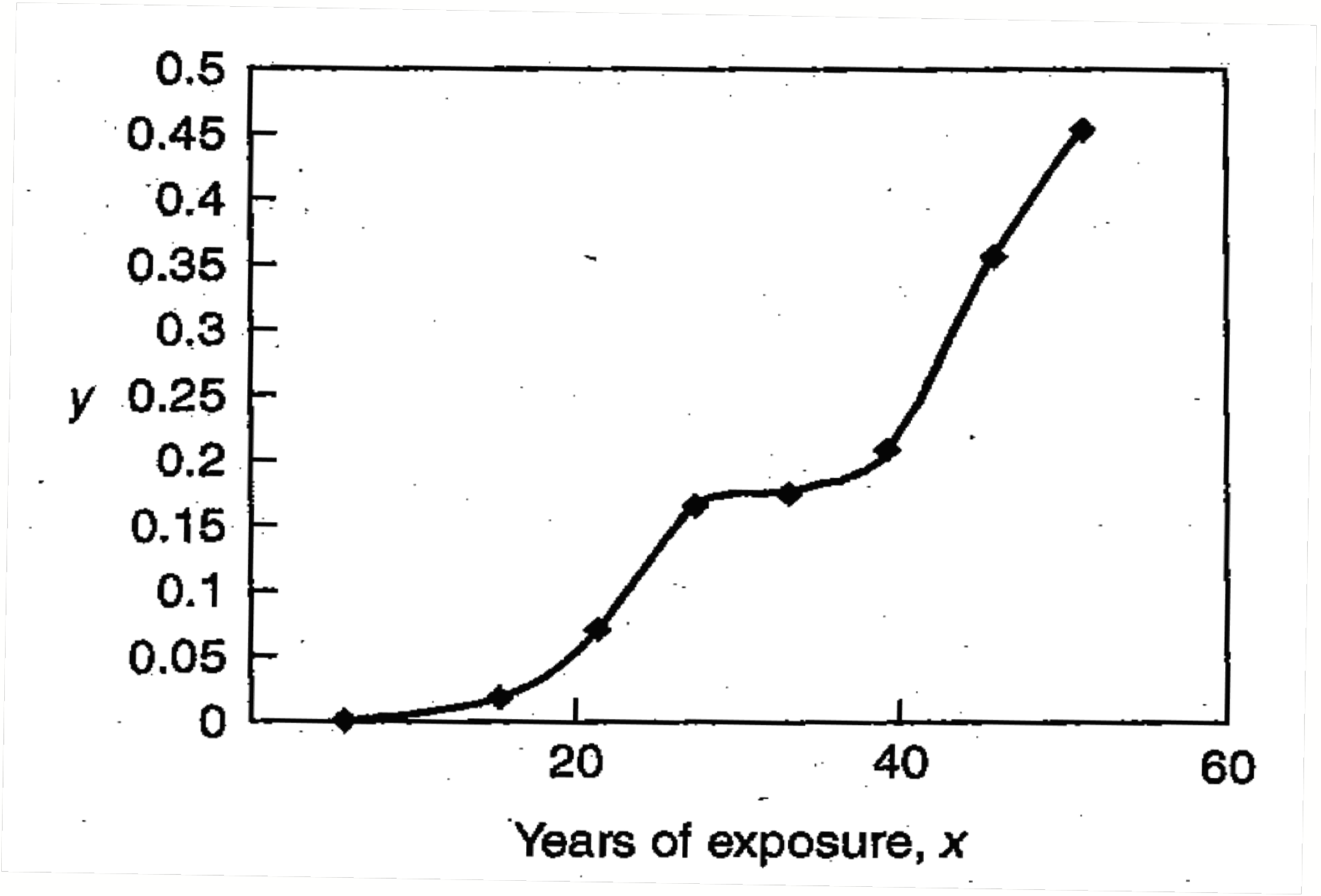
Relate the proportion of coal miners who developed severe pneumoconiosis symptom to number of years of exposure.

