MA677_HW5
Sky Liu
MLE Problem

1. To estimate the proportion $p$ of butterflies that have a special marking on their wings. Consider two approaches:

   (a) Capture butterflies one at a time until five with the special marking have been collected. A total of 43 butterflies are required to collect the five. what is the M.L.E. of p?

   (b) Collect butterflies all day and count those with the special mark. 58 are captured. Three have the Mark. What is the M.L.E. of p?

(a) $\hat{p} = \frac{5}{43}$

(b) $\hat{p} = \frac{3}{58}$

2. Consider a random sample $X_1, \cdots, X_n \sim Uniform(0, \theta), \theta$ unknown. Show that the sequence of MLE's of $\theta$ is a consistent sequence.

$X_1 \ldots X_n \sim U(0, \theta)$

$f(x) = \frac{1}{\theta}$

$\mathcal{L}(\theta, X_1 \ldots X_n) = f(x_1) \ldots f(x_n) = \frac{1}{\theta^n} = \theta^{-n}, \quad \theta \geq \max(X_n)$

$= \ell \; -n \log(\theta)$

$\frac{d\ell}{d\theta} = -\frac{n}{\theta}$

The sequence of MLE of $\theta$ is a function with no x, therefore it's a constant sequence.

3. Consider a random sample $X_1, \cdots, X_n \sim N(\mu, \sigma^2), \mu, \sigma$ unknown. Find the MLE of the 0.95 quantile.

$X_1 \dots X_n \sim N(\mu, \sigma^2)$ . $\mu$ is known. $\sigma^2$ is unknown

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\mathcal{L}(\mu, \sigma; x_1 \dots x_n) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2\right)}$$

$$= \ell - n\log(\sigma) - \frac{n}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{d\mathcal{L}(\mu, \sigma; x_1 \dots x_n)}{d\sigma} = -\frac{n}{\sigma} + \sigma^{-3}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$95\%\ \text{MLE} = \hat{\mu} + 1.64\,\hat{\sigma}$$

$$= \bar{X} + 1.64 \cdot \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

4. Consider again a random sample $X_1, \cdots, X_n \sim N(\mu, \sigma^2), \mu, \sigma$ unknown. Find the MLE of $v = P(X > 2)$.

$X_1 \dots X_n \sim N(\mu, \sigma^2)$

MLE of $v = P(X > 2)$

$$v = P\left(Z > \frac{2-\mu}{\sigma}\right)$$

$$= 1 - \Phi\,\frac{2-\mu}{\sigma}$$

5. Find the MLE estimator for $\theta$ in the Cauchy distribution given the sample X = (-22.33, -10.29, -1.35, -1.73, 6.91, -0.52, 0.43, -0.00, -8.66, -7.16, 1.15, 1.15, -3.75, 2.54, 7.31, 0.65, 6.66, 5.52, 2.02, -1.48).

$X \sim \text{Cauchy}(\theta, 1)$

$$f(x) = \frac{1}{\pi(1+(x-\theta)^2}$$

$$\mathcal{L}(\theta, 1; X) = \frac{1}{\pi^n \cdot \prod_{i=1}^{n} (1+(x_i-\theta)^2}$$

$$= \frac{1}{\pi^n} - \frac{1}{(1+(x-\theta)^2)^n}$$

$$\ell \; -n\log(\pi) - \sum_{i=1}^{n} n\log(1+(x-\theta)^2)$$

$$\frac{d\ell}{d\theta} = \sum_{i=1}^{n} \frac{2(x_i-\theta)}{1+(x_i-\theta)^2} = 0$$

$$2\theta(\theta^2+(1-x^2)) = 0$$

$$\hat{\theta} = \pm\sqrt{x^2-1}$$

Given $\bar{x} = -1.15$

$$\hat{\theta} = \pm\sqrt{0.3225} \approx \pm 0.57$$

6. Consider a random sample of 21 observations from exponential($\lambda$). Mean ($\mu > 0$). 20 of the observations are collected without incident and have a mean of 6. The 21st observation was not measured exactly except that it is greater than 15.

   Find the MLE of $\mu$.

$X_1 \dots X_n \sim \exp(\lambda)$

$f(x) = \lambda e^{-\lambda x}$

$F(x) = 1 - e^{-\lambda x}$

$F(x > 15) = e^{-15\lambda}$

$L(\mu, X_{20}) = \frac{1}{\mu^{20}} \exp\left(\frac{\sum^{n} x}{\mu}\right) \exp\left(\frac{-15}{\mu}\right)$

$= \frac{1}{\mu^{20}} \exp\left(\frac{-135}{\mu}\right)$

$\ell \quad 20 \log(\mu) - 135\mu$

$\frac{d\ell}{d\mu} = \frac{20}{\mu} - 135 = 0$

$\hat{\mu} = 6.75$

7. $X_1, \cdots, X_n$ form a random sample from a Poisson distribution for which the mean is unknown. Determine the MLE of the standard deviation of the distribution.

