AMS 206B – Exam 2
Mickey Warner

We fit a data set with $I = 5$ time series and $T = 300$ observations each. The AR(1) model is given by

$$
\begin{aligned}
y_{t,i} &= \phi_i y_{t-1,i} + \epsilon_{t,i}, \quad \epsilon_{t,i} \sim N(0, \nu), \\
\phi_i | \phi &\overset{iid}{\sim} N(\phi, \tau^2), \\
p(\phi, \nu) &\propto 1/\nu
\end{aligned}
$$

with $\tau^2 = 0.1$. The full likelihood is given by the I products of

$$
p(y_{1:T,i} | \phi_i, \nu) = \frac{(1 - \phi_i^2)^{1/2}}{(2\pi\nu)^{T/2}} \exp\left\{ -\frac{1}{2\nu} \left[y_{1,i}^2 (1 - \phi_i^2) + \sum_{t=2}^{T} (y_{t,i} - \phi_i y_{t-1,i})^2 \right] \right\}.
$$

This expression can be simplified by approximating it with the conditional likelihood given by

$$
p(y_{2:T,i} | \phi_i, \nu, y_{1,i}) = \frac{1}{(2\pi\nu)^{T/2}} \exp\left\{ -\frac{1}{2\nu} \sum_{t=2}^{T} (y_{t,i} - \phi_i y_{t-1,i})^2 \right\}.
$$

The approximation is more accurate as T increases since the first observation $y_{1,i}$ places less and less of a role in the estimation process. We sample from the posterior using both likelihoods and compare these results and approaches. We also look at how choices for τ^2 affect our posterior results.

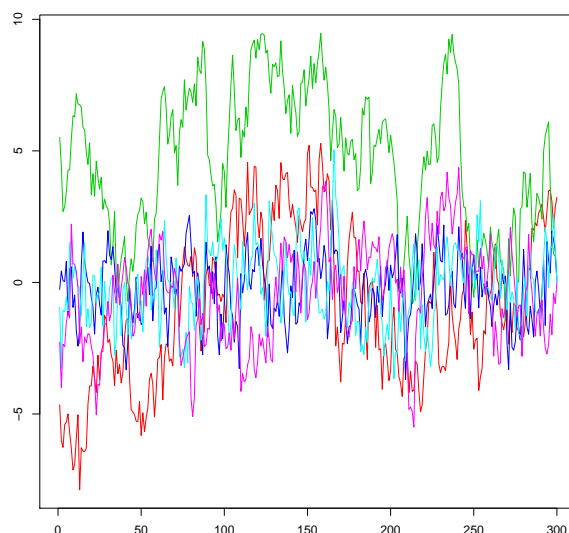A plot of the data is given in Figure 1.

Figure 1: The $I = 5$ time series, each with $T = 300$ observations.

1 Posterior based on conditional likelihood

Denote $\mathbf{y}$ as the time series data. We obtain samples using Gibbs updates from the full conditionals of each parameter. These are given by:

$$
\nu|\phi_1,\ldots,\phi_I,\phi,\mathbf{y} \sim IG\left(\frac{IT}{2},\frac{1}{2}\sum_{i=1}^{I}\left[\sum_{t=2}^{T}(y_{t,i}-\phi_i y_{t-1,i})^2\right]\right)
$$

$$
\phi_i|\phi_{-i},\phi,\nu,\mathbf{y} \sim N\left(\frac{\tau^2\sum_{t=2}^{T}y_{t,i}y_{t-1,i}+\nu\phi}{\tau^2\sum_{t=2}^{T}y_{t-1,i}^2+\nu},\frac{\tau^2\nu}{\tau^2\sum_{t=2}^{T}y_{t-1,i}^2+\nu}\right)
$$

$$
\phi|\phi_1,\ldots,\phi_I,\nu,\mathbf{y} \sim N\left(\frac{1}{I}\sum_{i=1}^{I}\phi_i,\frac{\tau^2}{I}\right),
$$

where ϕ_{-i} denotes the vector of $(\phi_1,\ldots,\phi_I)$ excepting the ith element. These are easy to sample from using standard methods. After a burn-in of 5000 draws, we keep 20000 samples to calculate density estimates and other functionals. Density estimates of the posterior samples are shown in Figure 2 and a numerical summary is given in Table 1.

