Review (2.2,2.3)

- Mathematical expectation (mean,variance,standard deviation, moment)
- Properties of mathematical expectation (mean,variance)
- Hypergeometric distribution

Today's Lecture (2.4)

- Bernoulli distribution and its mean, variance.
- Binomial distribution and its p.m.f., properties.
- Cumulative distribution function and its properties.
- Geometric distribution and Negative binomial distribution

Bernoulli distribution

Definition: A random experiment is called a set of Bernoulli trials if it consists of several trials such that

- Each trial has only 2 possible outcomes (usually called "Success" and "Failure").
- The probability of success p, remains constant for all trials;
- The trials are independent, i.e. the event "success in trial i" does not depend on the outcome of any other trials.

Examples: Repeated tossing of a fair die: success="6", failure="not 6". Each toss is a Bernoulli trial with

$$P(\text{success}) = \underline{1/6}$$

Definition: The random variable X is called a **Bernoulli random variable** if it takes only 2 values, 0 and 1.

The p.m.f. $f_X(x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = 0 \end{cases}$ → success

- $E(X) = p \times 1 + (1-p) \times 0 = p$
- $Var(X) = EX^2 - (EX)^2$

$$= p - p^2 = \underbrace{p}_{\text{success probability}} \times \underbrace{(1-p)}_{\text{failure probability}}$$

Binomial distribution

Definition: Let X be the number of successes in n independent Bernoulli trials each with probability of success$=p$. Then X has the Binomial distribution with parameters n and p. We write $X \sim b(n,p)$.

Example:

- number of heads in 7 coin tosses

  Let X= # of heaads, then "success" means head.
  $X \sim b(7, \frac{1}{2})$

- number of correct guesses on a multiple-choice exam with 25 questions and 4 choices each

  $X \sim b(25, \frac{1}{4})$

Binomial or not?

1. $X=$ number of boys before the first girl in a family.

   No, → "geometric"

2. $X=$ number of girls among the next 50 children born in Champaign County hospital

   Yes. girl → "success"
   $X \sim b(50, \frac{1}{2})$
   if we believe boys and girls are equally likely

The p.m.f. of Binomial distribution

If $X \sim b(n, p)$, then the p.m.f. of X is

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \text{ for } x = 0, 1, \cdots, n$$

Explanation:
An outcome with x successes and $n - x$ failures has probability,

$$p^x (1-p)^{n-x}$$

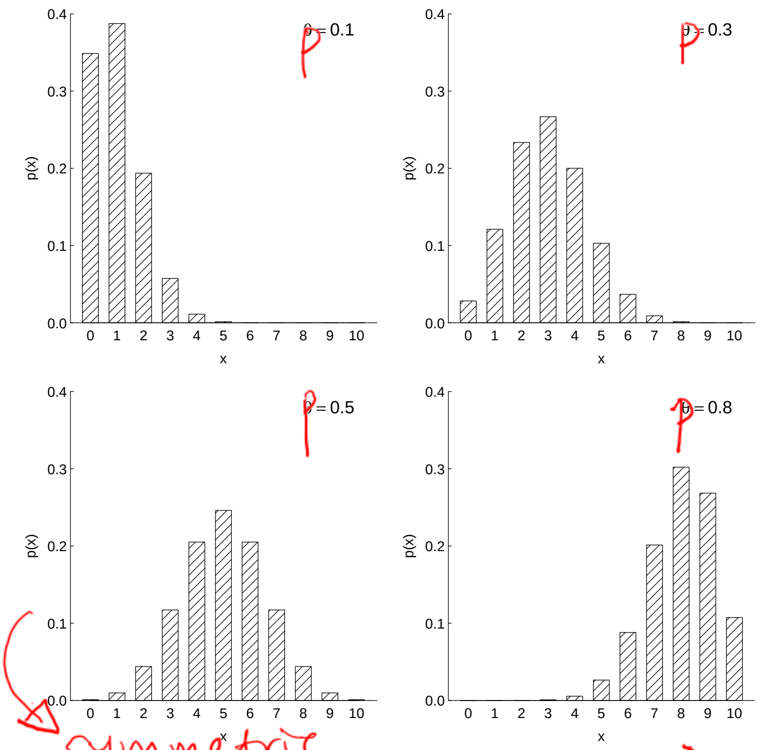
There are $\binom{n}{x}$ possible outcomes with x successes and $(n-x)$ failures.

$$\begin{aligned} P(\#\text{ successes} = x) &= (\#\text{ outcomes with } x \text{ successes}) \\ &\quad \times (\text{prob. of each outcome}) \\ &= \binom{n}{x} p^x (1-p)^{n-x} \end{aligned}$$

Note: $f_X(x) = 0$ if $x \notin \{0, 1, 2, \cdots, n\}$. Check that $\Sigma_{x=0}^{n} f_X(x) = 1$.

Binomial distribution for $n = 10$

skew to left



symmetric

skew to the right

Example: Let X be the number of times I get a '6' out of 10 rolls of a fair die.

$X \sim b(10, \frac{1}{6})$

- What is the distribution of X?
- What is the probability that $X \geq 2$?

$$P(X \geq 2) = 1 - P(X=0) - P(X=1)$$
$$= 1 - \binom{10}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{10} - \binom{10}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^9$$
$$= 0.515$$

compare cdf with p.m.f.

