Bayesian Methods in Finance

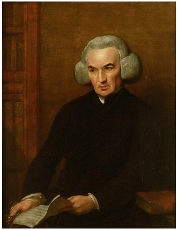
Amina Hussein & Andy Ton & Arav Agarwal

December 13, 2023
Bowdoin College

Beginning of Bayesian Stats In Finance

Bayesian statistics has been tied to finance since its inception.

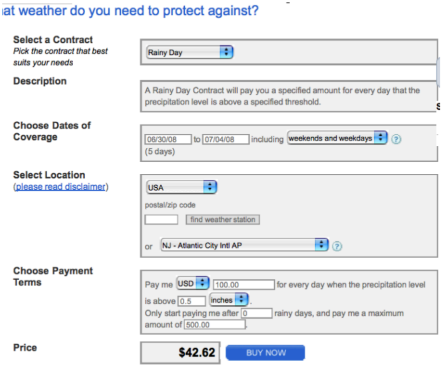
Richard Price published Bayes' Theorem in 1763. He was an actuary who went on to apply Bayes' Theorem in his work and developed level premium life insurance.



Years	Age	Death Probability	Present Value of Total Future Death Benefits Payable (1)	Present Value of Total Future Premiums Receivable (2)	Policy Reserves: (1) - (2)	Expected Losses
1	40	0.002084	238,294,011	10,142,468	**10,142,468**	20,840,000
2	41	0.002241	229,368,711	22,241,861	22,241,861	23,970,696
3	42	0.002439	218,473,849	33,350,246	33,350,246	25,685,031
4	43	0.002686	205,112,915	43,013,795	43,013,795	27,824,664
5	44	0.002975	188,689,833	50,679,716	50,679,716	30,332,091
6	45	0.003297	168,654,470	55,842,029	55,842,029	33,060,837
7	46	0.003639	144,524,824	58,064,113	58,064,113	35,813,723
8	47	0.003997	115,929,472	57,021,660	57,021,660	38,520,846
9	48	0.004366	82,523,452	52,418,774	52,418,774	41,141,584

Current Research - Commodities Pricing

How to model volatile assets like commodities?



Year	Payout	Year	Payout
2007	$0	1992	$0
2006	$0	1991	$0
2005	$100	1990	$0
2004	$0	1989	$0
2003	$100	1988	$0
2002	$0	1987	$100
2001	$100	1986	$100
2000	$100	1985	$0
1999	$0	1984	$0
1998	$0	1983	$0
1997	$0	1982	$100

We can hedge against fluctuations like the weather!

How is Bayesian Statistics Used Today?

Examples:

- Portfolio Allocation (Black-Litterman - 1990)
- Options Pricing (Black-Scholes - 1973)
- Market Making (Logarithmic Market Scoring Rule - 2002)

Notice that they all share a similar feature...

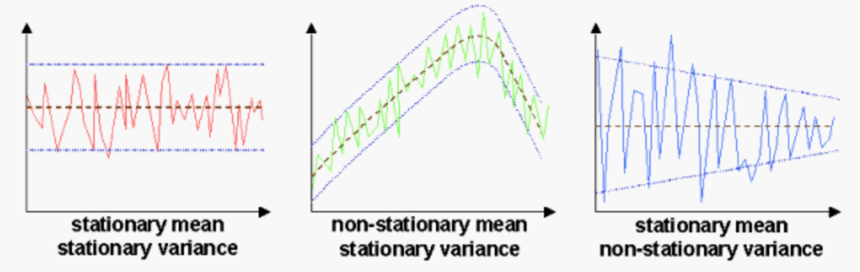
Time series models!

- Used to forecast based on verified historical data
- Examples of time series data: stock prices, temperature readings, sales data, etc.
- Examples of time series models are AR, MA, ARCH, GARCH, ARMA, ARIMA, etc.

