COLUMBIA UNIVERSITY
Statistical Inference
Practice exercises (solutions)

Prof. Marco Avella
TA: George Chu, Joe Suk
Summer 2020

Question 1

Let $X_1, \ldots, X_n$ denote the observations in the random sample, and let α and β denote the parameters of the prior gamma distribution of θ. It was shown in Theorem 7.3.4 that the posterior distribution of θ will be the gamma distribution with parameters $\alpha + n$ and $\beta + n\overline{X}_n$. The Bayes estimator, which is the mean of this posterior distribution is, therefore,

$$\frac{\alpha + n}{\beta + n\overline{X}_n} = \frac{1 + (\alpha/n)}{\overline{X}_n + (\beta/n)}.$$

Since the mean of the exponential distribution is $1/\theta$, it follows from the law of large numbers that $\overline{X}_n$ will converge in probability to $1/\theta$ as $n \to \infty$. It follows, therefore, that the Bayes estimators will converge in probability to θ as $n \to \infty$. Hence, the Bayes estimators form a consistent sequence of estimators of θ.

Question 2

(a) $E(\hat{\theta}) = \alpha E(\overline{X}_m) + (1-\alpha)E(\overline{Y}_n) = \alpha\theta + (1-\alpha)\theta = \theta$. Hence, $\hat{\theta}$ is an unbiased estimator of θ for all values of α, m and n.

(b) Since the two samples are taken independently, $\overline{X}_m$ and $\overline{Y}_n$ are independent. Hence,

$$\operatorname{Var}(\hat{\theta}) = \alpha^2 \operatorname{Var}(\overline{X}_m) + (1-\alpha)^2 \operatorname{Var}(\overline{Y}_n) = \alpha^2 \left(\frac{\sigma_A^2}{m}\right) + (1-\alpha)^2 \left(\frac{\sigma_B^2}{n}\right).$$

Since $\sigma_A^2 = 4\sigma_B^2$, it follows that

$$\operatorname{Var}(\hat{\theta}) = \left[\frac{4\alpha^2}{m} + \frac{(1-\alpha)^2}{n}\right]\sigma_B^2.$$

By differentiating the coefficient of σ_B^2, it is found that $\operatorname{Var}(\hat{\theta})$ is a minimum when $\alpha = m/(m+4n)$.

Question 3

MLE for σ^2 is $\hat{\sigma}^2_{ML} = \frac{1}{n}\sum_{i=1}^n X_i^2$, so by the invariance property, we have $\hat{\theta}_{ML} = \frac{1}{2}\log(\hat{\sigma}^2_{ML}) = \frac{1}{2}\log(\frac{1}{n}\sum_{i=1}^n X_i^2)$.

$$
\begin{aligned}
f(x \mid \sigma) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\} \\
\lambda(x \mid \sigma) &= -\frac{1}{2}\log(2\pi) - \log\sigma - \frac{x^2}{2\sigma^2} \\
\lambda'(x \mid \sigma) &= -\frac{1}{\sigma} + \frac{x^2}{\sigma^3} \\
\lambda''(x \mid \sigma) &= \frac{1}{\sigma^2} - \frac{3x^2}{\sigma^4}
\end{aligned}
$$

Therefore,

$$
I(\sigma) = -E_\sigma\left[\lambda''(X \mid \sigma)\right] = -\frac{1}{\sigma^2} + \frac{3E\left(X^2\right)}{\sigma^4} = -\frac{1}{\sigma^2} + \frac{3}{\sigma^2} = \frac{2}{\sigma^2}.
$$

Recall by chain rule that

$$
E\left(\frac{\partial L}{\partial \sigma}\right)^2 = E\left(\frac{\partial L}{\partial \theta}\right)^2\left(\frac{\partial \theta}{\partial \sigma}\right)^2,
$$

i.e.

$$
I(\sigma) = I(\theta)\left(\frac{\partial \theta}{\partial \sigma}\right)^2.
$$

Now, $\left(\frac{\partial \theta}{\partial \sigma}\right)^2 = \frac{1}{\sigma^2}$, so we have $I(\theta) = 2$ and $I_n(\theta) = 2n$.