2 Posterior based on full likelihood

When using the full likelihood, a small adjustment is made to the full conditional of ν and Metropolis steps are used to draw ϕ_i. This leaves us with closed-form expressions for ν and ϕ, but not for ϕ_i:

$$
\nu|\phi_1,\ldots,\phi_I,\phi,\mathbf{y} \sim IG\left(\frac{IT}{2},\frac{1}{2}\sum_{i=1}^{I}\left[y_{1,i}^2(1-\phi_i^2)+\sum_{t=2}^{T}(y_{t,i}-\phi_i y_{t-1,i})^2\right]\right)
$$

$$
p(\phi_i|\phi_{-i},\phi,\nu,\mathbf{y}) \propto (1-\phi_i^2)^{1/2}\exp\left\{-\frac{1}{2\nu}\left[y_{1,i}^2(1-\phi_i^2)+\sum_{t=2}^{T}(y_{t,i}-\phi_i y_{t-1,i})^2\right]\right\}\times
$$

$$
\exp\left\{-\frac{1}{2\tau^2}(\phi_i-\phi)^2\right\}
$$

$$
\phi|\phi_1,\ldots,\phi_I,\nu,\mathbf{y} \sim N\left(\frac{1}{I}\sum_{i=1}^{I}\phi_i,\frac{\tau^2}{I}\right).
$$

This formulation necessarily requires $\phi_i \in [-1,1]$; we define $p(\phi_i,\phi_{-1},\phi,\nu,\mathbf{y})=0$ outside of this interval. We use normal proposal distributions for the Metropolis updates and each are tuned individually to get desirable acceptance rates.

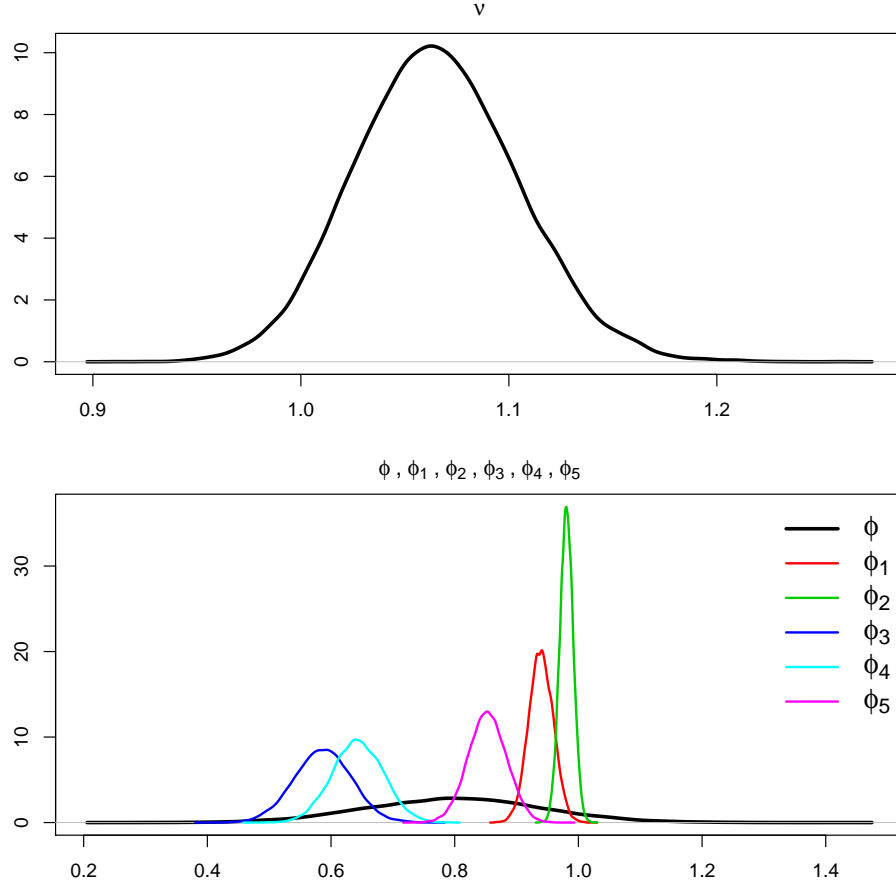Results are shown in Figure 3 and Table 2.

