LECTURE SEVEN

PLAN FOR TODAY

1. Basic facts about Beta Densities
2. Previous Lecture Recap
3. Beta priors & posteriors (Bernoulli or Binomial likelihoods)
4. Real Data Application

Beta Density

Beta (a, b)

$$f(u) = \frac{I\{0 \le u \le 1\}\, u^{a-1}(1-u)^{b-1}}{\int_0^1 v^{a-1}(1-v)^{b-1}\,dv}$$

Beta $(3, 5)$ $\quad u^2(1-u)^4$

1. $a = 1, b = 1$

   Beta $(1,1)$ = Uniform $\quad I\{0 \le u \le 1\}$

2. Mean corresponding to Beta (a, b)

   $$= \frac{a}{a+b}$$

Beta(1, 4) $\left(\text{mean} = \frac{1}{1+4} = 0.2\right)$

Beta(4, 1) $\left(\text{mean} = \frac{4}{4+1} = 0.8\right)$

③ Variance corresponding to Beta(a, b)

$$= \frac{a}{a+b} \times \frac{b}{a+b} \times \frac{1}{a+b+1}$$

Corollary: a+b large ⇒ variance small

⇓

skinny Beta

LAST LECTURE

Quality of a product based on reviews: Pos, Neg

Pos + Neg = Tot

θ

$$\text{Proportion of positive reviews} = \frac{\text{Pos}}{\text{Pos} + \text{Neg}}$$

→ ① Naive Estimate of θ

② Proportion

③ Naive Proportion

④ MLE ⑤ Frequentist Estimate

Likelihood: $\mathbb{P}(\text{observing the data} \mid \theta = u)$

$$= \underbrace{\theta^{Pos}(1-\theta)^{Neg}}$$

Prior: Uniform prior.

Uniform prior for θ + Bernoulli likelihood → Posterior Beta (Pos+1, Neg+1)

Uniform = Beta(1,1)

$$\text{mean} = \frac{Pos+1}{Tot+2}$$

Beta(1,1) prior + Bernoulli likelihood ⇒ Beta(1+Pos, 1+Neg) Posterior

Beta(a,b) prior + Bernoulli likelihood ⇒ Beta(a+Pos, b+Neg)

Eg: Very rare for the quality to be below 0.8. Typically quality is 0.9.

$$\frac{a}{a+b} = 0.9 \Rightarrow a = 9b$$

$$\text{standard deviation} = \sqrt{\frac{a}{a+b}\,\frac{b}{a+b}\,\frac{1}{a+b+1}} = 0.05$$

$$\sqrt{\frac{0.9 \times 0.1}{a+b+1}} = 0.05$$

$$\Rightarrow a+b+1 = \frac{0.09}{0.0025} = \frac{900}{25} = 36$$

$$\left.\begin{aligned} a+b &= 35 \\ a &= 9b \end{aligned}\right\}$$ Solve for a & b.

$b = 3.5$

$a = 31.5$

Beta $(31.5, 3.5)$

$\downarrow$

Beta $(31.5 + Pos,\ 3.5 + Neg)$

$$u^{a-1}(1-u)^{b-1} \times u^{Pos}(1-u)^{Neg}$$

Comparison Between Frequentist & Bayes Estimate

$\theta \longrightarrow \dfrac{Pos}{Pos+Neg}$ (Frequentist)

$\theta \downarrow$ Posterior

Beta$(a+Pos,\ b+Neg)$

Bayes Estimate = Posterior Mean

$$= \frac{a+Pos}{a+b+Pos+Neg}$$

Qn: When are $\frac{Pos}{Tot}$ & $\frac{a+Pos}{a+b+Tot}$ same or different?

$$\frac{a+Pos}{a+b+Tot} = \left(\frac{a+b}{a+b+Tot}\right)\frac{a}{a+b} + \left(\frac{Tot}{a+b+Tot}\right)\frac{Pos}{Tot}$$