Problem Set 3
Eco 517

1) The likelihood of obtaining a sequence of T zeros is

$$p(X|\varepsilon) = (1-\varepsilon)^T$$

Taking a prior on ε that is $Beta(p,q)$, then we have

$$p(\varepsilon|X) \propto p(X|\varepsilon)\, p(\varepsilon) \propto (1-\varepsilon)^T \varepsilon^{p-1} (1-\varepsilon)^{q-1} = \varepsilon^{p-1} (1-\varepsilon)^{T+q-1}$$

This is the kernel of a $Beta(p, T+q)$ distribution. The exact density of $\varepsilon|X$ is therefore

$$p(\varepsilon|X) = \frac{\Gamma(p+T+q)}{\Gamma(p)\Gamma(T+q)} \varepsilon^{p-1} (1-\varepsilon)^{T+q-1}$$

The expectation of interest is therefore given by

$$\begin{aligned} E\left[\frac{1}{\varepsilon}|\{X_1,...,X_T\}\right] &= \int_0^1 \frac{1}{\varepsilon} \frac{\Gamma(p+T+q)}{\Gamma(p)\Gamma(T+q)} \varepsilon^{p-1} (1-\varepsilon)^{T+q-1} d\varepsilon \\ &= \int_0^1 \frac{\Gamma(p+T+q)}{\Gamma(p)\Gamma(T+q)} \varepsilon^{p-2} (1-\varepsilon)^{T+q-1} d\varepsilon \\ &= \frac{\Gamma(p+T+q)\Gamma(p-1)}{\Gamma(p)\Gamma(p-1+T+q)} \int_0^1 \frac{\Gamma(p-1+T+q)}{\Gamma(p-1)\Gamma(T+q)} \varepsilon^{p-2} (1-\varepsilon)^{T+q-1} d\varepsilon \end{aligned}$$

Note that the integrand is the denstiy of a $\beta(p-1, T+q)$ random variable and therefore integrates to 1. We then have

$$E\left[\frac{1}{\varepsilon}|\{X_1,...,X_T\}\right] = \frac{\Gamma(p+T+q)\Gamma(p-1)}{\Gamma(p)\Gamma(p-1+T+q)}$$

Using the fact that $\Gamma(k) = (k-1)\Gamma(k-1)$ we can simplify this expression to find that

$$E\left[\frac{1}{\varepsilon}|\{X_1,...,X_T\}\right] = \frac{p+T+q-1}{p-1}$$

a) The β distribution nests the uniform$(0,1)$ distribution and this occurs when $p = q = 1$. This expectation is infinite in this case regardless of the value of T.

b) When $p = 3$, and $q = 2$ we get

$$E\left[\frac{1}{\varepsilon}|\{X_1,...,X_T\}\right] = \frac{T+4}{2}$$

and as $T \to \infty$ so does $E\left[\frac{1}{\varepsilon}|\{X_1,...,X_T\}\right]$.

1 Obtaining posterior predictive densities for the SNLM with noninformative priors

In vector notation, the standard normal linear model is given by

$$Y|\beta,\sigma^2,X\tilde{}N\left(X\beta,\sigma^2 I\right)$$

The prior that we are asked to use is

$$p\left(\beta,\sigma^2\right)\propto\frac{1}{\sigma}$$

The posterior is then proportional to

$$\begin{aligned}
& \frac{1}{\sigma}\left(\frac{1}{\sigma}\right)^T \exp\left(-\frac{1}{2\sigma^2}\left(Y-X\beta\right)'\left(Y-X\beta\right)\right) \\
= & \left(\frac{1}{\sigma}\right)^{T+1} \exp\left(-\frac{1}{2\sigma^2}\left(Y-X\widehat{\beta}+X\left(\widehat{\beta}-\beta\right)\right)'\left(Y-X\widehat{\beta}+X\left(\widehat{\beta}-\beta\right)\right)\right) \\
= & \left(\frac{1}{\sigma^2}\right)^{\frac{T+1}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(Y-X\widehat{\beta}\right)'\left(Y-X\widehat{\beta}\right)+\left(\beta-\widehat{\beta}\right)'X'X\left(\beta-\widehat{\beta}\right)\right) \\
= & \left(\frac{1}{\sigma^2}\right)^{\frac{T+1}{2}} \exp\left(-\frac{e'e}{2\sigma^2}\right)\left(\sigma^k\left|X'X\right|^{\frac{-1}{2}}\right)\sigma^{-k}\left|X'X\right|^{\frac{1}{2}} \exp\left(\frac{-1}{2\sigma^2}\left(\beta-\widehat{\beta}\right)'X'X\left(\beta-\widehat{\beta}\right)\right)
\end{aligned}$$

