國立清華大學
NATIONAL TSING HUA UNIVERSITY

EE 306001
Probability

Lecture 7: discrete random variable

李祈均

Quiz

- Wednesday is the quiz
- You can bring a cheat sheet (hand written), but we will collect it

We have covered all the basics foundations of probability

- Now, we are going to introduce a concept of random variable
  - Ways to assign numerical results to the outcomes of an experiment
- We will define what are random variables, describe random variables using probability mass function (PMF)
  - PMF: a compact way of representing how some numerical values are more likely to occur compared to others

Random variables

- An assignment of a value (number) to every possible outcome
  - Imagine the sample space is this class (many students in this class)
  - We are interested in the height of a random student
  - Measuring the height using ‘real number’
  - Then, carry out the experiment

- Mathematically:
  - A function from the sample space Ω to the real numbers
  - Imaging this function takes an argument, which is the outcome of the experiment, i.e., a typical student, and produces the value of that function, which is the height of that particular student
- Notation-wise (note the distinction)
  - Random variable, $\boldsymbol{X}$ – this is a function
  - Numerical value, $\boldsymbol{x}$ – this is a value that the function takes
- Can have multiple random variables for the same sample space
  - In this example, say, the weight of a student

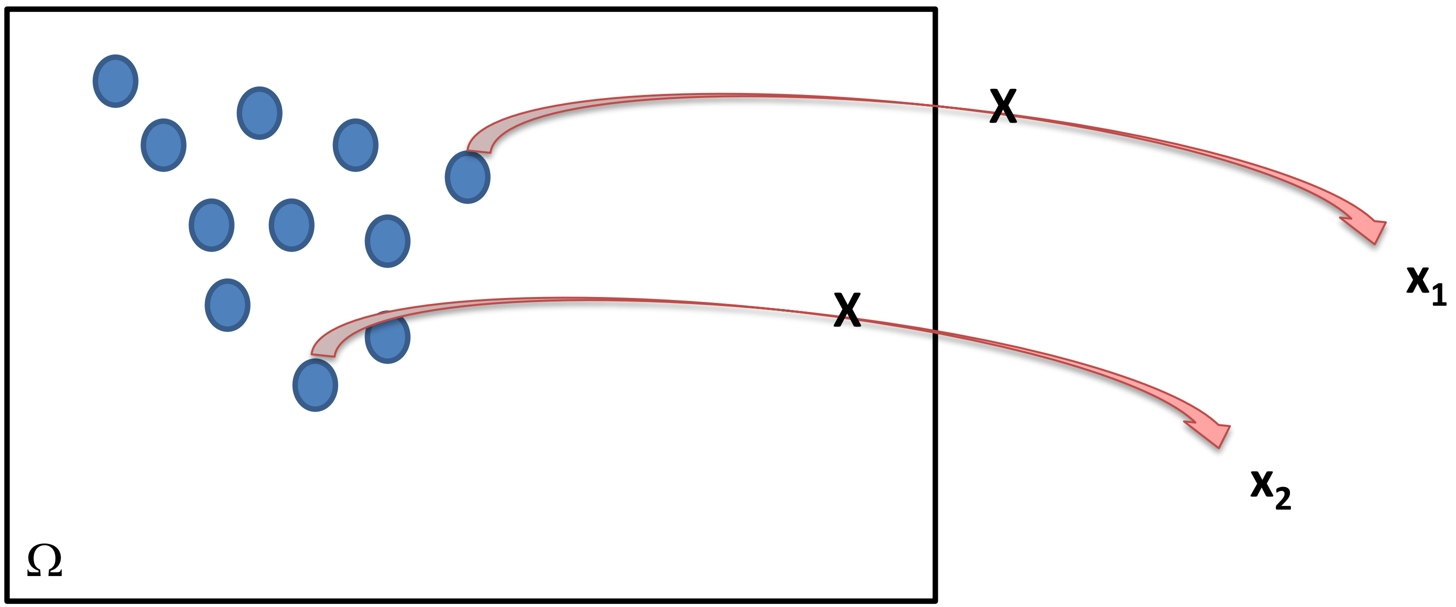
We can be clear that:

- Random variable is not random, is not a variable, an appropriate way to think about it:
- it's just a function from the sample space to the real numbers!

