- Project 2 due Monday, April 24
- Exam 2 on Wednesday (April 12)
    - Bring calculator
    - Closed notes
    - Similar format to last exam

Question 4 (review activity)

Q: is there a relationship between income & # doctor visits?

Population model
$$\begin{cases} \text{Visits}_i \sim \text{Poisson}(\lambda_i) \\ \log(\lambda_i) = \beta_0 + \beta_1 \text{Female}_i + \beta_2 \text{Age}_i + \beta_3 \text{Income}_i + \\ \qquad \beta_4 \text{Private}_i + \beta_5 \text{Freepoor}_i + \beta_6 \text{Freerepat}_i \end{cases}$$

$H_0: \beta_3 = 0$ $\qquad$ $H_A: \beta_3 \neq 0$

Wald test $\qquad$ or LRT

$\hat{\beta}_3 = -0.30$ $\qquad$ $SE(\hat{\beta}_3) = 0.09$

$$\text{test stat} = \frac{-0.30 - 0}{0.09} \approx -3.46$$

p-value $= 0.00055$

Wald test: $\quad Z = \dfrac{\hat{\beta} - b}{SE(\hat{\beta})}$

$H_0: \beta = b$

compare to $N(0,1)$

LRT: $2\log L_{full} - 2\log L_{reduced}$

(for Poisson, $=$ deviance$_{reduced}$ $-$ deviance$_{full}$)

compare to χ^2_q $\quad$ $q = \#$ parameters tested

Question 10 : what is a potential latent variable?

whether the study participant ever goes to the doctor

Test hypotheses with ZIP models

HW 9:

$$\log\left(\frac{\alpha_i}{1-\alpha_i}\right) = \gamma_0 + \gamma_1 \text{Child}_i + \gamma_2 \text{Persons}_i + \gamma_3 \text{Camper}_i$$

$\alpha_i = P(\text{do not go fishing})$

$$\log(\lambda_i) = \beta_0 + \beta_1 \text{Child}_i + \beta_2 \text{Persons}_i + \beta_3 \text{Camper}_i + \log(\text{LOS}_i)$$

4(a): Are groups with more children less likely to go fishing?

$H_0: \gamma_1 = 0$ $\quad H_A: \gamma_1 > 0$

(equiv: $H_0: \gamma_1 \leq 0$)

Wald test:

$$\text{test stat} = \frac{\hat{\gamma}_1 - 0}{SE(\hat{\gamma}_1)}$$

(compare to $N(0,1)$)

4(b): Do groups with more children, who do go fishing, catch fewer fish per day?

$H_0: \beta_1 = 0$ $\quad H_A: \beta_1 < 0$

$$\text{test stat: } \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)}$$

Question 12: want estimated # of doctor visits for a male participant, age 20, no health insurance, income is \$5000 annually

If $Z_i = 1$ (patient never goes to the doctor)

$\Rightarrow$ average # visits = 0 $\qquad P(Z_i = 1) = \alpha_i$

If $Z_i = 0$ (sometimes go to the doctor)

$\Rightarrow$ average # visits = λ_i $\qquad P(Z_i = 0) = 1 - \alpha_i$

So expected # visits (w/out knowing Z_i) = $\lambda_i(1-\alpha_i)$

$$\log\left(\frac{\hat{\alpha}_i}{1-\hat{\alpha}_i}\right) = 1.70 - 0.80(0.2) - 0.2(0.5)$$

(age = $\frac{\text{years}}{100}$, income = $\frac{\$}{10{,}000}$)

$\Rightarrow$ solve to get $\hat{\alpha}_i$

$$\log(\hat{\lambda}_i) = 0.17 + 0.37(0.2) - 0.41(0.5)$$

$\Rightarrow$ solve to get $\hat{\lambda}_i$

Question 13: probability that male participant, age 20, no health insurance, \$5000 annually, did not go to the doctor in past 2 weeks.

visits in past 2 weeks

$$P(Y_i = 0) = e^{-\lambda_i}(1-\alpha_i) + \alpha_i$$

$$P(Y_i = y) = \begin{cases} e^{-\lambda_i}(1-\alpha_i) + \alpha_i & y = 0 \\ \dfrac{e^{-\lambda_i}\lambda_i^{y}}{y!}(1-\alpha_i) & y > 0 \end{cases}$$

- Find $\hat{\lambda}_i, \hat{\alpha}_i$ (see previous slide)

$$\hat{P}(Y_i = 0) = e^{-\hat{\lambda}_i}(1-\hat{\alpha}_i) + \hat{\alpha}_i$$

$$P(Y_i = 2) = \frac{e^{-\lambda_i}\lambda_i^{2}}{2!}(1-\alpha_i)$$

$$\log\left(\frac{\alpha_i}{1-\alpha_i}\right) = \gamma_0 + \gamma_1 \ldots$$

$$\log(\lambda_i) = \beta_0 + \beta_1 \ldots$$

<u>if</u> there were an offset, then Poisson part would be

$$\log(\lambda_i) = \beta_0 + \beta_1 \ldots \quad + \underbrace{\log(\ldots)}_{\text{offset}}$$

Question 8: (Poisson vs. quasi-Poisson)

Same:
- $\log(\lambda_i) = \beta_0 + \beta_1 \ldots$ (same systematic component)
- estimated coefficients $\hat{\beta}_0, \hat{\beta}_1$, etc. are the same

Different: $SE_{\text{quasi-Poisson}} = \sqrt{\hat{\phi}}\, SE_{\text{Poisson}}$