Agenda
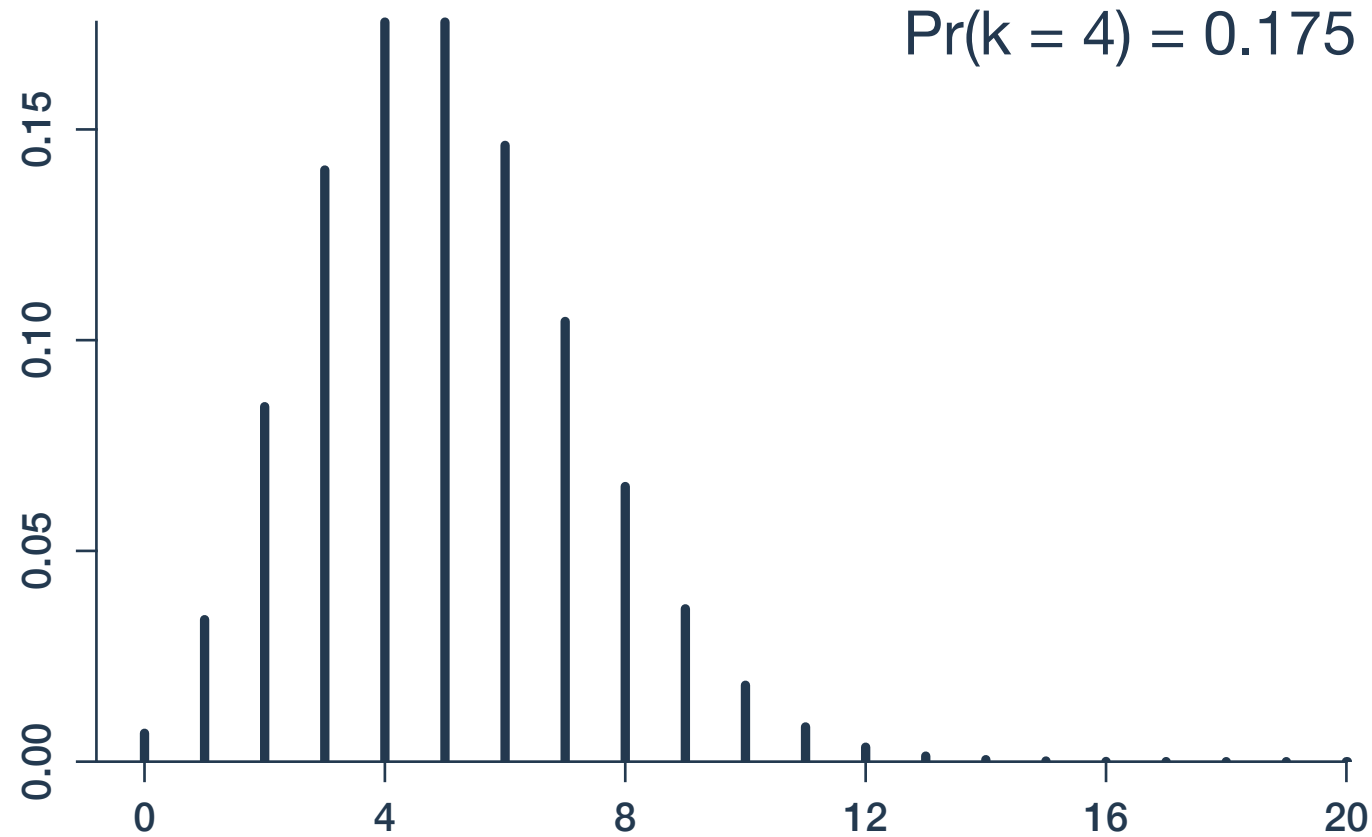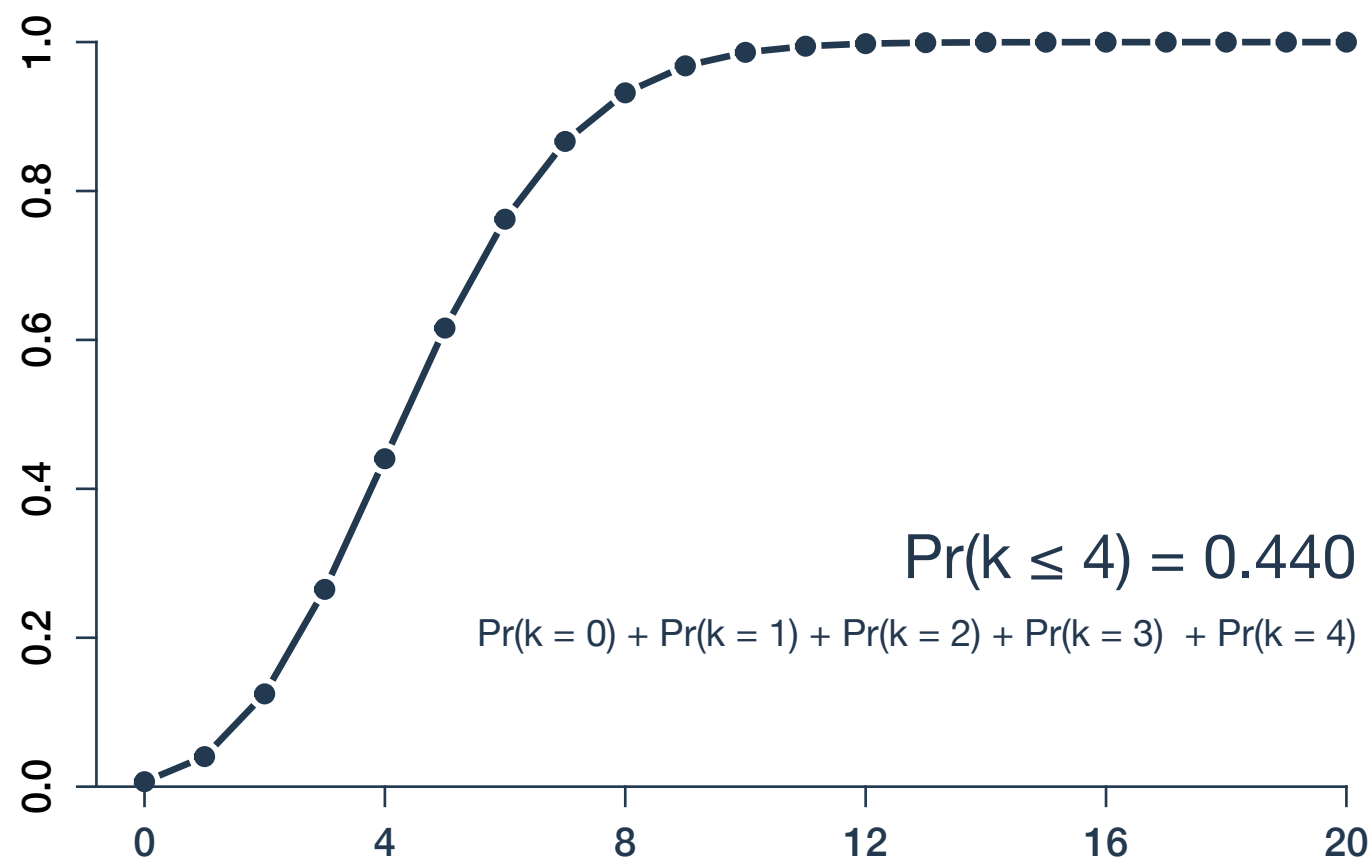
1. Cumulative probability distributions

