# Bayesian analysis using Stata

Yulia Marchenko

Director of Biostatistics
StataCorp LP

2015 Nordic and Baltic Stata Users Group meeting

## 7 More examples (extra)

- Normal linear regression
- Random-intercept model
- Random-coefficient model
- Meta analysis
- Nonlinear Poisson model: Change-point analysis
- Bioequivalence in a crossover trial

Bayesian analysis is a statistical paradigm that answers research questions about unknown parameters using probability statements.

- What is the probability that a person accused of a crime is guilty?
- What is the probability that treatment A is more cost effective than treatment B for a specific health care provider?
- What is the probability that the odds ratio is between 0.3 and 0.5?
- What is the probability that three out of five quiz questions will be answered correctly by students?

You may be interested in Bayesian analysis if

- you have some prior information available from previous studies that you would like to incorporate in your analysis. For example, in a study of preterm birthweights, it would be sensible to incorporate the prior information that the probability of a mean birthweight above 15 pounds is negligible. Or,
- your research problem may require you to answer a question: What is the probability that my parameter of interest belongs to a specific range? For example, what is the probability that an odds ratio is between 0.2 and 0.5? Or,
- you want to assign a probability to your research hypothesis. For example, what is the probability that a person accused of a crime is guilty?
- And more.

- Observed data sample $D$ is fixed and model parameters $\boldsymbol{\theta}$ are random.
- $D$ is viewed as a result of a one-time experiment.
- A parameter is summarized by an entire distribution of values instead of one fixed value as in classical frequentist analysis.

- There is some prior (before seeing the data!) knowledge about $\boldsymbol{\theta}$ formulated as a **prior distribution** $p(\boldsymbol{\theta})$.
- After data $D$ are observed, the information about $\boldsymbol{\theta}$ is updated based on the **likelihood** $f(D|\boldsymbol{\theta})$.
- Information is updated by using the Bayes rule to form a **posterior distribution** $p(\boldsymbol{\theta}|D)$:

$$p(\boldsymbol{\theta}|D) = \frac{f(D|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(D)}$$

where $p(D)$ is the **marginal distribution** of the data $D$.

- Estimating a posterior distribution $p(\boldsymbol{\theta}|D)$ is at the heart of Bayesian analysis.
- Various summaries of this distribution are used for inference.
- Point estimates: posterior means, modes, medians, percentiles.
- Interval estimates: **credible intervals** (CrI)—(fixed) ranges to which a parameter is known to belong with a pre-specified probability.
- Monte-Carlo standard error (MCSE)—represents precision about posterior mean estimates.
- Hypothesis testing—assign probability to any hypothesis of interest
- Model comparison: model posterior probabilities, Bayes factors

- Potential subjectivity in specifying prior information—noninformative priors or sensitivity analysis to various choices of informative priors.
- Computationally demanding—involves intractable integrals that can only be computed using intensive numerical methods such as Markov chain Monte Carlo (MCMC).

**Research problem**

- Prevalence of a rare infectious disease in a small city (Hoff 2009)
- A sample of 20 subjects is checked for infection
- Parameter $\theta$ is the proportion of infected individuals in the city
- Outcome $y$ is the # of infected individuals in the sample

## Model

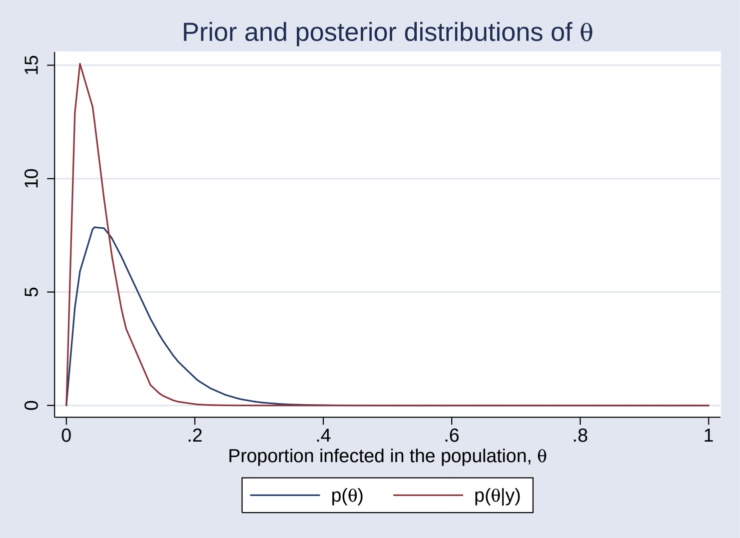
- Likelihood, $f(y|\theta)$: Binomial
- Prior, $p(\theta)$: Infection rate ranged between 0.05 and 0.20, with an average prevalence of 0.10, in other similar cities
- Bayesian model:

$$
\begin{aligned}
y|\theta &\sim \mathrm{Binomial}(20, \theta) \\
\theta &\sim \mathrm{Beta}(2, 20)
\end{aligned}
$$

- Posterior: $\theta|y \sim \mathrm{Beta}(2 + y, 20 + 20 - y)$

## Observed data

- We sample individuals and observe none who have an infection, $y = 0$
- Posterior: $\theta|y \sim \mathrm{Beta}(2, 40)$
- Prior mean: $E(\theta) = 2/(2{+}20) = 0.09$
- Posterior mean: $E(\theta|y) = 2/(2{+}40) = 0.048$
- Posterior probability: $P(\theta < 0.10) = 0.93$


Prior and posterior distributions of θ
15
10
5
0
0
.2
.4
.6
.8
1
Proportion infected in the population, θ
p(θ)
p(θ|y)

- Fit beta-binomial model using `bayesmh` (variable `y` has one observation equal to 0)
- MCMC method: adaptive Metropolis-Hastings (MH)

```
. set seed 14

. bayesmh y, likelihood(binlogit(20), noglmtransform) ///
>            prior({y:_cons}, beta(2,20))

Model summary
------------------------------------------------------------------------------
Likelihood:
  y ~ binomial({y:_cons},20)

Prior:
  {y:_cons} ~ beta(2,20)
------------------------------------------------------------------------------

Bayesian binomial regression                     MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling         Burn-in          =      2,500
                                                 MCMC sample size =     10,000
                                                 Number of obs    =          1
                                                 Acceptance rate  =      .4205
Log marginal likelihood = -1.1714402             Efficiency       =      .1401

------------------------------------------------------------------------------
             |                                                Equal-tailed
           y |      Mean   Std. Dev.     MCSE     Median  [95% Cred. Interval]
-------------+----------------------------------------------------------------
       _cons |  .0466517   .0316076   .000844   .0391639   .0058112   .1260038
------------------------------------------------------------------------------
```

- Compute posterior probability

```
. bayestest interval {y:_cons}, upper(0.1)
Interval tests      MCMC sample size =    10,000
       prob1 : {y:_cons} < 0.1
```

| | Mean | Std. Dev. | MCSE |
|---|---|---|---|
| prob1 | .9299 | 0.25533 | .006074 |

| *Command* | *Description* |
|---|---|
| **Estimation** | |
| `bayesmh` | Bayesian regression using MH |
| `bayesmh` *evaluators* | User-written Bayesian models using MH |
| | |
| **Postestimation** | |
| `bayesgraph` | Graphical convergence diagnostics |
| | |
| `bayesstats ess` | Effective sample sizes and more |
| `bayesstats summary` | Summary statistics |
| `bayesstats ic` | Information criteria and Bayes factors |
| | |
| `bayestest model` | Model posterior probabilities |
| `bayestest interval` | Interval hypothesis testing |

