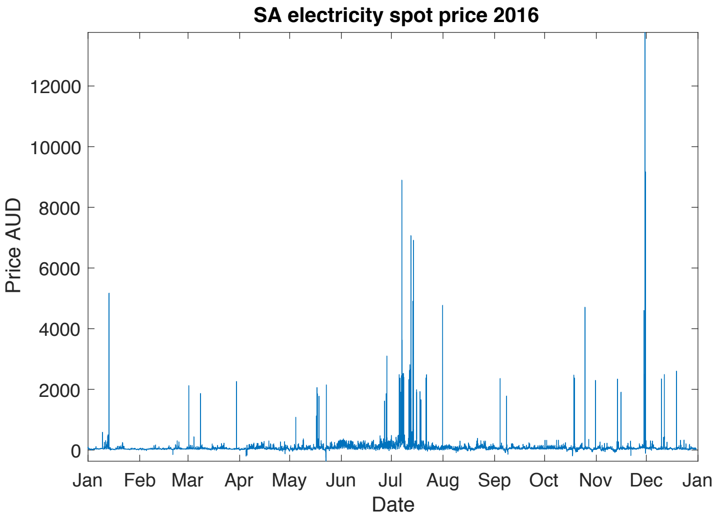
A model for South Australian electricity prices

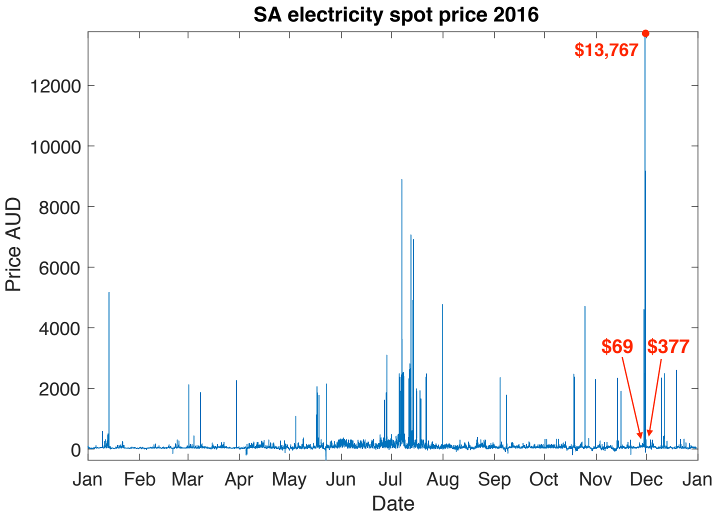
Angus Lewis
Supervised by Prof. Nigel Bean & Dr. Giang Nguyen

ANZIAM SA mini-meeting

South Australia has some of the highest and most volatile *wholesale* electricity prices



South Australia has some of the highest and most volatile *wholesale* electricity prices



Why do we need a model?

- Risk management
- Value contracts
- Value investments
- Describe the behaviour of the market (i.e. the distribution/occurence of spikes)

About the market

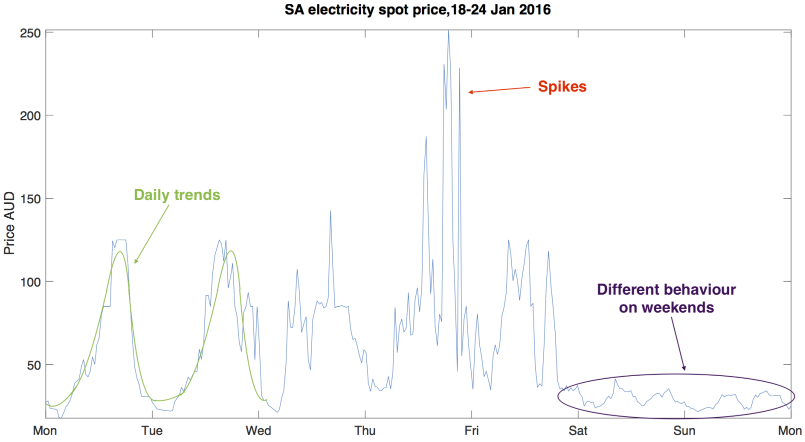
Every 5 minutes AEMO

- Aggregates supply bids - scheduled generators only
- Estimates wind generation
- Estimates demand
- Matches supply and demand
  - Dispatches generators
  - Sets the *dispatch* price
  - Spot price = average of 6 dispatch prices