2. Predicting educational attainment

3. Log cumulative odds link

4. Intercept-only ordered logit

5. Ordered logit with predictors

6. Estimating ordered logit in R

Cumulative probability distributions

Probability mass function (PMF)

Pois(5)
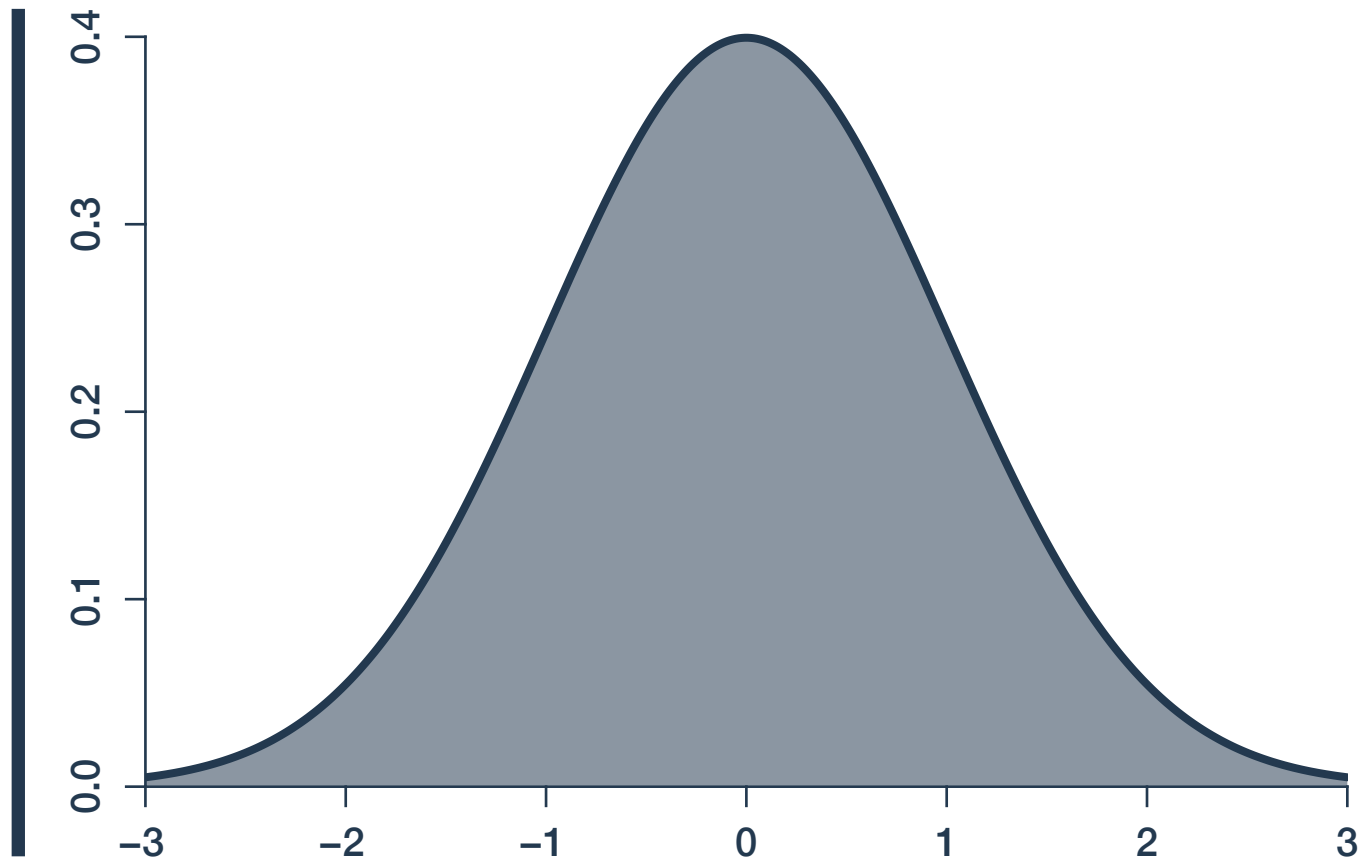
$Pr(k = 4) = 0.175$

Cumulative distribution function (CDF)

Pois(5)