## Models

- 10 built-in likelihoods: normal, logit, ologit, Poisson, ...
- 18 built-in priors: normal, gamma, Wishart, Zellner's $g$, ...
- Continuous, binary, ordinal, and count outcomes
- Univariate, multivariate, and multiple-equation models
- Linear, nonlinear, and canonical generalized nonlinear models
- Continuous univariate, multivariate, and discrete priors
- User-defined models

## MCMC methods

- Adaptive MH
- Adaptive MH with Gibbs updates—hybrid
- Full Gibbs sampling for some models

**Built-in models**

```
bayesmh ..., likelihood() prior() ...
```

**User-defined models**

```
bayesmh ..., {evaluator() | llevaluator() prior()} ...
```

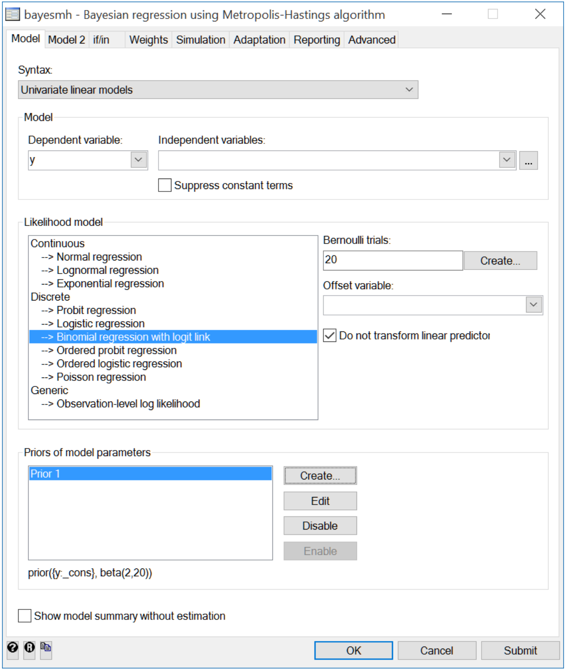
Postestimation features are the same whether you use a built-in model or program your own!

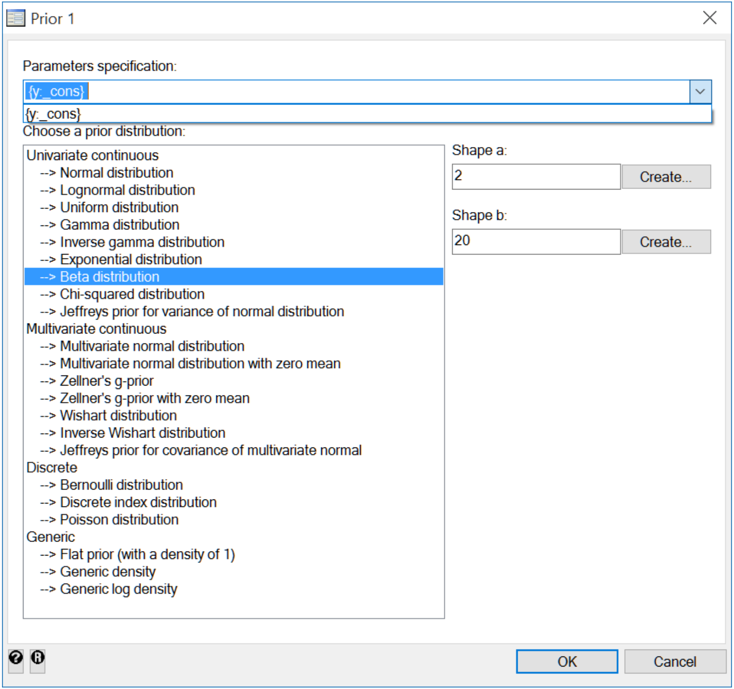
- Perform Bayesian analysis by using the command line
- Or, use a powerful point-and-click interface
- You can access the GUI by typing

```
. db bayesmh
```

or from the Statistics menu

(NEXT SLIDE)


bayesmh - Bayesian regression using Metropolis-Hastings algorithm
Model
Model 2
if/in
Weights
Simulation
Adaptation
Reporting
Advanced
Syntax:
Univariate linear models
Model
Dependent variable:
y
Independent variables:
Suppress constant terms
Likelihood model
Continuous
--> Normal regression
--> Lognormal regression
--> Exponential regression
Discrete
--> Probit regression
--> Logistic regression
--> Binomial regression with logit link
--> Ordered probit regression
--> Ordered logistic regression
--> Poisson regression
Generic
--> Observation-level log likelihood
Bernoulli trials:
20
Create...
Offset variable:
Do not transform linear predictor
Priors of model parameters
Prior 1
Create...
Edit
Disable
Enable
prior({y:_cons}, beta(2,20))
Show model summary without estimation
OK
Cancel
Submit

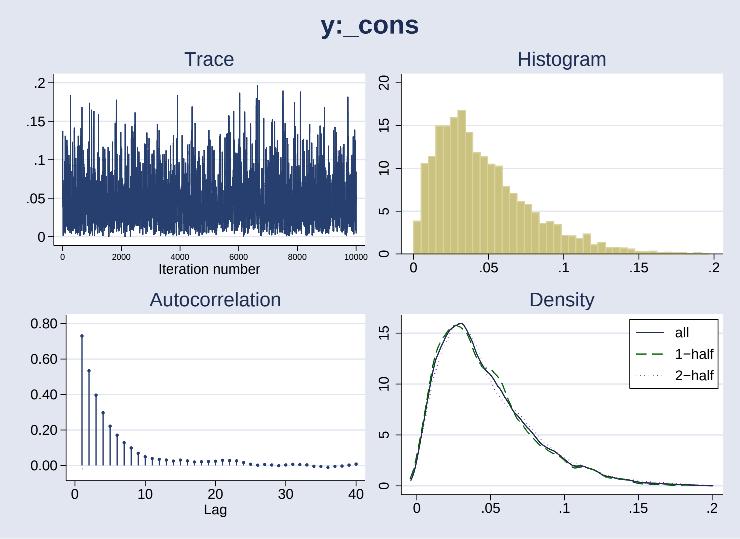

Prior 1
Parameters specification:
{y:_cons}
{y:_cons}
Choose a prior distribution:
Univariate continuous
--> Normal distribution
--> Lognormal distribution
--> Uniform distribution
--> Gamma distribution
--> Inverse gamma distribution
--> Exponential distribution
--> Beta distribution
--> Chi-squared distribution
--> Jeffreys prior for variance of normal distribution
Multivariate continuous
--> Multivariate normal distribution
--> Multivariate normal distribution with zero mean
--> Zellner's g-prior
--> Zellner's g-prior with zero mean
--> Wishart distribution
--> Inverse Wishart distribution
--> Jeffreys prior for covariance of multivariate normal
Discrete
--> Bernoulli distribution
--> Discrete index distribution
--> Poisson distribution
Generic
--> Flat prior (with a density of 1)
--> Generic density
--> Generic log density
Shape a:
2
Create...
Shape b:
20
Create...
OK
Cancel

- Recall the beta-binomial model from the motivating example.
- Let's store the estimation results for future comparison.
- `estimates store` requires first saving `bayesmh`'s MCMC data.
- Use option `saving()` during estimation or on replay:

```
. bayesmh, saving(betabin)
note: file betabin.dta saved
. estimates store betabin
```

