ベイズ統計

古谷知之

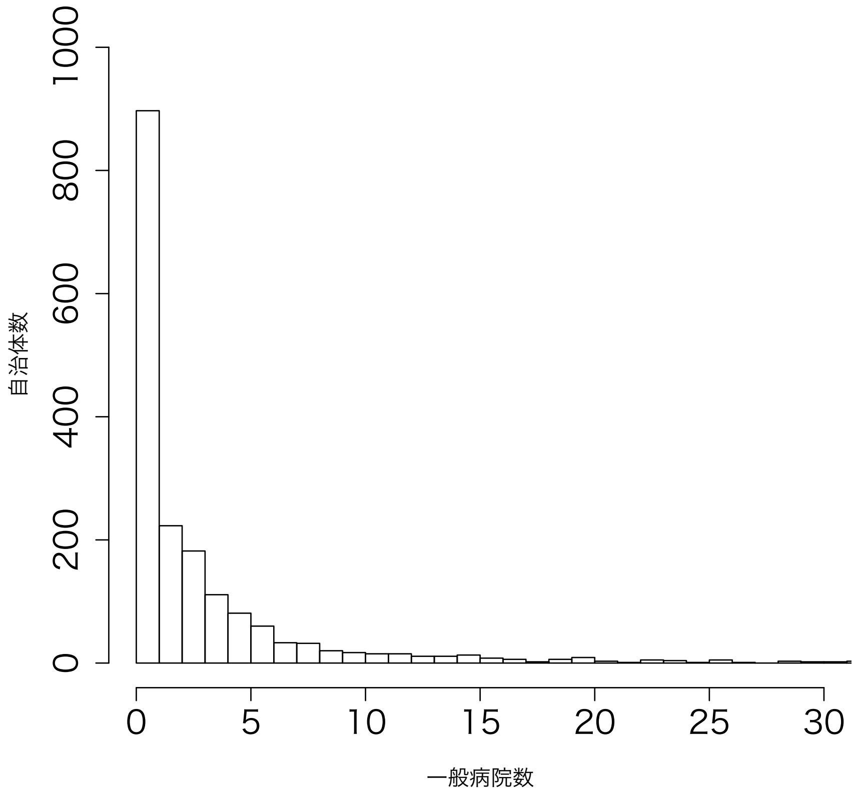
講義概要

- マルチレベル・モデルの演習(前回続き)
- ポアソン回帰モデル
- 負の二項分布モデル
- R演習

ポアソン分布

- 事象の発生数y、平均と分散がλのとき、ポアソン分布は以下のようになる

$$P(y) = \frac{\lambda^y \exp(-\lambda)}{y!}$$



ポアソン回帰モデルのベイズ推定

- ポアソン分布を説明する回帰モデルは次式のように表される

$$y_i = Po\left(\lambda_i\right)$$

$$\log \lambda_i = \beta_0 + \beta_1 x_{1i}$$

- 回帰係数が正規分布に従うので、例えば以下のように事前情報を設定する。

$$\beta_0, \beta_1 \sim N\left(0, 1 \times 10^{-6}\right)$$

としてベイズ推定する。

負の二項分布

- 二項分布：n回の試行でr回成功（ベルヌーイ試行）$P(y) = {}_nC_r \pi^r (1-\pi)^{n-r}$
- 負の二項分布：ある事象についてr回までその事象が生じなかった確率分布

$$P(y) = {}_{n-1}C_{r-1} \pi^r (1-\pi)^{n-r}$$

- 過大分散が生じるような現象に対して用いられる

負の二項分布

- n=y+rと置き換えると、負の二項分布はガンマ分布を用いて以下のように表せる

$$P(y) = {}_{y+r-1}C_{r-1}\pi^r(1-\pi)^y$$

$$= \frac{\Gamma(y+r)}{y!\Gamma(r)}\pi^r(1-\pi)^y$$

- このとき
  - 平均　$r(1-\pi)/\pi$
  - 分散　$r(1-\pi)/\pi^2$

負の二項分布

- さらにπ=r/(λ+r)と置き換えると、負の二項分布以下のように表せる

$$P(y)=\frac{\Gamma(y+r)}{y!\Gamma(r)}\left(\frac{r}{r+\lambda}\right)^r\left(\frac{\lambda}{r+\lambda}\right)^y$$

- このとき
  - 平均 λ
  - 分散 $(\lambda+r\lambda)/r$
- $(1+r)/r$を分散指標という

負の二項分布のベイズ推定

- 負の二項分布モデル

$$y_i \sim NB(r, \pi_i),$$

$$\pi_i \sim r/(\lambda_i + r),$$

$$\log(\lambda_i) = \beta_0 + \beta_1 x_{1i}$$

- 事前情報は例えば以下のようにする

$$\beta_0, \beta_1 \sim N(0, 0.01)$$

$$r \sim \Gamma(0.001, 0.001)$$