Streamlining the Practice of Statistics with Prior Predictive Analysis

John Geweke
University of Technology Sydney

Lancaster Lecture
The Statistical Society of Australia
New South Wales Branch
March 23, 2010

Centre for the Study of Choice at UTS

- Academic and applied research in
  - choice modeling
  - structural modeling and econometrics
  - emphasis on ultimately addressing significant public and private sector policy problems.
- Funding from ...
  - Breakthrough contract work with public and private sector clients
  - Grants from several governments and agencies
  - UTS priority investment funds
- Organization strives to combine
  - Efficiency and teamwork of private sector research and consulting
  - Academic freedom of university setting

Speaker

- Working at several interfaces
  - Economics and statistics (= Econometrics)
  - Theory, data and policy
- Recent history
  - Classroom to policy applications
  - Away from solutions in search of problems to problems in search of solutions
  - From US to Australia
- This evening's presentation draws from
  - *Contemporary Bayesian Econometrics and Statistics* (2005, Wiley)
  - *Complete and Incomplete Econometric Models* (2010, Princeton University Press)

Motivation

- Decision problem
  - Design a choice experiment to learn about consumer preferences for a new product
  - Produce a forecast for tax revenues in the next fiscal year
- Data are available, or some aspects of the data are known
- Decide whether one or more existing models are adequate
- Decide whether to invest in a new model
- There are time and budget constraints

Three components of a complete model A

- Prior density
$$p\left(\boldsymbol{\theta}_{A,T} \mid A\right)$$
- Conditional density of observables
$$p\left(\mathbf{y}_T \mid \boldsymbol{\theta}_{A,T}, A\right);$$
    - $L\left(\boldsymbol{\theta}_{A,T}; \mathbf{y}_T^o\right) = p\left(\mathbf{y}_T^o \mid \boldsymbol{\theta}_{A,T}, A\right)$ is the likelihood function.
- Vector of interest density
$$p\left(\boldsymbol{\omega} \mid \mathbf{y}_T^o, \boldsymbol{\theta}_{A,T}, A\right)$$

Simulating from the prior predictive distribution

- A vector of features of the observables

$$\begin{aligned} \mathbf{z}_T &= h(\mathbf{y}_T) \quad \text{(random variable)} \\ \mathbf{z}_T^o &= h(\mathbf{y}_T^o) \quad \text{(observed value)} \end{aligned}$$

- Predictive density is

$$p(\mathbf{z}_T | A) = \int p(\boldsymbol{\theta}_{A,T} \mid A)\, p\left(\mathbf{z}_T \mid \boldsymbol{\theta}_{A,T}, A\right) d\boldsymbol{\theta}_{A,T}.$$

- Simulation
  - $\boldsymbol{\theta}_{A,T}^{(m)} \sim p(\boldsymbol{\theta}_{A,T} \mid A)$
  - $\mathbf{y}_T^{(m)} \sim p\left(\mathbf{y}_T \mid \boldsymbol{\theta}_{A,T}^{(m)}, A\right)$
  - $\mathbf{z}_T^{(m)} = h\left(\mathbf{y}_T^{(m)}\right)$

Prior predictive analysis

- Recall the function

$$\mathbf{z}_T = h(\mathbf{y}_T)$$

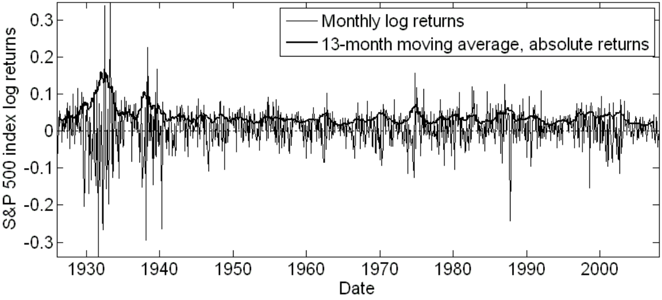
- Compare $p(\mathbf{z}_T | A)$, represented by

$$\mathbf{z}_T^{(m)} \sim p(\mathbf{z}_T \mid A) \quad (m = 1, \ldots, M)$$

  - with

$$\mathbf{z}_T^o = h(\mathbf{y}_T^o)$$

Data

- S&P 500 log returns (monthly)
- 1926:2 - 2007:12 ($T = 983$)
- 

Some observed features

Feature	Data
Return mean $\times$ 100	0.810
Return standard deviation $\times$ 100	5.515
Months in bear markets	350
Largest bear market decline	0.846
Return skewness	-0.437
Return excess kurtosis	8.166
Ratio of range to standard deviation	12.451
Return autocorrelation, lag 1	0.078
Squared return autocorrelation, lag 1	0.242
Squared return autocorrelation, lag 12	0.182
Absolute return long memory parameter	0.693