Time Series Models: An Overview

1. Stationary Check:

- Ensure constant statistical properties over time
- Apply transformations if needed (e.g., differencing)



Time Series Models: An Overview

TEX

2. Decomposition:

- ***Trend***: Long-term movement or direction in the data
- ***Seasonality***: Regular and predictable variations in the data
- ***Residual components***: Noise component, random fluctuations or variability in the data

It informs model choices!

Time Series Models: An Overview

3: Autocorrelation Analysis

- Examine correlation with past observations
- Identify lags and patterns
- Utilize ACF and PACF (to identify the order of AR and MA)

Time Series Models: An Overview

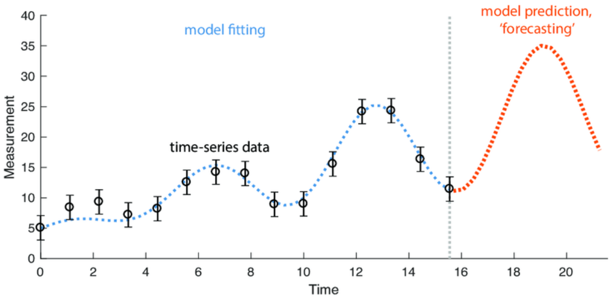
4: Model Selection, Fitting, and Validation

- Choose appropriate model based on data characteristics
- Split data into training and testing sets
- Train the model
- Assess model performance using metrics (MAE, MSE)

Time Series Models: An Overview

4: Forecasting

- Use the validated model for future predictions
- Visualize the forecasted values along with the historical data to assess the model's performance



Autoregressive models: AR(p)

An autoregressive model of order p is defined as

$$y_t = \sum_{i=1}^{p} \varphi_i y_{t-i} + \epsilon_t,$$

where $\varphi_1, \ldots, \varphi_p$ are parameters of the model, and ϵ_t is white noise.

So, we have a stochastic difference equation.

Autoregressive models: AR(p)

Together with the moving-average (MA) model, it is a special case and key component of the more general autoregressive–moving-average (ARMA) and autoregressive integrated moving average (ARIMA) models of time series, which have a more complicated stochastic structure.
Also vector autoregressive model (VAR), which consists of a system of more than one interlocking stochastic difference equation in more than one evolving random variable.

The notation MA(q) refers to the moving average model of order q:

$$X_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q} = \mu + \epsilon_t + \sum_{i=1}^{q} \epsilon_{t-i}.$$

Different from AR(p) how?

AR(1)

A first-order autoregressive model (AR(1)) with normal noise takes each point y_n in a sequence to be generated according to

$$y_n \sim \mathcal{N}(\alpha + \beta y_{n-1}, \sigma).$$

So, the expected value of y_n is $\alpha + \beta y_{n-1}$, with noise scaled as σ.

AR(1) inference

Data: Apple stock prices

Prior: Improper flat priors on the regression coefficients α and β and on the positively-constrained noise scale σ

Likelihood: Specified by

$$y_n \sim \mathcal{N}(\alpha + \beta y_{n-1}, \sigma)$$

AR(1) inference

Model:

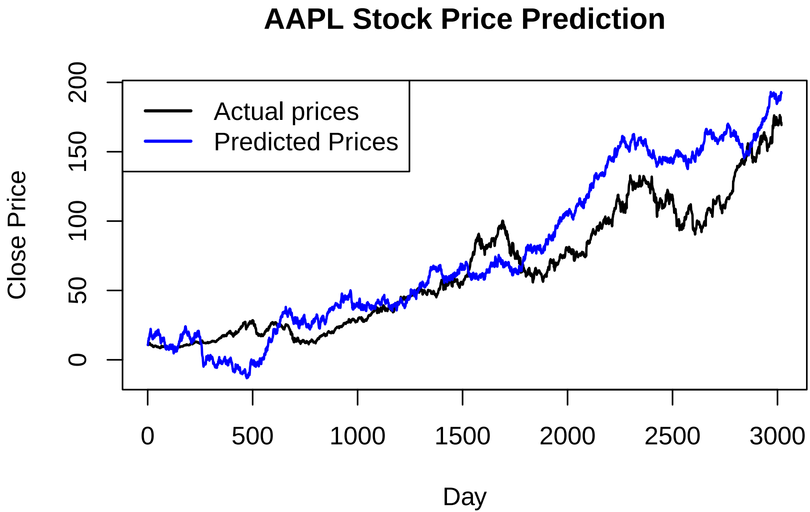
$$y_n \sim \mathcal{N}(\alpha + \beta y_{n-1}, \sigma)$$

