Conditional Probability (Chapter 2)

Consider X, Y discrete taking values in $\{0, 1, \dots\}$ for example.

$$p_{X|Y}(x|y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Defined if $p_Y(y) > 0$

The conditional pmf defines a distribution/pmf since $p_{X|Y}(x|y) \geq 0$ if $p_Y(y) \neq 0$

and

$$\sum_x p_{X|Y}(x|y) = \sum \frac{p_{X,Y}(x,y)}{p_Y(y)} = \frac{p_Y(y)}{p_Y(y)} = 1$$

<u>Ex</u> $X \sim Bin(N, p)$ $N \sim Bin(M, q)$

$$p_{X|N}(k|n) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \dots, n$$

$$p_N(n) = \binom{M}{n} q^n (1-q)^{M-n} \quad (Bin(M, q))$$

$$P(X=k) = \sum_{i=0}^{M} P(X=k, N=i)$$

$$= \sum_{i=0}^{M} P(X=k \mid N=i)\, P(N=i)$$

$$= \sum_{i=0}^{M} p_{X|N}(k,i)\, p_N(i)$$

$$= \sum_{i=0}^{M} \binom{i}{k} p^k (1-p)^{i-k} \binom{M}{i} q^i (1-q)^{M-i}$$

where $\binom{i}{k} = 0$ when $i < k$

So we might as well sum

$$\sum_{i=k}^{M} \binom{i}{k} p^k (1-p)^{i-k} \binom{M}{i} q^i (1-q)^{M-i}$$

$$\binom{i}{k}\binom{M}{i} = \frac{\cancel{i!}}{(i-k)!\,k!} \frac{M!}{\cancel{i!}\,(M-i)!} = \frac{M!}{k!} \frac{1}{(i-k)!\,(M-i)!}$$

$$\left(\frac{q}{1-q}\right)^k p^k \frac{M!}{k!} (1-q)^M \underbrace{\sum_{i=k}^{M} q^{i-k} \frac{(1-p)^{i-k}}{(1-q)^{i-k}} \frac{1}{(i-k)!\,(M-i)!}}$$

The inner sum after relabeling is

$i \geq i-k$

$M-i \geq M-$

$$\sum_{i=0}^{M-k} \alpha^i \frac{1}{i!\,(M-k-i)!} = \frac{(1+\alpha)^{M-k}}{(M-k)!}$$

using the Binomial theorem.

$$= \left(\frac{q}{1-q}\right)^{k} p^{k} \frac{M}{k!} \frac{(1-q)^{M}}{(M-k)!} \left(1 + \frac{(1-p)q}{1-q}\right)^{M-k}$$

$$= (pq)^{k} (1-pq)^{M-k} \binom{M}{k}$$

$$\Rightarrow \quad X \sim \mathrm{Binom}(M, pq)$$

$$X \mid N \sim \mathrm{Binom}(N, p)$$

$$N \sim \mathrm{Bin}(M, q)$$

Negative Binomial

$X \sim \text{Negbin}(n, p)$ represents the # of independent trials required to get exactly n successes where p is the probability of success.

$$P(X = k) = \binom{n-1}{k-1}(1-p)^{n-k} p^{k}$$

Example 2:

Suppose $X \sim \text{Neg Bin}(p, N)$ $\quad N \sim \text{Geom}(\beta)$

Given $p_N(n) = (1-\beta)^{n-1}\beta \qquad n = 1, 2, \ldots$

$$p_{X|N}(i|n) = \binom{i-1}{n-1}(1-p)^{i-n} p^n \qquad i \geq n$$

Find $p_X(k)$

$$p_X(k) = \sum_{t=k}^{\infty} p_{X|N}(k|t)\, p_N(t)$$

$$= \sum_{t=1}^{k} \binom{k-1}{t-1}(1-p)^{k-t} p^t (1-\beta)^{t-1}\beta \qquad \text{COV: } u = t-1$$

$$= \sum_{u=0}^{k-1} \binom{t-1}{u}(1-p)^{k-1-u} p^{u+1}(1-\beta)^u \beta$$

$$= (p\beta)\sum_{u=0}^{k-1} \binom{t-1}{u}(1-p)^{k-1-u} p^u (1-\beta)^u$$

$$= (p\beta)\left(1-p+p(1-\beta)\right)^{k-1} = (p\beta)(1-p\beta)^{k-1}$$

$\Rightarrow \quad X \sim \text{Geom}(p\beta)$

Why? Toss a coin until you see your first heads. Let N be the # of trials required. Then toss a coin until you see N successes.

