UCLA School of Education & Information Studies
Department of Education

EDUC 231D
Advanced Quantitative Methods: Multilevel Analysis
Winter 2025

Binary Outcomes

Lecture 14 Presentation Slides

February 25, 2025

Today's Topics

- Overview of binary outcomes
- Single-level logistic regression model
- Two-level logistic regression model

Overview of Binary Outcomes

Binary Outcomes

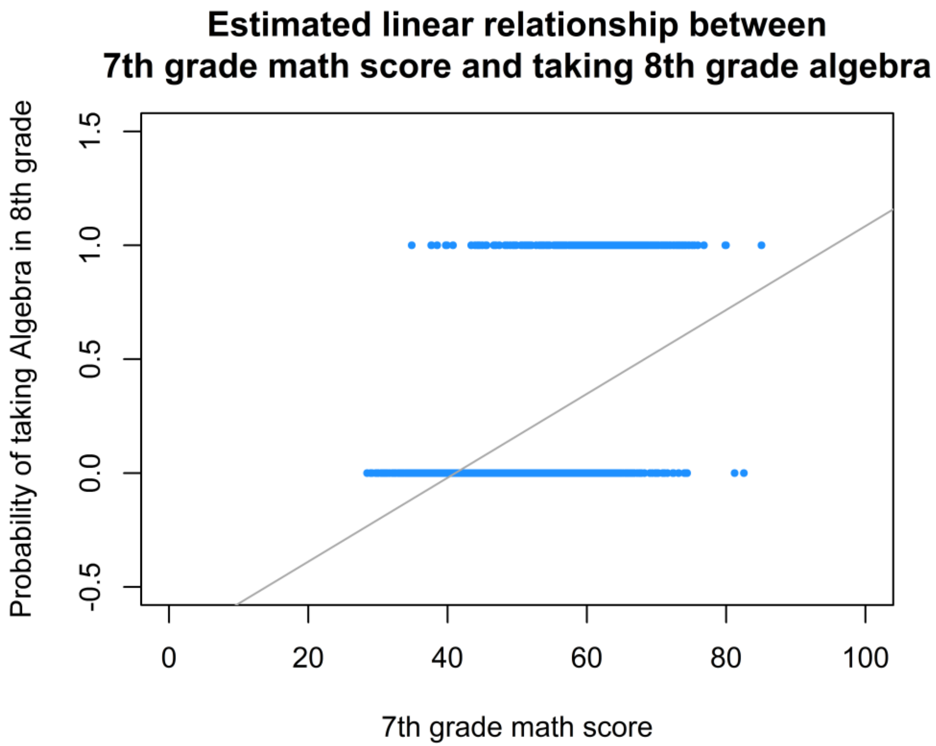
- Across most fields, including education, many outcomes of interest take on binary (or dichotomous) values of 1 or 0
  - Retained in a grade (1) or not (0)
  - Ever suspended from school (1) or not (0)
  - Graduated high school (1) or not (0)
  - Attended college (1) or not (0)
  - Employed (1) or not (0)
  - Have health insurance (1) or not (0)
  - Vaccinated (1) or not (0)
- Linear regression models do not work well with binary outcomes because linear model predictions are not bound to 0 and 1 values

Binary Outcomes

- Consider the example of whether a student takes Algebra 1 in 8th grade or not (from LSAY data)

Proportion of students who take Algebra 1 in 8th grade by parental education level

	Less than college degree (N=1181)	**College degree** (N=566)	**Overall** (N=1747)
Proportion taking Algebra 1	0.145	0.318	0.201



Binary Outcomes

- To model binary outcomes, we need to "alter" the linear regression model ($y_i = \beta_0 + \beta_1 X_i$) in two ways:
  - Add a nonlinear transformation that bounds the output between 0 and 1
  - Treat the model estimates as probabilities that map to a binary outcome