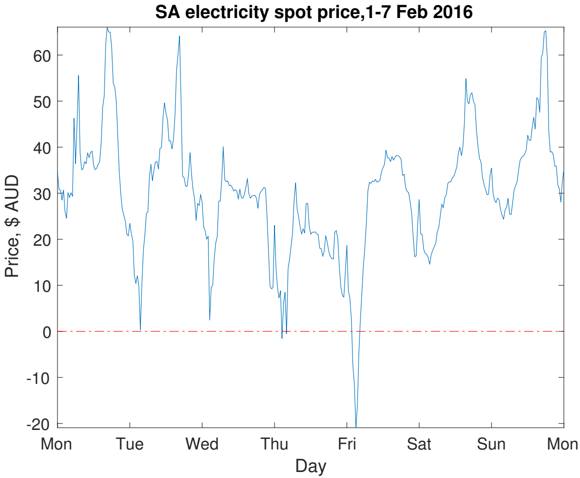
Spikes and drops are caused by a mismatch between supply and demand

- Unexpected demand
- Unexpected wind

Price data - Characteristics



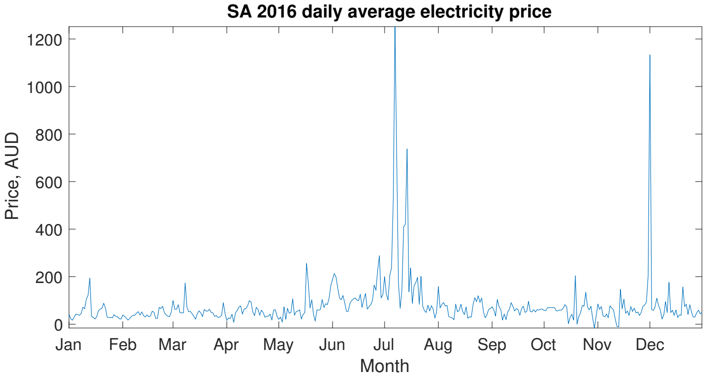
Price data - Characteristics



- Mean reversion - to a trend line
- Negative prices!

Price data - Simplifying the problem

- Model average daily price
  - Justification: Some contracts are valued on daily average prices
  - Drawback: Not all contracts are valued in this way
  - Drawback: Lose a lot of information



The Model

$$\overbrace{P_t}^{\text{price process}} = \overbrace{y_t}^{\text{seasonal component}} + \overbrace{X_t}^{\text{stochastic component}}$$

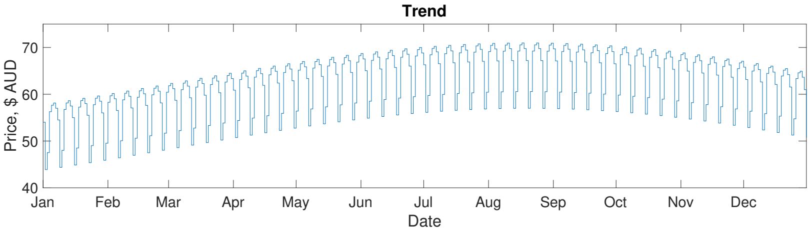
$$y_t = \text{weekly trend} + \text{yearly trend}$$

$$X_t = \text{regime-switching model}$$

The Model - Seasonal components

$$\overbrace{P_t}^{\text{price process}} = \overbrace{y_t}^{\text{seasonal component}} + \overbrace{X_t}^{\text{stochastic component}}$$

$$y_t = \text{weekly trend} + \text{yearly trend}$$



The Model - Stochastic component

Model 1: 3-regime - shifted log-normal spikes and drops

$$X_t = \begin{cases} B_t & \text{when } R_t = 1, \\ S_t & \text{when } R_t = 2, \\ D_t & \text{when } R_t = 3. \end{cases}$$

The Model - Stochastic component

Model 1: 3-regime - shifted log-normal spikes and drops

$$X_t = \begin{cases} B_t & \text{when } R_t = 1, \\ S_t & \text{when } R_t = 2, \\ D_t & \text{when } R_t = 3. \end{cases}$$

