Lecture 25: Order Statistics and Conditional Expectation

~~Bank and~~

Bank-post office example

Let's say we go to the bank and we wait for a total of X minutes to be served and suppose

waiting time to be served at bank $\to X \sim \text{Gamma}(a, \lambda)$,

wait at PO $\to Y \sim \text{Gamma}(b, \lambda)$

X, Y independent.

Find joint distribution of $X + Y = T$ (total),

$\frac{X}{X+Y} = W$ (of our total waiting time, what fraction of that was spent waiting at the bank).

Note: Gamma$(5, \lambda)$ $\downarrow$ we can think of it as the sum of five i.i.d. exponential λ. ~~lambd~~

Example – There are 5 people in line before person A and each of them takes an ~~exponential~~ Expo(λ) amount of time to be served. There is only 1 line. Person A waits in line and then eventually his turn.

Solu: Let $\lambda = 1$ to simplify notation. (doesn't lose generality)

Joint PDF:

$$f_{T,W}(t,w) = f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(t,w)}\right|$$

$$= f_X(x)\, f_Y(y)\left|\frac{\partial(x,y)}{\partial(t,w)}\right| \quad [\because X, Y \text{ independent}]$$

$$= \frac{1}{\Gamma(a)\Gamma(b)} x^a e^{-x} y^b e^{-y} \frac{1}{xy}\left|\frac{\partial(x,y)}{\partial(t,w)}\right|$$

absolute of determinant of Jacobian

$$x+y=t, \quad \frac{x}{x+y}=w$$

$$\boxed{x=tw}$$

$$\boxed{y=t(1-w)}$$

$$\left|\frac{\partial(x,y)}{\partial(t,w)}\right| = \begin{vmatrix} w & t \\ 1-w & -t \end{vmatrix} = -tw - t(1-w) = -t$$

$$\therefore f_{T,W}(t,w) = \frac{1}{\Gamma(a)\Gamma(b)} x^a e^{-x} y^b e^{-y} \frac{1}{xy} t$$

Joint PDF

$$= \frac{1}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1} t^{a+b} e^{-t} \, {}^{1}\!/_{t}$$

$\boxed{W, T \text{ are independent}}$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1} \underbrace{\frac{t^{a+b} e^{-t}}{\Gamma(a+b)}}_{\text{Gamma}(a+b, 1)}$$

Gamma PDF

Marginal PDF of W

$$f_W(w) = \int_{-\infty}^{+\infty} f_{T,W}(t,w)\, dt$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1} \underbrace{\int_{-\infty}^{\infty} \frac{t^{a+b} e^{-t}}{\Gamma(a+b)} dt}_{1}$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1}$$

Normalizing constant of Beta(a,b):

$T \sim \text{Gamma}(a+b, 1)$,

$W \sim \text{Beta}(a,b)$

T & W are independent

Example find $E(W)$, $W \sim \text{Beta}(a,b)$.

Solu: Ways of doing this : (1) LOTUS/definition.

(2) $E\left(\frac{X}{X+Y}\right) = \frac{E(X)}{E(X+Y)}$ [special property not by Linearity]

Why is $E\left(\frac{X}{X+Y}\right) E(X+Y) = E(X)$ in this special problem of Gamma–Beta?

$\rightarrow$ For this problem, $\frac{X}{X+Y}$ is indep. of $X+Y$.

Therefore, they are uncorrelated since independence implies uncorrelation.

$E\left(\frac{X}{X+Y}\right) = \frac{E(X)}{E(X+Y)} = \frac{a}{a+b}$

since, mean of ~~the~~ $\text{Gamma}(a,1) = a$

$\text{Gamma}(a+b, 1) = (a+b)$.

or, $E(X+Y) = E(X) + E(Y)$ (by Linearity)

$= (a+b)$

Order Statistics

Let $X_1, \ldots, X_n$ be i.i.d. random variables. The order statistics are $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$, where

$$X_{(1)} = \min(X_1, \ldots, X_n), \ldots, X_{(n)} = \max(X_1, \ldots, X_n).$$

e.g. if n is odd, the median is $X_{\left(\frac{n+1}{2}\right)}$. Get other "quantiles".

