

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

Stata webinar

Introduction to Bayesian Analysis in Stata

Gustavo Sánchez

StataCorp LLC

June 16, 2026



Outline

Overview

Fundamental
equation
MCMC
Convergence
Stata tools

Linear
regression

Model fit
Postestimation
bayesstats ess
bayesgraph
Multiple Markov
Chains
bayestestmodel
Forecasting
probabilities

Summary

References

Outline

1 Bayesian analysis: Overview

- Fundamental equation
- Convergence
- The Stata tools

2 A couple of examples

- Linear regression
- Multivariate regression

Overview

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

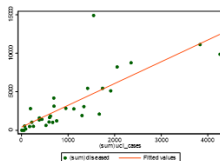
References

Frequentist

Data hypothetically repeatable

```
. list month defunciones casos_uci, abbreviate(12)
```

	month	defunciones	casos_uci
1.	2022m11	631	524
2.	2022m12	1942	1298
3.	2022m1	5453	1748
4.	2022m2	4183	691
5.	2022m3	1688	382
6.	2022m4	1422	436
7.	2022m5	1848	628
8.	2022m6	1663	681
9.	2022m7	3133	696
10.	2022m8	1846	219



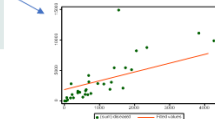
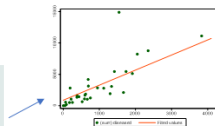
Theoretical
Model

Bayesian

Data known

```
. list month defunciones casos_uci, abbreviate(12)
```

	month	defunciones	casos_uci
1.	2022m11	631	524
2.	2022m12	1942	1298
3.	2022m1	5453	1748
4.	2022m2	4183	691
5.	2022m3	1688	382
6.	2022m4	1422	436
7.	2022m5	1848	628
8.	2022m6	1663	681
9.	2022m7	3133	696
10.	2022m8	1846	219



Theoretical
Model

Bayesian Analysis vs. Frequentist Analysis

Outline

Overview

Fundamental
equation
MCMC
Convergence
Stata tools

Linear regression

Model fit
Postestimation
bayesstats ess
bayesgraph
Multiple Markov
Chains
bayestestmodel
Forecasting
probabilities

Summary

References

Frequentist Analysis

- Estimates unknown fixed parameters.
- The data come from a random sample (hypothetical repeatable).
- Uses data to estimate unknown fixed parameters.
- p -values are conditional probability statements that assume H_0 to be true.

"Conclusions are based on the distribution of statistics derived from random samples, assuming unknown but fixed parameters."

Bayesian Analysis

- Probability distributions for unknown random parameters.
- The data are fixed.
- Combines data with prior beliefs to get updated probability distributions for the parameters.
- It allows formulating probabilistic statements for the hypothesis of interest.

"Bayesian analysis answers questions based on the distribution of parameters conditional on the observed sample."

Fundamental equation for Bayesian analysis

$$p(\theta|y) \propto L(\theta; y) \pi(\theta)$$

Where:

$p(\theta|y)$: posterior probability distribution for θ

$L(\theta; y)$: likelihood function of the parameters given the data.

$\pi(\theta)$: prior probability distribution for θ

- How do we get it? Using the Bayes' Theorem:

$$p(\theta|y) = \frac{p(y|\theta) p(\theta)}{p(y)} = \frac{f(y; \theta) \pi(\theta)}{f(y)}$$

Where:

$f(y; \theta)$: probability density function for y given θ .

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

We need to solve the fundamental equation

- Let's assume that both the data and the prior beliefs are normally distributed:
 - The data:** $y \sim N(\theta, \sigma_d^2)$
 - The prior:** $\theta \sim N(\mu_p, \sigma_p^2)$
- The posterior distribution would be normal with the parameters specified below (see, for example, Cameron & Trivedi 2005):
 - The posterior:** $\theta|y \sim N(\mu, \sigma^2)$

Where:

$$\begin{aligned}\mu &= \sigma^2 (N\bar{y}/\sigma_d^2 + \mu_p/\sigma_p^2) \\ \sigma^2 &= (N/\sigma_d^2 + 1/\sigma_p^2)^{-1}\end{aligned}$$

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

The previous example has a closed-form solution.

- What if there is not closed form solution, or it is a complex distribution?
 - Use simulation to get the posterior distribution.
 - We can use Markov chain Monte Carlo (MCMC). E.g.:
 - Metropolis–Hastings (MH) algorithm
 - Gibbs sampling

Outline

Overview

Fundamental
equation

MCMC

Convergence
Stata tools

Linear
regression

Model fit

Postestimation
bayesstats ess

bayesgraph

Multiple Markov
Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Links for Bayesian analysis and MCMC on our YouTube channel:

- Introduction to Bayesian statistics, part 1: The basic concepts.

<https://www.youtube.com/watch?v=0F0QoMCSKJ4&feature=youtu.be>

- Introduction to Bayesian statistics, part 2: MCMC and the Metropolis–Hastings algorithm.

<https://www.youtube.com/watch?v=OTO1DygELpY&feature=youtu.be>

Monte Carlo simulation

Outline

Overview

Fundamental equation

MCMC

Convergence

Stata tools

Linear regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

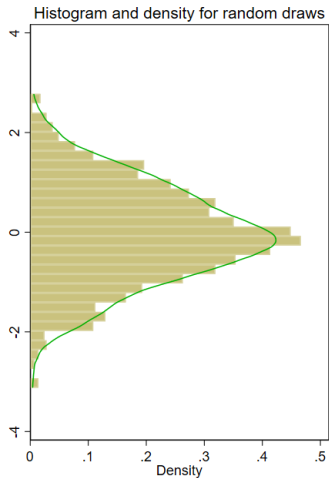
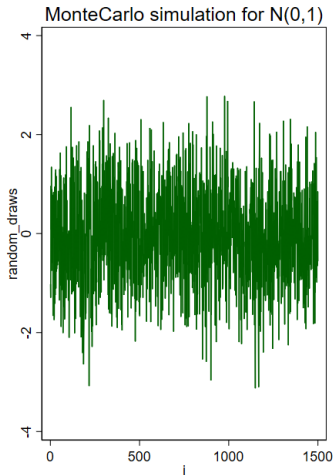
bayestestmodel

Forecasting

probabilities

Summary

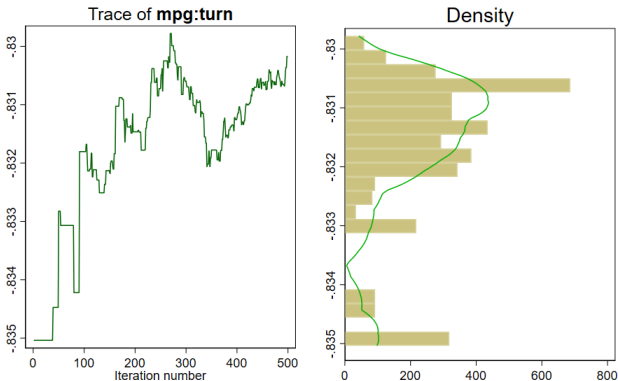
References



Metropolis–Hastings simulation (Not convergent)

- Trace plot illustrates the sequence of accepted proposal states (with insufficient burn-in iterations).

