tabular

Uvod v Bayesovsko statistiko

Gregor Gorjanc

Univerza v Edinburgh-u / Univerza v Ljubljani

2025-05-27

Sodobni statistični pristopi na doktorskem študiju Statistike

Ljubljana

Okvirni vozni red

Uvod

Osnove verjetnosti

Verjetje

Bayesovski pristop

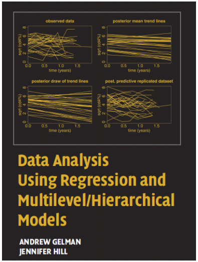
Metode MCMC

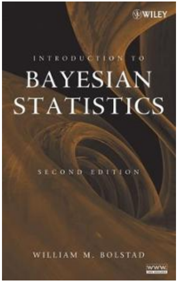
Splošni Bayes programi/paketi (demo)

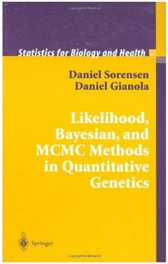
0. Uvod

- Zgodovina (Laplace in Bayes)
- Moje izkušnje (statistična/kvantitativna genetika)
- Poimenovanje?
  - Bayesova
  - Bayesovska
- Bayesian = "probabilistic modelling"

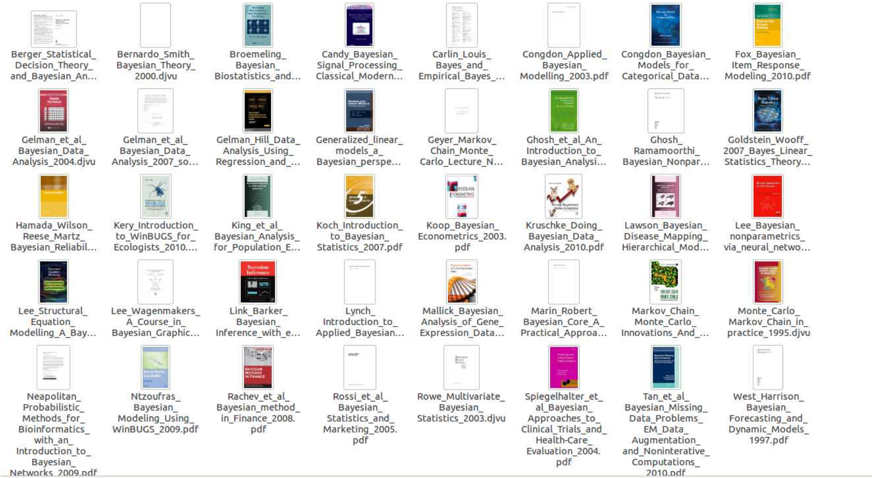
Nekaj literature (moj študij)







Literature je več kot dovolj



Priporočam

Richard McElreath

Anthropology, Evolutionary Ecology, Bayesian Data Analysis

HOME PUBLICATIONS SOFTWARE SR2 FOR INSTRUCTORS **STATISTICAL RETHINKING** TALKS TEACHING

Statistical Rethinking

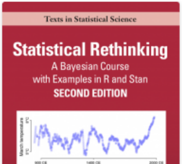
A Bayesian Course with Examples in R and Stan (& PyMC3 & brms & Julia too, see links below)

Most recent set of free lectures: Statistical Rethinking 2023

Second Edition

The second edition is now out in print. Publisher information on the CRC Press page. For more detail about what is new, look here.

Materials



2nd Edition

- Book: CRC Press
- Book sample: Chapters 1 and 2 (2MB PDF)
- Lectures and slides:

FOLLOW ME ON TWITTER

Tweets from @rlmcelreath

Richard McElr...
@. · Aug 16, 2023

I used to teach game theo
both undergrad & phd lev
One game I would do at s
is version of Keynesian
beauty contest: everyone
picks a number 0-100,
person closest to 2/3 of
average wins. Nash is 0. E
anyone choosing 0 loses,
the class aren't (yet) gam
theorists. >

100 6K

Richard McElr...
@ · Nov 30, 2022

Priporočam (naslednja stopnja)

Michael Betancourt

Once and future physicist masquerading as an applied statistician.

New York, NY

Email

Twitter

GitHub

Part II: Modeling and Inference

Modeling and Inference
Generative Modeling
Prior Modeling
Introduction to Stan
Identifiability and Degeneracies
Principled Model Building Workflow

Part III: Modeling Techniques

Hierarchical Modeling
Factor Modeling
Modeling Sparsity
Modeling Ordinal Outcomes
Survival Modeling
Modeling Consensus HTML PDF
Variate-Covariate Modeling
Modeling Functional Behavior I: Linearized Models
Modeling Functional Behavior II: General Linearized Models
Underdetermined Linearized Regression
The QR Decompostion for Linearized Regression
Modeling Functional Behavior III: Gaussian Processes
Modeling Functional Behavior IV: Stochastic Differential Equations
Modeling Functional Behavior V: The Brownian Bridge

1. Osnove verjetnosti

Verjetnost

- Mečemo kocko

- Dva pogleda na verjetnost:
  - verjetnost dogodka, da pade 6 pik: $\mathrm{Pr}(6\,pik) = 1/6$
  - mečemo kocko in spremljamo frekvenco (= verjetnost) dogodkov

- Porazdelitev verjetnosti dogodkov, $\sum_i \mathrm{Pr}\,(i\;pik) = 1$:

Št. pik	1	2	3	4	5	6
Verjetnost	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

Vaja

- Na namizju poiščite ikono za program R in zaženite program
- Uporabite spodnjo kodo

```
## Žepni računalnik
1 + 1
log(1)

## Mečimo kocko 10-krat
sample(x=1:6, size=10, replace=TRUE)
sample(x=1:6, size=10, replace=TRUE)

## ... 10.000-krat
x <- sample(x=1:6, size=10000, replace=TRUE)
t <- table(x)
p <- prop.table(x=t)
p
```

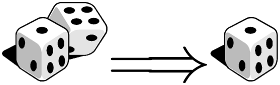
Skupna verjetnost (ang. joint prob.)

- Mečemo **dve** kocki
- Ali je št. pik na prvi kocki povezano s št. pik na drugi kocki? NE.
  –> Dogodka sta torej neodvisna.
- Verjetnost skupnega dogodka, da pade dvakrat 6 pik:
  $\Pr(6\,pik, 6\,pik) = \Pr(6\,pik) \times \Pr(6\,pik) = 1/6 \times 1/6 = 1/36$
- Porazdelitev verjetnosti dogodkov, $\sum_{i,j} \Pr(i\,pik, j\,pik) = 1$:

Št. pik	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

Robna verjetnost (ang. marginal prob.)

- Mečemo **dve** kocki



- Porazdelitev verjetnosti za število pik 1, 2, ... 6:

Št. pik	1	2	3	4	5	6	Skupaj
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
Skupaj	1/6	1/6	1/6	1/6	1/6	1/6	1

$$\Pr(A) = \sum_B \Pr(A, B) \qquad p(a) = \int p(a, b)\, db$$

Vaja

- Uporabite spodnjo kodo

```
## Skupna verjetnost
x1 <- sample(x=1:6, size=10000, replace=TRUE)
x2 <- sample(x=1:6, size=10000, replace=TRUE)
t <- table(x1, x2)
p <- prop.table(x=t)
p

## Robna verjetnost
rowSums(x=p)
colSums(x=p)
```

Skupna verjetnost II

- Mečemo **eno** kocko
- Ali je št. pik na kocki povezano z naravo števila (sodo ali liho)? DA. –> Dogodka sta torej odvisna.
- Verjetnost dogodka, da je število sodo in da pade 6 pik:

$$
\begin{aligned}
Pr(6\,pik, sodo\,št.) &= Pr(6\,pik|sodo\,št.) \times Pr(sodo\,št.) \\
Pr(6\,pik, sodo\,št.) &= 1/6 \\
Pr(sodo\,št.) &= 1/2 \\
Pr(6\,pik|sodo\,št.) &= \frac{Pr(6\,pik, sodo\,št.)}{Pr(sodo\,št.)} = \frac{1/6}{1/2} = 1/3
\end{aligned}
$$

- $Pr(6\,pik|sodo\,št.)$ –> Pogojna verjetnost (ang. conditional prob.)

