Economics 690
Problem Set: Hypothesis Testing

(1) Consider a set of data $y_1, y_2, \cdots y_T$ giving rise to likelihood function $L(\theta)$. Suppose two competing models are of interest which impose the following restrictions on θ: $\mathcal{M}_1 : \theta = c_1$ and $\mathcal{M}_2 : \theta = c_2$ for given constants c_1 and c_2.

Note: $\mathcal{M}_j$ can be obtained by imposing a *dogmatic* or *point prior* of the form

$$p(\theta) = I(\theta = \theta_j), \ j = 1, 2.$$

Suppose you believe both models are equally likely *a priori*. That is $p(\mathcal{M}_1) = p(\mathcal{M}_2) = 1/2$.

Show, under these assumptions, that the posterior odds of model 1 in favor of model 2,denoted K_{12}, is simply the likelihood ratio:

$$K_{12} = \frac{p(y|\theta = c_1)}{p(y|\theta = c_2)} = \frac{L(c_1)}{L(c_2)}.$$

(2) Using the result in (1), consider the following problem. Suppose that *only one* observation is available from a $N(\theta, 1)$ distribution, and let us call that single observation y.

Additionally, suppose that two models are of interest: $\mathcal{M}_1 : \theta = -1$ and $\mathcal{M}_2 : \theta = 1$ and both models are believed to be equally likely *a priori*.

Derive the Bayes factor in this case as a function of y and comment on the reasonableness of this result. If $y = 1$ (and these two models are the only ones under consideration), what is the *posterior probability* associated with $\mathcal{M}_1$? with $\mathcal{M}_2$?

(3) In our lecture notes on hypothesis testing, I provided an example of marginal likelihood calculation in our log wage equation. Specifically, we found that the posterior odds supporting a non-zero return to education were 6×10^{36}.

I want you to reproduce these results on your own. To do this, it will be useful to note:

- The prior hyperparameters were $\nu = 6$, $\lambda = (2/3)(.2)$, $\mu = 0$ and $V_\beta = 10I_2$.
- To calculate the marginal likelihood for a given model, you need to calculate the ordinate of a Student-t density. The lecture notes related to distributions provide the functional form of the Student-t.
- With this large sample size, there are a couple of things we need to do. First, we will need to work on the *log scale*. That is, we will need to calculate the log of the Student-t ordinates for both models, subtract them, and then exponentiate the result to get the desired posterior odds. [Note: This is how I proceeded in the lecture notes].
- When working on the log scale note that "gammaln" will calculate the log of the gamma function in MATLAB.
- Second, calculating the inverse and determinant of

  $$\Sigma \equiv \lambda(I_n + XV_\beta X')$$

  can be difficult. To this end, you should make use of the linear algebra formulas:

  $$(A + BCD)^{-1} = A^{-1} - A^{-1}B[DA^{-1}B + C^{-1}]^{-1}DA^{-1}$$

  and

  $$|A + BCD| = |A||C||C^{-1} + DA^{-1}B|.$$

  where A and C are nonsingular matrices, to substantially simplify these calculations.

(3b) Repeat your marginal calculations above, this time changing $V_\beta = 10I_2$ to $V_\beta = 1.0 \times 10^{100} I_2$. (This would correspond to a much "flatter" prior for β.)

Note: In Matlab, you can enter 1.0×10^{100} as "1.0e100"

How do your posterior odds change and how does this relate to Bartlett's paradox?