stat 134 lec 10

Last time

Sec 3.1 Random Variables

A RV, X, is the outcome of an experiment.

(X,Y) is the joint outcome of two RVs X, Y.

The event $(X=x, Y=y)$ is the intersection of events $X=x$ and $Y=y$.

The probability X and Y satisfies some condition is the sum of $P(x,y)$ over all pairs (x,y) that satisfy that condition.

ex $$P(X=Y) = \sum_{(x,y):\, x=y} P(x,y) = \sum_{\text{all } x} P(x,x)$$

ex $$P(X+Y=s) = \sum_{(x,y):\, x+y=s} P(x,y) = \sum_{\text{all } x} P(x, s-x)$$

Independence of (X,Y,Z) means

$P(X=x, Y=y, Z=z) = P(X=x)P(Y=y)P(Z=z)$ for all $x \in X$, $y \in Y$, $z \in Z$.

Today

1. Student responses from Concept test last time.
2. Sec 3.1 Sums of independent Poissons is Poisson
3. Sec 3.2 Expectation of a RV.

(i)

1. The joint distribution of X and Y is drawn below:

	3/8	1/2	1/8	P(x) \ P(y)
1	1/4	1/3	1/12	2/3
0	1/8	1/6	1/24	1/3
Y \ X	0	1	2	

a X and Y are independent

b $P(X = x|Y = 0) = P(X = x|Y = 1)$, for all x.

c More than one of the above

d None of the above

c	X and Y are independent since each joint probability is the product of the two respective marginal probabilities. Since X and Y are independent, P(X=x) should not depend on Y, making B true. Thus, C is the correct answer.

c	Since this is independent, whatever y is, it does not affect the chance of x.
d	God help me

a	The two rows Y=0 and Y=1 give different probabilities for each value of X

if we divide both row by their marginal prob we get the same answer

② Sum of independent Poisson is Poisson

informal arguement:

$X_1 \sim Bin(1000, \frac{1}{1000})$ ≈ Pois(1)

$X_2 \sim Bin(2000, \frac{1}{1000})$ ≈ Pois(2)

indep

$X_1 + X_2 \sim ?$ $Bin(3000, \frac{1}{1000}) \approx Pois(3)$

$X_1 + X_2$ = # heads in 1000+2000 = 30[00]

$p = \frac{1}{1000}$ coin tosses.

So sum of two indep binomials with the same p is Binomial and this example suggest that sum of 2 indep Poisson is Poisson.

Lets prove this rigorously:

Recall binomial theorem

$$(a+b)^3 = \binom{3}{3}a^3b^0 + \binom{3}{2}a^2b^1 + \binom{3}{1}a^1b^2 + \binom{3}{0}a^0b^3$$
$$= a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Recall $X \sim Pois(\mu)$

$$P(X=k) = \frac{e^{-\mu}\mu^k}{k!}$$

Claim If $X \sim Pois(\mu)$ and $Y \sim Pois(\lambda)$ are independent then

$$S = X + Y \sim Pois(\mu + \lambda).$$

Pf/
$$P(S=s) = P(X=0, Y=s) + P(X=1, Y=s-1) + \dots P(X=s, Y=0)$$

addition rule

$$= \sum_{k=0}^{s} P(X=k, Y=s-k)$$

independence of X, Y

$$= \sum_{k=0}^{s} P(X=k) P(Y=s-k)$$

$$= \sum_{k=0}^{s} \frac{e^{-\lambda} \mu^{k}}{k!} \cdot \frac{e^{-\mu} \lambda^{s-k}}{(s-k)!}$$

$\frac{s!}{s!} = 1$

$$= e^{-(\lambda+\mu)} \frac{1}{s!} \sum_{k=0}^{s} \underbrace{\frac{s!}{k!(s-k)!}}_{=\binom{s}{k}} \mu^{k} \lambda^{s-k}$$

