STAT 153 & 248 - Time Series
Lab One

Spring 2025, UC Berkeley

January 24, 2025

Part 2: Bayesian analysis in a simple estimation problem

Problem 1. *Suppose we have six observations*

$$Y_1 = 26.6, Y_2 = 38.5, Y_3 = 34.4, Y_4 = 34, Y_5 = 31, Y_6 = 23.6,$$

which we model as

$$Y_1, \ldots, Y_6 \overset{i.i.d}{\sim} N(\theta, \sigma^2),$$

where θ and σ^2 are unknown parameters. Conduct Bayesian inference on the unknown parameters θ and σ^2. Note that the answer depends on your choice of priors for θ and σ^2.

Suggestion. *Try using the following priors*

$$\theta, \log\sigma \overset{i.i.d}{\sim} \mathit{Unif}(-C, C)$$

with a large constant C, as in the last lecture.

Solution. *The first step is to choose priors for θ and σ^2. Similar to the analysis of linear regression in class, we shall assume that*

$$\theta, \log\sigma \overset{i.i.d}{\sim} \mathit{Unif}(-C, C)$$

for a large constant C. These priors are supposed to capture our large prior uncertainty on the values of θ and σ. In density form, the prior density becomes:

$$\begin{aligned} f_{\theta,\sigma}(\theta,\sigma) &= f_\theta(\theta) f_\sigma(\sigma) = f_\theta(\theta) f_{\log\sigma}(\log\sigma)\frac{1}{\sigma} \\ &= \frac{I\{-C < \theta < C\}}{2C} \cdot \frac{I\{-C < \log\sigma < C\}}{2C} \cdot \frac{1}{\sigma} = \frac{I\{-C < \theta, \log\sigma < C\}}{4\sigma C^2}. \end{aligned}$$

The likelihood is

$$\begin{aligned} f_{Y_1,\ldots,Y_n|\theta,\sigma}(y_1,\ldots,y_n) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i-\theta)^2}{2\sigma^2}\right) \\ &= (2\pi)^{-n/2}\sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\theta)^2\right). \end{aligned}$$

The posterior density of θ, σ given the observed data $Y_1, \dots, Y_n$ is therefore

$$f_{\theta,\sigma|data}(\theta, \sigma) \propto \frac{I\{-C < \theta, \log\sigma < C\}}{4\sigma C^2} \cdot (2\pi)^{-n/2}\sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right)$$

$$\propto I\{-C < \theta, \log\sigma < C\} \cdot \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right). \quad (1)$$

This gives the joint posterior of θ and σ. If we want only the posterior of θ (i.e., the conditional density of θ given the data $Y_1, \dots, Y_n$), we need to integrate the above with respect to σ. This gives the following:

$$f_{\theta|data}(\theta) \propto \int I\{-C < \theta, \log\sigma < C\} \cdot \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right) d\sigma$$

$$= I\{-C < \theta < C\} \cdot \int_{e^{-C}}^{e^C} \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right) d\sigma.$$

When C is large, the integral will be basically be the same as 0 to ∞, leading to:

$$\int_{e^{-C}}^{e^C} \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right) d\sigma \approx \int_0^\infty \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right) d\sigma.$$

The change of variable

$$\frac{\sigma}{\sqrt{\sum_{i=1}^n (y_i - \theta)^2}} = s$$

gives

$$\int_0^\infty \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right) d\sigma = \frac{1}{[\sum_{i=1}^n (y_i - \theta)^2]^{n/2}} \int_0^\infty s^{-n-1} \exp\left(-\frac{1}{2s^2}\right) ds$$

$$\propto \frac{1}{[\sum_{i=1}^n (y_i - \theta)^2]^{n/2}}.$$

We thus get

$$f_{\theta|data}(\theta) \propto \frac{I\{-C < \theta < C\}}{[\sum_{i=1}^n (y_i - \theta)^2]^{n/2}}.$$

When C is large, the indicator above will play no role, so we just drop it to obtain:

$$f_{\theta|data}(\theta) \propto \left[\sum_{i=1}^{n}(y_i - \theta)^2\right]^{-n/2}. \quad (2)$$

It turns out that this posterior density is related to the t-density with $n-1$ degrees of freedom. To see this, recall (from wikipedia for example) first that t-density with $n-1$ degrees of freedom is proportional to

$$\left(1 + \frac{x^2}{n-1}\right)^{-n/2}. \quad (3)$$

To see the connection between (2) and (3), first write (below $\bar{y} = (y_1 + \cdots + y_n)/n$)

$$\begin{aligned}
f_{\theta|data}(\theta) &\propto \left[\sum_{i=1}^{n}(y_i - \theta)^2\right]^{-n/2} \\
&= \left(\sum_{i=1}^{n}(y_i - \bar{y})^2 + n(\bar{y} - \theta)^2\right)^{-n/2} \\
&= \left(\sum_{i=1}^{n}(y_i - \bar{y})^2\right)^{-n/2} \left(1 + \frac{n(\bar{y} - \theta)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right)^{-n/2} \\
&\propto \left(1 + \frac{n(\bar{y} - \theta)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right)^{-n/2} \\
&= \left(1 + \frac{n(\bar{y} - \theta)^2}{(n-1)s^2}\right)^{-n/2} \qquad (4)
\end{aligned}$$

where

$$s^2 := \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2. \qquad (5)$$

The expression (4) looks very much like the t-density (3). To see this more explicitly, let us compute the posterior density of $\sqrt{n}(\theta - \bar{y})/s$:

$$f_{\frac{\sqrt{n}(\theta-\bar{y})}{s}|data}(x) = f_{\theta|data}\left(\bar{y} + \frac{sx}{\sqrt{n}}\right) \propto \left(1 + \frac{x^2}{n-1}\right)^{-n/2}.$$

This proves that

$$\frac{\sqrt{n}(\theta - \bar{y})}{s} \mid data \sim t_{n-1}$$

where t_{n-1} is the t-distribution with $n-1$ degrees of freedom.

