

Feb. 16

Expanding on
Poisson regressions

1. Interpreting coefficients from Poisson regressions
2. Over-dispersed Poisson regressions
3. Zero-inflated Poisson regressions
4. Over-dispersion and zero-inflation in R

Interpreting coefficients

Model from last time

$$H_i \sim \text{Pois}(\lambda_i)$$
$$\log(\lambda_i) = \alpha + \beta_M M_i + \beta_G G_i$$

$$\alpha \sim \text{Norm}(3, 1)$$

$$\beta_M \sim \text{Norm}(0, 0.5)$$

$$\beta_G \sim \text{Norm}(0, 0.3)$$

H_i
Hours of video games played last week

M_i
Student gender (indicator for boys)

G_i
Student's grade, centered at 10th

	<i>Mean</i>	<i>90% credible interval</i>		<i>exp(Mean)</i>
α	0.27	0.25	0.30	1.32
β_M	1.11	1.09	1.14	3.05
β_G	-0.14	-0.15	-0.13	0.87

Interpreting coefficients

Model from last time

$$H_i \sim \text{Pois}(\lambda_i)$$
$$\log(\lambda_i) = \alpha + \beta_M M_i + \beta_G G_i$$
$$\alpha \sim \text{Norm}(3, 1)$$
$$\beta_M \sim \text{Norm}(0, 0.5)$$
$$\beta_G \sim \text{Norm}(0, 0.3)$$

	<i>Mean</i>	<i>exp(Mean)</i>
α	0.27	1.32
β_M	1.11	3.05
β_G	-0.14	0.87

α (baseline)

A student who is not a boy ($M_i=0$) and is in grade 10 ($G_i=0$) is predicted to play about 1.32 hours of games per week.

β_M (gender)

Boys ($M_i=1$) are expected to spend about 3.05 times more time than non-boys ($M_i=0$) playing games.

β_G (grade)

A one-year increase in grade is associated with playing 0.87 times as much. This is a decrease of 13% each year.

