

ANOVA Bayesiana

Antônio Oss Boll e Letícia Maria Puttlitz

Introdução e contextualização

A ANOVA Bayesiana é uma abordagem estatística para testar hipóteses em modelos ANOVA que permite incorporar informações prévias sobre os parâmetros do modelo e produz distribuições de probabilidade para os parâmetros. A ANOVA Bayesiana pode ser útil em situações com poucos dados disponíveis ou em modelos mais complexos. Nesta apresentação, vamos explorar os conceitos básicos da ANOVA Bayesiana e sua aplicação.

Motivação

A ANOVA Bayesiana é uma abordagem alternativa à ANOVA frequentista. Ela permite:

- Modelar a incerteza dos parâmetros usando distribuições de probabilidade.
- Incorporar conhecimento prévio sobre os parâmetros e hipóteses.
- A ANOVA Bayesiana pode ser usada para lidar com dados desbalanceados, outliers e heterocedasticidade.
- Pode ser uma opção melhor do que a ANOVA frequentista quando se trabalha com um número limitado de dados.

Tutorial

Banco de dados

O banco de dados foi criado a partir do seguinte código:

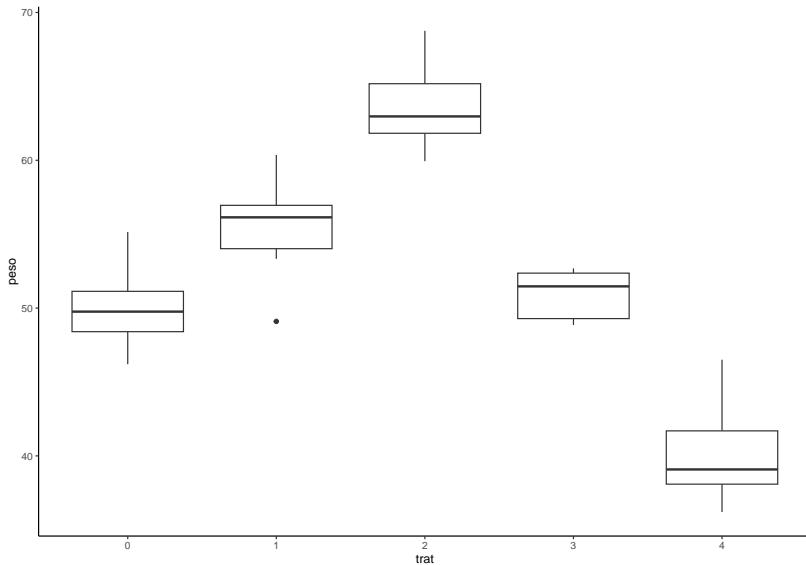
```
set.seed(123)

nsample <- 10; ngroups <- 5
trat <- as.factor(rep(0:4, rep(nsample, ngroups)))
peso <- c(rnorm(nsample, 50, 3), # grupo controle
          rnorm(nsample, 55, 3), # trat 1
          rnorm(nsample, 65, 3), # trat 2
          rnorm(nsample, 50, 3), # trat 3
          rnorm(nsample, 40, 3)) # trat 4
dados <- data.frame(peso, trat)
```

O banco de dados apresenta 2 variáveis:

- ▶ peso: peso
- ▶ trat: Controle, tratamento 1 - 4

Análise gráfica



JAGS

O modelo JAGS foi definido da seguinte forma:

```
model {  
  #Likelihood  
  for (i in 1:N) {  
    y[i]~dnorm(mean[i],tau)  
    mean[i] <- inprod(beta[],X[i,]) }  
  #Priors  
  for (i in 1:ngroups) {  
    beta[i] ~ dnorm(0, 1.0E-6)  
  }  
  sigma ~ dunif(0, 100)  
  tau <- 1 / (sigma * sigma)  
}
```

Código

```
X <- model.matrix(~trat, dados)
data <- with(dados, list(y = dados$peso, # variavel resp
                        X = X, # grupos
                        N = nrow(dados), # n observacoes
                        ngroups = length(levels(dados$trat)) # n grupos
))

params <- c("beta", "sigma")

library(jagsUI)

anova_jags_adt <- jags(data, NULL, params, "anova.jags",
                       n.chains=2, n.iter=10000,
                       DIC=FALSE)
```