$Pr(k \leq 4) = 0.440$

$Pr(k = 0) + Pr(k = 1) + Pr(k = 2) + Pr(k = 3) + Pr(k = 4)$

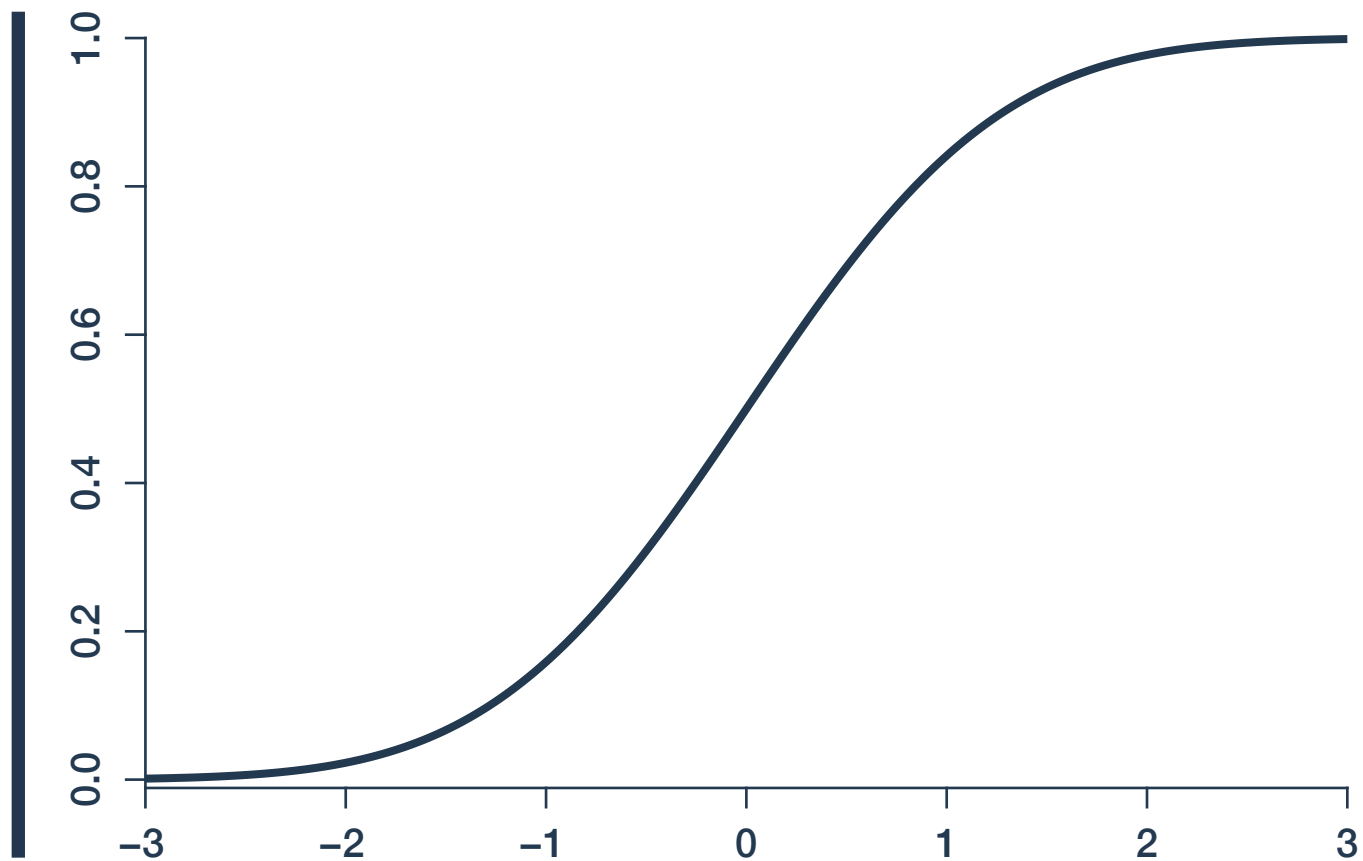# Cumulative probability distributions

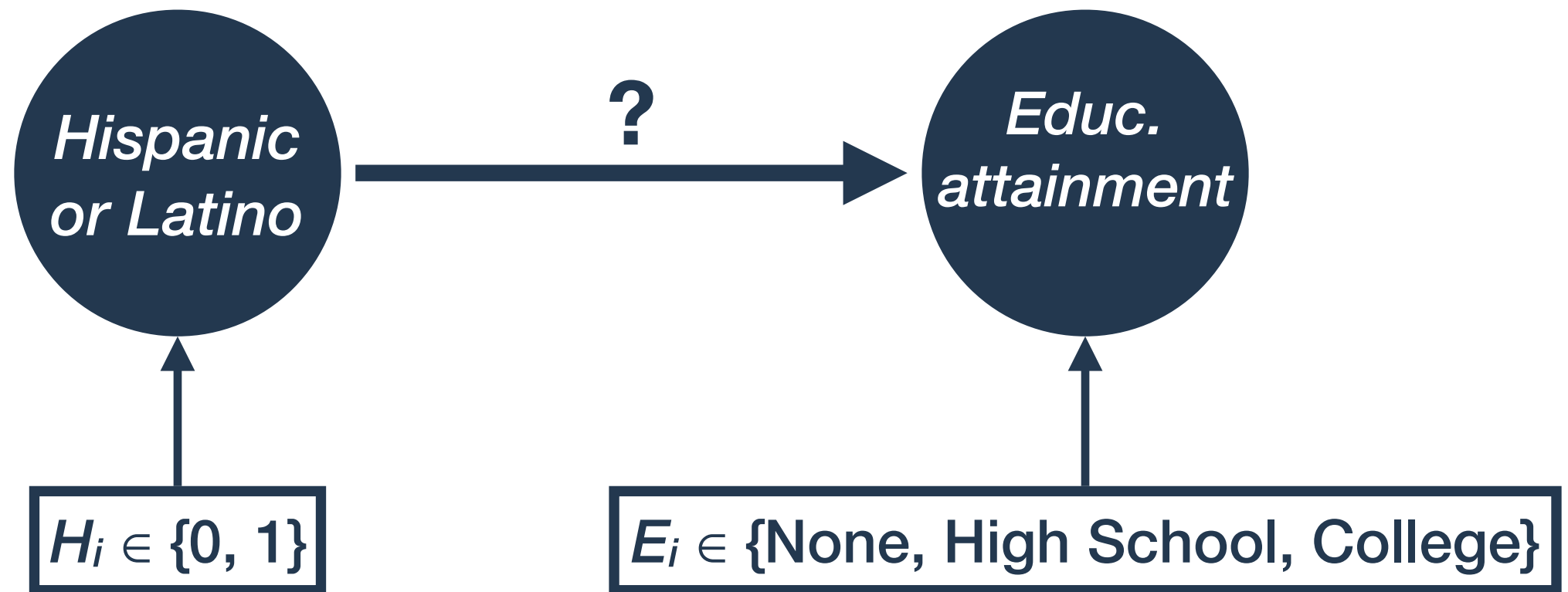**Probability density function (PDF)**

Norm(0, 1)

