Stat 134 lec 35

Last time

Sec 6.2 Conditional expectation (discrete case)

X discrete.

$$g(x) = E(Y|X=x) \quad \sum_{y \in Y} y P(Y=y|X=x)$$

function of X

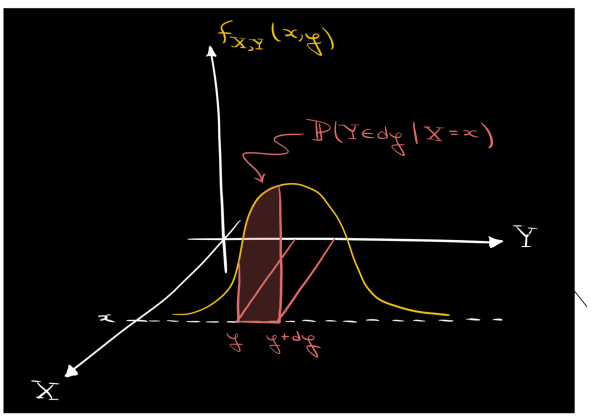
$E(Y|X)$ is a RV

$E(Y) = E(E(Y|X))$ law of iterated expectation

sec 6.3 Conditional probability (continuous case).

$$f_{Y|X=x}(y) = \frac{f_{(X,Y)}(x,y)}{f_X(x)}$$ ← constant

conditional density



ex

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

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

integral conditioning formula.

Today ① Go over student explanations from Concept test
② Sec 6.3 Bayesian statistics
Posterior density.
Conjugate Pairs ← Beta, Binomial / Gamma, Poisson

1

Stat 134

Wednesday November 7 2018

1. Let N have a Poisson(μ) distribution. Suppose that given $N = n$, random variable X follows a Binomial(n,p) distribution. E(X) is:

   a np

   b μ

   c $n\mu$

   d none of the above

Take a minute and discuss how you solved this.

<table>
<tr><td>d</td><td></td></tr>
<tr><td>d</td><td>Taking the sum of npP(N=n), we can rearrange terms so that we get e^{-\mu}\mu p times the Taylor series for e^{\mu}. Cancelling terms yields \mu p.</td></tr>
<tr><td>d</td><td>1.) We know from before that X follows Poison(mu*p)
2.) E[X] = E[E(X|N)] = pE[n] = p*mu</td></tr>
<tr><td>c</td><td>the expectation of X given N = n is n. the expectation of N is mu. therefore the expectation of X depends entirely on the distribution of N. The expectation of X should therefore be directly proportional to N, with the peak of the expectation being when N = mu. The value of the graph at that point would be mu * n</td></tr>
<tr><td>d</td><td></td></tr>
<tr><td>a</td><td>E(X) = E(X=x| N=n). when you expand this, eventually the mu's cancel out and you are left with the formula for expectation of a binomial distribution, which is np</td></tr>
<tr><td>d</td><td>μp sum of npe (-μ) μ (n) /n!</td></tr>
<tr><td>d</td><td>Cant have n in value.</td></tr>
<tr><td>d</td><td>Plug into formula (for given n, E(X) is np), then rearrange sum, which becomes Taylor's series for e^mu, entire thing simplifies to mu*p</td></tr>
<tr><td></td><td>If N=n then E(X) = np conditioned my the expecttation of N which is m. So E(X) is np/m</td></tr>
<tr><td>d</td><td>Not too sure, need more practise</td></tr>
<tr><td>d</td><td>We should be taking p and mu into account</td></tr>
<tr><td>d</td><td>up</td></tr>
<tr><td></td><td></td></tr>
<tr><td>d</td><td></td></tr>
<tr><td>d</td><td>miu*p</td></tr>
<tr><td>d</td><td></td></tr>
<tr><td>d</td><td>E(x) = E(E(X/N)) Since for a given n, we have E(X) = np, but n here is a R.V. that has a expectation of u, so the final answer should be up.</td></tr>
<tr><td>d</td><td>Sum of np*P(N=n) over all n. Taking p out, we have expectation of poisson. So it should be p*mu</td></tr>
<tr><td>d</td><td>Doesn't account for poisson distribution of N</td></tr>
<tr><td>d</td><td>E(X)=E(X|N=n)=p/miu</td></tr>
<tr><td>c</td><td>Expectation of poisson is mu and by the binomial distribution E(X) = np which makes the conditional expectation =n * mu</td></tr>
<tr><td>d</td><td></td></tr>
<tr><td>a</td><td></td></tr>
<tr><td>d</td><td></td></tr>
<tr><td>c</td><td>np is too obvious, no way it's just mew, so that leaves 2 options, and in this class it's rarely none of the above. Therefore nmew</td></tr>
<tr><td>d</td><td></td></tr>
<tr><td>d</td><td>You do magical algebra to get p*mu^2</td></tr>
<tr><td>a</td><td>Lanlan told me its right</td></tr>
<tr><td>a</td><td>Andy told me it's A</td></tr>
<tr><td>a</td><td></td></tr>
<tr><td>d</td><td>E(X) =E(E(X|N))=E(E(Bin(n,p))=E(np) = SUM(P(N=n) np) = mu*p</td></tr>
<tr><td>a</td><td>E (Np) =Np =n p</td></tr>
<tr><td>d</td><td>E(X|N=n)=E(X, N=n)/P(N=n) = np/(mu^n * e^{-mu} / n!)</td></tr>
<tr><td>d</td><td>E(X) = E(E(X|N=n)). E(X|N=n) = sum(x*P(X|N=n)) = sum_n(np) = Np.
E(Np) = pE(N) (p constant) = mu*p</td></tr>
<tr><td>d</td><td></td></tr>
<tr><td>d</td><td></td></tr>
<tr><td>d</td><td>should be up</td></tr>
<tr><td>d</td><td>E(X | N) = Np. E(Np) = pE(N). E(N) = mu. Therefore E(X) = pmu</td></tr>
<tr><td>d</td><td></td></tr>
</table>

