Coursework 4: STAT 570

Philip Pham

October 22, 2018

1. Consider the so-called Neyman-Scott problem in which $Y_{ij} \mid \mu_i, \sigma_2 \sim_{\text{ind}} \mathcal{N}\left(\mu_i, \sigma^2\right), i = 1, \ldots, n, j = 1, 2$.

   (a) Obtain the MLE of σ^2 and show that it is inconsistent. Why does this inconsistency arise in this example?

   Solution: The likelihood is

$$
\begin{aligned}
L\left(\mu, \sigma\right) &= \prod_{i=1}^{n}\prod_{j=1}^{2} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left(Y_{ij} - \mu_i\right)^2\right) \\
&= \prod_{i=1}^{n} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2}\left[\left(Y_{i1} - \mu_i\right)^2 + \left(Y_{i2} - \mu_i\right)^2\right]\right),
\end{aligned} \tag{1}
$$

   so the log-likelikehood is

$$
l\left(\mu, \sigma\right) = -n\log\left(2\pi\right) - n\log\left(\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[\left(Y_{i1} - \mu_i\right)^2 + \left(Y_{i2} - \mu_i\right)^2\right]. \tag{2}
$$

   Taking the derivative with respect to σ^2, we have

$$
\frac{\partial}{\partial\sigma^2} l\left(\mu, \sigma^2\right) = -\frac{n}{\sigma^2} + \frac{1}{2\left(\sigma^2\right)^2}\sum_{i=1}^{n}\left[\left(Y_{i1} - \mu_i\right)^2 + \left(Y_{i2} - \mu_i\right)^2\right]. \tag{3}
$$

   Solving Equation 3, where $\frac{\partial}{\partial\sigma^2} l\left(\hat{\mu}, \hat{\sigma}^2\right) = 0$, we have

$$
\hat{\sigma}^2 = \frac{1}{2n}\sum_{i=1}^{n}\left[\left(Y_{i1} - \hat{\mu}_i\right)^2 + \left(Y_{i2} - \hat{\mu}_i\right)^2\right]. \tag{4}
$$

   Taking the derivative of Equation 2 with respect to μ_i, we have

$$
\frac{\partial}{\partial\mu_i} l\left(\mu, \sigma^2\right) = \frac{1}{\sigma^2}\left(Y_{i1} + Y_{i2} - 2\mu_i\right). \tag{5}
$$

   Solving Equation 5, where $\frac{\partial}{\partial\mu_i} l\left(\hat{\mu}, \hat{\sigma}^2\right) = 0$, we have

$$
\hat{\mu}_i = \frac{Y_{i1} + Y_{i2}}{2}. \tag{6}
$$

   Substituting Equation 6 into Equation 4, we have

$$
\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_{i1} - Y_{i2}}{2}\right)^2. \tag{7}
$$

Taking the expected value of Equation 7, we have

$$\begin{aligned}\mathbb{E}\left[\hat{\sigma}^2\right] &= \frac{1}{4n}\sum_{i=1}^{n}\left(\mathbb{E}\left[Y_{i1}^2\right] + \mathbb{E}\left[Y_{i2}^2\right] - 2\mathbb{E}\left[Y_{i1}Y_{i2}\right]\right) \\ &= \frac{1}{4n}\sum_{i=1}^{n}\left(\left(\sigma^2+\mu_i^2\right) + \left(\sigma^2+\mu_i^2\right) - 2\mu_i^2\right) \\ &= \frac{\sigma^2}{2}. \end{aligned} \tag{8}$$

Clearly, $\mathbb{E}\left[\hat{\sigma}^2\right] = \sigma^2/2 \nrightarrow \sigma^2$, so the estimator is not consistent.
This is because the MLE estimate of σ^2 depends on $\mu_1, \ldots, \mu_n$, so the number of parameters being estimated increases with n. Thus, the model is not well-defined.

