COLUMBIA UNIVERSITY | Prof. Marco Avella
Statistical Inference | TA: George Chu, Joe Suk
Final solutions | Summer 2020

---

1. We have $\mathbb{P}(X_1 = 0) = e^{-\lambda}$ so that since the MLE of λ is $\overline{X}_n$, the MLE of $\mathbb{P}(X_1 = 0)$ is $\delta := e^{-\overline{X}_n}$. Then, by CLT and the Delta Method, we have

$$\sqrt{n}(e^{-\overline{X}_n} - e^{-\lambda}) \xrightarrow{d} N(0, \lambda \cdot e^{-2\lambda}).$$

By Slutsky and continuous mapping theorem, we have

$$\frac{\sqrt{n}(e^{-\overline{X}_n} - e^{-\lambda})}{\sqrt{\overline{X}_n} \cdot e^{-\overline{X}_n}} \xrightarrow{d} N(0, 1).$$

It is also possible to arrive at this via asymptotic normality of the MLE and a Fisher's information computation. Thus, our approximate C.I. is

$$e^{-\overline{X}_n} \pm z_{1-\alpha/2} \cdot \frac{\sqrt{\overline{X}_n} \cdot e^{-\overline{X}_n}}{\sqrt{n}}.$$

A second solution is to convert an approximate $1-\alpha$-C.I. for λ (obtained via CLT) into a $1-\alpha$-C.I. for $e^{-\lambda}$ via the monotone transformation $x \mapsto e^{-x}$:

$$\exp\left(-\overline{X}_n \pm z_{1-\alpha/2}\sqrt{\frac{\overline{X}_n}{n}}\right).$$

2. The log-likelihood is

$$\log L(\theta | X_1, X_2, X_3) = -3\log(\theta) - \frac{1}{\theta}\sum_{k=1}^{3}\frac{X_k}{k} + \text{const.}$$

so that

$$\frac{\partial}{\partial\theta}\log L(\theta | X_1, X_2, X_3) = -\frac{3}{\theta} + \frac{1}{\theta^2}\sum_{k=1}^{3}\frac{X_k}{k} = 0 \implies \hat{\theta}_{\text{MLE}} = \frac{1}{3}\sum_{k=1}^{3}\frac{X_k}{k}.$$

Next, the posterior of ψ has pdf

$$\begin{aligned}\pi(\psi | X_1, X_2, X_3) &\propto f(X_1, X_2, X_3 | \psi)\pi(\psi) \\ &\propto \psi^3 e^{-\psi\sum_{k=1}^{3}\frac{X_k}{k}}\beta^\alpha\psi^{\alpha-1}e^{-\beta\psi} \\ &\propto \psi^{(\alpha+3)-1}e^{-\psi(\beta+\sum_{k=1}^{3}X_k/k)}.\end{aligned}$$

Thus, $\psi | X_1, X_2, X_3 \sim \text{Gamma}(\alpha + 3, \beta + \sum_{k=1}^{3} X_k/k)$.

3. The MOM estimator using the first moment is $\hat{p}_1 = \overline{N}_n^{-1}$ and, for the second moment estimator, we solve the quadratic equation

$$\frac{1-p}{p^2} + \frac{1}{p^2} = \frac{2-p}{p^2} = \frac{1}{n}\sum_{i=1}^{n} N_i^2 \implies \hat{p}_2 = \frac{-1 + \sqrt{1 + \frac{8}{n}\sum_{i=1}^{n} N_i^2}}{\frac{2}{n}\sum_{i=1}^{n} N_i^2}.$$

Next, the joint likelihood is

$$f(N_1, \dots, N_n) = p^n (1-p)^{\sum_{i=1}^{n} N_i - n},$$

so that $\sum_{i=1}^{n} N_i$ is sufficient for p. Thus, we would expect $\hat{p}_1$ to be better based on sufficiency. Furthermore, it is straightforward to show $\hat{p}_1$ is also the MLE of p. Thus, it is asymptotically efficient.

4. The MLE of μ and σ are respectively $\bar{X}$ and $\sqrt{\frac{n-1}{n}}s$. Also, $\sqrt{n}\frac{\bar{X}-\mu}{s} \sim t_{n-1}$, so we have

$$P\left(\sqrt{n}\frac{\hat{\mu}-\mu}{\sqrt{\frac{n}{n-1}}\hat{\sigma}} > t_{n-1,0.05}\right) = 0.95$$

$$\implies P\left(\hat{\mu} > \mu + \frac{t_{n-1,0.05}}{\sqrt{n-1}}\hat{\sigma}\right) = 0.95.$$

It follows that $k = \frac{t_{n-1,0.05}}{\sqrt{n-1}}$. For $n = 17$, we have $k = \frac{-1.746}{\sqrt{16}} = -0.4365$.

5. $\pi(\theta) = P(X \leq 5) = 1 - P(X \geq 6) = 1 - (1-\theta)^5$.
or $\theta + \theta(1-\theta) + \theta(1-\theta)^2 + \theta(1-\theta)^3 + \theta(1-\theta)^4$.

6. (a) It corresponds to the t-test of $H_0 : \beta_1 = 0$ vs $H_0 : \beta_1 \neq 0$. The p-value is very close to zero, so we will reject the null, i.e. the data suggests that there is an association between the atmospheric pressure and the boiling point of water.

   (b) The 0.95 prediction CI is given by

$$(\hat{\beta}_0 + \hat{\beta}_1 x) \pm t_{n-2,0.975}\, s\sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\sum(x_i - \bar{x})^2}}$$
$$= 3.1519538 \pm 2.145 \cdot 0.002616 \cdot 1.0382131$$
$$= [3.1461, 3.1578].$$

   (c) Denote the MLE in the restricted case by $\tilde{\beta}_0$ and $\tilde{\sigma}^2$ respectively, then $\tilde{\beta}_0 = \bar{y} = \hat{\beta}_0 + \hat{\beta}_1\bar{x} = 3.2104$, and $\tilde{\sigma}^2 = \frac{1}{n}\sum(y_i - \bar{y})^2 = \frac{1}{n}SST = \frac{1}{n}\frac{SSE}{1-R^2} = \frac{1}{16}\frac{14(0.002616^2)}{1-0.9996} = 0.014970$.

   (d) Note that $\hat{\sigma}^2 = \frac{1}{n}SSE$ and $\tilde{\sigma}^2 = \frac{1}{n}SST$, so the likelihood ratio is

$$\frac{(2\pi\tilde{\sigma}^2)^{-n/2}\exp\{-SST/(2\tilde{\sigma}^2)\}}{(2\pi\hat{\sigma}^2)^{-n/2}\exp\{-SSE/(2\hat{\sigma}^2)\}} = \left(\frac{SSE}{SST}\right)^{\frac{n}{2}} = (1-R^2)^{\frac{n}{2}}.$$