- Logistic regression (or the *logit link function*) is one popular way to model binary outcomes:

$$\Pr(y_i = 1) = p_i$$

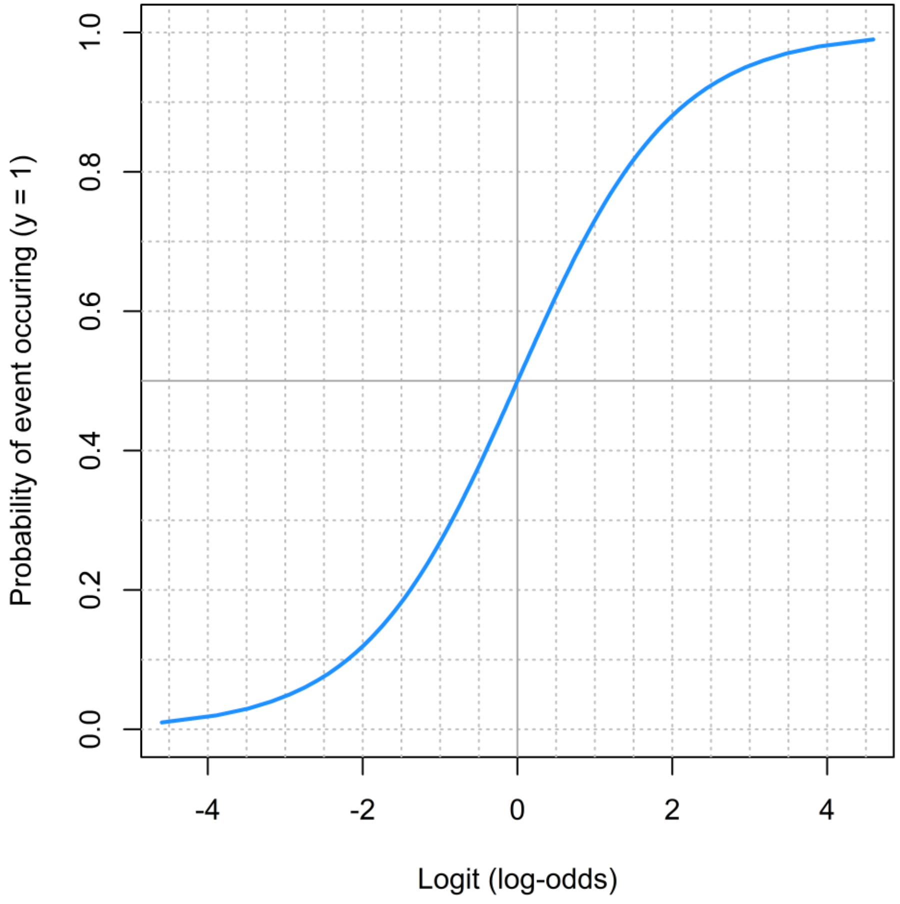
$$logit(p_i) = \ln\left(\frac{p_i}{1 - p_i}\right), \qquad p_i = \frac{e^{logit(p_i)}}{1 + e^{logit(p_i)}}$$

Binary Outcomes

- Relationship between probability (p_i) and logit (log-odds)

$$logit(p_i) = ln\left(\frac{p_i}{1 - p_i}\right)$$

$$p_i = \frac{e^{logit(p_i)}}{1 + e^{logit(p_i)}}$$



Single-Level Logistic Regression Model

Logistic Regression

- The logistic regression model:

$$logit(p_i) = \beta_0 + \beta_1 X_i$$

$$p_i = logit^{-1}(\beta_0 + \beta_1 X_i) = \frac{e^{(\beta_0 + \beta_1 X_i)}}{1 + e^{(\beta_0 + \beta_1 X_i)}}$$

- Coefficient estimates are interpreted in the log-odds scale, but can be converted to an odds ratio or predicted probabilities (for different values of X_i)

Logistic Regression: Algebra 1 Example

- Model relationship between taking Algebra 1 in 8th grade and parent education (college degree or not)

$$logit(p_i) = \beta_0 + \beta_1 PCOLGED_i$$

```
m0 <- glm(ALGIN8 ~ PCOLGED, data = lsayx,
          family = binomial(link = "logit"))
```

