**BOSTON UNIVERSITY QUESTROM SCHOOL OF BUSINESS**

**MF 840 - Spring 2020**

**Eric Jacquier**

**Take Home Exam**

**Due Saturday May 2nd at 5:00pm by email**

The only missing problem now is the Gibbs sampling exercise

* This exam is "open everything", to do in groups of two, turn in one output per group
* The two members of the group to contribute equally to the problem set. Groups cannot communicate about the problem set.

**Problem 1: Testing and power of the test**

a) Look at the top of Page 15, in the Bayes Factor lecture note; the BF0/1 equality. I ask you to “make sure that the bottom term is simply …”. Prove it. That is prove that the terms inside the denominator exponential are equal across the second equal sign.

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

b) Now we will use Formula (3) P.15 to discuss the discrepancy between standard testing / p-value and posterior odds ratios.

Say that you find a z-statistic of 2.0, your prior on μ is p(μ) ∼ N(μ0 , σ2/T0) with T0 = 5.

* Recompute a table similar to the table on page 9, but using (3) p. 15 as the odds ratio.
* Also compute the table p. 16.
* For what sample sizes does the Odds ratio approach contradict the p-value approach

c) Fill in the table P. 17. Again use T0 =5 for all situations.

Add one column to the table: the sample size where the 5% test agrees exactly with the Odds

Ratio

e) In general we find that classical testing rejects too often relative to odds ratios. A critique is that the odds ratio is always a function of the prior. So the answer could change if we change the prior. Let’s see this.

For the three p-values in the table on P.16, find the prior that would make the odds ratio always agree with the test. Specifically, write the value T0/T of that prior. That is, the prior variance uses a T0 which is a constant ratio of T, find the ratio.

So for what type of priors (more precise, less precise than ?) will the Odds ratios disagree with the z-test/p-value approach, say for T=100.

**Problem 2: Precision of MCMC estimates**

You simulated the posterior density of β|D with a preliminary number of draws S=1000. The sample mean and standard deviation of your draws is 0.92 and 0.46. You will report the 0.92 and 0.46 as your MC estimates of E(β|D) and V(β|D). You want to make sure that reported digits are accurate. You assume asymptotic conditions given the large sample size. By “accuracy” we mean that a 95% confidence interval does not affect the last reported digit (the 2 and the 6)

a) Assuming independent MC draws, how many MC draws should you have to report 0.92 with total accuracy, to report 0.46 with total accuracy?

b) Oops ! You just realized that your MC simulation scheme has autocorrelation, ρ1 = 0.7, ρ2 = 0.4. 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 mean 0.92 to be accurate? What is

your RNE?

**Problem 3: Savage Density Ratio Method**

The file FF-3fac-mon.csv contains the monthly Market excess return and Fama-French Factors returns until 2019. Use only the last 5 years, 60 observations, from 2015 to 2019.

a)You estimate the market model regression, of the size factor (smb) Small-Minus-Big on the excess return on the market excess return (name is rmrf):

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

where RM is the excess return on the market model. Because the smb factor is a long short, you do not need to subtract the risk free rate. You want to test if the **market beta of the smb factor is zero**.

* Give the OLS estimate, t-statistic, p-value, for β.

b) you want to compute the posterior odds BF0/1 of the restricted vs unrestricted model. You will use the Savage Density ratio methods. You use standard conjugate Normal – InvertedGamma priors. Specifically the vector **β** = (α,β) contains the intercept α and the slope β. You only care about the slope but need to specify a prior for the vector.

Your prior is: ***β*** | σ ∼ ( (0,0) , σ2g(X’X)-1) with g = 5, That is a g-prior with mean vector 0 and g=5.

Your prior for σ is the standard ItG ∼ (ν0 = 4 , ν0 s02).

* Choose s0 to be equal to the sample estimate of the standard deviation of smb for 2014. This is in the spirit of Bayesian updating. What is your s0?
* Write the formula of the **exact univariate prior density for p(**β**)** the slope coefficient. You may want to refer to a specific element of the (X’X)-1 matrix. Use AZ’s formulas to write the normalization constants.
* Write the **exact univariate posterior density** of p(β|D), the slope coefficient. Be sure to specify every parameter, degrees of freedom, posterior covariance matrix.
* On a figure plot exactly both densities together. Mark the point on the X axis where the null H0 is (β=0)
* Compute the Bayes Factor by the Savage density ratio method.

**Problem 4: Savage Density Ratio Method**

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 standard devition σ = 1/. But as we say, Page 7, Bayes 2, h is Gamma distributed, then σ is inverted Gamma. So there is no difference really.

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

**a)** Note that p(β) in 4.1 is not function of σ. We will use 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) =

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

**b)** 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:

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!

**c)** 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. Write what’s left, it 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 =

**d)** 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 (0,0). That is, you expect the msb portfolio will have no abnormal performance (intercept α). And, since msb is a long short portfolio, you expect its β to be zero.

You specify a prior covariance matrix A-1 with the following criteria.

1) Intercept (α) and slope are independent

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

3) With a normal, you want 95% chances that the β is between (-0.4,-0.4).

