BBA182 Applied Statistics
Week 7 (1)Discrete random variables – expected variance and standard deviation
Discrete Probability Distributions

DR SUSANNE HANSEN SARAL

EMAIL: SUSANNE.SARAL@OKAN.EDU.TR

HTTPS://PIAZZA.COM/CLASS/IXRJ5MMOX1U2T8?CID=4#

WWW.KHANACADEMY.ORG

Cumulative Probability Function, F(x_0)

Practical application

The cumulative probability distribution, F(x_0) can be used for example in inventory planning?

Example:

Based on an analysis of it's sales history, the manager of a Toyota car sales department knows that on any single day the number of cars sold can vary from 0 to 5.

Cumulative Probability Function, $F(x_0)$ Practical application: Car dealer

The random variable, X, is the number of possible cars sold in a day:

Table 4.2 Probability Distribution Function for Automobile Sales

x	$P(x)$	$F(x)$
0	0.15	0.15
1	0.30	0.45
2	0.20	0.65
3	0.20	0.85
4	0.10	0.95
5	0.05	1.00

Cumulative Probability Function, $F(x_0)$
Practical application

Example: If there are 3 cars in stock. The car dealer will be able to satisfy 85% of the customers

Table 4.2 Probability Distribution Function for Automobile Sales

x	$P(x)$	$F(x)$
0	0.15	0.15
1	0.30	0.45
2	0.20	0.65
3	0.20	0.85
4	0.10	0.95
5	0.05	1.00

Cumulative Probability Function, $F(x_0)$
Practical application

Example: If only 2 cars are in stock, then 35 % [(1-.65) x 100] of the customers will not have their needs satisfied.

Table 4.2 Probability Distribution Function for Automobile Sales

x	$P(x)$	$F(x)$
0	0.15	0.15
1	0.30	0.45
2	0.20	0.65
3	0.20	0.85
4	0.10	0.95
5	0.05	1.00

Properties of discrete random variables:

Expected value

The expected value, E[X], also called the mean, μ, of a discrete random variable is found by multiplying each possible value of the random variable by the probability that it occurs and then summing all the products:

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

The expected value of **tossing two coins** simultaneously is :

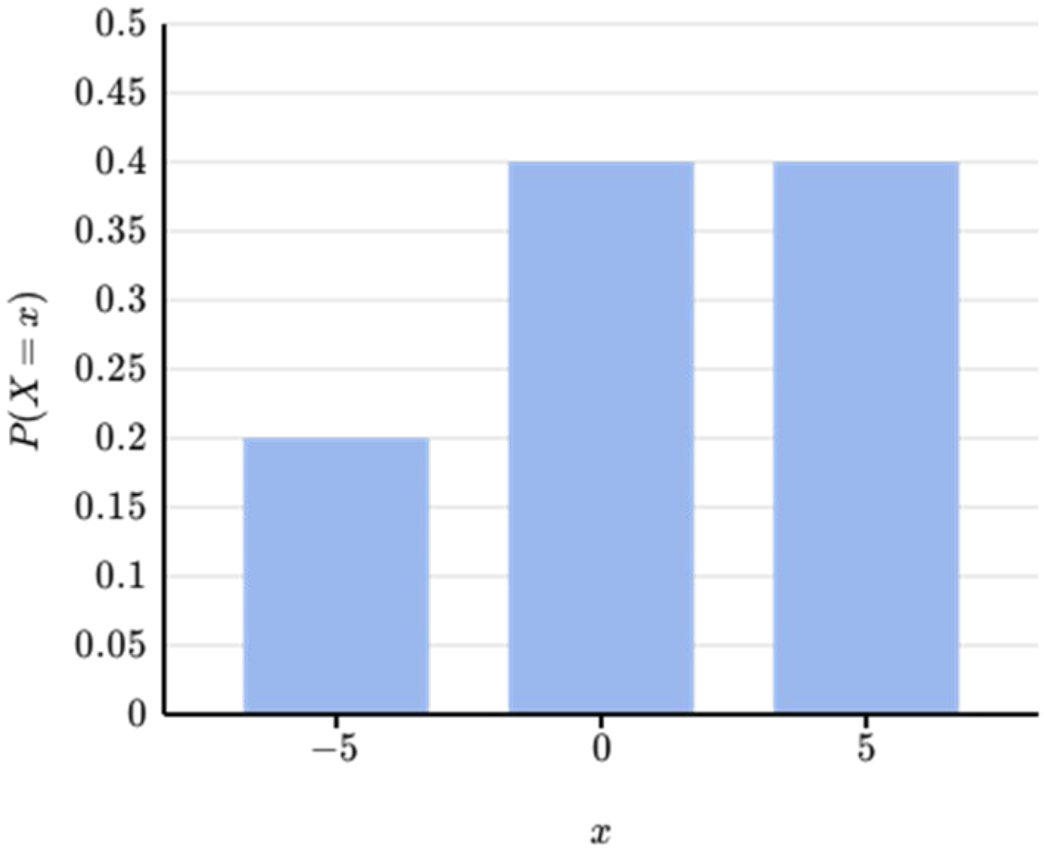
E[x] = (0 x .25) + (1 x .50) + (2 x .25) = 1.0

