Exemplos biparamétricos: comparação de duas distribuições de Poisson (Intervalos de credibilidade)

Prof. Caio Azevedo

Dados reais: comparação do número de acidentes

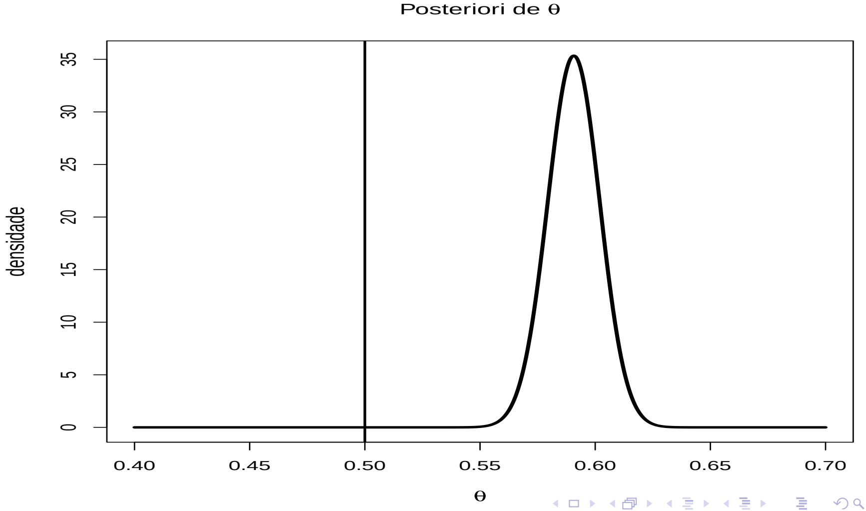
- Descrição: número de acidentes (com algum tipo de trauma para as pessoas envolvidas) em 92 dias (correspondentes) em dois anos distintos (1961 e 1962), medidos em algumas regiões da Suécia.
- Considerou-se apenas 43 dias, correspondendo a dias de 1961 em que não havia limite de velocidade e de 1962 em que havia limites de velocidade (90 ou 100 km/h).
- Vamos assumir que
  $X_i|\lambda_1 \overset{i.i.d.}{\sim} Poisson(\lambda_1), i = 1, ..., 43$ (1961).
  $Y_i|\lambda_2 \overset{i.i.d.}{\sim} Poisson(\lambda_2), i = 1, ..., 43$ (1962).

Dados reais: comparação do número de acidentes

- Assumiremos também que $p(\lambda_1, \lambda_2) \propto \mathbb{1}_{(0,\infty)}(\lambda_1)\mathbb{1}_{(0,\infty)}\mathbb{1}(\lambda_2)$.
- Dessa forma, teremos que (como já foi visto)

  $\lambda_1|\boldsymbol{x} \sim gama(n\overline{x} + 1, n^{-1}) \perp \lambda_2|\boldsymbol{y} \sim gama(n\overline{y} + 1, n^{-1})$ (exercício).
- Assim, $\theta = \frac{\lambda_1}{\lambda_1+\lambda_2}|(\boldsymbol{x}, \boldsymbol{y}) \sim beta(n\overline{x} + 1, n\overline{y} + 1)$.
- Esses dados já foram analisados de uma forma um pouco diferente aqui.

Dados reais: comparação do número de acidentes (cont.)



Dados reais: comparação do número de acidentes (cont.)

- Podemos facilmente obter um $IC_B(\theta,\gamma)$ simétrico, através do R, uma vez que $\theta = \frac{\lambda_1}{\lambda_1+\lambda_2}|(\boldsymbol{x},\boldsymbol{y}) \sim beta(n\overline{x}+1, n\overline{y}+1)$
- Basta digitar

```
ICsim <- c(qbeta((1-gammac)*0.5,shape1=n*xb+1,
shape2=m*yb+1),qbeta((1+gammac)*0.5,
shape1=n*xb+1,shape2=m*yb+1))
```

em que "gammac" é o grau de credibilidade desejado.

- Para o intervalo $HPD(\theta,\gamma)$ basta usar a função "hpd", ou seja

```
hpd(qbeta,shape1=n*xb+1,shape2=m*yb+1,conf=gammac)
```

Dados reais: comparação do número de acidentes (cont.)