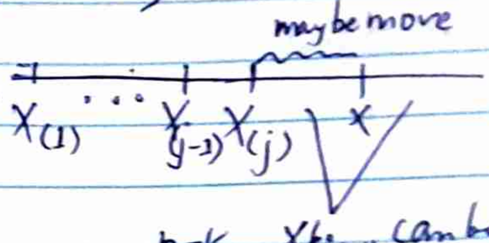
Difficult to work with, since dependent.

Tricky in discrete case because of "ties" (i.e., 2 random variables being exactly equal).

Continuous Case

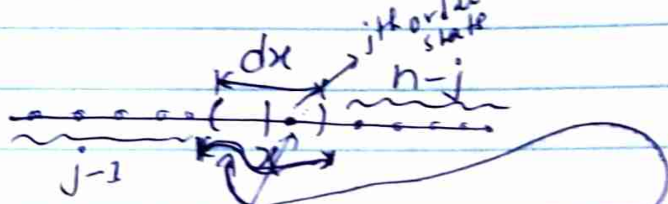
Let $X_1, \ldots, X_n$ be i.i.d. with PDF f, CDF F. Find the CDF and PDF of $X_{(j)}$ (jth order statistics)

CDF: $P(X_{(j)} \leq x) = P(\text{at least } j \text{ of } X_{(i)}\text{'s are} \leq x)$



$$= \sum_{k=j}^{n} \binom{n}{k} F(x)^k (1-F(x))^{n-k}$$

PDF:



$f_{X_{(j)}}(x)\,dx$ = prob. that jth order statistics in "dx" tiny interval

$$= n\binom{n-1}{j-1}(f(x)\,dx)\,F(x)^{j-1}(1-F(x))^{n-j}$$

jth order statistics; $F(x)$ → prob. of success (left of x)

$$\left[f_{X_{(j)}}(x) = n\binom{n-1}{j-1} F(x)^{j-1}(1-F(x))^{n-j} f(x)\right]$$

Example $U_1, \ldots, U_n$ i.i.d. Unif(0,1). Find the distribution of the jth order statistics.

Solu: $f_{U_{(j)}}(x) = n\binom{n-1}{j-1} x^{j-1}(1-x)^{n-j}$, for $0 < x < 1$ | PDF is just 1, $0<x<1$, 0, otherwise

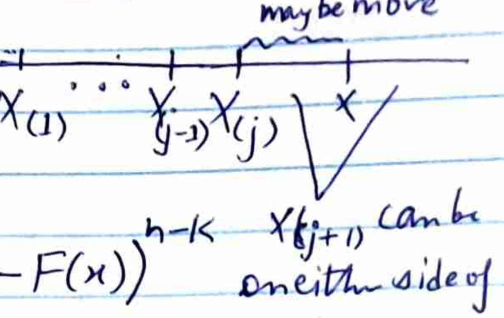
→ CDF for Uniform increases linearly, since the PDF is const.

$\Rightarrow U_{(j)} \sim \text{Beta}(j, n-j+1)$

Difficult to work with, since dependent.
Tricky in discrete case because of "ties" (i.e., 2 random variables being exactly equal).

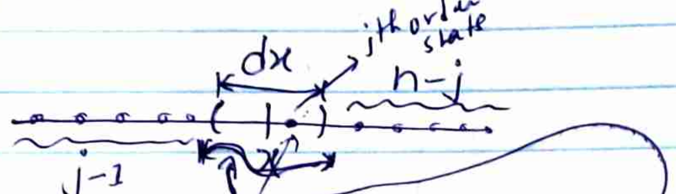
Continuous Case:
Let $X_1, \dots, X_n$ be i.i.d. with PDF f, CDF F. Find the CDF and PDF of $X_{(j)}$ (jth order statistics)

CDF: $P(X_{(j)} \leq x) = P($at least j of $X_{(i)}$'s are $\leq x)$



$$= \sum_{k=j}^{n} \binom{n}{k} F(x)^k (1-F(x))^{n-k}$$

PDF:



$f_{X_{(j)}}(x)\,dx$ = prob. that jth order statistics in "dx" tiny interval

$$= n \binom{n-1}{j-1} (f(x)\,dx) F(x)^{j-1} (1-F(x))^{n-j}$$

jth order statistics

→ prob. of success (left of x)

$$\left[f_{X_{(j)}}(x) = n \binom{n-1}{j-1} F(x)^{j-1} (1-F(x))^{n-j} f(x) \right]$$

Example: $U_1, \dots, U_n$ i.i.d. Unif$(0,1)$. Find the distribution of the jth order statistics.