Question 4

14. This result follows from previous exercises in two different ways. First, by Exercise 6, the statistic $T' = \prod_{i=1}^n X_i$ is a sufficient statistic. Hence, by Exercise 13, $T = \log\, T'$ is also a sufficient statistic. A second way to establish the same result is to note that, by Exercise 24(g) of Sec. 7.3, the gamma distributions form an exponential family with $d(x) = \log\, x$. Therefore, by Exercise 11, the statistic $T = \sum_{i=1}^n d(X_i)$ is a sufficient statistic.

Question 5

The shortest CI is given by $\bar{X} \pm t_{1-\alpha/2}(n-1)s/\sqrt{n}$, so we have $L^2 = 4t^2_{1-\alpha/2}(n-1)s^2/n$, and thus $E(L^2) = 4t^2_{1-\alpha/2}(n-1)\sigma^2/n$. For $n = 10$ and $\alpha = 0.05$, we have $E(L^2) = 4(2.262)^2\sigma^2/10 = 2.047\sigma^2$.

1. As n increases, $t_{1-\alpha/2}(n-1)$ decreases as well, so $E(L^2)$ decreases.

2. As α increases, $t_{1-\alpha/2}(n-1)$ decreases, so $E(L^2)$ decreases.

Question 6

$$\begin{aligned}
f(x \mid \alpha) &= \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x) \\
\lambda(x \mid \alpha) &= \alpha \log \beta - \log \Gamma(\alpha) + (\alpha - 1)\log x - \beta x \\
\lambda'(x \mid \alpha) &= \log \beta - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \log x \\
\lambda''(x \mid \alpha) &= -\frac{\Gamma(\alpha)\Gamma''(\alpha) - [\Gamma'(\alpha)]^2}{[\Gamma(\alpha)]^2}
\end{aligned}$$

Therefore,

$$I(\alpha) = \frac{\Gamma(\alpha)\Gamma''(\alpha) - [\Gamma'(\alpha)]^2}{[\Gamma(\alpha)]^2}$$

The distribution of the M.L.E. of α will be approximately the normal distribution with mean α and variance $1/[nI(\alpha)]$. Note that we have determined this distribution without actually determining the M.L.E. itself.

Question 7

Exercise 2. 1) Recall $T = \frac{\sqrt{n}(\overline{X}-\mu_0)}{\sqrt{\frac{\sum(X_i-\overline{X})^2}{n-1}}} = \frac{\sqrt{n}\overline{X}}{\sqrt{\frac{\sum X_i^2 - n\overline{X}^2}{n-1}}}$

Pluggling in the values, we get $T = \frac{\sqrt{8}\left[-\frac{11.2}{8}\right]}{\sqrt{\frac{43.7-8\left(\frac{11.2}{8}\right)^2}{7}}} = \frac{\sqrt{8}(-1.4)}{\sqrt{\frac{28.02}{7}}} = -1.97919$

As we want a level $\alpha = 0.05$ test, we need $0.05 = \mathbb{P}_{\mu_0=0}(T \geq k)$ or $0.95 = \mathbb{P}_{\mu_0=0}(T < k)$

As under null, $T \sim t_7$, we get from the table $k = 1.895$

We should reject when $T \geq 1.895$, as we got $T = -1.97919$, we **cannot reject**

2) Recall t-test rejects when $T \geq k$ such that $\alpha = \mathbb{P}_{\text{null}}(T \geq k)$, where

$T = \frac{\sqrt{n}(\overline{X}-\mu_0)}{S}$ and $S = \sqrt{\frac{\sum(X_i-\overline{X})^2}{n-1}}$

Note $\pi(\mu, \sigma^2|\delta) = \mathbb{P}(T \geq c|\mu, \sigma^2)$

If we set $Z = \frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}$ and $W = \frac{\sum(X_i-\overline{X})^2}{\sigma^2}$ then $Z \sim \mathcal{N}(0,1)$, $W \sim \chi^2_{n-1}$ and $Z \perp W$