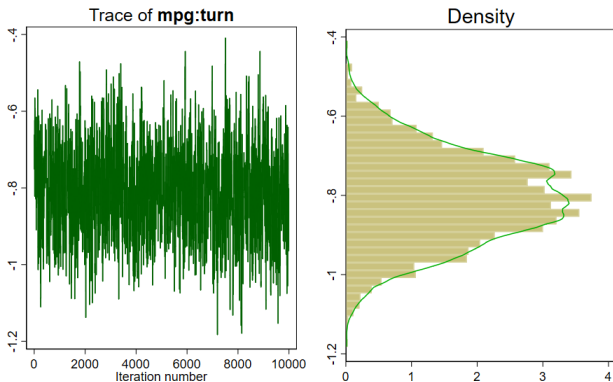
Metropolis-Hastings simulation
with insufficient 'iterations and burning iterations'



Metropolis–Hastings simulation (Convergent)

- Trace plot illustrates the sequence of accepted proposal states (with enough burn-in iterations).

Metropolis–Hastings simulation
convergent result



Stationary sequences indicate converge is achieved

- Trace patterns for parameter θ

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

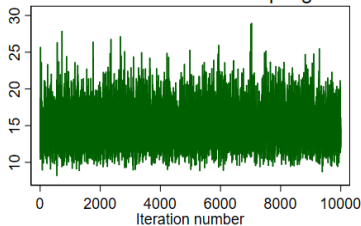
bayestestmodel

Forecasting
probabilities

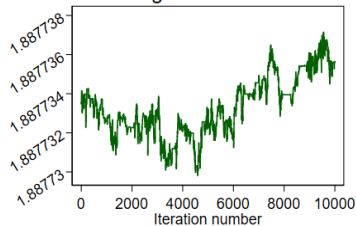
Summary

References

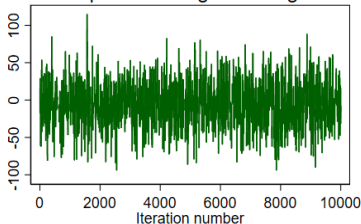
Pattern for Gibbs sampling



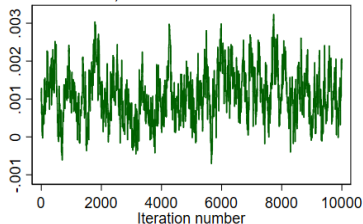
Convergence not achieved



Metropolis-Hastings convergence



M-H, not well mixed trace



Small autocorrelation indicate efficient MCMC

- Autocorrelation patterns for parameter θ .

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

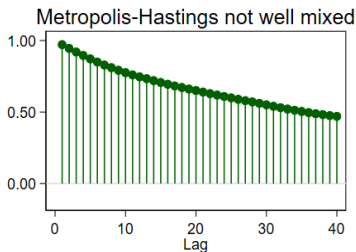
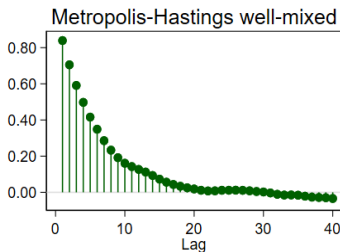
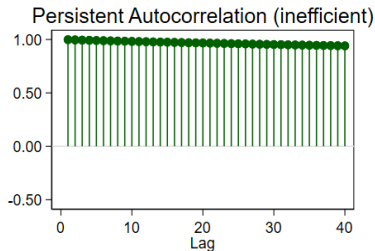
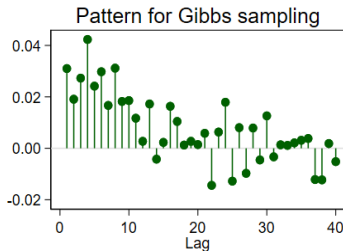
Chains

bayestestmodel

Forecasting
probabilities

Summary

References



Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

The Stata tools for Bayesian regression

Stata's convenient syntax: `bayes:`

```
regress y x1 x2 x3
```

```
bayes: regress y x1 x2 x3
```

```
logit y x1 x2 x3
```

```
bayes: logit y x1 x2 x3
```

```
var x1 x2 x3, lags(1/3)
```

```
bayes: var y x1 x2 x3, lags(1/3)
```


Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Estimation

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Example 1: Inflation in Spain

- Let's work with a simple linear regression for inflation as a function of the dollar-euro exchange rate and oil prices
- Let's consider the following model specification:

$$\textit{inflation} = \alpha_1 + \beta_{\textit{dollar_euro}} * \textit{dollar_euro} + \beta_{\textit{oil_price}} * \textit{oil_price} + \epsilon_1$$

Where:

inflation : Monthly interannual change for the CPI.

dollar_Euro : Monthly interannual change in the Dollar-Euro exch.rate.

oil_price : Monthly interannual change for the spot Crude Oil Price
West Texas Intermediate (WTI).

import fred: Dialog box

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Import Federal Reserve Economic Data

Search FRED

Keywords:
euro exchange rate monthly

Full text Series ID

Tags:
Sources
Releases
Seasonal Adjustment
Frequencies
Geography Types
Geographies
Concepts