Figure 2: *Conditional likelihood.* Kernel density estimates of the posterior for each parameter (marginally). Top: the posterior for ν. Bottom: posteriors for the group mean ϕ (black) and the individual autoregressive parameters (colored).

	mean	sd	2.5%	97.5%
ν	1.07	0.04	0.99	1.15
ϕ_1	0.94	0.02	0.90	0.98
ϕ_2	0.98	0.01	0.96	1.00
ϕ_3	0.59	0.05	0.49	0.68
ϕ_4	0.64	0.04	0.56	0.72
ϕ_5	0.85	0.03	0.79	0.91
ϕ	0.80	0.14	0.52	1.08

Table 1: *Conditional likelihood.* Summary statistics from the posterior samples. Sample means, standard deviations, and equal-tailed 95% probability intervals are given.
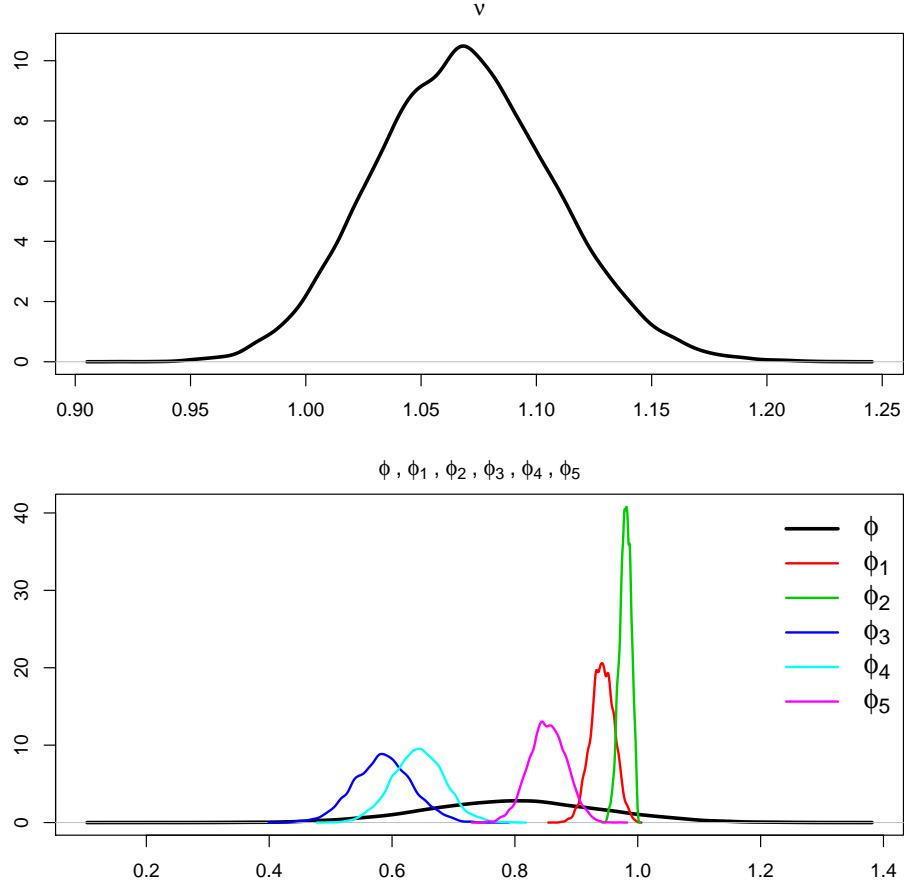
Figure 3: *Full likelihood.* Kernel density estimates of the posterior for each parameter (marginally). Top: the posterior for ν. Bottom: posteriors for the group mean ϕ (black) and the individual autoregressive parameters (colored).

	mean	sd	2.5%	97.5%	ar	effSize
ν	1.07	0.04	1.00	1.15	—	—
ϕ_1	0.94	0.02	0.90	0.98	0.27	3417.27
ϕ_2	0.98	0.01	0.96	1.00	0.26	3171.56
ϕ_3	0.59	0.05	0.49	0.68	0.28	3672.59
ϕ_4	0.64	0.04	0.56	0.73	0.28	3523.27
ϕ_5	0.85	0.03	0.79	0.91	0.28	3762.62
ϕ	0.80	0.14	0.52	1.08	—	—

Table 2: *Full likelihood.* Summary statistics from the posterior samples. Sample means, standard deviations, equal-tailed 95% probability intervals, and Metropolis acceptance rate and effective sample size (when applicable) are given.