Note $T = \frac{\sqrt{n}(\overline{X}-\mu_0)}{\sqrt{\frac{\sum(X_i-\overline{X})^2}{n-1}}} = \frac{Z+\sqrt{n}\left(\frac{\mu-\mu_0}{\sigma}\right)}{\sqrt{\frac{W}{n-1}}}$. Hence, if $\frac{\mu_1-\mu_0}{\sigma_1} = \frac{\mu_2-\mu_0}{\sigma_2} \implies T_1 = T_2$

$\implies \pi(\mu_1, \sigma_1^2|\delta) = \pi(\mu_2, \sigma_2^2|\delta)$

Question 8

Exercise 3. Recall $T = \frac{S_X^2/(m-1)\sigma_1^2}{S_Y^2/(n-1)\sigma_2^2} \sim F_{m-1,n-1}$. Note $V = T\left(\frac{\sigma_1^2}{\sigma_2^2}\right)$

Then $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2|\delta) = \mathbb{P}(V \leq c_1) + \mathbb{P}(V \geq c_2) = \mathbb{P}(T \leq c_1\frac{\sigma_2^2}{\sigma_1^2}) + \mathbb{P}(T \geq c_2\frac{\sigma_2^2}{\sigma_1^2}) = G_{m-1,n-1}\left(c_1\frac{\sigma_2^2}{\sigma_1^2}\right) + 1 - G_{m-1,n-1}\left(c_2\frac{\sigma_2^2}{\sigma_1^2}\right)$

Now we find c_1 and c_2 for $\alpha_0 = 0.05$

Using hint, $\mathbb{P}_{\sigma_1=\sigma_2}(V \leq c_1) = \frac{0.05}{2} \iff \mathbb{P}_{\sigma_1=\sigma_2}(T \leq c_1) = 0.025 \iff \mathbb{P}_{\sigma_1=\sigma_2}\left(\frac{1}{T} \geq \frac{1}{c_1}\right) = 0.025 \iff \mathbb{P}_{\sigma_1=\sigma_2}\left(\frac{1}{T} \leq \frac{1}{c_1}\right) = 0.975$.

Recall $\frac{1}{T} \sim F_{20,10} \implies$ from table $\frac{1}{c_1} = 3.42$. Then $\boldsymbol{c_1 = 0.2924}$

For c_2, $\mathbb{P}_{\sigma_1=\sigma_2}(V \geq c_2) = \frac{0.05}{2} \iff \mathbb{P}_{\sigma_1=\sigma_2}(T \geq c_2) = 0.025 \iff \mathbb{P}_{\sigma_1=\sigma_2}(T \leq c_2) = 0.975 \iff G_{10,20}(c_2) = 0.975$. Using table, we get $\boldsymbol{c_2 = 2.77}$

Question 9

Exercise 5. 1)To find $\hat{\beta}$, $(X^T X)^{-1} X^T Y = (0.7668, 2.6704, -0.5319)^T$

2)To find $\hat{\sigma}^2 = \frac{S^2}{n} = \frac{2.32}{20} = 0.116$

3)To perform the test $\beta_1 = 2$ vs $\beta_1 \neq 2$, we need to find $U = \frac{\hat{\beta}_1 - 2}{\sqrt{\zeta_{11}\tilde{\sigma}^2}}$ where $\tilde{\sigma}^2 = \frac{S^2}{n-p}$. Recall that under null $U \sim t_{n-p}$, i.e. $U \sim t_{17}$

$U = \frac{2.6704-2}{\sqrt{\frac{(2.32)(0.52066)}{17}}} = \frac{0.6704}{\sqrt{0.07105}} = \frac{0.6704}{0.2665} = 2.514996$

We should reject if $|U| \geq T_{17}^{-1}\left(1 - \frac{\alpha_0}{2}\right)$. Setting $\alpha_0 = 0.05$ and using the table, we find $T_{17}^{-1}(0.975) = 2.11$. Hence, **we reject** H_0