Stat 134 lec 35

Warmup 8:00 – 8:10 AM

rule of average conditional probability:

ex $X \sim \text{Unif}(0,1)$

$I_1 | X=x, I_2 | X=x \overset{iid}{\sim} \text{Ber}(x)$

Find $P(I_2=1 \mid I_1=1)$

Are I_1 I_2 independent?

$$P(I_2=1 \mid I_1=1) = \frac{P(I_2=1, I_1=1)}{\underset{\overset{\parallel}{1/2}}{P(I_1=1)}}$$

$$P(I_2=1, I_1=1) = \int_{x=0}^{x=1} \underbrace{P(I_2=1, I_1=1 \mid X=x)}_{\underset{P(I_2=1|X=x)P(I_1=1|X=x)}{\parallel}} \overset{1}{\overbrace{f_X(x)}} dx = \int_0^1 x^2 dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}$$

$P(I_2=1|X=x) = x$, $P(I_1=1|X=x) = x$

$$P(I_2=1 \mid I_1=1) = \frac{1/3}{1/2} = \frac{2}{3} > \frac{1}{2}$$

$\frac{1}{2} = P(I_2=1)$

so I_2 and I_1 are dependent.

or

$$P(I_2=1 \mid I_1=1) = \int_{x=0}^{x=1} \underset{\underset{\overset{\parallel}{x}}{P(I_2=1|X=x)}}{\underset{\parallel}{P(I_2=1 \mid I_1=1, X=x)}} \underset{\underset{2x}{\parallel}}{f_{X|I_1=1}(x)} dx = \int_0^1 2x^2 dx = \left.\frac{2x^3}{3}\right|_0^1 = \frac{2}{3}$$

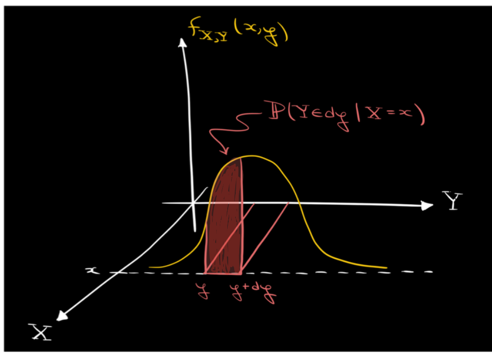
Stat 134 lec 35

Quiz 6 Wednesday Dec 4 Sec 5.3 – 6.3

Last time

sec 6.3 Conditional probability (continuous case).

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{\underbrace{f_X(x)}_{\leftarrow \text{ constant}}} \qquad \text{conditional density}$$



ex

$$P(Y\in dy \mid X=x) \approx f_{Y|X=x}(y)\,dy$$

$$P(Y\in dy) = \int_{x\in X} P(Y\in dy \mid X=x) f_X(x)\,dx$$

rule for average conditional probability

In frequentist statistics we interpret probability as a long run average constant known only to Tyche, the goddess of fortune.

In Bayesian statistics we interpret probability as a RV

ex When probability a coin lands head is a RV X rather than an unknown constant, we are doing Bayesian statistics. The posterior density, $f_{X|I_1=1}(x)$ is an example of a conditional density

Posterior $\propto$ likelihood $\cdot$ Prior

$$f_{X|I_1=1}(x) = \frac{\overbrace{P(I_1=1 \mid X=x)}^{\text{likelihood}} \cdot \overbrace{f_X(x)}^{\text{Prior}}}{\underbrace{P(I_1=1)}_{\text{constant}}}$$

Today Sec 6.3

1. Bayesian statistics
2. Conjugate Pairs

① Sec 6.3 Bayesian Stats

Review ~~Beta~~ Distribution

variable part,

$X \sim Beta(r,s)$

$$f_X(x) = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} x^{r-1}(1-x)^{s-1}, \quad 0 < x < 1$$

where for $r \in \mathbb{Z}^+$, $\Gamma(r) = (r-1)!$

ex If $0 < x < 1$,

$$f_X(x) \propto 1 \Rightarrow X \sim Beta(1,1)$$

$$f_X(x) \propto x \Rightarrow X \sim Beta(2,1)$$

$$f_X(x) \propto x(1-x) \Rightarrow X \sim Beta(2,2)$$

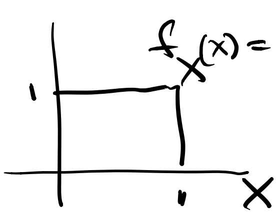
ex $X \sim Unif(0,1)$

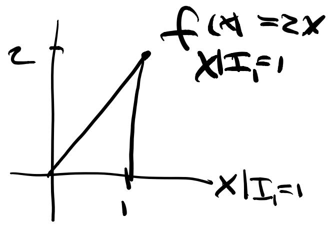
$I_1 | X=x, I_2 | X=x \overset{iid}{\sim} Ber(x)$

Prior density $f_X(x) = 1 \Rightarrow X \sim Beta(1,1)$

Posterior density $f_{X|I_1=1}(x) = 2x \Rightarrow X | I_1 = 1 \sim Beta(2,1)$

Prior $X \sim Unif(0,1)$ Posterior





ex Let A be an event and let $X \sim \text{Unif}(0,1)$

Suppose that conditional of $X=x$, A has probability x (i.e $P(A|X=x)=x$).