Not immediately obvious (at least to me).

There is another way of defining negative binomial:

$Y \sim \text{Negbin}(p, n)$

Y = represents the # of FAILURES you see until you see n successes.

This is the definition used in the textbook, and this also matches the Wikipedia definition.

$P(Y=k)$

$$\_\ \underline{S}\ \_\ \_\ \_\ \underline{S}\ \_\ \_\ \underline{S} \qquad (Y=6)$$

$$X = 6+3$$

3 successes.

Recall our definition:

$X \sim \text{Negbin}(n,p)$ represents the # of independent trials required to get exactly n successes where p is the probability of success.

If $X=k$, then $Y = k-n$ Or if $Y=k$, $X = n+k$

$$\Rightarrow P(Y=k) = P(X=n+k) = \binom{n+k-1}{k-1}(1-p)^k p^n$$

$$= \binom{n+k-1}{n}(1-p)^k p^n$$

Similarly, consider:

$X \sim \text{Geometric}(p)$ $\quad X =$ Trial # on which 1st success is seen.

$$P(X=k) = (1-p)^{k-1}p$$

$Y \sim \text{Geometric}_0(p)$ $\quad Y =$ # of failures before you see your 1st success.

Ex: Find

1) range (Y)
2) Relationship between X and Y
3) $P(Y=k)$

Conditional Expectation.

Let g be a function. Then

$$E[g(X) \mid Y=y] := \sum g(x) \underbrace{p_{X|Y}(x|y)}_{\text{conditional pmf.}} \quad \text{if } p(y)>0$$

$$E[g(X)] = \sum_{x} g(x)\, p_X(x) \qquad \text{by definition.}$$

Ex: $Y \sim$ Roll a die, $X \sim \text{Unif}(\{1,2,\dots,Y\})$

Find $E[X|Y]$

$$p_{X|Y}(i|k) = \frac{1}{k}$$

$$E[X|Y=k] = \sum_{i=0}^{k} i\, p_{X|Y}(i|k) = \frac{\cancel{k}(k+1)}{2\cancel{k}} = \frac{k+1}{2}$$

This is true for $Y = 1, 2, \dots, 6$. So we write

$$E[X|Y] = \frac{Y+1}{2} = h(Y)$$

Conditional expectation is a RANDOM VARIABLE!

This is true in general: $E[g(X)|Y] = h(Y)$.

So now, continue

$$E\left[E[X \mid Y]\right] = E\left[\frac{Y+1}{2}\right] = \sum_{i=1}^{6} \frac{i+1}{2} \cdot \frac{1}{6}$$

$$= \frac{\cancel{6} \cdot 7}{2 \cdot 2 \cdot \cancel{6}} + \frac{\cancel{6}}{2 \cdot \cancel{6}} = \frac{9}{4} = 2.25$$

Let us find

1) $\text{range}(X) = \{1, 2, \dots, 6\}$

2) $$p_X(i) = \sum_{k=1}^{6} p_{X|Y}(i \mid k)\, p_Y(k)$$

Note $$p_{X|Y}(i \mid k) = \begin{cases} 0 & i > k \\ \frac{1}{k} & i \in \{1, 2, \dots, k\} \end{cases}$$

$$= \sum_{k=i}^{6} \frac{1}{k} \cdot \frac{1}{6} = \frac{1}{6}\left(\frac{1}{i} + \frac{1}{i+1} + \dots + \frac{1}{6}\right)$$

$$E[X] = \sum_{i=1}^{6} i\, p_X(i) = \sum_{i=1}^{6} \frac{1}{6}\, i \left(\frac{1}{i} + \frac{1}{i+1} + \dots + \frac{1}{6}\right)$$

$$= \frac{1}{6}\left[\begin{array}{l} 1 + \frac{1}{2} + \dots + \frac{1}{6} \\ 2\left(\frac{1}{2} + \dots + \frac{1}{6}\right) \\ \\ 6 \cdot \frac{1}{6} \end{array}\right] = \frac{1}{6}\left[1 + \frac{1}{2}(1+2) + \frac{1}{3}(1+2+3) + \dots + \frac{1}{6}(1+2+\dots+6)\right]$$

$$= \frac{1}{6} \sum_{i=1}^{6} \frac{1}{i}(1+2+\dots+i) = \frac{1}{6} \sum_{i=1}^{6} \frac{i(i+1)}{2i} = \frac{1}{12}\left[\frac{6 \cdot 7}{2} + 6\right]$$

$$= \frac{1}{2}\left[\frac{7}{2}+1\right] = \frac{9}{4}$$

COINCIDENCE? I THINK NOT!

