

Problem 6.1

Suppose that instead of the Poisson model (6.1), we assume a binomial model

$$Pr\{x_k = x\} = \binom{n}{x} \theta_k^x (1 - \theta_k)^{n-x},$$

n some fixed and known integer such as $n = 10$. What is the equivalent of Robbins' formula (6.5)?

Solution**Problem 6.2**

Define $\mathbb{V}\{\theta|x\}$ as the variance of θ given x . In the Poisson situation (6.1), show that

$$\mathbb{V}\{\theta|x\} = \mathbb{E}\{\theta|x\} \cdot (\mathbb{E}\{\theta|x+1\} - \mathbb{E}\{\theta|x\}),$$

where $\mathbb{E}\{\theta|x\}$ is as given in (6.5).

Solution By Robbins' formula, $\mathbb{E}(\theta|x) = (x+1) \frac{f_g(x+1)}{f_g(x)}$, where $f_g(x) = \int_0^\infty p_\theta(x) g(\theta) d\theta$ is marginal distribution of x . Then, we know that $\mathbb{V}\{\theta|x\} = \mathbb{E}(\theta^2|x) - (\mathbb{E}(\theta|x))^2$. So we'll show that $\mathbb{E}(\theta^2|x) = \mathbb{E}(\theta|x) \cdot \mathbb{E}(\theta|x+1)$.

$$\begin{aligned} \mathbb{E}(\theta^2|x) &= \frac{\int_0^\infty \theta^2 p_\theta(x) g(\theta) d\theta}{f_g(x)} = \frac{\int_0^\infty \theta^2 \frac{\theta^x e^{-\theta}}{x!} g(\theta) d\theta}{f_g(x)} \\ &= (x+2)(x+1) \frac{\int_0^\infty \theta^{x+2} \frac{e^{-\theta}}{(x+2)!} g(\theta) d\theta}{f_g(x)} = (x+2)(x+1) \frac{f_g(x+2)}{f_g(x)} = (x+2)(x+1) \frac{f_g(x+2)}{f_g(x)} \\ &= (x+2)(x+1) \frac{f_g(x+2)}{f_g(x+1)} \frac{f_g(x+1)}{f_g(x)} = \mathbb{E}(\theta|x+1) \cdot \mathbb{E}(\theta|x) \end{aligned}$$

Therefore, we have shown that the above equation holds.

Problem 6.3

Instead of (6.8), assume $g(\theta) = (1/\sigma)e^{-\theta/\sigma}$ for $\theta > 0$.

- Numerically find the maximum likelihood estimate $\hat{\sigma}$ for the Poisson model (6.1) fit to the count data in Table 6.1.
- Calculate the estimates of $\hat{E}\{\theta|x\}$, as in the third row of Table 6.1.

$$g(\theta) = \frac{\theta^{\nu-1} e^{-\theta/\sigma}}{\sigma^\nu \Gamma(\nu)}, \text{ for } \theta \geq 0, \quad (6.8)$$

Solution

- We'll find the maximum likelihood estimates of σ . The marginal probability density function of x is

$$f_g(x) = \int_0^\infty \frac{\theta^x e^{-\theta}}{x!} g(\theta) d\theta = \int_0^\infty \frac{\theta^x e^{-\theta}}{x!} (1/\sigma) e^{-\theta/\sigma} d\theta = \frac{\gamma^{x+1}}{\sigma} \int_0^\infty \frac{1}{\Gamma(x+1) \gamma^{x+1}} \theta^x e^{-\theta/\gamma} d\theta, \text{ where } \gamma = \frac{\sigma}{\sigma+1}.$$

Since the function in the integral follows Gamma($x+1$, γ) distribution, $f_g(x) = \frac{\gamma^{x+1}}{\sigma} = \frac{\sigma^x}{(\sigma+1)^{x+1}}$. So the marginal likelihood function of x is $L(\sigma) = \prod_{i=1}^N f_g(x_i)$. Define the log likelihood function $l(\sigma) = \log(L(\sigma)) = \sum_{i=1}^N (x_i \log \sigma - (1+x_i) \log(\sigma+1))$. The score function is $S(\sigma) = \dot{l}(\sigma) = -\frac{n}{\sigma+1} + \frac{n\bar{x}}{\sigma} - \frac{n\bar{x}}{\sigma+1}$ where $\bar{x} = \sum_{i=1}^N x_i$. So we'll find the mle of σ satisfying $S(\sigma) = 0$.

$$S(\sigma) = 0 \Leftrightarrow -n\sigma + n\bar{x}(\sigma+1) - n\bar{x}\sigma = 0 \Leftrightarrow -n\sigma + n\bar{x} = 0 \Leftrightarrow \sigma = \bar{x}$$

Therefore, the mle of σ is \bar{x} , denoted $\hat{\sigma} = \bar{x}$. So, we plug $\hat{\sigma}$ in $f_g(x)$. Then,

$$f_{\hat{g}}(x) = \frac{\hat{\sigma}^x}{(\hat{\sigma} + 1)^{x+1}}$$

Actually, $\hat{\sigma} = \bar{x} = \frac{1}{9461} \sum_{i=1}^{9461} x_i = \frac{1}{9461} (1317 \cdot 1 + 239 \cdot 2 + 42 \cdot 3 + 14 \cdot 4 + 4 \cdot 5 + 4 \cdot 6 + 1 \cdot 7) = 0.2143537$

Then, $\hat{y}_x = 9461 f_{\hat{g}}(x)$ where $x = 0, 1, \dots, 7$ By using R, we can calculate $\forall y_x$.

In summary,

\hat{y}_x	Fitting value
\hat{y}_0	7790.976
\hat{y}_1	1375.237
\hat{y}_2	242.7523
\hat{y}_3	42.84982
\hat{y}_4	7.563708
\hat{y}_5	1.33512
\hat{y}_6	0.235671
\hat{y}_7	0.04159986

- (b) Since $\hat{E}\{\theta|x\} = (x+1) \frac{f_{\hat{g}}(x+1)}{f_{\hat{g}}(x)} = (x+1) \frac{\hat{\sigma}}{1+\hat{\sigma}}$. So we can calculate the estimates of posterior mean for all $x = 0, 1, \dots, 6$

$\hat{E}(\theta x)$	Fitting value
$\hat{E}(\theta x=0)$	0.1765167
$\hat{E}(\theta x=1)$	0.3530333
$\hat{E}(\theta x=2)$	0.5295500
$\hat{E}(\theta x=3)$	0.7060667
$\hat{E}(\theta x=4)$	0.8825833
$\hat{E}(\theta x=5)$	1.0591000
$\hat{E}(\theta x=6)$	1.2356167
$\hat{E}(\theta x=7)$.

The R code related to the above is attached to the last page.

Problem 6.7

The nodes data of Section 6.3 consists of 844 pairs (n_i, x_i) .

- Plot x_i versus n_i
- Perform a cubic regression of x_i versus n_i and add it to the plot.
- What would you expect the plot to look like if the values of n_i were assigned randomly before surgery?

Solution

Problem 2

Show that the marginal distribution of "x" (in the missing species problem) is negative binomial.

Problem 3

Verify $E(t) = e_1 \frac{1-(1-\gamma t)^{-\nu}}{\gamma \nu}$ where $\gamma = (1/\sigma + 1)^{-1}$. (See the lecture note Chapter 6-2, page 5, parametric

approach)