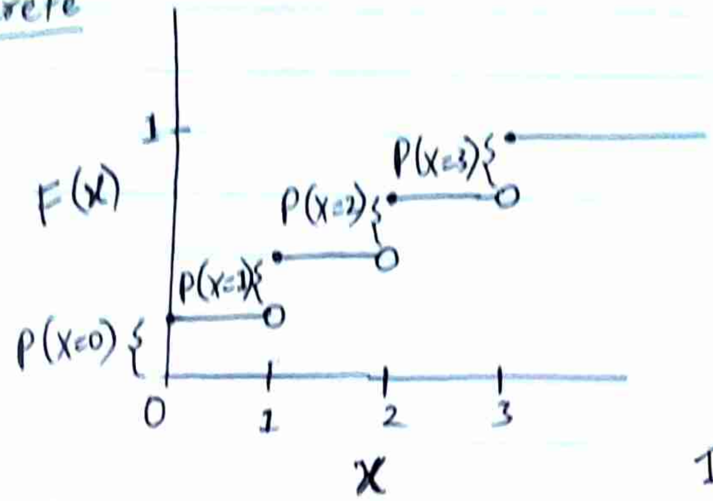
Lecture 9: Expectation, Indicator Random Variables, Linearity

CDF (~~Com~~ Cumulative Distribution Function)

$F(x) = P(X \le x)$ as a function of real x.

→ cdf gives us the entire distribution, from it we can calculate any probability we want for x.

discrete



Find $P(1 < X \le 3)$ using F(cdf).

Solu:

$P(X \le 1) + P(1 < X \le 3) = P(X \le 3)$

$P(1 < X \le 3) = P(X \le 3) - P(X \le 1)$

$= F(3) - F(1)$

In General,

$$\boxed{P(a < X \le b) = F(b) - F(a)}$$

Note - from cdf, we can recover pmf, from pmf, we can recover cdf just by summing things up.

Properties of CDF:

1. increasing (not necessarily strictly increasing)
2. right continuous
3. $F(x) \to 0$ as $x \to -\infty$
   $F(x) \to 1$ as $x \to \infty$

This is "if and only if".

Independence of Random Variables (r.v.s.)

X, Y are independent random variables if

$$P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y), \ \forall x, y$$

Discrete case: $P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$

(Won't work in continuous case)

Note: In discrete case, it's easier to work with PMF than the CDF.

Averages in random variables (Means, Expected Value)

Example

$1, 1, 1, 1, 1, 3, 3, 5$

Method 1: Add all numbers & divide by 8.

Method 2: $\frac{5}{8} \cdot 1 + \frac{2}{8} \cdot 3 + \frac{1}{8} \cdot 5$ (Weighted Average)

$$\frac{1}{n}\sum_{j=1}^{n} j = \frac{(n+1)}{2}$$ (Arithmetic Series) (Unweighted Average)

Average of a discrete random variable X

$$E(X) = \sum_{x} \underbrace{x}_{\text{value}} \underbrace{P(X=x)}_{\text{PMF}}$$, summed over x with $P(X=x) > 0$

Example $X \sim \text{Bern}(p)$

$$E(X) = 1 \cdot P(X=1) + 0 \cdot P(X=0)$$
$$= P(X=1)$$

$$X = \begin{cases} 1, & \text{if } A \text{ occurs} \\ 0, & \text{otherwise} \end{cases}$$

Indicator Random Variable

Then, $\boxed{E(X) = P(A)}$ fundamental bridge

→ The bridge is between expected values and probabilities, this says any problem we want in probability (i.e., if we have any event A, we want $P(A)$ – if we want we can always reinterpret that as the expected value of an indicator.

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

$$E(X) = \sum_{k=0}^{n} k\binom{n}{k}p^k q^{n-k}$$

$$= \sum_{k=0}^{n} n\binom{n-1}{k-1}p^k q^{n-k}$$

$$= n\sum_{k=0}^{n}\binom{n-1}{k-1}p^k q^{n-k}$$

$$= np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1} q^{n-k}$$

$$= np\underbrace{\sum_{j=0}^{n-1}\binom{n-1}{j}p^j q^{n-1-j}}_{=1 \text{ (Binomial theorem)}}$$

$$k\binom{n}{k} = n\binom{n-1}{k-1}$$

→ n people choose committee of size k, with 1 person as president

1. Choose president first, then committee $\binom{n}{1}$
2. Choose committee $\binom{n}{k}$ first, then president $\binom{k}{1}$

$k-1=j$

$k=1, j=0$

$k=n, j=n-1$

$n-(j+1) = n-j-1$

$$\boxed{E(X) = np}$$

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

Single most property of expectation is Linearity.