Some alternative models

- Let y_t denote the monthly return
- Simplest model (known to be deficient in 1920 if not earlier):

$$y_t \overset{iid}{\sim} N\left(\mu, \sigma^2\right)$$

- Genearalized autoregressive conditional heteroscedasticity (GARCH):

$$y_t \sim N\left(\mu, \sigma_t^2\right); \;\; \sigma_t^2 = \alpha_0 + \alpha_1\left(y_{t-1} - \mu\right)^2 + \beta_1 \sigma_{t-1}^2$$

- Stochastic volatility (SV):

$$y_t \sim N\left(\mu, \sigma_t^2\right), \quad \log\sigma_t^2 = \alpha + \delta\log\sigma_{t-1}^2 + \sigma_v v_t, \;\; v_t \overset{iid}{\sim} N(0,1)$$

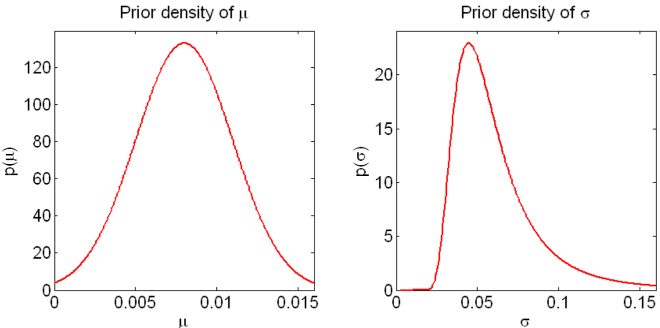
Prior distributions: Mean and variance

- Mean and variance in all three models:

$$\mu \sim N\left(.008, .003^2\right)$$

$$\frac{0.01}{\sigma^2} \sim \chi^2(4) \iff \frac{1}{\sigma^2} \sim \mathrm{Gamma}(2, 200)$$

- The densities:



Prior distribution: Excess kurtosis
(GARCH and stochastic volatility models)

- Excess kurtosis:

κ exponentially distributed with mean 8

- GARCH model:

$$y_t \sim N\left(\mu, \sigma_t^2\right); \quad \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$
$$\kappa = \frac{6\alpha_1^2}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2} \geq 0$$

- Stochastic volatility model:

$$y_t \sim N\left(\mu, \sigma_t^2\right), \quad \log\sigma_t^2 = \alpha + \delta\log\sigma_{t-1}^2 + \sigma_v v_t, \quad v_t \overset{iid}{\sim} N(0,1)$$
$$\kappa = 3\exp\left(\frac{\sigma_v^2}{1-\delta^2}\right) - 3 \geq 0$$

Prior distribution: Volatility persistence
(GARCH and stochastic volatility models)

- Correlation of successive squared returns:

$$\rho \text{ uniformly distributed on } (0,1)$$

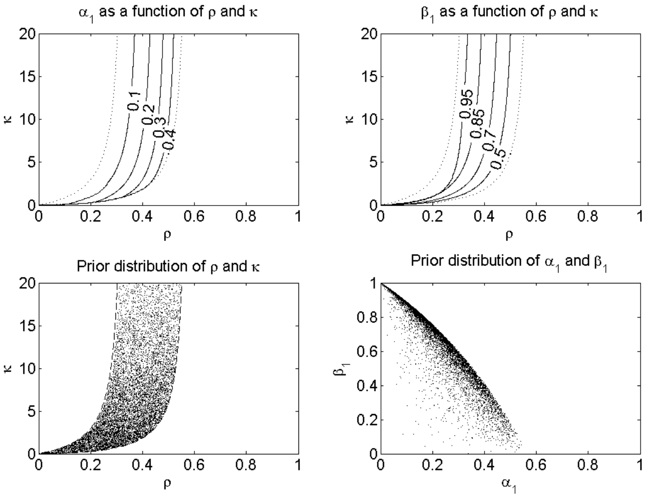
- GARCH model:

$$\begin{aligned} y_t &\sim N\left(\mu, \sigma_t^2\right); \;\; \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\ \rho &= \frac{\alpha_1\left(1-\beta_1^2-\alpha_1\beta_1\right)}{1-\beta_1^2-2\alpha_1\beta_1} \end{aligned}$$