TOWER PROPERTY OF CONDITIONAL EXPECTATION

Recall: $E[g(X)|Y] = \sum_i g(i)\, p_{X|Y}(i|Y) = h(Y)$

Now, $E[E[g(X)|Y]] \quad E[h(Y)] = \sum_j h(j)\, p_Y(j)$

$$= \sum_j \sum_i g(i)\, p_{X|Y}(i|j)\, p_Y(j)$$

$$= \sum_i \sum_j g(i) \frac{p_{X,Y}(i,j)}{p_Y(j)}\, p_Y(j) = \sum_i g(i)\, p_X(i)$$

$$= E[g(X)]$$

Called the tower property of conditional expectation

Notice that

$\sum_y E[g(X)\,|\,Y=y]\, p_Y(y)$ looks like we're taking the expectation of a random variable.

In fact, we may write

$E[g(X)\,|\,Y]$ as this random variable.

Once the random Y has been fixed $E[g(X)|Y]$ is known so we may think of it is a random variable of the form $h(Y)$ where

$h(y) = E[g(X)\,|\,Y=y]$.

Properties of Conditional Expectation

1) Linearity:

$$E[2X^2+3X^4 \mid Y=j] = \sum_i (2i^2+3i)\, p_{X|Y}(i|j) = E[2X^2|Y=j] + E[3X^4|Y=j]$$

$$E[a\,g(X)+b\,h(Y) \mid Y] = E[a\,g(X)|Y] + E[b\,h(X)|Y]$$

2) Positivity for positive fns.

$$E[2X^2 \mid Y=j] = \underbrace{\sum 2i^2\, p_{X|Y}(i|j)}_{\text{positive terms.}}$$

3) $E[v(X,Y) \mid Y=j] = E[v(X,j) \mid Y=j]$

4) $E[g(X)h(Y) \mid Y=y] = h(y)\,E[g(X)|Y=y]$

Can prove with examples.

5) If X and Y are independent

$$E[g(X) \mid Y=y] = E[g(X)]$$

Exercise

6) $E[g(X)\,h(Y)] = \sum E[g(X)h(Y) \mid Y=y]\, P_Y(y)$

y is like a constant inside the expecta.

$$= \sum_y h(y)\, E[g(X)|Y=y]\, P_Y(y)$$

$$= E\big[h(Y)\, E[g(X)|Y]\big]$$

2.2 The Dice Game Craps

I used to think this was a boring section, but I quite like it now, since it reinforces ideas from conditioning.

The key NEW IDEA is "conditioning on the first step"

2 dice are rolled and the sum is recorded. Let $Z = X+Y$

$$Z = \begin{cases} 2,3,12 & \longrightarrow \text{Player loses} \\ 7,11 & \longrightarrow \text{Player wins} \\ 4,5,6,8,9,10 & \longrightarrow \text{Player keeps rolling until } Z \text{ or } 7 \text{ is seen. If} \\ & \quad Z \text{ appears then LOSE, if } 7 \text{ appears } Z \text{ wins.} \end{cases}$$

Recall p =

2	3	4	5	6	7	8	9	10	11	12
$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$								

Let A be the event that PLAYER wins the game. Let the rolls be

$Z_1, Z_2, Z_3, \ldots$ until the game ends.

$$P(A) = \sum_{i=2}^{12} P(A \mid Z_1 = i)\, p(i)$$

$$= 1 \cdot \left(p(11)+p(7)\right) + \sum_{i \in 4,5,6,8,9,10} P(A \mid Z_1 = i)\, p(i) + 0 \cdot \left(p(2)+p(3)+p(12)\right)$$

$$P(A \mid Z_1 = i) = P(B_i)$$

where B_i is the event that i is rolled before a 7 (starting from roll 2)

$$P(B_i) = \sum_{\substack{Z_2 = j \\ j \notin \{i,7\}}} P(B_i \mid Z_2 = j)\, p(j) + P(B_i \mid Z_2 = i)\, p(i) + P(B_i \mid Z_2 = 7)\, p(7)$$