	Estimate	Standard Error	z value	Pr(>|z|)	
(Intercept)	-1.776	0.083	-21.478	0.0000	***
PCOLGED	1.013	0.122	8.277	0.0000	***

*Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05*

Logistic Regression: Algebra 1 Example

- Calculate predicted probabilities:
  - $(\hat{p} \mid PCOLGED = 0) = logit^{-1}(\beta_0 + \beta_1 0) = logit^{-1}(-1.776)$
  - $(\hat{p} \mid PCOLGED = 1) = logit^{-1}(\beta_0 + \beta_1 1) = logit^{-1}(-1.776 + 1.013)$

```
library("arm")
p0 <- invlogit(m0$coef[1])
p1 <- invlogit(m0$coef[1] + m0$coef[2])
```

  - $(\hat{p} \mid PCOLGED = 0) = 0.145$
  - $(\hat{p} \mid PCOLGED = 1) = 0.318$

Logistic Regression: Algebra 1 Example

- Model relationship between taking Algebra 1 in 8th grade and 7th grade math score

$$logit(p_i) = \beta_0 + \beta_1 MTHSCORE_i$$

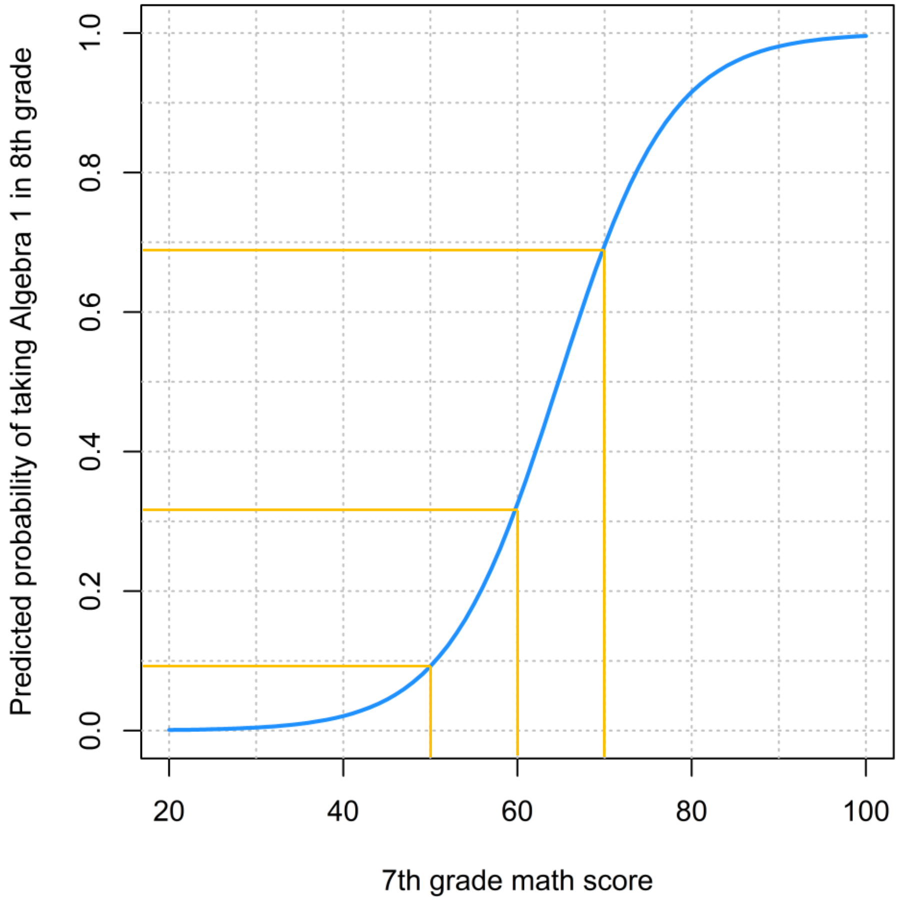
```
m0 <- glm(ALGIN8 ~ MTHSCORE, data = lsayx,
          family = binomial(link = "logit"))
```