Pogojna verjetnost (ang. conditional prob.)

- Verjetnost dogodka, da pade 6 pik, če vemo, da je padlo sodo število: $\Pr(6\,pik|sodo\,\check{s}tevilo) = {}^1\!/_3$
- Porazdelitev verjetnosti dogodkov, $\sum_i \Pr\left(i\,pik|liho\,\check{s}tevilo\right) = 1$:

Liho število

Št. pik	1	2	3	4	5	6
Verjetnost	$1/3$	0	$1/3$	0	$1/3$	0

- Porazdelitev verjetnosti dogodkov, $\sum_i \Pr\left(i\,pik|sodo\,\check{s}tevilo\right) = 1$:

Sodo število

Št. pik	1	2	3	4	5	6
Verjetnost	0	$1/3$	0	$1/3$	0	$1/3$

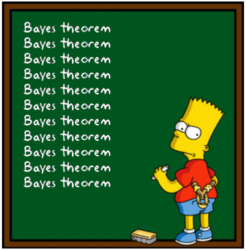
Bayesov teorem

$$
\begin{aligned}
\Pr(A,B) &= \Pr(A|B) \times \Pr(B) = \Pr(B|A) \times \Pr(A) \\
\Pr(A|B) &= \frac{\Pr(B|A) \times \Pr(A)}{\Pr(B)} \\
&= \frac{\Pr(B|A) \times \Pr(A)}{\sum_A \Pr(B|A) \times \Pr(A)} \\
\Pr(A|B) &\propto \Pr(B|A) \times \Pr(A)
\end{aligned}
$$

Vaja

Frekvenca obolelih v populaciji znaša 0.008. Obstaja test za določitev obolelosti, ki je lažno pozitiven v 10 % in lažno negativen v 5 %.

- Koliko znaša verjetnost, da ima posameznik bolezen, če je rezultat testa zanj pozitiven?
- Kakšen je vpliv predhodne (apriorne) informacije?
- Koliko znaša verjetnost, da ima posameznik bolezen, če je rezultat ponovljenega testa zanj ponovno pozitiven?



Vaja (rešitev)

Dogodki:

- status: zdrav, bolan
- test: negativen, pozitiven

Predhodna verjetnost obolelosti brez testa:

$$\Pr\left(bolan\right) = 0.008$$

Test:

	Status	
Test	Zdrav	Bolan
Negativen	$\Pr\left(negativen \mid zdrav\right) = 0.90$	$\Pr\left(negativen \mid bolan\right) = 0.05$
Pozitiven	$\Pr\left(pozitiven \mid zdrav\right) = 0.10$	$\Pr\left(pozitiven \mid bolan\right) = 0.95$
Skupaj	1	1

Vaja (rešitev) II

- Koliko znaša verjetnost, da ima naključni posameznik bolezen, če je rezultat testa zanj pozitiven?

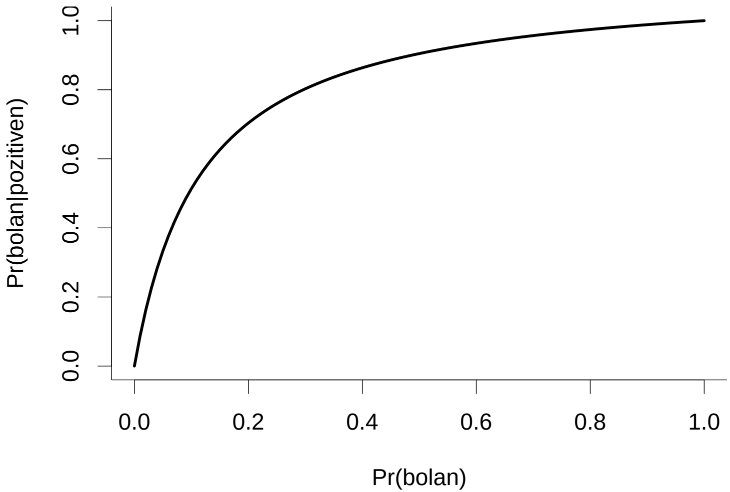
$$
\begin{aligned}
\Pr(bolan|pozitiven) &= \frac{\Pr(pozitiven|bolan)\times\Pr(bolan)}{\Pr(pozitiven)} \\
&= \frac{\Pr(pozitiven|bolan)\times\Pr(bolan)}{\Pr(pozitiven|bolan)\times\Pr(bolan)+\Pr(pozitiven|zdrav)\times\Pr(zdrav)} \\
&= \frac{0.95\times 0.008}{0.95\times 0.008+0.1\times 0.992} = \frac{0.0076}{0.1068} = 0.071
\end{aligned}
$$

$$
\begin{aligned}
\Pr(zdrav|pozitiven) &= \frac{\Pr(pozitiven|zdrav)\times\Pr(zdrav)}{\Pr(pozitiven)} \\
&= \frac{\Pr(pozitiven|zdrav)\times\Pr(zdrav)}{\Pr(pozitiven|bolan)\times\Pr(bolan)+\Pr(pozitiven|zdrav)\times\Pr(zdrav)} \\
&= \frac{0.10\times 0.992}{0.95\times 0.008+0.1\times 0.992} = \frac{0.0992}{0.1068} = 0.929
\end{aligned}
$$

$$
\begin{aligned}
\Pr(bolan|pozitiven)+\Pr(zdrav|pozitiven) &= 0.071+0.929=1 \\
\Pr(pozitiven) &= 0.0076+0.0992=0.1068
\end{aligned}
$$

Vaja (rešitev) III

- Kakšen je vpliv predhodne (apriorne) informacije?



Vaja (rešitev) IV

- Koliko znaša verjetnost, da ima naključni posameznik bolezen, če je rezultat ponovljenega testa zanj pozitiven?

$$
\begin{aligned}
\Pr(\mathit{bolan}|\mathit{pozitiven}) &= 0.071 \Longrightarrow \Pr(\mathit{bolan}) \\
\Pr(\mathit{bolan}|2x\ \mathit{pozitiven}) &\Longrightarrow \Pr(\mathit{bolan}|\mathit{pozitiven}) \\
\Pr(\mathit{bolan}|\mathit{pozitiven}) &= \frac{\Pr(\mathit{pozitiven}|\mathit{bolan}) \times \Pr(\mathit{bolan})}{\Pr(\mathit{pozitiven})} \\
&\frac{\Pr(\mathit{pozitiven}|\mathit{bolan}) \times \Pr(\mathit{bolan})}{\Pr(\mathit{pozitiven}|\mathit{bolan}) \times \Pr(\mathit{bolan}) + \Pr(\mathit{pozitiven}|\mathit{zdrav}) \times \Pr(\mathit{zdra}} \\
&= \frac{0.95 \times 0.071}{0.95 \times 0.071 + 0.1 \times 0.929} = \frac{0.06745}{0.16035} = 0.421
\end{aligned}
$$

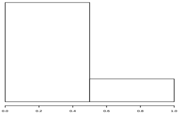
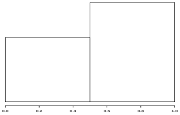
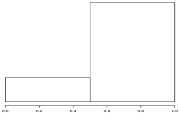
Vprašanja?