Solu: $f_{U_{(j)}}(x) = n \binom{n-1}{j-1} x^{j-1} (1-x)^{n-j}$, for $0 < x < 1$

PDF is just 1 $0<x<1$, 0, otherwise

→ CDF for Uniform increases linearly, since the PDF is const.

$\Rightarrow U_{(j)} \sim \text{Beta}(j, n-j+1)$

$E|U_1 - U_2| = E(\max) - E(\min)$

$\quad\quad\quad\quad\quad\quad \text{Beta}(2,1) \quad \text{Beta}(1,2)$

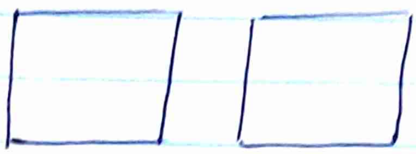
$= 2/3 - 1/3 = 1/3$

Conditional Expectation

$E(X|A)$ event

$E(X) = E(X|A)\,P(A) + E(X|A^c)\,P(A^c)$

Two Envelope Paradox



$E(X) = \sum_x x\,P(X=x)$

Expand with LOTP

We are given two envelopes. Each envelope contains a check for some amount of money. We are given that one envelope contains exactly twice as much money as the other one. It's symmetrical. We have to pick one. After we pick one, we are given option to switch. Should we switch or not?

Lecture 26: Conditional Expectation continued...

$\boxed{\$X}$ $\boxed{\$Y}$ One envelope has twice as much as the other.

Argument 1: $E(Y) = E(X)$ (by symmetry)

Argument 2: $E(Y) = E(Y \mid Y = 2X)\,P(Y = 2X) + E(Y \mid Y = X/2)\,P(Y = X/2)$ (LOTP)

$$\neq E(2X) \cdot 1/2 + P(X/2) \cdot 1/2$$

$$= 2E(X) \cdot 1/2 + 1/2\,P(X) \cdot 1/2$$

$$= 5/4\,E(X)$$

Corrected version:

$$E(Y) = E(Y|Y=2x)P(Y=2x) + P(Y|Y=x/2)P(Y=x/2)$$

$$= E(2x|Y=2x)\,P(Y=2x) + P(x/2|Y=x/2)\,P(Y=x/2)$$

$$= E(2x|Y=2x)\cdot 1/2 + P(x/2|Y=x/2)\cdot 1/2$$

$P(Y=2x) = P(Y=x/2) = 1/2$ (equally likely)

$$E(Y|Y=2x) \neq E(2x)$$

Let I be an indicator of which envelope has more money.

I ← indicator random variable

or, indicator of $Y=2x$.

Then X, I are dependent. → have to be, if expected value have to be finite.

Example

Patterns in Coin flips

Repeated fair coin flips.

How many flips until HT? find $E(W_{HT})$

How many flips until HH? find $E(W_{HH})$

→ random variables

Solu: $E(W_{HT}) = E(W_1) + E(W_2)$

Partial progress

$\underbrace{TTTTH}_{W_1}\underbrace{H(HT)}_{W_2}$

wait for the first time the coin lands head. Let's call that W_1

W_2 → time for first tail after first head.