- R_t evolves with probabilities $p_{ij} = \mathbb{P}(R_t = j | R_{t-1} = i)$

The Model - Stochastic component

Model 1: 3-regime - shifted log-normal spikes and drops

$$X_t = \begin{cases} B_t & \text{when } R_t = 1, \\ S_t & \text{when } R_t = 2, \\ D_t & \text{when } R_t = 3. \end{cases}$$

- R_t evolves with probabilities $p_{ij} = \mathbb{P}(R_t = j | R_{t-1} = i)$
- $B_t = \alpha + \phi B_{t-1} + \sigma \varepsilon_t$ - AR(1) base regime
  - to capture mean reversion

The Model - Stochastic component

Model 1: 3-regime - shifted log-normal spikes and drops

$$X_t = \begin{cases} B_t & \text{when } R_t = 1, \\ S_t & \text{when } R_t = 2, \\ D_t & \text{when } R_t = 3. \end{cases}$$

- R_t evolves with probabilities $p_{ij} = \mathbb{P}(R_t = j | R_{t-1} = i)$
- $B_t = \alpha + \phi B_{t-1} + \sigma \varepsilon_t$ - AR(1) base regime
  - to capture mean reversion
- $\log(S_t - q_2) \sim \mathcal{N}(\mu_S, \sigma_S^2)$ - Shifted log-normal spikes
  - $q_2 = 75^{th}$ percentile of the data

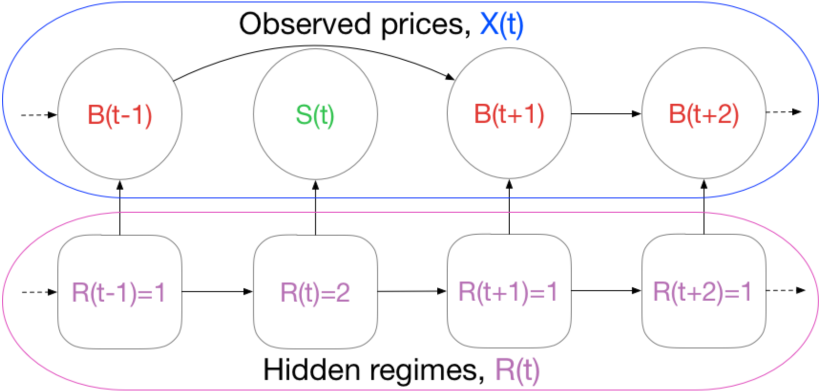
The Model - Stochastic component

Model 1: 3-regime - shifted log-normal spikes and drops

$$X_t = \begin{cases} B_t & \text{when } R_t = 1, \\ S_t & \text{when } R_t = 2, \\ D_t & \text{when } R_t = 3. \end{cases}$$

- R_t evolves with probabilities $p_{ij} = \mathbb{P}(R_t = j | R_{t-1} = i)$
- $B_t = \alpha + \phi B_{t-1} + \sigma \varepsilon_t$ - AR(1) base regime
  - to capture mean reversion
- $\log(S_t - q_2) \sim \mathcal{N}(\mu_S, \sigma_S^2)$ - Shifted log-normal spikes
  - $q_2 = 75^{th}$ percentile of the data
- $\log(q_3 - D_t) \sim \mathcal{N}(\mu_D, \sigma_D^2)$ - Shifted log-normal drops
  - $q_3 = 25^{th}$ percentile of the data

The Model - stochastic component



Inference

We use data-augmented block-wise MCMC for Bayesian inference.

Inference

We use data-augmented block-wise MCMC for Bayesian inference.

Bayesian inference

- Goal:

$$\overbrace{p(\boldsymbol{\theta}|\text{data})}^{\text{posterior}} = \frac{p(\text{data}|\boldsymbol{\theta})\,\overbrace{p(\boldsymbol{\theta})}^{\text{prior}}}{p(\text{data})}$$

- We use uniform (improper) priors

Inference

We use data-augmented block-wise MCMC for Bayesian inference.

MCMC

- Markov chain Monte Carlo - Generate a Markov chain that has the posterior as its stationary distribution

Inference

We use data-augmented block-wise MCMC for Bayesian inference.