- They can be discrete or continuous
  - Round the height to the nearest centimeter (that's discrete)
  - Infinite precision weight recorder (that's continuous)

- We will start with discrete random variables

- Important concept fundamentally (seems trivial)
  - Random variable is a function (can be thought to be a subroutine in a programming language)
  - X: random variable
  - x: numerical values



We want to describe something about the relative likelihoods of the different numerical values that the random variable can take

We are asking: how likely is it that we obtain an outcome that leads to that particular numerical value

Probability mass function (PMF)

- "probability law, probability distribution of discrete X"
- Let's write out the notation of this probability

$$p_X(x) = P(X = x) = P(\{\omega \in \ \Omega, \text{s.t.}\, X(\omega) = x\}$$

- Since, these are probabilities:
  - $p_X(x) \geq 0$
  - $\sum_x p_X = 1$
- Example: X (random variable) is the number of coin tosses until first head
  - First toss, X =1, second toss X = 2, n^{th} toss X=n
  - Assume independent toss:
    - P(H) = p
    - $p_X(k) = P(X = k) = P(TT \ldots TH) = (1 - p)^{k-1}p, \ \mathrm{k} = 1,2, \ldots$

Geometric PMF

$$p_X(k) = P(X = k) = P(TT \ldots TH) = (1-p)^{k-1}p, \qquad \mathrm{k} = 1,2,\ldots$$

The shape of this PMF is the graph of geometric sequence,

So we call X is a ‘geometric random variable’

This is a simple example, this event could be realized in *one and only one way*: so find the probability of this, we just need to find the probability of this particular outcome

More generally, there’s going to be many outcomes that can lead to the same numerical values…

How to compute PMF

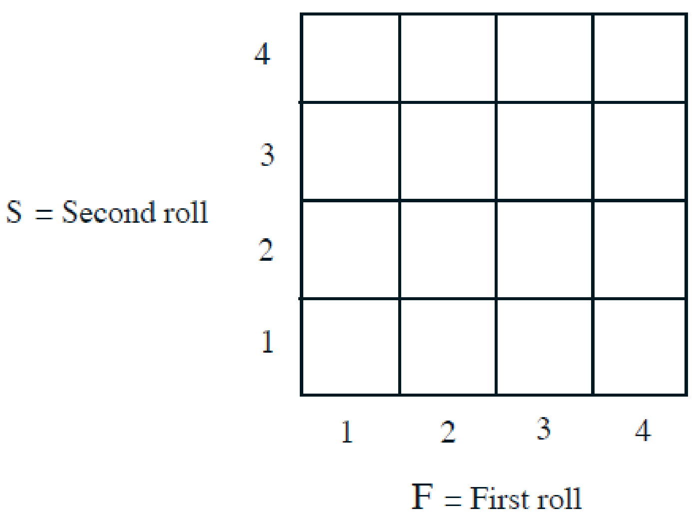
- Collect all possible outcomes for which X is equals to x
- Add their probability
- Repeat for all x

Example: two independent rolls of a fair tetrahedral die

- F: outcome of first throw
- S: outcome of second throw
- $X = \min(F, S)$

Example: two independent rolls of a fair tetrahedral die

- F: outcome of first throw -> random variable
- S: outcome of second throw -> random variable
- X = min(F, S) -> random variable