Results:

$$\alpha = 0.03, \;\; \beta = 1, \;\; \sigma = 1.17$$

Interpretation?

AR(1) inference



AR(2) inference

Model:

$$y_n \sim \mathcal{N}(\alpha + \beta y_{n-1} + \gamma y_{n-2}, \sigma)$$

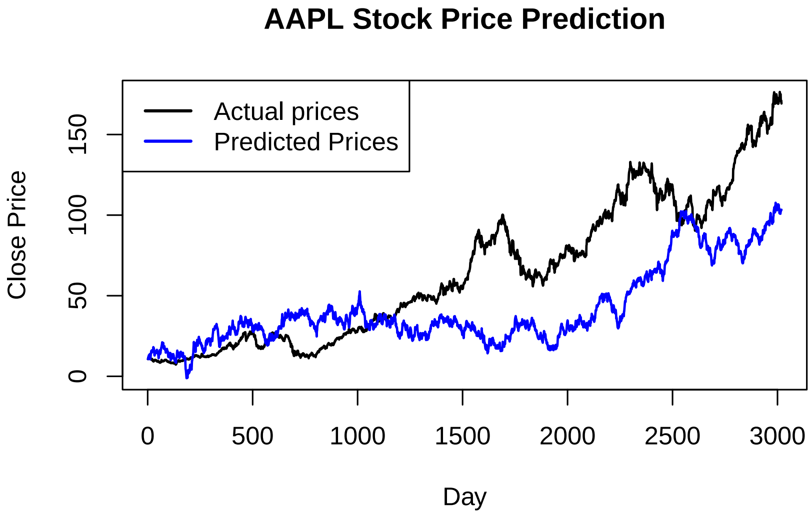
Results:

$$\alpha = 0.03, \;\; \beta = 1.02, \;\; \gamma = -0.02, \;\; \sigma = 1.17$$

Took 5 minutes!

Interpretation?

AR(2) inference



AR(3) inference

Model:

$$y_n \sim \mathcal{N}(\alpha + \beta y_{n-1} + \gamma y_{n-2} + \delta y_{n-3}, \sigma)$$

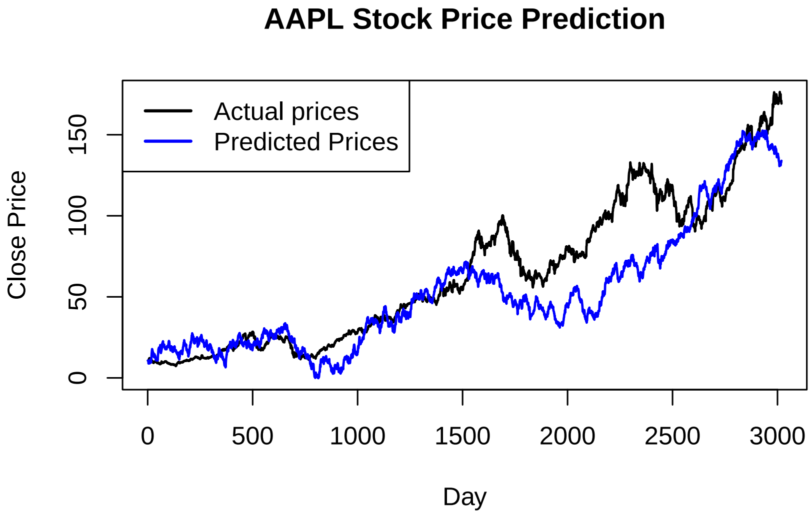
Results:

$$\alpha = 0.03, \;\; \beta = 1.02, \;\; \gamma = -0.05, \;\; \delta = 0.02, \;\; \sigma = 1.17$$

Took 15 minutes!