Data-augmentation

- Recall:

$$p(\boldsymbol{\theta}|\text{data}) = \frac{\overbrace{p(\text{data}|\boldsymbol{\theta})}^{\text{likelihood}}\, p(\boldsymbol{\theta})}{p(\text{data})}$$

- Likelihood:

$$p(\text{data}|\boldsymbol{\theta}) = \sum_{\boldsymbol{R}} p(\text{data}|\boldsymbol{R}, \boldsymbol{\theta}) p(\boldsymbol{R}|\boldsymbol{\theta})$$

where $\boldsymbol{R} = (R_0, R_1, ..., R_T)$ is a sequence of hidden regimes

- Solution:

$$p(\boldsymbol{R}, \boldsymbol{\theta}|\text{data}) = \frac{\overbrace{p(\text{data}|\boldsymbol{R}, \boldsymbol{\theta}) p(\boldsymbol{R}|\boldsymbol{\theta})}^{\text{augmented likelihood}}\, p(\boldsymbol{\theta})}{p(\text{data})}$$

Inference

We use data-augmented block-wise MCMC for Bayesian inference.

Block-wise aka Metropolis within Gibbs

- The support of the posterior $p(\boldsymbol{R}, \boldsymbol{\theta} | \text{data})$ is large:

$$\mathcal{S} = \underbrace{(-\infty, \infty)^4}_{\alpha,\phi,\mu_S,\mu_D} \times \underbrace{[0, \infty)^6}_{\sigma,\sigma_S,\sigma_D} \times \underbrace{[0, 1]^3}_{p_{ij}} \times \underbrace{\{1, 2, 3\}^T}_{\text{regime sequence}}$$

  where T is the length of the data

$\Longrightarrow$ Difficult to find a proposal that

1. simultaneously updates all parameters and
2. allows the Markov chain to explore the space well

- Solution: partition $\mathcal{S}$ into blocks and update each block iteratively

Inference

We use data-augmented block-wise MCMC for Bayesian inference.

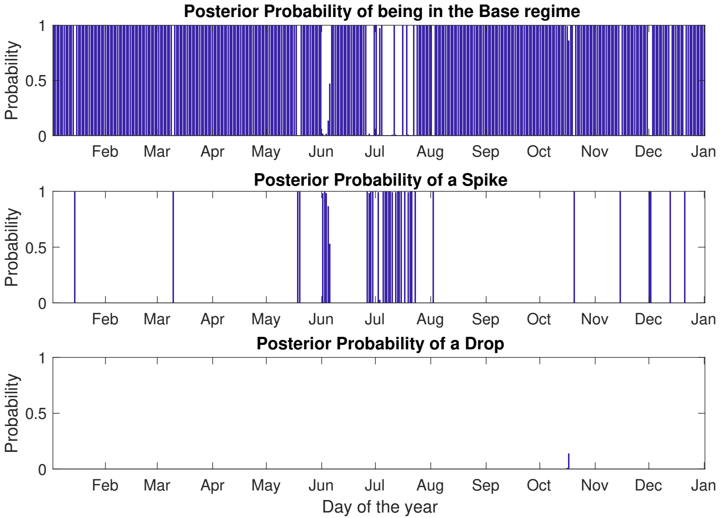
Algorithm: To create an MCMC chain of length N

- For $n = 1, 2, 3, ..., N$:
  - Sample $p_{ij}^{(n)}$ using a Gibbs sampler
  - Data-augmentation: sample $\boldsymbol{R}^{(n)}$
  - Apply Metropolis-Hastings steps to update each of

$$\{\alpha^{(n)}, \phi^{(n)}, \sigma^{(n)}, \mu_D^{(n)}, \sigma_D^{(n)}, \mu_S^{(n)}, \sigma_D^{(n)}\}$$

