Introduction to Bayesian Analysis in Stata

Gustavo Sánchez

StataCorp LLC

October 24 , 2018
Barcelona, Spain

Outline

1. Bayesian analysis: Basic Concepts
   - The general idea
   - The Method
2. The Stata Tools
   - The general command `bayesmh`
   - The `bayes` prefix
   - Postestimation Commands
3. A few examples
   - Linear regression
   - Panel data random effect probit model
   - Change point model

The general idea

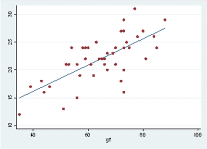
Frequentist

Theoretical Model →

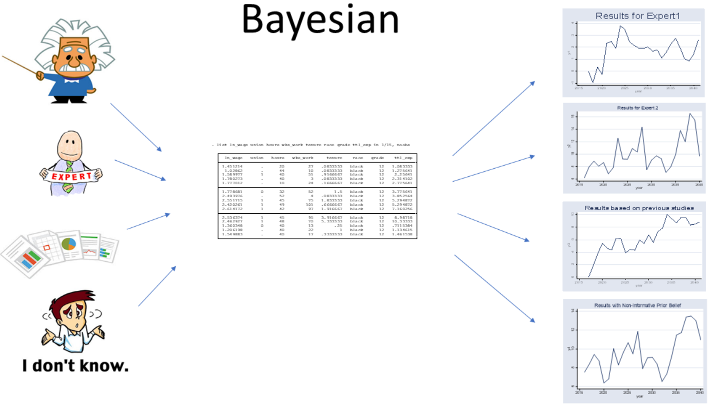
```
. list ln_wage union hours wks_work tenure race grade ttl_exp in 1/15, noobs
```

ln_wage	union	hours	wks_work	tenure	race	grade	ttl_exp
1.451214	.	20	27	.0833333	black	12	1.083333
1.02862	.	44	10	.0833333	black	12	1.275641
1.589977	1	40	51	.9166667	black	12	2.25641
1.780273	.	40	3	.0833333	black	12	2.314102
1.777012	.	10	24	.1666667	black	12	2.775641
1.778681	0	32	52	1.5	black	12	3.775641
2.493976	.	52	4	.0833333	black	12	3.852564
2.551715	1	45	75	1.833333	black	12	5.294872
2.420261	1	49	101	.6666667	black	12	5.294872
2.614172	1	42	97	1.916667	black	12	7.160256
2.536374	1	45	95	3.916667	black	12	8.98718
2.462927	1	48	70	5.333333	black	12	10.33333
1.360348	0	40	13	.25	black	12	.7115384
1.206198	.	40	22	1	black	12	1.134615
1.549883	.	40	17	.3333333	black	12	1.461538

→



The general idea



Bayesian Analysis vs Frequentist Analysis

Frequentist Analysis

- Estimate unknown fixed parameters.
- Data for a (hypothetical) repeatable random sample.
- Uses data to estimate unknown fixed parameters.
- Data expected to satisfy the assumptions for the specified model.

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

Bayesian Analyis

- Probability distributions for unknown random parameters
- The data is fixed.
- Combines data with prior beliefs to get probability distributions for the parameters.
- Posterior distribution is used to make explicit probabilistic statements.

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

Stata's simple syntax: `bayes:`

```
regress y x1 x2 x3

bayes: regress y x1 x2 x3


logit y x1 x2 x3

bayes: logit y x1 x2 x3


mixed y x1 x2 x3 || region:

bayes: mixed y x1 x2 x3 || region:
```

The Method

The Method

- Inverse law of probability (Bayes' Theorem):

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

- Marginal distribution of y, f(y), does not depend on (θ)

- We can then write the fundamental equation for Bayesian analysis:

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

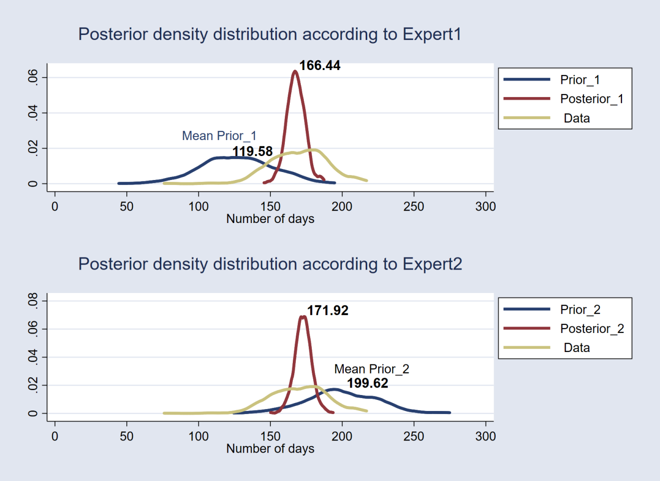
The Method

- Let's assume that both, the data and the prior beliefs, are normally distributed:
  - **The data**: $y \sim N\left(\theta, \sigma_d^2\right)$
  - **The prior**: $\theta \sim N\left(\mu_p, \sigma_p^2\right)$
- Homework...: Doing the algebra with the fundamental equation we find that the posterior distribution would be normal with (see for example Cameron & Trivedi 2005):
  - **The posterior**: $\theta | y \sim N\left(\mu, \sigma^2\right)$

  Where:

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

Example (Posterior distributions)