(b) Derive the posterior distribution corresponding to the prior

$$\pi\left(\mu_1, \ldots, \mu_n, \sigma^2\right) \propto \sigma^{-n-2} \tag{9}$$

and show that

$$\mathbb{E}\left[\sigma^2 \mid Y\right] = \frac{1}{2(n-1)}\sum_{i=1}^{n}\frac{\left(Y_{i1}-Y_{i2}\right)^2}{2}. \tag{10}$$

Solution: Using the likelihood in Equation 1 and the prior in Equation 9. We have that

$$p\left(\mu, \sigma^2 \mid Y\right) \propto L\left(\mu, \sigma^2\right)\pi\left(\mu_1, \ldots, \mu_n, \sigma^2\right). \tag{11}$$

We have that

$$\begin{aligned} p\left(Y\right) &= \int_0^\infty \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty L\left(\mu, \sigma^2\right)\pi\left(\mu_1, \ldots, \mu_n, \sigma^2\right)\,\mathrm{d}\mu_1\cdots\mathrm{d}\mu_n\,\mathrm{d}\sigma^2 \\ &= \int_0^\infty \frac{1}{2^n\pi^n\left(\sigma^2\right)^{(3n+2)/2}}\left(\pi\sigma^2\right)^{n/2}\prod_{i=1}^{n}\exp\left(-\frac{1}{4\sigma^2}\left(Y_{i1}-Y_{i2}\right)^2\right)\mathrm{d}\sigma^2 \\ &= \int_0^\infty \frac{1}{2^n\pi^{n/2}\left(\sigma^2\right)^{n+1}}\exp\left(-\frac{1}{4\sigma^2}\sum_{i=1}^{n}\left(Y_{i1}-Y_{i2}\right)^2\right)\mathrm{d}\sigma^2 \\ &= -\frac{2^n}{\pi^{n/2}}\left(\sum_{i=1}^{n}\left(Y_{i1}-Y_{i2}\right)^2\right)^{-n}\int_\infty^0 u^{n-1}\exp\left(-u\right)\,\mathrm{d}u \\ &= \frac{1}{\pi^{n/2}}\left(\sum_{i=1}^{n}\frac{\left(Y_{i1}-Y_{i2}\right)^2}{2}\right)^{-n}\Gamma(n). \end{aligned} \tag{12}$$

Normalizing Equation 11 with the evidence Equation 12, we have the posterior

$$p\left(\mu, \sigma^2 \mid Y\right) = \frac{\left(\sigma^2\right)^{-(3n+2)/2}}{2^n\pi^{n/2}\Gamma(n)}\left(\sum_{i=1}^{n}\frac{\left(Y_{i1}-Y_{i2}\right)^2}{2}\right)^n\prod_{i=1}^{n}\prod_{j=1}^{2}\exp\left(-\frac{1}{2\sigma^2}\left(Y_{ij}-\mu_i\right)^2\right). \tag{13}$$

Marginalizing μ in Equation 13, we get

$$\begin{aligned} p\left(\sigma^2 \mid Y\right) &= \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty p\left(\mu, \sigma^2 \mid Y\right)\mathrm{d}\mu_1\cdots\mathrm{d}\mu_n \\ &= \frac{\left(\sigma^2\right)^{-n-1}}{2^n\Gamma(n)}\left(\sum_{i=1}^{n}\frac{\left(Y_{i1}-Y_{i2}\right)^2}{2}\right)^n\exp\left(-\frac{1}{4\sigma^2}\sum_{i=1}^{n}\left(Y_{i1}-Y_{i2}\right)^2\right). \end{aligned} \tag{14}$$

Taking the expectation over the distribution in Equation 14, we have that

$$\begin{aligned}
\mathbb{E}\left[\sigma^2 \mid Y\right] &= \int_0^\infty \sigma^2 p\left(\sigma^2 \mid Y\right) \, \mathrm{d}\sigma^2 \\
&= \frac{1}{\Gamma(n)} \int_0^\infty \left(\frac{1}{4\sigma^2} \sum_{i=1}^{n} \left(Y_{i1} - Y_{i2}\right)^2\right)^n \exp\left(-\frac{1}{4\sigma^2} \sum_{i=1}^{n} \left(Y_{i1} - Y_{i2}\right)^2\right) \\
&= \frac{\sum_{i=1}^{n} \left(Y_{i1} - Y_{i2}\right)^2}{4\Gamma(n)} \int_0^\infty u^{n-1-1} \exp\left(u\right) \, \mathrm{d}u \\
&= \frac{\Gamma(n-1)}{2\Gamma(n)} \sum_{i=1}^{n} \frac{\left(Y_{i1} - Y_{i2}\right)^2}{2} \\
&= \frac{1}{2(n-1)} \sum_{i=1}^{n} \frac{\left(Y_{i1} - Y_{i2}\right)^2}{2},
\end{aligned} \tag{15}$$

which is the desired result.