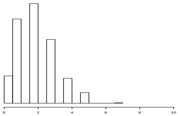
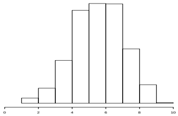
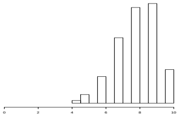
Porazdelitve

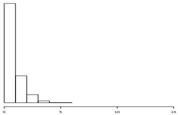
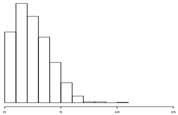
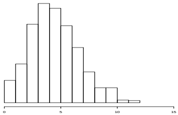
Prikaz nekaj pogosto uporabljenih porazdelitev v statistiki in njihovih lastnosti

Porazdelitve - diskretne

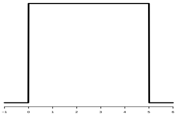
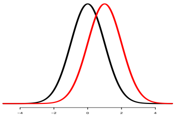
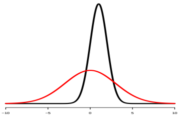
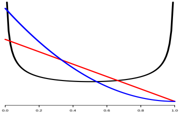
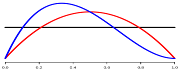
Porazdelitev	Zapis	$\Pr(x)$	$E(x)$	$Var(x)$
Bernoullijeva	$B(1,\theta)$	$\theta^{x}(1-\theta)^{1-x}$	θ	$\theta(1-\theta)$
Binomska	$B(n,\theta)$	$\binom{n}{x}\theta^{x}(1-\theta)^{n-x}$	$n\theta$	$n\theta(1-\theta)$
Poissonova	$P(\lambda)$	$\frac{\lambda^{x}\exp(-\lambda)}{x!}$	λ	λ

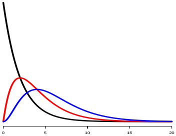
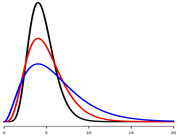
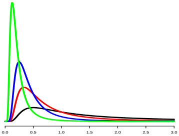




Porazdelitve - zvezne

Porazdelitev	Zapis	$p(x)$	$E(x)$	$Var(x)$
Enakomerna	$U(\alpha, \beta)$	$\frac{1}{\beta-\alpha}$	$\frac{\alpha+\beta}{2}$	$\frac{(\beta-\alpha)^2}{12}$
				
Gaussova	$N(\mu, \sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	μ	σ^2
				
Beta[1]	$Be(\alpha, \beta)$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
				

[1] $\Gamma(x)$ = gama funkcija

Porazdelitve - zvezne II

Porazdelitev	Zapis	$p(x)$	$E(x)$	$Var(x)$
Gama[2]	$G(\alpha, \beta)$	$\beta^{\alpha} \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x)$	α/β	α/β^2
	$G(\alpha, \theta = 1/\beta)$	$\frac{1}{\theta^{\alpha}} \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\theta}\right)$	$\alpha\theta$	$\alpha\theta^2$
				
Inverzna gama	$IG(\alpha, \beta)$	$\beta^{\alpha} \frac{1}{\Gamma(\alpha)} (x)^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right)$	$\frac{\beta}{\alpha-1}$	$\frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$
				

[2] $\Gamma(x)$ = gama funkcija

Vaja

Razložimo gostoto verjetnosti za binomsko porazdelitev:

$$\Pr(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

Vaja (rešitev)

Gostota verjetnosti za Bernoullijevo porazdelitev $B(1, \theta)$:

$$\Pr(x) = \theta^{x} (1 - \theta)^{1-x}$$

- zaloga vrednosti je $[0, 1]$
  (0=cifra ali 1=glava pri metu kovanca, 0=zdrav ali 1=bolan, ...)
- verjetnost, da se zgodi dogodek 1 je enaka: $\Pr(x = 1) = \theta$
  ($\theta = 1/2$ pri poštenem kovancu)
- splošni zapis (glej zgoraj):
  - $\Pr(x = 0) = \theta^{0} (1 - \theta)^{1-0} = 1 (1 - \theta)^{1} = 1 - \theta$
  - $\Pr(x = 1) = \theta^{1} (1 - \theta)^{1-1} = \theta^{1} (1 - \theta)^{0} = \theta$

Če spremljamo več zaporednih dogodkov iz $B(1, \theta)$ je število možnih kombinacij dogodkov enako $\binom{n}{x}$. Tako je gostota verjetnosti za binomsko porazdelitev $B(n, \theta)$:

$$\Pr(x) = \binom{n}{x} \theta^{x} (1 - \theta)^{n-x}$$

Vaja

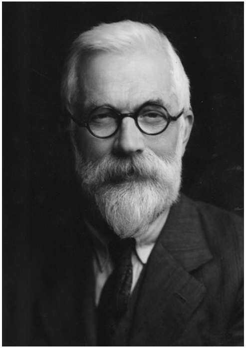
- Uporabite spodnjo kodo

```
n <- 10; theta <- 0.5
## Verjetnost, da dobimo x=1, x=5 ali x=9?
x <- c(1, 5, 9)
choose(n, x)*theta^x*(1-theta)^(n-x)

## Vzorčimo vrednosti iz Binomske porazdelitve
x <- rbinom(n=100, size=n, prob=theta)

## Pregled vzorca
table(x); hist(x)

## Izračun povprečja in variance
n*theta; n*theta*(1-theta)  ## pričakovano
mean(x); var(x)             ## opaženo
```

- Povečajte vzorec s 100 na 100.000 in ponovite. Kaj opazite?
- DN Preglejte dokumentacijo: ?rpois, ?runif, ?rnorm, ...

Vprašanja?
(pavza)

2. Verjetje
(ang. likelihood)



Primer

Obiščemo 7 neodvisnih lokacij in opravimo po 5 neodvisnih zanesljivih testov za določitev obolelosti. Rezultati so:

Lokacija	1	2	3	4	5	6	7	Skupaj
Št. pozitivnih	2	0	1	1	1	1	2	8

- Koliko znaša delež obolelih v populaciji?
  - Število pozitivnih testov na lokaciji i: y_i (poznamo)
  - Delež obolelih: θ (ne poznamo)
  - Predpostavljeni model: $\Pr(y_i|n=5,\theta) \sim B(5,\theta)$
  - Princip verjetja (Fisher): "**Poiščimo takšno vrednost parametrov** (θ), **da bodo zbrani podatki** ($\mathbf{y}$), **kar se da najbolj verjetni.**"
  - Gostota verjetnosti $\Pr(\mathbf{y}|n=5,\theta)$ in verjetje $L(\mathbf{y}|n=5,\theta)$:

$$\Pr(\mathbf{y}|n=5,\theta) = \prod_{i=1}^{7} \binom{n}{y_i} \theta^{y_i} (1-\theta)^{n-y_i} \qquad L(\mathbf{y}|n=5,\theta) = \prod_{i=1}^{7} \binom{n}{y_i} \theta^{y_i} (1-\theta)^{n-y_i}$$

Primer II

Iščemo vrednost parametra (θ), da bodo zbrani podatki ($\mathbf{y}$), kar se da najbolj verjetni. –> Poiskati moramo maksimum funkcije $L\left(y|n,\theta\right)$.
Standarden postopek:

1. "Računalniki ne marajo zelo majhnih števil"
   $\log\left(L\left(y|n,\theta\right)\right) = l\left(y|n,\theta\right)$
2. "Zmasiramo" izraz $l\left(y|n,\theta\right)$
3. Odvajamo in odvod enačimo z nič
4. Rešimo enačbo –> izračunamo oceno parametra $\left(\hat{\theta}_{MLE}\right)$ (direktno ali z numeričnimi metodami)
5. Varianco ocene izračunamo na osnovi informacijske matrike $\left(Var\left(\hat{\theta}_{MLE}\right) = \left(-E\left(\frac{\partial^2 \log(L(y|n,\theta))}{\partial\theta\partial\theta'}\right)\right)^{-1}\right)$

MLE = ang. Maximum Likelihood Estimate, slo. cenilka po metodi največjega verjetja