- Check MCMC convergence

```
. bayesgraph diagnostics {y:_cons}
```



- Check MCMC sampling efficiency

```
. bayesstats ess {y:_cons}
Efficiency summaries    MCMC sample size =    10,000
```

| y | ESS | Corr. time | Efficiency |
|---|---|---|---|
| _cons | 1400.87 | 7.14 | 0.1401 |

- Test an interval hypothesis

```
. bayestest interval {y:_cons}, upper(0.1)
Interval tests      MCMC sample size =    10,000
       prob1 : {y:_cons} < 0.1
```

| | Mean | Std. Dev. | MCSE |
|---|---|---|---|
| prob1 | .9299 | 0.25533 | .006074 |

- Motivating example used a beta prior for $\theta$
- Sensitivity analysis to the choice of the priors is very important in Bayesian analysis
- Consider an alternative prior—a power prior

- Based on similar historical data $y_0$
- Idea: raise the likelihood function of the historical data to the power $\alpha_0$, where $0 \leq \alpha_0 \leq 1$.
- $\alpha_0$ quantifies the uncertainty in $y_0$ by controlling the heaviness of the tails of the prior distribution.
- $\alpha_0 = 0$ means no information from the historical data and $\alpha_0 = 1$ means that the historical data have as much weight as the observed data.

- Suppose that in another similar city, a random sample of 15 subjects was selected and 1 subject had a disease.
- Let's consider a competing power prior:

$$p(\theta) \sim \{\mathrm{BinomialPMF}(15, 1, \theta)\}^{\alpha_0}$$

- Let $\alpha_0 = 0.5$.

- bayesmh does not have built-in power priors but we can use prior()'s suboption density() to specify our own prior.

```
. set seed 14

. bayesmh y, likelihood(binlogit(20), noglmtransform)                    ///
>         prior({y:_cons}, density(sqrt(binomialp(15,1,{y:_cons})))) ///
>         saving(powerbin)

Model summary
------------------------------------------------------------------------------
Likelihood:
  y ~ binomial({y:_cons},20)

Prior:
  {y:_cons} ~ density(sqrt(binomialp(15,1,{y:_cons})))
------------------------------------------------------------------------------
```

```
Bayesian binomial regression                    MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling        Burn-in          =      2,500
                                                MCMC sample size =     10,000
                                                Number of obs    =          1
                                                Acceptance rate  =      .4294
Log marginal likelihood = -3.4630512            Efficiency       =      .1228
```

| y | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| _cons | .0501507 | .0392846 | .001121 | .0401686 | .0040134 | .1521774 |

```
file powerbin.dta not found; file saved
. estimates store powerbin
```

- Compute model posterior probabilities

```
. bayestest model powerbin betabin
Bayesian model tests
```

| | log(ML) | P(M) | P(M\|y) |
|---|---|---|---|
| powerbin | -3.4631 | 0.5000 | 0.0918 |
| betabin | -1.1714 | 0.5000 | 0.9082 |

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

- Compute the Bayes factor—the ratio of the marginal likelihoods of the two models calculated using the same data.

```
. bayesstats ic powerbin betabin
Bayesian information criteria
```

| | DIC | log(ML) | log(BF) |
|---|---|---|---|
| powerbin | 2.129576 | -3.463051 | . |
| betabin | 1.956201 | -1.17144 | 2.291611 |

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

- In addition to the many built-in models, you can also program your own models.
- Program log likelihood and use one of the built-in priors:

  `. bayesmh ..., llevaluator(`*llprogname*`) prior() ...`

- Or, program the log posterior:

  `. bayesmh ..., evaluator(`*lpprogname*`) ...`

- One of the questions we received shortly after releasing `bayesmh` is "How do I fit Bayesian hurdle models?"
- A hurdle model (Cragg model) is used to model a bounded dependent variable. It combines a selection model that determines the boundary points of the dependent variable with an outcome model that determines its nonbounded values.
- Hurdle models are not currently among the built-in `bayesmh` models.
- But, we can program them using `bayesmh`'s used-defined evaluators.
- Below I provide two types of log-likelihood evaluators: one using Stata's command `churdle` (new in Stata 14) to compute the log likelihood and the other programming the log likelihood from scratch.

- We consider a subset of the `fitness` data from **[R] churdle**.
- We consider a simple linear hurdle model.
- We model the decision to exercise or not as a function of an individual's average commute to work.
- Once a decision to exercise is made, we model the number of hours an individual exercises per day as a function of age.

```
. webuse fitness
. set seed 17653
. sample 10
(17,848 observations deleted)
```

- We use `churdle` to compute the log-likelihood values at each MCMC iteration.

```
. program mychurdle1
          version 14.0
          args llf
          tempname b
          mat `b´ = ($MH_b, $MH_p)
          cap churdle linear $MH_y1 $MH_y1x1 if $MH_touse, ///
>                      select($MH_y2x1) ll(0) from(`b´) iterate(0)
          if _rc {
                  if (_rc==1) { // handle break key
                          exit _rc
                  }
                  scalar `llf´ = .
          }
          else {
                  scalar `llf´ = e(ll)
          }
    end
```

```
. set seed 14
. gen byte hours0 = (hours==0)
. bayesmh (hours age) (hours0 commute),         ///
>   llevaluator(mychurdle1, parameters({lnsig})) ///
>   prior({hours:} {hours0:} {lnsig}, flat) dots

Burn-in 2500 aaaaaaaaa1000aaaaaaaaa2000aaaa. done
Simulation 10000 .........1000.........2000.........3000.........4000.........5
> 000.........6000.........7000.........8000.........9000.........10000 done
Model summary
------------------------------------------------------------------------------
Likelihood:
  hours hours0 ~ mychurdle1(xb_hours,xb_hours0,{lnsig})
Priors:
       {hours:age _cons} ~ 1 (flat)                                        (1)
  {hours0:commute _cons} ~ 1 (flat)                                        (2)
                 {lnsig} ~ 1 (flat)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_hours.
(2) Parameters are elements of the linear form xb_hours0.
```

```
Bayesian regression                                MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling           Burn-in          =      2,500
                                                   MCMC sample size =     10,000
                                                   Number of obs    =      1,983
                                                   Acceptance rate  =      .2752
                                                   Efficiency:  min =     .04197
                                                                avg =     .06659
Log marginal likelihood = -2772.4136                            max =     .08861
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| hours | | | | | | |
| age | .0051872 | .0027702 | .000093 | .0052248 | -.0002073 | .0104675 |
| _cons | 1.163384 | .1219417 | .005135 | 1.16685 | .9203519 | 1.388663 |
| hours0 | | | | | | |
| commute | -.0716184 | .1496757 | .005623 | -.0758964 | -.3733355 | .2181717 |
| _cons | .1454332 | .084041 | .003066 | .1451574 | -.0222543 | .3128047 |
| lnsig | .1341657 | .034162 | .001668 | .1336526 | .0634106 | .2021694 |

- This model took 25 minutes

- The corresponding log likelihood programmed from scratch

```
. program mychurdle2
          version 14.0
          args lnf xb xg lnsig
          tempname sig
          scalar `sig´ = exp(`lnsig´)
          tempvar lnfj
          qui gen double `lnfj´ = normal(`xg´)  if $MH_touse
          qui replace `lnfj´    = log(1 - `lnfj´) if $MH_y1 <= 0 & $MH_touse
          qui replace `lnfj´    = log(`lnfj´) - log(normal(`xb´/`sig´)) ///
>                              + log(normalden($MH_y1,`xb´,`sig´))       ///
>                                if $MH_y1 > 0 & $MH_touse
          summarize `lnfj´ if $MH_touse, meanonly
          if r(N) < $MH_n {
              scalar `lnf´ = .
              exit
          }
          scalar `lnf´ = r(sum)
     end
```