	Estimate	Standard Error	z value	Pr(>|z|)	
(Intercept)	-10.067	0.556	-18.119	0.0000	***
MTHSCORE	0.156	0.009	16.480	0.0000	***

*Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05*

Logistic Regression: Algebra 1 Example

- Visualize relationship between math score and predicted probability of taking Algebra 1 in 8th grade
- Compare predicted probability for students with a math score of 50, 60, and 70
  - $(\hat{p} \mid MTHSCORE = 50) = 0.092$
  - $(\hat{p} \mid MTHSCORE = 60) = 0.326$
  - $(\hat{p} \mid MTHSCORE = 70) = 0.696$



Two-Level Logistic Regression Model

Unconditional Two-Level Logistic Regression

- Level 1 (student-level):

$$logit(p_{ij}) = \beta_{0j}$$

Notice that there's no level-1 residual term. For logistic regression, the level-1 variance is $\frac{\pi^2}{3} \approx 3.29$

- Level 2 (school-level):

$$\beta_{0j} = \gamma_{00} + u_{0j}\,, \qquad u_{0j} \sim N(0\,, \tau_{00})$$

Unconditional Two-Level Logistic Regression: Algebra 1 Example

- Estimate between-school variation in probability of taking Algebra 1 in 8th grade

$$logit(p_{ij}) = \gamma_{00} + u_j$$

```
m1 <- glmer(ALGIN8 ~ 1 + (1 | SCHOOLID),
            data = lsayx, family = binomial(link = "logit"))
```

Unconditional Two-Level Logistic Regression: Algebra 1 Example

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: ALGIN8 ~ 1 + (1 | SCHOOLID)
   Data: lsayx

     AIC       BIC    logLik deviance df.resid
  1680.0    1690.9    -838.0   1676.0     1745

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.8648 -0.5378 -0.4045 -0.2079  3.4966

Random effects:
 Groups   Name        Variance Std.Dev.
 SCHOOLID (Intercept) 0.843    0.9182
Number of obs: 1747, groups:  SCHOOLID, 50

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.6401     0.1559  -10.52   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Unconditional Two-Level Logistic Regression: Algebra 1 Example

- To what extent does the probability of taking Algebra 1 differ across schools?
- Intraclass correlation (ICC) calculation for multilevel logistic regression:

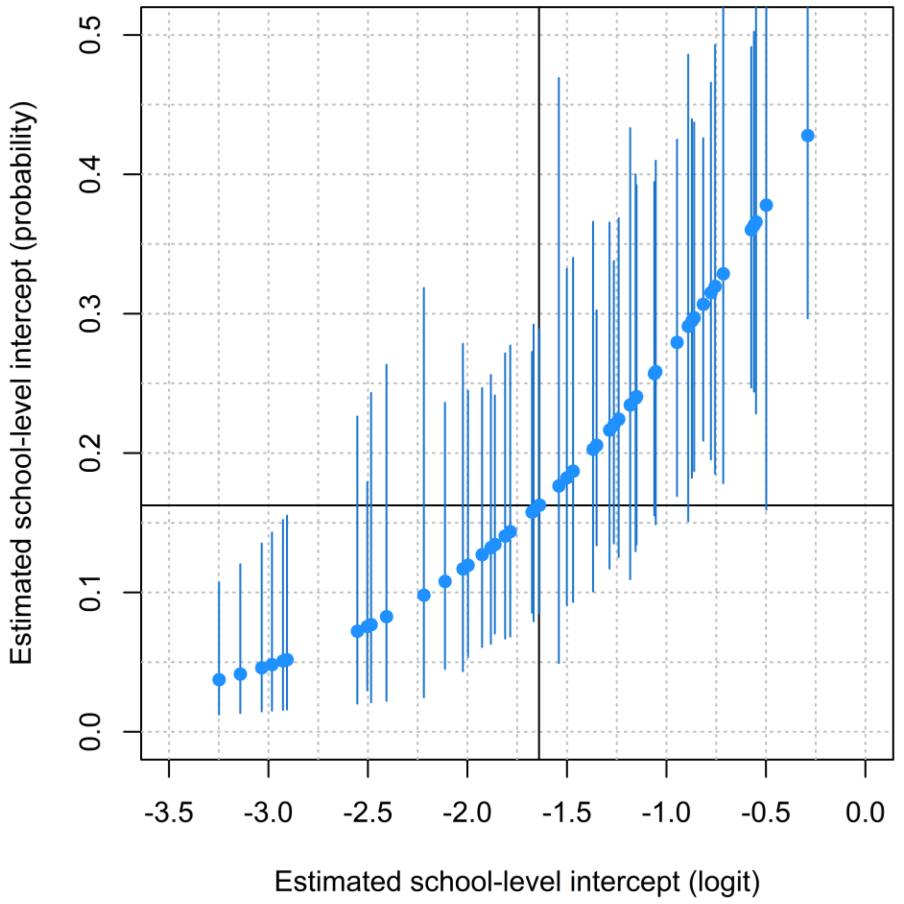
$$\frac{\tau_{00}}{\tau_{00} + \frac{\pi^2}{3}} = \frac{0.843}{0.843 + 3.29} = 0.204$$