The Method

- The previous example has a closed form solution.
- What about the cases with non-closed solutions, or more complex distributions?
  - Integration is performed via simulation
  - We need to use intensive computational simulation tools to find the posterior distribution in most cases.
  - Markov chain Monte Carlo (MCMC) methods are the current standard in most software. Stata implement two alternatives:
    - Metropolis-Hastings (MH) algorithm
    - Gibbs sampling

The Method

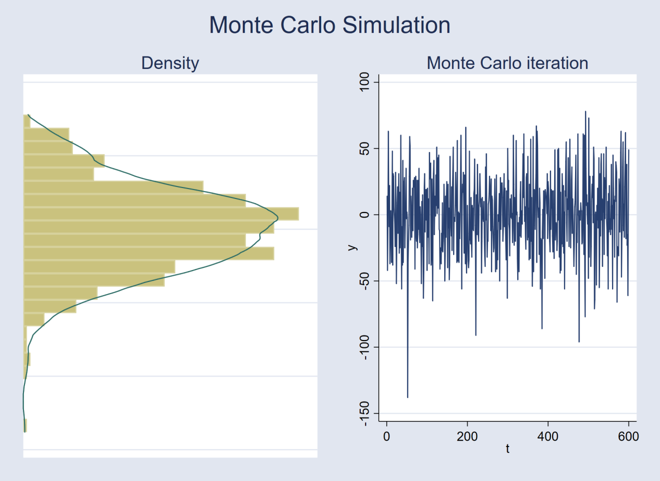
- Links for Bayesian analysis and MCMC on our youtube channel:
  - Introduction to Bayesian statistics, part 1: The basic concepts

    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

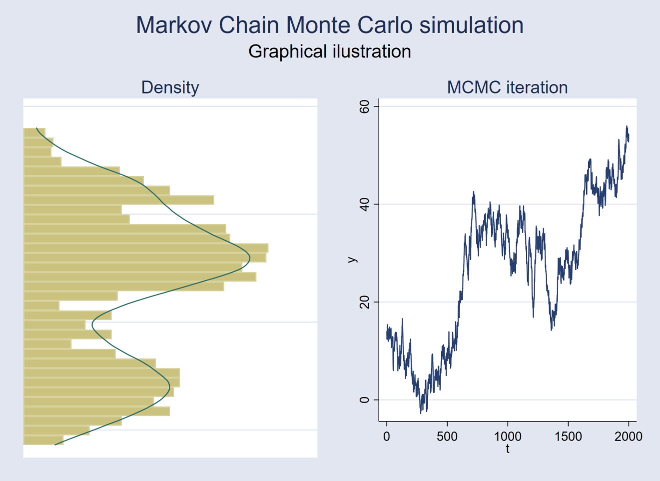
The Method

- Monte Carlo Simulation



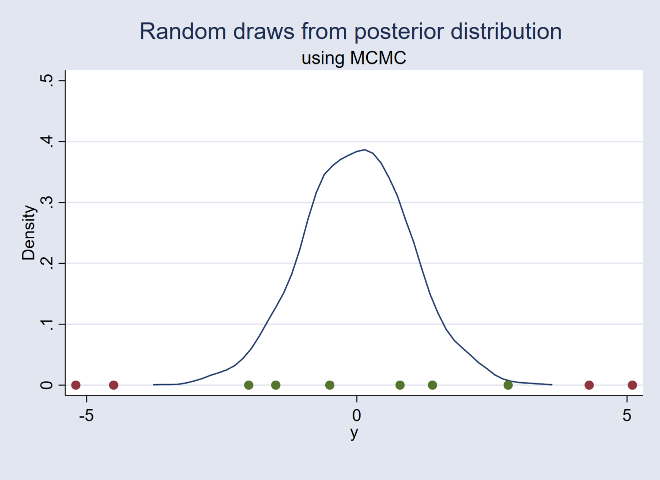
The Method

- Markov Chain Monte Carlo Simulation



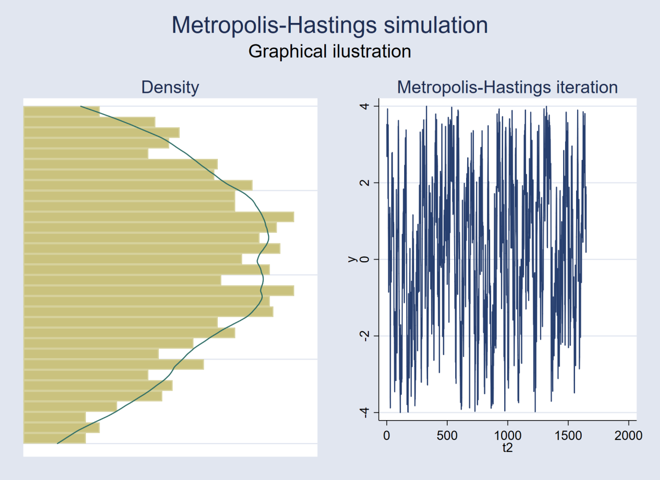
The Method

- Metropolis Hastings intuitive idea
  - Green points represent accepted proposal states and red points represent rejected proposal states.



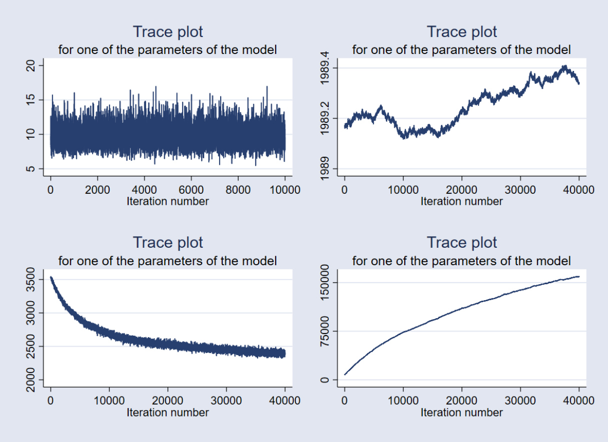
The Method

- Metropolis Hastings simulation
  - The trace plot illustrates the sequence of accepted proposal states.



