Universidad de Talca, FEN

Discrete Choice Models: Problem Set 1

Professor: Mauricio Sarrias

2020

1 Theory: Maximum Likelihood Estimator

1. Assume that you have a sample of n independent observations from a Poisson distribution with density function.:

$$f(y;\lambda) = \frac{\exp(-\lambda)\,\lambda^y}{y!} \quad y = 0, 1, 2, ...$$

   where $\mathbb{E}(y) = \text{Var}(y) = \lambda$.

   a) (1 pt) Find the ML estimator λ.
   b) (1 pt) Show that the expectation of the score is zero.
   c) (1 pt) Derive the Hessian and the information matrix.
   d) (2 pt) Does the information equality hold?
   e) (2 pt) Derive the variance of the ML estimator.

2. Suppose that $y|x$ is Poisson distributed with parameter $\lambda(x)$ in a sample of n independent observations. Furthermore, assume that $\lambda(x)$ is specified as $\lambda(x;\alpha,\beta) = \alpha+\beta x$. (Hint: what is the expected value and variance of the Poisson distribution?)

   a) (2 pt) Does this specification make sense?
   b) (2 pt) Derive the score function.

3. Consider ML estimation of the parameter α of the Pareto distribution with density function:

$$f(y;\alpha) = \alpha / y^{\alpha+1} \quad \alpha > 0, y \geq 0$$

a) (1 pt) Find the ML estimator of α.

b) (3 pt) Find the LR, Wald and Score test statistics for $H_0 : \alpha = \alpha_0$ against $H_1 : \alpha \neq \alpha_0$.

4. Consider ML estimation of the parameter θ with density function:

$$f(y;\theta) = \theta y^{\theta-1} \quad \theta > 0, 0 < y < 1$$

a) (1 pt) Find the ML estimator of θ.

b) (2 pt) Find the asymptotic distribution of the ML estimator.

5. (Trinity for linear regression model) Consider the linear regression model with normal errors, whose conditional density for observation i is

$$\log f(y_i|\boldsymbol{x}_i;\boldsymbol{\beta},\sigma^2) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - \frac{(y_i - \boldsymbol{x}_i^\intercal \boldsymbol{\beta})^2}{2\sigma^2}.$$

Let $(\widehat{\boldsymbol{\beta}}, \widehat{\sigma}^2)$ be the unrestricted ML estimate of $\boldsymbol{\theta} = (\boldsymbol{\beta}^\intercal, \sigma^2)^\intercal$ and let $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\beta}}^\intercal, \widetilde{\sigma}^2)^\intercal$ be the restricted ML estimate subject to the constraint $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{c}$, where $\boldsymbol{R}$ is an $r \times K$ matrix of known constant. Assume that $\boldsymbol{\Theta} = \mathbb{R}^K \times \mathbb{R}_{++}$ and that $\mathbb{E}(\boldsymbol{x}_i \boldsymbol{x}_i^\intercal)$ is nonsingular. Also, let

$$\widehat{\boldsymbol{H}} = \begin{pmatrix} \frac{1}{\widehat{\sigma}^2}\frac{1}{n}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i^\intercal & \boldsymbol{0} \\ \boldsymbol{0}^\intercal & \frac{1}{2(\widehat{\sigma}^2)^2} \end{pmatrix}; \; \widetilde{\boldsymbol{H}} = \begin{pmatrix} \frac{1}{\widetilde{\sigma}^2}\frac{1}{n}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i^\intercal & \boldsymbol{0} \\ \boldsymbol{0}^\intercal & \frac{1}{2(\widetilde{\sigma}^2)^2} \end{pmatrix}$$

a) (2 pt) Verify that $\widehat{\boldsymbol{\beta}}$ minimizes the sum of squared residuals. So it is the OLS estimator. Verify that $\widetilde{\boldsymbol{\beta}}$ minimizes the sum of squared residuals subject to the constraint $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{c}$. So it is the restricted least squares estimators.

b) (2 pt) Let $Q_n(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^b \log f(y_i|\boldsymbol{x}_i;\boldsymbol{\beta},\sigma^2)$. Show that

$$Q_n(\widehat{\boldsymbol{\theta}}) = -\frac{1}{2}\log(2\pi) - \frac{1}{2} - \frac{1}{2}\log\left(\frac{SSR_U}{n}\right),$$
$$Q_n(\widetilde{\boldsymbol{\theta}}) = -\frac{1}{2}\log(2\pi) - \frac{1}{2} - \frac{1}{2}\log\left(\frac{SSR_R}{n}\right),$$

where $SSR_U \equiv \sum_i (y_i - \boldsymbol{x}_i^\intercal \widehat{\boldsymbol{\beta}})^2$ is the unrestricted sum of squared residuals and $SSR_R \equiv \sum_i (y_i - \boldsymbol{x}_i^\intercal \widetilde{\boldsymbol{\beta}})^2$ is the restricted sum of squared residuals.

c) (2 pt) Verify that the $\widehat{\boldsymbol{H}}$ given here, although not the same as $-\frac{1}{n}\sum_{i=1}^n \boldsymbol{H}(\boldsymbol{w}_i;\widehat{\boldsymbol{\theta}})$, is consistent for $-\mathbb{E}\left[\boldsymbol{H}(\boldsymbol{w}_i;\boldsymbol{\theta}_0)\right]$. Verify that the $\widetilde{\boldsymbol{H}}$ given here, although not the same as $-\frac{1}{n}\sum_{i=1}^n \boldsymbol{H}(\boldsymbol{w}_i;\widetilde{\boldsymbol{\theta}})$, is consistent for $-\mathbb{E}\left[\boldsymbol{H}(\boldsymbol{w}_i;\boldsymbol{\theta}_0)\right]$.

d) (6 pt) Show that the Wald, LM, and LR statistics using $\widehat{\boldsymbol{H}}$ and $\widetilde{\boldsymbol{H}}$ given here can be written as:

$$
\begin{aligned}
W &= n \cdot \frac{\left(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{c}\right)^{\intercal} \left[\boldsymbol{R}\left(\boldsymbol{X}^{\intercal}\boldsymbol{X}\right)^{-1}\boldsymbol{R}^{\intercal}\right]^{-1} \left(\boldsymbol{R}\widehat{\boldsymbol{\beta}} - \boldsymbol{c}\right)}{SSR_U} \\
LM &= n \cdot \frac{\left(\boldsymbol{y} - \boldsymbol{X}\widetilde{\boldsymbol{\beta}}\right)^{\intercal} \boldsymbol{P} \left(\boldsymbol{y} - \boldsymbol{X}\widetilde{\boldsymbol{\beta}}\right)}{SSR_R} \\
LR &= n \cdot \left[\log\left(\frac{SSR_R}{n}\right) - \log\left(\frac{SSR_U}{n}\right)\right]
\end{aligned}
$$

where $\boldsymbol{P} = \boldsymbol{X}\left(\boldsymbol{X}^{\intercal}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\intercal}$.