

Bayesian statistics

Hồi Quy Tuyến Tính t test & ANOVA

Khương Quỳnh Long

Hà Nội, 08/2019

<https://gitlab.com/LongKhuong/adhere-bayesian-statistics>

Nội dung

1

- Bayes prefix

2

- Các bước xây dựng hồi quy tuyến tính

3

- Chẩn đoán chuỗi MCMC

4

- Kiểm định giả thuyết

5

- Bayes T test và Bayes ANOVA

“bayes” prefix

```
bayes [ , bayesopts]: command [ , estopts]
```

Bayes prefix và bayes options

Tương tự MLE

Bayes options:

- Block về Prior (giới thiệu trong bài sau)
- Block về chuỗi MCMC (giới thiệu trong bài sau)
- Block về initialization
- Block về reporting
- nchains(#): song song nhiều chuỗi MCMC (Stata ≥ 16)
-

Block về initialization

- ▶ Mặc định Stata sử dụng MLE “dẫn đường” cho điểm bắt đầu chuỗi MCMC
- ▶ “**initsummary**” khai báo thông tin giá trị bắt đầu
- ▶ “**nomleintitial**” ngăn sử dụng MLE làm điểm bắt đầu
- ▶

Block về reporting

- ▶ “**credible** (#)”: tùy chỉnh Credible Interval, mặc định 95%
- ▶ “**hpd**”: dùng Highest Probability Density Interval thay cho Credible Interval

Một số option khác

- ▶ “`saving(filename[, replace])`”: lưu trữ kết quả mô phỏng vào file “`filename.dta`”
- ▶ “`showeffects[(reref)]`”: trong multilevel → hiển thị các random effects
- ▶ “`melabel`”: hiển thị nhãn kết quả trong multilevel giống như phương pháp MLE
- ▶ ...

Các bước xây dựng hồi quy tuyến tính Bayes

- Câu hỏi nghiên cứu, biến số...
- Xác định Prior trước khi phân tích
- (Sử dụng sensitivity analysis để đánh giá Prior)
- Tính toán mô hình ban đầu
- Chẩn đoán chuỗi MCMC
- Kiểm tra các giá trị như ESS, DIC...
- Khai thác thông tin từ pp hậu nghiệm
- Kiểm định giả thuyết và kết luận

Data

Body Fat data

- ▶ Dữ liệu từ nghiên cứu của Roger W. Johnson và Carleton College
<https://ww2.amstat.org/publications/jse/v4n1/dataset/s.johnson.html>
- ▶ Mục tiêu nghiên cứu là xây dựng mô hình hồi quy tuyến tính tiên lượng % mỡ cơ thể (biến bodyfat) từ các chỉ số nhân trắc

des

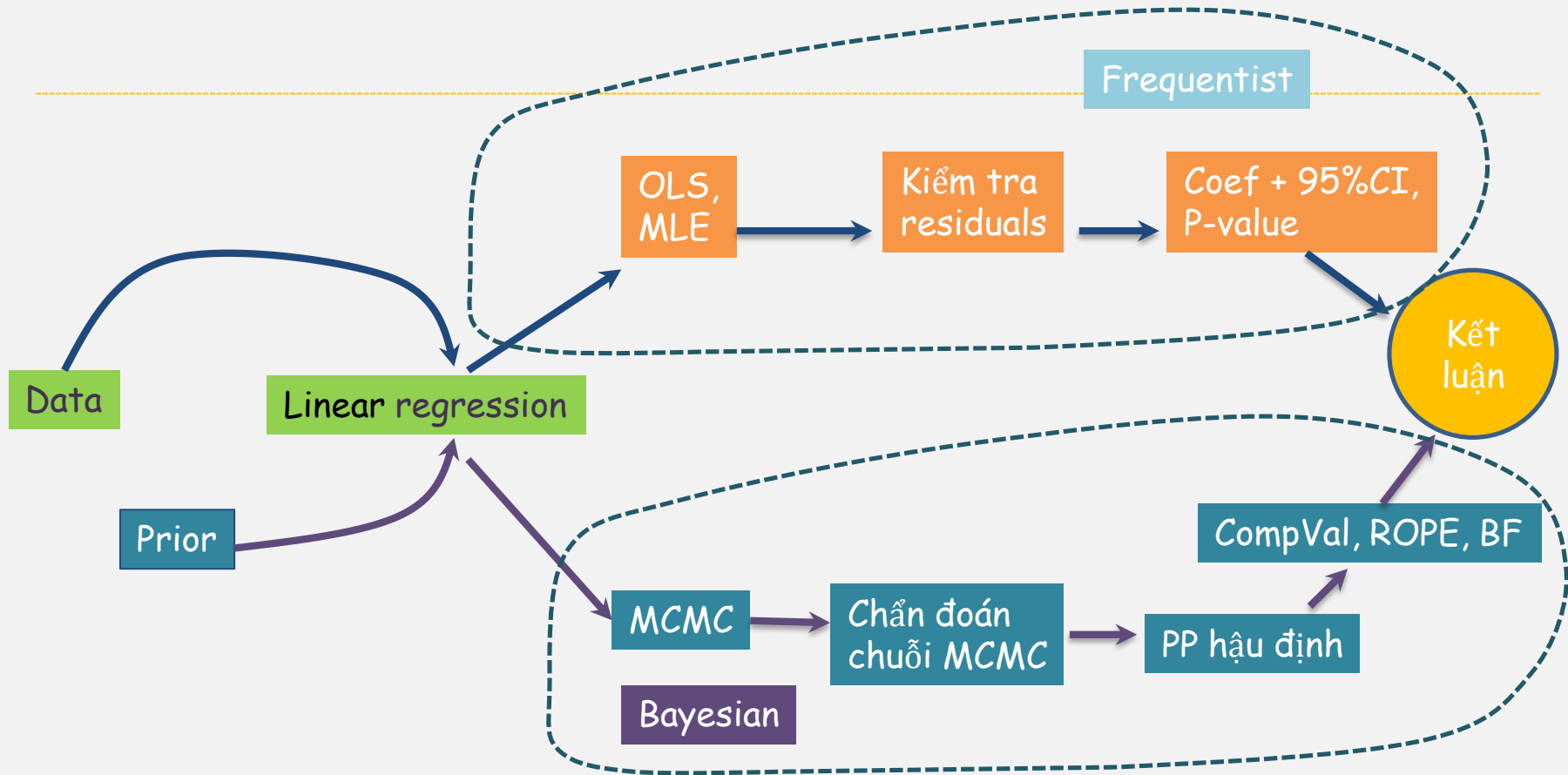
```
obs:      251
vars:      15
size:     26,857
3 Oct 2018 15:56
```

variable name	storage type	display format	value label	variable label
id	int	%10.0g		id
bodyfat	double	%10.0g		%
age	byte	%10.0g		yrs
weight	double	%10.0g		lbs (1 pound = 0,45359237 kg)
height	double	%10.0g		inches (1 inch =2.54 cm)
neck	double	%10.0g		Neck circumference (cm)
chest	double	%10.0g		Chest circumference (cm)
abdomen	double	%10.0g		Abdomen circumference (cm)
hip	double	%10.0g		Hip circumference (cm)
thigh	double	%10.0g		Thigh circumference (cm)
knee	double	%10.0g		Knee circumference (cm)
ankle	double	%10.0g		Ankle circumference (cm)
biceps	double	%10.0g		Extended biceps circumference (cm)
forearm	double	%10.0g		Forearm circumference (cm)
wrist	double	%10.0g		Wrist circumference (cm)

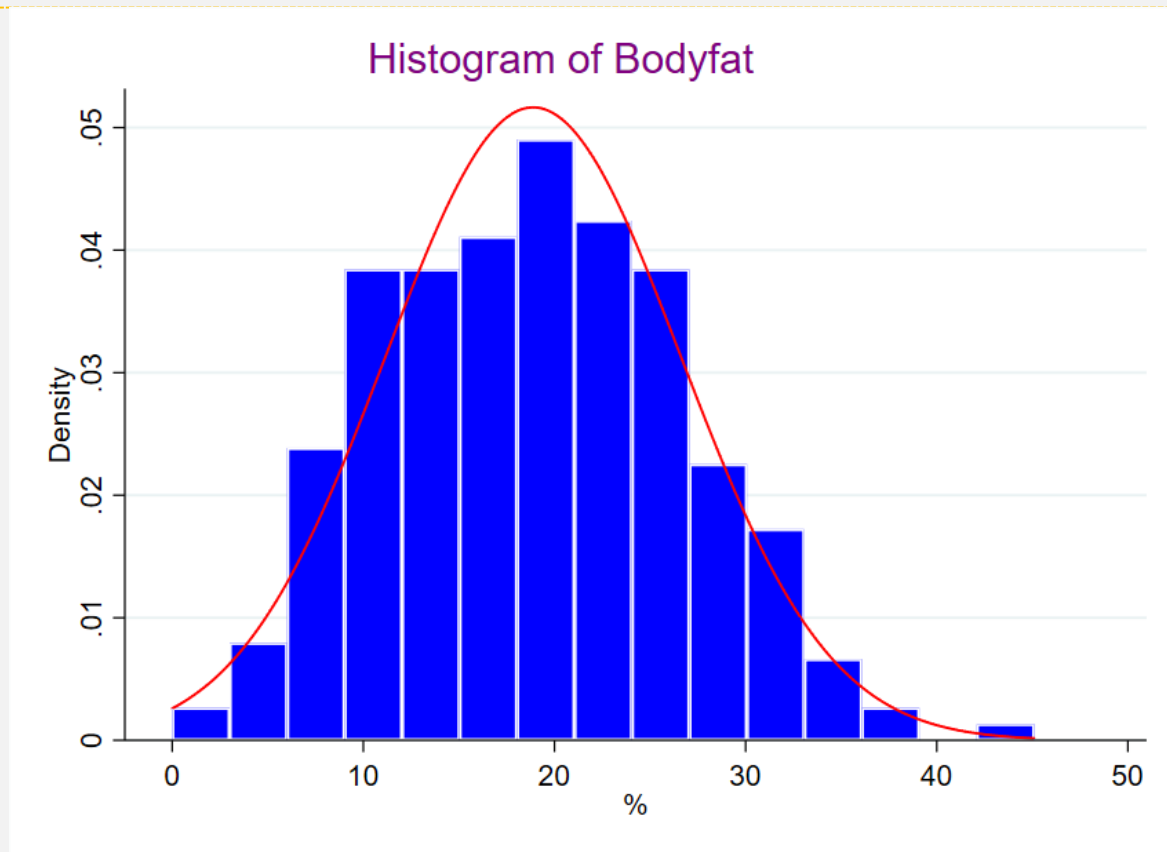
Sorted by:

Xây dựng hồi quy tuyến tính bayes nhằm xác định mối liên quan giữa % mỡ cơ thể (bodyfat) và cân nặng (weight) (pound)

$$\text{bodyfat} = a + b * \text{weight} + e$$



Phân phối của Bodyfat



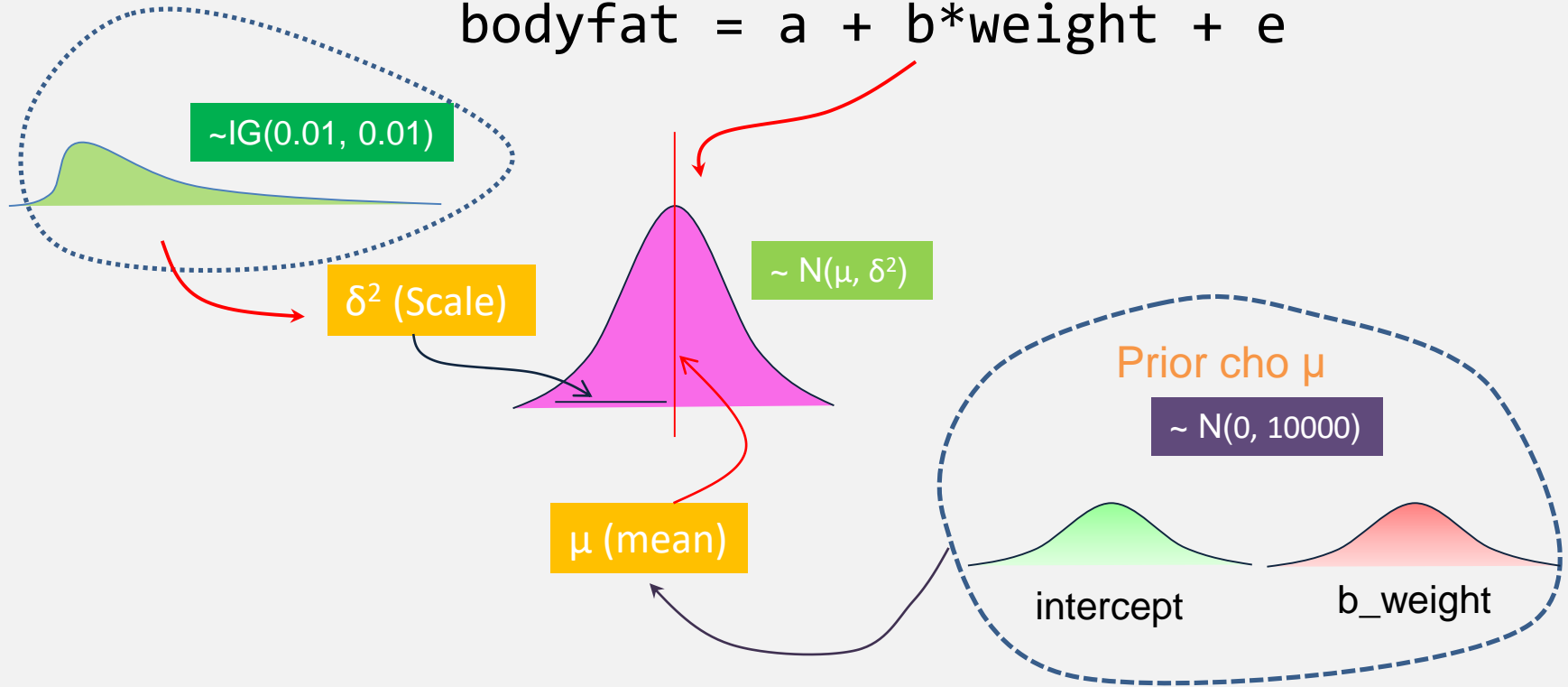
Xác định Prior trước khi phân tích

- ▶ (Đề cập trong bài sau)
 - ▶ Trong ví dụ này Sử dụng Prior mặc định trong stata
 - Prior của intercept
 - Prior của weight $\sim N(0, 10000)$
 - Prior của $\sigma^2 \sim IG(0.01, 0.01)$
- } Location của Bodyfat
- } Phương sai của Bodyfat

$$\text{bodyfat} = a + b * \text{weight} + e$$

Prior cho δ^2

$$\text{bodyfat} = a + b * \text{weight} + e$$



▶ Hồi quy tuyến tính bằng OLS

reg bodyfat weight

▶ Xây dựng mô hình bayesian bằng prefix “bayes”

bayes: reg bodyfat weight

- ▶ Stata dùng mặc định chuỗi MCMC gồm:
 - 12500 iterations, burn-in 2500 iterations
 - Thinning = 1

Model summary

Likelihood:

```
bodyfat ~ regress(xb_bodyfat,{sigma2})
```

Priors:

```
{bodyfat:weight _cons} ~ normal(0,10000)
```

```
{sigma2} ~ igamma(.01,.01)
```

Prior cho intercept và weight

Prior cho σ^2

(1)

(1) Parameters are elements of the linear form xb_bodyfat.

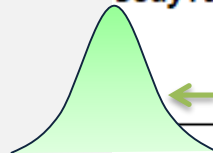
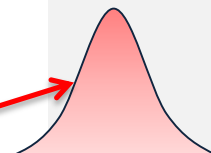
Bayesian linear regression

Random-walk Metropolis-Hastings sampling

Thông số chuỗi MCMC

MCMC iterations	=	12,500
Burn-in	=	2,500
MCMC sample size	=	10,000
Number of obs	=	251
Acceptance rate	=	.3335
Efficiency: min	=	.1092
avg	=	.1432
max	=	.2096

Log marginal likelihood = -830.68517

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred. Interval]	
bodyfat						
weight	.1609436	.0130591	.000392	.1611812	.1344475	.1869115
_cons	-9.876793	2.355601	.071298	-9.944696	-14.54835	-5.160426
sigma2	37.6969	3.434352	.075021	37.53565	31.47561	44.78007

Note: Default priors are used for model parameters.

► Cần chẩn đoán MCMC (MCMC diagnostics) trước khi diễn giải kết quả !

bayesgraph

- ▶ bayesgraph cung cấp các biểu đồ tóm tắt kết quả và chẩn đoán chuỗi MCMC
- **histogram** (pp hậu định cho tham số mô hình)
- **trace**
- **ac** (autocorrelation)
- **kdensity** (pp hậu định cho tham số + chia nửa trước, nửa sau của chuỗi MCMC)
- **cusum** (đánh giá xu hướng chuỗi MCMC)
- **R hat (Gelman-Rubin)**: Khi chạy song song nhiều chuỗi MCMC (Stata ≥ 16)

Chẩn đoán chuỗi MCMC

- ▶ MCMC chứa nội dung của toàn bộ mô hình và phải đại diện cho pp hậu nghiệm
- ▶ Sử dụng trace plot (chủ yếu)
- ▶ Mục đích:
 - Chuỗi MCMC có ổn định hay không?
 - Chuỗi MCMC có bị “mắc kẹt” tại điểm nào không?
 - Xu hướng chuỗi MCMC?

Trace plot

vẽ trace plot cho tham số weight

```
bayesgraph trace {bodyfat: weight}
```

vẽ trace plot cho tham số intercept

```
bayesgraph trace {bodyfat: _cons}
```

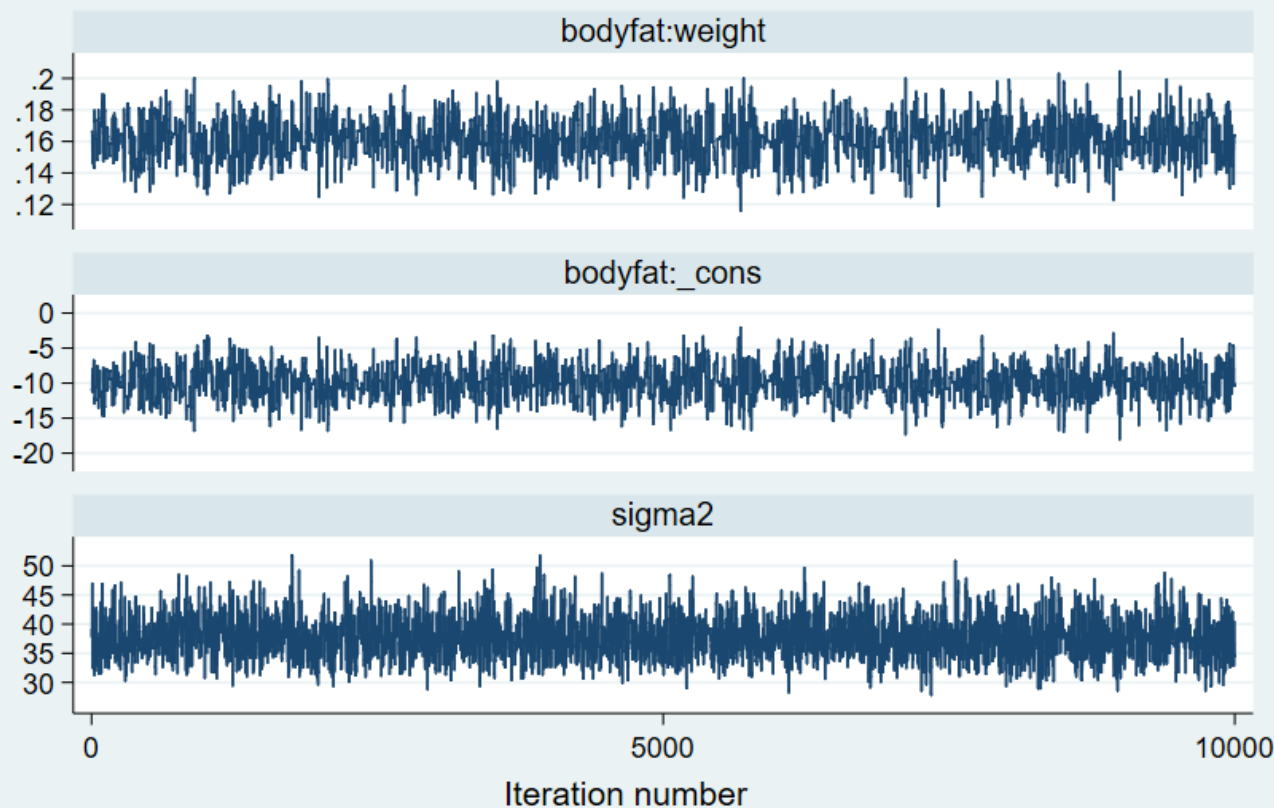
vẽ trace plot cho cả 2 tham số trong 1 hình

```
bayesgraph trace {bodyfat: _cons} {bodyfat: weight}, byparm
```