- Adapt proposals every N^* steps to 'optimise' the algorithm

Results - 3-regimes



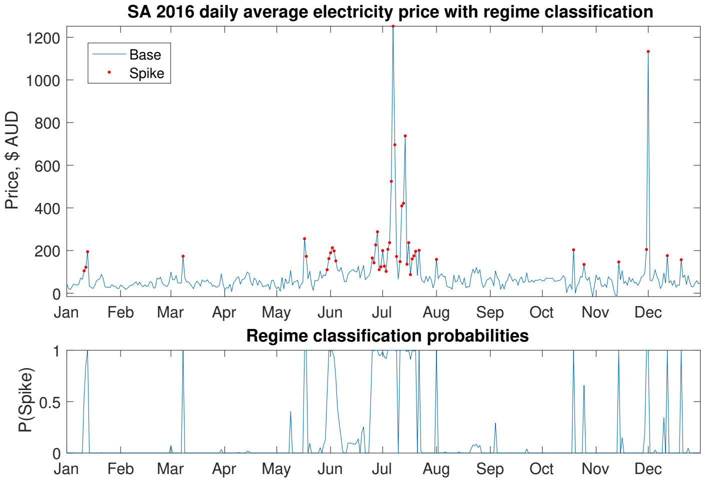
2-regime model

Model 2: 2-regime - shifted log-normal spikes

- $B_t = \alpha + \phi B_{t-1} + \sigma\varepsilon_t$ - AR(1) base regime
- $\log(S_t - q_2) \sim \mathcal{N}(\mu_S, \sigma_S^2)$ - Shifted log-normal spikes
  - $q_2 = 75^{th}$ percentile of the data

$$X_t = \begin{cases} B_t & \text{when } R_t = \text{Base}, \\ S_t & \text{when } R_t = \text{Spike}, \end{cases}$$

Results - 2-regimes



Results - 2-regimes

- $B_t = \alpha + \phi B_{t-1} + \sigma\varepsilon_t$ - AR(1) base regime
- $\log(S_t - q_2) \sim \mathcal{N}(\mu_S, \sigma_S^2)$ - Shifted log-normal spikes

Table: Posterior means for parameters

Parameter	
p_{11}	0.9452
p_{22}	0.6573
α	-0.5974
ϕ	0.5027
σ^2	423.8374
μ_S	4.5003
σ_S^2	1.2291

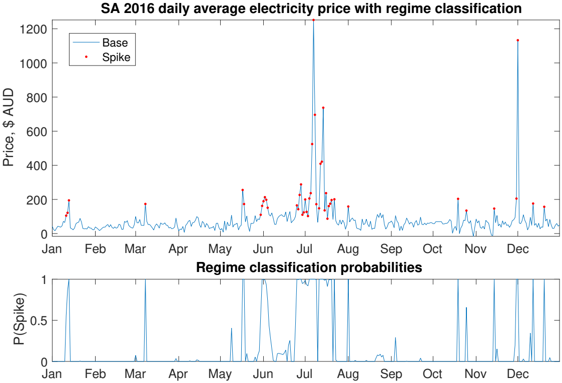
Results - Interpretation

- Mean spike size $=$ \$314.90 (above the trend)
- Spike std. dev. $= \$1,363.89$
- Stationary dist. of regimes

$$(\pi_1, \pi_2) \approx (0.862, 0.138)$$

Parameter	Mean
p_{11}	0.9452
p_{22}	0.6573
α	-0.5974
ϕ	0.5027
σ^2	423.8374
μ_S	4.5003
σ_S^2	1.2291

Is regime-switching really time-homogeneous?



In addition it is known that prices are dependent on

- weather
- day of week
- transmission outages

Model extension - exogenous variables

- $B_t = \alpha + \phi B_{t-1} + \sigma \varepsilon_t$
- $\log(S_t - q_2) \sim \mathcal{N}(\mu_S, \sigma_S^2)$
- Stochastic component

$$X_t = \begin{cases} B_t & \text{when } R_t = \text{Base}, \\ S_t & \text{when } R_t = \text{Spike}, \end{cases}$$

- R_t evolves with probabilities

$$\begin{aligned} p_{ij} &= \mathbb{P}(R_t = j | R_{t-1} = i, \mathbf{w}_t) \\ &= \frac{\exp(\boldsymbol{\beta}_j \mathbf{w}_t)}{\sum_k \exp(\boldsymbol{\beta}_k \mathbf{w}_t)} \end{aligned}$$

where $\boldsymbol{\beta}_j \mathbf{w}_t = \beta_{j,0} w_0 + \beta_{j,1} w_1 + \beta_{j,2} \mathbb{I}(R_{t-1} = 2) + \ldots$

