STA 250: Statistics
Lab 5

This lab overviews prediction for the two-parameter normal model with conjugate prior

Sampling from a $\mathsf{N}\chi^{-2}(m, k, r, s^2)$

We saw in class that $(W, V) \sim \mathsf{N}\chi^{-2}(m, k, r, s^2)$ if and only if $rs^2/V \sim \chi^2(r)$ and $W|(V = v) \sim \mathsf{Normal}(m, v/k)$. This makes it easy to draw samples from any $\mathsf{N}\chi^{-2}(m, k, r, s^2)$ pdf. For example, to draw samples from $\mathsf{N}\chi^{-2}(0, 1, 3, 1)$ we can run the following code:

```
n.samp <- 1e4
chisq.samp <- rgamma(n.samp, 3/2, 1/2)
V.samp <- 3 / chisq.samp                  ## note: r * s^2 = 3
W.samp <- rnorm(n.samp, 0, sqrt(V.samp)) ## note: cond var = V.samp / 1 = V.samp
```

It will be interesting to get a visual verification of the other result we learned: $\sqrt{k}(W - m)/s \sim t(r)$. See the following code and the resulting plot:

```
## note: sqrt(1) * (W - 0) / 1 = W
## to compare against t(3) pdf
##
w.range <- qt(c(1e-2/2, 1 - 1e-2/2), 3)
w.breaks <- c(-1e16, seq(w.range[1], w.range[2], length = 20), 1e16)
hist(W.samp, freq = FALSE, border = "white", col = "gray",
     xlim = w.range, breaks = w.breaks, ann = FALSE)
w.grid <- seq(w.range[1], w.range[2], .01)
lines(w.grid, dt(w.grid, 3))
title(xlab = "W.samp", ylab = "Density")
```

Task 1. For the NSW study, with a $\mathsf{N}\chi^{-2}(0, 4, 2, 9000^2)$ prior on (μ, σ^2), we obtained a $\mathsf{N}\chi^{-2}(4163.547, 189, 187, 8878.862^2)$ posterior pdf. Draw 10,000 samples of (μ, σ^2) from this pdf and verify that the posterior pdf of $\sqrt{189}(\mu - 4163.547)/8878.862$ matches $t(187)$.

Sampling future observation

Suppose data $X = (X_1, \cdots, X_n)$ and a future observable X^* are modeled as: $X_1, \cdots, X_n, X^* \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu, \sigma^2)$, with (μ, σ^2) assigned a $\mathsf{N}\chi^{-2}(m, k, r, s^2)$ distribution. In order to draw samples of X^* from the posterior predictive given $X = x$, we could adopt the strategy we learned last week: first draw a $(\mu, \sigma^2) \sim \xi(\mu, \sigma^2|x)$ and then draw an X^* from the corresponding $\mathsf{Normal}(\mu, \sigma^2)$. Again work with $\xi(\mu, \sigma^2) = \mathsf{N}\chi^{-2}(0, 4, 2, 9000^2)$ so that $\xi(\mu, \sigma^2|x) = \mathsf{N}\chi^{-2}(\tilde{m}, \tilde{k}, \tilde{r}, \tilde{s}^2)$ with $\tilde{m} = 4163.547$, $\tilde{k} = 189$, $\tilde{r} = 187$ and $\tilde{s}^2 = 8878.862^2$.

Task 2. For the NSW study, draw 10,000 samples of a future observable from its posterior predictive pdf.

TASK 3. Give a visual verification of the result that $\frac{X^*-\tilde{m}}{\tilde{s}\sqrt{1+1/\tilde{k}}} \sim t(\tilde{r})$.

Sampling two future observables

Extend the above setting to include two future observables X^* and Y^* which are modeled along with data $X = (X_1, \cdots, X_n)$ as: $X_1, \cdots, X_n, X^*, Y^* \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu, \sigma^2)$ again with $\xi(\mu, \sigma^2) = \mathsf{N}\chi^{-2}(0, 4, 2, 9000^2)$. We can draw a sample of (X^*, Y^*) from their joint posterior predictive by first drawing a (μ, σ^2) from $\xi(\mu, \sigma^2|x) = \mathsf{N}\chi^{-2}(\tilde{m} = 4163.547, \tilde{k} = 189, \tilde{r} = 187, \tilde{s}^2 = 8878.862^2)$ and then drawing X^* and Y^* from the corresponding $\mathsf{Normal}(\mu, \sigma^2)$.

TASK 4. For the NSW study, draw 10,000 samples of (X^*, Y^*) pairs.

TASK 5. Give a visual verification of the result that $\frac{X^*-Y^*}{\tilde{s}\sqrt{2+1/\tilde{k}}} \sim t(\tilde{r})$.