Expected value for a discrete random able
Exercise

X is a discrete random variable. The graph below defines a probability distribution, P(X) for X.

What is the expected value of X?

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

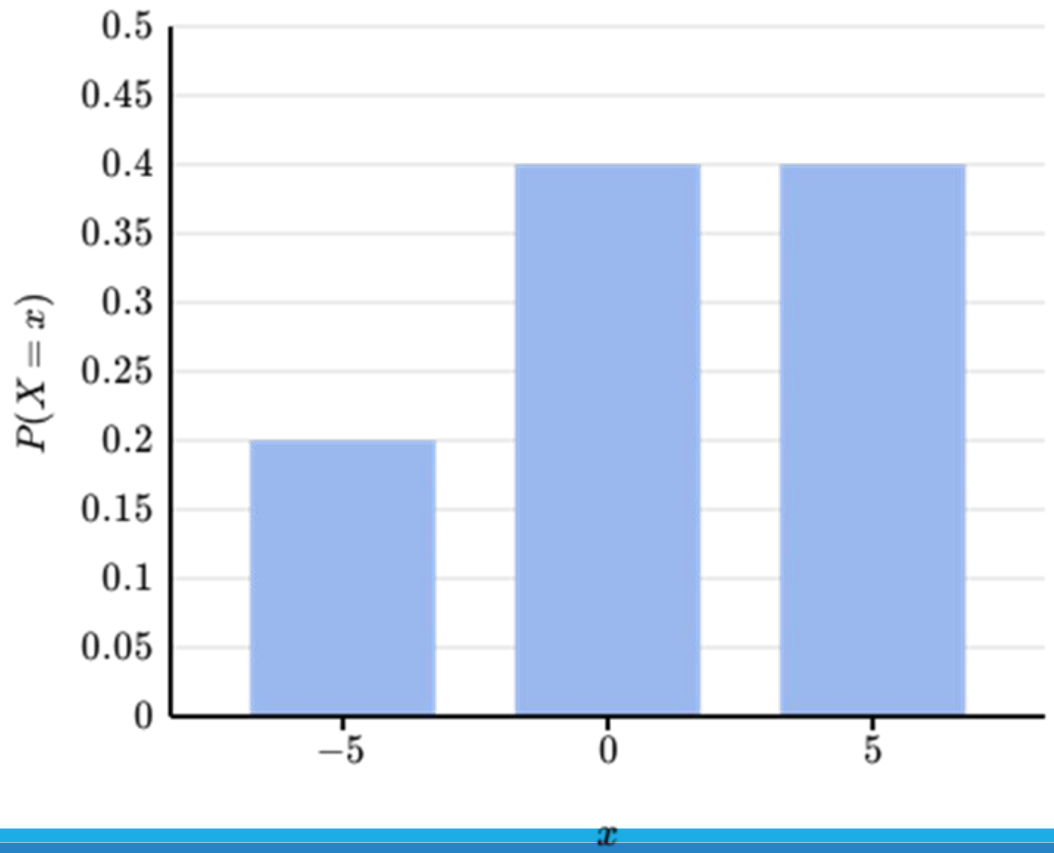


Expected value for a discrete random iable

X is a discrete random variable. The graph below defines a probability distribution, P(X) for X.