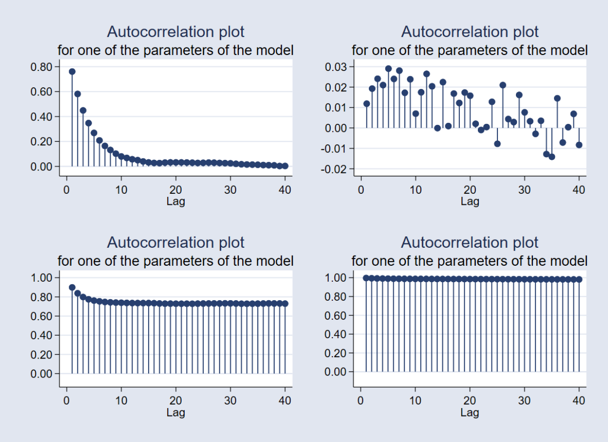
The Method

- We expect to obtain a stationary sequence when convergence is achieved.



The Method

- An efficient MCMC should have small autocorrelation.
- We expect autocorrelation to become negligible after a few lags.



The Stata tools for Bayesian regression

The Stata tools: `bayesmh` & `bayes:`

- `bayesmh` General purpose command for Bayesian analysis
  - You need to specify all the components for the Bayesian regression: Likelihood, priors, hyperpriors, blocks, etc
- `bayes:` Simple syntax for Bayesian regressions
  - Estimation command defines the likelihood for the model.
  - Default priors are assumed to be "noninformative"'.
  - Other model specifications are set by default depending on the model defined by the estimation command.
  - Alternative specifications may need to be evaluated.

The Stata tools: Postestimation commands

- Bayesstats ess
- Bayesgraph
- Bayesstats ic
- Bayestest model
- Bayestest interval
- Bayesstats summary

Examples

Example 1: Linear Regression

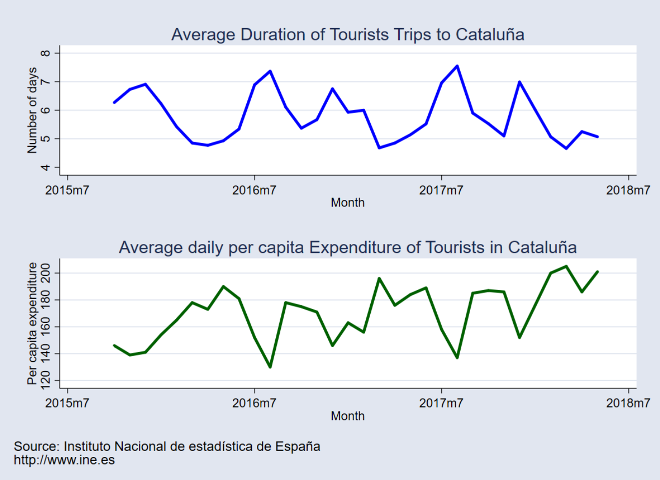
- Let's look at our first example:
  - We have stats on the average number of days tourists spend in Cataluña and their average per capita expenditure.
  - We fit a linear regression for the average number of days.
  - Let's consider two specifications:

$$\begin{aligned} tripdays &= \alpha_1 + \beta_{day} * capexp\_day + \epsilon_1 \\ tripdays &= \alpha_2 + \beta_{avg} * avgexp\_cap + \epsilon_2 \end{aligned}$$

Where:

tripdays : Number of days tourists spend in Cataluña.
capexp_day: Tourists' daily per capita expenditure.
avgexp_cap: Tourists' total per capita expenditure.

Example 1: Linear Regression



Example 1: Linear Regression

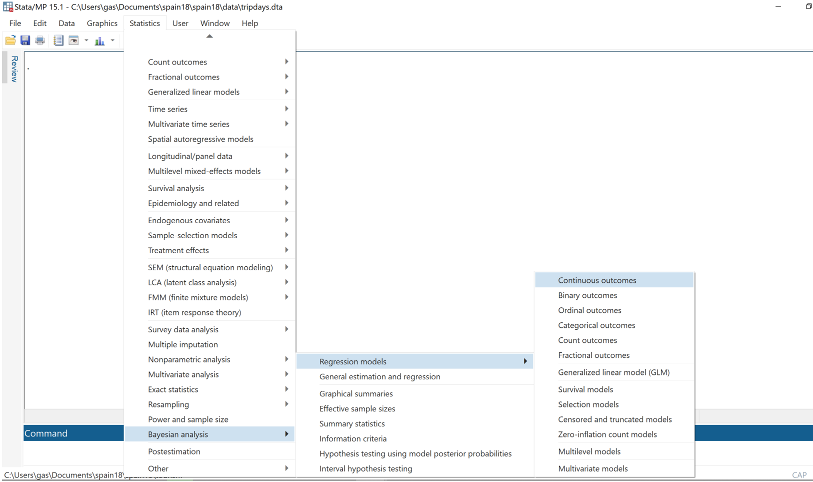
- Linear regression with the `bayes`: prefix

```
bayes ,rseed(123): regress tripdays capex_day
```