$$= P(B_i)\left(1 - p(i) - p(7)\right) + p(i)$$

$$\Rightarrow P(B_i) = \frac{p(i)}{p(i)+p(7)}$$

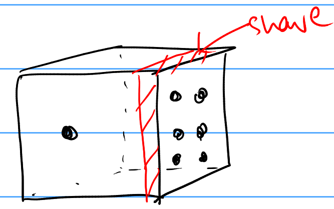
Putting this together we get

$$P(A) = p(11) + p(7) + \sum_{i \in \{4,5,6,8,9,10\}} \frac{p(i)}{1-\alpha_p}$$

If we plug in the numbers, we should get $P(A) \approx 0.49$

So did they know all of this when they invented the game? CURIOUS.

Shaved dice: Suppose we change the pmf of Z by shaving dice.



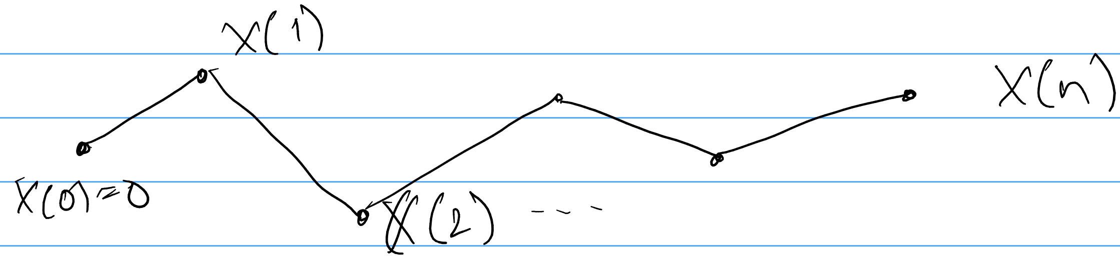
Problem: Write a computer program that computes a win probability as a function of the faces shaved and ϵ, the change in the corresponding win probability.

Hidden bonus: send me your code and graphs for a +1 bonus on your final.

You may collaborate with one other person on this.

2.3 Random Sums

Recall SRW



$\{\xi_i\}_{i=1}^{n}$ collection of steps, iid ± 1 valued

$$X_n = \xi_1 + \xi_2 + \cdots + \xi_n$$

Find the pmf $P(X_n = a) = p(a)$

This is our prototypical stochastic process.

<u>Stochastic Process</u>: Collection of rvs $\{X(i)\}_{i=0}^{n}$ describing the evolution (in discrete time) of some quantity.

<u>General random walk</u> $\{\xi_i\}_{i=1}^{n}$ may not be ± 1 valued but are iid. May also take values in $\mathbb{Z}^d$.

Random sum

What if the # of terms in the sum is random?

let $N \in \{0, 1, 2, \cdots\}$ $\{\xi_i\}$ are iid
and independent of N

$$X = \begin{cases} \xi_1 + \xi_2 + \cdots + \xi_N & \text{if } N > 0 \\ 0 & N = 0 \end{cases}$$

Examples:

1) Queuing: Let N be the # of customers arriving at the DMV
Let X_i be the amount of time required to service.
$X_1 + X_2 + \cdots + X_N$ is the service time.

2) Risk: Let N be the # of claims arriving at an insurance company. Let $\{X_i\}_{i=1}^{N}$ be the amount of each claim.

Then $X_1 + X_2 + \cdots + X_N$ is the total claim amount in that time period.

3) Let ξ_i be the number of children a woman has. Suppose there are N individuals in a population. How many individuals will we have in the next generation?

This is represented by X_n as well.

Find $E[X]$ and $Var(X)$

$$E[X] = E[E[X|N]]$$

$$E[X|N=k] = k E[\xi_1] \quad - (\star 3)$$

$$\Rightarrow E[X] = E[N E[\xi_1]] = E[N] E[\xi]$$

How to show $(\star 3)$? Suppose $\{\xi_i\}$ are discrete.

