STA 360/602L: MODULE 2.3

MARGINAL LIKELIHOOD AND POSTERIOR PREDICTION

DR. OLANREWAJU MICHAEL AKANDE



MARGINAL LIKELIHOOD

Recall that the marginal likelihood is

$$L(y)=f(y_1,\ldots,y_n)=\int_{\Theta}f(y_1,\ldots,y_n| heta)\pi(heta)\mathrm{d} heta=\int_{\Theta}L(heta|y)\pi(heta)\mathrm{d} heta.$$

• For clarity, when dealing with a single y instead of y_1, \ldots, y_n , we can write

$$L(y)=f(y)=\int_{\Theta}f(y| heta)\pi(heta)\mathrm{d} heta=\int_{\Theta}L(heta|y)\pi(heta)\mathrm{d} heta.$$

- When this is the case, for example in the case of the binomial distribution, I will often write
 - the marginal likelihood as L(y) or f(y), and
 - the sampling (conditional) likelihood as $f(y|\theta)$ or $L(\theta|y)$.



MARGINAL LIKELIHOOD

- What is the marginal likelihood for the beta-binomial?
- We have

$$egin{aligned} L(y) &= \int_{\Theta} p(y| heta) \pi(heta) \mathrm{d} heta \ &= \int_{0}^{1} inom{n}{y} heta^{y} (1- heta)^{n-y} rac{1}{B(a,b)} heta^{a-1} (1- heta)^{b-1} d heta \ &= inom{n}{y} rac{B(a+y,\,b+n-y)}{B(a,b)}, \end{aligned}$$

by the integral definition of the Beta function.

- Marginal likelihood for the beta-Bernoulli follows directly.
- Deriving the marginal likelihood for conjugate distributions is often relatively straightforward.



PRIOR PREDICTIVE DISTRIBUTION

- We may care about making predictions before we even see any data.
- This is often useful as a way to see if the sampling distribution we have chosen is appropriate, after integrating out all unknown parameters.
- The prior predictive distribution is

$$egin{aligned} p(y) &= \int_{\Theta} p(y, heta) \, d heta \ &= \int_{\Theta} p(y| heta) \cdot \pi(heta) \, d heta \end{aligned}$$

- In some sense, the prior predictive distribution marginalizes the sampling distribution (for a single y) over the prior.
- When dealing with a single y instead of y_1, \ldots, y_n , this is just the marginal likelihood of the data.



POSTERIOR PREDICTIVE DISTRIBUTION

- We often care about making predictions for new data points, given the current pbserved data.
- For example, suppose $y_1, \ldots, y_n \stackrel{iid}{\sim} \text{Bernoulli}(\theta)$.
- We may wish to predict a new data point y_{n+1} .
- We can do so using the posterior predictive distribution $p(y_{n+1}|y_{1:n})$.
- Why are we not including the parameter in the posterior predictive distribution?



POSTERIOR PREDICTIVE DISTRIBUTION

• Recall that we have conditional independence of the y's given θ .

■ So,

$$egin{aligned} p(y_{n+1}|y_{1:n}) &= \int_{\Theta} p(y_{n+1}, heta|y_{1:n}) \, d heta \ &= \int_{\Theta} p(y_{n+1}| heta,y_{1:n}) \cdot \pi(heta|y_{1:n}) \, d heta \ &= \int_{\Theta} p(y_{n+1}| heta) \cdot \pi(heta|y_{1:n}) \, d heta. \end{aligned}$$

 So, in some sense, the posterior predictive distribution marginalizes the sampling distribution over the posterior.



POSTERIOR PREDICTIVE DISTRIBUTION

• When we talk about the posterior predictive distribution for Bernoulli data, we are really referring to $p(y_{n+1} = 1|y_{1:n})$ and $p(y_{n+1} = 0|y_{1:n})$.

Then,

$$egin{aligned} p(y_{n+1} = 1 | y_{1:n}) &= \int_{\Theta} p(y_{n+1} = 1 | heta) \cdot \pi(heta | y_{1:n}) \, d heta \ &= & \dots \ &= & \dots \ &= & \dots \end{aligned}$$

which simplifies to what? To be done on the board!

- What then is $p(y_{n+1}=0|y_{1:n})$?
- Posterior predictive pmf therefore takes the form

$$p(y_{n+1}|y_{1:n})=rac{a_n^{y_{n+1}}b_n^{1-y_{n+1}}}{a_n+b_n}; \hspace{0.2cm} y_{n+1}=0,1.$$

• What are a_n and b_n ?



Going forward...

- From here on, we will often deal with multiple data points y_1, \ldots, y_n frequently.
- To make that obvious, we will write the Bayes rule as one of the following

$$egin{aligned} \pi(heta|y_1,\ldots,y_n) &= rac{\pi(heta)\cdot p(y_1,\ldots,y_n| heta)}{\int_{\Theta}\pi(heta)\cdot p(y_1,\ldots,y_n| heta)\mathrm{d} heta} \ \pi(heta|y_1,\ldots,y_n) &= rac{\pi(heta)\cdot p(y_1,\ldots,y_n| heta)}{p(y_1,\ldots,y_n)} \ \pi(heta|y) &= rac{\pi(heta)\cdot L(heta|y)}{L(y)}, \end{aligned}$$

where $y=(y_1,\ldots,y_n).$



WHAT'S NEXT?

Move on to the readings for the next module!