[W_1 & W_2 indep. because the coin is memoryless. Even if they were not indep., we could still apply linearity]

$= 2 + 2$

$= 4$

Symmetry

$E(W_{TT}) = E(W_{HH})$

$E(W_{HT}) = E(W_{TH})$

But, symmetry does not tell us that

$E(W_{HT}) = E(W_{HH})$

→ Expected value $= 1$

Since, $W_j - 1 \sim \text{Geom}(1/2)$ (our convention)

(W is geometry, careful that we defined the geometric to not include the success)

$$E(W_{HH}) = E(W_{HH} \mid \text{1st toss is } H)\cdot\tfrac{1}{2} + E(W_{HH} \mid \text{1st toss is } T)\cdot\tfrac{1}{2}$$

|TTTTH T/H

TTTTH T

No partial progress we have to restart again

$$= \Big[\underbrace{2\cdot\tfrac{1}{2}}_{\substack{\text{1st toss is H,}\\ \text{2nd toss is H}\\ \text{(only 2 flips)}}} + \underbrace{\{2 + E(W_{HH})\}\cdot\tfrac{1}{2}}_{\substack{\text{1st toss is H,}\\ \text{2nd toss is T}\\ \text{(we waste 2 toss from HT and then}\\ \text{same problem again as before}}}\Big]\cdot\tfrac{1}{2} + \underbrace{[1 + E(W_{HH})]\tfrac{1}{2}}_{\substack{\text{1st toss is T}\\ \text{(we waste 1 toss,}\\ \text{then it's same problem again.}}}$$

$$E(W_{HH}) = \frac{1}{2} + \frac{1}{2} + \frac{E(W_{HH})}{4} + \frac{1}{2} + \frac{E(W_{HH})}{2}$$

$$= \frac{3}{2} + \frac{3E(W_{HH})}{4}$$

or, $E(W_{HH}) = 6$. ✓

~~Intuition~~

~~TTHHHT... HHHTT~~

TED Talk

Peter Donnelly

✓ $E(Y \mid \underbrace{X=x}_{\text{event}})$

$$\overset{Y \text{ discrete}}{=} \sum_{y} y\, P(Y=y \mid X=x)$$

(we learned that $X=x$, so we conditioned on that information)

$$E(Y \mid X=x) \overset{Y \text{ Continuous}}{=} \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\, dy$$

$(X,Y \text{ cont.})$ $\overset{Y \text{ Continuous}}{=} \int_{-\infty}^{\infty} y\, \frac{f_{X,Y}(x,y)}{f_X(x)}\, dy$

Let $g(x) = E(Y|X=x)$. Then define $E(Y|X) = g(X)$.

eg. if $g(x) = x^2$, then $g(X) = X^2$

so, $E(Y|X)$ is a r.v., a function of X.

$E(Y|X) \longrightarrow$ it says assuming we get to pretent that X is Known, then what's our best prediction of Y.

$E(h(X)|X) = h(X)$

Example Let X, Y be i.i.d. Pois(λ).

$E(X+Y|X) = ?$, $E(X|X+Y) = ?$

Solu: $E(X+Y|X) = E(X|X) + E(Y|X)$ [By linearity]

$= \underbrace{X}_{\substack{X \text{ is a} \\ \text{function of} \\ \text{itself (i.e., } X)}} + E(Y)$

$= X + \lambda$

i.i.d.
X, Y independent
(If we have independence, we can drop the stuff we are conditioning on)

$E(X|X+Y)$

Let $T = X+Y$, find Conditional PMF.

$P(X=k|T=n) = P(T=n|X=k)P(X=k) / P(T=n)$ (Bayes' Rule)

$= P(Y=n-k|X=k)P(X=k) / P(T=n)$

X,Y indep. $= \dfrac{\dfrac{e^{-\lambda}\lambda^{n-k}}{(n-k)!}\, e^{-\lambda}\lambda^{k}/k!}{e^{-2\lambda}(2\lambda)^n/n!}$

Sum of indep. Poissons is poisson.

$= \binom{n}{k}\frac{1}{2^n}$

$\therefore X|T=n \sim \text{Bin}(n, 1/2)$

$E(X|T=n) = n/2$ (Binomial, Conditioned on an event)

$E(X|T) = T/2$.

Alternative way:

$E(X|X+Y) = E(Y|X+Y)$ [by symmetry, since i.i.d.]

$E(X|X+Y) + E(Y|X+Y) = E(X+Y|X+Y)$ (by linearity)

$= E(X+Y)$

$= X+Y$

$\Rightarrow E(X|T) = T/2$.

<u>Iterated Expectation (Adam's Law)</u> ↳ single most important property of conditional expectation

$E(E(Y|X)) = E(Y)$