◦ What predictors? Weather, load, maintenance indicators

Preliminary results - Temperature-dependent probabilities

$$
\begin{aligned}
p_{ij} &= \mathbb{P}(R_t = j | R_{t-1} = i, \mathbf{w}_t) \\
&= \frac{\exp(\boldsymbol{\beta}_j \mathbf{w}_t)}{\sum_k \exp(\boldsymbol{\beta}_k \mathbf{w}_t)}
\end{aligned}
$$

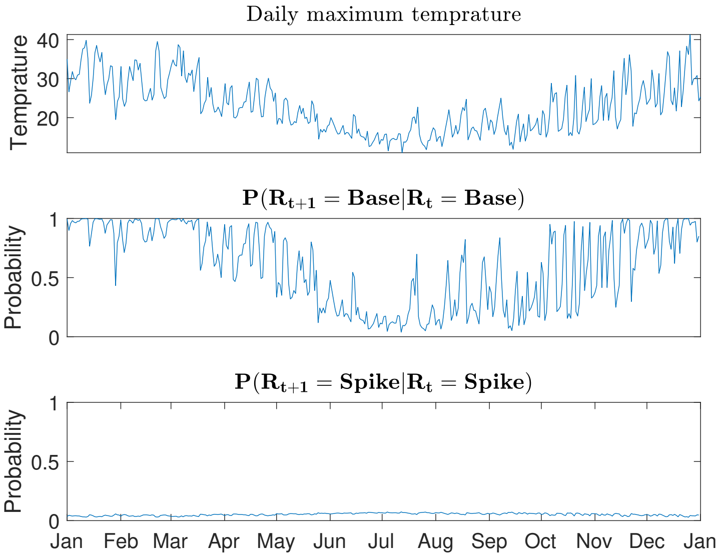
where

$$
\begin{aligned}
\boldsymbol{\beta}_1 \mathbf{w}_t &= \beta_{1,0} + \beta_{1,1} C_t^\circ + \beta_{1,2} \mathbb{I}(R_{t-1} = 2) + \beta_{1,3} \mathbb{I}(R_{t-1} = 2) C_t^\circ \\
\boldsymbol{\beta}_2 &= \mathbf{0}
\end{aligned}
$$

Parameter	
$\beta_{1,0}$	-7.0309
$\beta_{1,1}$	0.3466
$\beta_{1,2}$	9.1939
$\beta_{1,3}$	-0.3136

Parameter	
α	-1.041
ϕ	0.5169
σ^2	421.0808
μ_S	4.4682
σ_S^2	1.2942

Interpretation



Future research

- 30-minute prices
  - Investigate trend models
  - Test the AR(1) assumption
- Investigate predictors for regime-switching probabilities
- Use exogenous variables in the stochastic component

$$X_t = \begin{cases} \beta \boldsymbol{w_t} + B_t & \text{when } R_t = \text{Base}, \\ S_t & \text{when } R_t = \text{Spike}, \end{cases}$$

Some model specifications

Model 0: 2-regime - Normal spikes.

- $B_t = \alpha + \phi B_{t-1} + \sigma\varepsilon_t$ - AR(1) base regime.
- $Z_t \sim \mathcal{N}(\mu_S, \sigma_S^2)$ - Normal spikes.

$$X_t = \begin{cases} B_t & \text{when } R_t = \text{Base}, \\ Z_t & \text{when } R_t = \text{Spike}, \end{cases}$$