Over-dispersion

$$H_i \sim \text{Pois}(\lambda_i)$$

$$\log(\lambda_i) = \alpha + \beta M_i$$

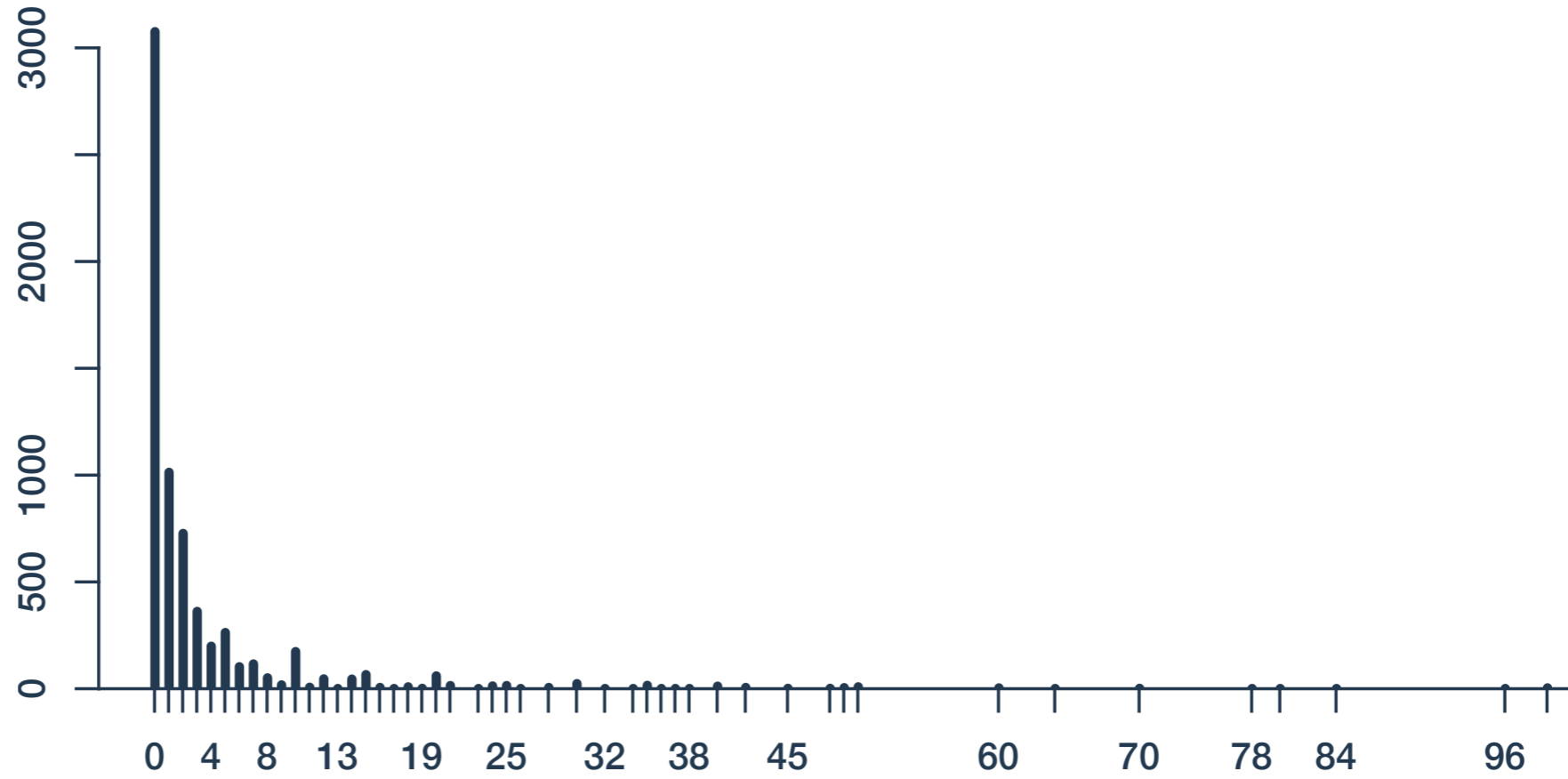
