Practice Final Solution

STA 200B
Winter 2020
University of California, Davis

Exam Rules: This exam is closed book and closed notes. Use of calculators, cell phones or any other communication or electronic devices is not allowed. You must show all of your work to receive credit. You are allowed to use 2 sheets of notes, two-sided. Your exam will be conducted online. The exam will be provided at the same time through piazza to all. You will then display it on your screen. We will open a zoom session in which you need to participate and where you enable your camera on laptop or cell phone to record your test taking, in lieu of the proctoring and monitoring that is happening in an in-class final. As if it was an in-class exam, books and cell phones need to be packed away (except if you use the camera on the cell phone to monitor your exam taking, in which case you need to have the zoom app on your cell phone). With your signature you ascertain that you abide by these rules and the UC Davis Code of Academic Conduct.

Name : ____________________

ID : ____________________

Signature : ____________________

1. Let $X_1, \ldots, X_n$ be a random sample from a normal distribution with unknown mean μ and unknown variance σ^2. Suppose you only observe $\sum_{i=1}^n X_i$ and $\sum_{i=1}^n X_i^2$.

   (a) Derive a γ confidence interval for μ with the shortest length and interpret the meaning of this confidence interval.

   Solution: $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1}(\sum_{i=1}^n X_i^2 - n\bar{X}^2)$.
   Since $\frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}} \sim t_{n-1}$, a γ confidence interval for μ with the shortest length is

   $$\left(\bar{X} - t_{(1+\gamma)/2,n-1}\hat{\sigma}/\sqrt{n}, \bar{X} + t_{(1+\gamma)/2,n-1}\hat{\sigma}/\sqrt{n}\right).$$

   (b) Derive a γ confidence interval for σ^2.

Solution: Since $\frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{\sigma^2} \sim \chi^2_{n-1}$, a γ confidence interval for σ^2 can be derived from

$$\begin{aligned}
\gamma &= P\left(\chi^2_{n-1}(\frac{1-\gamma}{2}) \le \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{\sigma^2} \le \chi^2_{n-1}(\frac{1+\gamma}{2})\right) \\
&= P\left(\frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{\chi^2_{n-1}(\frac{1-\gamma}{2})} \le \sigma^2 \le \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{\chi^2_{n-1}(\frac{1+\gamma}{2})}\right)
\end{aligned}$$

hereafter $\chi^2_m(\alpha)$ is the α quantile of the χ^2 distribution with m degrees of freedom. It follows that a γ confidence interval for σ^2 is

$$\left(\frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{\chi^2_{n-1}(\frac{1-\gamma}{2})}, \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{\chi^2_{n-1}(\frac{1+\gamma}{2})}\right).$$

(c) Derive jointly sufficient statistics for μ and σ.

Solution:

$$\begin{aligned}
f_n(\mathbf{x}|\mu,\sigma) &= \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left\{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}\right\} \\
&= \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left\{-\frac{\sum_{i=1}^n x_i^2}{2\sigma^2} + \frac{\mu\sum_{i=1}^n x_i}{\sigma^2} - \frac{n\mu^2}{2\sigma^2}\right\}
\end{aligned}$$

$(T_1, T_2) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ are jointly sufficient statistics by the factorization theorem with $u(\mathbf{x}) = 1$ and $v((T_1,T_2),(\mu,\sigma)) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left\{-\frac{T_2}{2\sigma^2} + \frac{\mu T_1}{\sigma^2} - \frac{n\mu^2}{2\sigma^2}\right\}$.

2. Suppose X_1 and X_2 are independent normal random variables with the same mean μ and variance σ^2. Let

$$\begin{aligned}
Y_1 &= X_1 + 2X_2 \\
Y_2 &= 2X_1 - X_2.
\end{aligned}$$

(a) Prove that Y_1 and Y_2 are independent and find their joint distribution.

Solution: Let $Z_1 = (Y_1 - 3\mu)/\sqrt{5\sigma^2}$ and $Z_2 = (Y_2 - \mu)/\sqrt{5\sigma^2}$. Then

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} \frac{X_1-\mu}{\sigma} \\ \frac{X_1-\mu}{\sigma} \end{bmatrix} = A \begin{bmatrix} \frac{X_1-\mu}{\sigma} \\ \frac{X_2-\mu}{\sigma} \end{bmatrix}$$

Since the matrix A of the transformation is orthogonal, it follows that Z_1 and Z_2 are i.i.d. and have the standard normal distribution. Thus, Y_1 and Y_2 are independent, $Y_1 \sim N(3\mu, 5\sigma^2)$ and $Y_2 \sim N(\mu, 5\sigma^2)$. Since Y_1 and Y_2 are independent, their joint distribution is also normal and is defined by these marginal distributions.