Interpretation?

AR(3) inference



Bayes factors

- $BF_{12} = 20.7$, in favor of AR(1) over AR(2).
- $BF_{23} = 25.3$, in favor of AR(2) over AR(3).
- $BF_{13} =??$

Bayes factors

- $BF_{12} = 20.7$, in favor of AR(1) over AR(2)
- $BF_{23} = 25.3$, in favor of AR(2) over AR(3)
- $BF_{13} = 524.3$, in favor of AR(1) over AR(3)

Notice!

AR models assume **homoskedasticity**. They assume that the variance of the error term at each time point is constant over time.

$$X_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t$$

$$\text{where} \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2)$$

How do we model time series data exhibiting changing volatility?

ARCH(p)

ARCH(p), Autoregressive Conditional Heteroskedasticity model with p lags, is a generalized form of the ARCH model that captures changing volatility over time.

It is defined as:

$$r_t = \mu + a_t$$

$$a_t = \sigma_t \epsilon_t$$

$$\epsilon_t \sim \mathcal{N}(0, 1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2$$

ARCH(p)

ARCH models primarily focus on the immediate impact of past shocks on the current volatility, i.e When there are large deviations (shocks) in the time series, ARCH acknowledges and models how these past shocks affect the variability of the data.

What about the the persistence of volatility over an extended period?

GARCH(p,q)

The Generalized Autoregressive Conditional Heteroskedasticity is designed to capture both *short-term shocks* and the *persistence of volatility over time.*

It is defined as:

$$\sigma_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}$$

where:

- h_t: Conditional variance at time t.
- p: Order of AR terms
- q: Order of MA terms

GARCH(1,1)

A first-order GARCH model (GARCH(1,1)) is generated according to a

$$y_t \sim \mathcal{N}(\mu_t, \sigma_t).$$

μ_t is the estimated mean and σ_t is the conditional volatility at time t.

$$\sigma_t = \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2}.$$

GARCH(1,1) inference

Data: Apple stock prices

Prior: Improper flat priors on μ_t, coefficients α, β and on the conditional variance σ_t

Likelihood: Specified by

$$y_t \sim \mathcal{N}(\mu_t, \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}}).$$

GARCH(1,1) inference

Model:

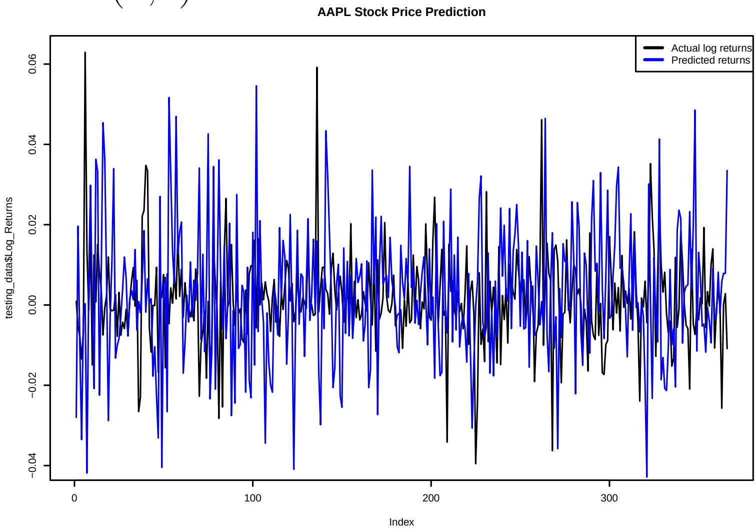
$$y_t \sim \mathcal{N}(\mu_t, \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2}).$$

Results:

$$\mu_t = 0.0016, \quad \sigma_t = 0.0214$$

Interpretation?