$$\alpha \sim \text{Norm}(3, 0.5)$$

$$\beta \sim \text{Norm}(0, 0.3)$$

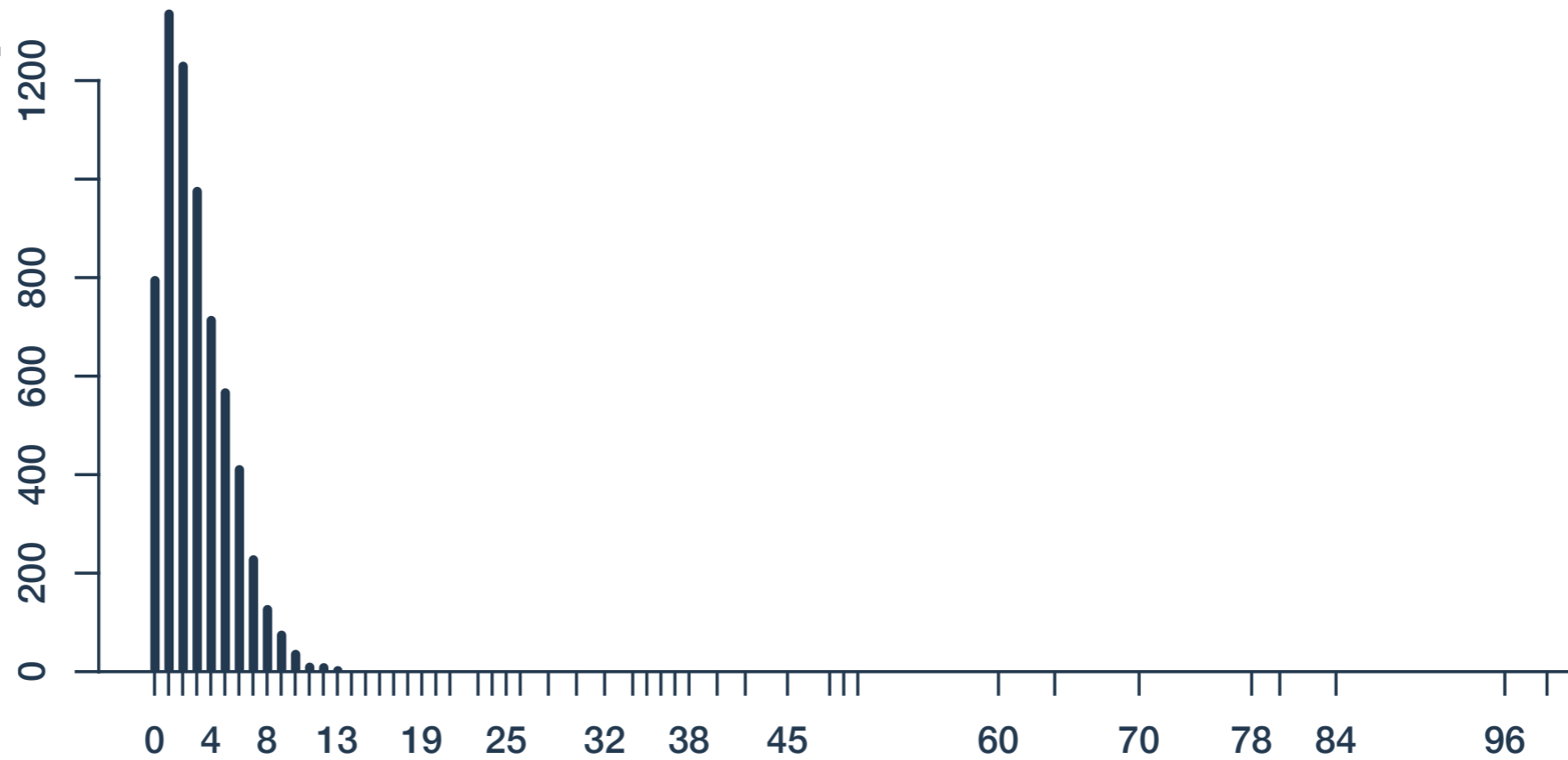
	<i>Mean</i>	<i>exp(Mean)</i>
α	0.27	1.32
β	1.11	3.05