What is the expected value of X?

$$E[X] = \mu = \sum_x xP(x)$$



E[X] = (-5)(0.2) + (0)(0.4) + (5)(0.4) = -1 + 0 + 2 = 1

:xpected variance
of a Discrete Random Variables

The measurements of **central tendency** and **variation for discrete random variables:**

- **Expected value E[X]** of a discrete random variable - **expectations**
- **Expected Variance, σ^2**, of a discrete random variable
- **Expected Standard deviation, σ**, of a discrete random variable

Variance of a discrete random variable

The variance is the measure of the spread of a set of numerical observations to the expected value, E[X].

For a **discrete random variable** we define the variance as the **weighted average** of the squares of its possible deviations $(x - \mu)$:

Variance and Standard Deviation

Let X be a discrete random variable. **The expectation** of the average of squared deviations about the mean, $(X-\mu)^2$, is called the **expected variance,** denoted σ^2 and given by:

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 P(x)$$

Expected Standard Deviation of a discrete random variable X

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_x (x-\mu)^2 P(x)}$$

Exercise:
Expected value,E[X], and variance, σ, of car sales

At a car dealer the number of cars sold daily could vary between 0 and 5 cars, with the probabilities given in the table. Find the expected value and variance for this probability distribution

Table 4.2 Probability Distribution Function for Automobile Sales

x	$P(x)$	$F(x)$
0	0.15	0.15
1	0.30	0.45
2	0.20	0.65
3	0.20	0.85
4	0.10	0.95
5	0.05	1.00

Calculation of variance of discrete random variable. Car sales – example

Calculating the expected value:

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

E(x) = (0)(.15)+(1)(.3)+(2)(.2)+(3)(.2)+(4)(.1)+(5)(.05)= 1.95 rounded up to 2 (discrete random variable)

Calculating the expected variance:

$$\sigma^2 = E[(X-\mu)^2] = \sum_{x} (x-\mu)^2 P(x)$$

$$\sigma^2 = (.15)(0-1.95)^2 + (.3)(1-1.95)^2 + (.2)(2-1.95)^2 + (.2)(3-1.95)^2 + (.1)(4-1.95)^2 + (.05)(5-1.95)^2 = \mathbf{2.57}$$

Class exercise

A car dealer calculates the proportion of new cars sold that have been returned a various number of times for the correction of defects during the guarantee period. The results are as follows:

Number of returns	0	1	2	3	4
Proportion P(x)	0.28	0.36	0.23	0.09	0.04

a) Graph the probability distribution function

b) Calculate the cumulative probability distribution

c) What is the probability that cars will be returned for corrections more than two times? $P(x > 2)$

d) $P(x < 2)$?

e) Find the expected value of the number of a car for corrections for defects during the guarantee period

f) Find the expected variance

Dan’s computer Works – class exercise

The number of computers sold per day at Dan’s Computer Works is defined by the following probability distribution:

x	0	1	2	3	4	5	6
P(x)	0.05	0.1	0.2	0.2	0.2	0.15	0.1

Calculate the expected value of number of computer sold per day:

$$E(X) = \sum_{i=1}^{n} X_i P(X_i)$$

$$= X_1 P(X_1) + X_2 P(X_2) + X_3 P(X_3) + X_4 P(X_4) + X_5 P(X_5)$$

Dan's computer Works – class exercise

The number of computers sold per day at Dan's Computer Works is defined by the following probability distribution:

x	0	1	2	3	4	5	6
P(x)	0.05	0.1	0.2	0.2	0.2	0.15	0.1

Calculate the expected value of number of computer sold per day:

$$E(X) = \sum_{i=1}^{n} X_i P(X_i)$$

$$= X_1 P(X_1) + X_2 P(X_2) + X_3 P(X_3) + X_4 P(X_4) + X_5 P(X_5)$$

E[x]= (0 x 0.05) + (1 x 0.1) + (2 x 0.2) + (3 x 0.2) + (4 x 0.2) + (5 x 0.15) + (6 x 0.1) = 3.25 rounded to 3