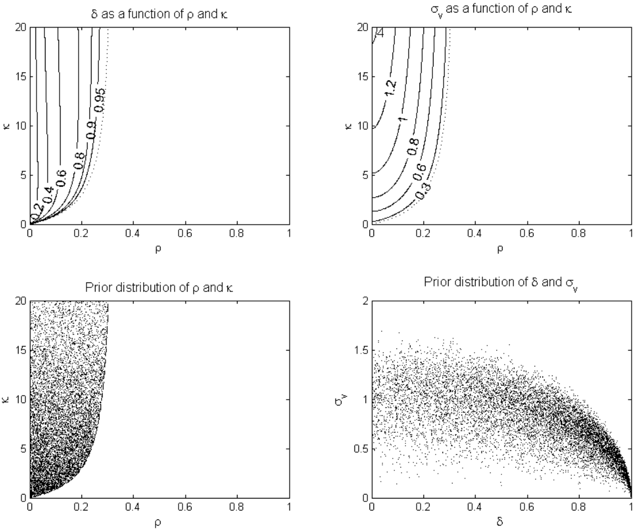
- Stochastic volatility model:

$$\begin{aligned} y_t &\sim N\left(\mu, \sigma_t^2\right), \;\; \log \sigma_t^2 = \alpha + \delta \log \sigma_{t-1}^2 + \sigma_v v_t, \;\; v_t \overset{iid}{\sim} N(0,1) \\ \rho &= \left[\exp\left(\frac{\delta\sigma_v^2}{1-\delta^2}\right)-1\right] \Big/ \left[3\exp\left(\frac{\sigma_v^2}{1-\delta^2}\right)-1\right] \end{aligned}$$

Priors in the GARCH model

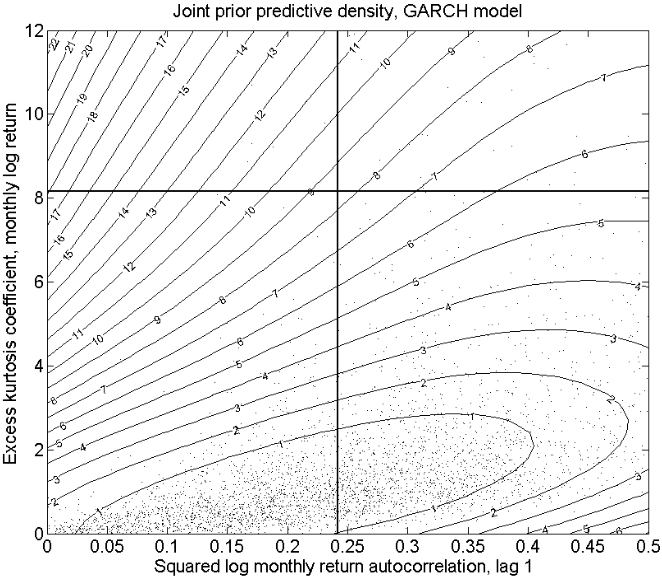


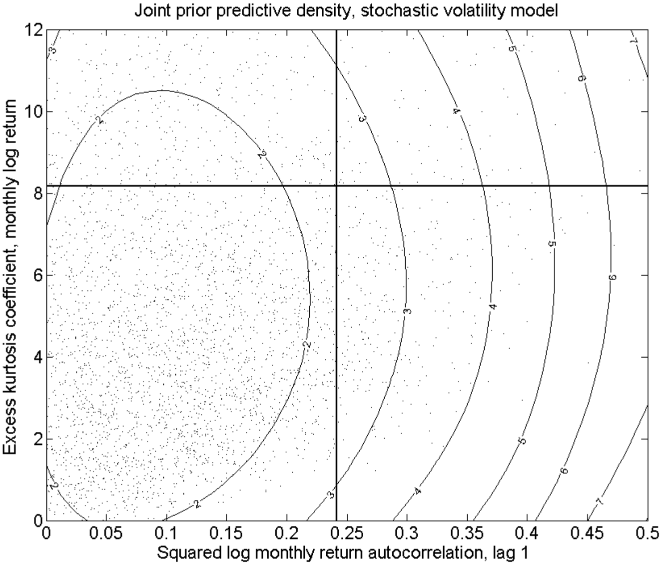
Priors in the stochastic volatility model



Prior predictive analysis: Overview

	Inverse cdf at data		
Feature	Gauss	GARCH	SV
Return mean	0.50	0.50	0.52
Return standard deviation	0.52	0.54	0.51
Months in bear markets	0.49	0.55	0.53
Largest bear market decline	0.91	0.90	0.91
Return skewness	0.00	0.03	0.16
Return excess kurtosis	1.00	0.98	0.74
Ratio of range to stan. dev.	1.00	0.66	0.68
Return autocorrelation, lag 1	0.99	0.96	0.96
Squared return AC, lag 1	1.00	0.51	0.88
Squared return AC, lag 12	1.00	0.91	0.98
$\lvert y_t \rvert$ long memory parameter	1.00	0.97	0.99


Joint prior predictive density, GARCH model
Excess kurtosis coefficient, monthly log return
Squared log monthly return autocorrelation, lag 1
0
2
4
6
8
10
12
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5


Joint prior predictive density, stochastic volatility model
Excess kurtosis coefficient, monthly log return
Squared log monthly return autocorrelation, lag 1
0
2
4
6
8
10
12
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5