Over-dispersion

**Actual
distribution**



**Posterior
predicted
distribution
(Poisson
regression)**



Over-dispersion

Poisson regression

$$H_i \sim \text{Pois}(\lambda_i)$$

$$\log(\lambda_i) = \alpha + \beta M_i$$

$$\alpha \sim \text{Norm}(3, 0.5)$$

$$\beta \sim \text{Norm}(0, 0.3)$$

Gamma-Poisson regression

$$H_i \sim \text{Pois}(\lambda_i)$$

$$\lambda_i \sim \text{Gamma}(\mu_i, \theta)$$

$$\log(\mu_i) = \alpha + \beta M_i$$

$$\alpha \sim \text{Norm}(3, 0.5)$$

$$\beta \sim \text{Norm}(0, 0.3)$$

$$\theta \sim \text{Unif}(0, 10)$$

Extra “dispersion”
from gamma

Two students who
look identical based
on covariates can
have different
Poisson rates λ_i .

One more prior

A.K.A.

Negative-binomial regression

Over-dispersed Poisson regression

Over-dispersion

$$H_i \sim \text{GammaPois}(\lambda_i, \theta)$$

$$\log(\lambda_i) = \alpha + \beta M_i$$

$$\alpha \sim \text{Norm}(3, 0.5)$$

$$\beta \sim \text{Norm}(0, 0.3)$$

$$\theta \sim \text{Unif}(0, 10)$$

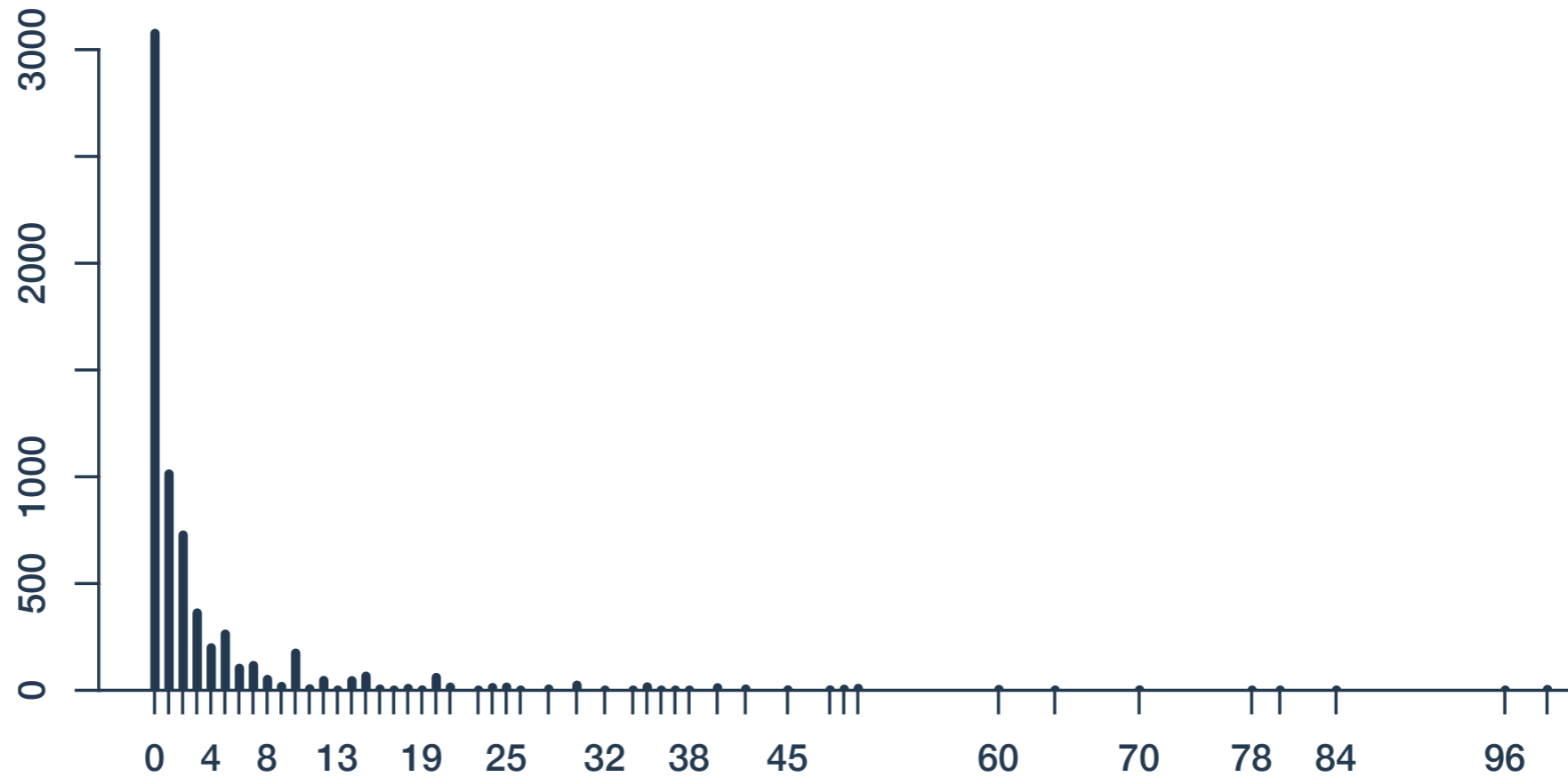
	<i>Mean</i>	<i>90% credible interval</i>		<i>exp(Mean)</i>
α	0.55	0.50	0.61	1.74
β	0.86	0.80	0.91	2.35
θ	8.56	8.09	9.03	—

