

# Coursework 4: STAT 570

Philip Pham

October 22, 2018

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

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

**Solution:** The likelihood is

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

so the log-likelihood is

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

Taking the derivative with respect to  $\sigma^2$ , we have

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

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

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

Taking the derivative of Equation 2 with respect to  $\mu_i$ , we have

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

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

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

Substituting Equation 6 into Equation 4, we have

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

Taking the expected value of Equation 7, we have

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

Clearly,  $\mathbb{E}[\hat{\sigma}^2] = \sigma^2/2 \not\rightarrow \sigma^2$ , so the estimator is not consistent.

This is because the MLE estimate of  $\sigma^2$  depends on  $\mu_1, \dots, \mu_n$ , 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

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

and show that

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

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

$$p(\mu, \sigma^2 | Y) \propto L(\mu, \sigma^2) \pi(\mu_1, \dots, \mu_n, \sigma^2).\tag{11}$$

We have that

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

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

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

Marginalizing  $\mu$  in Equation 13, we get

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

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

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

which is the desired result.

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

**Solution:** From Equation 15, we have that

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

so the posterior mean is inconsistent.

- (d) Examine the posterior distribution corresponding to the prior

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

**Solution:** If we use Equation 17, Equation 12 becomes

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

With Equation 18, the posterior becomes

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

Marginalizing Equation 19 over  $\mu$ , we have

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

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  $\sigma^2$  consistent in this case?

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

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

Taking the limit of Equation 21, we have

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

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,  $\lambda \sim \text{Gamma}(a, b)$ . Derive the form of the posterior distribution for  $\lambda$ .

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

Length (mm)	$a'$	$b'$	Mean	Standard error
1	16.634063	48.275125	0.344568	0.084484
10	16.634063	37.320125	0.445713	0.109284
20	16.634063	30.025125	0.554005	0.135836
50	16.634063	28.940125	0.574775	0.140928

Table 2: The results of updating the prior belief in Part 2b with the data.

the likelihood function

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

From Equation 23, we have the posterior

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

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

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

- (b) Choose  $a$  and  $b$  so that the prior probability that  $\lambda$  lies between 0.05 and 1 is 0.95.

**Solution:** For a specified mean  $\mu$ , 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

$$F(1) - F(0.05) = 0.95. \quad (26)$$

Choosing  $a \approx 3.634$  and  $b \approx 7.268$  satisfies Equation 26. These values were obtained numerically in `failure_stresses.ipynb`.

- (c) Obtain the posterior means and posterior standard deviations for  $\lambda$  and give histogram representations of  $p(\lambda | y)$  for each of the groups in Table 1. Compare inference with the frequentist analyses. Also give histogram representations of the posterior for  $\lambda^{-1}$ , again for each of the four groups.

**Solution:** The posterior parameters, means, and standard deviations can be seen in Table 2. The first two columns are the shape and rate parameters of the gamma posterior, respectively.

The means are similar to the estimates from frequentist analyses in the previous homework but are drawn towards the prior mean of 0.5. The