$X_1 \dots X_n \sim \text{Poisson}(\lambda)$

$f(x) = \frac{e^{-\lambda} \lambda^x}{x!}$

$L(\lambda; x_1 \dots x_n) = e^{-\lambda n} \lambda^{\sum x_i} \left(\prod_{i=1}^{n} \frac{1}{x_i}\right)$

$\ell \quad -n\lambda + \sum_{i=1}^{n} x_i \log(\lambda) + \log\left(\prod_{i=1}^{n} \frac{1}{x_i}\right)$

$\frac{d\ell}{d\lambda} = -n + \frac{\sum x_i}{\lambda} = 0$

$\hat{\lambda} = \frac{\sum x_i}{n} = \bar{x} \Rightarrow \hat{sd} = \sqrt{\hat{\lambda}} = \sqrt{\bar{x}}$

8. Consider a random sample $X_1, \cdots, X_n \sim exp(\beta), \beta$ is unknown. Determine the MLE of the median of the distribution.

$$X_1 \cdots X_n \sim exp(\beta)$$

$$f(x) = \lambda e^{-\lambda x}$$

$$\mathcal{L}(\beta, X_1 \ldots X_n) = \lambda^n e^{-\sum_1^n x_i \lambda}$$

$$\ell \quad n \log(\lambda) - \lambda \sum_1^n x_i$$

$$\frac{d\ell}{d\lambda} = \frac{n}{\lambda} - \sum_1^n x_i = 0$$

$$\hat{\lambda} = \frac{n}{\sum x_i} = \frac{1}{u} = \frac{1}{\bar{x}}$$

$$\text{median} = \lambda^{-1} \log(2) = \bar{x} \log(2)$$

Sufficiency Statistics

For each of these distributions show that the specified statistic T is suficient for the parameter.

1. The Bernoulli distribution with parameter $p$.
   $(0 < p < 1)$, $T = \sum_{i=i} X_i$.

$$f(P, X_1 \dots X_n) = p^{\Sigma x_i} (1-p)^{n^2 - \Sigma x_i}$$

$$U(X_1 \dots X_n) = 1, \quad V(T, P) = P^T (1-p)^{n^2 - T}$$

$$f(P, X_1 \dots X_n) = U(X_1 \dots X_n) \cdot V(T, P)$$

$T = \Sigma X_i$ is sufficient

2. The geometric distribution with parameter $p$.
   $(0 < p < 1)$, $T = \sum_{i=i} X_i$.

$$f(P, X_1 \dots X_n) = P^n (1-P)^{\Sigma X_i}$$

$$U(X_1 \dots X_n) = 1, \quad V(T, P) = P^n (1-P)^T$$

$T = \Sigma X_i$ is sufficient.

3. The negative binomial distribution with parameters $r$ and $p$.
   r is known. $(0 < p < 1)$, $T = \sum_{i=i} X_i$.

$$f(r, p, X_1 \dots X_n) = \prod_{i=1}^{n} C_{X_i + r - 1}^{r-1} \, p^r (1-p)^{X_i}$$

$$= \prod_{i=1}^{n} C_{X_i + r - 1}^{r-1} \, P^{nr} (1-P)^{\Sigma X_i}$$

$$U(X_1 \dots X_n) = \prod_{i=1}^{n} C_{X_i + r - 1}^{r-1}, \quad V(T, P) = P^{nr} (1-P)^{\Sigma X_i}$$

$T = \Sigma X_i$ is sufficient

4. The gamma distribution with parameters $\alpha$ and $\beta$. $\alpha$ is known. $(\beta > 0)$, $T = \sum_{i=i} X_i$.

$$f(\alpha, \beta, x_1 \ldots x_n) = \prod_{i=1}^{n} \frac{\beta^{\alpha}}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i}$$

$$= \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \left(\prod_{i=1}^{n} x_i^{\alpha-1}\right) e^{-\beta \Sigma x_i}$$

$$U(x_1 \ldots x_n) = \prod_{i=1}^{n} x_i^{\alpha-1}, \quad V(T, \beta) = \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} e^{-\beta T}$$

$T = \Sigma x_i$ is sufficient

5. The gamma distribution with parameters $\alpha$ and $\beta$. $\beta$ is known. $(\alpha > 0)$, $T = \prod_{i=i} X_i$.

$$f(\alpha, \beta, x_1 \ldots x_n) = \prod_{i=1}^{n} \frac{\beta^{\alpha}}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\beta x_i}$$

$$= \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \left(\prod_{i=1}^{n} x_i^{\alpha-1}\right) e^{-\beta \Sigma x_i}$$

$$= \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \exp(\alpha-1) \sum_{i=1}^{n} \ln(x_i)\, e^{-\beta \Sigma x_i}$$

$$U(x_1 \ldots x_n) = e^{\beta \Sigma x_i}, \quad v(T, \alpha) = \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \exp(\alpha-1) \sum_{i}^{n} \ln(T)$$

$T = \Sigma x_i$ is sufficient.