GARCH(1,1) inference



GARCH(1,2) inference

Model:

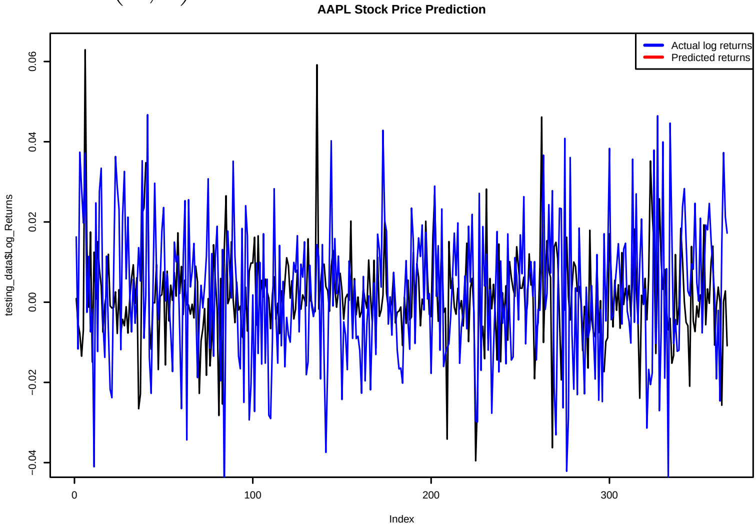
$$y_t \sim \mathcal{N}(\mu_t, \sigma_t = \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2}).$$

Results:

$$\mu_t = 0.003, \quad \sigma_t = 0.0161$$

Interpretation?

GARCH(1,2) inference



GARCH(2,2) inference

Model:

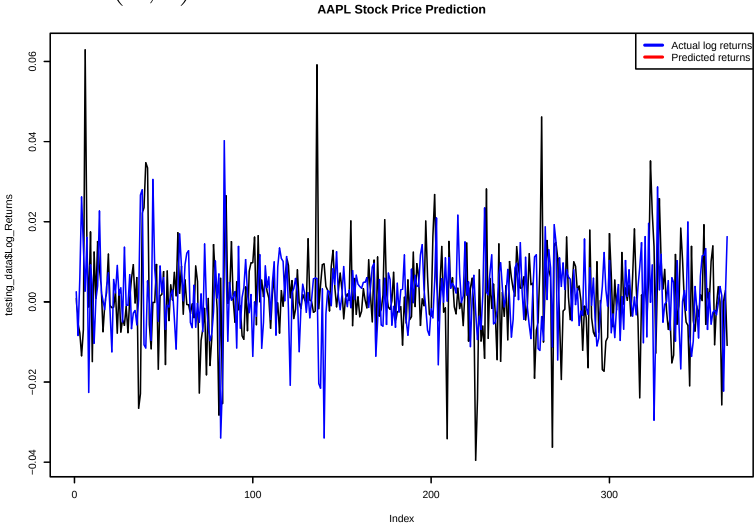
$$y_t \sim \mathcal{N}(\mu_t, \sigma_t = \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2}).$$

Results:

$$\mu_t = 0.0024, \quad \sigma_t = 0.0088$$

Interpretation?

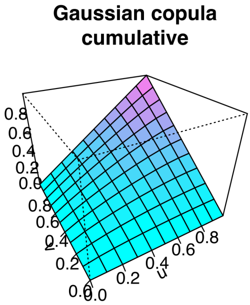
GARCH(2,2) inference

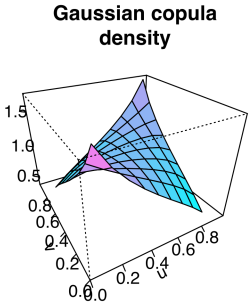


Bayes factors

- $BF_{12} = 58.37$, in favor of GARCH(1,1) over GARCH(1,2).
- $BF_{13} = 7.67$, in favor of GARCH(1,1) over GARCH(1,3).