$$g(n) = E(X|N=n) = np$$

$$E(g(n)) = \sum_{n=1}^{\infty} g(n) \frac{e^{-\mu}\mu^n}{n!}$$

$$= p \sum_{n=1}^{\infty} \frac{n e^{-\mu}\mu^n}{n!}$$

$$= \mu p \sum_{n=1}^{\infty} \frac{e^{-\mu}\mu^{n-1}}{(n-1)!} = \boxed{\mu p}$$

$= 1$

$$E(X|N) = Np$$

$$E(Np) = \boxed{\mu p}$$

② Review ~~Beta~~ distribution

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

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

variable part.

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

ex For $0<p<1$ if

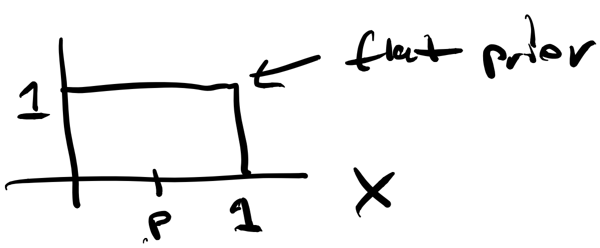
$$f_X(p) \propto 1 \Rightarrow X \sim \text{Beta}(1,1)$$

$$f_X(p) \propto p \Rightarrow X \sim \text{Beta}(2,1)$$

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

sec 6.3 Bayesian Statististics

ex imagine you have a prob p coin. Instead of thinking of p as a fixed constant, p has a distribution $X \sim \text{Unif}(0,1)$, called a flat prior. Allowing our parameter p to have a distribution is called Bayesian statistics



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

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

$$P(I_1=1) = \int_{p=0}^{p=1} P(I_1=1 \mid X=p)\,1\,dp = \int_0^1 p\,dp = \frac{p^2}{2}\Big|_0^1 = \boxed{\tfrac{1}{2}}$$

integral conditioning formula,

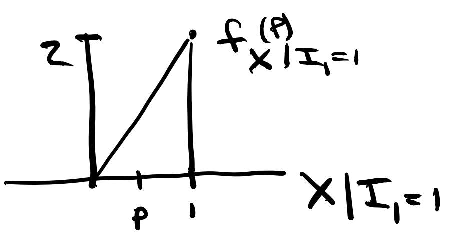
Knowing about conditional densities allows us to compute the posterior distribution where we update the distribution of p given that $I_1=1$.

posterior $\propto$ likelihood $\cdot$ prior

$$\Rightarrow f_{X|I_1=1}(p) = \overset{\text{constant}}{C}\ \overset{\text{likelihood}}{P(I_1=1 \mid X=p)}\ \overset{\text{prior}}{f_X(p)} = C \cdot p \cdot 1$$

$$\Rightarrow X|I_1=1 \sim \text{Beta}(2,1) \Rightarrow C = \frac{\Gamma(3)}{\Gamma(2)\Gamma(1)} = 2$$

$$\Rightarrow f_{X|I_1=1}(p) = 2p$$



Compare with flat prior above. After getting one head you think prob of getting another head is greater than 50%.

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

given $X=p$, let $I_1, I_2 \overset{iid}{\sim} \text{Ber}(p)$

If $I_1 = 1$, what is prob ~~that~~ $I_2 = 1$?

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

Note $I_1 \mid X=p$ and $I_2 \mid X=p$ are independent.