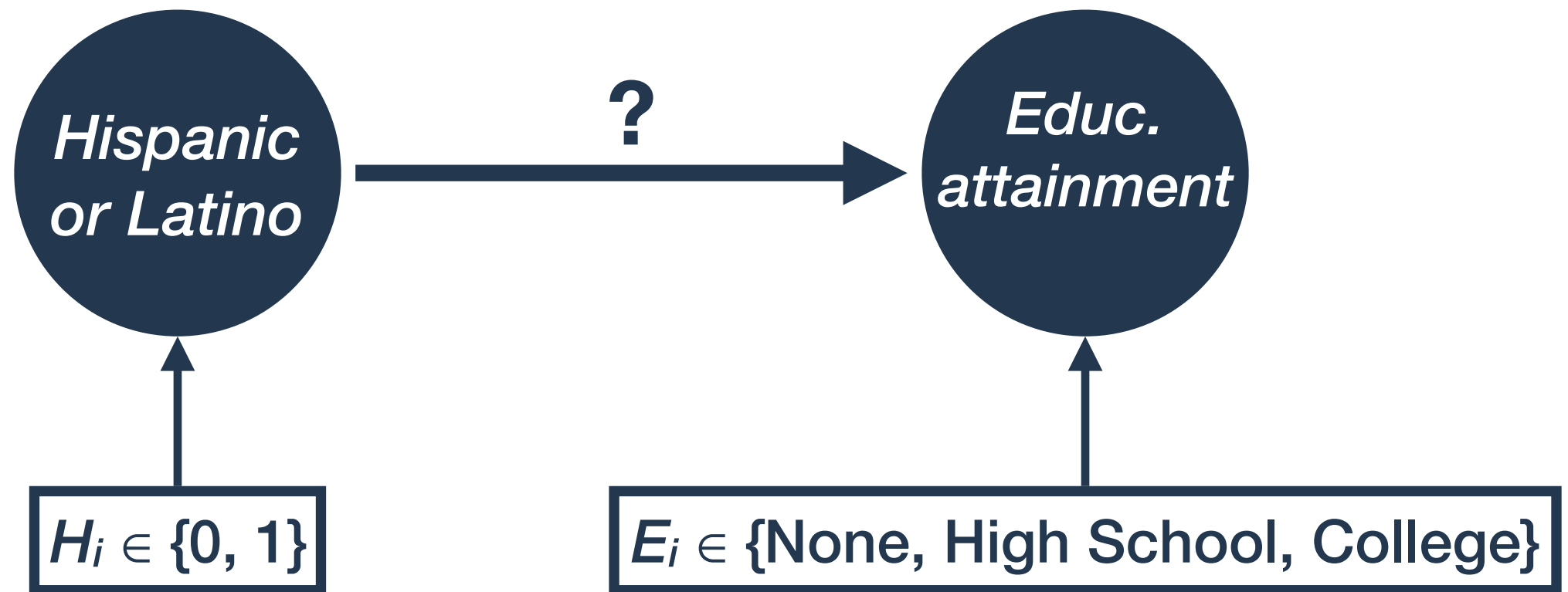**Cumulative distribution function (CDF)**

Norm(0, 1)

