**BOSTON UNIVERSITY QUESTROM SCHOOL OF BUSINESS**

**MF 840 - Spring 2021**

**Eric Jacquier**

**Final Exam**

**Due Wednesday May 5nd at 3:00pm on Gradescope**

* This exam is "open everything", to do individually.
* You can not communicate with others.
* You can use this WORD documents to put your answers exactly in the space provided. Then upload your exam on Gradescope.

**Problem 1: Short Odds Ratio questions**

1) In the Bayes Factor lecture note, consider the proof of [3] p. 15. Prove the second equality at the top of P. 15, where I asked you “make sure that the bottom term is simply …”. Simply prove that the terms inside the denominator exponential are equal across the second equal sign. Your proof must be handwritten and copied in the space below.

You can hand write your proof and copy a pic. of it in this space.

b) Use (3) P.15 for this question. You use a Normal Prior on μ with T0=10. You test H0: μ = μ0 in the model y = μ + ε, ε ∼ N(0,σ). Your computed z-statistic is just about ready to reject the null, with a p-value of exactly 5%. Your Bayes Factor agrees with the classical test. What is the sample size T?

**c)** Bayes factors are difficult to compute but we can get a good idea of the null vs the alternative by comparing the prior and the posterior densities of the parameter. For example, it is clear from this plot that the Null hypothesis H0: β =1 is contradicted by the data. The plot shows how the data moved the posterior to a different location from the prior.

Answer True or False and explain exactly what is True or False and why.

****

**Answer in the space provided here**

**Problem 2: Precision of MCMC estimates**

You simulated the posterior density of β|D with a preliminary number of draws S=2000. The sample mean and standard deviation of your draws is 0.84 and 0.38. You want to report the 0.84 as your MC estimates of E(β|D). You want to make sure that reported digits are *totally* accurate.

1) Assuming independent MC draws, how many MC draws should you have to report 0.84 with total accuracy?

2) Oops ! You just realized that your Gibbs MCMC simulation scheme has autocorrelation, ρ1 = 0.8, ρ2 = 0.5, ρ3 = 0.2. Write the theoretical formula you use for this question.

* What is your new estimate of the standard deviation of your MC estimate of the posterior mean?
* How many MC draws should you now need for the reported sample mean 0.84 to be totally accurate?
* What is your RNE?

**Problem 3: Savage Density Ratio Method**

You want to test the performance of the DFABC fund, the monthly returns are in file fund-mon.csv. The Market excess return and. the risk-free rate are in file FF-3fac-mon.csv. You use the period 2013-2020, T=96.

You estimate the standard market model regression, of course in excess returns form.

RDFABC,t = α + β RM,t + εt, εt ~N (0,σ) [1]

where RDFABC and RM are the excess returns on the market and the fund.

You are interested whether the Fund had (or not) an abnormal Jensen’s α over the 10 years.

In questions, 3)4), answer must be handwritten, show your computation and the exact AZ formula number you use.

1) Give the OLS estimate, t-statistic, p-value, for the annualized α and for β.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Dev. | t-stat | p-value |
| Monthly α |  |  |  |  |
| β |  |  | xxx |  |

2) you want to compute the posterior odds BF0/1 of the restricted vs unrestricted model. You use the Savage Density ratio method and the usual standard conjugate Normal G-prior – Inverted Gamma priors. Specifically the vector **β** = (α,β) contains the intercept α and the slope β. You only care about the Jensen’s α but need to specify a prior for the entire vector of course.

Your prior is: ***β*** | σ ∼N ( (0,1) , σ2g(X’X)-1) with g = 6. For σ, it is the standard ItG ∼ (ν0 = 12 , ν0 s02).

Choose s0 to be equal to the estimate of the noise standard deviation of regression [1] for 2011-2012, in the spirit of Bayesian updating. What is your s0?

s0 =

3) Write the theoretical formula of the **exact univariate prior density for p(**α**)** the intercept coefficient as a function of the prior hyper-parameters. You may want to refer to a specific element of the (X’X)-1 matrix such as {(X’X)-1}ij. Use AZ’s formulas to write the normalization constants. Hint: First write the bivariate prior p(**β**).

4) Write the **exact univariate posterior density** of p(α |D), the intercept, as a function of the relevant sufficient statistics, again probably starting with the bivariate **β|D**. Be sure to specify every parameter, degrees of freedom, posterior covariance matrix.