Sort by: Popularity Descend

#	ID	Title	Frequency
1	BAMLHE00EHYL...	ICE BofA Euro High Yield Index Option-Adjusted Spre...	Daily, Close
2	BAMLHE00EHYEY	ICE BofA Euro High Yield Index Effective Yield	Daily, Close
3	EXUSEU	U.S. Dollars to Euro Spot Exchange Rate	Monthly
4	BAMLEMCBPIOAS	ICE BofA Emerging Markets Corporate Plus Index Opti...	Daily
5	BAMLHE00EHYL...	ICE BofA Euro High Yield Index Total Return Index Val...	Daily, Close
6	BAMLEMCBPIEY	ICE BofA Emerging Markets Corporate Plus Index Effec...	Daily
7	RBXMBIS	Real Broad Effective Exchange Rate for Euro Area	Monthly
8	CCUSMA02EZM...	Currency Conversions: US Dollar Exchange Rate: Avera...	Monthly
9	BAMLEMCBPITRIV	ICE BofA Emerging Markets Corporate Plus Index Tota...	Daily, Close
10	TWEXBPA	Real Broad Dollar Index (Goods Only) (DISCONTINUED)	Monthly
11	EXGEUS	Germany / U.S. Foreign Exchange Rate (DISCONTINUE...	Monthly
12	TWEXMMTH	Nominal Major Currencies U.S. Dollar Index (Goods O...	Monthly
13	CCEUSP01GBM...	Currency Conversions: EURO Exchange Rate: Spot, En...	Monthly
14	BAMLHE00EHYL...	ICE BofA Euro High Yield Index Semi-Annual Yield to ...	Daily, Close
15	BAMLEMEBCRPL...	ICE BofA Euro Emerging Markets Corporate Plus Index...	Daily
16	EXFRUS	France / U.S. Foreign Exchange Rate (DISCONTINUED)	Monthly
17	EXBEUS	Belgium / U.S. Foreign Exchange Rate (DISCONTINUED)	Monthly
18	CCEUSP01USM...	Currency Conversions: EURO Exchange Rate: Spot, En...	Monthly
19	NBXMBIS	Broad Effective Exchange Rate for Euro Area	Monthly
20	TWEXBMTH	Nominal Broad U.S. Dollar Index (Goods Only) (DISCO...	Monthly
21	EXITUS	Italy / U.S. Foreign Exchange Rate (DISCONTINUED)	Monthly
22	EXNEUS	Netherlands / U.S. Foreign Exchange Rate (DISCONTI...	Monthly
23	EXGRUS	Greece / U.S. Foreign Exchange Rate (DISCONTINUED)	Monthly
24	BAMLEMCCLCRP...	ICE BofA US Emerging Markets Liquid Corporate Plus I...	Daily, Close
25	CCUSP01EZM6...	Currency Conversions: US Dollar Exchange Rate: Spot...	Monthly
26	EXSPUS	Spain / U.S. Foreign Exchange Rate (DISCONTINUED)	Monthly
27	CCUSMA02DEM...	Currency Conversions: US Dollar Exchange Rate: Avera...	Monthly
28	BAMLEMEBCRPL...	ICE BofA Euro Emerging Markets Corporate Plus Index...	Daily
29	TWEXMPA	Real Major Dollar Index (Goods Only) (DISCONTINUED)	Monthly
30	EXAUIUS	Austria / U.S. Foreign Exchange Rate (DISCONTINUED)	Monthly
31	CCUSMA02ITM...	Currency Conversions: US Dollar Exchange Rate: Avera...	Monthly
32	EXUSEC	Foreign Exchange Rate: Euro Community (DISCONTIN...	Monthly

1-82 of 82

Describe Add

Remove

Import Cancel

Ready

Data for exchange rate, cpi, and oil price

Outline

Overview

Fundamental

equation

MCMC

Convergence

Stata tools

Linear

regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

```
. describe
```

```
Contains data from data\cpi_spain_monthly.dta
```

```
Observations:      243
```

```
Variables:         7
```

```
12 Jun 2026 12:44
```

Variable name	Storage type	Display format	Value label	Variable label
month	float	%tm		
dollar_euro	float	%9.0g		U.S. Dollars to Euro Spot Exchange Rate
cpi_spain	float	%9.0g		Consumer Price Indices (CPIs, HICPs), COICOP 1999: Consumer Price Index: Tot..
oil_price	float	%9.0g		Spot Crude Oil Price: West Texas Intermediate (WTI)
ldollar_euro	float	%9.0g		ln(dollar_euro)
lcpi_spain	float	%9.0g		ln(cpi_spain)
loil_price	float	%9.0g		ln(oil_price)

```
Sorted by: month
```

```
.
```

```
.
```

```
. tsset
```

```
Time variable: month, 2005m1 to 2025m3
```

```
Delta: 1 month
```

Plot the log-levels of the variables

Outline

Overview

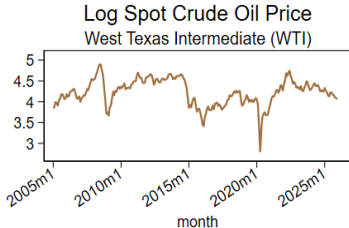
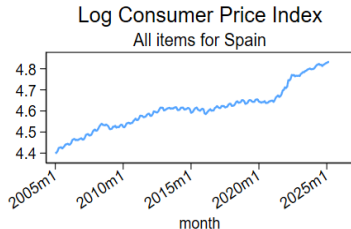
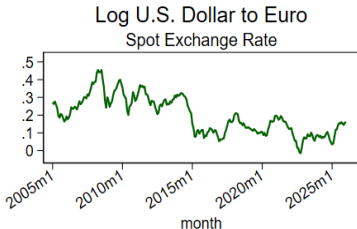
Fundamental
equation
MCMC
Convergence
Stata tools

Linear regression

Model fit
Postestimation
bayesstats ess
bayesgraph
Multiple Markov
Chains
bayestestmodel
Forecasting
probabilities

Summary

References



Source: Federal Reserve Economic Data (FRED)
Downloaded using -import fred- in Stata

Plot the first difference of the log-levels

Outline

Overview

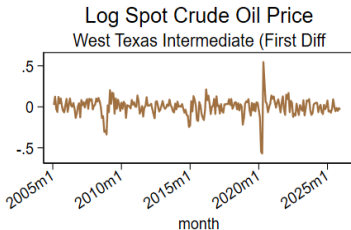
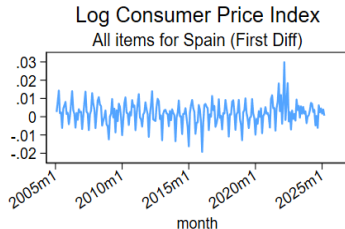
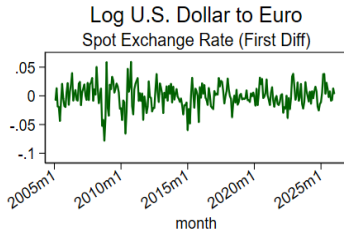
Fundamental
equation
MCMC
Convergence
Stata tools

Linear regression

Model fit
Postestimation
bayesstats ess
bayesgraph
Multiple Markov
Chains
bayestestmodel
Forecasting
probabilities

Summary

References



Source: Federal Reserve Economic Data (FRED)

Downloaded using `-import fred-` in Stata

Plot the interannual change of the log-levels

Outline

Overview

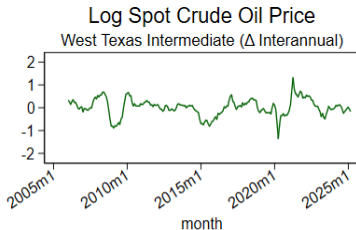
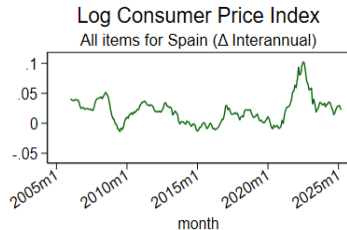
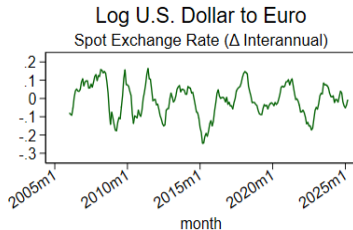
Fundamental
equation
MCMC
Convergence
Stata tools

Linear regression

Model fit
Postestimation
bayesstats ess
bayesgraph
Multiple Markov
Chains
bayestestmodel
Forecasting
probabilities