```
. set seed 14
. bayesmh (hours age) (hours0 commute),                         ///
>         llevaluator(mychurdle2, parameters({lnsig}) )   ///
>         prior({hours:} {hours0:} {lnsig}, flat) dots

Burn-in 2500 aaaaaaaaa1000aaaaaaaaa2000aaaa. done
Simulation 10000 .........1000.........2000.........3000.........4000.........5
> 000.........6000.........7000.........8000.........9000.........10000 done
Model summary
------------------------------------------------------------------------------
Likelihood:
  hours hours0 ~ mychurdle2(xb_hours,xb_hours0,{lnsig})
Priors:
       {hours:age _cons} ~ 1 (flat)                                       (1)
  {hours0:commute _cons} ~ 1 (flat)                                       (2)
                 {lnsig} ~ 1 (flat)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_hours.
(2) Parameters are elements of the linear form xb_hours0.
```

```
Bayesian regression                               MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling          Burn-in          =      2,500
                                                  MCMC sample size =     10,000
                                                  Number of obs    =      1,983
                                                  Acceptance rate  =      .2752
                                                  Efficiency:  min =     .04197
                                                               avg =     .06659
Log marginal likelihood = -2772.4136                           max =     .08861
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| **hours** | | | | | | |
| age | .0051872 | .0027702 | .000093 | .0052248 | -.0002073 | .0104675 |
| _cons | 1.163384 | .1219417 | .005135 | 1.16685 | .9203519 | 1.388663 |
| **hours0** | | | | | | |
| commute | -.0716184 | .1496757 | .005623 | -.0758964 | -.3733355 | .2181717 |
| _cons | .1454332 | .084041 | .003066 | .1451574 | -.0222543 | .3128047 |
| lnsig | .1341657 | .034162 | .001668 | .1336526 | .0634106 | .2021694 |

- This model took only 20 seconds!

- Bayesian analysis is a powerful tool that allows you to incorporate prior information about model parameters into your analysis.
- It provides intuitive and direct interpretations of results by using probability statements about parameters.
- It provides a way to assign an actual probability to any hypothesis of interest.

- Use `bayesmh` for estimation: choose one of the built-in models or program your own.
- Use postestimation features for checking MCMC convergence, estimating functions of model parameters, and performing hypothesis testing and model comparison.
- Work interactively using the command line or use the point-and-click interface.
- Check out the "More examples" section and the **[BAYES] Bayesian analysis** manual for more examples and details about Bayesian analysis.

- More computationally efficient handling of multilevel ("random-effects") models—option `reffects()` for two-level models and option `block(, reffects)` for models with more than two levels.
- For example, Bayesian IRT 1PL models with more than 32,000 subjects are now feasible:

```
. bayesmh y i.item, noconstant reffects(id) likelihood(logit) ///
>         prior({y:i.id},   normal(0, {var}))                  ///
>         prior({y:i.item}, normal(0, 10))                     ///
>         prior({var},      igamma(0.01,0.01))                 ///
>         block({y:i.item}, reffects)                          ///
>         block({var})
```

- Straightforward specification of unstructured covariances between random-effects parameters—prior distribution `mvnormal()` is now row-column conformable.
- For example,

```
. bayesmh ..., ... prior({y:i.id i.id#c.x}, mvnormal(2,{b0},{b1},{Sigma,matrix}))
```

models the unstructured covariance between random intercepts and random coefficients for `x` associated with the levels of `id`.

Carlin, B. P., A. E. Gelfand, and A. F. M. Smith. 1992. Hierarchical Bayesian analysis of changepoint problems. *Journal of the Royal Statistical Society, Series C* 41: 389–405.

Diggle, P. J., P. J. Heagerty, K.-Y. Liang, and S. L. Zeger. 2002. *Analysis of Longitudinal Data*. 2nd ed. Oxford: Oxford University Press.

Gelfand, A. E., S. E. Hills, A. Racine-Poon, and A. F. M. Smith. 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association* 85: 972–985.

Hoff, P. D. 2009. *A First Course in Bayesian Statistical Methods*. New York: Springer.

Turner, R. M., R. Z. Omar, M. Yang, H. Goldstein, and S. G. Thompson. 2000. A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine* 19: 3417–3432.

- Data: `weight` measurements of 48 pigs on 9 successive `weeks` (e.g., Diggle et al. (2002)).
- Ignore the grouping structure of the data for now
- Likelihood model:

$$\begin{aligned} \text{weight}_{ij} &= \beta_0 + \beta_1 \texttt{week}_{ij} + \epsilon_{ij} \\ \epsilon_{ij} &\sim \text{Normal}(0,\sigma^2) \end{aligned}$$

  where $i = 1, \ldots, 9$ and $j = 1, \ldots, 48$.
- Noninformative prior distributions:

$$\begin{aligned} \beta_i &\sim \text{Normal}(0,100),\ i = 0, 1 \\ \sigma^2 &\sim \text{InvGamma}(0.001, 0.001) \end{aligned}$$

```
. webuse pig
(Longitudinal analysis of pig weights)

. set seed 14

. bayesmh weight week, likelihood(normal({var}))         ///
>                      prior({weight:}, normal(0,100))   ///
>                      prior({var},     igamma(0.001,0.001))

Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
Likelihood:
  weight ~ normal(xb_weight,{var})

Priors:
  {weight:week _cons} ~ normal(0,100)                                      (1)
                {var} ~ igamma(0.001,0.001)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_weight.
```

```
Bayesian normal regression                       MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling         Burn-in          =      2,500
                                                 MCMC sample size =     10,000
                                                 Number of obs    =        432
                                                 Acceptance rate  =      .2291
                                                 Efficiency:  min =      .0692
                                                              avg =     .08122
Log marginal likelihood = -1270.6848                          max =     .09538
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| weight | | | | | | |
| week | 6.214291 | .0787262 | .002549 | 6.214297 | 6.055505 | 6.364085 |
| _cons | 19.32917 | .4468276 | .015889 | 19.31526 | 18.47262 | 20.22465 |
| var | 19.50327 | 1.33882 | .050894 | 19.44994 | 17.09487 | 22.30596 |

```
. set seed 14

. bayesmh weight week, likelihood(normal({var}))          ///
>                      prior({weight:}, normal(0,100))    ///
>                      prior({var},     igamma(0.001,0.001)) ///
>                      block({weight:}, gibbs)            ///
>                      block({var},     gibbs) nomodelsummary

Burn-in ...
Simulation ...

Bayesian normal regression                       MCMC iterations  =     12,500
Gibbs sampling                                   Burn-in          =      2,500
                                                 MCMC sample size =     10,000
                                                 Number of obs    =        432
                                                 Acceptance rate  =          1
                                                 Efficiency:  min =          1
                                                              avg =          1
Log marginal likelihood = -1270.6434                          max =          1
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| weight | | | | | | |
| week | 6.216249 | .0816994 | .000817 | 6.216445 | 6.053813 | 6.377687 |
| _cons | 19.31436 | .4619975 | .004539 | 19.31138 | 18.41486 | 20.22794 |
| var | 19.3699 | 1.329478 | .013295 | 19.31951 | 16.93417 | 22.17573 |

- Measurements within a pig are correlated—introduce a random intercept
- Likelihood model:

$$\begin{aligned} \text{weight}_{ij} &= \beta_0 + u_{0j} + \beta_1 \texttt{week}_{ij} + \epsilon_{ij} \\ \epsilon_{ij} &\sim \text{Normal}(0,\sigma^2) \\ u_{0j} &\sim \text{Normal}(0,\sigma_0^2) \end{aligned}$$

where $i = 1, \ldots, 9$ and $j = 1, \ldots, 48$.