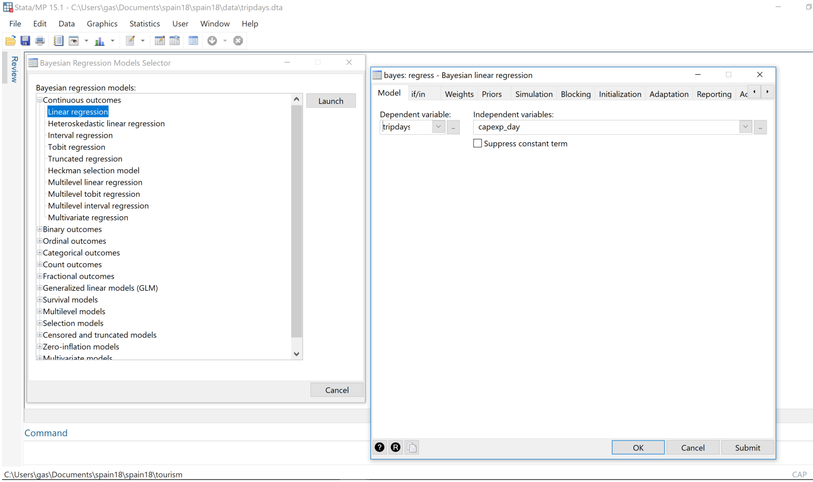
- Equivalent model with `bayesmh`

```
bayesmh tripdays capexp_day, rseed(123)              ///
    likelihood(normal(sigma2))                       ///
    prior(tripdays:capexp_day, normal(0,10000))      ///
    prior(tripdays:_cons, normal(0,10000))           ///
    prior(sigma2, igamma(.01,.01))                   ///
    block(tripdays:capexp_day _cons)                 ///
    block(sigma2)
```

Example 1: Menu for Bayesian regression



Example 1: Menu for Bayesian regression



Example 1: 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 (tripdays) and the explanatory variable (capex_day)
6. Click on 'OK'

Example 1: bayes: prefix

```
. bayes ,rseed(123) blocksummary:regress tripdays capexp_day

Burn-in ...
Simulation ...
Model summary
------------------------------------------------------------------------------
Likelihood:
  tripdays ~ regress(xb_tripdays,{sigma2})
Priors:
  {tripdays:capexp_day _cons} ~ normal(0,10000)
                     {sigma2} ~ igamma(.01,.01)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_tripdays.
Block summary
------------------------------------------------------------------------------
   1:  {tripdays:capexp_day _cons}
   2:  {sigma2}
------------------------------------------------------------------------------
```

Example 1: `bayes:` prefix

```
. bayes ,rseed(123) blocksummary:regress tripdays capexp_day

Bayesian linear regression                        MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling          Burn-in          =      2,500
                                                  MCMC sample size =     10,000
                                                  Number of obs    =          5
                                                  Acceptance rate  =      .3799
                                                  Efficiency:  min =     .03477
                                                               avg =     .08801
Log marginal likelihood = -16.207649                           max =      .1146
```

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred.	Interval]
tripdays						
capexp_day	-.0383973	.0128253	.000379	-.0377857	-.0647811	-.0122725
_cons	12.64544	2.484331	.073378	12.52498	7.610854	17.84337
sigma2	.0926729	.0928459	.004979	.0616775	.0151486	.3563017

Note: Default priors are used for model parameters.

Example 1: bayesstats ess

- Let's evaluate the effective sample size

```
. bayesstats ess

Efficiency summaries    MCMC sample size =    10,000
```

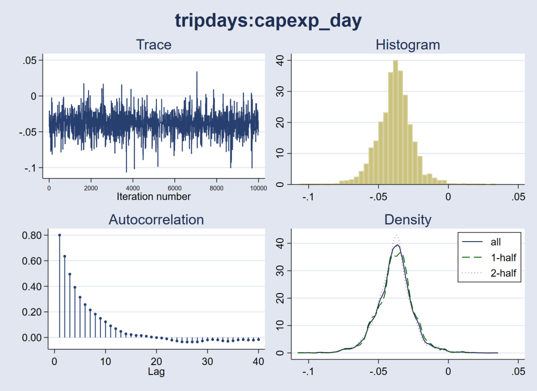
	ESS	Corr. time	Efficiency
tripdays			
capexp_day	1146.26	8.72	0.1146
_cons	1146.27	8.72	0.1146
sigma2	347.72	28.76	0.0348

- We expect to have an acceptance rate (see previous slide) that is neither to small nor too large.
- We also expect to have low correlation
- Efficiencies over 10% are considered good for MH. Efficiencies under 1% would be a source of concern.