#Vẽ tất cả các tham số

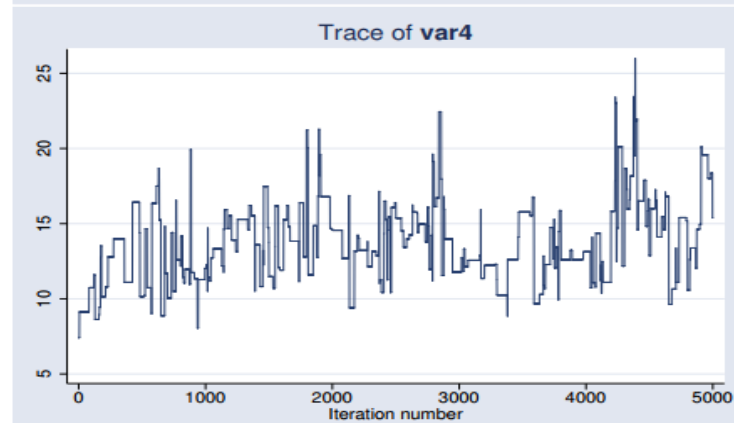
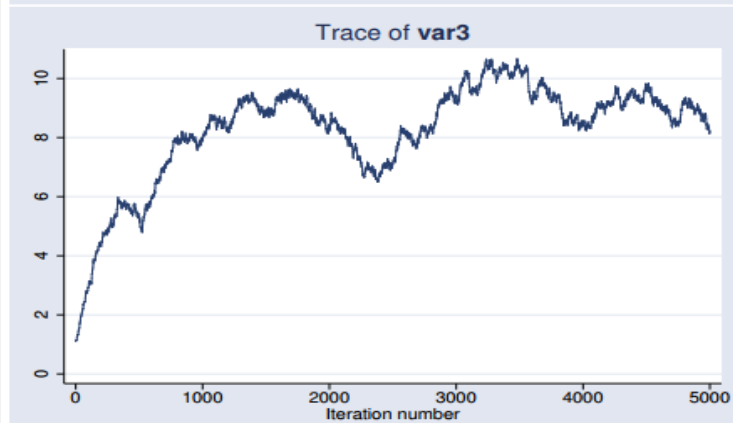
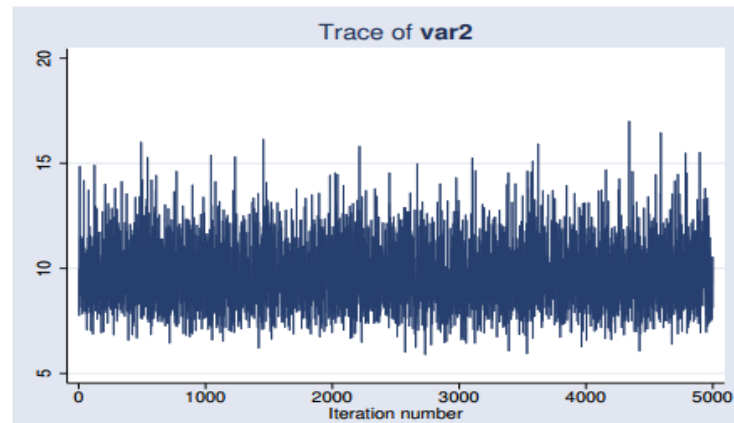
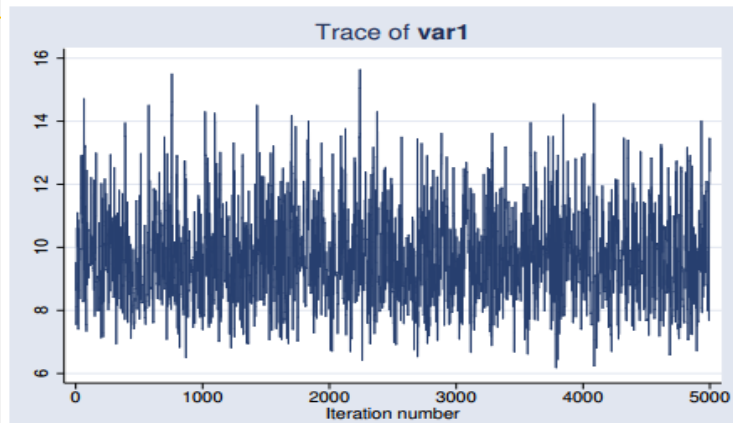
```
bayesgraph trace _all, byparm
```

Trace plots cho các tham số của mô hình



Graphs by parameter

Một số ví dụ



Auto correlation plot

vẽ autocorrelation plot cho tham số weight

```
bayesgraph ac {bodyfat: weight}
```

vẽ autocorrelation plot cho tham số intercept

```
bayesgraph ac {bodyfat: _cons}
```

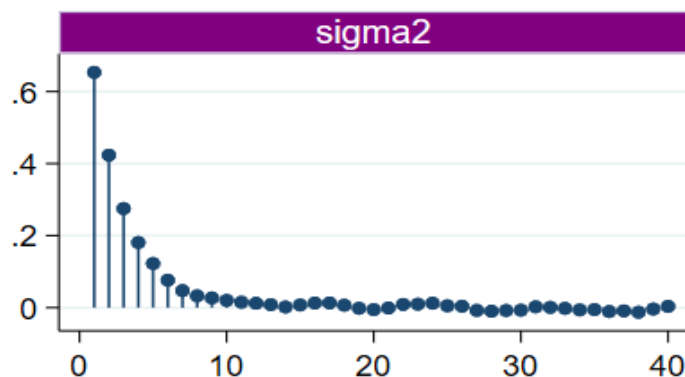
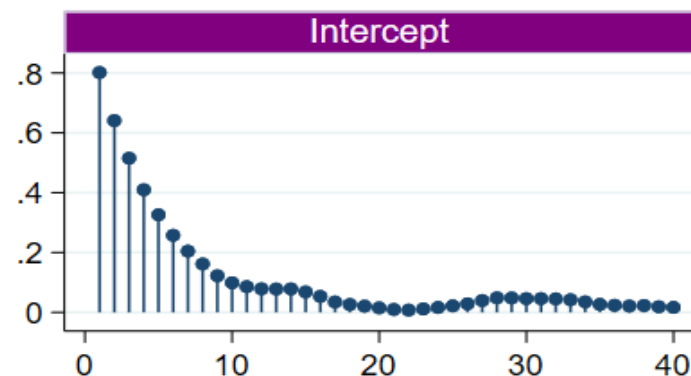
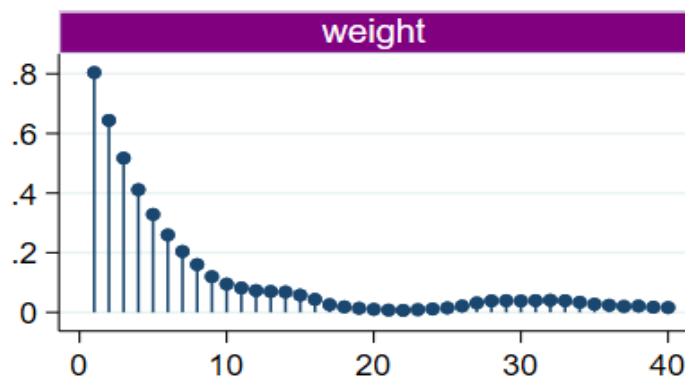
vẽ autocorrelation plot cho cả 2 tham số trong 1 hình

```
bayesgraph ac {bodyfat: _cons} {bodyfat: weight}, byparm
```

#Vẽ tất cả các tham số

```
bayesgraph ac _all, byparm
```

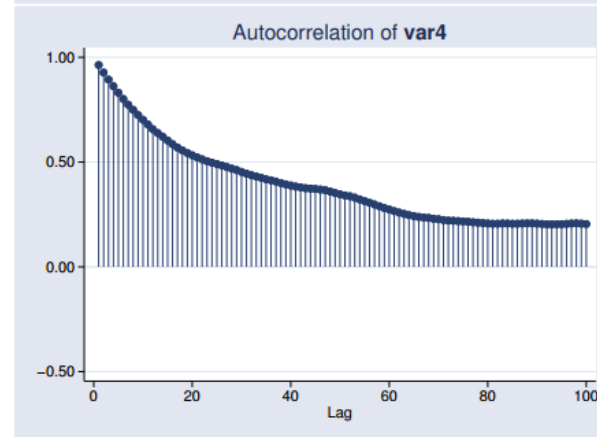
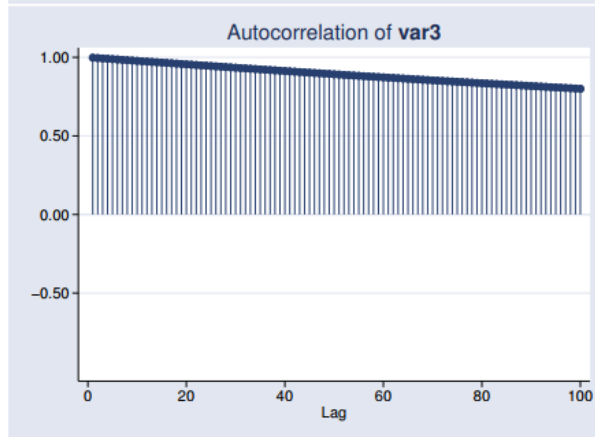
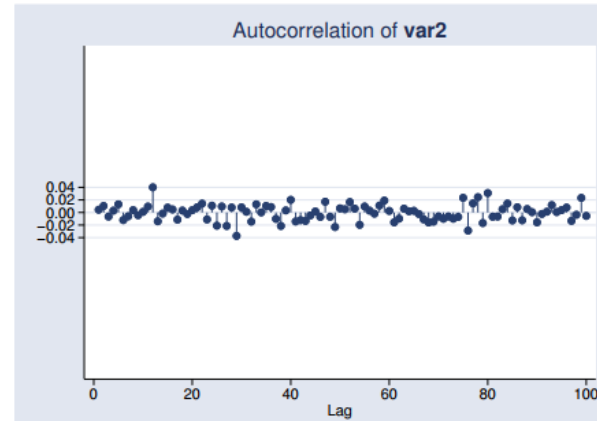
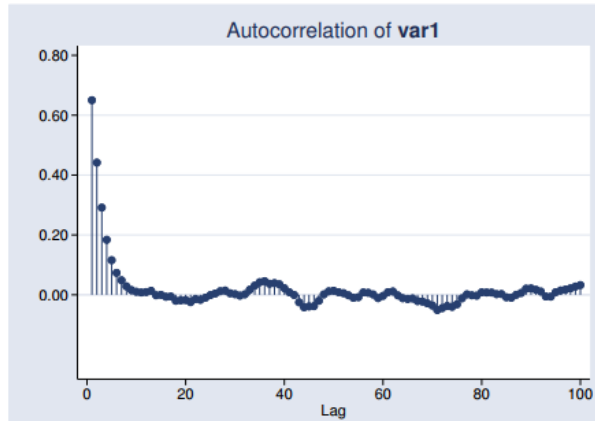