Summary

References



Source: Federal Reserve Economic Data (FRED)

Downloaded using -import fred- in Stata

Linear Regression

- Linear regression with the `bayes:` prefix

```
. bayes, rseed(1) : ///
> regress S12.lcpi_spain S12.ldollar_euro S12.loil_price
```

- Equivalent model with `bayesmh`

```
. bayesmh S12.lcpi_spain S12.ldollar_euro S12.loil_price, ///
> likelihood(normal({sigma2})) ///
> prior({S12.lcpi_spain:S12.ldollar_euro}, normal(0,10000)) ///
> prior({S12.lcpi_spain:S12.loil_price}, normal(0,10000)) ///
> prior({S12.lcpi_spain:_cons}, normal(0,10000)) ///
> prior({sigma2}, igamma(.01,.01)) ///
> block({S12.lcpi_spain:}) block({sigma2}) rseed(1)
```

Outline

Overview

Fundamental

equation

MCMC

Convergence

Stata tools

Linear regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

Menu for Bayesian regression

Outline

Overview

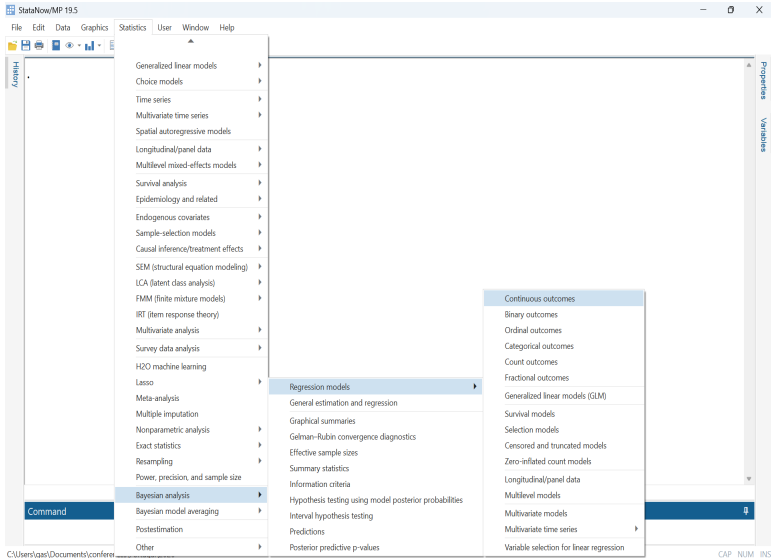
Fundamental equation
MCMC
Convergence
Stata tools

Linear regression

Model fit
Postestimation
bayesstats ess
bayesgraph
Multiple Markov Chains
bayestestmodel
Forecasting probabilities

Summary

References



Menu for Bayesian regression

Outline

Overview

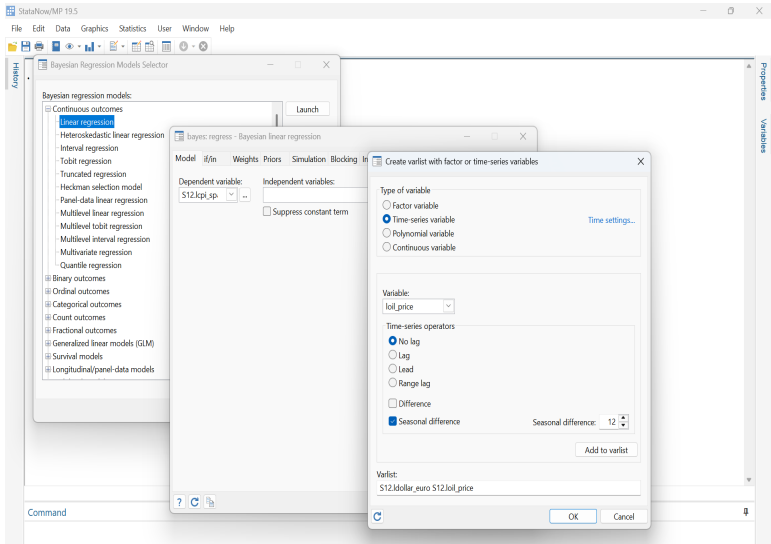
Fundamental equation
MCMC
Convergence
Stata tools

Linear regression

Model fit
Postestimation
bayesstats ess
bayesgraph
Multiple Markov Chains
bayestestmodel
Forecasting probabilities

Summary

References



Menu for Bayesian regression

- 1 Make the following sequence of selection from the main menu:

Statistics > Bayesian analysis > Regression models

- 2 Select 'Continuous outcomes'
- 3 Select 'Linear regression'
- 4 Click on 'Launch'
- 5 Specify the dependent variable (S12.lcpi) and the explanatory variables (S12.ldollar_euro S12.loil_price)
- 6 Click on 'OK'

bayes : prefix

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

```
. bayes, rseed(1) dryrun blocksummary:      ///
> regress S12.lcpi_spain S12.ldollar_euro S12.loil_price
```

Model summary**Likelihood:**

```
S12.lcpi_spain ~ regress(xb_S12.lcpi_spain, {S12e.lcpi_spain:sigma2})
```

Priors:

```
{S12.lcpi_s~n:S12.ldollar_euro S12.loil_price _cons} ~ normal(0,10000)
{S12e.lcpi_~n:sigma2} ~ igamma(.01,.01)
```

(1) Parameters are elements of the linear form `xb_S12.lcpi_spain`.

Block summary

```
1: {S12.lcpi_spain:S12.ldollar_euro S12.loil_price _cons}
2: {S12e.lcpi_spain:sigma2}
```

Expect an acceptance rate neither too small nor too large.

```
. bayes, rseed(1) nomodelsummary:      ///
> regress S12.lcpi_spain S12.l-dollar_euro S12.loil_price
```

Burn-in ...

Simulation ...

Bayesian linear regression

Random-walk Metropolis-Hastings sampling

```
MCMC iterations = 12,500
Burn-in         = 2,500
MCMC sample size = 10,000
Number of obs   = 231
Acceptance rate = .3295
Efficiency: min = .07405
               avg = .1152
               max = .2197
```

Log marginal-likelihood = 550.31151

	Mean	Std. dev.	MCSE	Median	Equal-tailed [95% cred. interval]	
S12.lcpi_s~n						
ldollar_euro						
S12.	-.0849912	.0164017	.000603	-.0849749	-.1163033	-.0524402
loil_price						
S12.	.0422695	.0038324	.00013	.042151	.0346927	.0497049
_cons	.0193729	.0012802	.000045	.0193937	.0168696	.0218113
S12e.lcpi__n						
sigma2	.0003725	.0000344	7.3e-07	.0003711	.0003117	.0004445

Note: Default priors are used for model parameters.