$$p_{X|N}(i|k) = \frac{p_{X,N}(i,k)}{p_N(k)} = \frac{P\left(\sum_{i=1}^{k} \xi_i = i, N=k\right)}{P(N=k)}$$

$$= P\left(\sum_{i=1}^{k} \xi_i = i\right)$$

$$\Rightarrow E[X|N=k] = \sum_{a \in range(X)} a \, p_{X|N}(a|k) = \sum a \, P\left(\sum_{i=1}^{k} \xi_i = a\right)$$

$$= E\left[\sum_{i=1}^{k} \xi_i\right]$$

Similarly compute $E[X^2]$

Then

$$Var(X) = E[N] Var(\xi) + E[\xi] Var(N)$$

So the variances of both quantities contribute.

$$E[X_N^2] = E[E[X_N^2 | N]]$$

$$E[X_N^2 | N=k] = \sum a^2 p_{X|N}(a|k)$$

$$= \sum a^2 \frac{P(X_k = a, N=k)}{P(N=k)} = E[X_k^2]$$

$$= E\left[\sum_{i=1}^{k} \sum_{j=1}^{k} \xi_i \xi_j\right] = k\, E[\xi_1^2] + k(k-1) E[\xi_1]^2$$

$$= k \operatorname{Var}(\xi_1^2) + k^2 E[\xi_1]^2$$

$$\Rightarrow E[X_N^2] - E[X_N]^2 = E[N] \operatorname{Var}(\xi_1^2) + E[N^2] E[\xi_1]^2$$

$$- E[N]^2 E[\xi_1]^2$$

$$= E[N] \operatorname{Var}(\xi_1^2) + \operatorname{Var}(N) E[\xi_1]^2$$

TWO contributions to the Variance: one from N and one from ξ_1.

Distribution of random sums

Suppose $\xi_1, \xi_2, \cdots \xi_n$ are continuous with pdf $f(z)$.

Does $X = \sum_{i=1}^{N} \xi_i$ have a pdf? Is it continuous?

Recall convolution:

$\xi_1 + \xi_2$ has pdf $f^{(2)}(z) = \int f(u) f(z-u) du$

$\xi_1 + \xi_2 + \xi_3$ has pdf $f^{(3)}(z) = \int f^{(2)}(u) f(z-u) du$

And in general $f^{(n)}(z) = \int f^{(n-1)}(u) f(z-u) du$

The cdf is $F^{(n)}(z) = \int_{-\infty}^{z} f^{(n)}(u) du$

$$F_{X_N}(t) = P(X_N \le t)$$

$$= \sum_{k=1}^{\infty} \underbrace{P(X \le t \mid N = k)}_{F_{X_k}(t)} P(N = k)$$

after supposing $P(N=0) = 0$

☆ What do you do if $P(N=0) > 0$?

$$= \sum_{k=1}^{\infty} F^{(k)}(t)\, p_N(k)$$ where $F^{(k)}$ is the k-convolution of ξ_1

⇒ The density of X is

$$f_X(t) = \sum_{k=1}^{\infty} f^{(k)}(t)\, p_N(k)$$

<u>Ex</u> Geometric sum of Exp

$\xi_i \sim Exp(\lambda)$ $\qquad N \sim Geom(p)$

$$f(z) = \lambda e^{-\lambda z} \qquad z \geq 0$$

$$f^{(n)}(z) = \begin{cases} \frac{\lambda^n}{(n-1)!} z^{n-1} e^{-\lambda z} & z \geq 0 \\ 0. & \text{other wise} \end{cases}$$

Gamma distribution. (One can prove this using mgfs)

$$f_X(z) = \sum_{n=1}^{\infty} \frac{\lambda^n}{(n-1)!} z^{n-1} e^{-\lambda z} \cdot \beta (1-\beta)^{n-1}$$

$$= \beta e^{-\lambda z} \lambda \sum_{n=0}^{\infty} \frac{[\lambda z(1-\beta)]^n}{n!} = \beta \lambda e^{-\lambda z} e^{\lambda z (1-\beta)}$$

$$= \beta \lambda e^{-\lambda \beta z} \qquad z \geq 0 \qquad (\text{Exp}(\lambda\beta) \text{ distribution})$$