Dan’s computer Works – class exercise

The number of computers sold per day at Dan’s Computer Works is defined by the following probability distribution:

x	0	1	2	3	4	5	6
P(x)	0.05	0.1	0.2	0.2	0.2	0.15	0.1

Calculate the variance of number of computer sold per day:

$$\sigma^2 = \text{Variance} = \sum_{i=1}^{n} [X_i - E(X)]^2 P(X_i)$$

Dan's computer Works – class exercise

The number of computers sold per day at Dan's Computer Works is defined by the following probability distribution:

x	0	1	2	3	4	5	6
P(x)	0.05	0.1	0.2	0.2	0.2	0.15	0.1

Calculate the variance of number of computer sold per day:

$$\sigma^2 = \text{Variance} = \sum_{i=1}^{n} [X_i - E(X)]^2 P(X_i)$$

$$\sigma^2 = (0 - 3.25)^2(0.05) + (1 - 3.25)^2(0.1) + (2 - 3.25)^2(0.2) + (3 - 3.25)^2(0.2) +$$

$$(4 - 3.25)^2(0.2) + (5 - 3.25)^2(0.15) + (6 - 3.25)^2(0.1) = 2.69$$

$$\sigma^2 = 2.69$$

Quizz

A small school employs 5 teachers who make between $40,000 and $70,000 per year.

One of the 5 teachers, Valerie, decides to teach part-time which decreases her salary from $40,000 to $20,000 per year. The rest of the salaries stay the same.

How will decreasing Valerie's salary affect the mean and median?

Please choose from one of the following options:

A) Both the mean and median will decrease.

B) The mean will decrease, and the median will stay the same.

C)The median will decrease, and the mean will stay the same.

D) The mean will decrease, and the median will increase.

Khan Academy – Empirical Rule

A company produces batteries with a mean life time of 1’300 hours and a standard deviation of 50 hours. Use the Empirical rule (68 – 95 – 99.7 %) to estimate the probability of a battery to have a lifetime longer than 1’150 hours. $P(x > 1'150 \text{ hours})$

Which of the following is the right answer?

95 %

84%

73%

99.85%

Stating that two events are statistically independent means that the probability of one event occurring is independent of the probability of the other event having occurred.

TRUE

FALSE

The time it takes a car to drive from Istanbul to Sinop is an example of a discrete random variable

True

False

Probability is a numerical measure about the likelihood that an event will occur.

TRUE

FALSE

Suppose that you enter a lottery by obtaining one of 20 tickets that have been distributed. By using the ***relative frequency method***, you can determine that the probability of your winning the lottery is 0.15.

TRUE

FALSE

If we flip a coin three times, the probability of getting three heads is 0.125.

TRUE

FALSE

The number of products bought at a local store is an example of a discrete random variable.

TRUE

FALSE

OKAN

Percent of mission complete

Student level: **Any level** Within mission: **Find topics or skills** Activity from: **All time**

Student Name	■ (red)	■ (blue)	⏲	Points
aayseciftcii	0	0	162	14,359
akbenizirem	6	3	413	28,049
Alsharef Abdalelah Alabdali	0	0	49	4,158
aysebese	0	0	252	19,708
BATUHAN KAYACAN	7	10	187	23,320
Bilal Türkmen	2	15	448	29,861
Bircan Taşkır	2	6	193	23,920
Burak Bülbül	4	0	307	14,450
dogukanatessal	1	0	381	17,273
dorukanelmaci	0	1	57	11,975
Duhan Çelik	0	0	0	0
Elçin Kurşun	3	0	213	21,814
Emre Koç	2	20	766	47,261
Emre Şen	1	0	443	36,846
Erkan Yildirim	0	14	124	17,586
fatih topal	0	0	19	9,100
Gizem Nur Gönüldaş	1	11	223	16,136
Harun Çakmak	2	6	277	24,097
kemal.ozcakir	1	0	294	19,647
Muhammed Bilal Aydin	2	28	419	65,701
ozgurmorcol	0	6	156	14,199
Sercan Gürgen	1	18	747	41,669
Simge Karakurt	6	0	701	40,280
Sıla Yurtseven	2	0	152	11,672
velat çay	0	0	32	2,232
y.enesyakan	0	0	0	0
Yaman Tiryakioğlu	0	0	0	0
zeid bunkheila	0	52	1,095	90,615
Çağatay Baykal	1	18	489	38,340

zeid bunkheila

Make a recommendation

Skills | Recommendations | Videos | Badges | Activity | Focus

Total High school statistics progress

100% progress

- 52 skills mastered
- 0 skills level two
- 0 skills level one
- 0 skills practiced
- 0 skills not started

