Universidade de Brasília

DEPARTAMENTO DE ESTATÍSTICA

04 julho 2024

Trabalho

Prof. Dr. Donald Matthew Pianto

Aluno: Bruno Gondim Toledo

Matrícula: 15/0167636

Inferência Bayesiana

1º/2024

Bayesiana 1

Lista 5

(5) $i = 1, \dots, n$, $S_i | \lambda_i \overset{indep}{\sim} \text{Poisson}(\lambda_i t_i)$, t_i fixo.

$\lambda_i \sim \text{Gama}(\alpha_0, \beta)$

$\beta \sim \text{Gama}(a, b)$

i) $\underset{\sim}{\lambda} | D, \beta$;

Sol: A verossimilhança de λ_i é

$$\mathbb{P}(S_i | \lambda_i) = \prod_{i=1}^{n} \frac{(\lambda_i t_i)^{S_i} e^{-(\lambda_i t_i)}}{S_i!}$$

f.d.p. Poisson:

$$\frac{\lambda^k e^{-\lambda}}{k!}$$

f.d.p. Gama:

$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

Já a prior de λ_i;

$$\mathbb{P}(\lambda_i | \beta) = \prod_{i=1}^{n} \frac{\beta^{\alpha_0}}{\Gamma(\alpha_0)} \lambda_i^{\alpha_0 - 1} e^{-\beta \lambda_i}$$

Como β é fixo;

$$\mathbb{P}(\underset{\sim}{\lambda} | D, \beta) \propto \mathbb{P}(S_i | \lambda_i) \mathbb{P}(\lambda_i | \beta)$$

Ignorando o produtório e constantes;

$$\mathbb{P}(\underset{\sim}{\lambda} | D, \beta) \propto (\lambda_i \overset{\text{fixo}}{t_i})^{S_i} e^{-(\lambda_i t_i)} \lambda_i^{\alpha_0 - 1} e^{-\beta \lambda_i}$$

$$\propto \lambda_i^{S_i} \lambda_i^{\alpha_0 - 1} e^{-\lambda_i t_i} e^{-\beta \lambda_i}$$

$$\propto \lambda_i^{S_i + \alpha_0 - 1} e^{-\lambda_i (t_i + \beta)}$$

2

Logo, isso será proporcional a uma dist. Gama com parâmetros $(\alpha = s_i + \alpha_0, \beta = t_i + \beta)$, t.q.

$$\underset{\sim}{\lambda} = (\lambda_1, \dots, \lambda_n) \mid D, \beta \sim \text{Gama}(s_i + \alpha_0, t_i + \beta).$$

ii)

a) $$w_i \propto \frac{\mathbb{P}(\beta_i \mid D)}{\mathbb{P}^*(\beta_i)}$$

Pela regra de Bayes, $\mathbb{P}(\beta_i \mid D) = \dfrac{\mathbb{P}(D \mid \beta)\,\mathbb{P}(\beta)}{\mathbb{P}(D)}$

Como $\mathbb{P}(D)$ é constante (fixo), temos que

$$w_i \propto \frac{\mathbb{P}(D \mid \beta_i)\,\mathbb{P}(\beta_i)}{\mathbb{P}^*(\beta_i)}$$

b) Usando os resultados anteriores, temos que

$$w_i \propto \frac{\mathbb{P}(D \mid \underset{\sim}{\lambda})\,\mathbb{P}(\underset{\sim}{\lambda} \mid \beta_i)\,\mathbb{P}(\beta_i)}{\mathbb{P}^*(\beta_i)\,\mathbb{P}(\underset{\sim}{\lambda} \mid \beta_i)} \propto \frac{\mathbb{P}(D \mid \beta_i)\,\mathbb{P}(\beta_i)}{\mathbb{P}^*(\beta_i)}$$

Que é igual aos pesos w_i encontrados em a), mostrando assim a proposição ⊠

Agora, basta aproximar $E_i(\lambda_i \mid D)$ e $Var(\lambda_i \mid D)$ por importância amostral, usando simulação computacional...

Lista 5

Questão 5

Aproximação por importância amostral para $\mathbb{E}(\lambda_i|D)$ e $Var(\lambda_i|D)$, utilizando $\alpha_0 = 0,166; a = 0,1; b = 0,01$, para o conjunto de dados *pump*, disponível no pacote *bang*

```r
# Dados (Disponível no pacote bang)
library(bang)
D <- as.data.frame(pump)

# Parâmetros
alfa0 <- 0.166
a <- 0.1
b <- 0.01
s <- D$failures
t <- D$time
n <- 1000 # Número de amostras

pesos <- numeric(n)
amostras_beta <- matrix(0, nrow=n, ncol=nrow(D))
beta <- rgamma(n, shape=a, rate=b)

for (i in 1:n) {
  beta_i <- beta[i]
  amostras_beta[i, ] <- rgamma(nrow(D), shape=s + alfa0, rate=t + beta_i)
  priori <- dgamma(beta_i, shape=a, rate=b)
  verossimilhanca <- prod(dpois(s, amostras_beta[i, ] * t))
  dist_proposta <- dgamma(beta_i, shape=a, rate=b)
  pesos[i] <- priori * verossimilhanca / dist_proposta
}

pesos <- pesos/sum(pesos)

medias <- colSums(amostras_beta * pesos)
variancias <- colSums((amostras_beta - medias)^2 * pesos)

D$esperanca_lambda <- medias
D$variancia_lambda <- variancias
D$bomba <- 1:10
D <- D[,c(5,1:4)]
colnames(D)[2:3] = c("n_falhas","tempo")
```

