Stat 134 lec 37 (no lec 36)

Warmup 1:00-1:10

The multiplication rule is

$$f(x,y) = f_{Y|X=x}(y)\, f_X(x)$$

let $X \sim \text{Gamma}(2,\lambda)$

$Y|X=x \sim \text{Unif}(0,x)$

a) Find $f_{Y|X=x}(y) = \frac{1}{x}$ $\quad 0<y<x<\infty$

b) Find $f(x,y)$

$X \sim \text{Gamma}(2,\lambda)$

$f_X(x) = \lambda^2 x e^{-\lambda x} \quad 0<x<\infty$

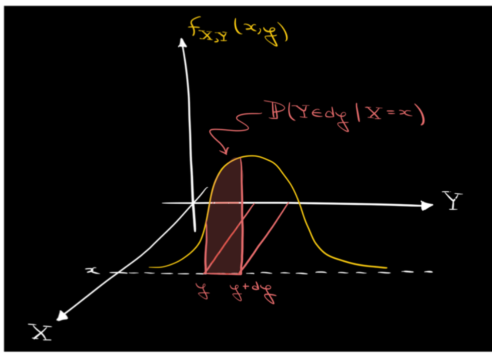
$f(x,y) = \frac{1}{x} \cdot \lambda^2 x e^{-\lambda x} = \lambda^2 e^{-\lambda x} \quad 0<x<\infty$

Last time

sec 6.3 Conditional probability (continuous case).

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

$$f(x,y) = f_{Y|X=x}(y)\, f_X(x)$$ multiplication rule



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

discrete case:

$$P(A) = \sum_{x \in X} P(A \mid X = x) P(X = x)$$

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

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

Find $P(I_1 = 1) = \int_{x=0}^{x=1} \underbrace{P(I_1 = 1 \mid X = x)}_{x} \underbrace{f_X(x)}_{1}\, dx$

$$= \int_0^1 x\, dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2}$$

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

$$P(I_2=1 \mid I_1=1) = \frac{P(I_2=1, I_1=1)}{\underset{\overset{\shortparallel}{1/2}}{P(I_1=1)}} = \frac{1/3}{1/2} = \boxed{2/3}$$

$$P(I_2=1, I_1=1) = \int_0^1 \underset{\overset{\shortparallel}{P(I_2=1 \mid X=x)\,P(I_1=1 \mid X=x)}}{P(I_2=1, I_1=1 \mid X=x)} \cdot \underset{\overset{\shortparallel}{1}}{f_X(x)}\,dx = \int_0^1 x^2\,dx = 1/3$$

$P(I_2=1 \mid X=x) = x$, $\quad P(I_1=1 \mid X=x) = x$

Are I_1, I_2 independent? NO

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, $f_{X \mid I_1=1}(x)$ is an example of a conditional density

conditional prob mass function (discrete)

$$P(I_1=1, X=x) = P(I_1=1 \mid X=x) f_X(x)$$
$$\shortparallel$$
$$f_{X \mid I_1=1}(x) \cdot P(I_1=1) \quad \Rightarrow$$

Posterior $\propto$ likelihood $\cdot$ Prior

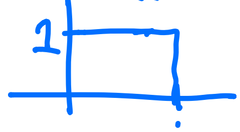
$$f_{X \mid I_1=1}(x) = \frac{\overbrace{P(I_1=1 \mid X=x)}^{\text{likelihood}} \cdot \overbrace{f_X(x)}^{\text{Prior}}}{\underbrace{P(I_1=1)}_{\text{constant}}}$$

CAUTION X is continuous and I_1 is discrete

We write $P(I_1 \mid X=x)$ for conditional probability mass function (pmf) of I_1 and $f_{X \mid I_1=1}(x)$ for the conditional density of X

ex Find $f_{X \mid I_1=1}(x) = \dfrac{x \cdot 1}{1/2} = 2x$

$f_X(x)$



$f_{X \mid I_1=1}(x) = 2x$

Today Sec 6.3

1. Bayesian statistics
2. Conjugate Pairs

① ~~Sec 6.3~~ Bayesian Stats

Review ~~Beta~~ Distribution

variable part.

$X \sim \text{Beta}(r, s)$

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

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

ex If $0 < x < 1$,

$f_X(x) \propto 1 \qquad X \sim \text{Beta}(1,1)$

$f_X(x) \propto x \qquad X \sim \text{Beta}(2,1)$

$f_X(x) \propto x(1-x) \qquad X \sim \text{Beta}(2,2)$

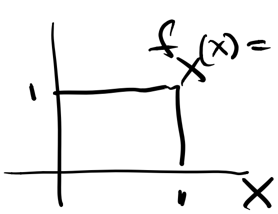
ex $X \sim \text{Unif}(0,1)$

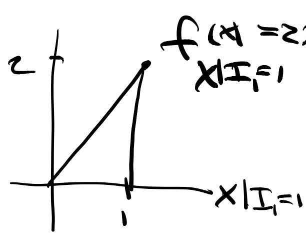
$I_1 | X = x, I_2 | X = x \overset{iid}{\sim} \text{Ber}(x)$

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

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

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





ex Let A be an event and
$X \sim \text{Unif}(0,1)$
Suppose $P(A \mid X=x) = x$
Find $f_{X|A^c}(x)$

$X \sim \text{Beta}(r,s)$

variable part.

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

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

$$= P(A^c \mid X=x) \cdot f_X(x) = 1-x$$

$$P(A^c \mid X=x) = 1-x, \quad f_X(x) = 1$$

var part of density

$X \mid A^c \sim \text{Beta}(1,2)$

tinyurl: http://tinyurl.com/nov23-pt1

Stat 134

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

$X \mid ABC^c \sim \ ?$

$f_{X\mid ABC^c} \propto$ likelihood · prior

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

$\propto x^2(1-x) \cdot 1$

$\Rightarrow X \mid ABC^c \sim$ 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 $X \sim \text{beta}(r, s)$

likelihood $Y \sim \text{Bin}(n, x)$

Posterior $\propto$ likelihood $\cdot$ prior

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

similar

$$= \underbrace{x^{j}(1-x)^{n-j} \cdot x^{r-1}(1-x)^{s-1}}$$

$$= x^{j+r-1}(1-x)^{n-j+s-1}$$

$$\Rightarrow X \mid Y=j \sim \text{Beta}(j+r,\, n-j+s)$$

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

$$f_{\Theta|N_1=n_1, N_2=n_2, N_3=n_3} \propto \text{likelihood} \cdot \text{prior}$$

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

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

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

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

Conclusion: Gamma for prior and Poisson for likelihood is a conjugate pair.

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.