(b) Let X_3 be another normal random variable that is i.i.d of X_1 and X_2. Find the constant c such that

$$\frac{c(\overline{X}_2 - \mu)}{\sqrt{(X_1 - X_2)^2 + 2(X_3 - \mu)^2}}$$

follows a t-distribution, where $\overline{X}_2$ is the average of X_1 and X_2. Also find the degrees of freedom of the t-distribution.

Solution: Note that $(X_1 - X_2)^2 = 2[(X_1 - \overline{X}_2)^2 + (X_2 - \overline{X}_2)^2]$, so $\overline{X}_2$ and $(X_1 - X_2)^2$ are independent.

Clearly,

$$\frac{\overline{X}_2}{\sqrt{(X_1 - X_2)^2 + 2X_3^2}} = \frac{\sqrt{2}\bar{X}_2}{2\sqrt{(X_1 - \overline{X}_2)^2 + (X_2 - \overline{X}_2)^2 + (X_3 - \mu)^2}}$$

We find that $\frac{\sqrt{2}}{\sigma}(\bar{X}_2 - \mu) \sim N(0,1)$ and $\frac{(X_1 - \overline{X}_2)^2 + (X_2 - \overline{X}_2)^2}{\sigma^2} \sim \chi_1^2$. Moreover, $\frac{(X_3 - \mu)^2}{\sigma^2} \sim \chi_1^2$ and is independent of $\overline{X}_2$ and $(X_1 - \overline{X}_2)^2 + (X_2 - \overline{X}_2)^2$. So $(X_1 - \overline{X}_2)^2 + (X_2 - \overline{X}_2)^2 + X_3^2 \sim \chi_2^2$ and is independent of $\overline{X}_2$.

This implies that

$$\frac{\sqrt{2}\bar{X}_2}{\sqrt{((X_1 - \overline{X}_2)^2 + (X_2 - \overline{X}_2)^2 + X_3^2)/2}} \sim t_2$$

Set the left hand side equal to $\frac{c(\overline{X}_2 - \mu)}{\sqrt{(X_1 - X_2)^2 + 2(X_3 - \mu)^2}}$ and we have $c = 2\sqrt{2}$.

3. Let $X_1, \ldots, X_n$ be a random sample from a uniform distribution on the interval $[\theta, \theta + 1]$, with unknown $\theta > 0$.

(a) Assuming that the prior distribution of θ has the pdf $f(x|\beta) = \frac{1}{\beta}e^{-x/\beta}$. Find the Bayes estimator of θ for the squared error loss function. Just provide an expression but do not work out the final answer.

Solution: The posterior p.d.f. of θ

$$\xi(\theta|\mathbf{x}) \propto f(\theta|\beta)g_n(\mathbf{x}|\theta) \propto \frac{1}{\beta}e^{-\theta/\beta}\mathbf{1}_{\{x_{(n)}-1 \leq \theta \leq x_{(1)}\}}$$

So the posterior p.d.f. of θ is

$$\xi(\theta|\mathbf{x}) = \frac{\frac{1}{\beta}e^{-\theta/\beta}}{\int_{x_{(n)}-1}^{x_{(1)}} \frac{1}{\beta}e^{-\theta/\beta}d\theta} \quad \text{for} \quad x_{(n)} - 1 \leq \theta \leq x_{(1)}$$

The Bayes estimator is the posterior mean $E(\theta|\mathbf{x})$, that is,

$$E(\theta|\mathbf{x}) = \int_{x_{(n)}-1}^{x_{(1)}} \theta\xi(\theta|\mathbf{x})d\theta.$$

(b) Derive a 95% Bayes confidence interval for θ.

Solution: Take the 2.5% quantile of the posterior distribution as the lower bound $L(\mathbf{x})$ and the 97.5% quantile as the upper bound $U(\mathbf{x})$. Then we have $P(L(\mathbf{x}) \leq \theta \leq U(\mathbf{x})|\mathbf{x}) = 0.95$.

(c) Derive a 95% upper Bayesian confidence bound. How does this confidence bound compare with the upper bound of the two-sided ci?

Solution: Take the 95% quantile as the upper bound $U_2(\mathbf{x})$. Then we have $P(\theta \leq U_2(\mathbf{x})|\mathbf{x}) = 0.95$. This upper bound is smaller than the upper bound of the two-sided ci.

(d) When would you use the upper confidence bound?

Solution: When we only cares how large the parameter can be. For example, the fatality rate of toxic.

4. Let $X_1, \ldots, X_n$ be a random sample from a Negative Binomial(r, θ) distribution, where r is known and $0 < \theta < 1$ is unknown. Assume θ has a prior distribution Beta(α, β), where $\alpha, \beta > 0$ are known.