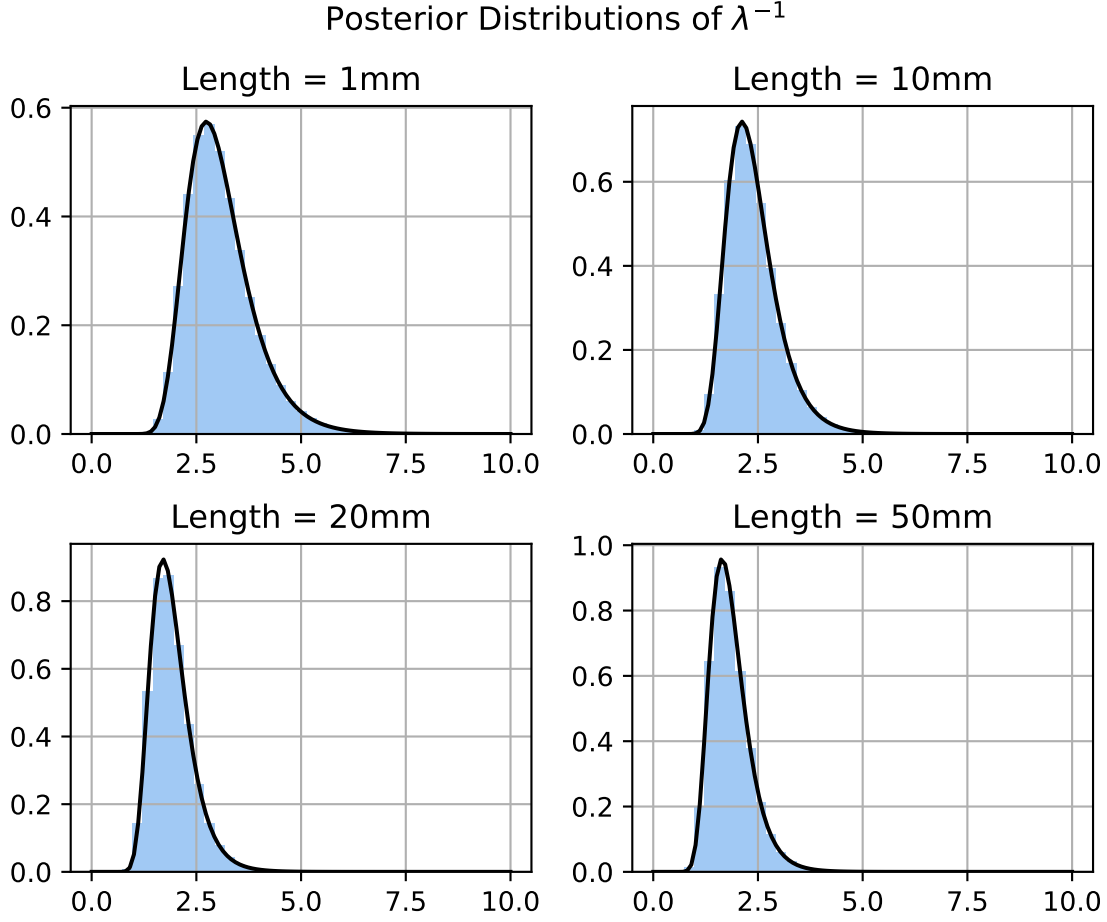


Figure 1: Histograms and probability density for samples drawn from the posteriors for  $\lambda^{-1}$ .

standard errors are slightly smaller than those obtained from fitting the exponential model but are much larger than either the quasi-likelihood model or sandwich estimates.

For the posterior distribution of  $\lambda^{-1}$ , we have that

$$\begin{aligned}
 p(\lambda^{-1} | Y) &= p(\lambda | Y) \left| \frac{d}{d\lambda^{-1}} (\lambda^{-1})^{-1} \right| = p(\lambda | Y) \lambda^2 \\
 &= \frac{(b + \sum_{i=1}^n Y_i)^{a+n}}{\Gamma(a+n)} \lambda^{a+n+1} \exp\left(-\left(b + \sum_{i=1}^n Y_i\right) \lambda\right), \quad (27)
 \end{aligned}$$

where we used Equation 25 and transformation of random variables, so  $\lambda^{-1} | Y \sim \text{InverseGamma}(a+n, b + \sum_{i=1}^n Y_i)$ .

Theoretical histograms along with their density are plotted in Figure 1.

Unlike the MLE, the posterior mean is not invariant under reparameterization:

$$\mathbb{E}[\lambda^{-1} | Y] = \frac{b'}{a' - 1} \geq \frac{b'}{a'} = \frac{1}{\mathbb{E}[\lambda | Y]}.$$

See `failure_stresses.ipynb` for code to reproduce plots.