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

**d1)** 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) =

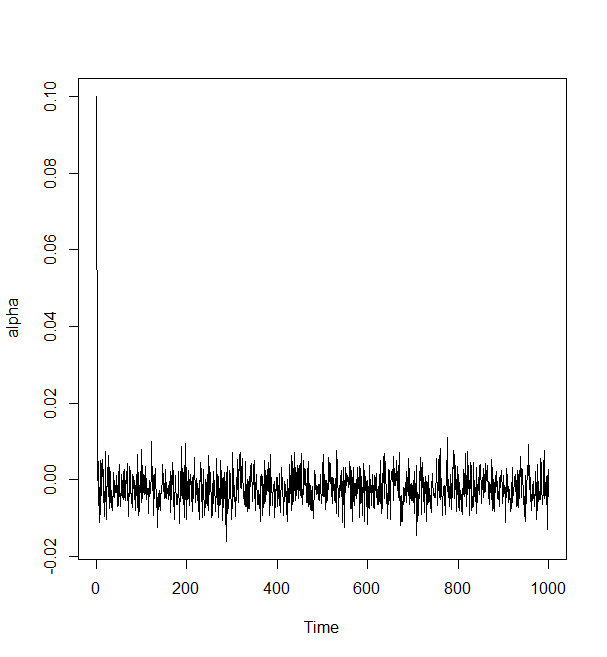
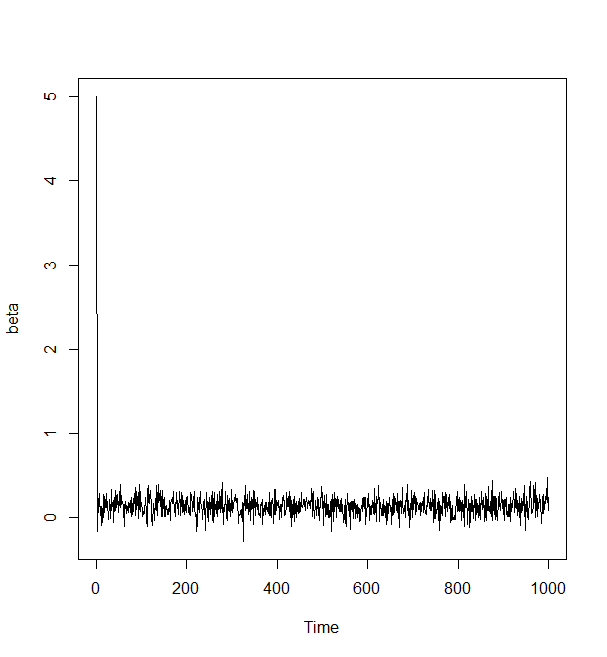
**d2)** Now run a Gibbs sampler. Start with an extreme value for β, such as +5. Watch the convergence of the first draws. Run 50000 draws of the sampler. Use Numeff to check for possible autocorrelation.

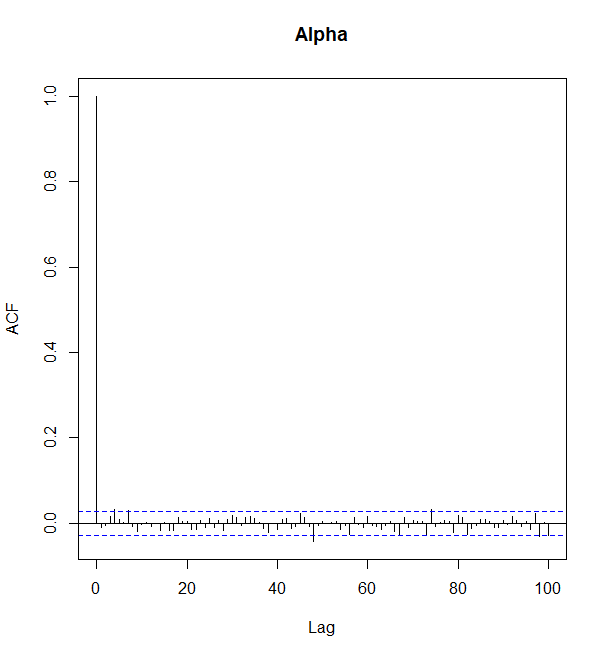
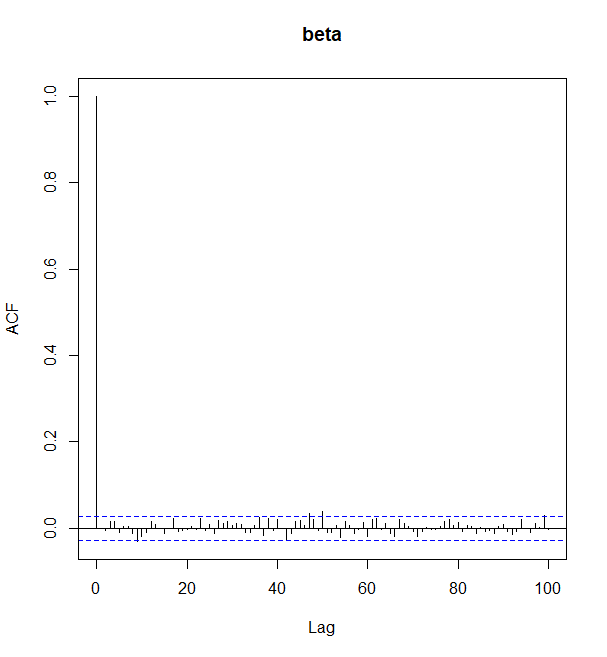
What was the RNE?

Was there any autocorrelation we should worry about?

numEff(alpha): std: 1.649828e-05; RNE: 0.7788013

numEff(beta): std: 0.0004962414; RNE: 1.058429

The autocorrelation could be ignored, the first draw drops down immediately.

**d3)** 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 | -0.002181991 | 0.004180329 | -0.0049892234 | 0.0006019328 |
| β | 0.136621 | 0.1078567 | 0.06586814 | 0.20749643 |
| σ | 0.03428518 | 0.03428518 | 0.03207423 | 0.03622378 |

**e3)** Make two figures. 1: Prior and posterior density of β (the smb beta). 2: Prior and posterior density of σ. For the priors use the exact density, use the AZ formula to plot the Inverted Gama density. For the posteriors, plot the empirical densities from your Gibbs draws (remove the first 100) with the plot(density) method.