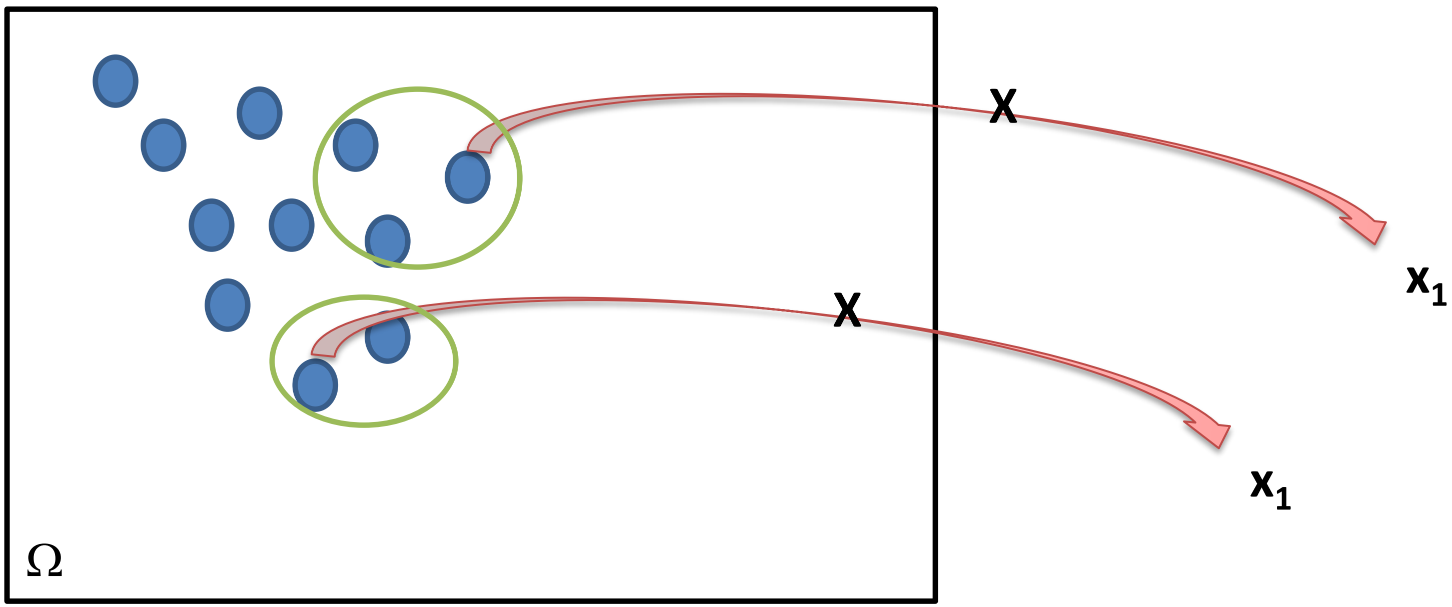


$$p_X(2) = ?$$

There are 5 ways this can happen,

Assuming this die is a fair die, tosses are independent,

We just add up the probabilities of these outcomes 5/16


X
x_1
X
x_1
Ω

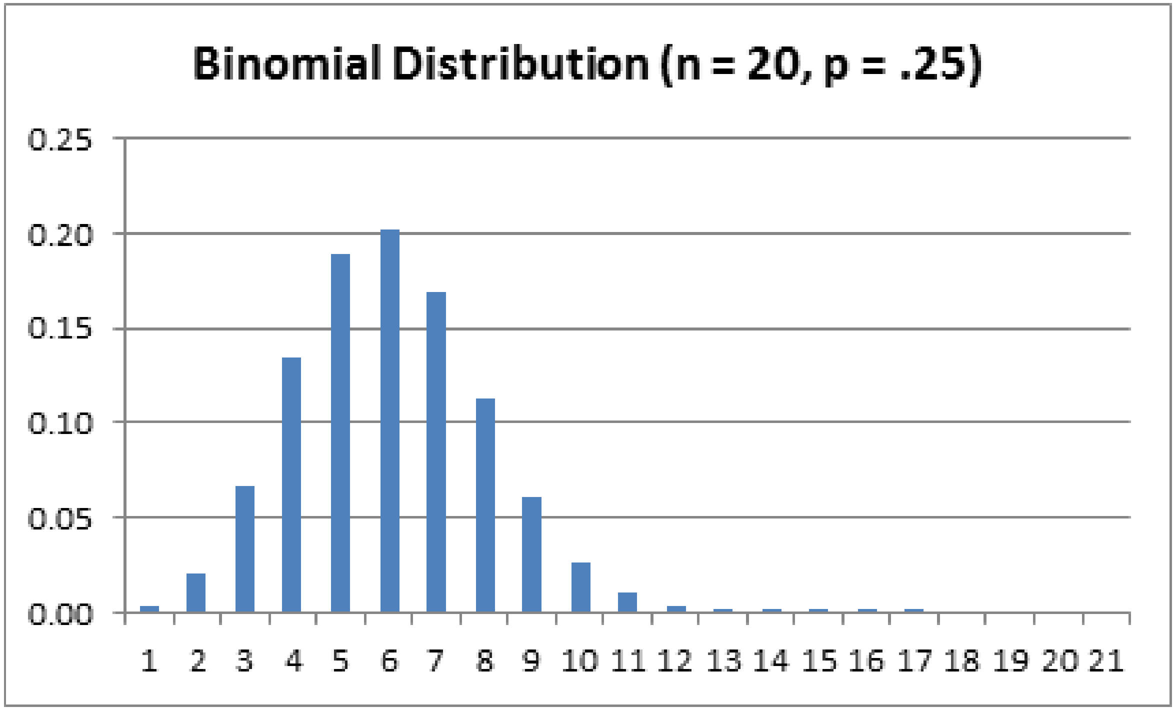
Binomial PMF

- Let's fix a number *n*, we decide to flip a coin *n* consecutive times, each time the coin tosses are independent, and each toss will have probability, *p*, of obtaining heads
  - P(H) = p

- Now consider the random variable, which is the total number of heads that have been obtained
  - Let's try a simple case, n = 4
  - $p_X(2) = P(HHTT) + P(HTHT) + P(HTTH) + P(THHT) + P(THTH) + P(TTHH) = 6p^2(1-p)^2 = \binom{4}{2}p^2(1-p)^2$

In general,

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \mathrm{k}=0,1,2,\ldots,\mathrm{n}$$



Look at this graph, a little like normal distribution? (limiting theorem)

Expectation

Loosely speaking: can be thought as 'average'

- Say you are playing a game, with 1/6 probability you will win 1 dollar, 2 dollars with probability ½, and 4 dollars with probability 1/3
- On average if you play many many time you will win?