(c) Hence, using Equation 15, show that $\mathbb{E}\left[\sigma^2 \mid Y\right] \to \sigma^2/2$ as $n \to \infty$, so that the posterior mean is inconsistent.

Solution: From Equation 15, we have that

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[\sigma^2 \mid Y\right] &= \lim_{n\to\infty} \frac{1}{2(n-1)} \sum_{i=1}^{n} \frac{\mathbb{E}\left[\left(Y_{i1} - Y_{i2}\right)^2\right]}{2} \\
&= \lim_{n\to\infty} \frac{1}{2(n-1)} \sum_{i=1}^{n} \frac{\operatorname{Var}\left(Y_{i1} - Y_{i2}\right)}{2} \\
&= \lim_{n\to\infty} \frac{n\sigma^2}{2(n-1)} \\
&= \frac{\sigma^2}{2} \neq \sigma^2,
\end{aligned} \tag{16}$$

so the posterior mean is inconsistent.

(d) Examine the posterior distribution corresponding to the prior

$$\pi\left(\mu_1, \ldots, \mu_n \sigma^2\right) \propto \sigma^{-2}. \tag{17}$$

Solution: If we use Equation 17, Equation 12 becomes

$$\begin{aligned}
p\left(Y\right) &= \int_0^\infty \frac{1}{2^n \pi^{n/2} \left(\sigma^2\right)^{n/2+1}} \exp\left(-\frac{1}{4\sigma^2} \sum_{i=1}^{n} \left(Y_{i1} - Y_{i2}\right)^2\right) \mathrm{d}\sigma^2 \\
&= \frac{\Gamma\left(\frac{n}{2}\right)}{\pi^{n/2}} \left(\sum_{i=1}^{n} \left(Y_{i1} - Y_{i2}\right)^2\right)^{-n/2}.
\end{aligned} \tag{18}$$

With Equation 18, the posterior becomes

$$p\left(\mu, \sigma^2 \mid Y\right) = \frac{\left(\sigma^2\right)^{-n-1}}{2^n \pi^{n/2} \Gamma(n/2)} \left(\sum_{i=1}^{n} \frac{\left(Y_{i1} - Y_{i2}\right)^2}{2}\right)^n \prod_{i=1}^{n} \prod_{j=1}^{2} \exp\left(-\frac{1}{2\sigma^2} \left(Y_{ij} - \mu_i\right)^2\right). \tag{19}$$

Marginalizing Equation 19 over μ, we have

$$p\left(\sigma^2 \mid Y\right) = \frac{1}{\sigma^2 \Gamma\left(n/2\right)} \left(\frac{\sum_{i=1}^{n} \left(Y_{i1} - Y_{i2}\right)^2}{4\sigma^2}\right)^{n/2} \exp\left(-\frac{\sum_{i=1}^{n} \left(Y_{i1} - Y_{i2}\right)^2}{4\sigma^2}\right). \tag{20}$$

Length (mm)	0	1	2	3	4	5	6	7	8	9	10	11	12
1	2.247	2.640	2.842	2.908	3.099	3.126	3.245	3.328	3.355	3.383	3.572	3.581	3.681
10	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.454	2.474
20	1.312	1.314	1.479	1.552	1.700	1.803	1.861	1.865	1.944	1.958	1.966	1.997	2.006
50	1.339	1.434	1.549	1.574	1.589	1.613	1.746	1.753	1.764	1.807	1.812	1.840	1.852

Table 1: Failure stress data for four groups of fibers.

Equation 20 is quite similar to Equation 14, but with n replaced by $n/2$ in the gamma function and the exponent of the sum of squares.