Activity from: All time — Only show attempted skills

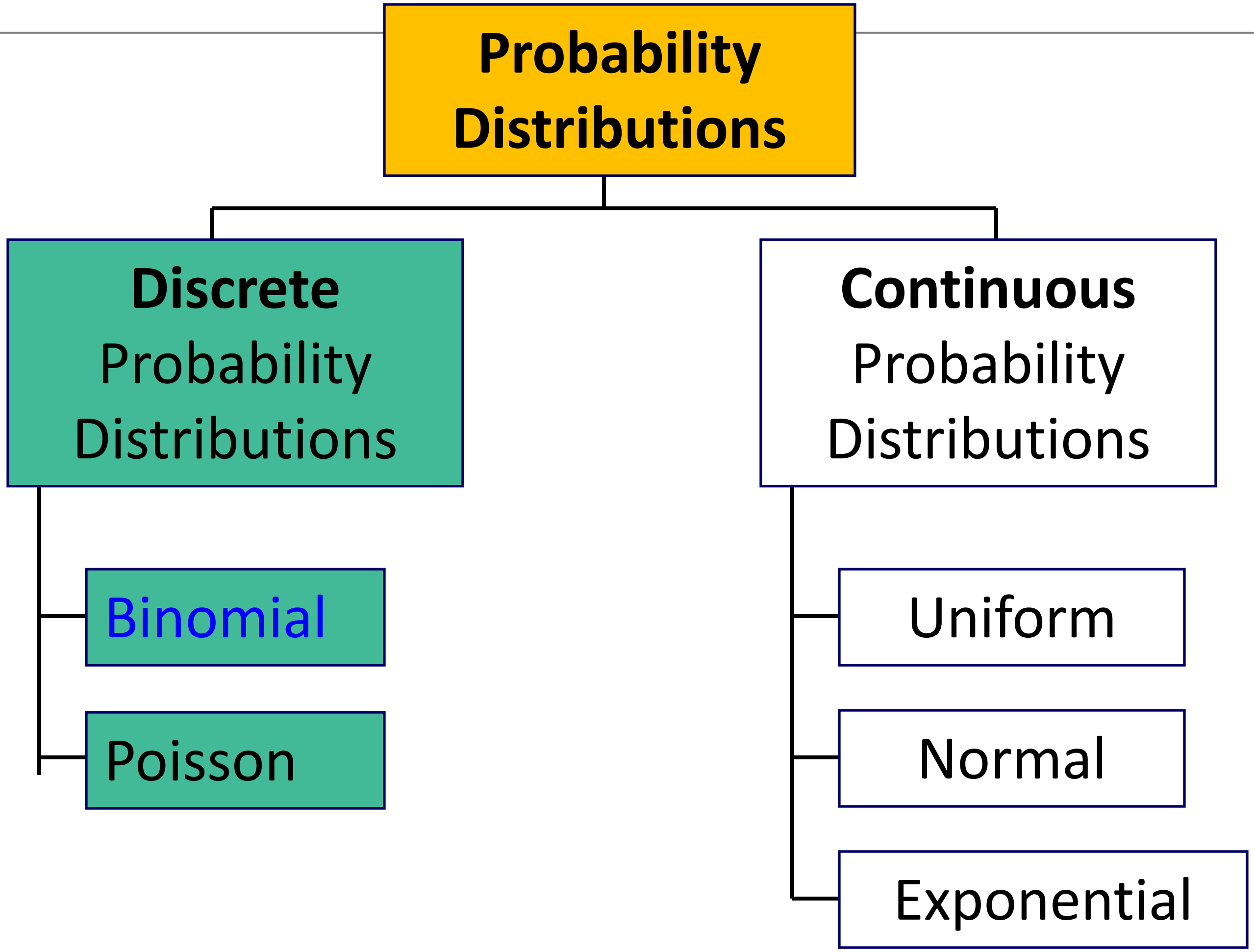
MISSION FOUNDATIONS + 10

Skill	Level	Questions	⏲
Create bar graphs	Mastered	4	3
Creating dot plots	Mastered	25	9
Create histograms	Mastered	35	36
Points on the coordinate plane	Mastered	3	3
Graph from slope-intercept form	Mastered	17	34
Percent word problems	Mastered	3	5
Rewrite decimals as fractions	Mastered	3	3
Comparing rational numbers	Mastered	8	9
Exponents	Mastered	4	1
Square roots	Mastered	3	1