Outline

Overview

Fundamental

equation

MCMC

Convergence

Stata tools

Linear

regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Postestimation

Evaluate efficiency `bayesstats ess`

- Correlation time estimates the lag after which autocorrelation in an MCMC sample is small.

```
. bayesstats ess, vsquish
```

```
Efficiency summaries      MCMC sample size =      10,000
                          Efficiency:  min =      .07405
                                      avg =      .1152
                                      max =      .2197
```

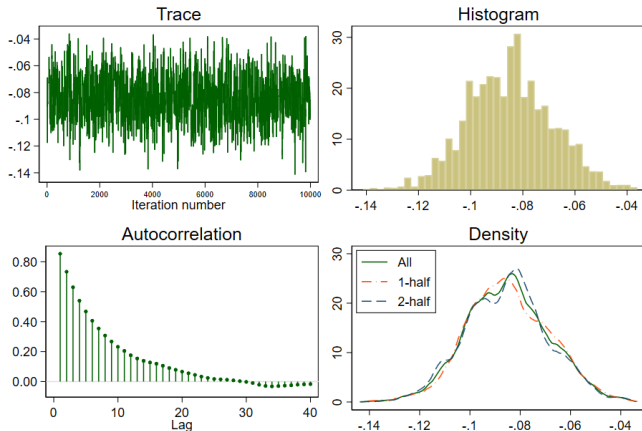
	ESS	Corr. time	Efficiency
S12.lcpi_s_n			
ldollar_euro			
S12.	740.49	13.50	0.0740
loil_price			
S12.	869.96	11.49	0.0870
_cons	801.31	12.48	0.0801
S12e.lcpi_n			
sigma2	2196.85	4.55	0.2197

- Efficiency around 10% or more is considered Ok for MH. Efficiencies under 1% would be a source of concern.

Look at the trace, correlation time, and density.

```
bayesgraph diagnostic {S12.l dollar_euro}
```

S12.lcpi_spain:S12.l dollar_euro

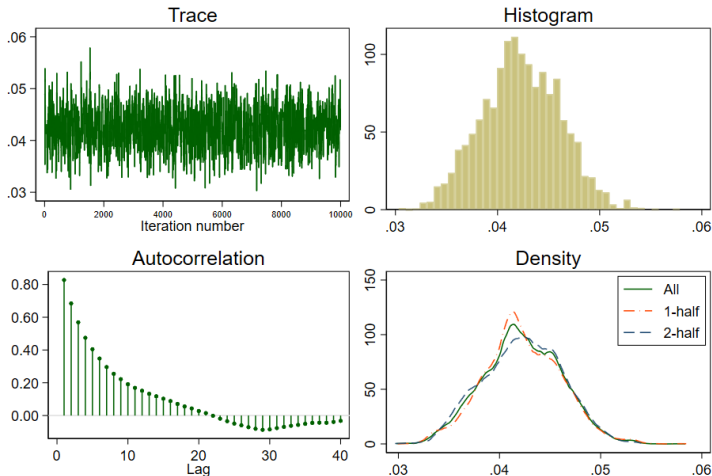


- Trace indicates that convergence was achieved
- Correlation becomes negligible after 15 periods

Diagnostic plots suggest convergence for S12.loil_price

```
. bayesgraph diagnostic {S12.loil_price}
```

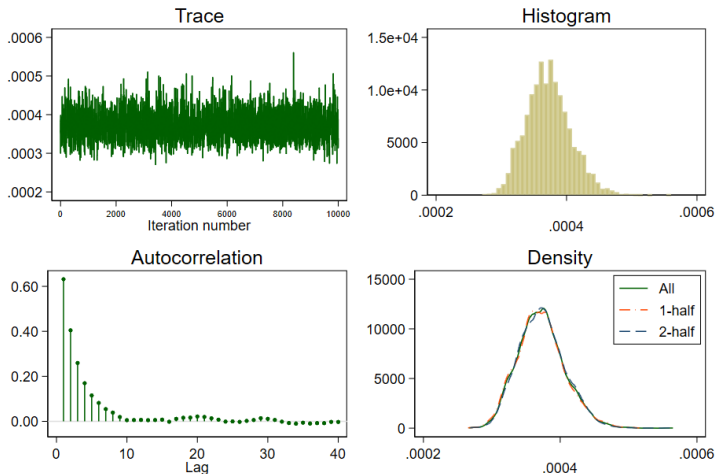
S12.lcpi_spain:S12.loil_price



Diagnostic plots indicate convergence for sigma2

```
. bayesgraph diagnostic {sigma2}
```

S12e.lcpi_spain:sigma2



Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

**Multiple Markov
Chains**

bayestestmodel

Forecasting
probabilities

Summary

References

Multiple Markov Chains

- Convergence requires the chains to be stationary and well-mixed.
- Performing the estimation on multiple chains allows checking for convergence
- In general, three to four chains should be enough to check for convergence

Outline

Overview

Fundamental

equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov
Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Gelman–Rubin convergence diagnostic statistic (R_c)

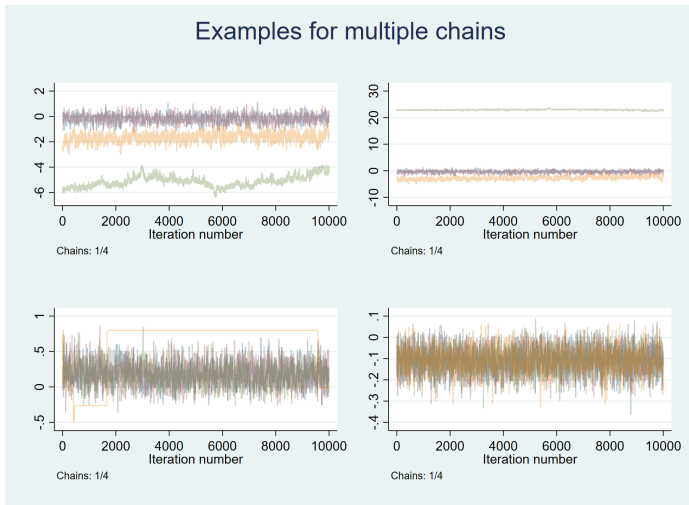
- Compares V (weighted average of between-chains, B , and within-chains variances, W) to variance W :

$$R_c \propto \sqrt{V/W}$$

- R_c greater than 1.1 indicates convergence problems.

Trace for multiple chains

- We expect to see similar trace plots for all the chains:



Multiple chains with `bayes : prefix`

```
. bayes, rseed(1) nchains(3) notable:      ///
> regress S12.lcpi S12.ldollar_euro S12.loil_price
```

```
Chain 1
  Burn-in ...
  Simulation ...
```

```
Chain 2
  Burn-in ...
  Simulation ...
```

```
Chain 3
  Burn-in ...
  Simulation ...
```

```
Model summary
```

Likelihood:

```
S12.lcpi_spain ~ regress(xb_S12.lcpi_spain, {S12e.lcpi_spain:sigma2})
```

Priors:

```
{S12.lcpi_s_n:S12.ldollar_euro S12.loil_price_cons} ~ normal(0,10000)    (1)
{S12e.lcpi_n:sigma2} ~ igamma(.01,.01)
```

(1) Parameters are elements of the linear form `xb_S12.lcpi_spain`.

