HW2 Answer Key

Jan 28 2016

Problem 1

For $y \geq 0$, rewrite the pdf as

$$f(y; \lambda) = \lambda \exp(-\lambda y) = \exp(-\lambda y + \log \lambda)$$

so, it belongs to canonical exponential family, with $b(\lambda) = -\lambda$, $c(\lambda) = \log \lambda$ and $d(y) = 0$. Then,

$$\mathrm{E}[Y] = -\frac{c'(\lambda)}{b'(\lambda)} = -\frac{1/\lambda}{-1} = \frac{1}{\lambda}$$

$$\mathrm{Var}[Y] = -\frac{b''(\lambda)\mathrm{E}[Y] + c''(\lambda)}{b'(\lambda)^2} = -\frac{-(1/\lambda^2)}{(-1)^2} = \frac{1}{\lambda^2}$$

3 for rewriting pdf, 3 for each part of b(), c() and d(), 2 for mean and 2 for variance. You are ok if you calculate the mean and variance in different way but get the same results.

Problem 2

The log-likelihood is

$$l = \sum_{i=1}^{N} \log f(y_i; k) = \sum_{i=1}^{N} [\log k + (k-1)\log y_i] = N \log k + (k-1)\sum_{i=1}^{N} \log y_i$$

Taking derivative with respect to k and we get the score function

$$U = \frac{dl}{dk} = \frac{N}{k} + \sum_{i=1}^{N} \log y_i$$

Let $U = 0$ we can solve for the value of k that maximizes this log-likelihood

$$k = -\frac{N}{\sum_{i=1}^{N} \log y_i}$$

4 for log likelihood, 2 for score (you may also use FOC) and 4 for solving k

Problem 3

From Problem 2,

$$U(k) = \frac{N}{k} + \sum_{i=1}^{N} \log y_i$$

and thus

$$U'(k) = -\frac{N}{k^2}$$

Hence the Newton-Raphson iteration equation is

$$k_{t+1} = k_t - \frac{U(k_t)}{U'(k_t)} = k_t - \frac{\frac{N}{k_t} + \sum_{i=1}^{N} \log y_i}{-\frac{N}{k_t^2}} = 2k_t + \frac{k_t^2}{N}\sum_{i=1}^{N} \log y_i$$

When the initial value $k_1 = 1$, the first step of iteration will be

$$k_2 = 2k_1 + \frac{k_1^2}{N} \sum_{i=1}^{N} \log y_i = 2 + \frac{1}{N} \sum_{i=1}^{N} \log y_i$$

2 each for U, U' and NR equation, 4 for result

Problem 4

For logit regression,

$$\mathrm{E}\left[Y|X\right] = \frac{e^{X'\beta}}{1 + e^{X'\beta}}$$

A male respondent without a college degree has the predicted probability of unemployment

$$\widehat{unem} = \frac{e^{-2.09414}}{1 + e^{-2.09414}} = 10.97\%$$

A female respondent with a college degree has the predicted probability of unemployment

$$\widehat{unem} = \frac{e^{-2.09414-0.03722-0.92560}}{1 + e^{-2.09414-0.03722-0.92560}} = 4.49\%$$

Therefore, a female respondent with a college degree has the predicted probability of being unemployed 6.48 percentage points less than a male respondent without a college degree.[1]

4 for showing procedure and 3 for each result

Problem 5

In `normalreg`, the link function is $g(\mu) = \mu$, which means $\mathrm{E}\left[Y|X\right] = X'\beta$. So, the marginal effect on *age* is

$$\frac{dY}{dage} = \beta_{age}$$

One more year in age yields a 0.09668 unit (hour) increase in working hours.

In `poissonreg`, the link function is $g(\mu) = \log \mu$, which means $\mathrm{E}\left[Y|X\right] = \exp(X'\beta)$. So, the marginal effect on *age* is

$$\frac{dY}{dage} = \exp(X'\beta)\beta_{age} = \beta_{age} Y$$

One more year in age yields a 0.28913 percentage increase in predicted hours worked.

4 for each result and 2 for showing equations. AT LEAST you need to mention the log as link function

Problem 6

Let $\mathbf{x}$ be the independent variables and $\pi_j \equiv \text{Prob}(y = j)$ be the outcome. Assume that the log odds ratio relative to the reference group (without losing generality, let group 1 be the reference group) is a linear function of independent variables, i.e.,

$$\log\left(\frac{\pi_j}{\pi_1}\right) = \mathbf{x}'\beta_j$$

for $j = 2, 3, \ldots, J$. Rewrite this equation and we get

$$\pi_j = \pi_1 e^{\mathbf{x}'\beta_j}$$

Note that the probability of outcomes sum up to 1, that is,

$$\sum_{j=1}^{J} \pi_j = 1$$

[1]Notes: You don't need to give the numbers. The second last step is good for full credit.

Replacing π_j with $\pi_1 e^{\mathbf{x}'\beta_j}$, we can solve for π_1

$$\pi_1 = \frac{1}{1 + \sum_{j=2}^{J} e^{\mathbf{x}'\beta_j}}$$

and for $j = 2, 3, \ldots, J$,

$$\pi_j = \pi_1 e^{\mathbf{x}'\beta_j} = \frac{e^{\mathbf{x}'\beta_j}}{1 + \sum_{j=2}^{J} e^{\mathbf{x}'\beta_j}}$$

5 for the log odds ratio, 1 for probability summing up and 4 for result. You are ok if you go with different subscript.