- Prior distributions:

$$\begin{aligned} \beta_i &\sim \text{Normal}(0,100),\ i = 0, 1 \\ \sigma^2 &\sim \text{InvGamma}(0.001, 0.001) \\ \sigma_0^2 &\sim \text{InvGamma}(0.001, 0.001) \end{aligned}$$

# Alternative model formulation

- Let $\tau_{0j} = \beta_0 + u_{0j}$
- Alternative likelihood model formulation:

$$
\begin{aligned}
\text{weight}_{ij} &= \tau_{0j} + \beta_1 \texttt{week}_{ij} + \epsilon_{ij} \\
\epsilon_{ij} &\sim \text{Normal}(0,\sigma^2) \\
\tau_{0j} &\sim \text{Normal}(\beta_0,\sigma_0^2)
\end{aligned}
$$

- Default MH sampling

```
. webuse pig
(Longitudinal analysis of pig weights)

. fvset base none id

. set seed 14

. bayesmh weight week i.id, likelihood(normal({var})) noconstant ///
>   prior({weight:i.id},  normal({weight:cons},{var_0}))        ///
>   prior({weight:week},  normal(0,100))                        ///
>   prior({weight:cons},  normal(0,100))                        ///
>   prior({var},          igamma(0.001,0.001))                  ///
>   prior({var_0},        igamma(0.001,0.001))                  ///
>   noshow({weight:i.id})
```

- Model summary

```
Burn-in ...
Simulation ...
Model summary
------------------------------------------------------------------------------
Likelihood:
  weight ~ normal(xb_weight,{var})
Priors:
  {weight:week} ~ normal(0,100)                                           (1)
  {weight:i.id} ~ normal({weight:cons},{var_0})                           (1)
          {var} ~ igamma(0.001,0.001)
Hyperpriors:
  {weight:cons} ~ normal(0,100)
        {var_0} ~ igamma(0.001,0.001)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_weight.
```

```
Bayesian normal regression                     MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling       Burn-in          =      2,500
                                               MCMC sample size =     10,000
                                               Number of obs    =        432
                                               Acceptance rate  =      .2341
                                               Efficiency:  min =    .001963
                                                            avg =    .005539
Log marginal likelihood = -1338.2346                        max =     .01159
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| weight | | | | | | |
| week | 6.257469 | .0273198 | .002538 | 6.256309 | 6.205179 | 6.309333 |
| var | 8.895206 | .6146577 | .138715 | 8.844657 | 7.799991 | 10.25156 |
| weight | | | | | | |
| cons | 13.75363 | .4025422 | .060251 | 13.75297 | 13.01862 | 14.56459 |
| var_0 | 12.36591 | .35361 | .054957 | 12.36093 | 11.66033 | 13.05275 |

Note: There is a high autocorrelation after 500 lags.

- Default MH sampling is very inefficient in this example
- Improve MCMC efficiency by blocking of parameters
- Use `block()`'s suboption `split` to block random-effects parameters—very important with many random effects

```
. set seed 14
. bayesmh weight week i.id, likelihood(normal({var})) noconstant ///
>   prior({weight:i.id}, normal({weight:cons},{var_0}))          ///
>   prior({weight:week}, normal(0,100))                          ///
>   prior({weight:cons}, normal(0,100))                          ///
>   prior({var},         igamma(0.001,0.001))                    ///
>   prior({var_0},       igamma(0.001,0.001))                    ///
>   block({var}) block({var_0})                                  ///
>   block({weight:week}) block({weight:cons})                    ///
>   block({weight:i.id}, split)                                  ///
>   nomodelsummary notable
```

- Blocking improved MCMC efficiency

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

Bayesian normal regression                         MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling           Burn-in          =      2,500
                                                   MCMC sample size =     10,000
                                                   Number of obs    =        432
                                                   Acceptance rate  =      .4447
                                                   Efficiency:  min =     .02386
                                                                avg =      .1491
Log marginal likelihood = -1052.2375                            max =      .1953
```

- Estimates are more similar to the frequentist results (see **[ME] mixed**)

```
. bayesstats summary {weight:week cons} {var_0} {var}
Posterior summary statistics                      MCMC sample size =     10,000
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| weight | | | | | | |
| week | 6.203559 | .0382251 | .002475 | 6.20247 | 6.132607 | 6.279994 |
| cons | 19.353 | .6176088 | .019352 | 19.3461 | 18.15131 | 20.57819 |
| var_0 | 15.88671 | 3.595539 | .094179 | 15.32318 | 10.62316 | 24.33477 |
| var | 4.427113 | .3264523 | .007969 | 4.404244 | 3.835123 | 5.102618 |

- Including random effects as a factor variable is not feasible with tens of thousands of random effects.
- Option `split` is very time consuming.
- Forthcoming option `reffects()` is an alternative.
- Replace `i.id` in the model formulation with option `reffects(id)` and remove `block(weight:i.id, split)`

```
. set seed 14
. bayesmh weight week, likelihood(normal({var})) noconstant reffects(id) ///
>   prior({weight:i.id}, normal({weight:cons},{var_0}))         ///
>   prior({weight:week}, normal(0,100))                         ///
>   prior({weight:cons}, normal(0,100))                         ///
>   prior({var},         igamma(0.001,0.001))                   ///
>   prior({var_0},       igamma(0.001,0.001))                   ///
>   block({var}) block({var_0})                                 ///
>   block({weight:week}) block({weight:cons})                   ///
>   nomodelsummary notable
```

- MCMC sampling efficiencies are slightly smaller

```
Bayesian normal regression                   MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling     Burn-in          =      2,500
                                             MCMC sample size =     10,000
                                             Number of obs    =        432
                                             Acceptance rate  =      .3788
                                             Efficiency:  min =     .01923
                                                          avg =      .0944
Log marginal likelihood = -1077.2283                      max =      .1566
```

- Estimates are similar to previous estimates

```
. bayesstats summary {weight:week cons} {var_0} {var}
Posterior summary statistics                      MCMC sample size =    10,000
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| weight | | | | | | |
| week | 6.215106 | .0378704 | .002731 | 6.214882 | 6.139693 | 6.290642 |
| cons | 19.25063 | .6306763 | .02307 | 19.24458 | 18.00894 | 20.48578 |
| var_0 | 16.00539 | 3.739944 | .104932 | 15.44782 | 10.336 | 24.8429 |
| var | 4.432357 | .3225202 | .00815 | 4.416106 | 3.836758 | 5.100198 |