SCATTERPLOTS + 8

Skill	Level	Questions	⏲
Constructing scatter plots	Mastered	6	9
Making good scatter plots	Mastered	20	29
Positive and negative linear correlations from scatter plots	Mastered	23	12
Describing trends in scatter plots	Mastered	13	9
Correlation coefficient intuition	Mastered	10	6

Probability Distributions



Binomial Probability Distribution

Bi-nominal (from Latin) means:
Two-names

- **A fixed number of observations, n**
  - e.g., 15 tosses of a coin; ten light bulbs taken from a warehouse
- **Only two mutually exclusive and collectively exhaustive possible outcomes**
  - e.g., head or tail in each toss of a coin; defective or not defective light bulb
  - Generally called “success” and “failure”
  - Probability of success is P , probability of failure is $1 – P$
- **Constant probability for each observation**
  - e.g., Probability of getting a tail is the same each time we toss the coin
- **Observations are independent**
  - The outcome of one observation does not affect the outcome of the other

Possible Binomial Distribution examples

- ✔ A manufacturing plant labels products as either defective or acceptable
- ✔ A firm bidding for contracts will either get a contract or not
- ✔ A marketing research firm receives survey responses of “yes I will buy” or “no I will not”
- ✔ New job applicants either accept the offer or reject it
- ✔ A customer enters a store will either buy a product or will not buy a product

The Binomial Distribution

The binomial distribution is used to find the probability of a **specific or cumulative number of successes in n trials**

We need to know:

n = number of trials
p = the probability of success on any single trial

We let:

r = number of successes
$q = 1 - p$ = the probability of a failure

The Binomial Distribution

The binomial formula is:

Probability of ***r*** success in ***n*** trials $= \dfrac{n!}{r!(n-r)!} p^r q^{n-r}$

The symbol ! means factorial, and $n! = n(n-1)(n-2)\ldots(1)$

$4! = (4)(3)(2)(1) = 24$

Also, 1! = 1 and 0! = 0 by definition

Example:
Calculating a Binomial Probability

What is the probability of one success in five observations if the probability of success is 0.1?

x = 1, n = 5, and P = 0.1

$$P(x=1) = \frac{n!}{x!(n-x)!}P^X(1-P)^{n-X}$$

$$= \frac{5!}{1!(5-1)!}(0.1)^1(1-0.1)^{5-1}$$

$$= (5)(0.1)(0.9)^4$$

$$= .32805$$

Binomial probability - Calculating binomial probabilities

Suppose that Ali, a real estate agent, has 5 people interested in buying a house in the area Ali’s real estate agent operates.

Out of the 5 people interested how many people will actually buy a house if the probability of selling a house is 0.40. $P(X = 4)$?

Solving Problems with the Binomial Formula

$$= \frac{n!}{r!(n-r)!} p^r q^{n-r}$$

Find the probability of 4 people buying a house out of 5 people, when the probability of success is .40

$n = 5$, $r = 4$, $p = 0.4$, and $q = 1 - 0.4 = 0.6$

P(X = 4)?

P(4 successes in 5 trials): $= \frac{5!}{4!(5-4)!} 0.4^4 0.6^{5-4}$

$$= \frac{5(4)(3)(2)(1)}{4(3)(2)(1)1!}(0.01536)(0.6) = .0768$$

Class exerise

$$= \frac{n!}{r!(n-r)!} p^r q^{n-r}$$

Find the probability of 3 people buying a house out of 5 people, when the probability of success is .40

P(X =3) ?

