STT 465
Single-parameter models: Beta-Binomial & Poisson

(G. de los Campos)

Beta-Binomial

OUTLINE:

$\Longrightarrow$ Elements of the models:

(1) Sampling model $p(Y|\theta)$

(2) Prior distribution $p(\theta)$

(3) From (1) and (2) and using Bayes Rule, we derive the posterior distribution of the parameter given the data, $p(\theta|Y)$

(4) In this case the posterior distribution has a recognizable form.

$\Longrightarrow$ Inference

- Posterior mean and posterior variance
- Is the Bayesian estimator unbiased?
- What happens as sample size increases?
  (consider both the effects on bias and variance)
- Posterior credibility regions (interpretation, types,…)

.

Beta-Binomial

⟹ Sampling model

$$p\left(y_i \mid \theta\right) = \theta^{y_i}\left(1-\theta\right)^{1-y_i}$$

[Assuming IID]

$$p\left(y_1,\ldots,y_n \mid \theta\right) = \prod_{i=1}^{n} \theta^{y_i}\left(1-\theta\right)^{1-y_i} = \theta^{n_1}\left(1-\theta\right)^{n_0} \qquad n_1 = \sum_{i=1}^{n} y_i \;\; ; \;\; n_0 = n - n_1$$

⟹ Prior: we will consider a Beta distribution (the uniform is a special case)

$$p\left(\theta \mid \alpha,\beta\right) = \frac{\theta^{\alpha-1}\left(1-\theta\right)^{\beta-1}}{B\left(\alpha,\beta\right)} \quad ; \; B\left(\alpha,\beta\right) = \int_0^1 x^{\alpha-1}\left(1-x\right)^{\beta-1} dx = \frac{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}{\Gamma\left(\alpha+\beta\right)}$$

$$p\left(\theta \mid \alpha,\beta\right) = \frac{\Gamma\left(\alpha+\beta\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}\theta^{\alpha-1}\left(1-\theta\right)^{\beta-1}$$

⟹ Discuss: Kernel vs Integrating constant.

Posterior Distribution

⇒ According to Bayes' rule

$$p(\theta \mid y_1,\ldots,y_n) = \frac{p(y_1,\ldots,y_n \mid \theta)\, p(\theta)}{p(y_1,\ldots,y_n)} \propto p(y_1,\ldots,y_n \mid \theta)\, p(\theta)$$

$$p(\theta \mid y_1,\ldots,y_n) \propto \theta^{n_1}(1-\theta)^{n_0}\frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)} \propto \theta^{\alpha+n_1-1}(1-\theta)^{\beta+n_0-1}$$

⇒ Finding the integrating constant

$$\int_0^1 \frac{\theta^{\alpha+n_1-1}(1-\theta)^{\beta+n_0-1}}{c(y)}d\theta = 1$$

$$\frac{1}{c(y)}\int_0^1 \theta^{\alpha+n_1-1}(1-\theta)^{\beta+n_0-1}d\theta = 1$$

$$\int_0^1 \theta^{\alpha+n_1-1}(1-\theta)^{\beta+n_0-1}d\theta = c(y) = B(\alpha+n_1,\beta+n_0,)$$

Posterior Distribution

⇒ Therefore

$$p\left(\theta \mid y_1,\ldots,y_n\right)=\frac{p\left(y_1,\ldots,y_n \mid \theta\right)p\left(\theta\right)}{p\left(y_1,\ldots,y_n\right)}\propto p\left(y_1,\ldots,y_n \mid \theta\right)p\left(\theta\right)$$

$$p\left(\theta \mid y_1,\ldots,y_n\right)=\frac{\theta^{\alpha+n_1-1}\left(1-\theta\right)^{\beta+n_0-1}}{B(\alpha+n_1,\beta+n_0)}=Beta\left(\theta \mid \alpha+n_1,\beta+n_0\right)$$

⇒ Because the posterior distribution has the same form as that of the prior, we say that the Beta prior is Conjugate to the Binomial likelihood.

⇒ Discuss:

- Posterior Vs. Prior Mean
- Posterior variance Vs sample size
- Posterior credibility regions (definition & interpretation).