binomial theorem

$$= e^{-(\lambda+\mu)} \frac{1}{s!} (\mu+\lambda)^{s}$$

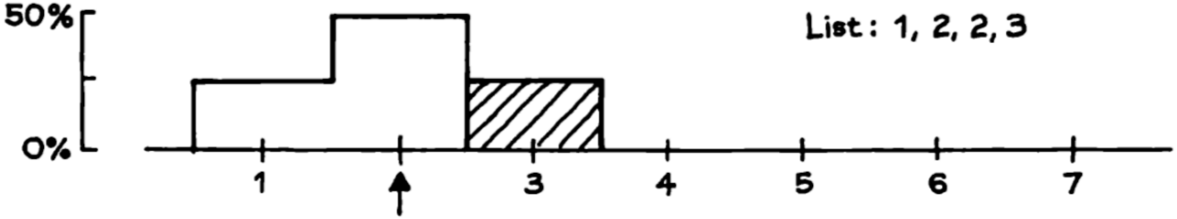
$$\Rightarrow S \sim Pois(\mu+\lambda).$$

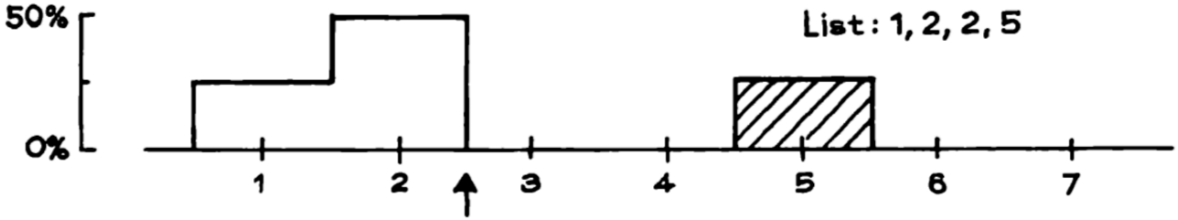
$\square$

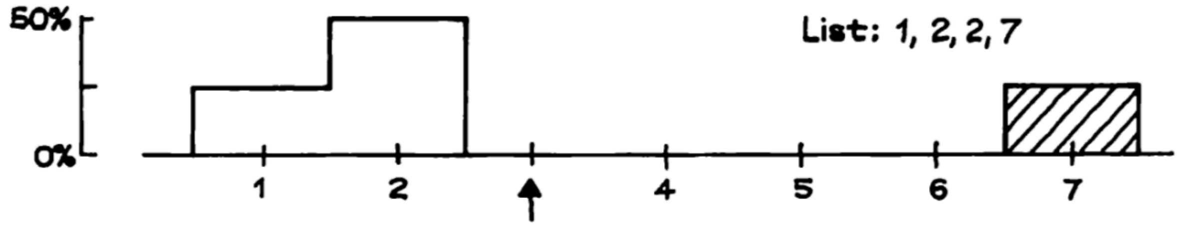
③ Sec 3.2 Expectation

$$E(X) = \sum_{x \in X} x \cdot P(X=x)$$

$$E(X) = \frac{1}{4}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{4}\cdot 2 + \frac{1}{4}\cdot 3 = 2$$







Properties of Expectation — P167 Pitman

1. $E(c) = c$
2. $E(X+Y) = E(X) + E(Y)$ (X, Y <u>don't</u> need to be independent)
3. $E(aX+b) = aE(X) + b$

Indicators

An indicator is a RV that has only 2 values 1 (w/ prob p) and 0 (with prob 1-p).

$$I = \begin{cases} 1 & \text{prob } p \\ 0 & \text{prob } 1-p \end{cases}$$ — same as a Bernoulli p trial.

$$E(I) = 1 \cdot p + 0 \cdot (1-p) = p$$

RV that are counts can often be written as a sum of indicators.

ex $X \sim Bin(n,p)$

successes in n Bernoulli p trials.

ex X = # heads in n flips of p coin

$$X = I_1 + I_2 + \cdots + I_n$$

where $I_j = \begin{cases} 1 & \text{if } j^{th} \text{ trial success} \\ 0 & \text{else} \end{cases}$ p

$$E(X) = \underbrace{E(I_1)}_{= p} + \cdots + \underbrace{E(I_n)}_{= p} = \boxed{np}$$

indicators are independent since trials are indep.

Ex $X \sim \text{Hyper}(N, G, n)$

ex $X = \#$ aces in a poker hand from a deck of cards

$N = 52$
$G = 4$
$n = 5$

a) what are the range of values of X?