Cumulative distribution function

We have defined the p.m.f. $f(x)$ as $f(x) = P(X = x)$. The *Cumulative distribution function,* c.d.f. or just distribution function $F(x)$, tells us everything there is to know about X.

cdf
$$F(x) = P(X \le x) \text{ for } -\infty < x < \infty$$

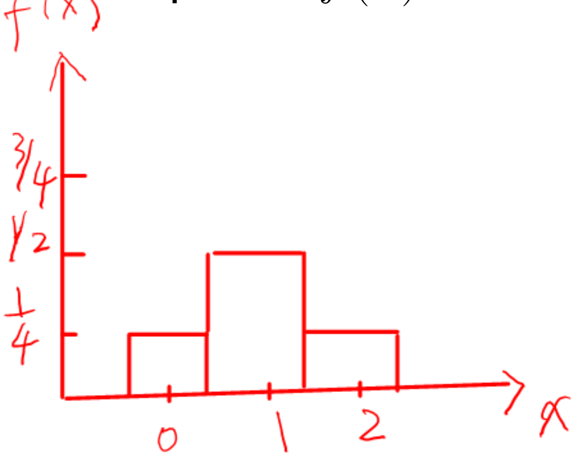
Example: Let $X \sim b(2, 1/2)$. Then

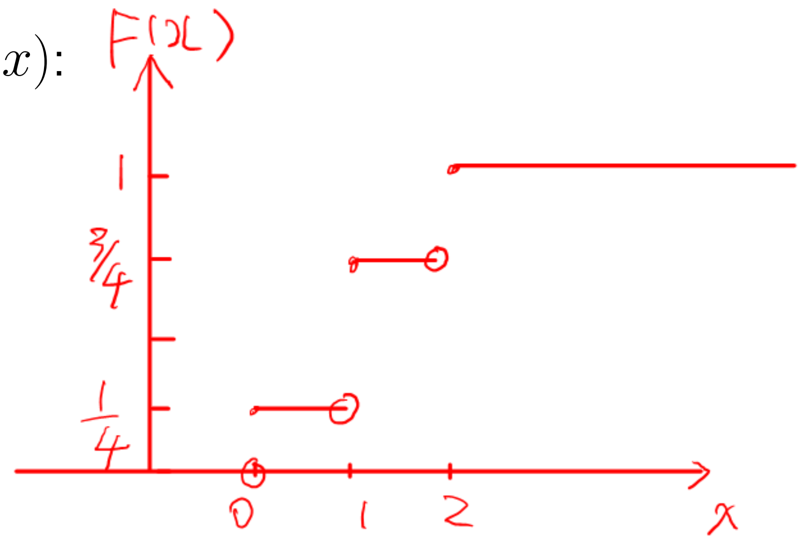
x	0	1	2
$f(x)$	1/4	1/2	1/4

Then

$$F(x) = P(X \le x) = \begin{cases} 0 & \text{if } x < 0 \\ 1/4 & \text{if } 0 \le x < 1 \\ 3/4 & \text{if } 1 \le x < 2 \\ 1 & \text{if } x \ge 2 \end{cases}$$

Graphs of $f(x)$ and $F(x)$:





property of cdf:
non-decreasing
≥ 0
≤ 1

Properties of distribution function

- $F(x)$ gives the cumulative probability up to and including point x. So

$$F(x) = \sum_{y \leq x} f(y)$$

- If X takes integer values, then

$$\begin{aligned} f(x) &= P(X = x) = P(X \leq x) - P(X \leq x - 1) \\ &= F(x) - F(x - 1) \end{aligned}$$

Note: Be careful of endpoints and the difference between $\leq$ and $<$. Assume X takes integer values, then

$$P(X < 10) = P(X \leq 9) = F(9)$$

Example: Let $X \sim b(6, 0.2)$, calculate the following quantities based on the table for Binomial distribution.

x	0	1	2	3	4	5	6
$F(x)$	.2621	.6553	0.9011	0.9830	0.9984	.9999	1.0000

- $P(X < 3)$ $= F(2) = 0.9011$
- $P(X = 2)$ $= F(3) - F(2) = 0.9830 - 0.9011$
- $P(2 < X \leq 5)$ $= F(5) - F(2) = 0.999 - 0.9011$
- $P(2 \leq X < 5) = ?$

be careful about the end point.

Geometric and Negative binomial distributions[1]

- Geometric Distribution:

X = number of independent trials until the 1st "success". Then

$$P(X = x) = (1-p)^{x-1}p, \ x = 1, 2, 3, \cdots.$$

And $EX = 1/p, Var(X) = (1-p)/p^2$.

Example. A slot machine at a casino randomly rewards 15% of the attempts. Assume that all attempts are independent. $\rightarrow p = 0.15$

a) What is the probability that your first reward occurs on your fourth trial?

$P(X=4) = (0.85)^3 \times 0.15$

- Negative Binomial Distribution:

X= number of independent trials until the r-th "success". Then

If $r=1$, then it's geometric!

$$P(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \ x = r, r+1, r+2, \cdots$$

And $EX = r/p, Var(X) = r(1-p)/p^2$.

Note the range of x values.

(b) What is the probability that your third reward occurs on your tenth trial?

$r=3$

$P(X=10) = \binom{9}{2} 0.15^3 \times 0.85^7$

(c) What is the probability that your get three rewards in ten trials?

binomial $X \sim b(10, 0.15)$

[1]Read the last two pages in Lecture 8.

$P(X=3) = \binom{10}{3} \times 0.15^3 \times 0.85^7$

probability in (b) should be smaller than in (c).