θ measures extra dispersion

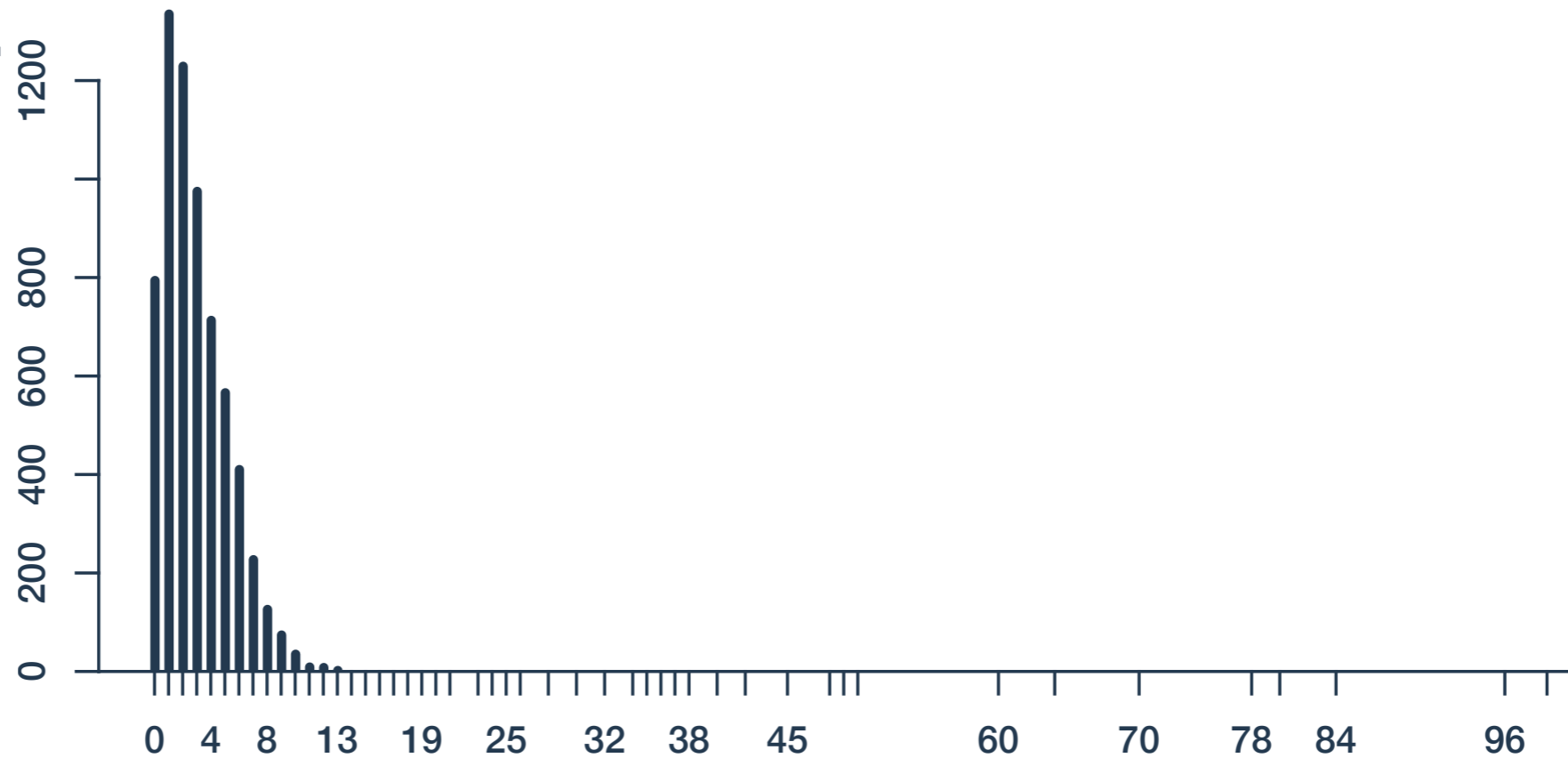


Zero-inflation

**Actual
distribution**



**Posterior
predicted
distribution
(Poisson
regression)**



Zero-inflation

Outcome variable is result of one of two processes

Either the student does not own a game console ($c_i = 1$) **or** the student does own a console and plays at some rate λ_i ($c_i = 0$).



$$H_i \begin{cases} = 0 & \text{if } c_i = 1 \\ \sim \text{Pois}(\lambda_i) & \text{if } c_i = 0 \end{cases}$$

Zero-inflation

Their chance of owning a console is modeled with p_i



$$H_i \begin{cases} = 0 & \text{if } c_i = 1 \\ \sim \text{Pois}(\lambda_i) & \text{if } c_i = 0 \end{cases}$$

$$c_i \sim \text{Bern}(p_i)$$

Zero-inflation

p_i is modeled as a linear function of family income



$$H_i \begin{cases} = 0 & \text{if } c_i = 1 \\ \sim \text{Pois}(\lambda_i) & \text{if } c_i = 0 \end{cases}$$

$$c_i \sim \text{Bern}(p_i)$$

$$\text{logit}(p_i) = \alpha_p + \beta_p W_i$$

Zero-inflation

λ_i is modeled as a linear function of gender



$$H_i \begin{cases} = 0 & \text{if } c_i = 1 \\ \sim \text{Pois}(\lambda_i) & \text{if } c_i = 0 \end{cases}$$

$$c_i \sim \text{Bern}(p_i)$$

$$\text{logit}(p_i) = a_p + \beta_p W_i$$

$$\log(\lambda_i) = a_\lambda + \beta_\lambda M_i$$

Zero-inflation

All four parameters
need priors



$$H_i \begin{cases} = 0 & \text{if } c_i = 1 \\ \sim \text{Pois}(\lambda_i) & \text{if } c_i = 0 \end{cases}$$