Resultados

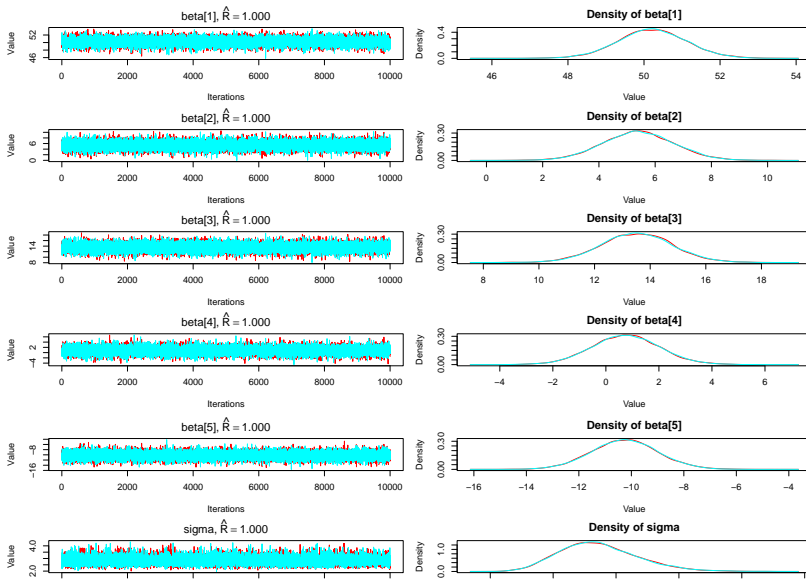
```
sum_anova <- round(anova_jags_adt$summary[,c(1:3,7,10)],3)
```

##	mean	sd	2.5%	97.5%	overlap0
## beta[1]	50.217	0.908	48.408	52.020	0
## beta[2]	5.404	1.279	2.919	7.925	0
## beta[3]	13.517	1.290	10.987	16.065	0
## beta[4]	0.754	1.290	-1.758	3.307	1
## beta[5]	-10.247	1.271	-12.738	-7.730	0
## sigma	2.864	0.312	2.332	3.541	0

- ▶ a média do grupo controle é 50.217
- ▶ a media do trat1 é 5.404 unidades maiores que o grupo 0
- ▶ a media do trat2 é 13.517 unidades maiores que o grupo 0
- ▶ a media do trat3 é 0.754 unidades maiores que o grupo 0
- ▶ a media do trat4 é 10.247 unidades menores que o grupo 0

Diagnóstico de convergência

```
plot(anova_jags_adt)
```



Diagnóstico de convergência

```
anova_jags_adt$summary[,c(8,9)]
```

```
##              Rhat n.eff
## beta[1] 0.9999978 20000
## beta[2] 1.0000213 20000
## beta[3] 0.9999747 20000
## beta[4] 0.9999934 20000
## beta[5] 0.9999739 20000
## sigma   1.0000159 20000
```

- ▶ Rhat - valores próximos de 1
- ▶ Tamanho amostral efetivo - valores altos

Comparação múltipla

```
mcmc <- matrix(c(anova_jags_adt$sims.list$beta[,1],
                 anova_jags_adt$sims.list$beta[,2],
                 anova_jags_adt$sims.list$beta[,3],
                 anova_jags_adt$sims.list$beta[,4],
                 anova_jags_adt$sims.list$beta[,5]),
               nrow = 20000, ncol = 5, byrow=F)
coefs <- as.matrix(mcmc)[, 1:5]
newdata <- data.frame(x = levels(dados$trat))
library(multcomp)
tuk.mat <- contrMat(n = table(newdata$x), type = "Tukey")
Xmat <- model.matrix(~x, data = newdata)
pairwise.mat <- tuk.mat %*% Xmat
```

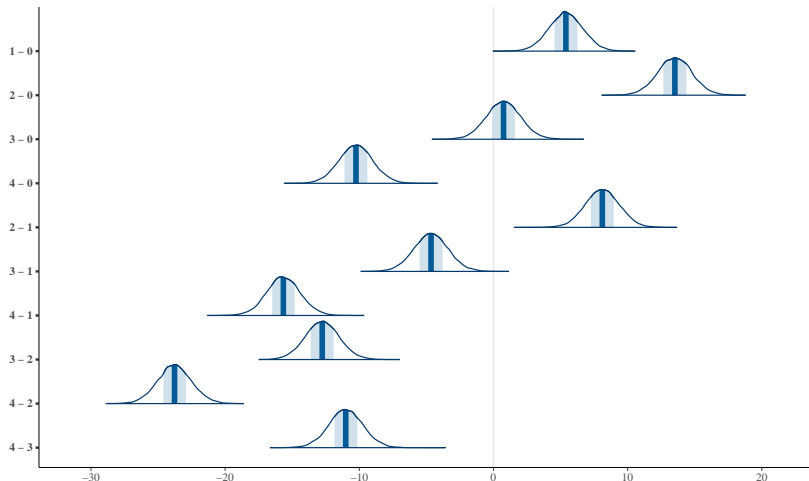
Comparação múltipla

Matriz de comparação resultante do código anterior:

##		(Intercept)	x1	x2	x3	x4
##	1 - 0		0	1	0	0
##	2 - 0		0	0	1	0
##	3 - 0		0	0	0	1
##	4 - 0		0	0	0	1
##	2 - 1		0	-1	1	0
##	3 - 1		0	-1	0	1
##	4 - 1		0	-1	0	1
##	3 - 2		0	0	-1	1
##	4 - 2		0	0	-1	1
##	4 - 3		0	0	0	-1

Comparação múltipla

```
library(bayesplot)  
mcmc_areas(coefs %*% t(pairwise.mat))
```



Comparaç o m ltipla

```
library(broom.mixed)
(comps = tidyMCMC(coefs %*% t(pairwise.mat),
  conf.int = TRUE, conf.method = "HPDinterval"))
```

```
## # A tibble: 10 x 5
```

##	term	estimate	std.error	conf.low	conf.high
##	<chr>	<dbl>	<dbl>	<dbl>	<dbl>
##	1 1 - 0	5.40	1.28	2.91	7.90
##	2 2 - 0	13.5	1.29	10.9	16.0
##	3 3 - 0	0.749	1.29	-1.79	3.26
##	4 4 - 0	-10.2	1.27	-12.7	-7.71
##	5 2 - 1	8.11	1.28	5.65	10.6
##	6 3 - 1	-4.65	1.30	-7.21	-2.12
##	7 4 - 1	-15.7	1.27	-18.2	-13.1
##	8 3 - 2	-12.8	1.28	-15.3	-10.2
##	9 4 - 2	-23.8	1.28	-26.3	-21.2
##	10 4 - 3	-11.0	1.28	-13.6	-8.53

Comparação múltipla

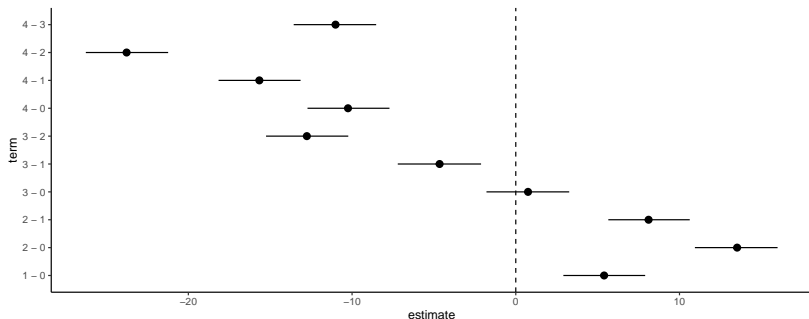
```
comps[c(3,5,7),]
```

```
## # A tibble: 3 x 5
##   term estimate std.error conf.low conf.high
##   <chr>      <dbl>      <dbl>    <dbl>    <dbl>
## 1 3 - 0      0.749        1.29    -1.79     3.26
## 2 2 - 1      8.11        1.28     5.65    10.6
## 3 4 - 1     -15.7        1.27   -18.2    -13.1
```

- ▶ A diferença entre o grupo 0 e grupo 3 não é significativa. Se o IC contem 0, a diferença não é significativa.
- ▶ a media do grupo 2 é 8.11 unid maiores que o grupo 1.
- ▶ a media do grupo 4 é 15.66 unid menores que o grupo 1.

Comparação múltipla

```
ggplot(comps, aes(y = estimate, x = term)) +  
  geom_pointrange(aes(ymin = conf.low,  
    ymax = conf.high)) +  
  geom_hline(yintercept = 0, linetype = "dashed") +  
  coord_flip() + theme_classic()
```



Diferença entre o trat3 e o grupo controle não é significativa, dado que o IC contém 0.

Conclusão

Há evidências para concluir que todos os grupos se diferem entre si, exceto o tratamento 3 do grupo controle. O grupo que apresentou maior valor de peso foi o tratamento 2 e com menor valor foi o tratamento 4.

Referências

- ▶ “<https://agabrioblog.onrender.com/jags/single-factor-anova-jags/single-factor-anova-jags/>”
- ▶ “https://web.pdx.edu/~joel8/resources/ConceptualPresentationResources/JAGS_ANOVA.pdf”
- ▶ “<https://www.r-bloggers.com/2020/12/general-code-to-fit-anova-models-with-jags-and-rjags/>”

Para acessar o repositório com os arquivos necessários para fazer esta análise, clique [aqui](#).