- We can use Gibbs sampling for some of the parameters to further improve MCMC sampling
- Average MCMC sampling efficiency increased from 9% to 30%

```
. set seed 14

. bayesmh weight week, likelihood(normal({var})) noconstant reffects(id) ///
>   prior({weight:i.id}, normal({weight:cons},{var_0}))          ///
>   prior({weight:week}, normal(0,100))                          ///
>   prior({weight:cons}, normal(0,100))                          ///
>   prior({var},         igamma(0.001,0.001))                    ///
>   prior({var_0},       igamma(0.001,0.001))                    ///
>   block({var}, gibbs) block({var_0}, gibbs)                    ///
>   block({weight:week}, gibbs) block({weight:cons}, gibbs)      ///
>   nomodelsummary notable

Burn-in ...
Simulation ...

Bayesian normal regression                       MCMC iterations  =     12,500
Metropolis-Hastings and Gibbs sampling           Burn-in          =      2,500
                                                 MCMC sample size =     10,000
                                                 Number of obs    =        432
                                                 Acceptance rate  =      .8235
                                                 Efficiency:  min =     .02439
                                                              avg =      .2851
Log marginal likelihood = -1077.0036                          max =      .6009
```

```
. bayesstats summary {weight:week cons} {var_0} {var}
Posterior summary statistics                      MCMC sample size =     10,000
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| weight | | | | | | |
| week | 6.216461 | .0383844 | .002458 | 6.217039 | 6.139121 | 6.291271 |
| cons | 19.24988 | .6046734 | .015102 | 19.24786 | 18.06586 | 20.46588 |
| var_0 | 15.78329 | 3.541348 | .045683 | 15.32768 | 10.28163 | 24.15133 |
| var | 4.423026 | .3241646 | .005444 | 4.409645 | 3.824604 | 5.100363 |

- Pig-specific slopes—random coefficient on `week`
- Likelihood model:

$$
\begin{aligned}
\text{weight}_{ij} &= \beta_0 + u_{0j} + (\beta_1 + u_{1j})\texttt{week}_{ij} + \epsilon_{ij} \\
\epsilon_{ij} &\sim \text{Normal}(0,\sigma^2) \\
u_{0j} &\sim \text{Normal}(0,\sigma_0^2) \\
u_{1j} &\sim \text{Normal}(0,\sigma_1^2)
\end{aligned}
$$

where $i = 1, \ldots, 9$ and $j = 1, \ldots, 48$.

- Prior distributions:

$$
\begin{aligned}
\beta_i &\sim \text{Normal}(0,100), \; i = 0, 1 \\
\sigma^2 &\sim \text{InvGamma}(0.001, 0.001) \\
\sigma_0^2 &\sim \text{InvGamma}(0.001, 0.001) \\
\sigma_1^2 &\sim \text{InvGamma}(0.001, 0.001)
\end{aligned}
$$

# Alternative model formulation

- Let $\tau_{0j} = \beta_0 + u_{0j}$ and $\tau_{1j} = \beta_1 + u_{1j}$
- Alternative likelihood model formulation:

$$\begin{aligned}
\text{weight}_{ij} &= \tau_{0j} + \tau_{ij}\texttt{week}_{ij} + \epsilon_{ij} \\
\epsilon_{ij} &\sim \text{Normal}(0,\sigma^2) \\
\tau_{0j} &\sim \text{Normal}(\beta_0,\sigma_0^2) \\
\tau_{1j} &\sim \text{Normal}(\beta_1,\sigma_1^2)
\end{aligned}$$

- Option `reffects()` supports only (two-level) random-intercept models
- Must use the factor-variable specification
- But can replace time-consuming `splitting` with the forthcoming suboption `reffects` in a `block()`

```
. webuse pig
(Longitudinal analysis of pig weights)

. fvset base none id

. set seed 14

. bayesmh weight i.id i.id#c.week, likelihood(normal({var})) noconstant ///
>   prior({weight:i.id},        normal({weight:cons},{var_0}))          ///
>   prior({weight:i.id#c.week}, normal({weight:week},{var_1}))          ///
>   prior({weight:week},        normal(0,100))                          ///
>   prior({weight:cons},        normal(0,100))                          ///
>   prior({var},                igamma(0.001,0.001))                    ///
>   prior({var_0},              igamma(0.001,0.001))                    ///
>   prior({var_1},              igamma(0.001,0.001))                    ///
>   block({weight:i.id},        reffects)                               ///
>   block({weight:i.id#c.week}, reffects)                               ///
>   block({var}, gibbs) block({var_0}, gibbs) block({var_1}, gibbs)     ///
>   block({weight:week}, gibbs) block({weight:cons}, gibbs)             ///
>   burnin(10000) noshow({weight:i.id i.id#c.week}) dots
```

- Model summary

```
Burn-in 10000 aaaaaaaaa1000aa......2000........3000........4000........5000
> ........6000........7000........8000........9000........10000 done
Simulation 10000 .........1000........2000........3000........4000........5
> 000........6000........7000........8000........9000........10000 done
Model summary
------------------------------------------------------------------------------
Likelihood:
  weight ~ normal(xb_weight,{var})
Priors:
         {weight:i.id} ~ normal({weight:cons},{var_0})                     (1)
  {weight:i.id#c.week} ~ normal({weight:week},{var_1})                     (1)
                 {var} ~ igamma(0.001,0.001)
Hyperpriors:
 {weight:week cons} ~ normal(0,100)
      {var_0 var_1} ~ igamma(0.001,0.001)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_weight.
```

```
. bayesstats summary {weight:week cons} {var_0} {var_1} {var}
Posterior summary statistics                      MCMC sample size =     10,000
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| weight | | | | | | |
| week | 6.206141 | .0934325 | .002277 | 6.206412 | 6.02124 | 6.388147 |
| cons | 19.33658 | .4127152 | .013154 | 19.33267 | 18.52088 | 20.14833 |
| var_0 | 7.192013 | 1.73689 | .080111 | 6.972026 | 4.541918 | 11.22479 |
| var_1 | .391377 | .0897799 | .00281 | .3801791 | .2507229 | .5967875 |
| var | 1.616059 | .1252948 | .004119 | 1.608114 | 1.389298 | 1.881644 |

- Relax the assumption of independence between random intercepts and random coefficients
- Likelihood model:

$$
\begin{aligned}
\text{weight}_{ij} &= \tau_{0j} + \tau_{ij}\texttt{week}_{ij} + \epsilon_{ij} \\
\epsilon_{ij} &\sim \mathrm{Normal}(0,\sigma^2) \\
\begin{pmatrix} \tau_{0j} \\ \tau_{1j} \end{pmatrix} &\sim \mathrm{MVN}\left\{ \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix} \right\}
\end{aligned}
$$

where $i = 1, \ldots, 9$ and $j = 1, \ldots, 48$.

- Prior distributions:

$$
\begin{aligned}
\beta_i &\sim \mathrm{Normal}(0{,}100),\ i = 0, 1 \\
\sigma^2 &\sim \mathrm{InvGamma}(0.001, 0.001) \\
\Sigma &\sim \mathrm{InvWishart}(3, I(2))
\end{aligned}
$$