Bayesian linear regression	Number of chains	=	3
Random-walk Metropolis-Hastings sampling	Per MCMC chain:		
	Iterations	=	12,500
	Burn-in	=	2,500
	Sample size	=	10,000
	Number of obs	=	231
	Avg acceptance rate	=	.3515
	Avg efficiency: min	=	.07516
	avg	=	.1096
	max	=	.1956
Avg log marginal-likelihood =	550.31207	Max Gelman-Rubin Rc =	1.002

Note: Default initial values are used for multiple chains.

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov
Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Multiple chains with `bayes :` prefix

```
. bayes, rseed(1) nchains(3) noheader:           ///
> regress S12.lcpi S12.ldollar_euro S12.loil_price
```

Chain 1

```
    Burn-in ...
    Simulation ...
```

Chain 2

```
    Burn-in ...
    Simulation ...
```

Chain 3

```
    Burn-in ...
    Simulation ...
```

	Mean	Std. dev.	MCSE	Median	Equal-tailed [95% cred. interval]	
S12.lcpi_s_n						
ldollar_euro						
S12.	-.0842184	.0164626	.000316	-.084313	-.1160499	-.0518913
loil_price						
S12.	.0421442	.003805	.000079	.0420887	.0346711	.0494671
_cons						
S12.	.0193879	.0012867	.000027	.0194076	.016854	.0218731
S12e.lcpi__n						
sigma2	.0003721	.0000345	4.5e-07	.00037	.0003103	.0004445

Note: Default initial values are used for multiple chains.

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov
Chains

bayestestmodel

Forecasting
probabilities

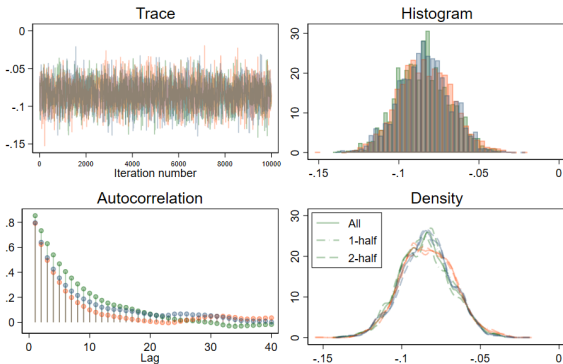
Summary

References

We expect to see similar diagnostic plots for all the chains:

```
. bayesgraph diagnostic {S12.1dollar_euro}
```

S12.lcpi_spain:S12.1dollar_euro



Chains: 1/3

- Trace indicates that convergence was achieved
- Correlation becomes negligible after 10 periods.

Postestimation command to compare models.

- `bayestest model` computes posterior probabilities for each model.
- The result indicates which model is more likely.
- It requires that the models use the same data and that they have proper posterior.
- It can be used to compare models with:
 - Different priors and/or different posterior distributions.
 - Different regression functions.
 - Different covariates.
- Verify MCMC convergence before comparing the models.

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

Higher posterior probability for model "lagged"

```
. bayestest model interannual lagged
```

Bayesian model tests

	log (ML)	P (M)	P (M y)
interannual	550.3115	0.5000	0.0122
lagged	554.7053	0.5000	0.9878

Note: Marginal likelihood (ML) is computed using Laplace-Metropolis approximation.

- What if we assign different priors for each model:

```
. bayestest model interannual lagged,prior(.75 .25)
```

Bayesian model tests

	log (ML)	P (M)	P (M y)
interannual	550.3115	0.7500	0.0357
lagged	554.7053	0.2500	0.9643

Note: Marginal likelihood (ML) is computed using Laplace-Metropolis approximation.

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov
Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Higher posterior probability for model "lagged"

```
. bayestest model interannual lagged
```

Bayesian model tests

	log (ML)	P (M)	P (M y)
interannual	550.3115	0.5000	0.0122
lagged	554.7053	0.5000	0.9878

Note: Marginal likelihood (ML) is computed using Laplace-Metropolis approximation.

- What if we assign different priors for each model:

```
. bayestest model interannual lagged,prior(.75 .25)
```

Bayesian model tests

	log (ML)	P (M)	P (M y)
interannual	550.3115	0.7500	0.0357
lagged	554.7053	0.2500	0.9643

Note: Marginal likelihood (ML) is computed using Laplace-Metropolis approximation.

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

**Forecasting
probabilities**

Summary

References

Forecasting probabilities

Can we compute probabilities for events associated to multiple equation forecasts?

Outline

Overview

Fundamental

equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

```
. collect preview
```

	2010Q1	2010Q1	2010Q1-Q2
Event			
inflation_over_5=1	.6778		
infl_over_5_exchrte_chg_over_3=1		.4554	
exchrte_change_over_3=1			.4303

- Frequentist approach: use Probability forecasting (Garraat et al. (2006)).
 - Forecasts subject to uncertainty about the future, the parameters, the model, and alternative policies.
 - See an example in Sanchez and Zavarce (2013) for probability forecast accounting for future uncertainty.
- The Bayesian approach allows obtaining probabilities for events based on parameters and future uncertainty.

Can we compute probabilities for events associated to multiple equation forecasts?

Outline

Overview

Fundamental

equation

MCMC

Convergence

Stata tools

Linear regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

```
. collect preview
```

	2010Q1	2010Q1	2010Q1-Q2
Event			
inflation_over_5=1	.6778		
infl_over_5_exchrte_chg_over_3=1		.4554	
exchrte_change_over_3=1			.4303

- Frequentist approach: use Probability forecasting (Garraat et al. (2006)).
 - Forecasts subject to uncertainty about the future, the parameters, the model, and alternative policies.
 - See an example in Sanchez and Zavarce (2013) for probability forecast accounting for future uncertainty.
- The Bayesian approach allows obtaining probabilities for events based on parameters and future uncertainty.

Bayesian forecasting

- We want to predict y^{new} given the observed data y .
- The distribution of y^{new} is called *posterior predictive distribution*,

$$\begin{aligned}\Pr(y^{\text{new}}|y) &= \int \Pr(y^{\text{new}}|\theta) \Pr(\theta|y) \\ &= \int f(y^{\text{new}}; \theta) p(\theta|y) d\theta\end{aligned}$$

- So, to obtain y^{new} , we can draw θ^* from its posterior distribution $p(\theta|y)$ and then draw y^{new} according to the likelihood model $f(y^{\text{new}}; \theta^*)$.

Outline

Overview

Fundamental

equation

MCMC

Convergence

Stata tools

Linear regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

Example 3: Predicted outcome for two-equation model

- Let's now fit a 2 equation linear regression for the Bitcoin price and the Standard and Poors US stock market index (S&P500) as a function of the Chicago Board Options Exchange Volatility Index (VIX).

$$\Delta \text{bitcoin} = \alpha_0 + \alpha_{vix} * \Delta \text{vix} + \epsilon_1$$

$$\Delta \text{sp500} = \beta_0 + \beta_{vix} * \Delta \text{vix} + \epsilon_2$$

Where:

bitcoin : daily change for the bitcoin price.

sp500 : daily change for the Standard & Poors 500 index.

vix : daily change for the Chicago Board Options Exchange Volatility Index (VIX).