3 Comparison between the two approaches

Figures 2 and 3 show very similar results. The only difference (though not seen from the figures) is that the conditional likelihood approach resulted in samples for ϕ_1 and ϕ_2 being greater than 1. This is impossible for the full likelihood approach and also causes some issues with autoregressive models. The summary statistics in both cases are nearly identical.

Since the conditional likelihood approximates the full likelihood, it isn't any surprise that the posteriors in each case are so similar. The main difference then is in the approach itself. Do we prefer an easier way to obtain samples over having a more accurate posterior? Or are we fine dealing with the challenges involved with the Metropolis algorithm?

For this data set, the Metropolis updates performed fairly well. We had no apparent issues in convergence or finding good proposal distributions. One area of concern is the effective sample size: out of 20000 iterations, we had in effect only 3500 samples from the posterior. Should the data set grow (considerably) in size, we may not have the resources to obtain a sufficient number of draws.

4 Prior sensitivity analysis

We repeat both analyses for $\tau^2 = 1$ and $\tau^2 = 10$ (recall the above results are for $\tau^2 = 0.1$). The parameter τ^2 is the prior variance that describes how each ϕ_i differs from the population mean ϕ.

The full conditionals given in Sections 1 and 2 reveal how much of a role τ^2 plays in the posterior. For ϕ, τ^2 interacts with the data only through the number of time series I, thus being more influential to the posterior of ϕ. For each ϕ_i, τ^2 can affect both the mean and variance, but it can also be swamped out by the data more easily. For ν, the affect of τ^2 only manifests itself by changes made to $\phi_1, \ldots, \phi_I$.

How τ^2 affects the posteriors can be seen in the figures on the next two pages. Figures 4 and 5 compare the conditional likelihood approach with the full likelihood approach, respectively. Again, the two posteriors are very similar. Compared to Figures 2 and 3, we see that the only difference is found in ϕ. For $\tau^2 = 1$, the posterior for ϕ has a much higher variance while the other parameters remain the same.

Figures 6 and 7 show posterior density estimates for $\tau^2 = 10$. We again see a wider ϕ (almost unseen in the figures) while $\phi_1, \ldots, \phi_I, \nu$ remain the same.
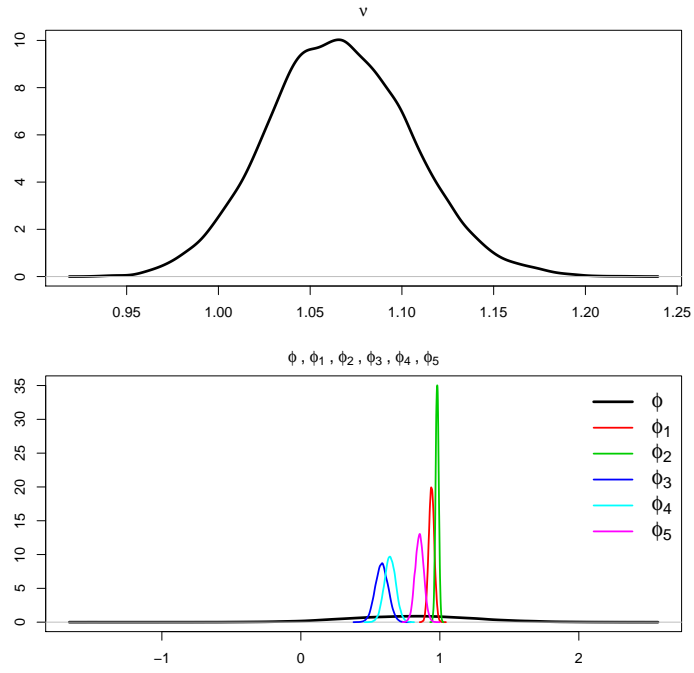
Figure 4: *Conditional likelihood.* Posterior distributions for $\tau^2 = 1$.
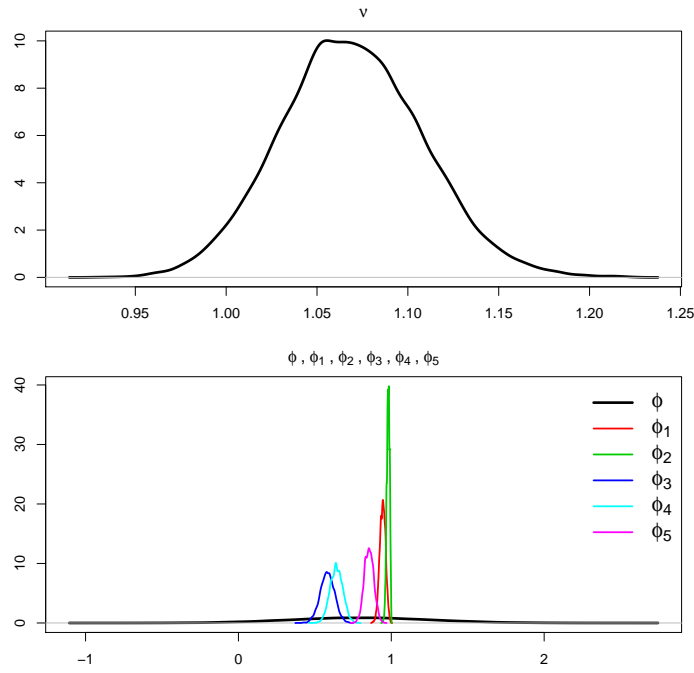