  - Naturally add up each dollar won with associated probability (2.5)

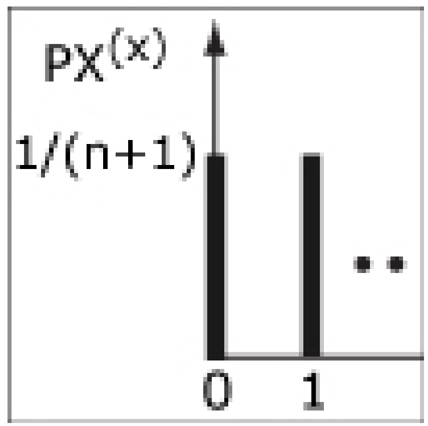
- Definition:

$$E[X] = \sum_{x} x p_X(x)$$

Interpretation of expected value

– Average in a large number of repetition of the experiment
– Center of gravity in the PMF

- Example: uniform distribution on 0,1,…n



$$E[X] = 0 * \frac{1}{N+1} + 1 * \frac{1}{N+1} + 2 * \frac{1}{N+1} + \cdots + N * \frac{1}{N+1} = \frac{N}{2}$$

Properties of expectation

Let X be a r.v., let Y = g(X)

- Hard: $E[Y] = \sum_y y p_Y(y)$
- Easy: $E[Y] = \sum_x g(x) p_X(x)$

- Properties:
  - if α and β are constants
  - $E[\alpha] = \alpha$
  - $E[\alpha X] = \alpha E[X]$
  - $E[\alpha X+\beta]= \alpha E[X] + \beta$

Variance

Recall: $E[g(X)] = \sum_x g(x)p_X(x)$

- First moment: $E[X] = \sum_x xp_X(x)$
- Second moment: $E[X^2] = \sum_x x^2 p_X(x)$
- Variance: second moment of 'centered random variable'
  - On average expected variation (fluctuation) measure

$$\mathrm{var}(X) = E[(X - E[X])^2] = \sum_x (x - E[X])^2 p_X(x)$$

$$= E[X^2] - (E[X])^2$$

Properties:

- var(X) $\geq$ 0
- var(αX + β) = α^2var(X)