Primer II

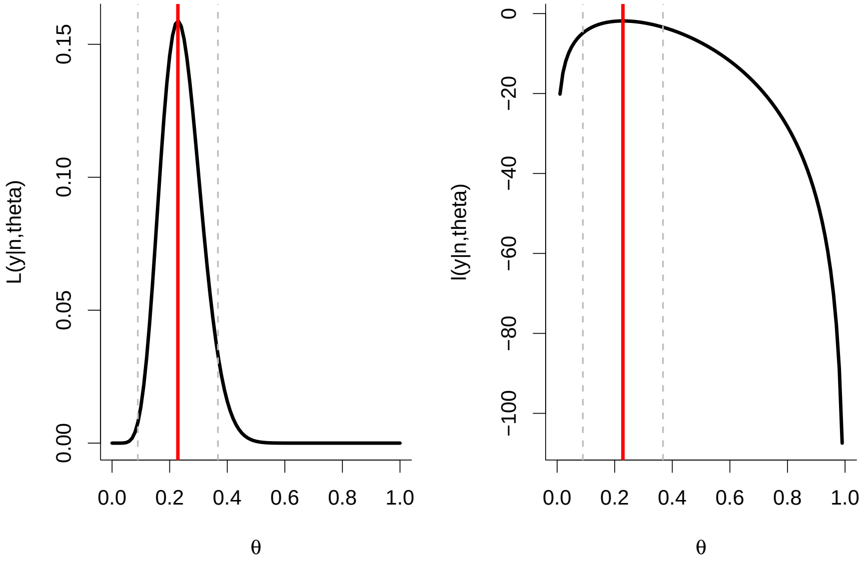
Ocena deleža obolelih glede na zbrane podatke:

$$
\begin{aligned}
\hat{\theta}_{MLE} &= \frac{y}{n} \\
&= \frac{8}{35} = 0.23
\end{aligned}
$$

Standardna napaka ocene:

$$
\begin{aligned}
SE\left(\hat{\theta}_{MLE}\right) &= \sqrt{\frac{-1}{-\frac{y}{p^2} - \frac{n-y}{(1-p)^2}}} \\
&= \sqrt{\frac{-1}{-\frac{8}{0.23^2} - \frac{35-8}{(1-0.23)^2}}} = 0.07
\end{aligned}
$$

Primer III



L(y|n,theta)
0.00
0.05
0.10
0.15
0.0
0.2
0.4
0.6
0.8
1.0
θ
l(y|n,theta)
0
−20
−40
−60
−80
−100
0.0
0.2
0.4
0.6
0.8
1.0
θ

Vaja

- Uporabite spodnjo kodo

```
n <- 35
y <- 8
p <- seq(from=0, to=1, by=0.01)
MLE <- y/n
MLE_SE <- sqrt(-1 / (-y/MLE^2 - (n-y)/(1-MLE)^2))
MLE_CI <- MLE + c(-1.96, 0, 1.96) * MLE_SE
L <- choose(n, y)*p^y*(1-p)^(n-y)
l <- log(L)
par(mfrow=c(1, 2))
plot(L ~ p, type="l"); abline(v=MLE_CI)
plot(l ~ p, type="l"); abline(v=MLE_CI)
```

- Kaj se spremeni, če imamo vzorec velikosti 10 in 2 pozitivna rezultata ali vzorec velikosti 1000 in 230 pozitivnih rezultatov?

3. Bayesovski pristop k statistiki

Bayesov teorem

- Koliko znaša delež obolelih v populaciji?
  - Število vseh in pozitivnih testov: $n = 35, y = 8$ (poznamo)
  - Delež obolelih: θ (ne poznamo)

$$
\begin{aligned}
p(\theta|y,n) &= \frac{p(y|n,\theta) \times p(\theta)}{p(y)} \\
&= \frac{p(y|n,\theta) \times p(\theta)}{\int_\theta p(y|n,\theta) \times p(\theta)} \\
p(\theta|y,n) &\propto p(y|n,\theta) \times p(\theta)
\end{aligned}
$$

"posterior is proportional to likelihood times prior"

- Slovar:
  - $p(\theta|y,n)$ = posteriorna porazdelitev (rezultat)
  - $p(y|n,\theta)$ = verjetje (predpostavljen model + podatki)
  - $p(\theta)$ = apriorna porazdelitev (predpostavka = predhodno znanje)

Tip porazdelitve?

- Koliko znaša delež obolelih v populaciji?
  - Število vseh in pozitivnih testov: $n = 35, y = 8$ (poznamo)
  - Delež obolelih: θ (ne poznamo)

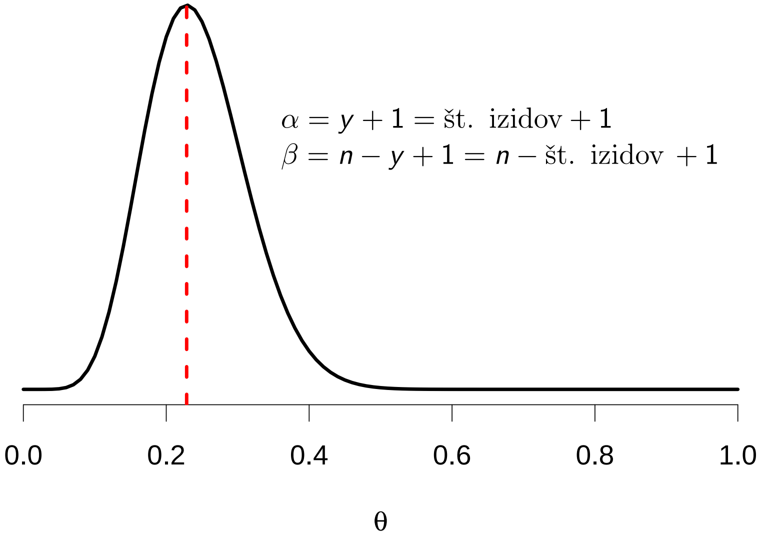
$$
\begin{aligned}
p(\theta|y,n) &\propto p(y|n,\theta) \times p(\theta) \\
&\propto \binom{n}{y} \theta^{y} (1-\theta)^{n-y} \times p(\theta) \\
&\propto \theta^{y} (1-\theta)^{n-y} \times p(\theta)
\end{aligned}
$$

"after the inspection beta distribution can be recognized"

$$x \sim Be(\alpha,\beta) \qquad p(x|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

$$p(y|n,\theta) = Be(y+1, n-y+1)$$

Tip porazdelitve $Be\left(y+1, n-y+1\right)$



Konjugirana apriorna porazdelitev

- Konjugirana apriorna porazdelitev ima enako "obliko" kot verjetje:

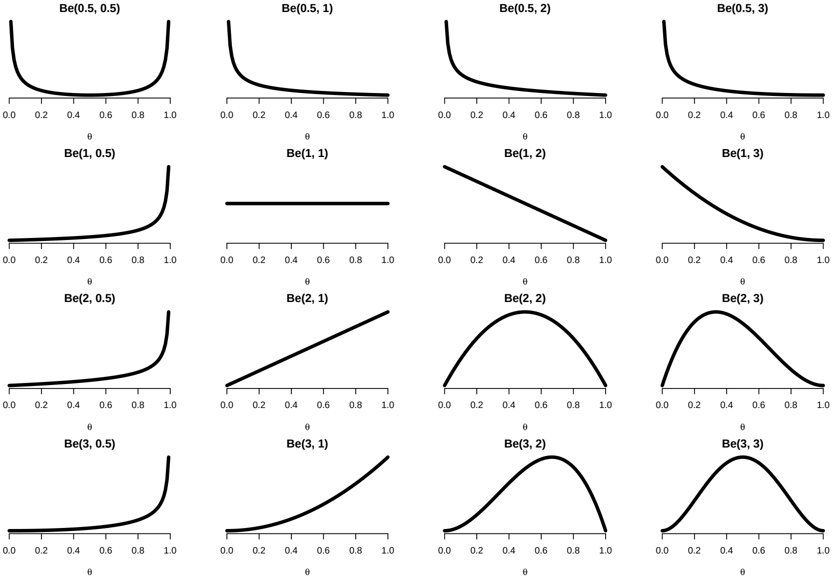
$$
\begin{aligned}
p(y|n,\theta) &= Be(y+1, n-y+1) \\
p(\theta) &= Be(\alpha,\beta) \\
p(\theta|y,n) &\propto \theta^{y}(1-\theta)^{n-y} \times \theta^{\alpha-1}(1-\theta)^{\beta-1} \\
&\propto \theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1}
\end{aligned}
$$

