STAT 238 - Bayesian Statistics
Lab Two

Aditya Guntuboyina

30 Jan 2026

1 Uniform Prior on $\log \sigma$ or σ

Consider the problem of estimating a scale parameter σ from observations $X_1, \dots, X_n \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$. We want to use an uniformative prior for σ. There are two options. Option A is $\sigma \sim \text{uniform}(0, +\infty)$ (or uniform$(0, C)$ for a large C). Option B is $\log \sigma \sim \text{uniform}(-\infty, +\infty)$ (or uniform$(-C, C)$ for a large C). The second prior (option B) is universally preferred to the first prior (option A). Here are some reasons for why this is the case.

1. **Relative Scale vs Absolute Scale**: Consider the two probabilities $\mathbb{P}\{1 < \sigma < 2\}$ and $\mathbb{P}\{101 < \sigma < 102\}$. Under the first prior (uniform$(0, C)$), both these probabilities are the same.

   Under the second prior ($\log \sigma$ is uniform on $(-C, C)$), we have:

   $$\mathbb{P}\{1 < \sigma < 2\} = \mathbb{P}\{0 < \log \sigma < \log 2\} = \frac{\log 2}{2C} = \frac{0.693}{2C}$$

   and

   $$\mathbb{P}\{101 < \sigma < 102\} = \mathbb{P}\{\log 101 < \log \sigma < \log 102\} = \frac{\log(102/101)}{2C} = \frac{0.00985}{2C}$$

   The second probability is therefore much smaller than the first. This reflects the fact that Prior B regards the values 101 and 102 as much closer to one another than the values 1 and 2. This behavior aligns well with intuition on a relative scale: moving from $\sigma = 1$ to $\sigma = 2$ corresponds to a doubling, whereas moving from $\sigma = 101$ to $\sigma = 102$ represents only a very small relative change.

2. **Prior Odds**: Under the first prior, consider the prior odds that $\sigma < 2$ versus $\sigma \geq 2$:

   $$\frac{\mathbb{P}\{\sigma < 2\}}{\mathbb{P}\{\sigma \geq 2\}} = \frac{2/C}{(C-2)/C} = \frac{2}{C-2}$$

   which is very small (as C is large). This is clearly an informative statement about σ (that it is much more likely to be more than 2 than smaller than 2). On the other hand, the same odds for the second prior becomes:

   $$\frac{\mathbb{P}\{\sigma < 2\}}{\mathbb{P}\{\sigma \geq 2\}} = \frac{\mathbb{P}\{\log \sigma < \log 2\}}{\mathbb{P}\{\log \sigma \geq \log 2\}} = \frac{C + \log 2}{C - \log 2}$$

   which is approximately equal to 1, reflecting prior ignorance between the events $\sigma < 2$ and $\sigma \geq 2$.

3. **Posterior Propriety for** $n = 1$: The posterior for σ corresponding to the first prior is:

$$\frac{2\left(\frac{S}{2}\right)^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}\right)}\sigma^{-n}\exp\left(-\frac{S}{2\sigma^2}\right)I\{\sigma > 0\} \tag{1}$$

while the posterior corresponding to the second prior is:

$$\frac{2\left(\frac{S}{2}\right)^{n/2}}{\Gamma\left(\frac{n}{2}\right)}\sigma^{-(n+1)}\exp\left(-\frac{S}{2\sigma^2}\right)I\{\sigma > 0\}. \tag{2}$$

In both these formulae, $S := \sum_{i=1}^{n} X_i^2$.

These two posteriors will behave similarly when n is large. However, when $n = 1$, the first posterior is actually ill-defined because the Gamma function term is $\Gamma(0)$ which is ∞. In other words, the posterior is improper when $n = 1$. On the other hand, the second posterior is proper even when $n = 1$.

This is another reason for preferring the second prior in this problem. We would intuitively expect to obtain some concrete information on σ when $n = 1$ so we want the posterior to be proper in that case. This is only true for the second prior but not for the first prior.

Now consider the problem of estimating both θ and σ from $X_1, \dots, X_n \overset{\text{i.i.d}}{\sim} N(\theta, \sigma^2)$. Here there are two options for the prior. Prior A is:

$$\theta, \sigma \text{ are independent with } \theta \sim \text{unif}(-\infty, \infty) \text{ and } \sigma \sim \text{unif}(0, \infty).$$

Prior B is

$$\theta, \log\sigma \overset{\text{i.i.d}}{\sim} \text{unif}(-\infty, \infty).$$

Again the universally preferred prior is the second one. All of the reasons previously mentioned apply here as well. For the third point, the marginal posterior of σ corresponding to the first prior is:

$$\frac{2\left(\frac{S}{2}\right)^{(n-2)/2}}{\Gamma\left(\frac{n-2}{2}\right)}\sigma^{-(n-1)}\exp\left(-\frac{S}{2\sigma^2}\right)I\{\sigma > 0\} \tag{3}$$

and the marginal posterior for σ corresponding to the second prior is:

$$\frac{2\left(\frac{S}{2}\right)^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}\right)}\sigma^{-n}\exp\left(-\frac{S}{2\sigma^2}\right)I\{\sigma > 0\}. \tag{4}$$

In both the above formulae, $S = \sum_{i=1}^{n}(X_i - \bar{X})^2$. The first posterior above is ill-defined when $n = 1, 2$ while the second posterior is ill-defined only for $n = 1$. It is well-known that in this problem we would need at least two observations to estimate σ, so we would like the posterior to be well-defined for $n = 2$ which is only true if we use the second prior.

For more, see Zellner [2, pages 41-47]. See also Jaynes [1, Section 12.4] for another justification for the $\log\sigma \sim \text{uniform}(-\infty, +\infty)$ prior based on invariance to certain parameter transformations.

References

[1] Jayes, E. T. (2003) *Probability theory: the logic of science.* Cambridge University Press, 2003.

[2] Zellner, A., (1971) *An introduction to Bayesian inference in Econometrics.* Wiley Classics Library, 1971 (reprint edition in 1996).