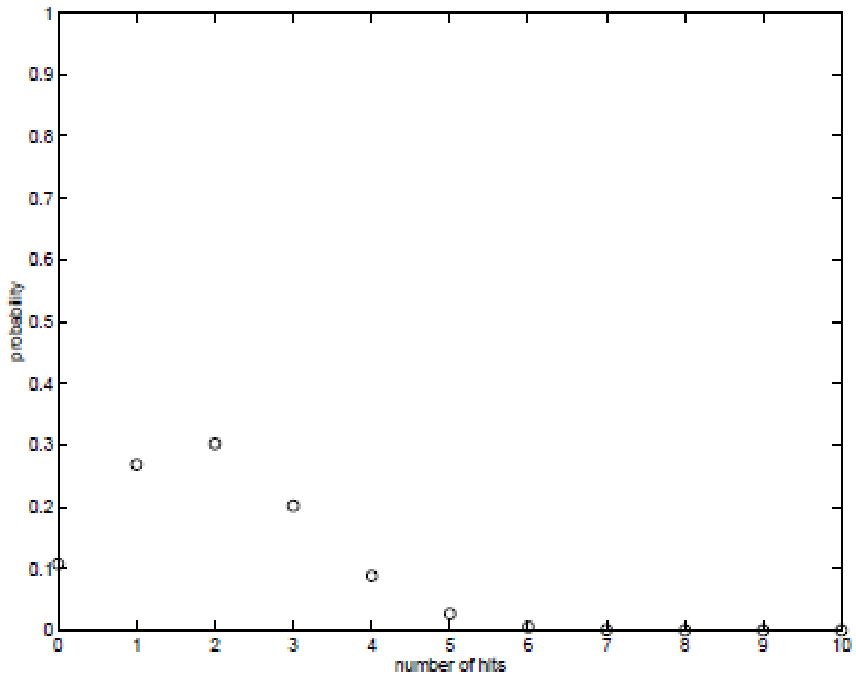
p1

A marksman takes 10 shots at a target and has probability of 0.2 hitting the target with each shot, independently of all other shots. Let X be the number of hits.

a) Calculate and sketch the PMF of X

- X is what? A Binomial random variable, n=10, p=0.2, therefore:

$$p_X(k) = \binom{10}{k} 0.2^k 0.8^{10-k}, \text{for k} = 0, \ \ldots, 10$$



b) What is the probability of scoring no hits

- P(No hits) = $P_X(0)$ = $(0.8)^{10}$ = 0.1074

c) What is the probability of scoring more hits than misses?

- P(more hits than misses) = $\sum_{k=6}^{10} p_X(k) = \sum_{k=6}^{10} \binom{10}{k} 0.2^k 0.8^{10-k} = 0.0064$

d) Find the expectation and variance of X

- Since X is Binomial random variable

$$E[X] = np = 10 \; x \; 0.2 = 2$$

$$var[X] = np(1-p) = npq = 10 \; x \; 0.2 \; x \; 0.8 = 1.6$$

(take it for granted for now, we will derive the formula later)

e) Suppose the marksman has to pay $3 dollar to enter the shooting range, and he gets $2 dollars for each hit. Let Y be his profit. Find the expectation and variance of Y

- So what is Y?
- $Y = 2X - 3$
- $E[Y] = 2E[X] - 3 = 1$
  - $E[aX + b] = aE[X] + b$

- Var[Y] = 4var(X) = 6.4
  - $var(aX + b) = a^2 var(X)$

- Random variable -> algebra -> many things that we can do now

p2

- 4 buses carrying 148 job-seeking students arrive at a job convention. The buses carry 40, 33,25, 50 students, respectively. One of students is randomly selected. Let X denote the number of students that were on the bus carrying this randomly selected student. One of the 4 bus drivers is also randomly selected. Let Y denote the number of students on his bus.

Which of E[X] or E[Y] do you think is larger?

- Give it in word reasoning
  - We expect E[X] to be higher than E[Y] since if we choose the student (just merely because the number of students), we are more likely to pick a bus with more students
  - More likely (higher probability)

Compute E[X], E[Y]

- First compute PMF of each random variable

- $p_X(x) = \begin{cases} \frac{40}{148}, x = 40 \\ \frac{33}{148}, x = 33 \\ \frac{25}{148}, x = 25 \\ \frac{50}{148}, x = 50 \\ 0, \text{otherwise} \end{cases}$

- $E[X] = 40 * \frac{40}{148} + 33 * \frac{33}{148} + 25 * \frac{25}{148} + 50 * \frac{50}{148} = 39.28$