Pneumoconiosis Data

TABLE 13.1 The Pneumoconiosis Data

Number of Years of Exposure	Number of Severe Cases	Total Number of Miners	Proportion of Severe Cases, y
5.8	0	98	0
15.0	1	54	0.0185
21.5	3	43	0.0698
27.5	8	48	0.1667
33.5	9	51	0.1765
39.5	8	38	0.2105
46.0	10	28	0.3571
51.5	5	11	0.4545

Pneumoconiosis Data

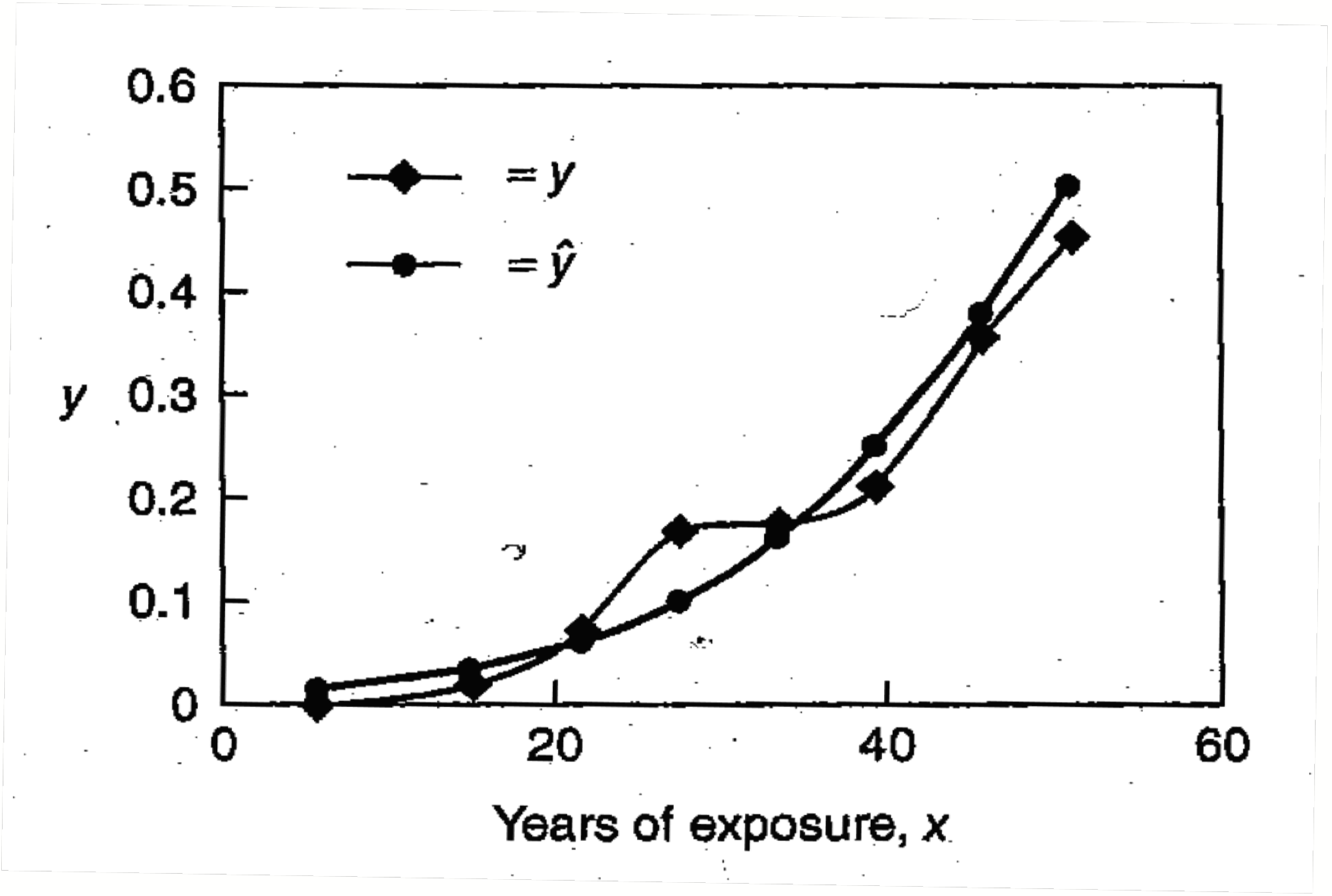


Pneumoconiosis Data

$$\hat{y} = \hat{\pi} = \frac{1}{1 + e^{+4.7965 - 0.0935x}}$$

$$\operatorname{Var}\left(\hat{\beta}\right) = \begin{bmatrix} 0.323283 & -0.0083480 \\ -0.0083480 & 0.0002380 \end{bmatrix}$$

Pneumoconiosis Data



JOHNS HOPKINS

WHITING SCHOOL *of* ENGINEERING