From this we can see that the posterior of $\beta|\sigma^2$ is $N\left(\widehat{\beta},\sigma^2\left(X'X\right)^{-1}\right)$. We can integrate the above with respect to β and obtain an expression proportional to the marginal of σ^2. That expression is

$$\left(\frac{1}{2\sigma^2}\right)^{\frac{T-k-2}{2}}\left(e'e\right)^{\frac{(T-k)}{2}}\exp\left(-\frac{e'e}{2\sigma^2}\right)$$

This means that under the prior $p\left(\beta,\sigma^2\right)\propto\frac{1}{\sigma}$, $\frac{1}{\sigma^2}$ has a $\Gamma\left(T-k,e'e\right)$ distribution.

In summary, we found that

$$\beta|\sigma^2\tilde{}N\left(\widehat{\beta},\sigma^2\left(X'X\right)^{-1}\right)$$
$$\frac{1}{\sigma^2}\tilde{}\Gamma\left(\frac{T-k}{2},e'e\right)$$

We are given new data $\widetilde{X}$ from 2004, and we want to use this to predict the outcome of 2004, $\widetilde{y}$. If we knew what the parameters of the model $\left(\beta,\sigma^2\right)$ are, then $\widetilde{y}$ would have a normal distribution with mean $\widetilde{X}\beta$ and variance σ^2. However, there is only uncertain knowledge of the model parameters, and this is summarized by their posterior distribution. In short, the sources of uncertainty in the posterior predictive density of the election outcome are

1. The variability of the model represented in the variance in Y represented by σ^2 not explained by $X\beta$.

2. The posterior uncertainty in β and σ^2 summarized by $p\left(\beta, \sigma^2|Y\right) = p\left(\beta|\sigma^2, Y\right) p\left(\sigma^2|Y\right)$ which we found before.

To draw from the posterior predictive distribution we can first draw $\left\{\beta_i, \sigma_i^2\right\}_{i=1}^N$ from $p\left(\beta, \sigma^2|Y\right)$, and then draw $\{\widetilde{y}_i\}_{i=1}^N$ from $N\left(\widetilde{X}\beta_i, \sigma_i^2\right)$. That is, for each draw of β_i, σ_i^2, take a draw from a $N\left(\widetilde{X}\beta_i, \sigma_i^2\right)$.

To draw from $p\left(\beta, \sigma^2|Y\right)$, first compute $\widehat{\beta}$ and $(X'X)^{-1}$. Second, compute $e'e$. Third, draw σ^2 from the inverse gamma distribution. Finally, draw β from the normal distribution.

We can also determine the posterior predective density analytically. First, consider $p\left(\widetilde{y}|\sigma^2, y\right)$. It can be shown that this is a normal distribution, and its mean can be found as follows

$$\begin{aligned} E\left[\widetilde{y}|\sigma^2, Y\right] &= E\left(E\left[\widetilde{y}|\beta, \sigma^2, y\right]|\sigma^2, Y\right) \\ &= E\left[\widetilde{X}\beta|\sigma^2, Y\right] \\ &= \widetilde{X}\widehat{\beta} \end{aligned}$$

The variance can be found as

$$\begin{aligned} var\left(\widetilde{y}|\sigma^2, Y\right) &= E\left[var\left(\widetilde{y}|\beta, \sigma^2, y\right)|\sigma^2, Y\right] + var\left[E\left\{var\left(\widetilde{y}|\beta, \sigma^2, y\right)\right\}|\sigma^2, Y\right] \\ &= E\left[\sigma^2 I|\sigma^2, Y\right] + var\left[\widetilde{X}\beta|\sigma^2, Y\right] \\ &= \left(I + \widetilde{X}\left(X'X\right)^{-1}\widetilde{X}'\right)\sigma^2 \end{aligned}$$

Therefore $\widetilde{y}|\sigma^2, Y\tilde{}N\left(\widetilde{X}\widehat{\beta}, \left(I + \widetilde{X}\left(X'X\right)^{-1}\widetilde{X}'\right)\sigma^2\right)$. If we average out σ^2, it can be shown that $\widetilde{y}|Y$ is distributed multivariate t with center $\widetilde{X}\widehat{\beta}$, scale matrix $\frac{e'e}{T-k}\left(I + \widetilde{X}\left(X'X\right)^{-1}\widetilde{X}'\right)$ and $T-k$ degrees of freedom.