$0, 1, 2, 3, 4$

b) write X as a sum of indicators

$X = I_1 + I_2 + I_3 + I_4 + I_5$

c) How is I_2 defined?

$I_2 = \begin{cases} 1 & \text{if } 2^{nd} \text{ card is an ace} \\ 0 & \text{else} \end{cases}$

d) Find $E(I_2)$

$E(I_2) = P(2^{nd} \text{ card is an ace}) = \boxed{\frac{4}{52}}$

e) Find $E(X)$

$E(X) = 5 \cdot E(I_1) = \boxed{5 \cdot \left(\frac{4}{52}\right)}$

Note One student in class suggested we define $I_2 = \begin{cases} 1 & \text{if get 2 aces} \\ 0 & \text{else} \end{cases}$

so $X = I_1 + 2 \cdot I_2 + 3 \cdot I_3 + 4 \cdot I_4$

This is also correct but more complicated that my solution.

We have

$$E(I_1) = \frac{\binom{4}{1}\binom{48}{4}}{\binom{52}{5}}$$

$$E(I_2) = \frac{\binom{4}{2}\binom{48}{3}}{\binom{52}{5}}$$

$$E(I_3) = \frac{\binom{4}{3}\binom{48}{2}}{\binom{52}{5}}$$

$$E(I_4) = \frac{\binom{4}{4}\binom{48}{1}}{\binom{52}{5}}$$

So $$E(X) = \frac{1}{\binom{52}{5}} \left[\binom{4}{1}\binom{48}{4} + 2 \cdot \binom{4}{2}\binom{48}{3} + 3\binom{4}{3}\binom{48}{2} + 4\binom{4}{4}\binom{48}{4} \right]$$

wow! $= 5 \cdot \left(\frac{4}{52}\right)$ ← I checked this using R

ex Suppose a fair die is rolled 10 times.
Let $X =$ number of different faces that appear in 10 rolls.

ex if roll 2, 3, 4, 2, 3, 5, 2, 3, 3, 2 then $X = 4$

a) what are the range of values of X?

1, 2, 3, 4, 5, 6

b) write X as a sum of indicators

$X = I_1 + I_2 + I_3 + I_4 + I_5 + I_6$

c) How is I_2 defined?

$$I_2 = \begin{cases} 1 & \text{if } 2 \text{ appears at least once,} \\ 0 & \text{else} \end{cases}$$

d) Find $E(I_2)$

$= 1 - P(2 \text{ never appears}) = 1 - \left(\frac{5}{6}\right)^{10}$

e) Find $E(X)$

$$\boxed{6 \cdot \left(1 - \left(\frac{5}{6}\right)^{10}\right)}$$

Stat 134

Chapter 3 Wednesday February 13 2019

1. A forgetful valet is attempting to return n cars to their n rightful owners. For each driver, the valet remembers the car correctly 5% of the time; otherwise the valet retrieves a car at random (possibly the correct car). What is the probability that the second driver retrieves his own car.

   a .05

   b $.05 + .95/n$

   c $.05 + .95/(n-1)$

   d none of the above

we have,

$= P(\text{correct, remembers}) + P(\text{correct, forgets})$

$= \underbrace{P(\text{correct} \mid \text{remember})}_{=1} \cdot \underbrace{P(\text{remember})}_{=.05} + \underbrace{P(\text{correct} \mid \text{forget})}_{=\frac{1}{n}} \cdot \underbrace{P(\text{forget})}_{=.95}$

we arent conditioning on the outcome of returning anyone elses car.

Everyone has the same chance of getting their car back!

2. A forgetful valet is attempting to return n cars to their n rightful owners. For each driver, the valet remembers the car correctly 5% of the time; otherwise the valet retrieves a car at random (possibly the correct car). Let N be the number of drivers who retrieve their own car. $E(N)$ is:

a $.05n + .95$

b $.05n$

c $.05n + 1$

d none of the above

N = # drivers who retrieve their own car

$0, 1, 2, \ldots n.$

$.05 + .95/n$

$N = I_1 + I_2 + \cdots + I_n$, $I_j = \begin{cases} 1 & \text{if } j^{th} \text{ driver retrieves their own car.} \\ 0 & \text{else.} \end{cases}$

$$E(N) = n\left(.05 + \frac{.95}{n}\right) = \boxed{.05n + .95}$$