<u>Stock prices changes</u>

stock price

$$X(0) \longrightarrow X(1) \longrightarrow X(2)$$

transaction

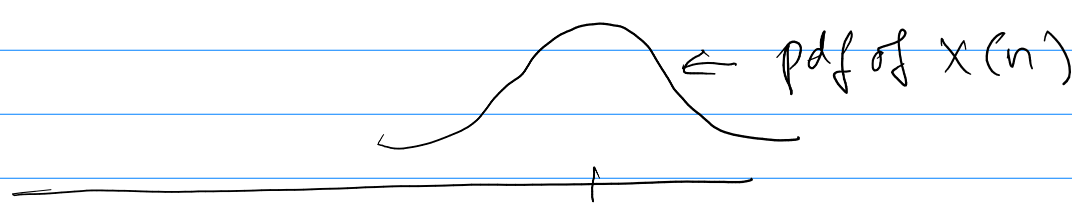
transaction (some one buys or sells a share)

Assume $\quad X(i) - X(i-1) = \xi_i$

$$X(n) = \xi_1 + \cdots + \xi_n + X(0)$$

So according to CLT

$$X(n) \approx N\left(X(0) + E[\xi_1]\, n,\ \mathrm{Var}(\xi_1)\, n\right)$$



$nE(\xi_1) + X(0)$

People studied this model and looked for its statistical properties. The density predicted by this model (Normal) didn't fit the pdf of the end-of-day price.

Let $Z = \xi_1 + \cdots + \xi_N$ ($X(0)$ is just a shift)

where N is the random # of transactions seen in a day.

Assume $\xi_i \sim N(0, \sigma^2)$

If $N = n$, $Z \sim N(0, n\sigma^2)$

$\Rightarrow$

$$f_{Z|N=n}(t) = \frac{1}{\sqrt{2\pi n \sigma^2}} \exp\left(-\frac{t^2}{2n\sigma^2}\right)$$

Assume $N \sim \text{Poisson}(\lambda)$

Then

$$f_Z(t) = \sum_{n=1}^{\infty} f_{Z|N=n}(t) \frac{\lambda^n e^{-\lambda}}{n!}$$

This density shows a much better fit with the data.

Now we want to generalize conditional expectation to absolutely continuous rvs.

<u>Continuous rvs</u>

Define similarly when X and Y are <u>jointly continuous</u>

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \qquad f_Y(y) > 0$$

$$F_{X|Y}(x|y) := \int_{-\infty}^{x} f_{X|Y}(u|y)\, du$$

Suppose you have a fn $g : \mathbb{R} \to \mathbb{R}$ and you want to describe its average value given some information on Y, we define

$$E[g(X) \mid Y=y] := \int_{-\infty}^{\infty} g(u) \underbrace{f_{X|Y}(u|y)}_{\text{conditional pdf}} du$$

when $f_Y(y) > 0$.

These are DEFINITIONS.

Notice that $P(Y=y) = 0$ if Y is continuous.

So we cannot use our previous definition for $p_{X|Y}(u|y)$ since $p_Y(u) = 0$

Ex: Suppose a woman arrives at a bus stop at $Y \sim \text{Unif}(0,1)$. Independently the bus arrives at $Z \sim \text{Unif}(0,1)$. Given $Y = 0.2$ what is the prob that she catches the bus.

Obviously the answer is $P(Z \in (0, 0.2)) = 0.2$

But $P(Z \in (0,Y) \mid Y = 0.2) = \dfrac{P(Z \in (0,Y), Y = 0.2)}{P(Y=0.2)}$

is UNDEFINED.

But the intuitive and correct answer is

$$F_{Z|Y=0.2}(0.2|0.2) - F_{Z|Y=0.2}(0|0.2)$$

$$= \int_0^{0.2} f_{Z|Y=0.2}(u|0.2)\, du$$

$$= \int_0^{0.2} \frac{f_{Z,Y}(u, 0.2)}{f_Y(0.2)}\, du = \int_0^{0.2} \frac{1}{1}\, du$$

$$= 0.2$$

Law of total probability if both X and Y jointly continuous.

$$P(a<X<b, c<Y<d) = \int_c^d \int_a^b f_{X|Y}(u|v)\, f_Y(v)\, dv$$

Recall the formula for discrete X and Y

$$E[g(X)h(Y)] = E[h(Y)\,E[g(X)|Y]]$$

We could take $h(Y)$ out since we "froze" Y first.