"after the inspection <u>beta distribution</u> can be recognized"

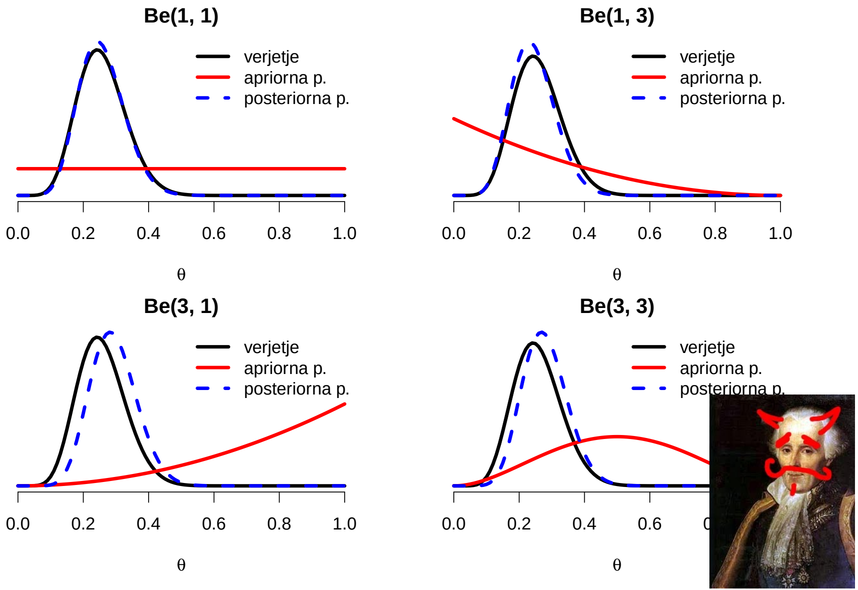
$$
p(\theta|y,n) = Be(y+1+\alpha, n-y+1+\beta)
$$

Moramo določiti parametra apriorne porazdelitve: α, β

Apriorna porazdelitev $Be(\alpha, \beta)$



Posteriorna porazdelitev $Be(y+1+\alpha, n-y+1+\beta)$



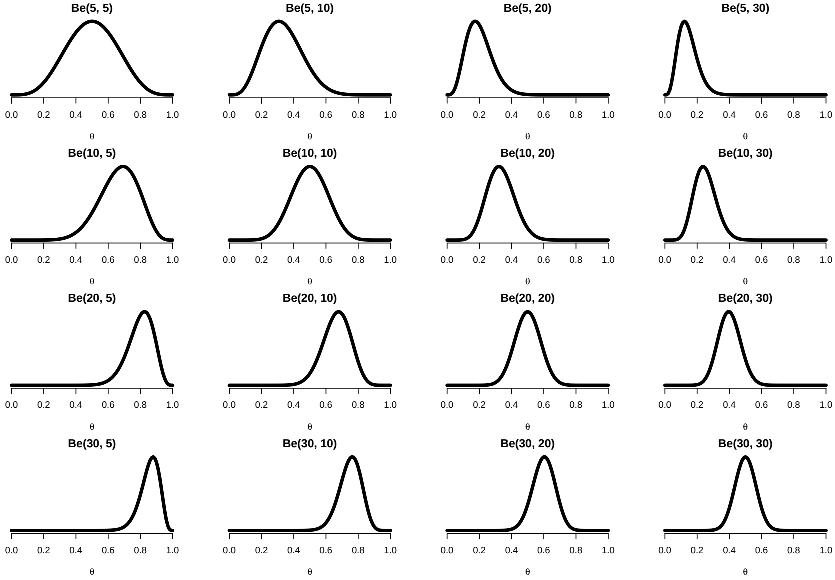
Posteriorna porazdelitev $Be\left(y+1+\alpha, n-y+1+\beta\right)$

$$E\left(x|\alpha,\beta\right)=\frac{\alpha}{\alpha+\beta} \qquad Var\left(x|\alpha,\beta\right)=\frac{\alpha\beta}{\left(\alpha+\beta\right)^2\left(\alpha+\beta+1\right)}$$

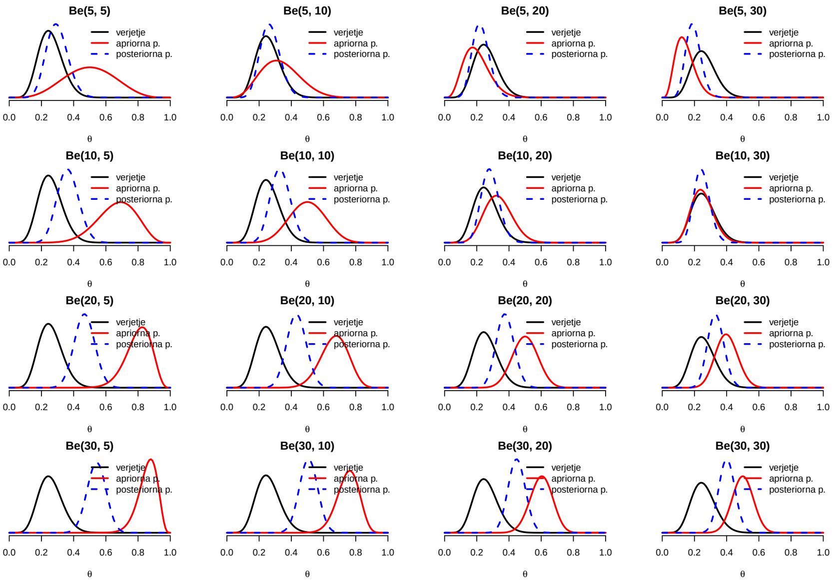
Apriorna p.	$E\left(\theta|y,n,\alpha,\beta\right)$	$SE\left(\theta|y,n,\alpha,\beta\right)$
$Be\left(1,1\right)$	0.256	0.069
$Be\left(1,3\right)$	0.243	0.066
$Be\left(3,1\right)$	0.293	0.070
$Be\left(3,3\right)$	0.279	0.068

$$\hat{\theta}_{MLE}=0.23, \quad SE\left(\hat{\theta}_{MLE}\right)=0.07$$

Apriorna porazdelitev $Be(\alpha, \beta)$



Posteriorna porazdelitev $Be\left(y+1+\alpha, n-y+1+\beta\right)$



Vaja

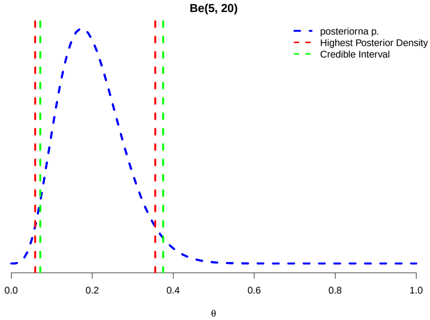
- S pomočjo spodnje kode izračunajte in narišite posteriorno porazdelitev za sledeče primere:
  - vzorec velikosti 10 in 2 pozitivna rezultata
  - vzorec velikosti 1000 in 230 pozitivnih rezultatov?

```
n <- 35; y <- 8
alpha <- 1; beta <- 10
a <- y + alpha; b <- n - y + beta
MLE <- y/n
MLE_SE <- sqrt(-1 / (-y/MLE^2 - (n-y)/(1-MLE)^2))
POS <- a / (a + b)
POS_SE <- sqrt(a * b / (((a + b)^2) * (a + b + 1)))
p <- seq(from=0, to=1, by=0.01)
dL   <- dbeta(x=p, shape1=y+1,   shape2=n-y+1)
dPRI <- dbeta(x=p, shape1=alpha, shape2=beta)
dPOS <- dbeta(x=p, shape1=a,     shape2=b)
matplot(y=cbind(dL, dPRI, dPOS), x=p, type="l",
        col=c("black", "red", "blue"))
abline(v=MLE)
```

Intervali

- Credible interval (analogija intervala zaupanja; v primeru, da poznamo posteriorno porazdelitev lahko enostavno izračunamo kvantile)
- Highest Posterior Density (HPD) (interval, ki je čim krajši in zajema določen % porazdelitve!)

$$CI\left(\theta|y,n,\alpha,\beta\right) = 0.07 - 0.38, \quad HPD\left(\theta|y,n,\alpha,\beta\right) = 0.06 - 0.36$$



Vaja II

- S pomočjo spodnje kode vzorčite iz izračunanih posteriornih porazdelitev in izračunajte:
  - povprečje
  - standardni odklon
  - credible interval
  - highest posterior density

```
n <- 35; y <- 8
alpha <- 1; beta <- 1
a <- y + 1 + alpha; b <- n - y + 1 + beta
x <- rbeta(n=100000, shape1=a, shape2=b)
hist(x)
mean(x)
sd(x)
mean(x) + c(-1, 1)*1.96*sd(x)
library(package="LaplacesDemon")
p.interval(x, plot=TRUE)
p.interval(x, plot=TRUE, HPD=FALSE)
```

Vprašanja?