Linearity

$$E(X+Y) = E(X) + E(Y)$$

Always true, even if X and Y are dependent

$E(cX) = cE(X)$, if c is a constant

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

$E(X) = ?$

Think as sum of n i.i.d. bernoulli P's, each of those bernoulli P's has expected value p and there is n of them. so, by linearity

$$E(X) = np$$

$$\begin{bmatrix} \text{Since, } X = X_1 + X_2 + \dots + X_n \\ X_j \sim Bern(p) \end{bmatrix}$$

(True, even if these were dependent)

Hypergeometric distribution

Example 5 card hand, $X = (\#\text{aces})$

Let X_j be indicator of jth card being an ace, $1 \leq j \leq 5$.

Indicator R.V.

$$\begin{aligned} E(X) &= E(X_1 + X_2 + \dots + X_5) \\ &= E(X_1) + \dots + E(X_5) \text{ [from Linearity]} \\ &= 5E(X_1) \text{ (by symmetry)} \\ &= 5P(\text{1st card is an ace}) \text{ [fundamental bridge]} \\ &= 5/13 \text{ , even though } X_j\text{'s are dependent} \end{aligned}$$

Important

This gives expected value of any hypergeometric distribution.

Geometric Distribution (p) :

Independent Bern(p) trials, count number of failures before the first success.

Let $X \sim Geom(p)$, $q = 1-p$

PMF : $P(X=k) = q^k p$, $k \in \{0,1,2,\dots\}$

FFFFFS

$P(X=5)$

5 failures before first success

$q^5 p$

Valid since

$$\sum_{k=0}^{\infty} p q^k = p/(1-q) = 1$$

$$E(X) = \sum_{k=0}^{\infty} k p q^k \quad \text{(PMF times value)}$$

$$= p \sum_{k=1}^{\infty} k q^k$$

$$= \cancel{p} \cdot q/p^{\cancel{2}}$$

$$= q/p$$

Start (without k)

$$\sum_{k=0}^{\infty} q^k = 1/(1-q) \quad \text{(we know)}$$

differentiate both side,

$$\sum_{k=1}^{\infty} k q^{k-1} = 1/(1-q)^2$$

$$\sum_{k=1}^{\infty} k q^k = q/p^2$$

Story proof: $X \sim Geom(p)$

Let $c = E(X)$

$$c = 0 \cdot p + (1 + c)\, q$$

1 failure

problem restarted (same prob. again)

success first time

failure first time (which happens with prob. q)

Recursion

[Flipping a coin with prob. of head over and over again until coin lands head for the first time, Count the no. of failure]

$$c = q + cq$$

$$c = q/p$$

Lecture 10: Expectation Continued

Proof of Linearity of Expectation

Let $T = X + Y$

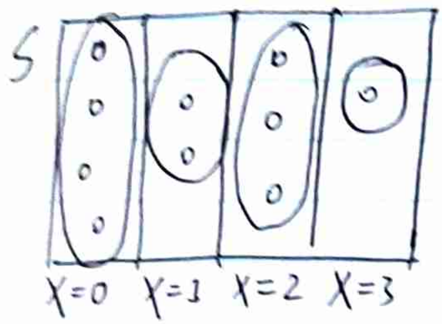
Show $E(T) = E(X) + E(Y)$.

Proof: (Discrete case) $$\sum_t t \underbrace{P(T=t)}_{PMF} = \sum_x x P(X=x) + \sum_y y P(Y=y)$$

Conditioning,

$$P(T=t) \stackrel{?}{=} \sum_x P(T=t \mid X=x) P(X=x) \quad [\text{condition on } X]$$

If X & Y are independent, then we can simplify this, but it is not the case here; it can also be dependent.