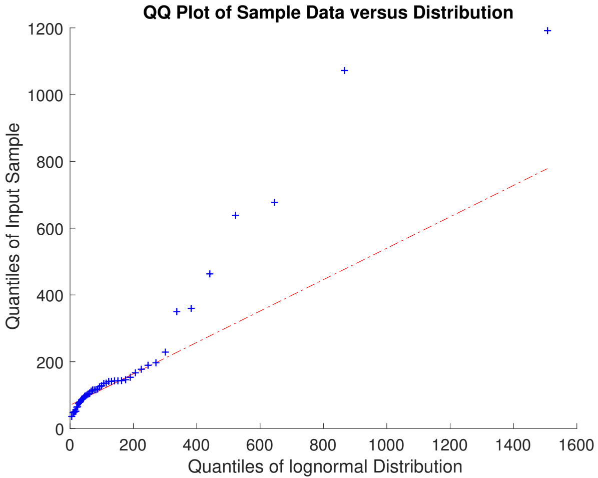
Results - Normal vs. log-normal spikes

- $B_t = \alpha + \phi B_{t-1} + \sigma\varepsilon_t$ - AR(1) base regime.
- $Z_t \sim \mathcal{N}(\mu_S, \sigma_S^2)$ - Normal spikes.
- $\log(S_t - q_2) \sim \mathcal{N}(\mu_S, \sigma_S^2)$ - Shifted log-normal spikes.

Table: Posterior means for parameters

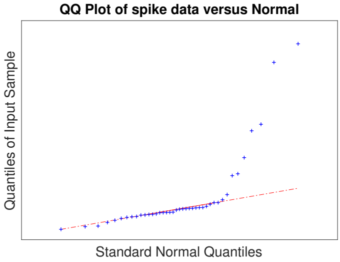
Parameter	Normal	LN
p_{11}	0.9543	0.9457
p_{22}	0.6551	0.5180
α	-4.9793	-2.1206
ϕ	0.5462	0.5290
σ^2	467.7538	464.3700
μ_S	191.9605	4.2605
σ_S^2	72852.8130	1.7273

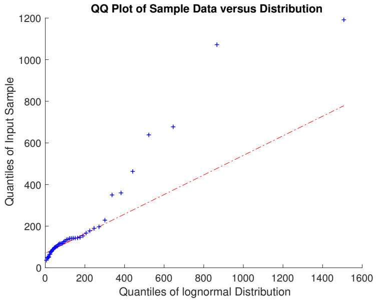
Diagnostics - But...



More work is needed here.

Diagnostics - Normal vs. log-normal spikes





Both aren't great, but the log-normal *might* be better.

Modelling 30-minute prices

- Much more complicated trend.
- Base prices may no longer be AR(1).
- Much more data,
  - Makes MCMC slower $\sim$ 3 hours (compared to 30 mins for a two regime model).

Metropolis Hastings algorithm

Generates a Markov chain $S_0, S_1, ..., S_N$ with $S_N \sim p(\boldsymbol{R}, \boldsymbol{\theta}|data)$ for large N.

Step 1. Set $n = 0$ and choose a starting point $S_0 = (\boldsymbol{R}^{(0)}, \boldsymbol{\theta}^{(0)})$.

Step 2. Given current state $S_n = (\boldsymbol{R}^{(n)}, \boldsymbol{\theta}^{(n)})$, generate

$$S' \sim \underbrace{q((\boldsymbol{\theta}', \boldsymbol{R}')|S_n)}_{\text{proposal}}.$$

Step 3. Generate $U \sim Uniform(0, 1)$ and let

$$S_{n+1} = \begin{cases} S' & ifU < \alpha(S_n, S') \\ S_n & \text{otherwise} \end{cases}$$

$$\alpha = \frac{p(data|\boldsymbol{R}', \boldsymbol{\theta}')(\boldsymbol{R}', \boldsymbol{\theta}')q((\boldsymbol{\theta}^{(n)}, \boldsymbol{R}^{(n)})|S')}{p(data|\boldsymbol{R}^{(n)}, \boldsymbol{\theta}^{(n)})(\boldsymbol{R}^{(n)}, \boldsymbol{\theta}^{(n)})q((\boldsymbol{\theta}', \boldsymbol{R}')|S_n)}$$

Estimation

- Observations: $\mathbf{x}_T = (x_0, x_1, ..., x_T)$.
- Parameters: $\boldsymbol{\theta}$.
- The likelihood:

$$L(\boldsymbol{\theta}) = \mathbb{P}(\mathbf{x}_T|\boldsymbol{\theta}) = \sum_{\mathbf{R}} \mathbb{P}(\mathbf{x}_T|\mathbf{R}, \boldsymbol{\theta})\mathbb{P}(\mathbf{R}|\boldsymbol{\theta}).$$

  But the sum is too big to calculate.

- Expectation-Maximisation.