Find the conditional density of X given A^c, $f_{X|A^c}(x)$

soln

posterior

$$f_{X|A^c}(x) \propto \text{likelihood} \cdot \text{prior} = 1-x$$

$$\text{likelihood} = P(A^c|X=x) = 1-x, \qquad \text{prior} = f_X(x) = 1$$

$$\boxed{\begin{aligned} &X \sim \text{Beta}(r,s) \\ &f_X(x) = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} x^{r-1}(1-x)^{s-1}, \quad 0<x<1 \end{aligned}}$$

variable part.

$$\Rightarrow X|A^c \sim \text{Beta}(1,2) \Rightarrow f_{X|A^c}(x) = \frac{\Gamma(3)(1-x)}{\Gamma(2)\Gamma(1)}$$

$\Gamma(3) = 2$, $\Gamma(2) = 1$, $\Gamma(1) = 1$

tinyurl: http://tinyurl.com/nov25-pt1

Stat 134

Monday April 23 2019

1. Let A, B and C be events and let X be a random variable uniformly distributed on (0,1). Suppose conditional on X=x, that A, B, and C are independent each with probability x. The conditional density of X given that A and B occurs and C doesn't is:

   a $Beta(2,2)$

   b $Beta(3,2)$

   c $Beta(2,3)$

   d none of the above

$f_{X|ABC^c}(x) \propto$ likelihood · prior

$\propto P(ABC^c | X=x) \cdot f_X(x)$

$\propto \underbrace{P(A|X=x)}_{=x} \underbrace{P(B|X=x)}_{=x} \underbrace{P(C^c|X=x)}_{=1-x} \underbrace{f_X(x)}_{=1} = x^2(1-x)$

$\Rightarrow \boxed{Beta(3,2)}$

The posterior can be difficult to calculate. One situation where it is very easy to calculate is when the variable part of the likelihood and prior are similar so that the posterior distribution is the same as the prior distribution.

② Sec 6.3 conjugate pairs

When the prior and the posterior belong the same distribution family we say that the prior and the likelihood are conjugate. The prior and likelihood being conjugate means their variable parts are similar.

ex prior $X \sim beta(r, s)$

likelihood $Y \sim Bin(n, x)$

Posterior $\propto$ likelihood $\cdot$ prior

$$f_{X|Y=j}(x) \propto P(Y=j|X=x)\, f_X(x)$$

$$\underbrace{\overbrace{x^j(1-x)^{n-j}}^{\text{similar}} \cdot \overbrace{x^{r-1}(1-x)^{s-1}}}_{\displaystyle = x^{j+r-1}(1-x)^{n-j+s-1}}$$

$\Rightarrow X|Y=j \sim Beta(j+r, n-j+s)$

Conclusion: beta for prior and Binomial for likelihood is a conjugate pair since the posterior is beta.

ex Suppose $\Theta \sim \text{Gamma}(r, \lambda)$ with r, λ known.

Let $(N_1|\Theta=\theta, N_2|\Theta=\theta, N_3|\Theta=\theta) \overset{iid}{\sim} \text{Pois}(\theta)$.

Find the posterior distribution of Θ.

Gamma: $f_\Theta(\theta) \propto \theta^{r-1} e^{-\lambda\theta}$

Poisson: $P(N=n|\Theta=\theta) \propto e^{-\theta}\theta^{n}$

$$f_{\Theta|N_1=n_1, N_2=n_2, N_3=n_3}(\theta) \propto \text{likelihood} \cdot \text{Prior}$$

$$\propto P(N_1=n_1, N_2=n_2, N_3=n_3 | \Theta=\theta) f_\Theta(\theta)$$

$$\propto \frac{e^{-\theta}\theta^{n_1}}{n_1!} \frac{e^{-\theta}\theta^{n_2}}{n_2!} \frac{e^{-\theta}\theta^{n_3}}{n_3!} \theta^{r-1} e^{-\lambda\theta}$$

$$\propto \theta^{(n_1+n_2+n_3)+r-1} e^{-(3+\lambda)\theta}$$

$$\sim \text{Gamma}(n_1+n_2+n_3+r, 3+\lambda)$$

$\Rightarrow$ Gamma and Poisson are conjugate pairs with posterior gamma.

If the likelihood is Poisson we will often try and model the prior to be gamma since then the posterior will be gamma and it is easy to see what the parameters are.

Furthermore, if we collect more data (the likelihood remains Poisson) and our old posterior can become our new prior. The new posterior will necessarily be a Gamma with different parameters.