Example 1: `bayesgraph`

- We can use `bayesgraph` to look at the trace, the correlation, and the density. For example:

. bayesgraph diagnostic {capex_day}

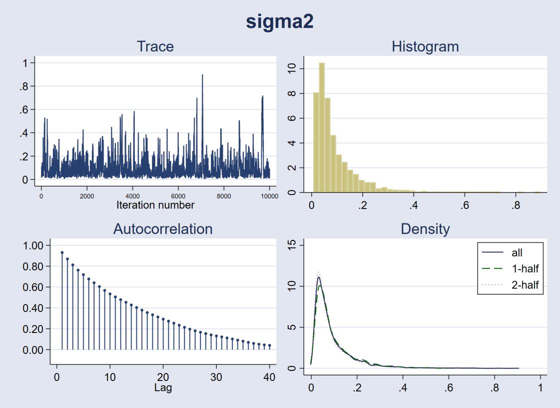


- The trace indicates that convergence was achieved
- Correlation becomes negligible after 10 periods

Example 1: `bayesgraph`

- We can use `bayesgraph` to look at the trace, the correlation, and the density. For example:

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



- Correlation is still persistent after 10 periods

Example 1: **thinning()**

- We can reduce autocorrelation by using thinning
- Save the random draws skipping a prespecified number of simulated values in the MCMC iteration process.
- Use the option 'thinning(#)' to indicate that Stata should save simulated values from every (1+k*#)th iteration (k=0,1,2,...).

```
bayes ,nomodelsummary nodots rseed(123) ///
thinning(4): regress tripdays capexp_day
```

Example 1: thining()

```
. bayes ,rseed(123) nomodelsummary thinning(4):   ///
>          regress tripdays capexp_day

note: discarding every 3 sample observations; using observations 1,5,9,...

Burn-in ...
Simulation ...
Bayesian linear regression                        MCMC iterations  =     42,497
Random-walk Metropolis-Hastings sampling          Burn-in          =      2,500
                                                  MCMC sample size =     10,000
                                                  Number of obs    =          5
                                                  Acceptance rate  =      .3773
                                                  Efficiency:  min =      .1052
                                                               avg =       .313
Log marginal likelihood = -16.191209                           max =       .418
```

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred.	Interval]
tripdays						
capexp_day	-.0384152	.0126655	.000196	-.0383658	-.0636837	-.0126029
_cons	12.64972	2.455834	.037984	12.62602	7.628034	17.57862
sigma2	.0917518	.0951007	.002932	.0605151	.0151486	.3519349

Note: Default priors are used for model parameters.

Example 1: bayesstats ess

- Let's evaluate again the effective sample size

```
. bayesstats ess

Efficiency summaries    MCMC sample size =     10,000
```

	ESS	Corr. time	Efficiency
tripdays			
capexp_day	4159.44	2.40	0.4159
_cons	4180.27	2.39	0.4180
sigma2	1051.71	9.51	0.1052

- The efficiency improved for all the parameters.
- Correlation time was significantly reduced.

Example 1: `bayestest model`

- `bayestest model` is another postestimation command to compare different models.
- `bayestest model` computes the 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
- MCMC convergence should be verified before comparing the models.

Example 1: bayestest model

- Let's fit now two other models and compare them to the one we already fitted.
- We store the results for the three models and we use the postestimation command `bayestest model` to select one of them.

```
quietly {
    bayes , rseed(123) saving(pcap,replace):        ///
          regress tripdays capexp_day
    estimates store daily

    bayes , rseed(123) saving(total,replace):       ///
          regress tripdays avgexp_cap
    estimates store total

    bayes , rseed(123) saving(media,replace)        ///
          prior(tripdays:_cons, normal(9,.4)):      ///
          regress tripdays
    estimates store mean
}
bayestest model daily total mean
```

Example 1: bayestest model

- Here is the output for `bayestest model`

```
. quietly {
. bayestest model daily total mean
Bayesian model tests
```

	log(ML)	P(M)	P(M|y)
daily	-16.2076	0.3333	0.4997
total	-18.6705	0.3333	0.0426
mean	-16.2955	0.3333	0.4577

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

- We could also assign different priors for the models:

```
. bayestest model daily total mean,  ///
>           prior(.15 0.75 0.1)
Bayesian model tests
```

	log(ML)	P(M)	P(M|y)
daily	-16.2076	0.1500	0.4910
total	-18.6705	0.7500	0.2092
mean	-16.2955	0.1000	0.2998

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

Example 1: `bayestest model`

- Here is the output for `bayestest model`

```
. quietly {
. bayestest model daily total mean
Bayesian model tests
```

	log(ML)	P(M)	P(M|y)
daily	-16.2076	0.3333	0.4997
total	-18.6705	0.3333	0.0426
mean	-16.2955	0.3333	0.4577

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

- We could also assign different priors for the models:

```
. bayestest model daily total mean,  ///
>          prior(.15 0.75 0.1)
Bayesian model tests
```

	log(ML)	P(M)	P(M|y)
daily	-16.2076	0.1500	0.4910
total	-18.6705	0.7500	0.2092
mean	-16.2955	0.1000	0.2998

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

Example 2: Random Effects Probit model

Example 2: Random effects probit model

- Let's use `bayes:` to fit a random effects for a binary variable, whose values depend on a linear latent variable.

$$y_{it} \quad = \quad \beta_0 + \beta_1 x1_{it} + \beta_2 x2_{it} + ... + \beta_k xk_{it} + \alpha_i + \epsilon_{it}$$

Where:

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}* > 0 \\ 0 & \text{otherwise} \end{cases}$$

$\alpha_i \sim N\left(0, \sigma^2_\alpha\right)$ is the individual random panel effect

$\epsilon_{it} \sim N\left(0, \sigma^2_e\right)$ is the idiosyncratic error term

- This is also referred as a two-level random intercept model.
- We can also fit this model with `meprobit` or `xtprobit,re`

Example 2: Random effects probit model

- This time we are going to work with simulated data.
- Here is the code to simulate the panel dataset:

```
    clear
    set obs 100
    set seed 1

* Panel level *
    generate id=_n
    generate alpha=rnormal()
    expand 5

* Observation level *
    bysort id:generate year=_n
    xtset id year
    generate x1=rnormal()
    generate x2=runiform()>.5
    generate x3=uniform()
    generate u=rnormal()

* Generate dependent variable *
    generate y=.5+1*x1+(-1)*x2+1*x3+alpha+u>0
```