(a) Find the posterior distribution of θ given $X_1, \ldots, X_n$.

Solution: Let $\propto$ denote the left hand side is a constant multiple (involving no θ) of the right hand side. Then

$$\xi(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}$$
$$f_n(\mathbf{X}|\theta) \propto \theta^{nr}(1-\theta)^{\sum_{i=1}^n X_i}$$
$$\xi(\theta|\mathbf{X}) \propto \xi(\theta) f_n(\mathbf{X}|\theta) \propto \theta^{nr+\alpha-1}(1-\theta)^{\sum_{i=1}^n X_i+\beta-1}.$$

Compare the form of $\xi(\theta|\mathbf{X})$ with a Beta distribution density and we conclude the posterior distribution of θ is Beta$(nr+\alpha, \sum_{i=1}^n X_i + \beta)$.

(b) Find the Bayes estimator for θ under squared error loss.

Solution: The Bayes estimator under squared error loss is the posterior mean, which is

$$\hat{\theta} = (\alpha + nr)/(\alpha + nr + \beta + \sum_{i=1}^n X_i).$$

(c) Show $\bar{X}$ is a minimal sufficient statistic for θ, by using the Bayes estimator found in the last part.

Solution: By $f_n(\mathbf{X}|\theta)$ found in (a) and the factorization theorem, $\sum_{i=1}^n X_i$ is a sufficient statistic for θ. Since it is a one-to-one function of $\bar{X}$, the latter is also sufficient. Since $\hat{\theta}$ is a one-to-one function of $\sum_{i=1}^n X_i$, it is itself sufficient and since $\hat{\theta}$ is a Bayes estimator, it is minimal sufficient by the criterion that Bayes estimators that are sufficient are minimal sufficient. Finally $\bar{X}$ is minimal sufficient since it is a one-to-one function of $\hat{\theta}$.

(d) Obtain a 90% Bayesian ci for θ. What is the interpretation of this ci and how does it differ from the interpretation of a frequentist ci?

Solution: Take the 5% quantile of the posterior distribution as the lower bound $L(\mathbf{x})$ and the 95% quantile as the upper bound $U(\mathbf{x})$. Then we have $P(L(\mathbf{x}) \leq \theta \leq U(\mathbf{x})|\mathbf{x}) = 0.9$. It means that given the observed data there is a 90% probability that the parameter falls into $(L(\mathbf{x}), U(\mathbf{x}))$. The data is treated as fixed here while in frequentist ci the randomness comes from the data. In other words, in frenquentist ci if we draw random sample multiple times, we expect the parameter be covered by the ci 90% times while in Bayesian ci, given the same data, if we draw the parameter multiple time, it will be in the ci 90%times.

5. Assume we have a sample from a Poisson(λ) distribution. Obtain the UMVUE of $P(X = 0)$.

Solution: The p.m.f is $f(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-\lambda}e^{\log \lambda x}}{x!}$. It belongs to exponential family with $a(\lambda) = e^{-\lambda}$, $b(x) = \frac{1}{x!}$, $c(\lambda) = \log \lambda$ and $d(x) = x$. The parameter space $\lambda > 0$ contains an open set. So by theorem on complete statistics of exponential family we have $\sum_{i=1}^n X_i$ is a complete statistics. It is easy to know $\mathbf{1}_{\{X_j=0\}}$ is an unbiased estimator of $p(X = 0)$. Then $E(\mathbf{1}_{\{X_j=0\}}|\sum_{i=1}^n X_i)$ is the UMVUE of $p(X = 0)$. By sum of k i.i.d Poisson(λ) random variables follows Poisson($k\lambda$)

$$
\begin{aligned}
E(\mathbf{1}_{\{X_j=0\}}|\sum_{i=1}^n X_i = t) &= P(X_j = 0|\sum_{i=1}^n X_i = t) \\
&= \frac{P(X_j = 0, \sum_{i=1}^n X_i = t)}{P(\sum_{i=1}^n X_i = t)} \\
&= \frac{P(X_j = 0)P(\sum_{i \neq j} X_i = t)}{P(\sum_{i=1}^n X_i = t)} \\
&= \frac{e^{-\lambda}e^{-(n-1)\lambda}((n-1)\lambda)^t/t!}{e^{-n\lambda}(n\lambda)^t/t!} \\
&= (\frac{n-1}{n})^t.
\end{aligned}
$$

So $(\frac{n-1}{n})^{\sum_{i=1}^n X_i}$ is the UMVUE of $P(X = 0)$ by Lehmann-Scheffé Theorem on UMVUE.