PROBLEM SET 1

AAEC 6564
INSTRUCTOR: KLAUS MOELTNER

General Instructions

Please type everything in LaTeX (including all Math) and hand in a pdf file. For problems involving Matlab, answer questions in LaTeX, and attach your script, log file, and any graphs to your main pdf file.

If you're ambitious, you can export figures from Matlab in .eps format (as shown in Matlab script `mod1s1a`), and load them directly into your LaTeX file, using something like

```
\begin{figure}[!ht]
\centering
\includegraphics[height=4in]{NameofYourFigure}
\caption{Example}
\label{figure1}
\end{figure}
```

If you do this, make sure the

```
\usepackage{epstopdf}
```

is called at the beginning of your LaTeX file (as it is in this example).

If you're extremely ambitious, you can copy your numeric output (log file) into Excel, add variable names, etc, convert them to LaTeX (via the downloadable "Excel2LaTeX" add-in), and import them directly into your LaTeX file. Detailed instructions for this are given at http://faculty.agecon.vt.edu/LaTeX.html.

Question 1

Consider the binomial(n, θ) likelihood and the beta(α, β) prior for θ we used in our class example. Show that the posterior distribution for θ is indeed another beta with parameters $\alpha + y, \beta + n - y$, where y is the number of successes in the n binomial trials.

Question 2 (Matlab)

In this problem you will show that the posterior results for the normal linear regression model with *conjugate* priors can also be derived via Gibbs Sampling. Please make sure to add your full script (including all functions) to your homework - hard copy is fine. Also make sure to hand in all graphs, either as hard copy or as embedded image in your LaTeX / pdf file.

PART A

(1) Re-run script `mod1s2a` and print the output log.

(2) Open script `mod2s1a` and save it as `ps1q2`. Modify the script as follows:

   (a) Change the folder designations for the log file and the "save" command as needed.
   (b) Make sure r1 (burn-ins) and r2 (keepers) are set to 5000 and 10,000, respectively (else change them to these settings).

(3) Open the function `gs_normal_independent` and save it as `gsPS1Q2`. Modify the function as follows:

   (a) change the name of the function in line 2 to fit the file name.
   (b) Modify your Gibbs Sampler by using the correct conditional posterior moments for the normal linear regression model with *conjugate* priors (see lecture notes).

(4) Run the model.

(5) Open and print the log file. Verify that the posterior moments for $\boldsymbol{\beta}$ and σ^2 are indeed identical to those produced by the analytical approach. Comment.

(6) Add the following task to your script: Draw 10,000 draws of $\boldsymbol{\beta}$ from its prior, by first drawing 1000 σ^2's, and for each of them 10 $\boldsymbol{\beta}$'s

(7) Plot the prior and posterior for β_3. The full code for your figure should read like this:

```
figure(1)
subplot(2,2,1)
plot(x1,f1,'-b',x2,f2,'-k','LineWidth',1);
title('prior and posterior distributions for $\beta_3$',...
    'interpreter','latex','fontsize',14);
xlabel('$\beta_3$','interpreter','latex','fontsize',12);
ylabel('density');
h=legend('$p\left(\beta_3\right)$','$p\left(\beta_3|y,X\right)$');
set(h,'interpreter','latex','fontsize',12);
set(gca,'Ylim',[0 70]); % limits for your y-axis, use same for all plots
set(gca,'Xlim',[-0.2 1.2]); % limits for your x-axis, use same for all plots
set(gca,'XTick',-5:0.2:5); % tick marks on x-axis, spaced in 0.2 intervals
text(0,40,['n = ' num2str(n)],'fontsize',10);
% position of text relative to x and y axis
text(0,35,['prior variance = ' num2str(V0(end,end))],'fontsize',10);
hold on; %this will allow you to add more subplots to the figure
```

PART B

(1) Save your script as `ps1q2partB`. Change the log and "save" paths accordingly.

(2) Reduce the prior variance for β_3 ONLY to 1e-6 (0.000001). You can do this by inserting the following code after defining $\mathbf{V}_0$:

```
V0(end,end)=1e-6;
```

(3) Begin the figure command with

```
subplot(2,2,2)
```

Drop the

```
hold on;
```

at the end of the figure command. In the figure command, change the text positioning coordinates to

```
text(0.4,40,['n = ' num2str(n)],'fontsize',10);
text(0.4,35,['prior variance = ' num2str(V0(end,end))],'fontsize',10);
```

(4) Re-run the script.

PART C

(1) Save your script as `ps1q2partC`. Change the log and "save" paths accordingly.
(2) Restore the original prior for β_3 , but reduce the sample size to 1000 in the data generation section.
(3) Begin the figure command with

```
subplot(2,2,3)
```

(4) Use the same text positioning coordinates as for the very first subplot.
(5) Re-run the script.

PART D

(1) Save your script as `ps1q2partD`. Change the log and "save" paths accordingly.
(2) Reduce the sample size further to 100 in the data generation section.
(3) Begin the figure command with

```
subplot(2,2,4)
```

(4) Use the same text positioning coordinates as for the very first subplot.
(5) Re-run the script.

For each new version, a new subplot should be added to your original figure. Compare these plots and ***comment on the effect of prior variance and sample size on the posterior***. Print your figure and add it to your homework. HINT: Use "File/ Print preview" from the figure window menu to get some nice formatting options - make sure the legends and added text show nicely in each sub-plot)

Question 3 (Matlab)

Consider data set `CTsales.raw`, which was created in STATA and saved as a delimited text file that is easy to read into Matlab. The data contain 7471 observations on residential home sales along the Connecticut coast between 2002 and 2014.

You will use the conjugate normal model to analyse these data. You can use your Gibbs Sampler from the previous question.

(1) Script `ps1q3` will get you started - just complete all the "Your turn" sections & adjust the file designations to fit your environment.

(2) Comment on the plot of marginal effects - are these effects unambiguously positive, negative, or does the posterior reach to both sides of zero? What are the (approximate) ranges of the posteriors? Are these effects as expected?

(3) Look at your log file (you can open it in Word and change your page margins to "narrow" so everything fits nicely). Extract the four marginal effects considered in the plot for the OLS results and paste them into an Excel sheet.

Now extract the posterior means and the "p> 0" statistics for the same four effects from the Bayesian output, and copy them to the Excel table. Make this small table look nice, print it, and add it to your homework hand-in.

(4) Comment on these effects - how do the Bayesian posterior means compare to the OLS estimates? Give a reason why they are (or are not) similar.

(5) Comment on the "p> 0" statistics and relate them to your graphical analysis (plot).