Age and education

The problem | Educational attainment is measured as categories, but those categories are ordered in an important way.
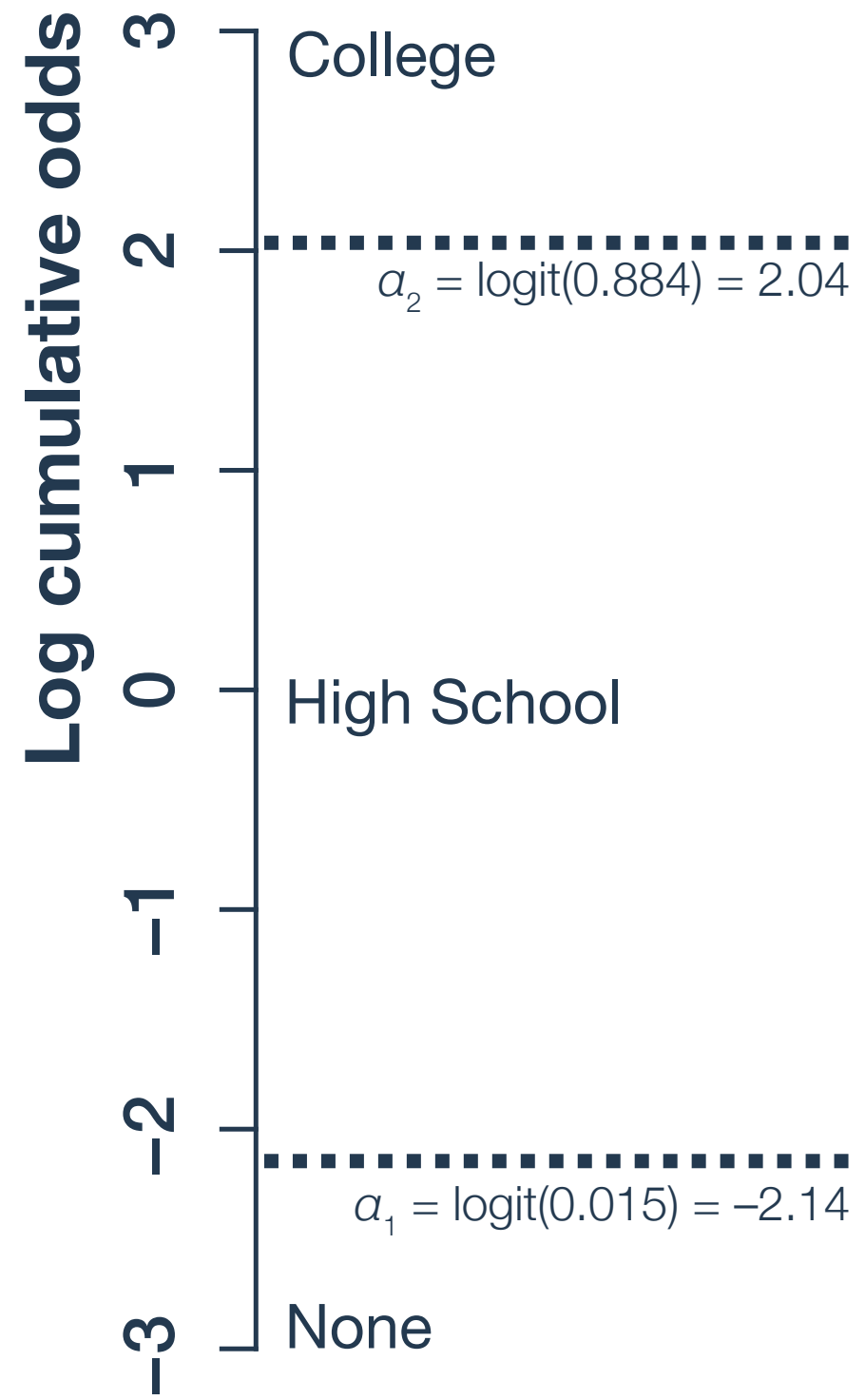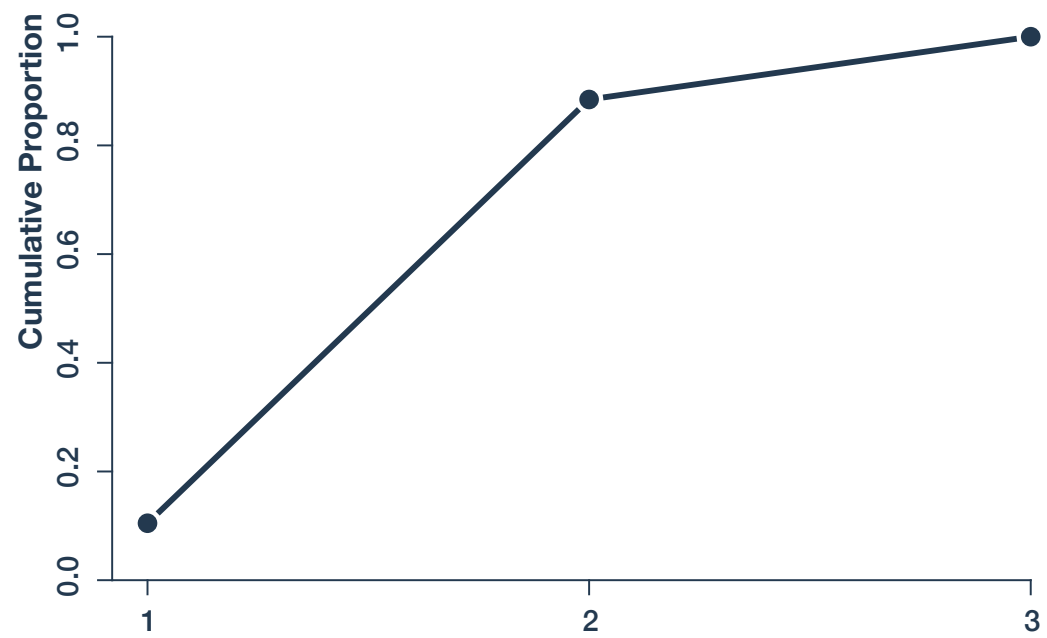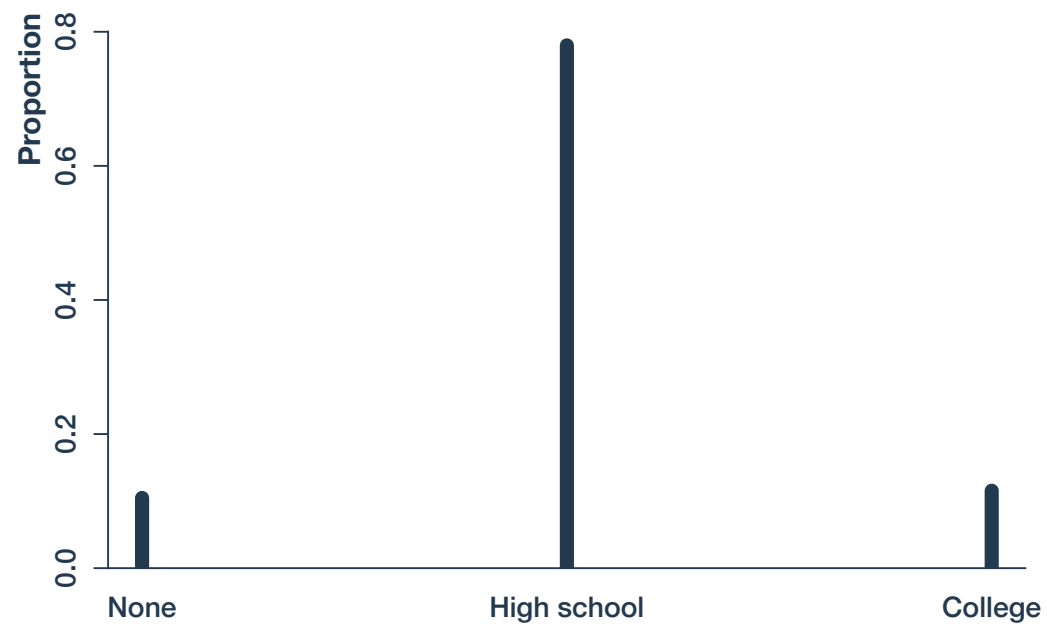
The solution | Treat education as a categorically-distributed variable, but constrain the probabilities to respect the order.
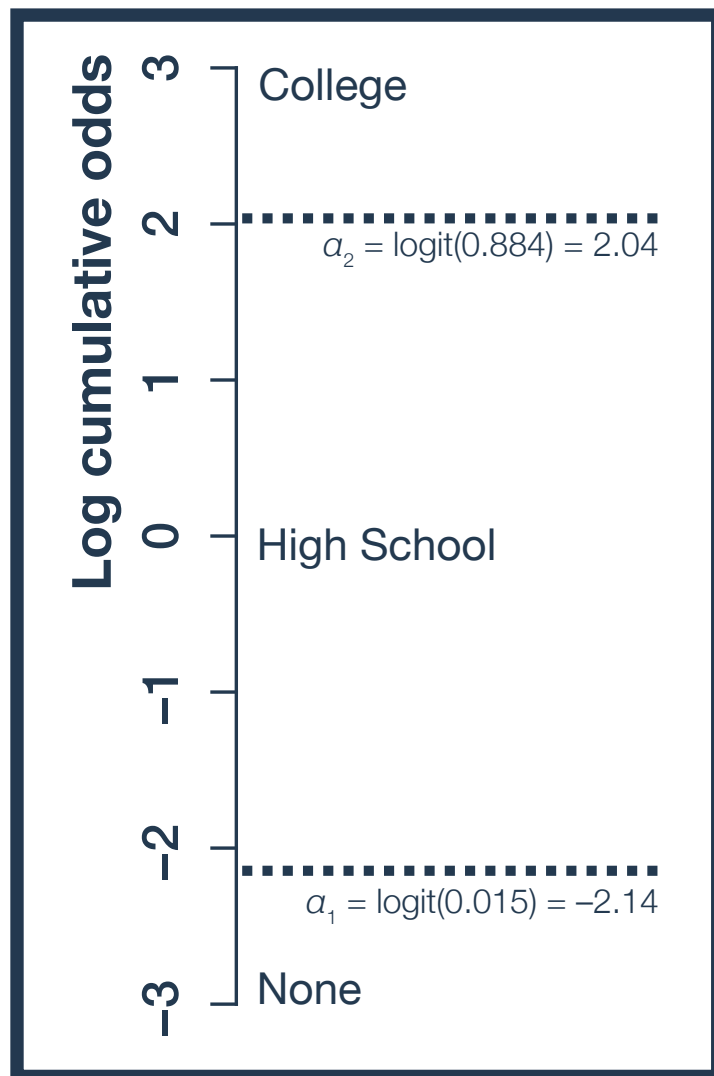
Age and education