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

We use `import fred` to load the data

```
. import fred SP500 CBBTCUSD VIXCLS, ///  
>      daterange(2017-01-03 2027-01-27) ///  
>      aggregate(daily) clear  
. rename SP500      sp500  
. rename VIXCLS      vix  
. rename CBBTCUSD    bitcoin  
. tsset daten,daily  
. generate bdate=bofd("bcaljan30",daten)  
. format bdate %tbbcaljan30  
. tsset bdate  
. keep if bdate~=.  
. tsappend,add(3)  
. tsset bdate  
. foreach var of varlist sp500 bitcoin vix {  
.      generate d_`var'=D.`var'  
. }  
.
```

[Outline](#)[Overview](#)[Fundamental](#)[equation](#)[MCMC](#)[Convergence](#)[Stata tools](#)[Linear](#)[regression](#)[Model fit](#)[Postestimation](#)[bayesstats ess](#)[bayesgraph](#)[Multiple Markov](#)[Chains](#)[bayestestmodel](#)[Forecasting](#)[probabilities](#)[Summary](#)[References](#)

Data for bitcoin price, S&P500, and VIX

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting

probabilities

Summary

References

```
. describe
```

Contains data

Observations: 2,289

Variables: 8

Variable name	Storage type	Display format	Value label	Variable label
bdate	float	%tb..		
daten	int	%td		numeric (daily) date
sp500	float	%9.0g		S&P 500
bitcoin	float	%9.0g		Coinbase Bitcoin
vix	float	%9.0g		CBOE Volatility Index: VIX
d_sp500	float	%9.0g		
d_bitcoin	float	%9.0g		
d_vix	float	%9.0g		

Sorted by: bdate

Note: Dataset has changed since last saved.

Outline

Overview

- Fundamental equation
- MCMC
- Convergence
- Stata tools

Linear regression

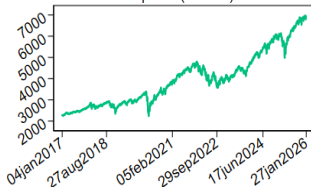
- Model fit
- Postestimation
- bayesstats ess
- bayesgraph
- Multiple Markov Chains
- bayestestmodel

- Forecasting probabilities

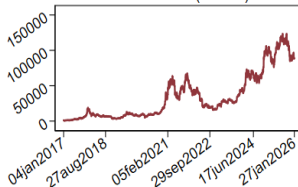
Summary

References

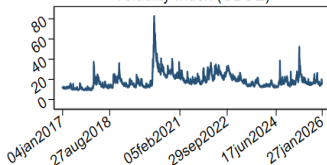
sp500 (Levels)



USD/Bitcoin (Levels)



Chicago Board Options Exchange Market
Volatility Index (CBOE)



Source: Federal Reserve Economic Data (FRED)

Downloaded using -import fred- in Stata

Bitcoin price, S&P500, and VIX (First difference)

Outline

Overview

Fundamental
equation
MCMC
Convergence
Stata tools

Linear regression

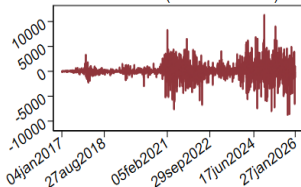
Model fit
Postestimation
bayesstats ess
bayesgraph
Multiple Markov
Chains
bayestestmodel

Forecasting probabilities

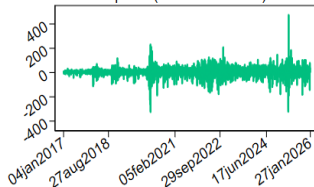
Summary

References

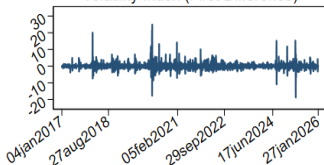
USD/Bitcoin (First Difference)



sp500 (First Difference)



Chicago Board Options Exchange Market
Volatility Index (First Difference)



Source: Federal Reserve Economic Data (FRED)
Downloaded using -import fred- in Stata

Multivariate regression for Bitcoin and S&P500 on vix

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov
Chains

bayestestmodel

Forecasting
probabilities

Summary

References

```
. bayesmh (d_bitcoin d_vix) (d_sp500 d_vix) if tin(02mar2020,20jan2026), ///
>    likelihood(mvnormal({Sigma,m}))          ///
>    prior({d_bitcoin:} {d_sp500:}, normal(0,10000))    ///
>    prior({Sigma,m}, jeffreys(2))                ///
>    block({d_sp500:} {d_bitcoin:}) block({Sigma,m})    ///
>    rseed(1) saving("$simul_dir\my_mcm",replace) nomodelsumm
```

Burn-in ...

Simulation ...

Bayesian multivariate normal regression	MCMC iterations	=	12,500
Random-walk Metropolis-Hastings sampling	Burn-in	=	2,500
	MCMC sample size	=	10,000
	Number of obs	=	1,477
	Acceptance rate	=	.2415
	Efficiency: min	=	.04936
	avg	=	.07376
	max	=	.105

Log marginal-likelihood = -20415.494

	Mean	Std. dev.	MCSE	Median	Equal-tailed [95% cred. interval]	
d_bitcoin						
d_vix	-142.3168	21.64446	.764288	-142.5587	-184.1331	-99.38673
_cons	43.10563	41.83143	1.88276	42.7842	-37.0697	125.0378
d_sp500						
d_vix	-18.33389	.3839057	.014775	-18.33276	-19.10893	-17.58398
_cons	2.551141	.8520962	.028017	2.540353	.7990881	4.196056
Sigma_1_1	3169163	116040.7	4323.5	3166113	2949056	3403165
Sigma_2_1	5639.605	1590.383	71.351	5615.904	2614.42	8803.371
Sigma_2_2	1082.912	40.57586	1.25199	1083.797	1005.431	1164.085

Note: Adaptation tolerance is not met in at least one of the blocks.

file C:\Users\gas\Documents\conferences\Portugal\2026\simul\my_mcm.dta saved.