What do we do if (X,Y) are jointly absolutely continuous?

$$E[g(X)h(Y)] = \int\int g(u)h(v)\,f_{X,Y}(u,v)\,du\,dv$$

$$= \int\int g(u)\,h(v)\,f_{X|Y}(u,v)\,f_Y(v)\,du\,dv$$

Fubini

$$= \int \underbrace{\left[\int g(u) f_{X|Y}(u,v)\,du\right]}_{E[g(X)\,|\,Y=v]} h(v)\,f_Y(v)\,dv$$

expectation over Y

$$= E\left[\,E[g(X)|Y]\,h(Y)\right]$$

Ex: $f_{XY}(x,y) = \frac{1}{y} e^{-(x/y)-y} \quad x,y>0$

$E[X \mid Y=y]$?

First find $f_{X|Y}(x|y) = \dfrac{\frac{1}{y} e^{-\left(\frac{x}{y}\right)-y}}{f_Y(y)} \quad x,y>0$

$$f_Y(y) = \int_0^\infty \frac{1}{y} e^{-\left(\frac{x}{y}\right)-y} dx = \frac{1}{y} e^{\left(\frac{-x}{y}\right)} \frac{1}{\left(\frac{-1}{y}\right)} e^{-y} \Big|_0^\infty$$

$$= e^{-y} \quad y>0$$

$$\Rightarrow E[X \mid Y=y] = \int_0^\infty x \frac{\frac{1}{y} e^{-\left(\frac{x}{y}\right)-y}}{e^{-y}} dx$$

$$= \int_0^\infty x \frac{1}{y} e^{-x/y} dx$$

Set $\lambda = \frac{1}{y}$ $\quad = \int_0^\infty x \underbrace{e^{-\lambda x} \lambda}_{Exp(\lambda)} dx = \frac{1}{\lambda} = y$

KP: (2.4.1), (2.4.3) (2.4.6) Good exercises.

Generalization to n variables:

$$E\left[g(X_n) \mid \underbrace{(X_1, X_2, \ldots, X_{n-1}) = (a_1, a_2, \ldots a_{n-1})}_{f(X_1, \ldots, X_n)}\right]$$

$$= \sum x_n \, p(x_n \mid a_1, \ldots a_{n-1})$$

$$= \sum x_n \frac{p_{X_1, \ldots X_n}(a_1, \ldots, a_{n-1}, x_n)}{p_{X_1, \ldots X_{n-1}}(a_1, \ldots, a_{n-1})}$$

★ Show (HW)

$$E[X_1 + X_2 + X_3 \mid X_2, X_3]$$

$$= X_2 + X_3 + E[X_1 \mid X_2 X_3]$$

Martingales

Recall random walk with mean 0 increments.

$\{\xi_n\}_{n=1}^{\infty}$, $X_n = \sum_{i=1}^{n} \xi_i$

Find $E[X_{n+1} \mid \underbrace{X_1, X_2, \dots, X_n}_{\mathcal{F}_n}]$

$= E[\xi_1 + \xi_2 + \dots + \xi_{n+1} \mid X_1, X_2, \dots, X_n]$

$= E[\xi_{n+1} + X_n \mid X_1 \dots X_n]$

$= X_n + E[\xi_{n+1} \mid X_1, \dots, X_n]$

using $E[h(Y) \mid Y] = h(Y)$

$= X_n + E[\xi_{n+1}] = X_n$

since ξ_{n+1} is independent.

This is called the Martingale property of the random walk.

$E[X_{n+1}] = E[E[X_{n+1} \mid X_1, \dots X_n]]$

$= E[E[X_{n+1} \mid X_n]] = E[X_n]$

Expectations of Martingales remains a constant.

☆ Show (HW) $E[X_m \mid X_0, X_1, \dots, X_n] = X_n \quad m \geq n+1$

In gambling: X_n might be the value of a players fortune after n plays of a game.

The game is FAIR if ~~their~~ the value neither inc. nor decreases.

Stock Prices in a Perfect Market:

X_n = Share price of Tesla at the end of day n.

People believe that in a perfect world that $\{X_n\}_{n=1}^{N}$ is a Martingale.

If you could predict that X_n would increase to X_{n+1} on average, then many buyers would come in, and then the current price X_n would instantly change.

This is sort of like a no-arbitrage principle.

Polya's Urn Model

An urn contains one red and one green ball. A ball is drawn at random and returned to the urn, together with ANOTHER ball of the same color.

let R_n = # of red balls $\quad R_0 = 1$

B_n = # of blue balls $\quad B_0 = 1$

let X_n = fraction of red balls $= \dfrac{R_n}{R_n + B_n}$

Show that X_n is a martingale. (Easy to just do in class)