Portanto, temos os valores de esperança e variância para cada λ_i:

```r
kable(D)
```

bomba	n_falhas	tempo	esperanca_lambda	variancia_lambda
1	5	94.320	0.0528686	0.8395437
2	1	15.720	0.0653743	0.8239147
3	5	62.880	0.0794960	0.8050125
4	14	125.760	0.1105495	0.7648497
5	3	5.240	0.5826058	0.4733171
6	19	31.440	0.6142293	0.4364407
7	1	1.048	0.9677066	0.8368096
8	1	1.048	0.9714465	0.9211895
9	4	2.096	1.9026566	2.3102390
10	22	10.480	2.0749824	2.4549429

Desta tabela, conseguimos tirar algumas conclusões. Vemos uma variabilidade para os valores de λ_i de 0.05 à 2.07, mostrando que este valor difere bastante em relação à bomba observada. Isso pode estar relacionado a diversos fatores, como idade da bomba ou fabricante. A variância deste parâmetro em geral foi baixa, salvo para as últimas bombas, o que indica que temos alguma precisão dessas estimativas. O algoritmo em sí funciona bem rápido para $n = 1000$, não sendo uma grande questão computar estes valores. A modelagem *Poisson* para este conjunto faz todo sentido, visto que temos visivelmente um processo de Poisson a tempo contínuo, com um intervalo de exposição dado e uma contagem de sucessos, porém vemos que a modelagem bayesiana é mais interessante neste caso que a frequentista por fornecer valores mais apropriados para os parâmetros, visto que numa aborgadem frequentista utilizariamos a propriedade da Poisson de que $\mathbb{E}(\lambda) = \mathbb{V}(\lambda)$, o que até observamos para algumas bombas do conjunto, porém algumas o valor da variância é significativamente diferente do valor da média, observada a escala dos números. É interessante também observar na prática o isomorfismo da Poisson com a Gama, visto que foi esta a distribuição utilizada para gerar os valores de λ_i.

Questão 6

Repetindo o exercício anterior, porém utilizando amostrador Gibbs

```
D <- as.data.frame(pump)

# Parâmetros
alfa0 <- 0.166
a <- 0.1
b <- 0.01
s <- D$failures
t <- D$time
n <- 5000 # Número de amostras
beta <- 0 # Valor inicial para iniciar o loop "não informativo" para Beta

amostras_lambda <- matrix(0, nrow=n, ncol=nrow(D))

for (i in 1:n) {
  lambda_i <- rgamma(nrow(D), shape=s + alfa0, rate=t + beta)
  beta <- rgamma(1, shape=a + nrow(D) * alfa0, rate=b + sum(lambda_i))
  amostras_lambda[i, ] <- lambda_i
}

df <- as.data.frame(t(amostras_lambda[1001:5000,])) # Removendo primeiras 1000 amostras (burn-in)
df$esperanca_lambda <- apply(df, 1, mean)
df$variancia_lambda <- apply(df, 1, var)
df = df[,4001:4002]

D = cbind(D,df)
D$bomba = 1:10
D = D[,c(5,1:4)]
colnames(D)[2:3] = c("n_falhas","tempo")
```

Portanto, temos os valores de esperança e variância para cada λ_i:

```
kable(D)
```

bomba	n_falhas	tempo	esperanca_lambda	variancia_lambda
1	5	94.320	0.0545004	0.0005816
2	1	15.720	0.0725302	0.0047355
3	5	62.880	0.0816517	0.0013078
4	14	125.760	0.1115865	0.0008862
5	3	5.240	0.5777698	0.1054447
6	19	31.440	0.6076317	0.0189502
7	1	1.048	0.9124393	0.7095151
8	1	1.048	0.8964294	0.7259171
9	4	2.096	1.7793163	0.7908399
10	22	10.480	2.0682979	0.1932916

A tabela é idêntica à anterior, senão pela estimativa dos parâmetros pelo método utilizado aqui. Este algoritmo foi até mais simples de implementar que o anterior. Foi definido um tamanho amostral maior $n = 5000$, por conta da necessidade de eliminar algumas amostras (*burn-in*, ou seja, eliminar algumas amostras até que o valor de β comece a "caminhar" pelo espaço paramétrico, visto que colocamos um valor inicial arbitrário $\beta = 0$). Foram eliminadas as primeiras 1000 amostras como *burn-in*. O tempo de execução é extremamente parcimonioso. Quanto as estimativas produzidas, a esperança de lambda é praticamente igual idependente do método, porém a variância aqui foi menor para as últimas bombas, mostrando que este método é bastante robusto para produzir estimativas sobre os parâmetros.