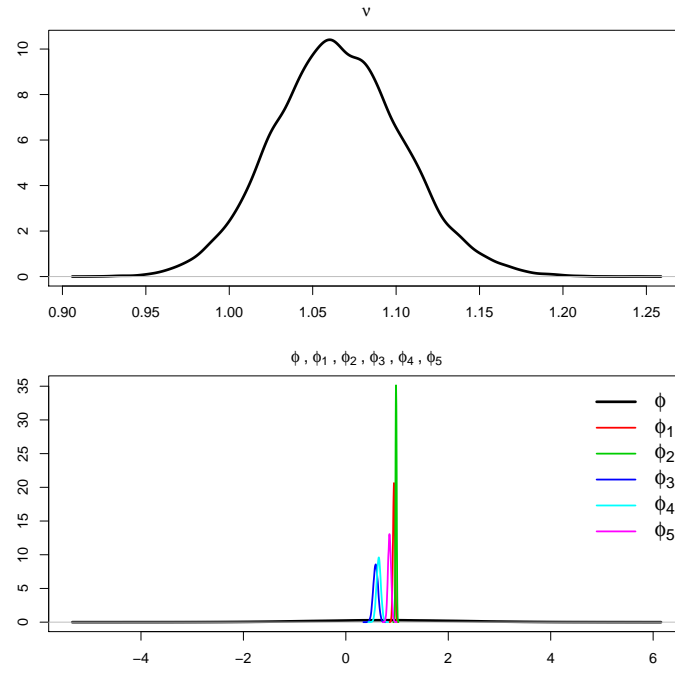Figure 5: *Full likelihood.* Posterior distributions for $\tau^2 = 1$.

6

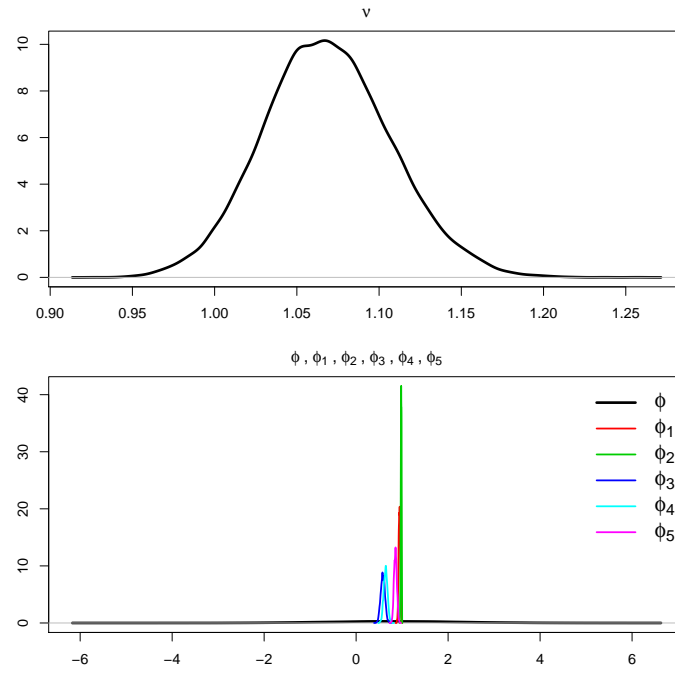Figure 6: *Conditional likelihood.* Posterior distributions for $\tau^2 = 10$.

Figure 7: *Full likelihood.* Posterior distributions for $\tau^2 = 10$.