Example 2: Random effects probit model

Let's show the results with `meprobit`:

```
. meprobit y x1 x2 x3 || id:,nolog
Mixed-effects probit regression                 Number of obs     =        500
Group variable:             id                  Number of groups  =        100

                                                Obs per group:
                                                              min =          5
                                                              avg =        5.0
                                                              max =          5

Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(3)      =      82.83
Log likelihood = -236.88589                     Prob > chi2       =     0.0000
```

y	Coef.	Std. Err.	P>|z|	[95% Conf.	Interval]
x1	.9769118	.1143889	0.000	.7527138	1.20111
x2	-.9896286	.1853433	0.000	-1.352895	-.6263625
x3	.9426958	.2941061	0.001	.3662584	1.519133
_cons	.5220418	.2187448	0.017	.0933098	.9507738
id					
var(_cons)	1.31	.3835866		.7379508	2.325494

```
LR test vs. probit model: chibar2(01) = 67.24    Prob >= chibar2 = 0.0000
```

Example 2: Random effects probit model

- We now fit the model with `bayes`:

```
bayes , nodots rseed(123) thinning(5) blocksummary: ///
        meprobit y x1 x2 x3 || id:
```

- Equivalent model with `bayesmh`

```
bayesmh y x1 x2 x3, thinning(5) rseed(123)    ///
    likelihood(probit)                         ///
    prior(y:i.id, normal(0,y:var))             ///
    prior(y:x1 x2 x3 _cons, normal(0,10000))   ///
    prior(y:var, igamma(.01,.01))              ///
    block(y:var)                               ///
    blocksummary dots
```

Example 2: Random effects probit model

```
. bayes , nodots rseed(123) thinning(5) blocksummary:
        meprobit y x1 x2 x3 || id:

note: discarding every 4 sample observations; using observations 1,6,11,...
Burn-in ...
Simulation ...
Multilevel structure
------------------------------------------------------------------------------
id
    {U0}: random intercepts
------------------------------------------------------------------------------

Model summary
------------------------------------------------------------------------------
Likelihood:
  y ~ meprobit(xb_y)
Priors:
  {y:x1 x2 x3 _cons} ~ normal(0,10000)                                     (1)
                {U0} ~ normal(0,{U0:sigma2})                               (1)
Hyperprior:
  {U0:sigma2} ~ igamma(.01,.01)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_y.
Block summary
------------------------------------------------------------------------------
   1:  {y:x1 x2 x3 _cons}
   2:  {U0:sigma2}
   3:  {U0[id]:1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
> 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
> 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76
> 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100}
------------------------------------------------------------------------------
```

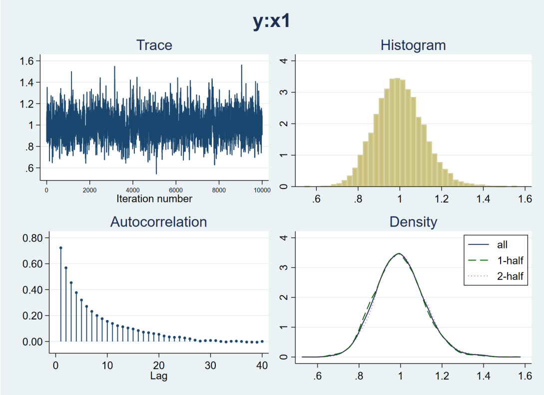
Example 2: Random effects probit model

```
. bayes , nodots rseed(123) thinning(5) blocksummary:
        meprobit y x1 x2 x3 || id:

Bayesian multilevel probit regression            MCMC iterations  =     52,496
Random-walk Metropolis-Hastings sampling         Burn-in          =      2,500
                                                 MCMC sample size =     10,000
Group variable: id                               Number of groups =        100
                                                 Obs per group:
                                                              min =          5
                                                              avg =        5.0
                                                              max =          5

Family : Bernoulli                               Number of obs    =        500
Link   : probit                                  Acceptance rate  =      .3268
                                                 Efficiency:  min =     .05399
                                                              avg =       .102
Log marginal likelihood                                       max =      .1628

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |      Mean   Std. Dev.     MCSE     Median  [95% Cred. Interval]
-------------+----------------------------------------------------------------
y            |
          x1 |  .9977099   .1181726   .003773   .9936143   .7810441   1.242439
          x2 | -1.018063   .1892596    .00557  -1.012598  -1.396798  -.6509636
          x3 |  .9539304   .2936949   .007279   .9514395   .3823801    1.52913
       _cons |  .5433822   .2205077    .00949   .5398387   .1216346   .9847166
-------------+----------------------------------------------------------------
id           |
   U0:sigma2 |  1.456558   .4384163   .015537   1.401461   .7611919   2.463175
------------------------------------------------------------------------------
Note: Default priors are used for model parameters.
```

Example 2: `bayesgraph diagnostic`

- We can look at the diagnostic graph for a couple of variables:

. bayesgraph diagnostic {y:x1}

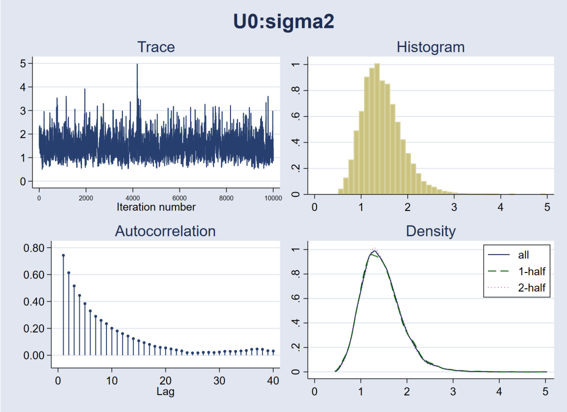


- The trace seems to indicate convergence this time.
- Autocorrelation decays quicker and becomes negligible after about 15 periods.