Poisson Model: Likelihood Analyses

⇒ Data: count (y_i=0,1,2,....)

⇒ Probability model $p\left(y_i \mid \theta\right) = \dfrac{\theta^{y_i} e^{-\theta}}{y_i!} \qquad > 0$

[Question: What part is the kernel and what part is the integrating constant?]
[Question: How would you find the integrating constant if are give the kernel?]

⇒ Expected value and variance: $E\left(y_i\right) = Var\left(y_i\right) = \theta$
[Question: how would you determine the E[] and Var[]?]
[Question: what is the coef. of variation sd(y)/|E(y)|?]

⇒ Joint distribution of n iid draws from a Poisson model (Sampling Model)

$$p\left(y_1,\ldots,y_n \mid \theta\right) = \prod_{i=1}^{n} \frac{\theta^{y_i} e^{-\theta}}{y_i!} = \left[e^{-n\theta}\theta^{\sum_{i=1}^{n} y_i}\right]\prod_{i=1}^{n} \frac{1}{y_i!} = e^{-n\theta}\theta^{n\bar{y}}c(y)$$

[Discuss: sufficient statistic]
[Discuss: how to get MLE of theta?; Does it have a closed form?]
[Discuss: what is the sampling variance of the MLE estimator?]
[Discuss: how would you provide an approximate 95% CI for the MLE estimate?]

Poisson Model: In search for a conjugate prior

⟹ Likelihood

$$p(y_1,\ldots,y_n \mid \theta) \propto e^{-n\theta}\theta^{n\bar{y}}$$

⟹ We may guess that a conjugate prior may have this form:

$$p(\theta) \propto e^{-\tau_1\theta}\theta^{\tau_2}$$

⟹ If we have something like that, the posterior will have the following kernel

$$\begin{aligned} p(\theta \mid y_1,\ldots,y_n) &\propto p(y_1,\ldots,y_n \mid \theta)\,p(\theta) \\ &\propto e^{-n\theta}\theta^{n\bar{y}} \times e^{-\tau_1\theta}\theta^{\tau_2} \\ &\propto e^{-\theta(\tau_1+n)}\theta^{\tau_2+n\bar{y}} \end{aligned}$$

⟹ We have such a form in the Gamma distribution (a=shape; b=rate)

$$p(\theta) = \frac{b^a}{\Gamma(a)}\theta^{(a-1)}e^{-\theta b}$$

[Discuss: (i) alternative parameterizations, (ii) mean and variance]

Poisson Model: Posterior Density

$\Rightarrow$ According to Bayes Rule

$$p(\theta \mid y_1,\ldots,y_n) = \frac{p(y_1,\ldots,y_n \mid \theta)\, p(\theta)}{p(y_1,\ldots,y_n)}$$

$$\propto p(y_1,\ldots,y_n \mid \theta)\, p(\theta)$$

$$p(\theta) = \frac{b^a}{\Gamma(a)} e^{-\theta(a-1)} \theta^{b}$$

$$p(y_1,\ldots,y_n \mid \theta) \propto e^{-n\theta} \theta^{n\bar{y}}$$

$\Rightarrow$ In the Gamma-Poisson Model we have

$$p(\theta \mid y_1,\ldots,y_n) \propto \left[e^{-n\theta} \theta^{n\bar{y}} \right] \times \left[\theta^{(a-1)} e^{-\theta b} \right]$$

$$\propto e^{-\theta(n+b)}\ \theta^{n\bar{y}+a-1}$$

$\Rightarrow$ This can be recognized as the kernel of a Gamma distribution with the following parameters

rate: $\tilde{a} = n\bar{y} + a$

shape: $\tilde{b} = n + b$

Poisson Model: Posterior Distribution

$\Rightarrow$ Posterior Mean

$$E\left(\theta \mid y_1,\ldots,y_n\right) = \frac{n\bar{y} + a}{n + b}$$

$\Rightarrow$ Posterior Variance

$$V\left(\theta \mid y_1,\ldots,y_n\right) = \frac{n\bar{y} + a}{\left(n + b\right)^2}$$

[Discuss: what happens as n$\rightarrow \infty$?]