$$E_i \sim \text{Categorical}(p_1, p_2, p_3)$$
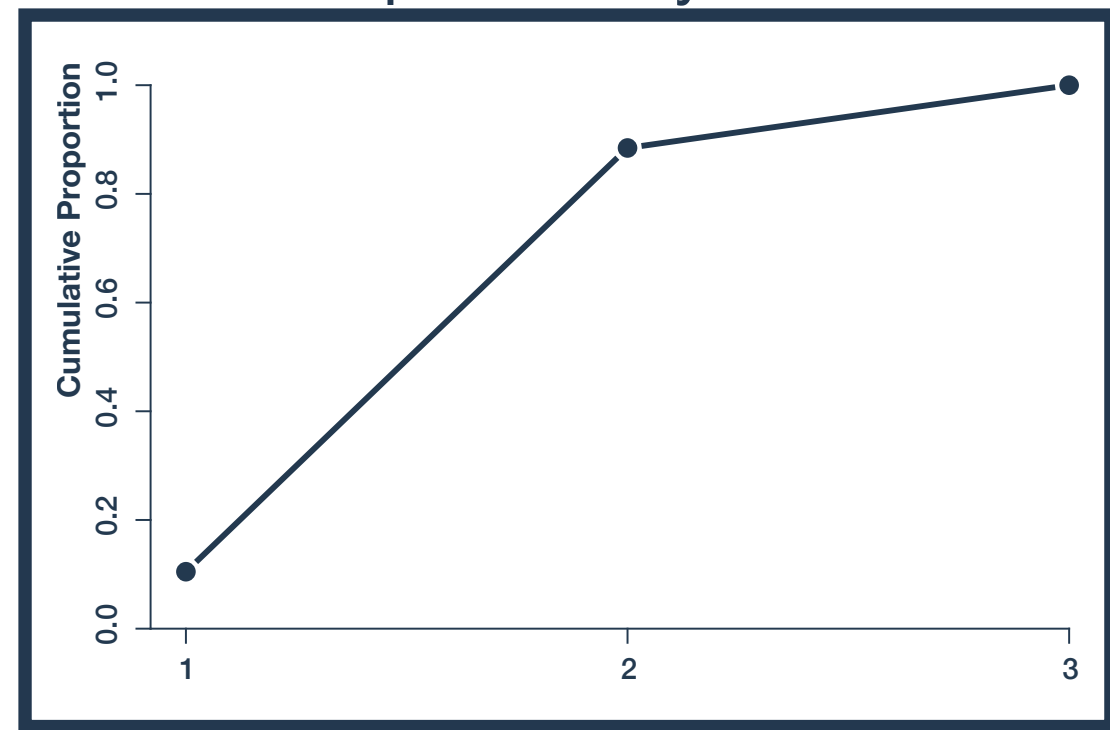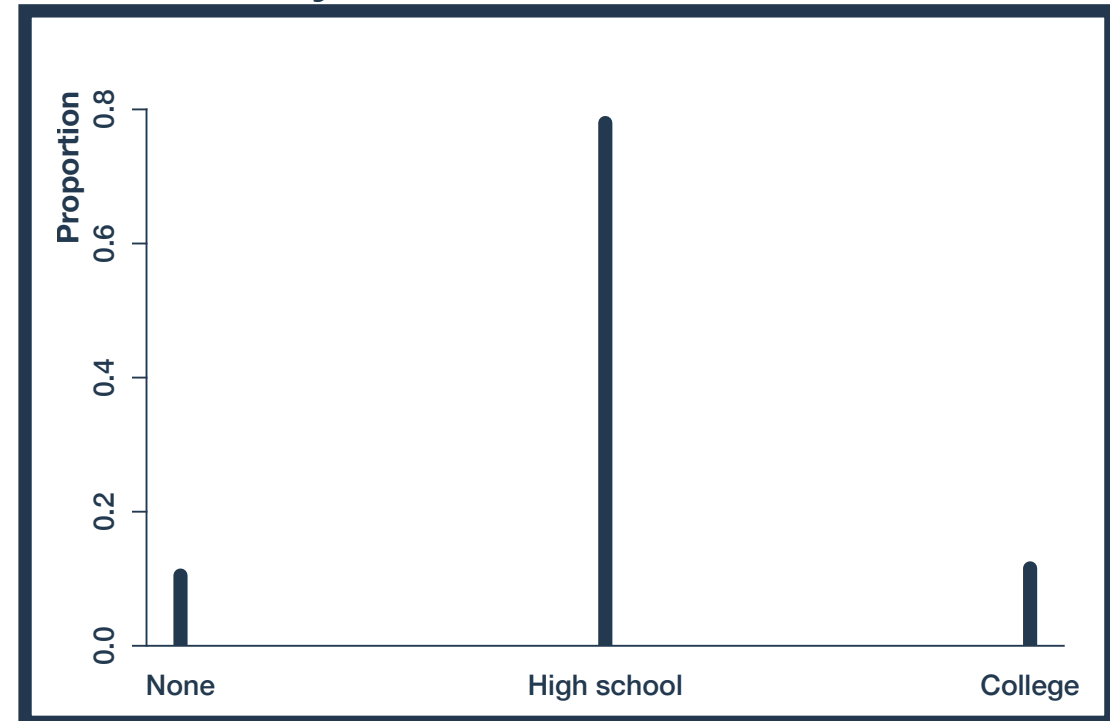
Log cumulative odds