Apriorna porazdelitev

"Tisti, ki uporablja Bayesovsko statistiko, na podlagi nejasnega/meglenega pričakovanja konja in bežnega pogleda na osla, trdno sklepa, da je videl mulo." Senn (1997)

- Neinformativna apriorna p. ne obstaja. Celo enakomerna apriorna porazdelitev pravi, da so vse vrednosti enako verjetne.
- Ni vse tako "črno", saj obstajajo pristopi za izpeljavo t.i. neinformativnih apriornih porazdelitev
  - Jefrreys-ove apriorne porazdelitve
  - referenčne apriorne porazdelitve (Bernardo)
  - ...
- Analiza občutljivosti
- Uporaba smiselne apriorne informacije vodi do pristranih (ang. biased) ocen, ki pa imajo manjšo varianco (kombinacija dveh virov informacije –> zmanjšana varianca)
- Prednost dajemo konjugiranim apriornim p., ki so podobne oblike kot posteriorna p.
- S povečevanjem števila podatkov se vpliv apriorne p. zmanjšuje

“Splošna” apriorna porazdelitev

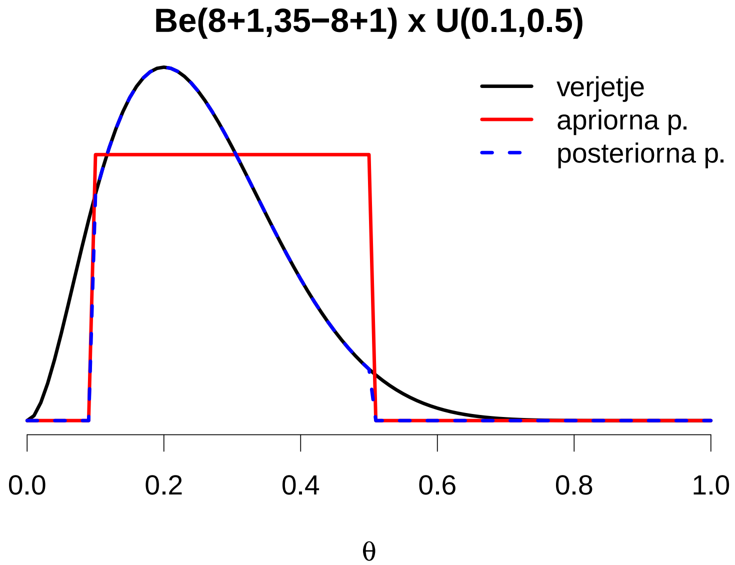
- Za apriorno porazdelitev lahko načeloma izberemo poljubni tip porazdelitve:

$$
\begin{aligned}
p(y|n,\theta) &= Be(y+1, n-y+1) \\
p(\theta) &= U(\alpha,\beta) \\
p(\theta|y,n) &\propto \theta^{y}(1-\theta)^{n-y} \times \frac{1}{\beta-\alpha} \\
&\propto \int_{\alpha}^{\beta} Be(y+1, n-y+1)\frac{1}{\beta-\alpha} d\theta \\
&\propto ????
\end{aligned}
$$

“after the inspection <u>no standard distribution</u> can be recognized”

Ne moremo direktno izračunati (včasih možno ampak bolj poredko).
Potreben drugačen pristop!

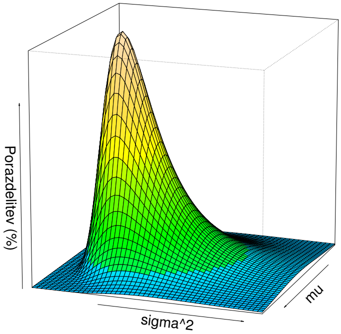
"Splošna" apriorna porazdelitev - enakomerna



Večje število parametrov?

- Statistični modeli imajo običajno več kot en parameter
- Npr. Gaussova porazdelitev

$$
\begin{aligned}
p\left(y|\mu,\sigma^2\right) &= N\left(\mu,\sigma^2\right)\\
p\left(\mu\right) &= N\left(\ldots\right)\\
p\left(\sigma^2\right) &= IG\left(\ldots\right)\\
p\left(\mu,\sigma^2|y\right) &\propto p\left(y|\mu,\sigma^2\right)p\left(\mu\right)p\left(\sigma^2\right)
\end{aligned}
$$

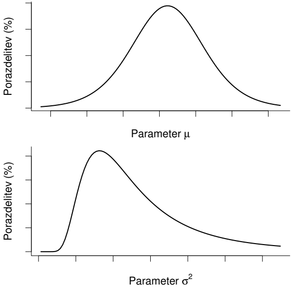


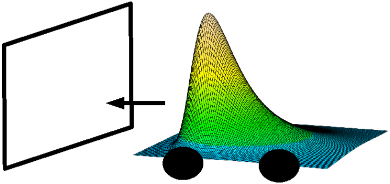
Ampak mi bi radi kot rezultat robne (enodimenzionalne) posteriorne porazdelitve: $p\left(\mu|y\right)$ in $p\left(\sigma^2|y\right)$

Večje število parametrov?

Ampak mi bi radi kot rezultat robne (enodimenzionalne) posteriorne porazdelitve!

$$p(\mu|y) = \int_0^{\infty} p(\mu, \sigma^2|y)\, d\sigma^2 \qquad p(\sigma^2|y) = \int_{-\infty}^{+\infty} p(\mu, \sigma^2|y)\, d\mu$$





Še večje število parametrov?

$$
\begin{aligned}
p(\theta_1, \theta_2, \ldots, \theta_p | y) &\propto p(y|\theta_1, \theta_2, \ldots, \theta_p)\, p(\theta_1)\, p(\theta_2) \ldots p(\theta_p) \\
p(\theta_1 | y) &= \int_{\theta_2} \int_{\ldots} \int_{\theta_p} p(\theta_1, \theta_2, \ldots, \theta_p | y)\, d\theta_2 d \ldots d\theta_p \\
p(\theta_2 | y) &= \int_{\theta_1} \int_{\ldots} \int_{\theta_p} p(\theta_1, \theta_2, \ldots, \theta_p | y)\, d\theta_1 d \ldots d\theta_p \\
\ldots & \\
\boldsymbol{\theta} &= \theta_1, \theta_2, \ldots, \theta_p \\
p(\theta_i | y) &= \int_{\boldsymbol{\theta}_{-i}} p(\boldsymbol{\theta} | y)\, d\boldsymbol{\theta}_{-i}
\end{aligned}
$$

Običajno analitično nerešljivo!!!

Moderni numerični pristopi za Bayesovsko statistiko

- MCMC
- Variational Bayes (Expectation Propagation, Mean-Field, ...)
- INLA (Integrated Nested Laplace Approximation) - za nekatere modele
- ...