- Forthcoming specification of the `mvnormal()` prior for specifying an unstructured covariance for multiple sets of random effects

```
. set seed 14
. bayesmh weight i.id i.id#c.week, likelihood(normal({var})) noconstant ///
>   prior({weight:i.id i.id#c.week},                                    ///
>     mvnormal(2,{weight:cons},{weight:week},{Sigma, matrix}))          ///
>   prior({weight:week cons},   normal(0,100))                          ///
>   prior({var},                igamma(0.001,0.001))                    ///
>   prior({Sigma,m},            iwishart(2,3,I(2)))                     ///
>   block({weight:i.id},        reffects)                               ///
>   block({weight:i.id#c.week}, reffects)                               ///
>   block({var},                gibbs)                                  ///
>   block({Sigma,m},            gibbs)                                  ///
>   burnin(10000) nomodelsummary notable dots

Burn-in ...
Simulation ...
Bayesian normal regression                        MCMC iterations  =     20,000
Metropolis-Hastings and Gibbs sampling            Burn-in          =     10,000
                                                  MCMC sample size =     10,000
                                                  Number of obs    =        432
                                                  Acceptance rate  =      .5005
                                                  Efficiency:  min =    .005916
                                                               avg =     .01594
Log marginal likelihood = -924.64857                           max =      .1389
```

```
. bayesstats summary {weight:week cons} {Sigma} {var}
Posterior summary statistics                      MCMC sample size =     10,000
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| weight | | | | | | |
| week | 6.212649 | .0965009 | .003403 | 6.214282 | 6.016377 | 6.390494 |
| cons | 19.33385 | .4098845 | .017624 | 19.32329 | 18.51801 | 20.15016 |
| Sigma_1_1 | 6.938195 | 1.637985 | .076558 | 6.735893 | 4.42667 | 10.71507 |
| Sigma_2_1 | -.0926991 | .2678663 | .009932 | -.0843172 | -.656516 | .4238284 |
| Sigma_2_2 | .3997822 | .0893766 | .002398 | .3879609 | .2610762 | .6069753 |
| var | 1.612773 | .1277831 | .004689 | 1.607633 | 1.385116 | 1.881754 |

- Meta analysis is a statistical analysis that involves summarizing results from similar but independent studies.
- Consider data from nine clinical trials that examined the effect of taking diuretics during pregnancy on the risk of preeclampsia (Tanner et al. 2000).
- Data contain estimates of treatment effects expressed as log odds-ratios (`lnOR`) and their respective variances (`var`).
- Negative `lnOR` values indicate that taking diuretics lowers the risk of preeclampsia.

- Likelihood model:

$$\begin{aligned} y_i &\sim \mathrm{Normal}(\mu_i, \mathtt{var}_i) \\ \mu_i &\sim \mathrm{Normal}(\theta, \tau^2) \end{aligned}$$

where $i = 1, \ldots, 9$.

- Prior distributions:

$$\begin{aligned} \theta &\sim \mathrm{Normal}(0{,}10000) \\ \tau^2 &\sim \mathrm{InvGamma}(0.0001, 0.0001) \end{aligned}$$

```
. use meta
(Meta analysis of clinical trials studying diuretics and pre-eclampsia)

. set seed 14

. fvset base none trial

. bayesmh lnOR i.trial, noconstant likelihood(normal(var))   ///
>         prior({lnOR:i.trial}, normal({theta},{tau2}))      ///
>         prior({theta},        normal(0,10000))             ///
>         prior({tau2},         igamma(0.0001,0.0001))       ///
>         block({lnOR:i.trial}, split)                       ///
>         block({theta},        gibbs)                       ///
>         block({tau2},         gibbs)

Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
Likelihood:
  lnOR ~ normal(xb_lnOR,var)

Prior:
  {lnOR:i.trial} ~ normal({theta},{tau2})                                  (1)

Hyperpriors:
  {theta} ~ normal(0,10000)
   {tau2} ~ igamma(0.0001,0.0001)
------------------------------------------------------------------------------
(1) Parameters are elements of the linear form xb_lnOR.
```

```
Bayesian normal regression                         MCMC iterations  =     12,500
Metropolis-Hastings and Gibbs sampling             Burn-in          =      2,500
                                                   MCMC sample size =     10,000
                                                   Number of obs    =          9
                                                   Acceptance rate  =      .6353
                                                   Efficiency:  min =     .01537
                                                                avg =      .0647
Log marginal likelihood =  8.2435069                            max =      .1798
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| lnOR | | | | | | |
| trial | | | | | | |
| 1 | -.2074594 | .3233577 | .014264 | -.2390982 | -.7840912 | .4732284 |
| 2 | -.7422326 | .3059792 | .014353 | -.7277104 | -1.352696 | -.2290158 |
| 3 | -.8101728 | .3579343 | .019156 | -.7938089 | -1.557279 | -.2024199 |
| 4 | -.8860118 | .4367827 | .027156 | -.8529495 | -1.824588 | -.1811792 |
| 5 | -1.032694 | .3685822 | .029732 | -1.046375 | -1.738105 | -.3787439 |
| 6 | -.3225829 | .0969534 | .003571 | -.3241207 | -.5102041 | -.1320317 |
| 7 | -.3476522 | .2873013 | .008138 | -.3712284 | -.8994376 | .2624625 |
| 8 | -.0831874 | .5189861 | .019312 | -.1686125 | -.9203838 | 1.128532 |
| 9 | -.0531772 | .268729 | .016447 | -.0631959 | -.5078684 | .5056795 |
| theta | -.499449 | .2307223 | .005441 | -.4849543 | -.9790357 | -.0413009 |
| tau2 | .3385446 | .4122769 | .016601 | .2325792 | .0003896 | 1.332994 |

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

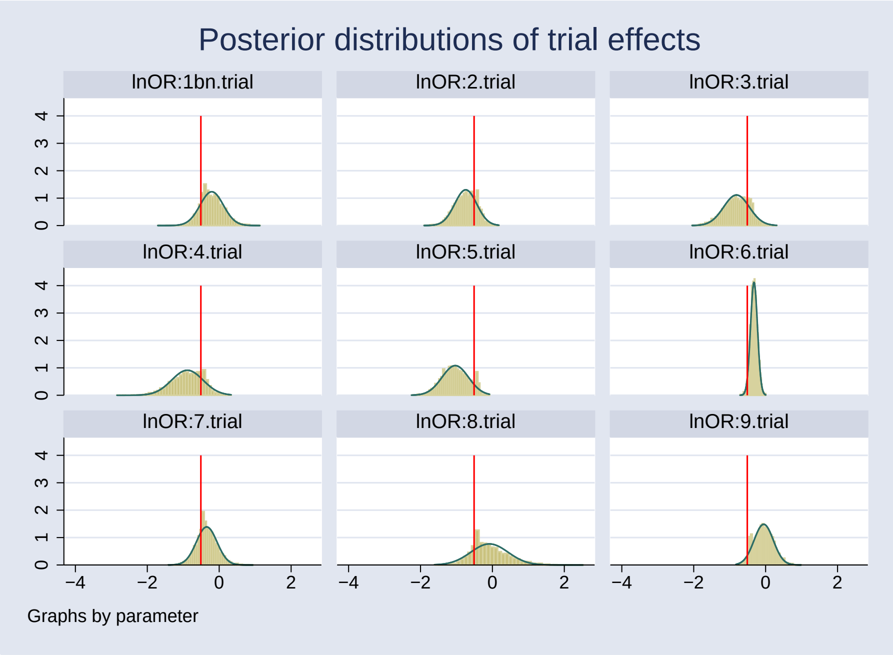
- Test whether taking diuretics reduces the risk of preeclampsia

```
. bayestest interval {theta}, upper(0)
Interval tests     MCMC sample size =    10,000
       prob1 : {theta} < 0
```

| | Mean | Std. Dev. | MCSE |
|---|---|---|---|
| prob1 | .9825 | 0.13113 | .0017971 |

- Plot posterior distributions of trial-specific effects

```
. bayesgraph histogram {lnOR:i.trial},                     ///
>         byparm(legend(off) noxrescale noyrescale         ///
>         title(Posterior distributions of trial effects)) ///
>         normal addplot(pci 0 -0.51 4 -0.51, lcolor(red))
```



Posterior distributions of trial effects
lnOR:1bn.trial
lnOR:2.trial
lnOR:3.trial
lnOR:4.trial
lnOR:5.trial
lnOR:6.trial
lnOR:7.trial
lnOR:8.trial
lnOR:9.trial
0 1 2 3 4
−4 −2 0 2
Graphs by parameter

- British coal mining disaster dataset from 1851 to 1962 (Carlin, Gelfand, and Smith 1992)
- Outcome `count`: number of disasters involving 10 or more deaths
- There was a fairly abrupt decrease in the rate of disasters around 1887–1895.
- Estimate the date, change point *cp*, when the rate of disasters changed.