Example 2: `bayesgraph diagnostic`

- We now look now at the diagnostic graphs for {U0:sigma2}

. bayesgraph diagnostic U0:sigma2



- The trace seems to indicate convergence this time.
- Autocorrelation decays quicker and becomes negligible after about 15 periods.

Example 2: `bayestest interval`

- We can perform interval testing with the postestimation command `bayestest interval`.
- It estimates the probability that a model parameter lies in a particular interval.
- For continuous parameters the hypothesis is formulated in terms of intervals.
- We can perform point hypothesis testing only for parameters with discrete posterior distributions.
- `bayestest interval` estimates the posterior distribution for a null interval hypothesis.
- `bayestest interval` reports the estimated posterior mean probability for Ho.

```
bayestest interval ({y:x1},lower(.9) upper(1.02)) ///
                   ({y:x2},lower(-1.1) upper(-.8))
```

Example 2: bayestest interval

- We can, for example, perform separate tests for different parameters:

```
. bayestest interval ({y:x1},lower(.9) upper(1.02)) ///
>                    ({y:x2},lower(-1.1) upper(-.8))

Interval tests      MCMC sample size =    10,000

       prob1 :   .9 < {y:x1} < 1.02
       prob2 : -1.1 < {y:x2} < -.8
```

	Mean	Std. Dev.	MCSE
prob1	.3888	0.48750	.0077073
prob2	.5474	0.49777	.0097517

- We can also perform a joint test:

```
. bayestest interval (({y:x1},lower(.9) upper(1.02)) ///
>               ({y:x2},lower(-1.1) upper(-.8)),joint)

Interval tests      MCMC sample size =    10,000

   prob1 : .9 < {y:x1} < 1.02, -1.1 < {y:x2} < -.8
```

	Mean	Std. Dev.	MCSE
prob1	.2249	0.41754	.0066399

Example 2: bayestest interval

- We can, for example, perform separate tests for different parameters:

```
. bayestest interval ({y:x1},lower(.9) upper(1.02)) ///
>                    ({y:x2},lower(-1.1) upper(-.8))

Interval tests     MCMC sample size =    10,000

       prob1 :   .9 < {y:x1} < 1.02
       prob2 : -1.1 < {y:x2} < -.8

-----------------------------------------------------
             |      Mean    Std. Dev.          MCSE
-------------+---------------------------------------
       prob1 |     .3888     0.48750       .0077073
       prob2 |     .5474     0.49777       .0097517
-----------------------------------------------------
```

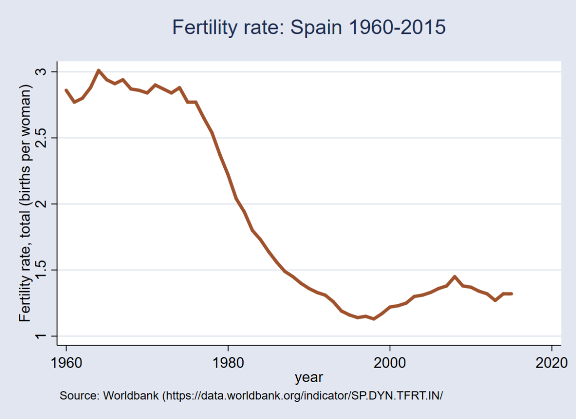
- We can also perform a joint test:

```
. bayestest interval (({y:x1},lower(.9) upper(1.02)) ///
>               ({y:x2},lower(-1.1) upper(-.8)),joint)

Interval tests     MCMC sample size =    10,000

   prob1 : .9 < {y:x1} < 1.02, -1.1 < {y:x2} < -.8

-----------------------------------------------------
             |      Mean    Std. Dev.          MCSE
-------------+---------------------------------------
       prob1 |     .2249     0.41754       .0066399
-----------------------------------------------------
```

Example 3: Change-point model

Example 3: Change-point model

- Let's work now with an example where we write our model using a substitutable expression.
- We have yearly data on fertility for Spain:



- The series has a significant change around 1980.
- We may consider fitting a change-point model.

Example 3: Gibbs sampling

Change point model specification with blocking

```
bayesmh fertil = ({mu1}*sign(year<{cp})           ///
                + {mu2}*sign(year>={cp})),        ///
      likelihood(normal({var}))                   ///
      prior({mu1}, normal(1,5))                   ///
      prior({mu2}, normal(5,5))                   ///
      prior({cp}, uniform(1960,2015))             ///
      prior({var}, igamma(2,1))                   ///
      initial({mu1} 5 {mu2} 1 {cp} 1960)          ///
      block(var, gibbs) block(cp) blocksummary    ///
      rseed(123) mcmcsize(40000)                  ///
      dots(500,every(5000))                       ///
      title(Change-point analysis)
```