- $E[Y] = 40 * \frac{1}{4} + 33 * \frac{1}{4} + 25 * \frac{1}{4} + 50 * \frac{1}{4} = 37$

p3

Consider a random variable X such that

$$p_X(x) = \frac{x^2}{a} \text{ for } x \in \{-3, -2, -1, 1, 2, 3\}$$

$$P(X = x) = 0 \text{ for x not in } \{-3, -2, 0, 1, 2, 3\}$$

Where a > 0 is a real parameter

a) Find a?

$$1 = \sum_{x=-3}^{3} p_X(x) = \frac{28}{a}$$

a = 28

b) What is PMF of the random variable Z=X^2

x	-3	-2	-1	1	2	3
$p_X(x)$	9/28	1/7	1/28	1/28	1/7	9/28
$Z\|X = x$	9	4	1	1	4	9

- We can see that Z can take only three possible values with non-zero probability, namely, 1,4,9. For each value, it each corresponds to two values of X. So we have,
- $p_Z(9) = P(Z = 9) = P(X = -3) + P(X = 3) = p_X(-3) + p_X(3)$

- Work out the rest of stuff

$$p_Z(z) = \begin{cases} \frac{1}{14}, z = 1 \\ \frac{2}{7}, z = 4 \\ \frac{9}{14}, z = 9 \end{cases}$$

p4

- St Petersburg paradox
- You toss independently a fair coin and you count the number of tosses until the first tail appears. If this number if n, you receive 2^n dollars. What is the expected amount that you will receive?
- How much would you be willing to pay to play this game?

$$\sum_{k=1}^{\infty} 2^k 2^{-k} = \sum_{k=1}^{\infty} 1 = \infty$$

So, if you are faced with the choice of playing for given fee $\boldsymbol{f}$ or not playing at all, and your objective is to make the choice that maximizes your expected net gain, you should be willing to pay any value of $\boldsymbol{f}$!

In experiment, people are willing to pay only about $20 to $30 to play the game.

There are other theory that tried to explain this paradox

Some PMF distribution

- Bernoulli random variable
  - Consider toss of a coin, which comes up with probability p, and a tail with probability 1-p
  - The Bernoulli random variable takes the two value
  - $X = \begin{cases} 1, & \text{X=0} \\ 0, & \text{X=1} \end{cases}$
  - PMF, $p_X(x) = \begin{cases} p, & \text{if k=1} \\ 1-p, & \text{if k=0} \end{cases}$
  - Cases:
    - State of telephone: free or busy
    - State of a person: healthy or sick

- Expectation of Bernoulli variable

$$E[X] = \sum_k X_k P(X = X_k) = 1 * p + 0 * (1 - p) = p$$

- Variance of Bernoulli variable

$$Var[X] = \sum_k \left(x_k \; - E(X)\right)^2 P(X = x_k)$$
$$= (1 - p)^2(p) + (0 - p)^2(1 - p) = p(1 - p)$$

What is parameter of this PMF? *p*

Binomial PMF

- n independent tosses, each with success probability of p
- X ~ B(n,p)
- Say Y is Bernoulli distribution
  - $X = \sum_{i=1}^{n} Y_i$

What the E[X]?

$$E(X) = E\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} E(Y_i) = \sum_{i=1}^{n} p = np$$

- Variance of X?

Let's first try something: (say X1 X2 are independent variables)