But is I_1 and I_2 independent?

recall,

$P(A) = \int_{x \in X} P(A \mid X=x) f_X(x)\,dx$ integral conditioning formula

$$P(I_2=1 \mid I_1=1) = \frac{P(I_2=1, I_1=1)}{P(I_1=1)}$$

$$= \frac{\int_{p=0}^{p=1} P(I_2=1, I_1=1 \mid X=p) f_X(p)\,dp}{\int_{p=0}^{p=1} P(I_1=1 \mid X=p) f_X(p)\,dp}$$

$$= \frac{\displaystyle\int_{p=0}^{p=1} \overbrace{P(I_2=1|X=p)}^{=p}\,\overbrace{P(I_1=1|X=p)}^{=p}\,\overbrace{f_X(p)}^{=1}\,dp}{\displaystyle\int_{p=0}^{p=1} \underbrace{P(I_1=1|X=p)}_{=p}\, f_X(p)\,dp}$$

$$= \frac{\left.\frac{p^3}{3}\right|_0^1}{\left.\frac{p^2}{2}\right|_0^1} = \frac{\frac{1}{3}}{\frac{1}{2}} \to 2/3$$

mult rule

so $P(I_1=1, I_2=1) = P(I_2=1|I_1=1)\,P(I_1=1)$

$= \frac{2}{3}\cdot\frac{1}{2} = 1/3.$

not equal.

$P(I_1=1)P(I_2=1) = \frac{1}{2}\,\frac{1}{2} = 1/4$

$\Rightarrow I_1, I_2$ not independent.

ex Let A and B be events and let $Y \sim U(0,1)$.
Suppose that conditional of $Y=p$, A and B are independent, each with probability p.
Find the conditional probability of A given that B occurs.

Soln

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\int_0^1 P(AB \mid Y=p) f_Y(p)\,dp}{\int_0^1 P(B \mid Y=p) f_Y(p)\,dp}$$

$$= \frac{\int_0^1 p^2\,dp}{\int_0^1 p\,dp} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

Soln

Stat 134

Wednesday November 14 2018

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

   a $beta(2,1)$

   b $beta(2,2)$

   c $beta(3,2)$

   d none of the above

$f_{Y|AB^c}(p) \propto$ likelihood $\cdot$ prior

$= c\, P(AB^c \mid Y=p) \cdot \underbrace{f_Y(y)}_{=1}$

$= c\, P(A \mid Y=p) P(B^c \mid Y=p) = c\, p(1-p)$

$\Rightarrow Y|AB^c \sim \boxed{\text{Beta}(2,2)}$

Let A and B be events and let Y be a random variable uniformly distributed on (0,1). Suppose that conditional on Y=p, A and B are independent each with probability p. The conditional density of Y given that A and B occurs is:

a $beta(2,1)$

b $beta(2,2)$

c $beta(3,2)$

d none of the above

$f_{Y|AB}(p) = c P(A|Y=p) P(B|Y=p) = cp^2 = c p^{3-1} (1-p)^{1-1}$

$\Rightarrow Y|AB \sim \boxed{Beta(3,1)}$

Sec 6.3 Conjugate pairs

When the prior and the posterior belong the same family we say that the prior and the likelihood are conjugate. This allows you to easily update your prior.

ex (see lec 29)

we have a p-coin with unknown probability p. We let p have distribution $X \sim \text{Beta}(r, s)$. We flip the coin n times and observe k heads. Find the posterior distribution of p.

Conjugate pair $\begin{cases} X \sim \text{beta}(r,s) & \text{prior} \\ S_n \sim \text{Bin}(n,p) & \text{likelihood} \end{cases}$ — # successes in n Bernoulli trials.

$$f_{X|S_n=k}(p) \propto \underset{\binom{n}{k} p^k (1-p)^{n-k}}{\text{likelihood}} \cdot \underset{\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} p^{r-1}(1-p)^{s-1}}{\text{prior}}$$

$$= c\, p^{k+r-1}(1-p)^{n-k+s-1} \Rightarrow X|S_n=k \sim \text{Beta}(k+r, n-k+s)$$

beta$(k+r, n-k+s)$

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

eg

Let $N=(N_1,\dots,N_5)\overset{iid}{\sim}\text{Pois}(\Theta)$ for unknown Θ.

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

Find the posterior distribution of Θ.

Soln

$\Theta \sim \text{Gamma}(r,\lambda)$ prior

$N_1 \sim \text{Poisson}(\Theta)$ likelihood.

$$f_{N|\Theta}(n_1,\dots,n_5) = f_{N_1|\Theta}(n_1)\cdots f_{N_5|\Theta}(n_5)$$

should write $p_{N|\Theta}(n_1,\dots,n_5)$

$$= \frac{e^{-\theta}\theta^{n_1}}{n_1!}\cdots\frac{e^{-\theta}\theta^{n_5}}{n_5!}$$

$$= \frac{e^{-5\theta}\theta^{n_1+\dots+n_5}}{n_1!\,n_2!\cdots n_5!}$$

likelihood prior

$$f_{\Theta|N}(\theta) \propto e^{-5\theta}\theta^{n_1+\dots+n_5}\;\theta^{r-1}e^{-\lambda\theta}$$

$$\propto \theta^{\sum n_i + r-1}e^{-(5+\lambda)\theta}$$

$$\sim \text{Gamma}\left(\sum n_i + r,\ 5+\lambda\right)$$

Conclusion: Poisson for likelihood and Gamma for prior is a <u>conjugate pair</u> since the posterior is Gamma.