Vprašanja?
(pavza)

4. Metode MCMC

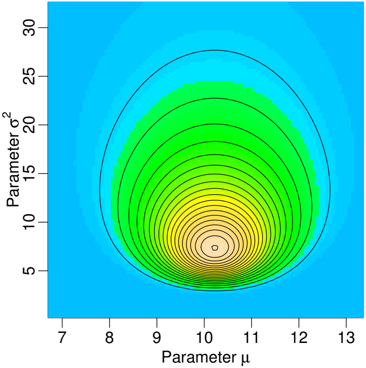
MCMC

- MCMC
  - ang. Monte Carlo Markov Chain
  - slo. Monte Carlo z Markovskimi verigami
- Monte Carlo - stohastična metoda za rešitev računsko/analitično zahtevnih primerov
- Markovske verige - sedanja vrednost parametra je odvisna le od predhodne vrednosti
- U okviru Bayesovske statistike uporabljamo metodo MCMC za vzorčenje iz (pogojnih) posteriornih porazdelitev
- Algoritmi
  - Metropolis in Metropolis-Hastings sampling
  - Gibbs sampling
  - NUTS
  - HMC
  - ...

Večje število parametrov?

- Statistični modeli imajo običajno več kot en parameter
- Npr. Gaussova porazdelitev

$$\begin{aligned} p\left(y|\mu,\sigma^2\right) &= N\left(\mu,\sigma^2\right) \\ p\left(\mu\right) &= N\left(\ldots\right) \\ p\left(\sigma^2\right) &= IG\left(\ldots\right) \\ p\left(\mu,\sigma^2|y\right) &\propto p\left(y|\mu,\sigma^2\right) p\left(\mu\right) p\left(\sigma^2\right) \end{aligned}$$



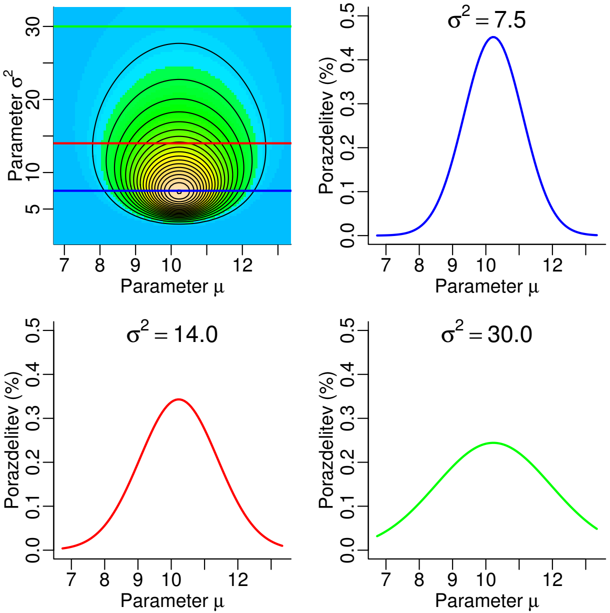
Ampak mi bi radi kot rezultat robne (enodimenzionalne) posteriorne porazdelitve: $p\left(\mu|y\right)$ in $p\left(\sigma^2|y\right)$

MCMC - pogojne porazdelitve

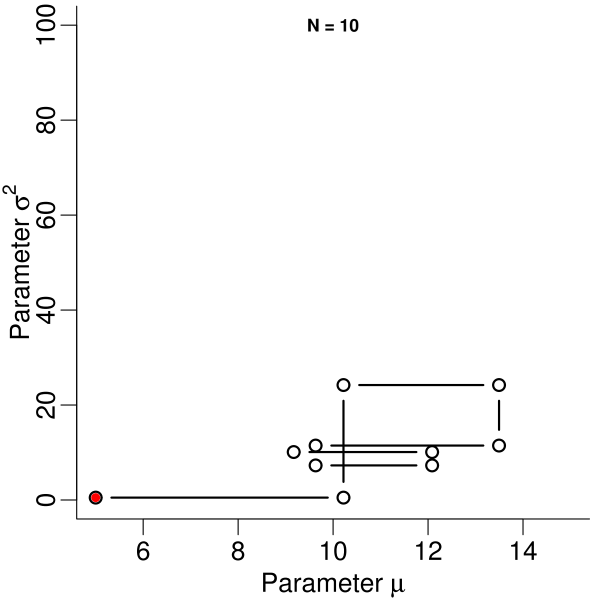
- Izpeljemo pogojne porazdelitve za vsak parameter posebej (lahko tudi za skupine/bloke parametrov skupaj)
- Iz enodimenzionalnih pogojnih porazdelitev je enostavno vzorčiti
- Pogojne porazdelitve za Gaussov model:

$$
\begin{aligned}
p\left(\mu, \sigma^2 | y\right) &\propto p\left(y | \mu, \sigma^2\right) p\left(\mu\right) p\left(\sigma^2\right) \\
p\left(\mu | y, \sigma^2\right) &\propto p\left(y | \mu, \sigma^2\right) p\left(\mu\right) \cancel{p\left(\sigma^2\right)} \\
p\left(\sigma^2 | y, \mu\right) &\propto p\left(y | \mu, \sigma^2\right) \cancel{p\left(\mu\right)} p\left(\sigma^2\right)
\end{aligned}
$$

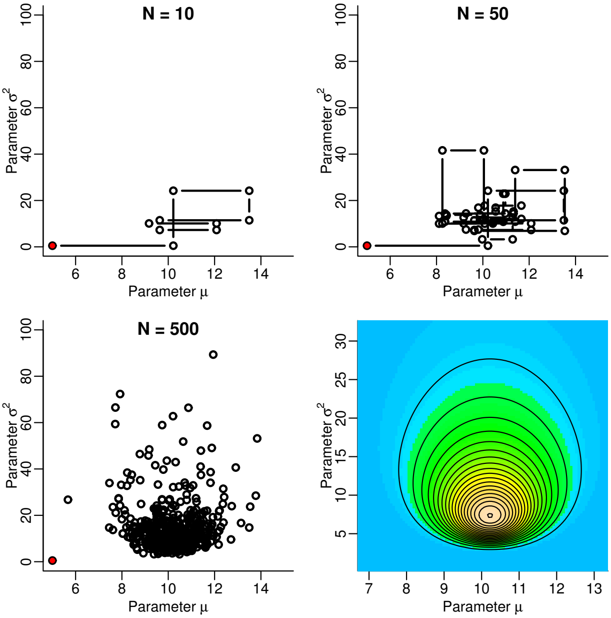
MCMC - pogojne porazdelitve (prikaz)