$$Var(X1 \ + X2) \ = \ E((X1 + X2)2) - \ (E(X1 + X2))2$$

$$(E(X1 + X2))2 = \ (E(X1) \ + E(X2))2 = \ E(X1)2 \ + E(X2)2 \ + 2E(X1)E(X2)$$

$$E((X1 + X2)2) = \ E(X12 \ + X22 \ + 2X1X2) \ = \ E(X12) \ + \ E(X22) \ + \ 2E(X1X2)$$

Independent of X1 X2, $P(X1, X2) \ = \ P(X1)P(X2)$

$$= \ E(X12) \ + \ E(X22) \ + \ 2E(X1)E(X2)$$

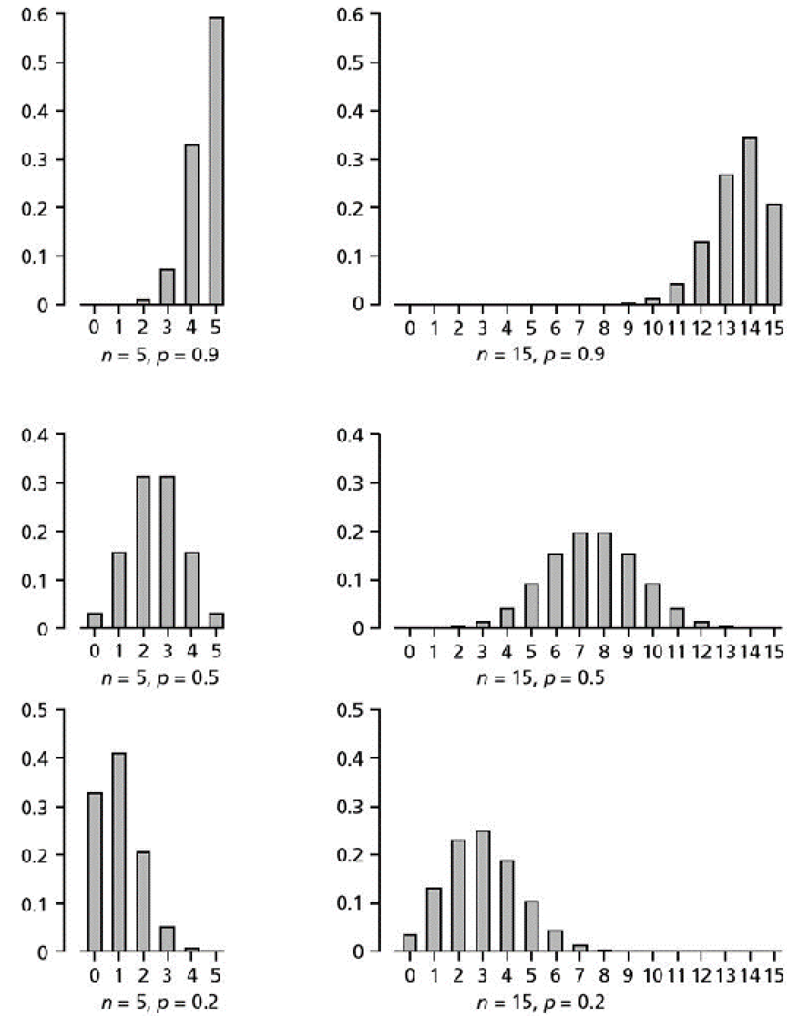
Put it together:

$$E(X12) \ + \ E(X22) \ + \ 2E(X1)E(X2) \ - \ E(X1)2 \ - E(X2)2 \ - \ 2E(X1)E(X2)$$

$$= \ E(X12) \ - \ E(X1)2 \ + \ E(X22) \ - E(X2)2$$

$$= Var(X1) \ + \ Var(X2)$$

- n independent tosses, each with success probability of p
- Say Y is Bernoulli distribution
  - $X = \sum_{i=1}^{n} Y_i$
- X ~ B(n,p)



- Var(Y) = ?
  - $\sum_i p(1-p)$
  - n p q , where q= (1-p)

Poisson distribution

- How to imagine this?
- Imagine this as binomial random variable with very small p and very large n
  - For example, let X be the number of typos in a book with a total of n words
  - Or the number of car involved in accidents in a city on a given day
  - The number of phone calls received by a call center per hour
  - The number of taxis passing particular street corner per hour

expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event.

Poisson distribution

- $p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 1,2,3 \ldots$
- λ is the parameter of this distribution (number of events occurring within an interval)

Approximates Binomial:

Take n=100, p=0.01 (=> k is 5, λ = np =1)

Using Binomial: $\frac{100!}{95!5!} 0.01^5 (1 - 0.01)^{95} = 0.00290$

Using Poisson: $e^{-1} \frac{1}{5!} = 0.00306$

Poisson distribution

- Expectation?

$$E[X] = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!}$$

Take m=k-1

$$= \lambda \sum_{m=0}^{\infty} e^{-\lambda} \frac{\lambda^m}{m!} = \lambda$$

Same thing can be used to derive variance: λ

Graphs!