- But the ICC can be a little misleading for logistic regression

Unconditional Two-Level Logistic Regression: Algebra 1 Example

- To what extent does the probability of taking Algebra 1 differ across schools?
- Visually inspect the school-level random effects and predicted probabilities

```
B0_j <- coef(m1)$SCHOOLID[,1]
p_j <- invlogit(B0_j)
```



Random Coefficients Logistic Regression

- Level 1 (student-level):

$$logit(p_{ij}) = \beta_{0j} + \beta_{1j}(X_{ij} - \bar{X}_{.j})$$

- Level 2 (school-level):

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Random Coefficients Logistic Regression : Algebra 1 Example

- Level 1 (student-level):

$$logit(p_{ij}) = \beta_{0j} + \beta_{1j}(PCOLGED_{ij} - \overline{PCOLGED}_{.j}) + \beta_{2j}(MTHSCORE_{ij} - \overline{MTHSCORE}_{.j})$$

- Level 2 (school-level):

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} + u_{2j} \end{aligned}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{pmatrix} \sim MVN(\mathbf{0}, \mathbf{T}), \qquad \mathbf{T} = \begin{pmatrix} \tau_{00} & . & . \\ . & \tau_{11} & . \\ . & . & \tau_{22} \end{pmatrix}$$

For this example, not estimating the covariances because of estimation issues

Random Coefficients Logistic Regression : Algebra 1 Example

```
m2 <- glmer(ALGIN8 ~ 1 + PCOLG.gpc + MATH.gpc
            + (1 + PCOLG.gpc + MATH.gpc || SCHOOLID),
            data = lsayx, family = binomial(link = "logit"))
```

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: ALGIN8 ~ 1 + PCOLG.gpc + MATH.gpc + (1 + PCOLG.gpc + MATH.gpc ||      SCHOOLID)
   Data: lsayx

     AIC      BIC   logLik deviance df.resid
  1205.4   1238.2   -596.7   1193.4     1741

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.6543 -0.3719 -0.1624 -0.0526 19.3572

Random effects:
 Groups     Name        Variance Std.Dev.
 SCHOOLID   (Intercept) 1.650480 1.28471
 SCHOOLID.1 PCOLG.gpc   0.100318 0.31673
 SCHOOLID.2 MATH.gpc    0.005543 0.07445
Number of obs: 1747, groups:  SCHOOLID, 50

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.50843    0.23665 -10.600  < 2e-16 ***
PCOLG.gpc    0.49585    0.19067   2.601  0.00931 **
MATH.gpc     0.20670    0.01813  11.402  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Random Coefficients Logistic Regression: Algebra 1 Example

- How does the relationship between math score and the probability of taking Algebra 1 differ across schools?
- Holding parent education constant at the school mean

```
B0_j <- coef(m2)$SCHOOLID[,1]
B2_j <- coef(m2)$SCHOOLID[,3]
p <- invlogit(B0_j[j] +
        B2_j[j]*(x - smmath[j]))
```