(e) Is the posterior mean for σ^2 consistent in this case?

Solution: Yes. Taking the expectation with Equation 20, we have

$$\begin{aligned}\mathbb{E}\left[\sigma^2 \mid Y\right] &= \int_0^\infty p\left(\sigma^2 \mid Y\right)\,\mathrm{d}\sigma^2 \\ &= \frac{1}{4\Gamma(n/2)}\sum_{i=1}^n (Y_{i1}-Y_{i2})^2 \int_0^\infty u^{n/2-1-1}\exp(-u)\,\mathrm{d}u \\ &= \frac{\Gamma(n/2-1)}{4\Gamma(n/2)}\sum_{i=1}^n (Y_{i1}-Y_{i2})^2 \\ &= \frac{1}{2(n/2-1)}\sum_{i=1}^n \frac{(Y_{i1}-Y_{i2})^2}{2} \\ &= \frac{1}{n-2}\sum_{i=1}^n \frac{(Y_{i1}-Y_{i2})^2}{2}. \end{aligned} \tag{21}$$

Taking the limit of Equation 21, we have

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\left[\sigma^2 \mid Y\right] &= \lim_{n\to\infty}\frac{1}{n-2}\sum_{i=1}^n \frac{\mathbb{E}\left[(Y_{i1}-Y_{i2})^2\right]}{2} \\ &= \lim_{n\to\infty}\frac{n}{(n-2)}\frac{1}{n}\sum_{i=1}^n \frac{\mathrm{Var}(Y_{i1}-Y_{i2})}{2} \\ &= \lim_{n\to\infty}\frac{n}{(n-2)}\frac{1}{n}\sum_{i=1}^n \frac{2\sigma^2}{2} \\ &= \lim_{n\to\infty}\frac{n}{(n-2)}\sigma^2 \\ &= \sigma^2, \end{aligned} \tag{22}$$

so the posterior mean is consistent when the prior doesn't depend on n.

2. The data in Table 1 contain data on a typical reliability experiment and give the failure stresses (in GPa) of four samples of carbon fibers of lengths 1, 10, 20 and 50mm.

(a) Consider a Bayesian analysis with an exponential likelihood and a gamma prior, $\lambda \sim \mathrm{Gamma}(a, b)$. Derive the form of the posterior distribution for λ.

Solution: Suppose that we observe independent and identically distributed $Y_i \sim \mathrm{Exponential}(\lambda)$ for $i = 1, \ldots, n$. Let $Y = (Y_1, \ldots, Y_n)$. Then, we have the likelihood function

$$L(\lambda) = p(Y \mid \lambda) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} Y_i\right). \tag{23}$$

From Equation 23, we have the posterior

$$\begin{aligned} p(\lambda \mid Y) &\propto p(Y \mid \lambda)\, p(\lambda) \\ &\propto \left(\lambda^n \exp\left(-\lambda \sum_{i=1}^{n} Y_i\right)\right)\left(\frac{b^a}{\Gamma(a)}\lambda^{a-1}\exp(-b\lambda)\right) \\ &\propto \lambda^{a+n-1}\exp\left(-\left(b + \sum_{i=1}^{n} Y_i\right)\lambda\right), \end{aligned} \tag{24}$$

which equal to the Gamma probability density function up to a constant factor independent of λ. Thus, we have that $\lambda \mid Y \sim \Gamma\left(a+n, b+\sum_{i=1}^{n} Y_i\right)$, and

$$p(\lambda \mid Y) = \frac{\left(b + \sum_{i=1}^{n} Y_i\right)^{a+n}}{\Gamma(a+n)}\lambda^{a+n-1}\exp\left(-\left(b + \sum_{i=1}^{n} Y_i\right)\lambda\right). \tag{25}$$

(b) Choose a and b so that the prior probability that λ lies between 0.05 and 1 is 0.95.

Solution: For a specified mean μ, we specify our prior as the gamma distribution $\text{Gamma}(k\mu, k)$ for some k. Let F be the cumulative distribution function. We choose k such that

$$F(1) - F(0.05) = 0.95. \tag{26}$$