Gaussian Copula





Used to model dependence between random variables.

Gaussian Copula

TEX

For a given correlation matrix $R \in [-1, 1]^{d\times d}$, the Gaussian copula with parameter matrix R can be written as

$$C_R^{\text{Gauss}} = \Phi_R\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_d)\right).$$

where Φ^{-1} is the inverse cumulative distribution function of a standard normal and Φ_R is the joint cumulative distribution function of a multivariate normal distribution with mean vector zero and covariance matrix equal to the correlation matrix R.

Copula inference

Data: Two commodities, namely, Tea and Coffee Robusta prices

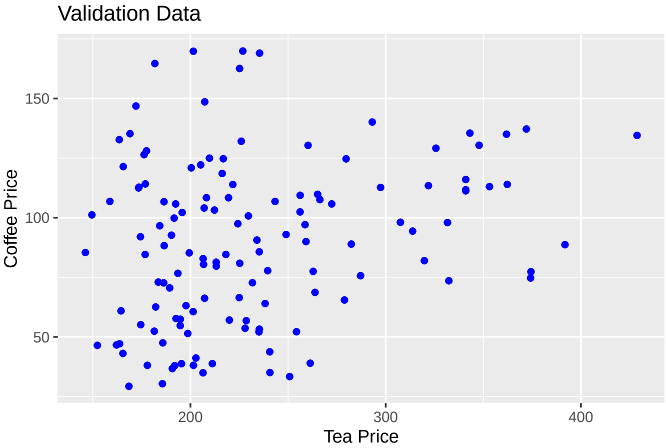
Prior: LKJ (Lewandowski-Kurowicka-Joe) on correlation matrix

Likelihood: Specified by

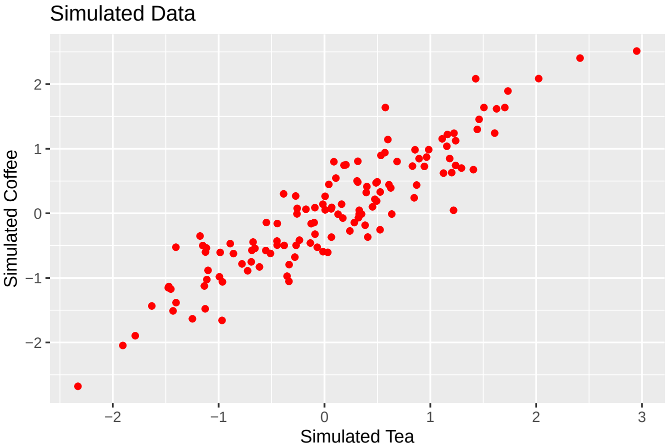
$$C_R^{\text{Gauss}} = \Phi_R\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_d)\right)$$

Note: used Cholesky decomposition to increase efficiency of computations

Copula inference



Copula inference



Possible Future Projects

- **Cross-model Evaluation**: Compare GARCH and AR performances.
- **Advanced GARCH Model Analysis**: Experiment with EGARCH and TGARCH to capture asymmetric volatility.

References

TEX

- Stan User's Guide (Version 2.33) - Stan Development Team
- Bayesian Methods in Finance - Eric Jacquier and Nicholas Polson
- Bayesian forecasting in economics and finance: A modern review - Martin, et al.
- A Bayesian Learning Method for Financial Time-Series Analysis - Zhu, et al.
- The use of Bayesian methods in financial research - Beck, et al.
- A Jump Diffusion Model for Agricultural Commodities with Bayesian Analysis - Schmitz, et al.
- Information Aggregation In Dynamic Markets With Strategic Traders - Michael Ostrovsky