

# STAT 151A Lecture 32

Henry Liev

10 November 2025

## Remark 0.1 (Linear Probability Model)

$y_i = \alpha + \beta x_i + \varepsilon_i$ , where  $y_i$  is binary

NLM:  $y_i = \alpha + \beta x_i + \varepsilon_i, \varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$

LPM: is  $\varepsilon_i \sim N(0, \sigma^2)$ , no

$\varepsilon_i = y_i - \mathbb{E}(y_i)$

$$\varepsilon_i = \begin{cases} 1 - (\alpha + \beta x_i) = 1 - \pi_i \\ -(\alpha + \beta x_i) = \pi_i \end{cases}$$

Not normally distributed and not homoskedastic

Bottom line: Inference is messy (HLM or bootstrap can help a bit)

**Remark 0.2 (Logistic Regression)**

What we need: a (nonlinear) mapping from the linear predictor to  $[0, 1]$

$$P(y_i = 1|x_i) = \pi_i = g(\alpha + \beta x_i)$$

Choose logistic function:

$$g(z) = \frac{1}{1 + \exp(-z)}$$

Why?

(1) It is nearly linear for  $\pi \in (0.2, 0.8)$

(2) Nice interpretation for  $\beta$ :

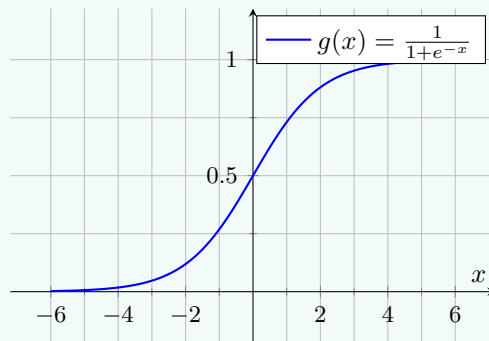
$$\pi_i = \frac{1}{1 + \exp(-\alpha - \beta x_i)} = \frac{\exp(\alpha + \beta x_i)}{\exp(\alpha + \beta x_i) + 1}$$

$$1 - \pi_i = \frac{\exp(\alpha + \beta x_i) + 1}{\exp(\alpha + \beta x_i) + 1} - \frac{\exp(\alpha + \beta x_i)}{\exp(\alpha + \beta x_i) + 1} = \frac{1}{\exp(\alpha + \beta x_i) + 1}$$

$$\frac{\pi_i}{1 - \pi_i} = \exp(\alpha + \beta x_i)$$

$$\log\left(\frac{\pi_i}{1 - \pi_i}\right) = \alpha + \beta x_i$$

A 1-unit difference in  $x_j$  is associated with a  $\beta_j$  difference in log odds  $\leftrightarrow e^{\beta_j}$  multiplicative difference in odds



---

**Remark 0.3** (How do we fit Logistic Regression?)

Maximum Likelihood

$$y_1, \dots, y_n \stackrel{iid}{\sim} \text{Bern}(\pi_i), \pi_i = \frac{1}{1 + \exp(-\beta x_i - \alpha)}$$

Likelihood

$$\begin{aligned}\mathcal{L}(y_1, \dots, y_n | \beta, \alpha, x_1, \dots, x_n) &= \prod_{i=1}^n f(y_i) \\ &= \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \\ &= \prod_{i=1}^n \left( \frac{\pi_i}{1 - \pi_i} \right)^{y_i} (1 - \pi_i) \\ &= \prod_{i=1}^n [\exp(\alpha + \beta x_i)]^{y_i} \frac{1}{1 + \exp(\alpha + \beta x_i)} \\ \log(L(y_1, \dots, y_n | \beta, \alpha, x_1, \dots, x_n)) &= \sum [y_i(\alpha + \beta x_i) - \log(1 + \exp(\alpha + \beta x_i))]\end{aligned}$$

Take derivatives wrt  $\beta, \alpha$  and set to 0

$$\begin{aligned}\frac{\partial}{\partial \alpha} &= \sum_{i=1}^n y_i - \sum_{i=1}^n \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)} \\ \frac{\partial}{\partial \beta} &= \sum_{i=1}^n y_i x_i - \sum_{i=1}^n \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)} x_i\end{aligned}$$

Need to solve

$$\begin{aligned}\sum y_i &= \sum \frac{1}{1 + \exp(-\hat{\alpha} - \hat{\beta} x_i)} \\ \sum y_i x_i &= \sum \frac{1}{1 + \exp(-\hat{\alpha} - \hat{\beta} x_i)} x_i\end{aligned}$$