With this formula, we can do any uncertainty quantification for θ based on the observed data $y_1 = 26.6, y_2 = 38.5, y_3 = 34.4, y_4 = 34, y_5 = 31, y_6 = 23.6$. For example, we can calculate the posterior probability that θ belongs to the interval $[28, 32]$ as

$$\begin{aligned}
&\mathbb{P}\left\{28 \le \theta \le 32 \middle| Y_1 = 26.6, Y_2 = 38.5, Y_3 = 34.4, Y_4 = 34, Y_5 = 31, Y_6 = 23.6\right\} \\
&= \mathbb{P}\left\{\frac{\sqrt{6}(28 - 31.35)}{5.48} \le t_5 \le \frac{\sqrt{6}(32 - 31.35)}{5.48}\right\} = 0.511.
\end{aligned}$$

We can also give an interval whose posterior probability is exactly 0.95. Indeed, using the fact that

$$\mathbb{P}\{-2.57 \le t_5 \le 2.57\} = 0.95,$$

we obtain

$$\mathbb{P}\left\{-2.57 \le \frac{\sqrt{6}(\theta - 31.35)}{5.48} \le 2.57 \middle| Y_1 = 26.6, Y_2 = 38.5, Y_3 = 34.4, Y_4 = 34, Y_5 = 31, Y_6 = 23.6\right\}$$
$$= 0.95$$

or

$$\mathbb{P}\left\{25.6 \le \theta \le 37.1 \middle| Y_1 = 26.6, Y_2 = 38.5, Y_3 = 34.4, Y_4 = 34, Y_5 = 31, Y_6 = 23.6\right\} = 0.95$$

If only a point estimate of θ is desired, then one can simply use the posterior mean which equals $\bar{y} = 31.35$ (this is also the posterior median and the posterior mode as the t-distribution is symmetric about 0).

Next, let us focus on inference for the parameter σ. To obtain the posterior density for σ, we need to integrate (1) with respect to θ. This gives

$$f_{\sigma|data}(\sigma) \propto I\{-C < \log\sigma < C\} \cdot \sigma^{-n-1} \int_{-C}^{C} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2\right) d\theta.$$

Writing $\sum_{i=1}^{n}(y_i - \theta)^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 + n(\bar{y} - \theta)^2$, we obtain

$$f_{\sigma|data}(\sigma) \propto I\{-C < \log\sigma < C\} \cdot \sigma^{-n-1} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2\right) \int_{-C}^{C} \exp\left(-\frac{n}{2\sigma^2}(\bar{y} - \theta)^2\right) d\theta.$$

Again assuming that C is large, we can calculate the integral above from $-\infty$ to ∞. Using the formula for the normalizing constant of a normal distribution, we get

$$\int_{-C}^{C} \exp\left(-\frac{n}{2\sigma^2}(\bar{y} - \theta)^2\right) d\theta = \int_{-\infty}^{\infty} \exp\left(-\frac{n}{2\sigma^2}(\bar{y} - \theta)^2\right) d\theta = \frac{\sigma\sqrt{2\pi}}{\sqrt{n}} \propto \sigma.$$

We thus get

$$f_{\sigma|data}(\sigma) \propto I\{-C < \log\sigma < C\}\sigma^{-n} \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2\right).$$

Although this expression can be left as is, we can relate this to the χ^2-density with a little extra simplification assuming that C is large. Indeed, if C is large, we can replace the indicator by $I\{\sigma > 0\}$ to get

$$f_{\sigma|data}(\sigma) \propto I\{\sigma > 0\}\sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2\right) = I\{\sigma > 0\} \cdot \sigma^{-n} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right)$$

where s is as in (5). By a simple change of variable formula, we can now deduce that

$$\begin{aligned} f_{\frac{(n-1)s^2}{\sigma^2}|data}(x) &= f_{\sigma|data}\left(\frac{\sqrt{n-1}s}{\sqrt{x}}\right) \cdot \frac{s\sqrt{n-1}}{2} \cdot x^{-3/2} \\ &\propto I\{x > 0\} \cdot x^{\frac{n-1}{2}-1} \exp\left(-\frac{x}{2}\right) \end{aligned}$$

which means that

$$\left.\frac{(n-1)s^2}{\sigma^2}\right| data \sim Gamma\left(\frac{n-1}{2}, \frac{1}{2}\right) = \chi^2_{n-1}. \tag{6}$$

From here, any inference can be obtained on σ. If a point estimate of σ is desired, one simple way is to just take the point estimate of $\frac{(n-1)s^2}{\sigma^2}$ to be equal to the mean of χ^2_{n-1} which is $n-1$. This gives

$$\frac{(n-1)s^2}{\hat{\sigma}^2} = n - 1 \iff \hat{\sigma} = s = 5.48.$$

One can also posterior probability of any event involving σ. One can also construct an interval for σ with posterior probability exactly 0.95.