5) On a figure plot exactly both exact densities together for a relevant range of α. Mark the point on the X axis where the null H0 is.

Your Figure here

6) Compute the Bayes Factor by the Savage density ratio method.

Numerator =

Denominator =

BF0/1 =

**Problem 4: Gibbs Sampling for the non-conjugate regression**

This looks long because I walk you step by step. Pages 60-62 in Gary-Koop show us what happens if we do not use a conjugate prior. He uses the “underbar” notation for priors, we have been using the subscript 0 notation mainly. For the regression noise, he uses the precision h rather than the σ = 1/. But as we say, Page 7, Bayes 2, h is Gamma distributed, then σ is inverted Gamma. So it’s all the same.

First you will rewrite his proofs with our notation and σ. Let’s walk step by step.

**1)** Note that the prior p(β) in 4.1 is not function of σ. It actually makes things more complicated. We will use our A instead of his . We can write GK’s 4.2 with σ instead of h as p(σ) ∼ ITG(ν0 , s02).

Given that, rewrite his 4.3 as p(β,σ | D):

(4.3) p( β , σ | D) =

Now rewrite his posterior covariance matrix and posterior mean of (β | σ , D)

(4.4) V(β | σ , D) =

(4.5) E(β | σ , D) =

2) On Page 9 of our conjugate proof, we wrote the two terms in the exponential where β appears. What is the big difference with the two key terms in the middle of GK P. 61?

Answer

Look at the posterior covariance matrix and posterior mean of β in 4.4 and 4.5. What is the major difference with the standard conjugate results posterior means and covariance page 10 and 11?

Answer

**3)** So we did find p(β | σ, D) in 4.3. Why can’t we just write the rest of 4.3 as p(σ |D) by definition?

Of course we can, but it is a complicated unknown distribution. Write p(σ |D) in 4.3, after removing the part which is p(β | σ ,D). There will be a proportional sign as we do not know this distribution. Use the fact that





Only Q will be left, and Q is in terms of our parameters:

Q =

The key is to **not forget any term that contains σ.** You get

p(σ | D) ∝

Boo. Ugly. We could do Metropolis on this, but there is much simpler, Gibbs!

**4)** Look at 4.3 and consider it as a function of (σ | β, D). That is, we are now looking at p(β, σ |D) as p(σ | β, D) p(β | D). Remove anything that does not have σ. Then what’s left is p(σ | β, D). It should be our version of GK’s equation just before 4.8

(4.8.0) p(σ | β , D) ∝

What distribution do you recognize? With what posterior ν1 and posterior ν1 s12

(4.9’) ν1 =

(4.10’) ν1 s12 =

**5)** Now let’s do it! We will do this on the regression we just run for Problem 3. For the **normal** prior on **β,** the prior mean will be (0,1). You specify a prior covariance matrix A-1 with the criteria.

1) Intercept (α) and slope (β) are independent

2) 95% chances that abnormal performance (α) is between -2% and + 2% **annualized**.

3) 95% chances that β is between ( 0.6, 1.6).

For the prior on σ, you use the same prior as in Problem 3.

What are the parameters **(numbers)** for the posteriors (β | σ ,D) and (σ | β, D)

ν1 =

ν1 s12 =

These will be functions of σ as well, but numbers everywhere you can:

2x2 matrix: V(**β** | σ , D) =

2x1 vector: E(β | σ , D) =

**6)** Now run the Gibbs sampler. Start with an extreme value for α β, such as (10% annnualized, 5). Watch the convergence of the first draws. Run 50000 draws of the sampler. Use Numeff to check for possible autocorrelation.

How many initial draws did you eliminate (burn-in) ?

What was the RNE?

Was there any autocorrelation we should worry about?

Fill in the table for your posterior distributions. Annualize the abnormal performance and the standard deviation of the regression noise.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean | Std.Dev | 25% | 75% |
| α |  |  |  |  |
| β |  |  |  |  |
| σ |  |  |  |  |

**7)** Plot side by side left: Prior and posterior density of α. right: Prior and posterior density of β . For the priors use the exact density, For the posteriors, plot the empirical densities from your Gibbs draws (without the burnin) with the plot(density) method.

Figure here (Left: α , Right: β

8) Diagnostic plots: Side by side: Time Series plot of the first 100 draws of β, ACF plot of the remaining draws of β

Figure here: left: TS plots of first 100 draws right: ACF plot of remaining draws