Autocorrelations cho các tham số của mô hình



Lag

Graphs by parameter

Một số ví dụ



Cumulative sums plot

vẽ cusum plot cho tham số weight

```
bayesgraph cusum {bodyfat: weight}
```

vẽ cusum plot cho tham số intercept

```
bayesgraph cusum {bodyfat: _cons}
```

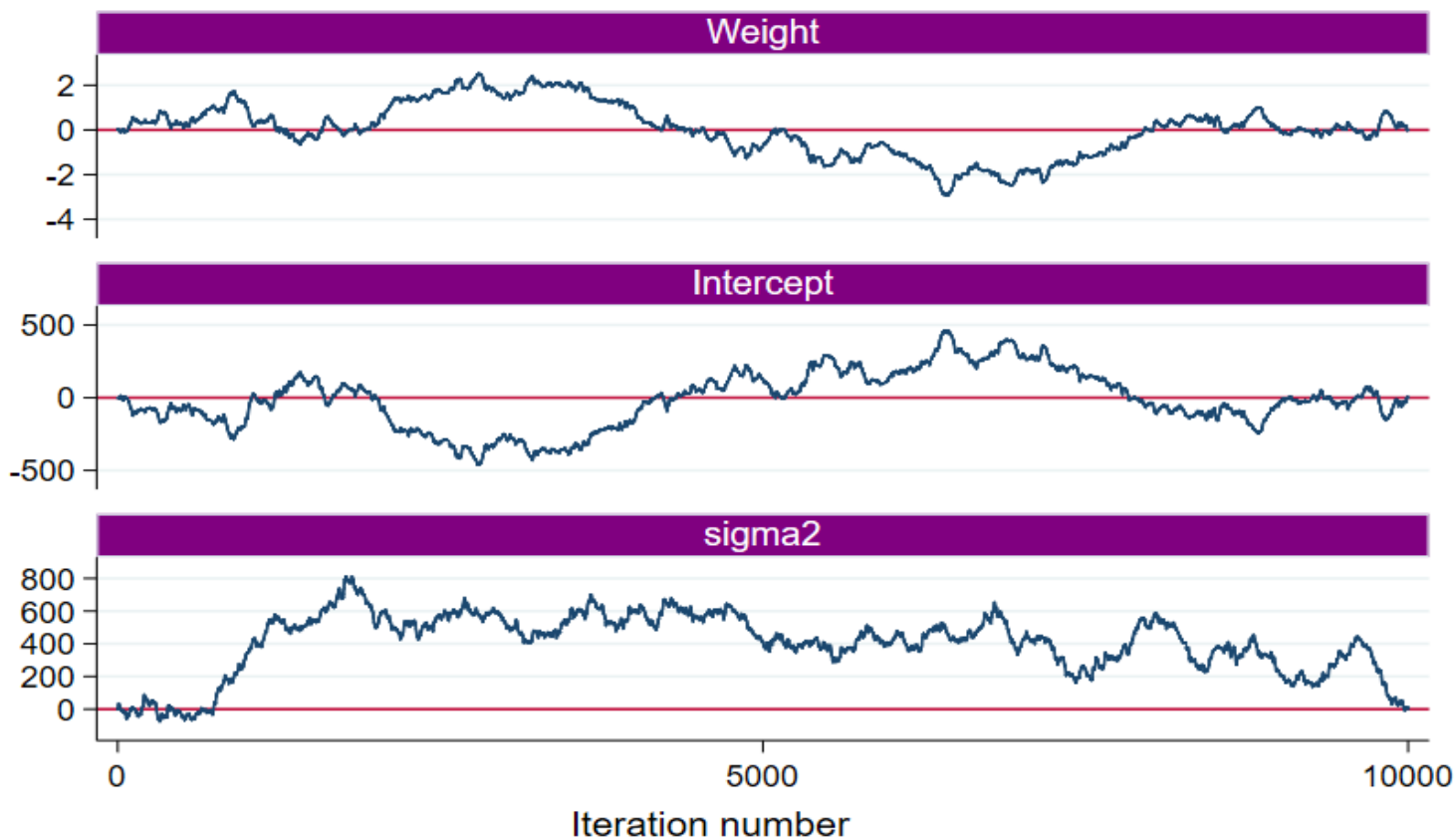
vẽ cusum plot cho cả 2 tham số trong 1 hình

```
bayesgraph cusum {bodyfat: _cons} {bodyfat: weight}, byparm
```

#Vẽ tất cả các tham số

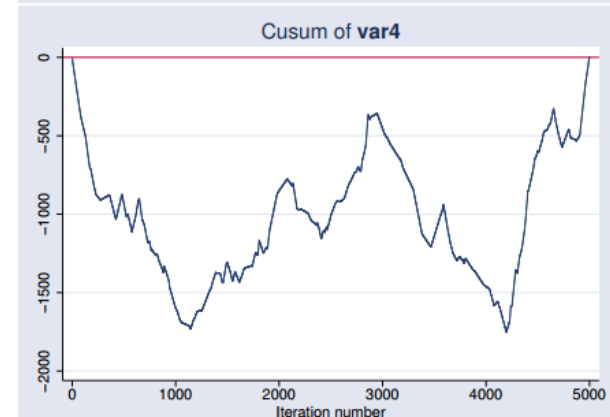
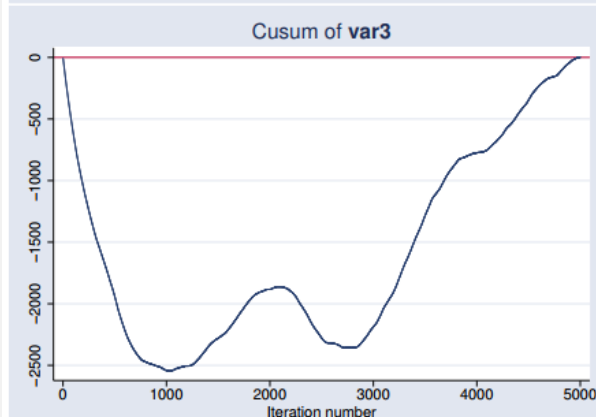
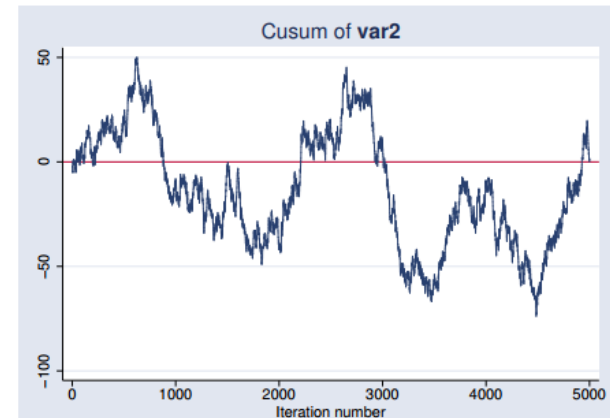
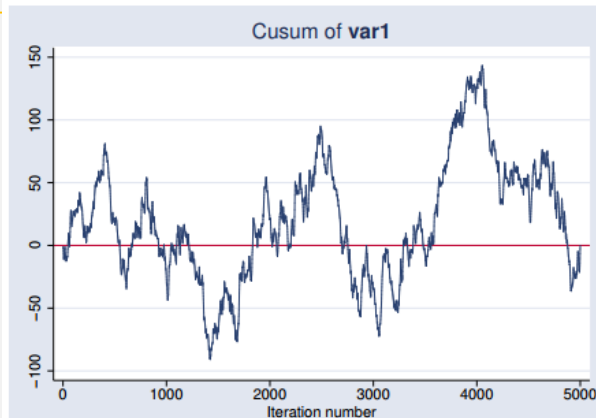
```
bayesgraph cusum _all, byparm
```

Cusum plots cho các tham số của mô hình



Graphs by parameter

Một số ví dụ



Vẽ phân phối hậu định

vẽ pp hậu định cho tham số weight

```
bayesgraph hist {bodyfat: weight}
```

vẽ pp hậu định cho tham số intercept

```
bayesgraph hist {bodyfat: _cons}
```

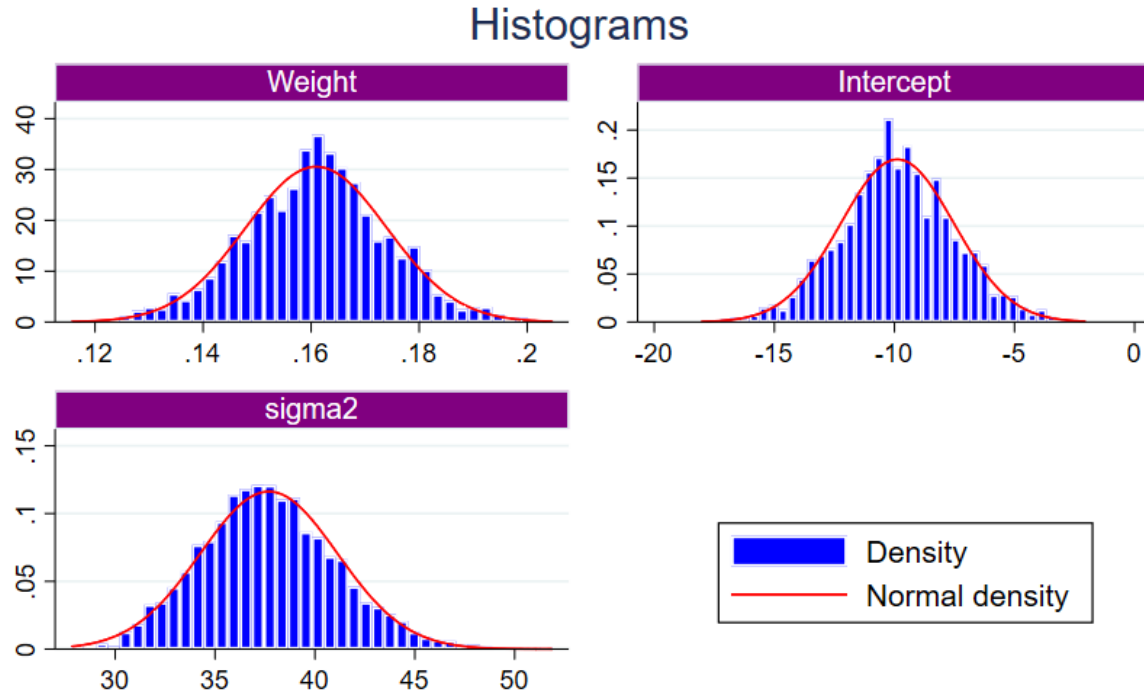
vẽ pp hậu định cho cả 2 tham số trong 1 hình

```
bayesgraph hist {bodyfat: _cons} {bodyfat: weight}, byparm
```

#Vẽ tất cả các tham số

```
bayesgraph hist _all, byparm norm
```

Vẽ phân phối hậu định



Vẽ 4 loại biểu đồ trong cùng 1 hình

`bayesgraph diagnostics {bodyfat: weight}`



Kiểm tra effect sample size

bayesstats ess

12,500 iteration – 2,500 = 10,000

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

Sample size*efficiency

	ESS	Corr. time	Efficiency
bodyfat			
weight	1108.91	9.02	0.1109
_cons	1091.58	9.16	0.1092
sigma2	2095.68	4.77	0.2096

1/ corr.time

Diễn giải kết quả

bayesstats summary

Posterior summary statistics					MCMC sample size = 10,000	
	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred. Interval]	
bodyfat						
weight	.1609436	.0130591	.000392	.1611812	.1344475	.1869115
_cons	-9.876793	2.355601	.071298	-9.944696	-14.54835	-5.160426
sigma2	37.6969	3.434352	.075021	37.53565	31.47561	44.78007