- Além disso, podemos calcular (motivados pelos objetivos e pela figura anterior):

$$P_{\theta|\boldsymbol{x}}(\theta > 1/2|\boldsymbol{x}) \equiv P(\theta > 1/2|\boldsymbol{x}).$$

- No R, basta digitar

```
1-pbeta(0.5,shape1=n*xb+1,shape2=m*yb+1)
```

Dados reais: comparação do número de acidentes (cont.)

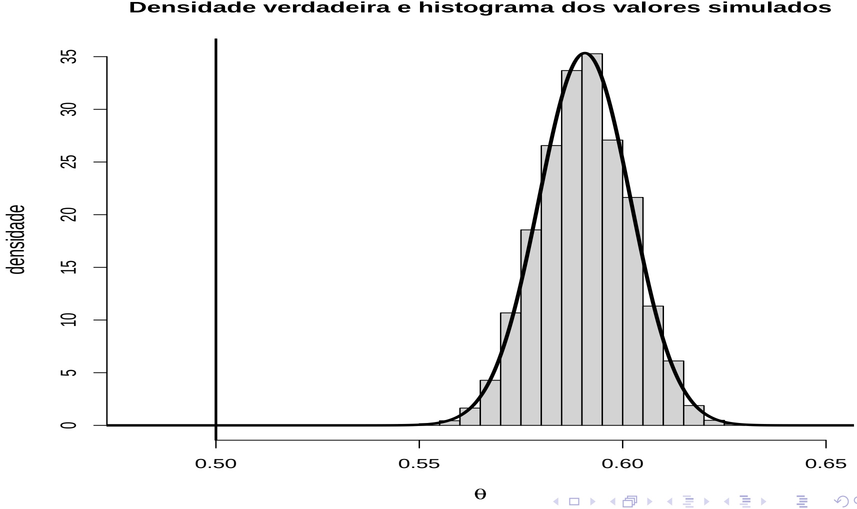
- Os resultados foram:
  - $IC_B(\theta; 0,95) = [0,568; 0,613]$
  - $HPD(\theta; 0,95) = [0,568; 0,613]$
  - $P_{\theta|\boldsymbol{x}}(\theta > 1/2|\boldsymbol{x}) > 0,9999$
- Os resultados acima nos levam à concluir que $\theta > 0,5$ e que, portanto, houve um redução no número médio de acidentes (eventualmente devido à imposição de limites de velocidade).

Obtendo a posteriori numericamente

- Uma vez que $\lambda_1|\boldsymbol{x} \sim gama(n\overline{x}+1, n^{-1})$, $\lambda_2|\boldsymbol{y} \sim gama(n\overline{y}+1, n^{-1})$ e $\lambda_1|(\boldsymbol{x},\boldsymbol{y}) \perp \lambda_2|(\boldsymbol{x},\boldsymbol{y})$, podemos simular "R" variáveis aleatórias, mutuamente independentes, com distribuições gama específicas, e calcular θ para cada par, ou seja:

  1. Simular $(\lambda_1^{(r)}, \lambda_2^{(r)}), r = 1, ..., R$ (das respectivas distribuições) e calcular $\theta^{(r)} = \frac{\lambda_1^{(r)}}{\lambda_1^{(r)}+\lambda_2^{(r)}}$.
  2. Dessa forma, teremos uma aproximação numérica da posteriori de interesse.
  3. Os resultados obtidos anteriormente, podem ainda sê-lo, uma vez que se tenha uma amostra aleatória de tamanho R da posteriori.

Dados reais: comparação do número de acidentes (cont.)



Dados reais: comparação do número de acidentes (cont.)

- Para uma amostra aleatória de tamanho R = 2.000 obtivemos
  - $IC_B(\theta; 0,95) = [0,569; 0,612]$
  - $HPD(\theta; 0,95) = [0,569; 0,612]$
  - $P(\theta > 1/2|\boldsymbol{x}) > 0,9999$

Densidade a posteriori e histogramas dos valores simulados