- Let's first get out of sample values for the exogenous variable VIX.

```
. arima d_vix, ar(1) nolog
```

ARIMA regression

Sample: 04jan2017 thru 04feb2026

Number of obs = 2285

Wald chi2(1) = 844.12

Log likelihood = -4801.732

Prob > chi2 = 0.0000

d_vix	OPG		z	P> z	[95% conf. interval]	
	Coefficient	std. err.				
d_vix						
_cons	.0026312	.0398296	0.07	0.947	-.0754335	.0806959
ARMA						
ar						
L1.	-.1925902	.0066288	-29.05	0.000	-.2055823	-.179598
/sigma	1.978771	.0082383	240.19	0.000	1.962624	1.994917

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

```
. predict dvix_hat, y dynamic(td(20jan2026))
```

```
. replace d_vix=dvix_hat if tin(21jan2026,23jan2026)
```

```
(3 real changes made)
```


- We now use `bayespredict` to get the posterior predictive distribution for our two outcome variables (monthly change in the bitcoin price and the S&P500 index)

```
. qui bayesmh (d_bitcoin d_vix)                                ///
>      (d_sp500 d_vix)   if tin(02mar2020,20jan2026), ///
>      likelihood(mvnormal({Sigma,m}))                ///
>      prior({d_bitcoin:} {d_sp500:}, normal(0,10000)) ///
>      prior({Sigma,m},                                jeffreys(2)) ///
>      block({d_sp500:} {d_bitcoin:}) block({Sigma,m}) ///
>      rseed(1) saving("$simul_dir\my_mcm", replace)

.
. bayespredict bt_change:{_ysim} sp_change:{_ysim2}  ///
>      if tin(21jan2026,23jan2026), rseed(1)          ///
>      saving("$simul_dir\mypred", replace)

Computing predictions ...
file C:\Users\gas\Documents\conferences\Portugal\2026\simul\mypred.dta saved.
file C:\Users\gas\Documents\conferences\Portugal\2026\simul\mypred.ster saved.
```

Outline

Overview

Fundamental

equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov

Chains

bayestestmodel

Forecasting
probabilities

Summary

References

- With the posterior predictive distributions we can estimate probabilities for events associated to our outcome variables.
- Let's look at the probabilities for single events like positive change in the bitcoin price and the S&P500 index for individual dates:

```
. bayesstats summary (Bt_Ch_Jan_21:{_ysim1[2276]}>0) ///
> (Bt_Ch_Jan_22:{_ysim1[2277]}>0) ///
> (SP_Ch_Jan_21:{_ysim2[2276]}>0) ///
> (SP_Ch_Jan_22:{_ysim2[2277]}>0) ///
> using "$simul_dir\mypred"
```

```
Posterior summary statistics                                MCMC sample size =    10,000
Bt_Ch_Jan_21 : _ysim1_2276>0
Bt_Ch_Jan_22 : _ysim1_2277>0
SP_Ch_Jan_21 : _ysim2_2276>0
SP_Ch_Jan_22 : _ysim2_2277>0
```

	Mean	Std. dev.	MCSE	Median	Equal-tailed [95% cred. interval]	
Bt_Ch_Jan_21	.5357	.4987488	.004987	1	0	1
Bt_Ch_Jan_22	.4868	.4998507	.004999	0	0	1
SP_Ch_Jan_21	.7048	.4561554	.004562	1	0	1
SP_Ch_Jan_22	.3934	.4885287	.004885	0	0	1

- We can also look at probabilities for combined events like joint positive changes in the bitcoin price and the S&P500 index for individual or multiple dates:

```
. bayesstats summary                                     ///
>   (Bt_Ch_21_23:{_ysim1[2276]}>0 & {_ysim1[2277]}>0 & {_ysim1[2278]}>0) ///
>   (SP_Ch_21_23:{_ysim2[2276]}>0 & {_ysim2[2277]}>0 & {_ysim2[2278]}>0) ///
>   (Bt_SP_ch_21:{_ysim1[2276]}>0 & {_ysim2[2277]}>0)      ///
>   using "$simul_dir\mypred"
```

```
Posterior summary statistics                                MCMC sample size =    10,000
Bt_Ch_21_23 : _ysim1_2276>0 & _ysim1_2277>0 & _ysim1_2278>0
SP_Ch_21_23 : _ysim2_2276>0 & _ysim2_2277>0 & _ysim2_2278>0
Bt_SP_ch_21 : _ysim1_2276>0 & _ysim2_2277>0
```

	Mean	Std. dev.	MCSE	Median	Equal-tailed [95% cred. interval]	
Bt_Ch_21_23	.1335	.3401313	.003396	0	0	1
SP_Ch_21_23	.126	.331866	.003319	0	0	1
Bt_SP_ch_21	.207	.4051758	.004126	0	0	1

Customized probability report

- Combine `bayesstats summary` with the `collect` commands to customize your report:

```
. collect clear
. forvalues i=2276/2278 {
.   local j=`i'-2275
.   collect
.   BT=r(summary) [1,1] SP=r(summary) [2,1]
.   BT_SP=r(summary) [3,1], tags(month[j']):
.   bayesstats summary ({_ysim1[`i']}>0) ({_ysim2[`i']}>0)
.   ({_ysim1[`i']}>0 & {_ysim2[`i']}>0) using "$simul_dir\mypred"
. }
. collect BT=r(summary) [1,1] SP=r(summary) [2,1]
. BT_SP=r(summary) [3,1], tags(month[4]): bayesstats summary
. ({_ysim1[2276]}>0 & {_ysim1[2277]}>0 & {_ysim1[2278]}>0)
. ({_ysim2[2276]}>0 & {_ysim2[2277]}>0 & {_ysim2[2278]}>0)
. ({_ysim1[2276]}>0 & {_ysim1[2277]}>0 & {_ysim1[2278]}>0 &
.   {_ysim2[2276]}>0 & {_ysim2[2277]}>0 & {_ysim2[2278]}>0)
. using "$simul_dir\mypred"
. quietly collect layout (result) (month)
. collect label levels month 1 "Jan 21" 2 "Jan 22"
. collect label levels result BT "Bitcoin change +"
. SP "S&P500 change +"
. BT_SP "Bitcoin and S&P500 change +"
. collect style header result, title(label)
. collect style cell result, nformat(%6.2f)
. collect style column, extraspace(1)
```

Customized probability report

```
. collect preview
```

	Jan 21	Jan 22	Jan 23	Jan 21-23
Result				
Bitcoin change +	0.54	0.49	0.50	0.13
S&P500 change +	0.70	0.39	0.47	0.13
Bitcoin and S&P500 change +	0.38	0.19	0.24	0.02

Outline

Overview

Fundamental
equation

MCMC

Convergence

Stata tools

Linear
regression

Model fit

Postestimation

bayesstats ess

bayesgraph

Multiple Markov
Chains

bayestestmodel

Forecasting
probabilities

Summary

References

Summing up

- Bayesian analysis: A statistical approach that can be used to answer questions about unknown parameters in terms of probability statements.
- It can be used when we have prior information on the distribution of the parameters involved in the model.
- Alternative approach or complementary approach to classic/frequentist approach?

Reference

Cameron, A. and Trivedi, P. 2005. *Microeconometric Methods and Applications*. Cambridge University Press, Section 13.2.2, 422–423.

Links

https://www.stata.com/meeting/uk17/slides/uk17_Marchenko.pdf

<https://www.stata.com/meeting/brazil16/slides/rising-brazil16.pdf>

https://www.stata.com/meeting/spain18/slides/spain18_Sanchez.pdf

https://www.stata.com/meeting/mexico21/slides/Mexico21_Sanchez.pdf