MATH/STAT 355: Problem Set 2

Prof. Taylor Okonek

Due: February 23, 2024

Maximum Likelihood Estimation

On your last Problem Set, you found the MLEs for β_0 and β_1 in Simple Linear Regression (did you realize this is what we were doing?!). Now we'll do Multiple Linear Regression, and more generally, derive the Normal Equations (it's easier than doing all the $\hat{\beta}_i$'s separately, so long as we remember linear algebra).

Let $\mathbf{Y}$ be a vector of dimension n, $\mathbf{X}$ be a matrix of dimension $n \times p$, $\boldsymbol{\beta}$ be a vector of unknown parameters of dimension p, and $\boldsymbol{\epsilon}$ be a random vector of dimension n such that

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I})$$

where $\sigma^2 > 0$ is known. Show that the MLE for β is given by

$$\hat{\boldsymbol{\beta}}_{MLE} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}$$

Method of Moments

1. Suppose $Y_1, \dots, Y_n \overset{iid}{\sim}$ Exponential(β), where $f_Y(y \mid \beta) = \beta e^{-\beta y}, \quad y \geq 0, \beta > 0$. Find the MOM estimator for β.

2. Suppose $Y_1, \dots, Y_n \overset{iid}{\sim} f_Y(y \mid \theta)$, where

$$f_Y(y \mid \theta) = (\theta^2 + \theta) y^{\theta - 1}(1 - y), \quad 0 \leq y \leq 1.$$

   Find the MOM estimator of θ.

3. Suppose $Y_1, \dots, Y_n \overset{iid}{\sim}$ Normal(μ, σ^2). Derive MOM estimates for μ and σ^2.

4. Suppose $X_1, \dots, X_n \overset{iid}{\sim}$ Negative Binomial(r, p), where $p_X(x \mid r, p) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, \dots$. Find the MOM estimates for r and p.

5. Suppose that $y_1 = 8.3$, $y_2 = 4.9$, $y_3 = 2.6$, and $y_4 = 6.5$ is a random sample of size 4 from the two parameter Uniform pdf,

$$f_Y(y \mid \theta_1, \theta_2) = \frac{1}{2\theta_2}, \quad \theta_1 - \theta_2 \leq y \leq \theta_1 + \theta_2.$$

   Use the method of moments to estimate θ_1 and θ_2 using the observed data.

Properties of Estimators

1. Let $X_1, X_2, \dots, X_n$ denote the outcome of a series of n independent trials, where $X_i = 1$ with probability p and $X_i = 0$ with probability $1 - p$, for $i = 1, 2, \dots, n$. Let $X = X_1 + X_2 + \cdots + X_n$.
   (a) Show that $\hat{p}_1 = X_1$ and $\hat{p}_2 = \frac{X}{n}$ are unbiased estimators for p.
   (b) Intuitively, $\hat{p}_2$ is a better estimator than $\hat{p}_1$ because $\hat{p}_1$ fails to include any of the information about the parameter contained in trials 2 through n. Verify that speculation by comparing the variances of $\hat{p}_1$ and $\hat{p}_2$.
2. Prove that $E[(\hat{\theta} - \theta)^2] = \text{Var}(\hat{\theta}) + \text{Bias}(\hat{\theta})^2$ (aka, that the two ways to write MSE are equivalent).
3. Suppose W is a random variable with $E(W) = \mu$ and $V(W) = \sigma^2$. Show that $\overline{W}^2 = \left(\frac{1}{n}\sum_{i=1}^n W_i\right)^2$ is an *asymptotically* unbiased estimator for μ^2.
4. Show that the two definitions of the information matrix are equivalent, i.e. that
$$E\left[\left(\frac{\partial}{\partial\theta}\log L(\theta \mid x)\right)^2\right] = -E\left[\frac{\partial^2}{\partial\theta^2}\log L(\theta \mid x)\right]$$
5. Let $X_1, \dots, X_n \overset{iid}{\sim} Poisson(\lambda)$. Show that $\hat{\lambda} = \frac{1}{n}\sum_{i=1}^n X_i$ is an *efficient* estimator for λ.