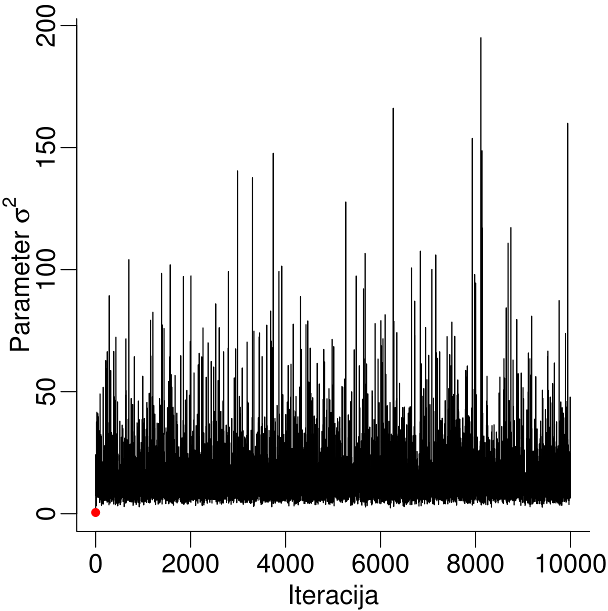
MCMC - vzorčenje



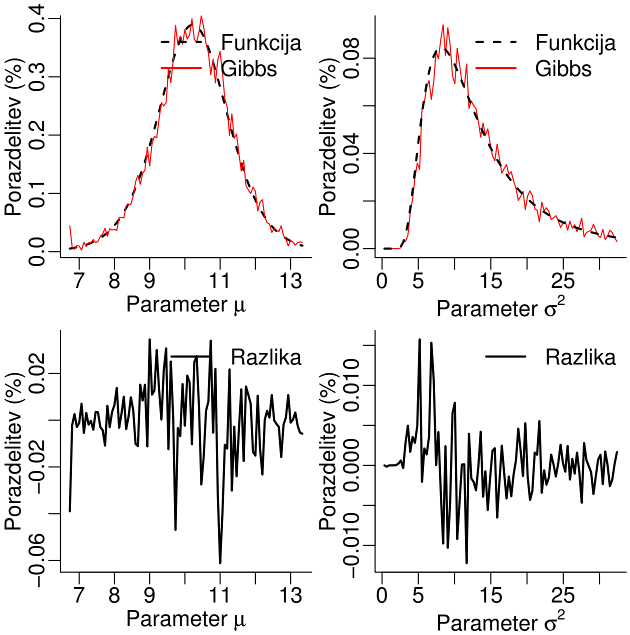
MCMC - vzorčenje II



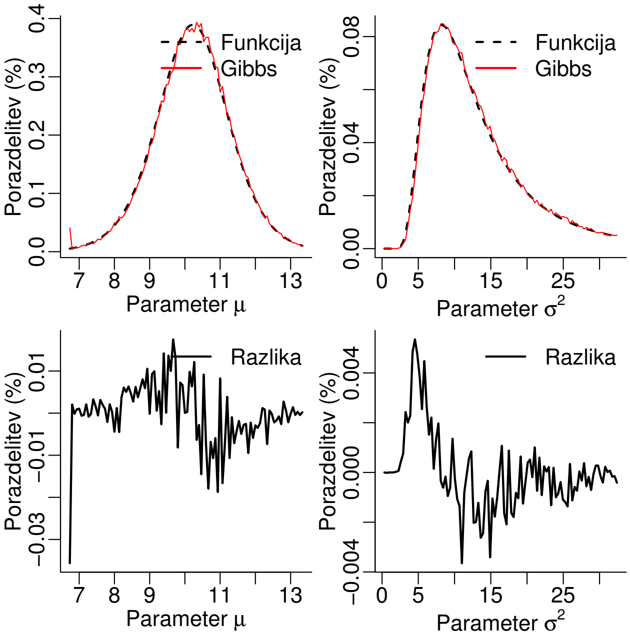
MCMC - Markovska veriga



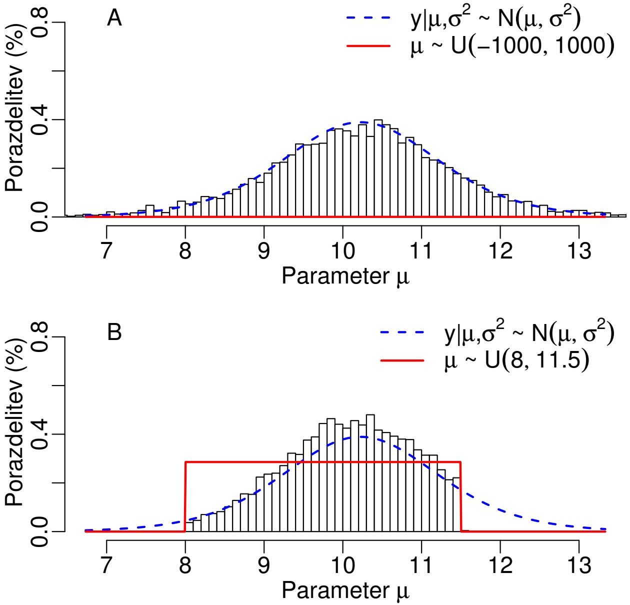
MCMC - robne porazdelitve



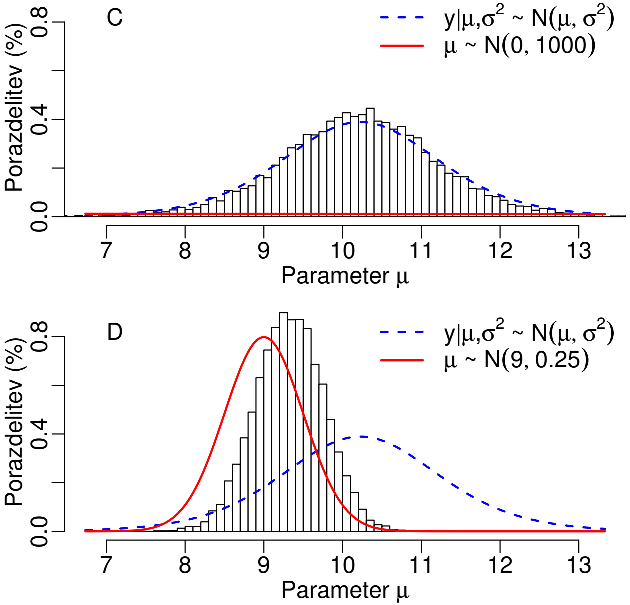
MCMC - robne porazdelitve II



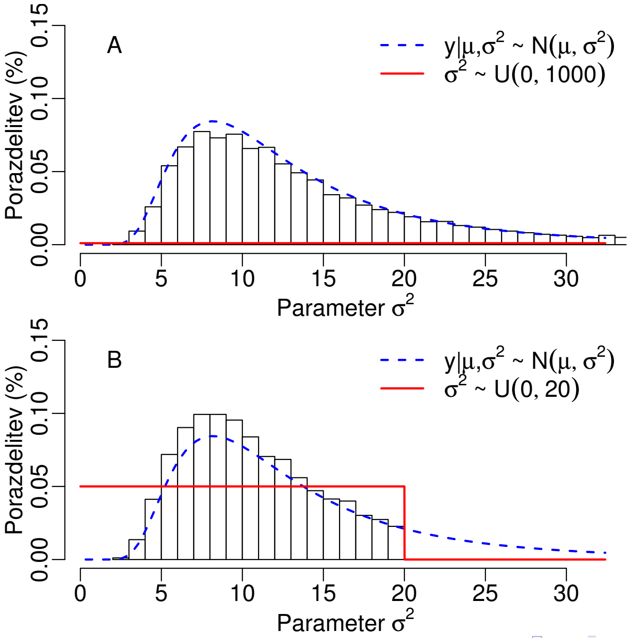
MCMC - različne apriorne p. za μ



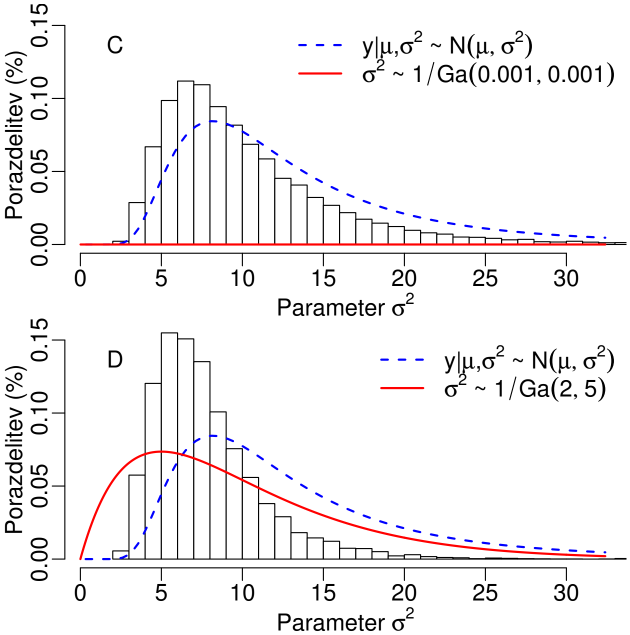
MCMC - različne apriorne p. za μ II



MCMC - različne apriorne p. za σ^2



MCMC - različne apriorne p. za σ^2 II



Vprašanja?

Splošni Bayes programi/paketi

- BUGS, JAGS, in Stan družina
- R paketi
  http://cran.r-project.org/web/views/Bayesian.html
  - brms
    https://cran.r-project.org/web/packages/brms/index.html
  - rstan https://cran.r-project.org/web/packages/rstan/index.html
  - R-INLA https://https://www.r-inla.org
  - ...