Example 3: Gibbs sampling

Change point model specification with blocking

```
bayesmh fertil = ({mu1}*sign(year<{cp})                    ///
                 + {mu2}*sign(year>={cp})),                ///
      likelihood(normal({var}))                            ///
      prior({mu1}, normal(1,5))                            ///
      prior({mu2}, normal(5,5))                            ///
      prior({cp}, uniform(1960,2015))                      ///
      prior({var}, igamma(2,1))                            ///
      initial({mu1} 5 {mu2} 1 {cp} 1960)                   ///
      block(var, gibbs) block(cp) blocksummary             ///
      rseed(123) mcmcsize(40000)                           ///
      dots(500,every(5000))                                ///
      title(Change-point analysis)
```

Example 3: Gibbs sampling

Change point model specification with blocking

```
. bayesmh fertil=({mu1}*sign(year<{cp})+{mu2}*sign(year>={cp})), ///
>         likelihood(normal({var}))                               ///
>         prior({mu1}, normal(0,5))                               ///
>         prior({mu2}, normal(5,5))                               ///
>         prior({cp}, uniform(1960,2015))                         ///
>         prior({var}, igamma(2,1))                               ///
>         initial({mu1} 5 {mu2} 1 {cp} 1960)                      ///
>         block(var, gibbs) block(cp) blocksummary                ///
>         rseed(123) mcmcsize(40000) dots(500, every(5000))       ///
>         title(Modelo de Cambio de Punto)

Burn-in 2500 aaaaa done
Simulation 40000 .........5000.........10000.........15000.........20000
> .........25000.........30000.........35000.........40000 done

Model summary
------------------------------------------------------------------------
Likelihood:
  fertility ~ normal({mu1}*sign(year<{cp})+{mu2}*sign(year>={cp}),{var})

Priors:
  {var} ~ igamma(2,1)
  {mu1} ~ normal(0,5)
  {mu2} ~ normal(5,5)
   {cp} ~ uniform(1960,2015)
------------------------------------------------------------------------

Block summary
------------------------------------------------------------------------
   1:  {var}                                                    (Gibbs)
   2:  {cp}
   3:  {mu1} {mu2}
------------------------------------------------------------------------
```

Example 3: Gibbs sampling

Change point model specification with blocking

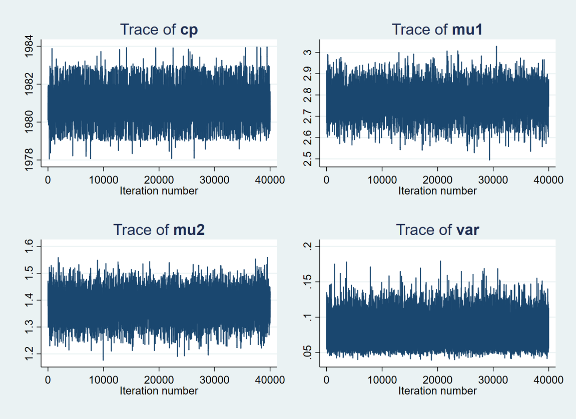
```
. bayesmh fertil=({mu1}*sign(year<{cp})+{mu2}*sign(year>={cp})), ///
>         likelihood(normal({var}))                                ///
>         prior({mu1}, normal(0,5))                                ///
>         prior({mu2}, normal(5,5))                                ///
>         prior({cp}, uniform(1960,2015))                          ///
>         prior({var}, igamma(2,1))                                ///
>         initial({mu1} 5 {mu2} 1 {cp} 1960)                      ///
>         block(var, gibbs) block(cp) blocksummary                ///
>         rseed(123) mcmcsize(40000) dots(500, every(5000))       ///
>         title(Modelo de Cambio de Punto)
Modelo de Cambio de Punto                          MCMC iterations  =     42,500
Metropolis-Hastings and Gibbs sampling             Burn-in          =      2,500
                                                   MCMC sample size =     40,000
                                                   Number of obs    =         56
                                                   Acceptance rate  =      .5704
                                                   Efficiency:  min =     .08572
                                                                avg =      .2629
Log marginal likelihood = -16.240692                            max =      .7203
```

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred.	Interval]
cp	1980.87	.7407595	.010454	1980.772	1979.439	1982.517
mu1	2.771024	.0654542	.001118	2.770196	2.64247	2.897339
mu2	1.376056	.0489823	.000706	1.375648	1.281815	1.472107
var	.078699	.0152773	.00009	.0768054	.0541305	.1136579

Example 3: `bayesgraph trace`

- Use bayesgraph trace to look at the trace for all the parameters.

. bayesgraph trace

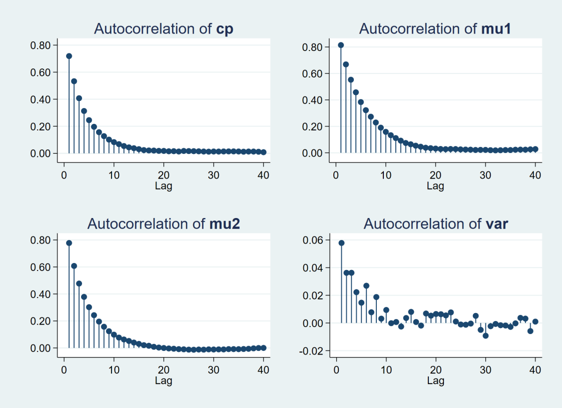


- The plots indicate that convergence seems to be achieved.

Example 3: `bayesgraph ac`

- Use bayesgraph ac to look at the autocorrelation for all the parameters.

. bayesgraph ac



- Autocorrelation decays and becomes negligible quickly for almost all the parameters.

Summing up

- Bayesian analysis: An 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.