- Likelihood model:

$$\begin{aligned} \texttt{counts}_i &\sim \text{Poisson}(\mu_1), \ \texttt{if} \ \texttt{year}_i < cp \\ \texttt{counts}_i &\sim \text{Poisson}(\mu_2), \ \texttt{if} \ \texttt{year}_i \geq cp \end{aligned}$$

where $i = 1, \ldots, 112$.

- Prior distributions:

$$\begin{aligned} \mu_1 &\sim 1 \\ \mu_2 &\sim 1 \\ cp &\sim \text{Uniform}(1851, 1962) \end{aligned}$$

```
. webuse coal
(British coal-mining disaster data, 1851-1962)

. set seed 14

. bayesmh count = ({mu1}*sign(year<{cp})+{mu2}*sign(year>={cp})), ///
>     likelihood(poisson, noglmtransform)     ///
>     prior({mu1} {mu2}, flat)                ///
>     prior({cp},        uniform(1851,1962))  ///
>     initial({mu1} 1 {mu2} 1 {cp} 1906)      ///
>     title(Change-point analysis)

Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
Likelihood:
  count ~ poisson({mu1}*sign(year<{cp})+{mu2}*sign(year>={cp}))

Priors:
  {mu1 mu2} ~ 1 (flat)
       {cp} ~ uniform(1851,1962)
------------------------------------------------------------------------------
```

- Estimate the ratio between the two means
- After 1890, the mean number of disasters decreased by a factor of about 3.4 with a 95% credible range of [2.47, 4.55].

```
Change-point analysis                          MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling       Burn-in          =      2,500
                                               MCMC sample size =     10,000
                                               Number of obs    =        112
                                               Acceptance rate  =       .228
                                               Efficiency:  min =     .03747
                                                            avg =     .06763
Log marginal likelihood = -173.29271                        max =      .1193
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| mu1 | 3.118753 | .3001234 | .015504 | 3.110907 | 2.545246 | 3.72073 |
| cp | 1890.362 | 2.4808 | .071835 | 1890.553 | 1886.065 | 1896.365 |
| mu2 | .9550596 | .1209208 | .005628 | .9560248 | .7311639 | 1.219045 |

```
. bayesstats summary (ratio:{mu1}/{mu2})
Posterior summary statistics                      MCMC sample size =     10,000
      ratio : {mu1}/{mu2}
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| ratio | 3.316399 | .5179103 | .027848 | 3.270496 | 2.404047 | 4.414975 |

- Crossover design is a repeated-measures design in which patients crossover from one treatment to another during the course of the study.
- Crossover designs are widely used for testing the efficacy of new drugs.
- Consider a two-treatment, two-period crossover trial comparing two Carbamazepine tablets: A—new and B—standard (Gelfand et al. 1990).
- 10 subjects were randomized to two treatment sequences: AB and BA.
- Outcome: logarithms of maxima of concentration-time curves.

- Likelihood model:

$$y_{i(jk)} = \mu + (-1)^{j-1}\frac{\phi}{2} + (-1)^{k-1}\frac{\pi}{2} + d_i + \epsilon_{i(jk)} = \mu_{i(jk)} + \epsilon_{i(jk)}$$

$$\begin{aligned} \epsilon_{i(jk)} &\sim \text{Normal}(0, \sigma^2) \\ d_i &\sim \text{Normal}(0, \tau^2) \end{aligned}$$

  where $i = 1, \ldots, 10$, $j = 1, 2$ is the treatment group (sequence), and $k = 1, 2$ is the period.

- Prior distributions:

$$\begin{aligned} \mu,\ \phi,\ \pi &\sim \text{Normal}(0, 10^6) \\ \sigma^2 &\sim \text{InvGamma}(0.001, 0.001) \\ \tau^2 &\sim \text{InvGamma}(0.001, 0.001) \end{aligned}$$

```
. webuse bioequiv
(Bioequivalent study of Carbamazepine tablets)

. set seed 14

. fvset base none id

. bayesmh y = ({mu}+(-1)^(treat-1)*{phi}/2+(-1)^(period-1)*{pi}/2+{y:i.id}), ///
>   likelihood(normal({var}))                                 ///
>   prior({y:i.id},       normal(0,{tau2}))                   ///
>   prior({tau2},         igamma(0.001,0.001))                ///
>   prior({var},          igamma(0.001,0.001))                ///
>   prior({mu} {phi} {pi}, normal(0,1e6))                     ///
>   block({y:i.id}, reffects)                                 ///
>   block({tau2},   gibbs)                                    ///
>   block({var},    gibbs)                                    ///
>   adaptation(every(200) maxiter(50)) burnin(10000)          ///
>   noshow({y:i.id})
```

```
Model summary
------------------------------------------------------------------------------
Likelihood:
  y ~ normal(<expr1>,{var})

Priors:
        {var} ~ igamma(0.001,0.001)
     {y:i.id} ~ normal(0,{tau2})
  {mu phi pi} ~ normal(0,1e6)

Hyperprior:
  {tau2} ~ igamma(0.001,0.001)

Expression:
  expr1 : {mu}+(-1)^(treat-1)*{phi}/2+(-1)^(period-1)*{pi}/2+({y:1bn.id}*1bn.i
          d+{y:2.id}*2.id+{y:3.id}*3.id+{y:4.id}*4.id+{y:5.id}*5.id+{y:6.id}*6
          .id+{y:7.id}*7.id+{y:8.id}*8.id+{y:9.id}*9.id+{y:10.id}*10.id)
------------------------------------------------------------------------------
```

```
Bayesian normal regression                          MCMC iterations  =     20,000
Metropolis-Hastings and Gibbs sampling              Burn-in          =     10,000
                                                    MCMC sample size =     10,000
                                                    Number of obs    =         20
                                                    Acceptance rate  =      .5959
                                                    Efficiency:  min =     .01359
                                                                 avg =     .03528
Log marginal likelihood = -8.6538165                             max =      .0511
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| mu | 1.444575 | .0492361 | .004224 | 1.444955 | 1.350906 | 1.54269 |
| phi | -.0092691 | .0537334 | .00255 | -.0087842 | -.1126505 | .0939082 |
| pi | -.1768478 | .0517259 | .002288 | -.1785769 | -.2839622 | -.0668874 |
| var | .0136361 | .0090926 | .000637 | .0109485 | .004295 | .0377165 |
| tau2 | .02173 | .0175663 | .000811 | .017856 | .0023005 | .0647257 |

- $\theta = \exp(\phi)$ is commonly used as a measure of bioequivalence.
- Bioequivalence is declared whenever $\theta$ lies in the interval (0.8, 1.2) with a high posterior probability.

```
. bayesstats summary (equiv:exp({phi})>0.8 & exp({phi})<1.2)
Posterior summary statistics                         MCMC sample size =    10,000
       equiv : exp({phi})>0.8 & exp({phi})<1.2
```

| | Mean | Std. Dev. | MCSE | Median | Equal-tailed [95% Cred. | Interval] |
|---|---|---|---|---|---|---|
| equiv | .9937 | .0791261 | .003951 | 1 | 1 | 1 |