Cân nặng càng cao tỉ lệ mỡ càng lớn, posterior median of coefficient = 0.16 (95% credible interval = 0.13 – 0.19)

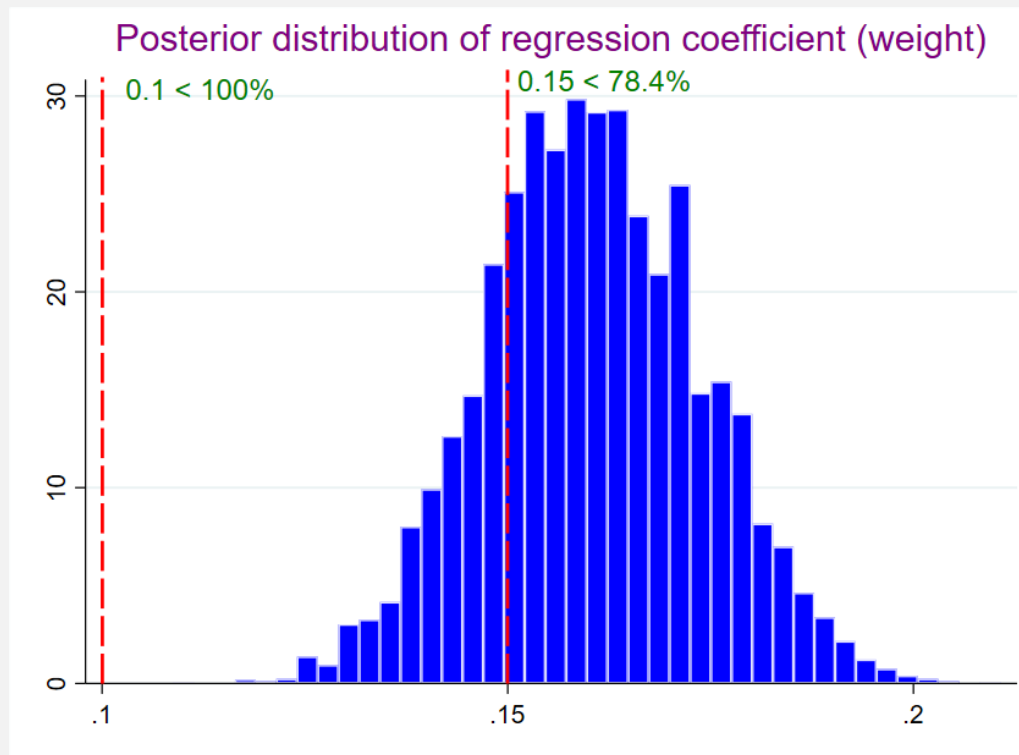
Kiểm định giả thuyết

- ▶ Bộ 3: CompVal, ROPE và Bayes Factor
 - CompVal: kiểm tra pp hậu nghiệm với ngưỡng 0
 - ROPE: [giả sử] trên thực hành lâm sàng, % mỡ thay đổi theo mỗi pound trọng lượng cơ thể trong khoảng 0 - 0.1 là không có ý nghĩa.
 - Bayes factor: tỷ trọng chứng cứ cho giả thuyết H_1 : Coef > 0.1 vs. H_2 : Coef < 0.1 hoặc H_1 : Coef > 0.15 vs. H_2 : Coef < 0.15....
 - Bayesian cho phép kiểm tra nhiều giả thuyết mà không gặp phải vấn đề kiểm định nhiều giả thuyết như Frequentist

CompVal

```
bayestest interval ({bodyfat: weight}, lower(0)) ///  
({bodyfat: weight}, lower(0.1)) ///  
({bodyfat: weight}, lower(0.15))
```

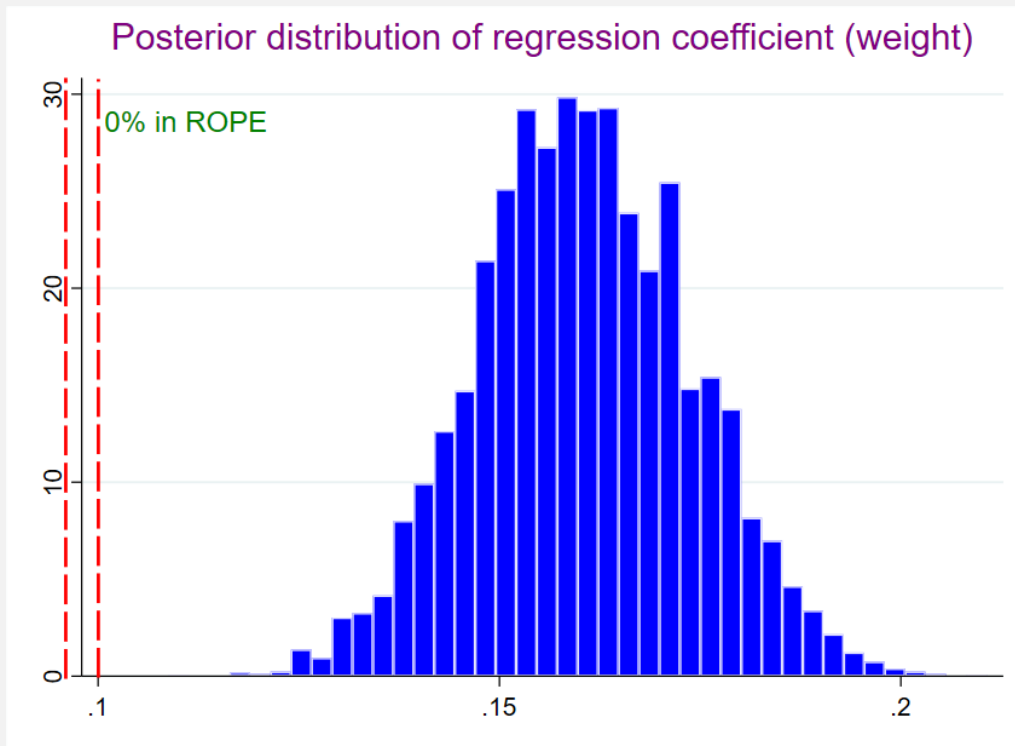
- ▶ 100% mật độ xác suất pp hậu nghiệm của hệ số phương trình nằm trên ngưỡng 0 và 0.1
- ▶ 78.4% nằm trên ngưỡng 0.15



ROPE

bayestest interval {bodyfat: weight}, lower(0) upper(0.1)

100% mật độ xác suất pp
hậu nghiệm của hệ số
phương trình nằm ngoài
(nằm ở vùng cao hơn)
khoảng ROPE (0 - 0.1)



Bayes Factor

▶ **Giả thuyết 1:** H_1 : Coef > 0.1 vs. H_2 : Coef < 0.1

$BF[H_1:H_2] = 100 / 0 = +\text{Inf} \rightarrow$ bằng chứng rất mạnh (gần như chắc chắn) nghiêng về giả thuyết H_1

▶ **Giả thuyết 2:** H_1 : Coef > 0.15 vs. H_2 : Coef < 0.15

$BF[H_1:H_2] = 78.4 / 21.6 = 3.6 \rightarrow$ bằng chứng yếu nghiêng về H_1

Tóm tắt

- Xây dựng mô hình bằng “**bayes**” prefix
- Chẩn đoán chuỗi MCMC (trace plot, autocorrelation, cusum và kdensity oplot) bằng “**bayesgraph**” prefix
- Kiểm tra các giá trị ESS, DIC...bằng “**bayesstats**” prefix
- Khai thác thông tin từ pp hậu nghiệm (ước lượng điểm, ước lượng khoảng)
- Kiểm định giả thuyết (CompVal, ROPE và BF) bằng “**bayestest**” prefix

Bayes t test và Bayes ANOVA

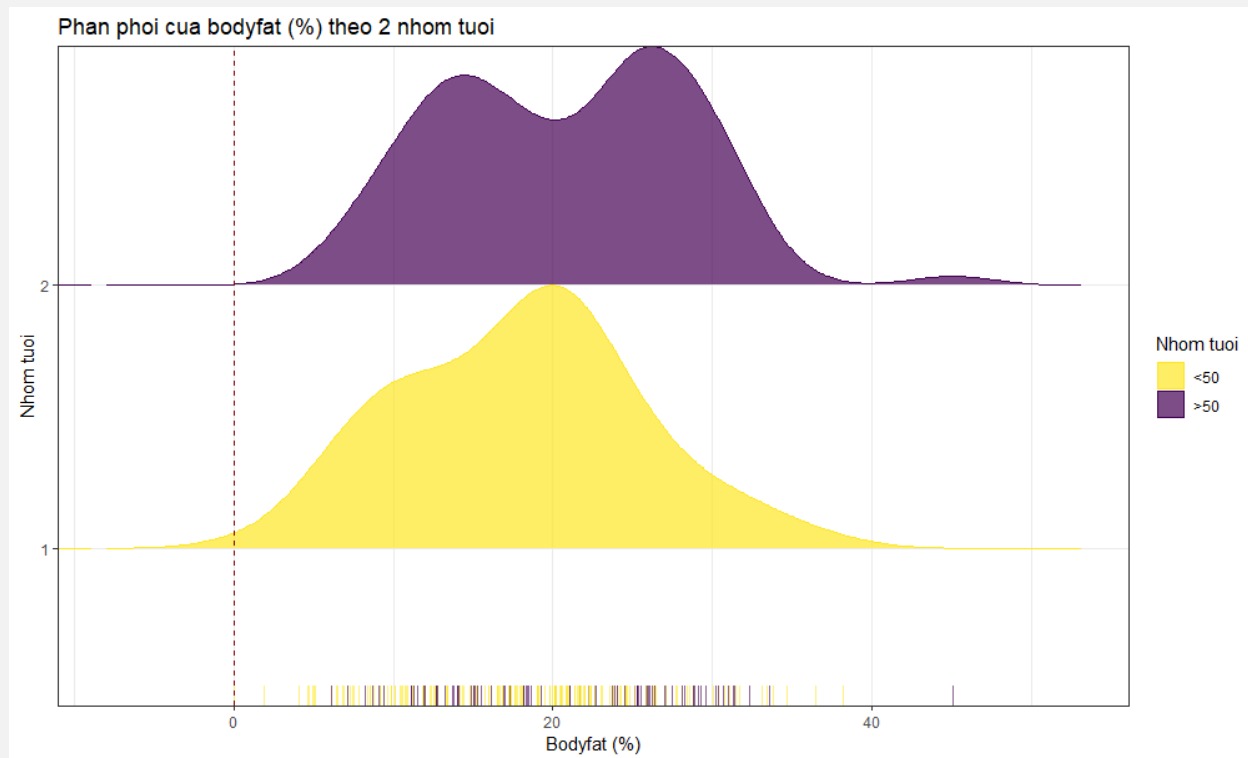
T test và ANOVA

- ▶ T test và ANOVA cũng là một mô hình hồi quy tuyến tính, mô tả biến phụ thuộc theo biến phân nhóm (2 cho t test và >2 cho ANOVA)
- ▶ Câu hỏi ?
 - Tỷ lệ mỡ cơ thể khác biệt như thế nào giữa 2 nhóm <50 tuổi và ≥ 50 tuổi
 - Tỷ lệ mỡ cơ thể khác biệt như thế nào giữa 3 nhóm <30 tuổi, $30 - <50$ tuổi và ≥ 50 tuổi

t test

tạo nhóm tuổi

gen age_group1 = age >=50



ttest bodyfat, by(age_group1)

```
. ttest bodyfat, by(age_group1)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	168	17.95893	.5841245	7.571119	16.80571	19.11215
1	83	20.76747	.8488953	7.733805	19.07875	22.45619
combined	251	18.88765	.4875422	7.724121	17.92744	19.84786

diff	-2.808541	1.023028	-4.823432	-.7936504
------	-----------	----------	-----------	-----------

diff = mean(0) - mean(1)

Ho: diff = 0

Ha: diff < 0

Pr(T < t) = 0.0032

Ha: diff != 0

Pr(|T| > |t|) = 0.0065

Ha: diff > 0

Pr(T > t) = 0.9968

t = -2.7453
degrees of freedom = 249

reg bodyfat age_group1

