Cranking through Bayesian Calculus, Part II

Ed Herbst

March 14, 2025

Let's get some more practice with the Bayesian machinery in a regression model.

$$y_t = x_t'\beta + u_t, \quad u_t \overset{i.i.d.}{\sim} N(0, \sigma^2). \tag{1}$$

We've practiced the Bayesian machinery already setting $\sigma^2 = 1$. Of course, that's a bad assumption for many problems. So let's incorporate estimation of σ^2 into our analysis. It is much more common in contempory analysis to use the *inverse gamma distribution* for σ^2 with parameters α and β. Specifically, a random variable σ^2 follows an inverse gamma distribution if and only if

$$p(\sigma^2|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma^2)^{-\alpha-1}\exp\left(\frac{-\beta}{\sigma^2}\right). \tag{2}$$

Problem 1: Plot the inverse gamma density for two different parameterizations $(\alpha, \beta) = (2, 1)$ and $(\alpha, \beta) = (2, 2)$. What is the difference between the two densities?

```
# your code here
```

The figures are helpful for gaining some intuition about the distribution. We can get some more insight from looking at moments of the distribution.

Problem 2: What are the expressions for the mean, median, variance of σ^2? You can do the integration (by parts) yourself, or just check them out on Wikipedia.

```
# your code here
```

In formulating a prior for a particular problem, it seems like it could be difficult to express beliefs about σ^2, given the expressions you've written down for Problem 2. Another way to parameterize the inverse gamma distribution comes from Gelman et al., who use a bijection from $(\nu_0, s_0^2) \to (\alpha, \beta)$ where

$$(\alpha, \beta) = \left(\frac{\nu_0}{2}, \frac{\nu_0}{2}s_0^2\right).$$

They refer to this distribution as a *scaled inverse chi-squared distribution* because, if $z \sim \chi^2(\nu_0)$, that is, if z follows a chi-squared distribution with ν_0 degrees of freedom, then $\sigma^2 = \nu_0 s_0^2/z$ follows the inverse gamma distribution.

Problem 3: Rewrite the density and the key moments of the distribution using this parameterization. What happens to the mean and median as $\nu_0 \longrightarrow \infty$?

```
# your code here
```

Let's get some intuition for this parameterization, and some practice constructing the posterior for σ^2. Assume that we have T independent observations of a normal random variables y_t,

$$y_t \sim iidN(0, \sigma^2), \quad t = 1, \ldots, T.$$

Problem 4: Write down the likelihood $p(Y|\sigma^2)$. Construct the average squared deviation of y_t from its mean. Call this s^2, and substitute this into the likelihood. Is this a sufficient statistic? Why or why not?

```
# your code here
```

Problem 5: Derive the posterior of $\sigma^2|Y$,

$$p(\sigma^2|Y) \propto p(Y|\sigma^2)p(\sigma^2),$$

using the (ν_0, s_0^2) parameterization of the prior distribution. Looking at the expression, how can you interpret ν_0 and s_0^2?