$\alpha_2 = \text{logit}(0.884) = 2.04$

$\alpha_1 = \text{logit}(0.015) = -2.14$

Log cumulative odds

Parameter scale

Log cumulative odds

College

$a_2 = \text{logit}(0.884) = 2.04$

High School

$a_1 = \text{logit}(0.015) = -2.14$

None

Inverse logit

Cumulative probability

Cumulative Proportion

Subtraction

Probability

Proportion

None   High school   College

Ordered logit model

Intercept-only ordered logit model
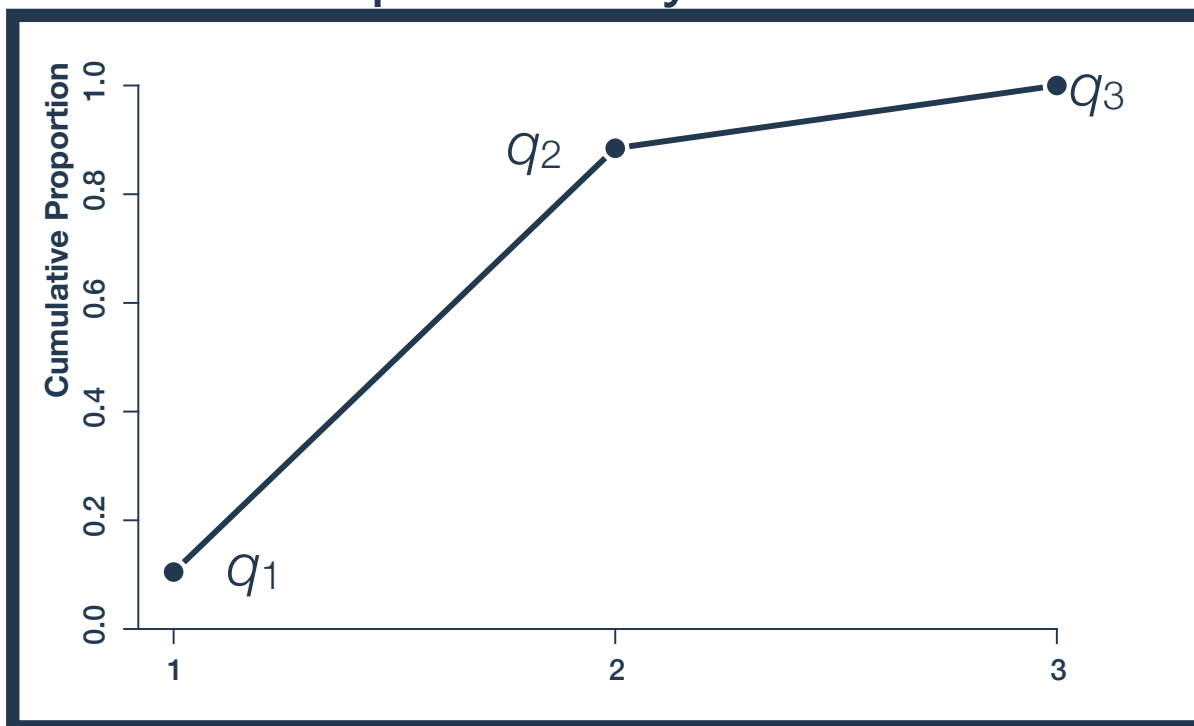
$$E_i \sim \text{Categorical}(\mathbf{p})$$