```
. reg bodyfat age_group1
```

Source	SS	df	MS	Number of obs	=	251
Model	438.202937	1	438.202937	F(1, 249)	=	7.54
Residual	14477.3088	249	58.1418023	Prob > F	=	0.0065
Total	14915.5117	250	59.6620469	R-squared	=	0.0294
				Adj R-squared	=	0.0255
				Root MSE	=	7.6251

bodyfat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age_group1	2.808541	1.023028	2.75	0.006	.7936504	4.823432
_cons	17.95893	.5882875	30.53	0.000	16.80027	19.11758

Hệ số phương trình chính là sự khác biệt của nhóm 1 so với nhóm 0 (intercept)

Bayes t test

bayes: reg bodyfat age_group1

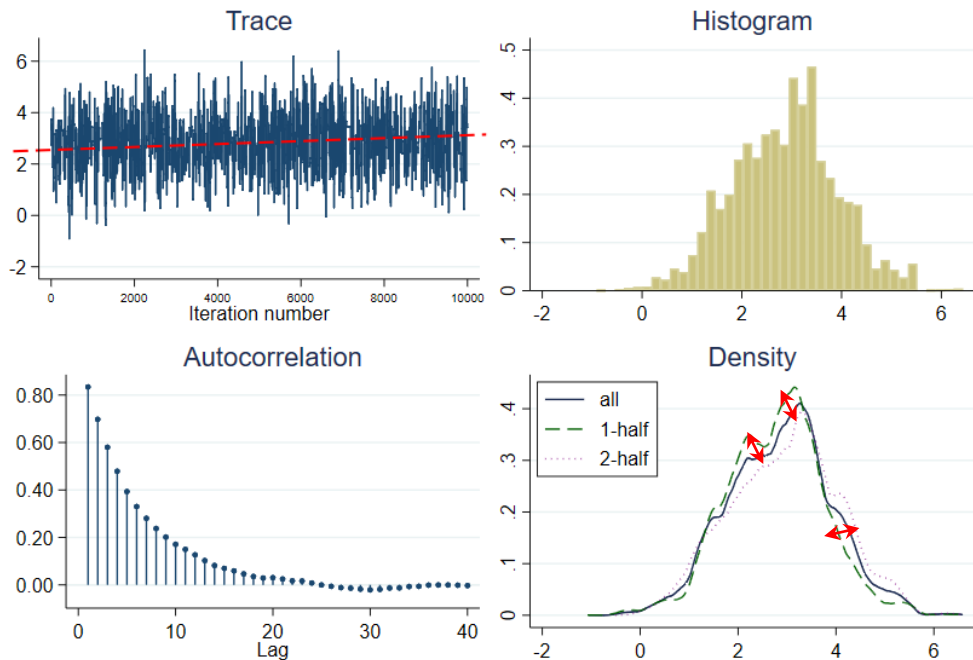
bayesgraph diagnostics {bodyfat: age_group1}

bayesstats ess

#ESS = 909.85, efficiency = 0.0910

Chưa “đủ” hội tụ !!

Sự khác biệt bodyfat giữa nhóm trên 50 so với nhóm dưới 50 tuổi



bayes, mcmcsize(15000) burnin(5000) thinning(5): reg bodyfat age_group1

15000 * 5 + 5000 = 80000 iterations

bayesgraph diagnostics {bodyfat: age_group1}

bayesstats ess

#ESS = 7944.8, efficiency = 0.53

Likelihood:

```
bodyfat ~ regress(xb_bodyfat,{sigma2})
```

Priors:

```
{bodyfat:age_group1_cons} ~ normal(0,10000)
{sigma2} ~ igamma(.01,.01)
```

(1) Parameters are elements of the linear form xb_bodyfat.

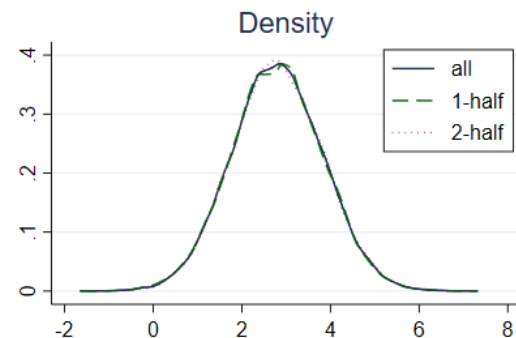
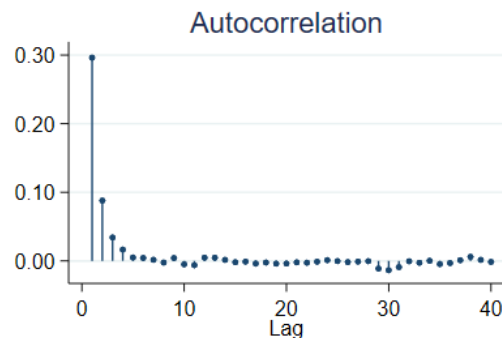
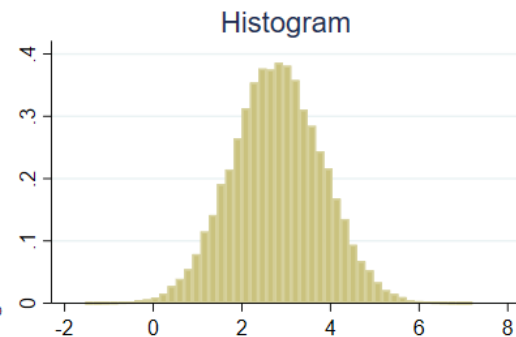
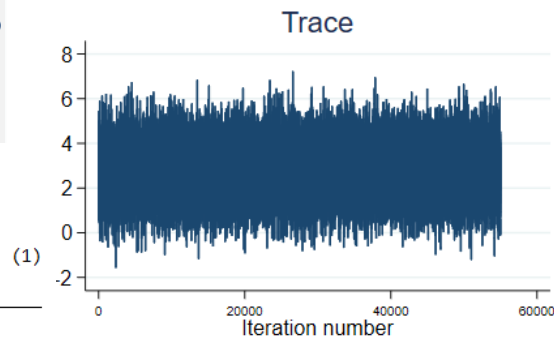
Bayesian linear regression

Random-walk Metropolis-Hastings sampling

MCMC iterations =	79,996
Burn-in =	5,000
MCMC sample size =	15,000
Number of obs =	251
Acceptance rate =	.3546
Efficiency: min =	.52
avg =	.6227
max =	.8183

Log marginal likelihood = -881.10404

Sự khác biệt bodyfat giữa hai nhóm trên 50 và dưới 50 tuổi



bayesstats summary