$$c_i \sim \text{Bern}(p_i)$$

$$\text{logit}(p_i) = a_p + \beta_p W_i$$

$$\log(\lambda_i) = a_\lambda + \beta_\lambda M_i$$

$$a_p \sim \text{Norm}(0, 1)$$

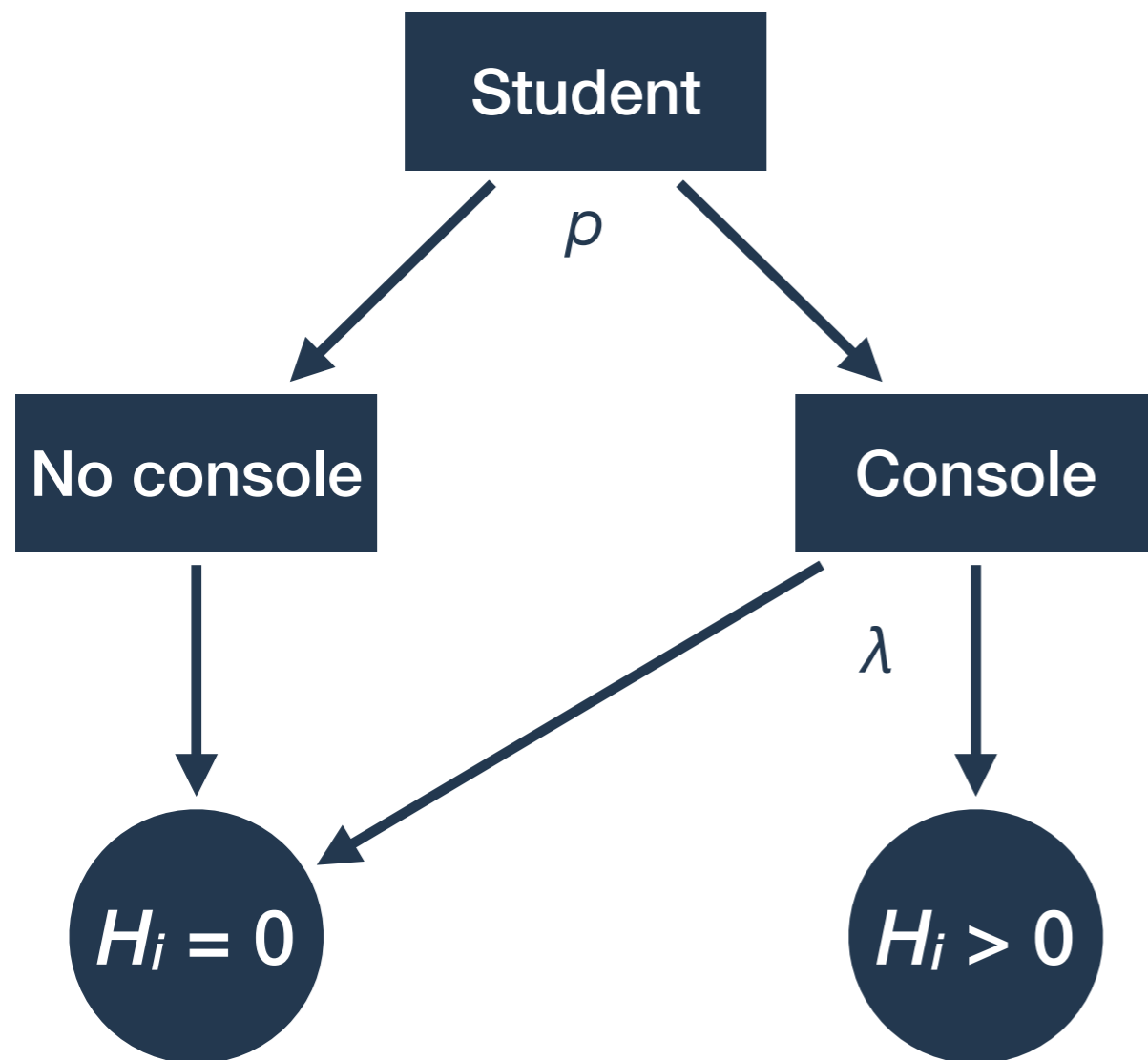
$$\beta_p \sim \text{Norm}(0, 2)$$

$$a_\lambda \sim \text{Norm}(3, 0.5)$$

$$\beta_\lambda \sim \text{Norm}(0, 0.3)$$

Zero-inflation

Data story:



$$H_i \begin{cases} = 0 & \text{if } c_i = 1 \\ \sim \text{Pois}(\lambda_i) & \text{if } c_i = 0 \end{cases}$$

$$c_i \sim \text{Bern}(p_i)$$

$$\text{logit}(p_i) = a_p + \beta_p W_i$$

$$\log(\lambda_i) = a_\lambda + \beta_\lambda M_i$$

$$a_p \sim \text{Norm}(0, 1)$$

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

$$a_\lambda \sim \text{Norm}(3, 0.5)$$

$$\beta_\lambda \sim \text{Norm}(0, 0.3)$$