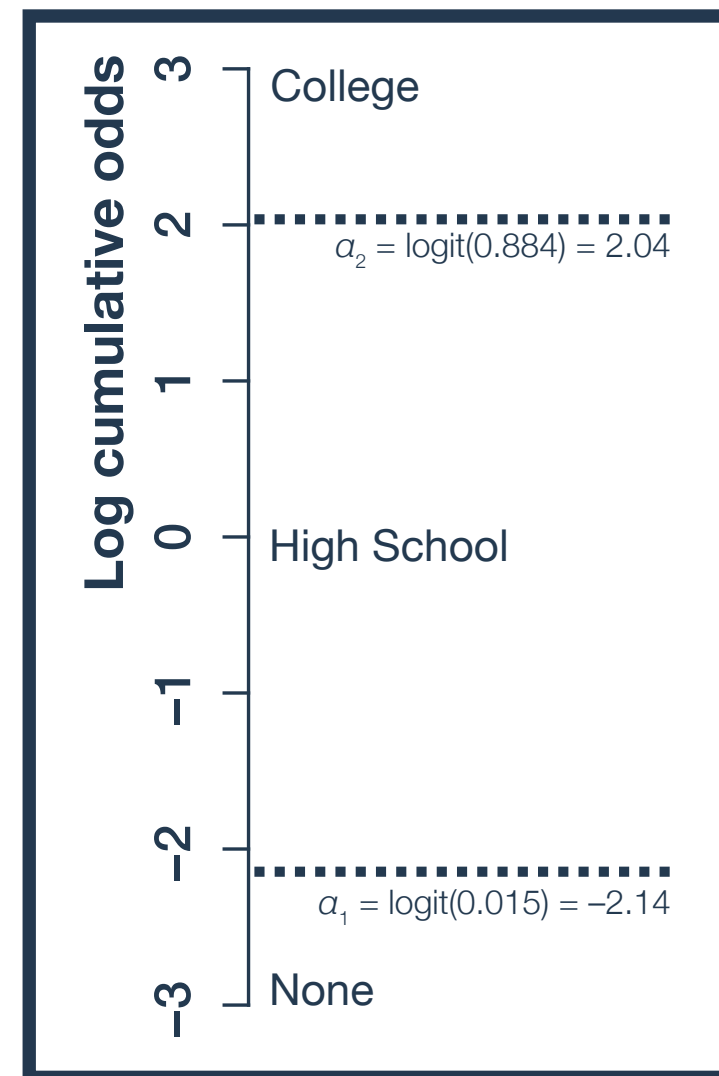$$p_k = q_k - q_{k-1}$$

$$\text{logit}(q_k) = a_k$$

q_k is the cumulative probability of category k

$q_0 = 0; q_1 = 1$

Cumulative probability

Parameter scale

College

$a_2 = \text{logit}(0.884) = 2.04$

High School

$a_1 = \text{logit}(0.015) = -2.14$

None

Ordered logit model

Ordered logit with predictors
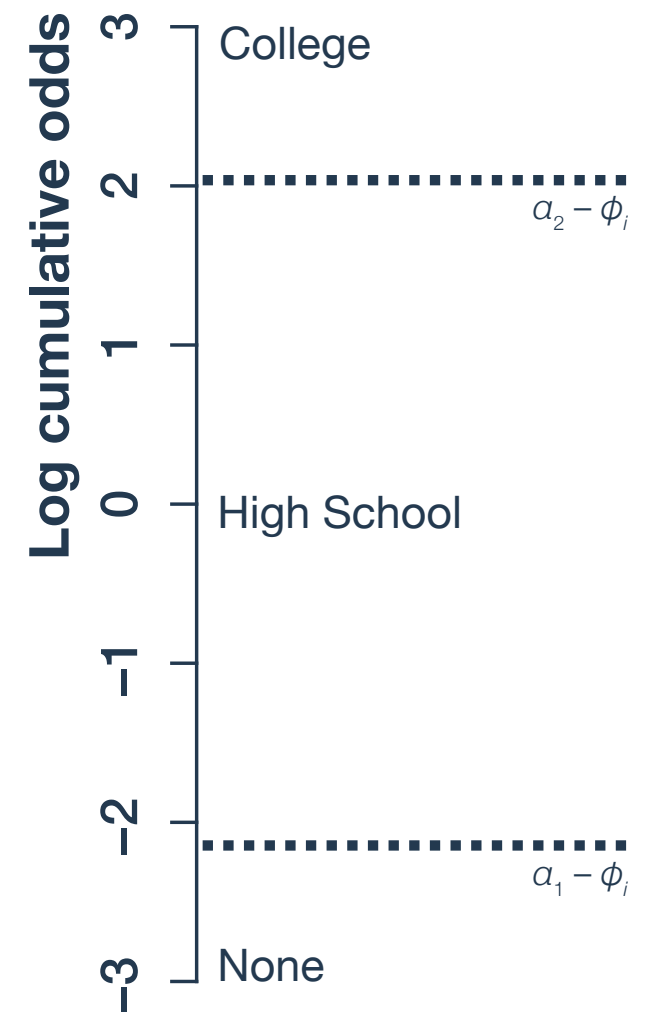
$$E_i \sim \text{Categorical}(\mathbf{p})$$

$$p_k = q_k - q_{k-1}$$

$$\text{logit}(q_k) = a_k - \phi_i$$

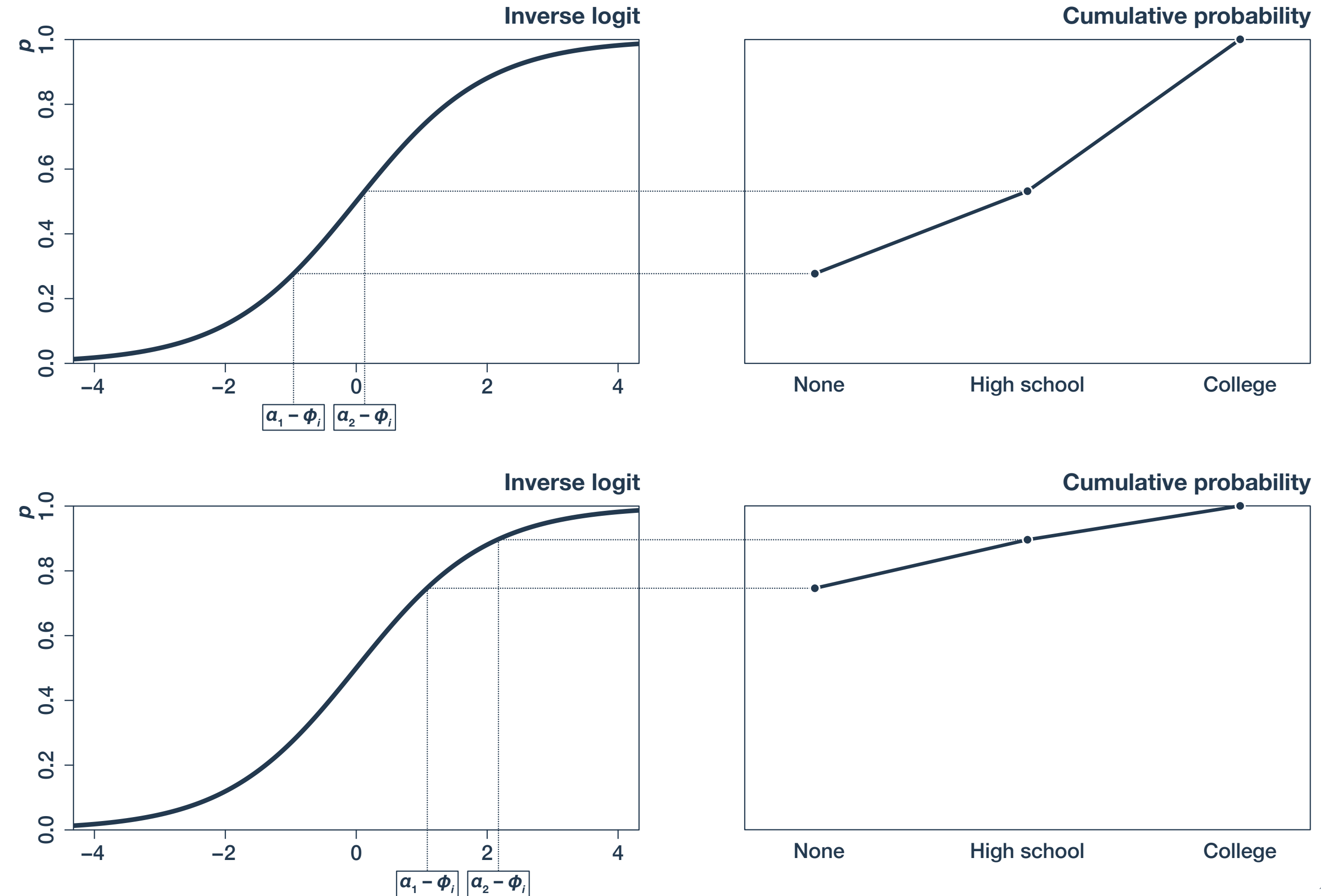$$\phi_i = \beta H_i$$

$$a_k \sim \text{Norm}(0, 1.5)$$

$$\beta \sim \text{Norm}(0, 2)$$

Ordered logit model

Ordered logit model