```
. bayesstats summary
```

Posterior summary statistics

MCMC sample size = 15,000

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred. Interval]	
bodyfat						
age_group1	2.804369	1.035819	.011621	2.803168	.7660664	4.824244
_cons	17.95994	.5975077	.006765	17.95501	16.80603	19.11968
sigma2	58.54031	5.280121	.047658	58.29911	49.05023	69.73186

Trung bình sự khác biệt về % mỡ cơ thể của nhóm trên 50 tuổi so với nhóm dưới 50 tuổi là 2.8, 95%Credible interval = 0.77 – 4.82

Kiểm định giả thuyết

- ▶ **[Giả sử]** trong thực hành lâm sàng, sự khác biệt % mỡ cơ thể từ 0 - 0.5 % là không có ý nghĩa
- **ROPE = [0 – 0.5]**
- **H1:** sự khác biệt giữa 2 nhóm > 0.5
- **H2:** sự khác biệt giữa 2 nhóm < 0.5
- **BF [H₁:H₂]**

▶ 0.35% pp hậu định nằm dưới ROPE

bayestest interval {bodyfat: age_group1}, lower(0) upper(0.5)

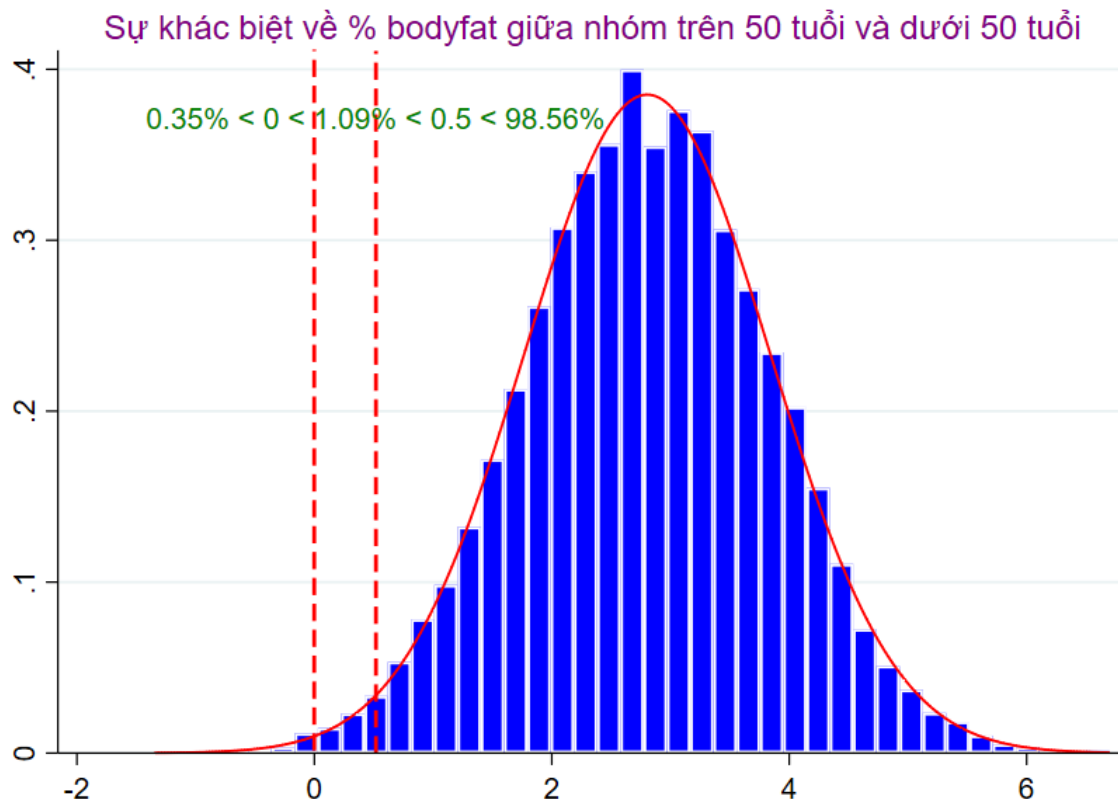
▶ 1.09% nằm trong ROPE

▶ 98.56% nằm trên ROPE

▶ $BF = 98.56 / 1.44 = 68.44$

→ Bằng chứng mạnh cho thấy TB % bodyfat nhóm trên 50 tuổi cao hơn trên 0.5% so với nhóm dưới 50 tuổi (BF)

Nhóm trên 50 tuổi có TB % bodyfat cao hơn nhóm dưới 50 tuổi có ý nghĩa trên thực hành lâm sàng (ROPE)

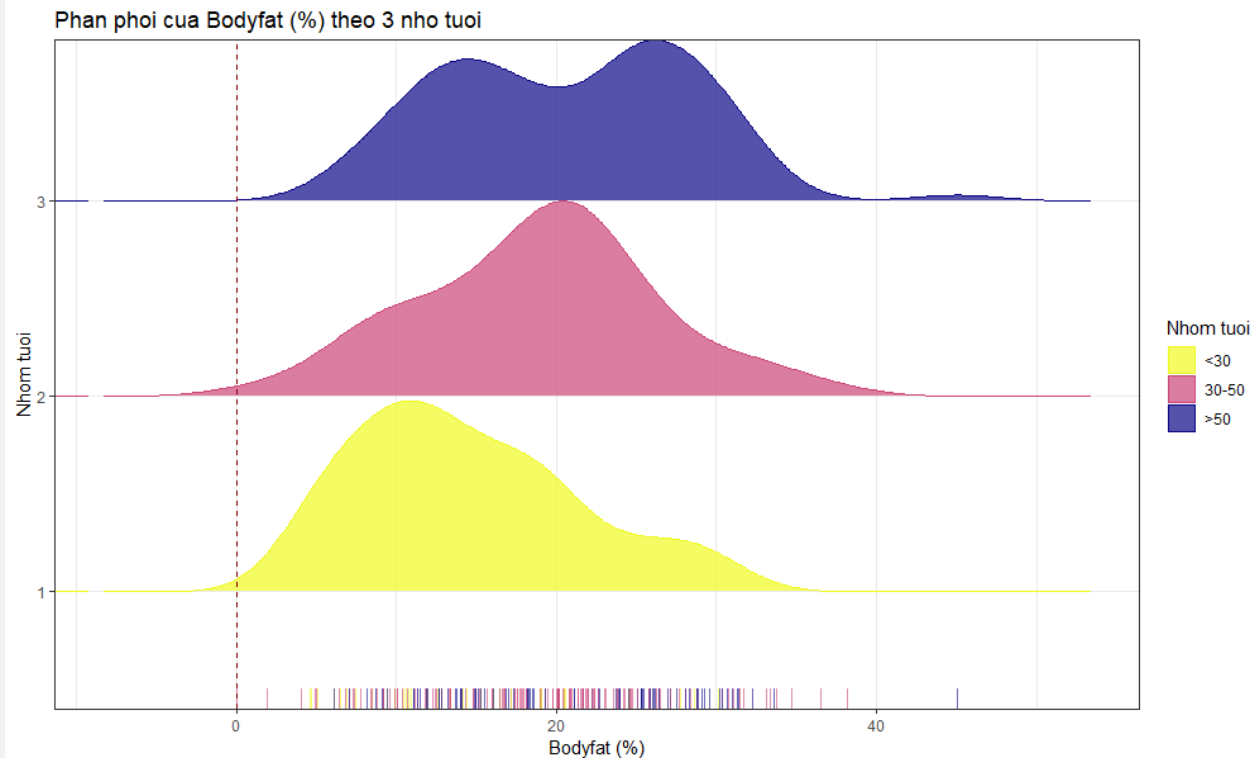


ANOVA

tạo nhóm tuổi 2

recode age (min/29.99 = 1) (30/49.99 = 2) (50/max = 3), gen(age_group2)

group	n
< 30	36
30 - 50	132
≥ 50	83



Frequentist ANOVA

oneway bodyfat age_group2, tab

. oneway bodyfat age_group2, tab					
RECODE of age (yrs)	Summary of %		Freq.		
	Mean	Std. Dev.			
1	14.422222	6.9497014	36		
2	18.923485	7.4696317	132		
3	20.76747	7.7338046	83		
Total	18.887649	7.7241211	251		
Analysis of Variance					
Source	SS	df	MS	F	Prob > F
Between groups	1011.31013	2	505.655063	9.02	0.0002
Within groups	13904.2016	248	56.065329		
Total	14915.5117	250	59.6620469		
Bartlett's test for equal variances: chi2(2) = 0.5433 Prob>chi2 = 0.762					

qui anova bodyfat age_group2
pwcompare age_group2 ,
mcompare(tukey) effects

reg bodyfat i. age_group2

Pairwise comparisons of marginal linear predictions

Margins : asbalanced

	Number of Comparisons
age_group2	3

post-hoc của ANOVA theo Tukey

	Contrast	Std. Err.	Tukey t	P> t	Tukey [95% Conf. Interval]
age_group2					
2 vs 1	4.501263	1.407873	3.20	0.004	1.181676 7.820849
3 vs 1	6.345248	1.494274	4.25	0.000	2.821939 9.868557
3 vs 2	1.843985	1.048915	1.76	0.186	-.6292238 4.317194

. reg bodyfat i.age_group2						
Source	SS	df	MS	Number of obs	=	251
Model	1011.31013	2	505.655063	F(2, 248)	=	9.02
Residual	13904.2016	248	56.065329	Prob > F	=	0.0002
Total	14915.5117	250	59.6620469	R-squared	=	0.0678
				Adj R-squared	=	0.0603
				Root MSE	=	7.4877

bodyfat	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age_group2						
2	4.501263	1.407873	3.20	0.002	1.728349	7.274176
3	6.345248	1.494274	4.25	0.000	3.402161	9.288334
_cons	14.42222	1.247946	11.56	0.000	11.9643	16.88015

Sử dụng hồi quy tuyến tính theo frequentist không so sánh trực tiếp được nhóm 2 với 3

Bayes ANOVA

bayes, mcmcsize(15000) burnin(5000) thinning(5): reg bodyfat i.age_group2

	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred. Interval]	
bodyfat						
age_group2						
2	4.502771	1.403737	.019137	4.493358	1.753398	7.266821
3	6.32463	1.484915	.019321	6.322721	3.439987	9.260628
_cons	14.43133	1.241195	.015841	14.43495	11.9848	16.85628
sigma2	56.47605	5.102202	.046836	56.14959	47.3546	67.13833

Note: Default priors are used for model parameters.

Không có nhóm 3
(>50 tuổi) vs. nhóm
2 (30-50 tuổi) ??

Trong bayesian, chuỗi MCMC chứa đựng toàn bộ thông tin của mô hình, và mỗi tham số có riêng 1 chuỗi. Chuỗi 3 chứa thông tin nhóm >50 vs < 30, chuỗi 2 chứa thông tin nhóm 30-50 vs < 30

→ Lấy thông tin nhóm 3 so với nhóm 2 (>50 vs 30-50) bằng cách lấy chuỗi 3 – chuỗi 2 !!!

Trace plot

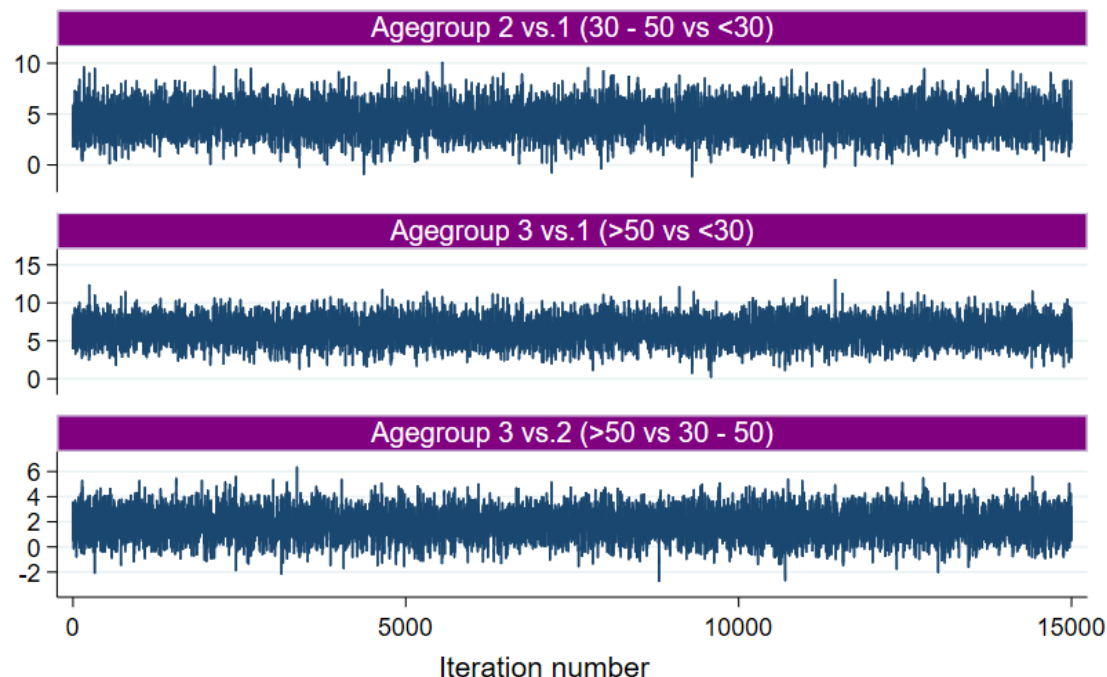
```
bayesgraph trace {bodyfat: 2.age_group2} {bodyfat: 3.age_group2}  
(age3_2: {bodyfat: 3.age_group2} - {bodyfat: 2.age_group2}), byparm
```

tinh effect sample size

