Statistics 2 Unit 1

Team 8

April 24, 2018

Contents

1 Task PI64

Let $X_1, ..., X_n$ be iid from a geometric distribution, i.e. $f(x) = p(1-p)^x,\ x = 0, 1, ...$.

1.1 a)

The parameter by the MME is

$$\hat{\mu_1} = \bar{X} = \frac{1-\hat{p}}{\hat{p}} \iff \hat{p} = \frac{1}{\bar{X}+1}.$$

1.2 b)

The parameter by the MLE is

$$l(p, x_i) = \prod_{i=0}^{n} p(1-p)^{x_i} = p^n (1-p)^{\sum x_i}$$

$$\frac{\partial log(l(x))}{\partial p} = \frac{1}{p} + \frac{x}{1-p} \stackrel{!}{=} 0$$

$$\hat{p} = \frac{1}{\bar{X}+1}.$$

1.3 c)

The Fisher information was already calculated in the exercise 62.

$$I(p) = \frac{1}{p^2(1-p)}$$

The asymptotic variance of the MLE is given by

$$\frac{1}{nI(p)} = n\ p^2\ (1-p).$$

1.4 d)

The posterior is defined as $f(p|x_i) \propto l(p, x_i)\ f(p)$.
So we get for the uniform distribution on $[0, 1]$ the following posterior:

$$f(p|x_i) \propto p^n (1-p)^{\sum x_i}\ 1 =\propto p^n (1-p)^{\sum x_i}.$$

We see the posterior distribution can be written as a beta distribution with parameters $(\alpha = 2, \beta = \sum x_i + 1)$.
The posterior mean can be then found by using the mean of a beta:

$$X = \frac{\alpha}{\alpha + \beta} = \frac{2}{\sum x_i + 3}.$$

2 Task PI66:

2.1 a)

Writing $X = (X_1, \ldots, X_n)$, the log-likelihood function is

$$\begin{aligned} \ell(\tau) &= \log \prod_{i=1}^{n} \frac{1}{\tau} e^{-x_i/\tau} \\ &= -n \log(\tau) - \frac{1}{\tau} \sum_{i=1}^{n} x_i \end{aligned}$$

Now, take the derivative wrt to τ

$$\ell'(\tau) = -\frac{n}{\tau} + \frac{1}{\tau^2} \sum_{i=1}^{n} x_i$$

Solving for $l'(\tau) = 0$, the MLE of τ is

$$\widehat{\tau} = \overline{X}$$

2.2 b)

Let $S = X_1 + \ldots + X_n \sim \Gamma\left(n, \frac{1}{\tau}\right)$
Thus, the PDF of $\overline{X} = \frac{S}{n}$ is:

$$f_{\overline{X}}(x) = \frac{s^{n-1}}{\tau^n \Gamma(n)} e^{-s/\tau} \left|\frac{ds}{dx}\right| = \frac{n^n x^{n-1}}{\tau^n \Gamma(n)} e^{-nx/\tau}, x > 0,$$

which is the PDF of the $\Gamma(n, n/\tau)$ distribution.

2.3 c) and d)

Since $\overline{X} \sim \Gamma(n, n/\tau)$, we have

$$E\left(\overline{X}\right) = \tau$$

$$Var\left(\overline{X}\right) = \frac{\tau^2}{n}$$

From the CLT it follows that $\dfrac{\overline{X} - \tau}{\sqrt{\dfrac{\tau^2}{n}}}$ is approximately distirbuted as $N(0, 1)$ for large n.

2.4 e)

The Cramer-Rao lower bound is $1/[nI(\tau)]$, where

$$I(\tau) = -E\left[\frac{\partial^2}{\partial \tau^2} \log\left(\frac{1}{\tau} e^{-\frac{X_1}{\tau}}\right)\right] = -\frac{1}{\tau^2} - E\left(\frac{2X_1}{\tau^3}\right) = \frac{1}{\tau^2}$$

since $E(X_1) = \tau$. This implies that the Cramer-Rao lower bound is

$$[nI(\tau)]^{-1} = \frac{\tau^2}{n}$$

This lower bound equals the variance of $\overline{X}$.
Thus, we conclude that there is no other unbiased estimate of τ with a smaller variance than $\overline{X}$.

2.5 f)

From part (c), we have $\dfrac{\overline{X} - \tau}{\sqrt{\dfrac{\tau^2}{n}}}$ is approximately distirbuted as $N(0, 1)$ for large n.
Thus, an approximate $100(1 - \alpha)\,\%$ CI for τ is:

$$\overline{X} \pm z_{1-\alpha/2} \frac{\tau}{\sqrt{n}} \approx \overline{X} \pm z_{1-\alpha/2} \frac{\overline{X}}{\sqrt{n}}$$

or equivalently the set of τ's satisfying

$$\tau - z_{1-\alpha/2} \frac{\tau}{\sqrt{n}} \leq \overline{X} \leq \tau + z_{1-\alpha/2} \frac{\tau}{\sqrt{n}}$$

2.6 g)

Note that $\overline{X}$ has exactly the $\Gamma\left(n, n/\tau\right)$ distribution.
Let $G_\tau\left(\alpha\right)$ denote the 100α percentile of $\Gamma\left(n, n/\tau\right)$ distribution, i.e.

$$P\left(\overline{X} \leq C_\tau\left(\alpha\right)\right) = \alpha$$

Then an exact $100\left(1-\alpha\right)\%$ CI for τ is given by the set of τ's satisfying

$$G_\tau\left(\alpha/2\right) \leq \overline{X} \leq G_\tau\left(1-\alpha/2\right)$$