Pebble world (Discrete Random variables)

$$\underbrace{E(X) = \sum_x x P(X=x)}_{\text{grouped}}$$

$$\underbrace{= \sum_s X(s) \underbrace{P(\{s\})}_{\text{mass of pebble } s}}_{\text{Ungrouped}}$$

Proof of Linearity (discrete case)

$$E(T) = \sum_s (X+Y)(s) P(\{s\})$$

$$= \sum_s [X(s) + Y(s)] P(\{s\})$$

$$= \sum_s X(s) P(\{s\}) + \sum_s Y(s) P(\{s\})$$

$$= E(X) + E(Y)$$

Similarly, $E(cX) = cE(X)$, if c is a constant.

Extreme case of dependence : $X = Y$

Then,

$$E(X+Y) = E(2X) = 2E(X)$$
$$= E(X) + E(Y)$$

Negative Binomial
parameters r, p $\rightarrow$ generalization of geometric distribution

Story : Independent Bern(p) trials, we want to know the # failures before the rth success.

PMF : $P(X=n)$
$$= \binom{n+r-1}{r-1} p^r (1-p)^n,$$
$$n = 0, 1, 2, \ldots$$

1 0 0 0 1 0 0 1 0 0 0 0 1 0 0 $\boxed{1}$ ← rth success

$r=5$, $n=11$ (11 failure)

(5 successes)

$$\binom{n+r-1}{r-1} p^r (1-p)^n, \; n = 0, 1, 2 \ldots$$

→ no. of terms except the last one (i.e., rth success)

$$E(X) = E(X_1 + \ldots + X_r)$$

X_j is the # failures ~~before~~ between $(j-1)$th and jth success.

Then, $X_j \sim \text{Geom}(p)$

By linearity,
$$E(X) = E(X_1) + \ldots + E(X_r) = r\, q/p$$

[Assume Geometric is starting at 0.]

First Success Distribution

$X \sim FS(p)$, time until 1st success, counting the success (→ Bern(p) trial)

Let $Y = X - 1$. Then $Y \sim \text{Geom}(p)$

$$E(X) = E(Y) + 1$$
$$= q/p + 1 = 1/p$$

→ subtracted 1 because we don't want to count success

✓ Expected value of first success time

$1/p$ is pretty intuitive. Suppose that our prob. of success is 1 in 10. Then, it says, on average it would take 10 trials to get the first success, if we include that success.

Example
(putnam Exam) → toughest maths exam in U.S.

Random permutation of $1, 2, \ldots, n$ (all values are equally likely) where $n \geq 2$.
Find the expected number of local maxima.

$$3 \; 2 \; 1 \; 4 \; 7 \; 5 \; 6$$

→ local maxima

Soln: Let I_j be indicator r.v. of position j having a local maxima, $1 \leq j \leq n$.

$$E(I_1 + \ldots + I_n) = E(I_1) + E(I_2) + \ldots + E(I_n)$$

(by linearity)

$$= \frac{(n-2)}{3} + 2 \cdot \frac{1}{2}$$

$(n-2)$ intermediate points. Each of them is $1/3$rd prob. Expected value of indicator is prob. of event. prob. of event is $1/3$ for each of the intermediate position.

$3 \; 2 \; 1 \; 4 \; 7 \; 5 \; 6$

~~that is~~ ~~1/3 prob~~ the largest number among three is equally likely to be in any of those 3 positions so, $\frac{1}{3}$rd chance that biggest no. is in the middle

2 end points

$$= \frac{(n+1)}{3}$$

[There are only two neighbors. ~~so there is~~ ~~1/2~~ It is equally likely that first number is bigger ~~number~~ than second number or second number is bigger than first.]

$\{3, 2\}, \{5, 6\}$

two end points

Example

St. Petersburg paradox (doesnot involve Indicator Random variable(s))

you get $\$2^X$, where X is the number of flips of fair coin until first H, including success.

How much should you pay to play?

Soln:

$Y = 2^X$, find $E(Y)$.

$$E(Y) = \sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k}$$

$\frac{1}{2^k}$ → tail $(k-1)$ times and then head

$$= \sum_{k=1}^{\infty} 1$$

$$= \infty$$

Bound at 1 trillion (We are not going to get more than trillion dollar from this)

$$10^{12} = (10^3)^4$$

$$\sim \$2^{40}$$

Then, $\sum_{k=1}^{40} 2^k \cdot \frac{1}{2^k} = 40$. $\$40$ we should pay $40 to play the game.

General case

$E(2^X) = \infty \neq 2^{E(X)} = 4$ (this is false)

(for the first success, the expected value is $1/p$ which in this case, is 2)

(Linearity is true)