```
bayesstats ess {bodyfat: 2.age_group2}  
{bodyfat: 3.age_group2} (age3_2:  
{bodyfat: 3.age_group2} - {bodyfat:  
2.age_group2})
```

	ESS	Corr. time	Efficiency
bodyfat			
age_group2			
2	5380.48	2.79	0.3587
3	5906.71	2.54	0.3938
age3_2	6062.48	2.47	0.4042

Trace plots



Graphs by parameter
age3_2: {bodyfat:3.age_group2} - {bodyfat:2.age_group2}

bayesstats summary {bodyfat: 2.age_group2} {bodyfat: 3.age_group2}
 (age3_2: {bodyfat: 3.age_group2} - {bodyfat: 2.age_group2})

Posterior summary statistics

MCMC sample size = 15,000

age3_2 : {bodyfat:3.age_group2} - {bodyfat:2.age_group2}

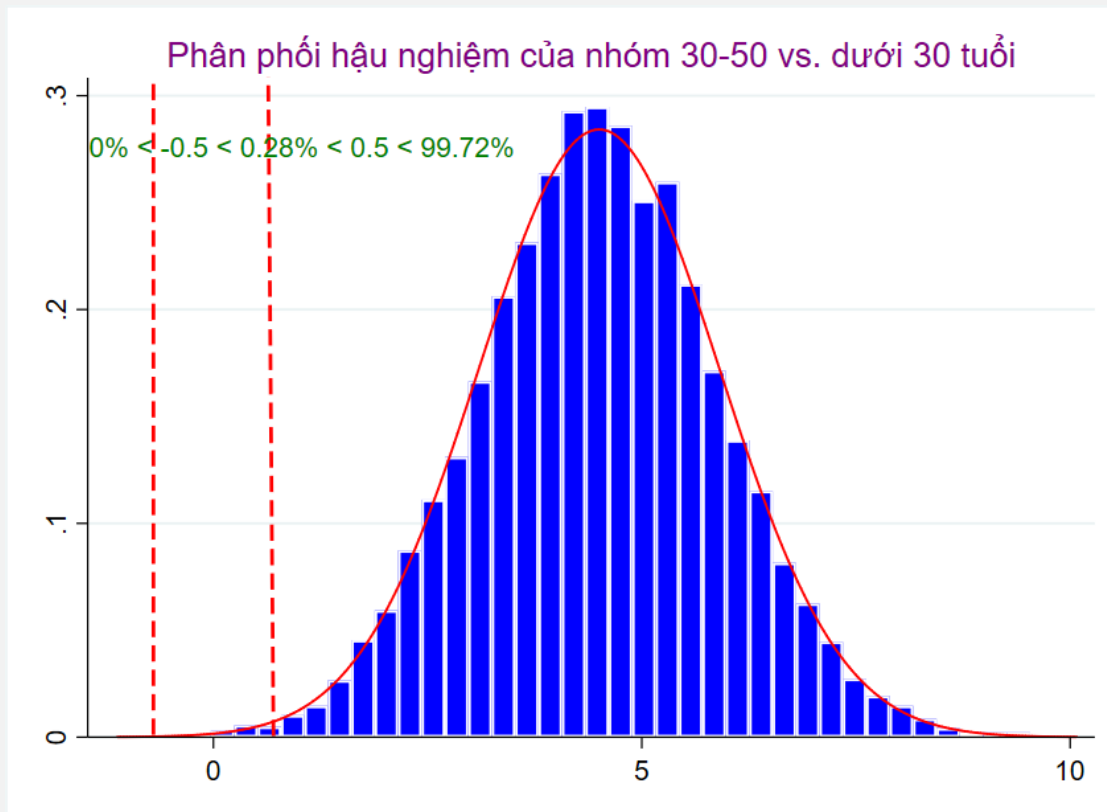
	Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred. Interval]	
bodyfat						
age_group2						
2	4.502771	1.403737	.019137	4.493358	1.753398	7.266821
3	6.32463	1.484915	.019321	6.322721	3.439987	9.260628
age3_2	1.821859	1.04413	.01341	1.826822	-.2584381	3.870804

Trung bình sự khác biệt về % mỡ cơ thể của nhóm:

- 30-50 vs. < 30 tuổi: 4.5 (95% credible interval = 1.75 – 7.27)
- > 50 vs. < 30 tuổi: 6.3 (95% CI = 3.44 – 9.26)
- > 50 vs. 30-50 tuổi: 1.8 (95% CI = -0.26 – 3.87)

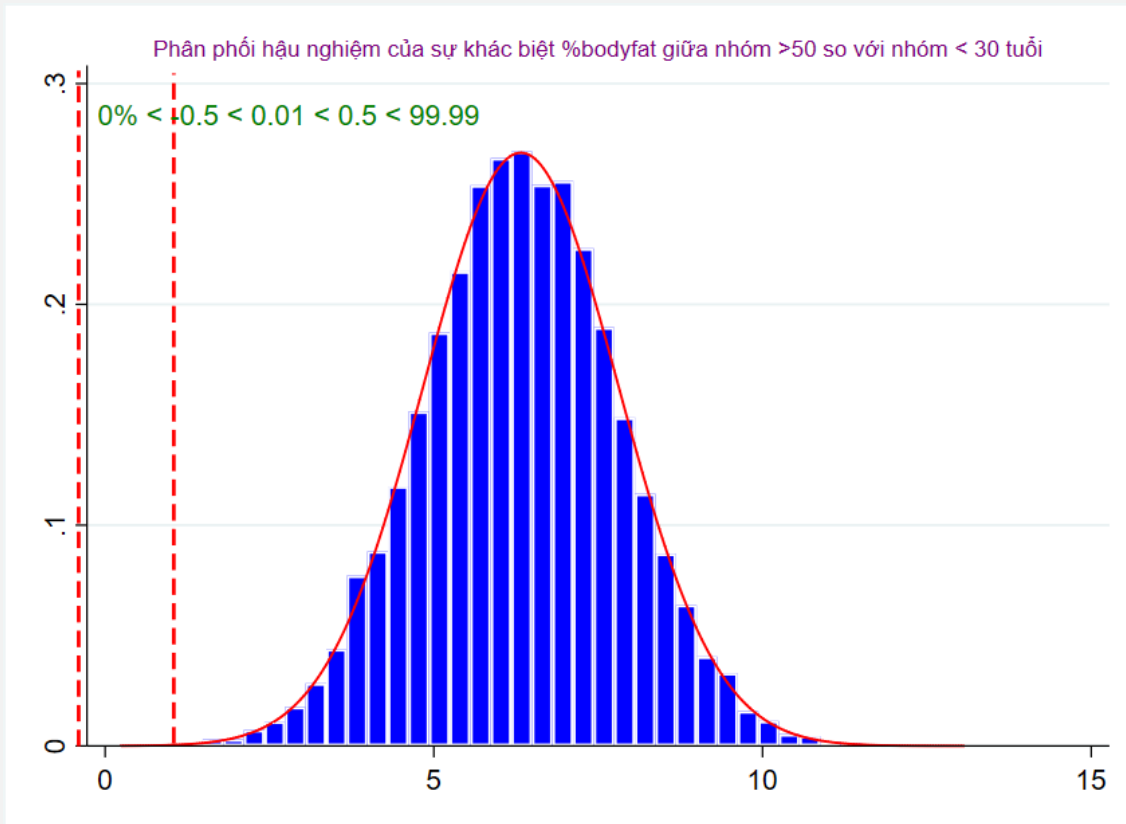
Kiểm định giả thuyết

bayestest interval {bodyfat: 2.age_group2}, lower(-0.5) upper(0.5)



Kiểm định giả thuyết (2)

bayestest interval {bodyfat: 3.age_group2}, lower(-0.5) upper(0.5)



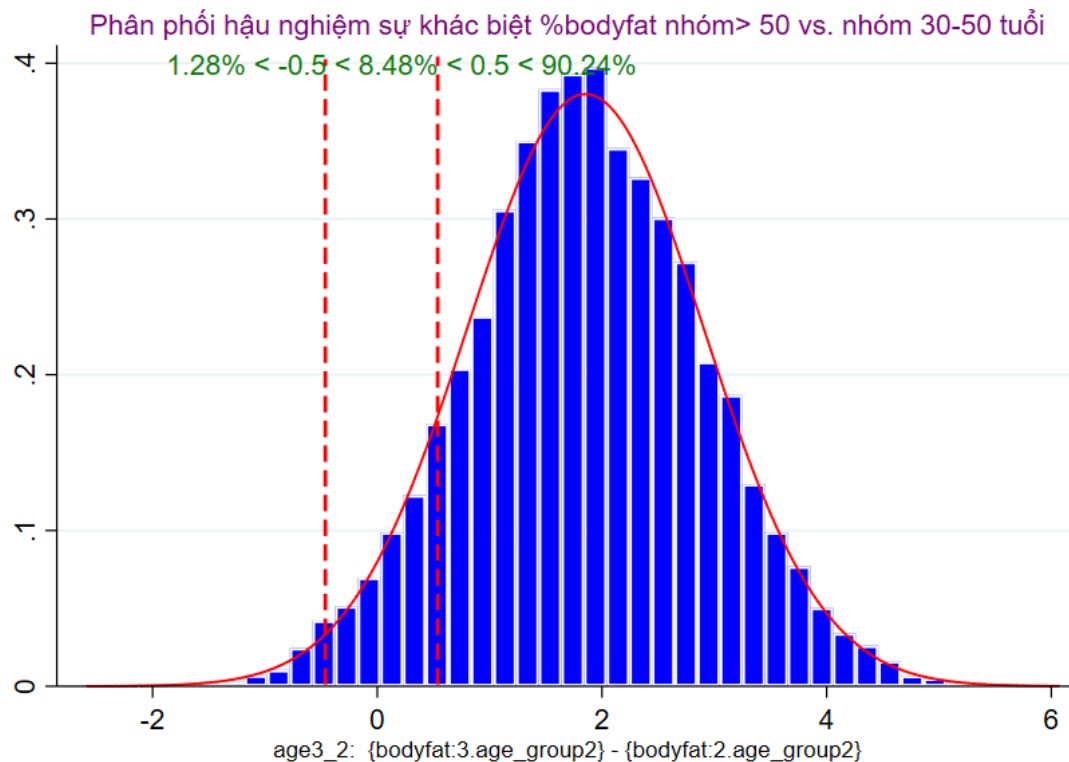
Kiểm định giả thuyết (3)

bayestest interval (age3_2: {bodyfat: 3.age_group2} - {bodyfat: 2.age_group2}),
lower(-0.5) upper(0.5)

$BF_1[>0.5 : < 0.5] = 90.24/9.76 = 9.25$
→ bằng chứng yếu ủng hộ $H(>0.5)$

bayestest interval (age3_2:
{bodyfat: 3.age_group2} -
{bodyfat: 2.age_group2}), lower(0)
95.99 %

$BF_1[>0 : < 0] = 95.99/4.01 = 23.94$
→ bằng chứng mạnh ủng hộ $H(>0)$



Tóm tắt

- ▶ Bayes t test và ANOVA có thể thực hiện thông qua hồi quy tuyến tính Bayes
- ▶ Tăng cỡ mẫu MCMC khi MCMC chưa hội tụ
- ▶ Thông tin chứa trong chuỗi MCMC có thể được sử dụng như một biến số thực sự (cộng, trừ, nhân, chia, chuyển dạng ...)

Thank you !