$n = 5$, $r = 3$, $p = 0.4$, and $q = 1 - 0.4 = 0.6$

$P(X = 3)$? $= \frac{n!}{r!(n-r)!} p^r q^{n-r}$

Find the probability of 3 people buying a house out of 5 people, when the probability of success is .40

$n = 5$, $r = 3$, $p = 0.4$, and $q = 1 - 0.4 = 0.6$

P(3 successes in 5 trials):

$$= \frac{5!}{3!(5-3)!} 0.4^3 0.6^{5-3}$$

$$= \frac{5(4)(3)(2)(1)}{(3)(2)(1)(2)(1)} (0.064)(0.36) = .2304$$

eating a probability distribution with the Binomial Formula – house sale example

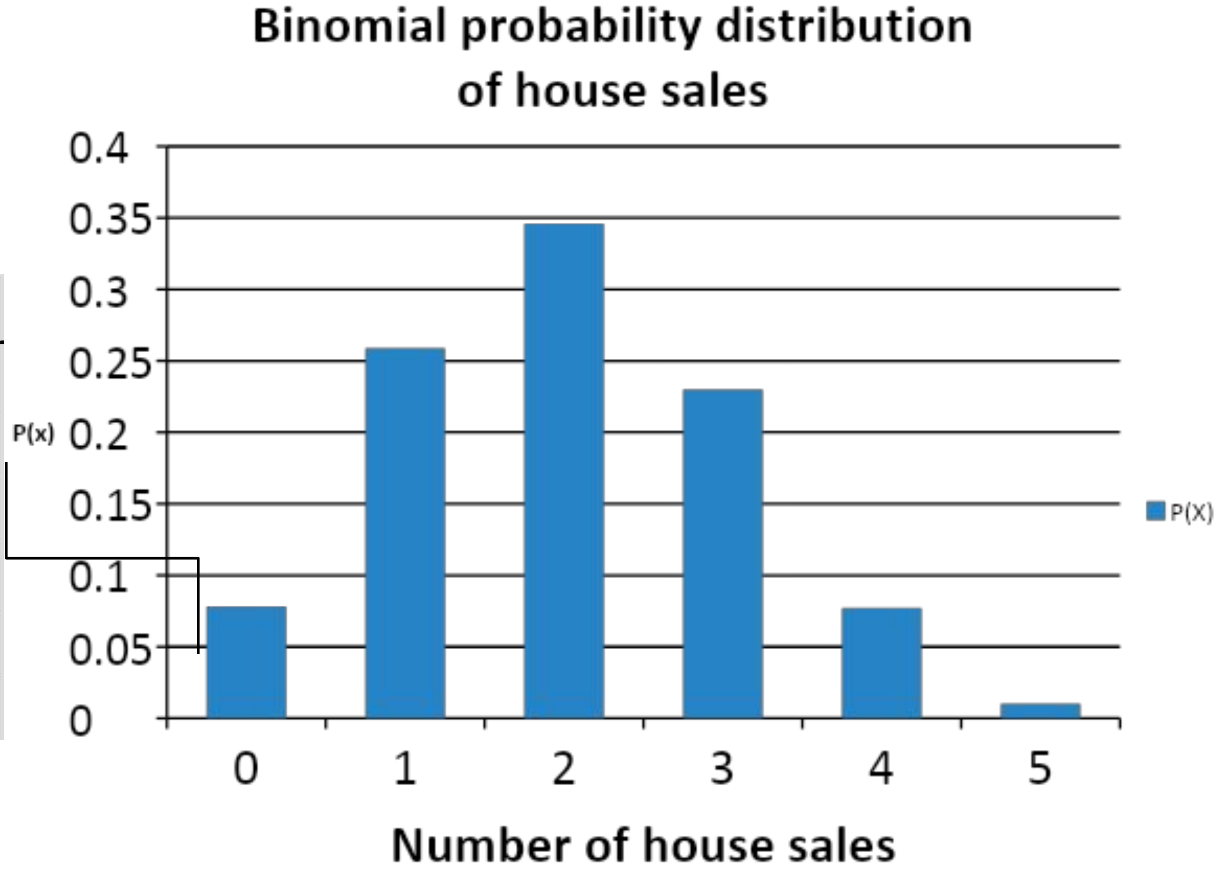
TABLE 2.8 – Binomial Distribution for $n = 5$, $p = 0.40$

NUMBER OF HEADS (r)		PROBABILITY =	$\frac{5!}{r!(5-r)!}(0.5)^r(0.5)^{5-r}$
0	$P(X = 0)$	0.0778 =	$\frac{5!}{0!(5-0)!}(0.5)^0(0.5)^{5-0}$
1	$P(X = 1)$	0.2592 =	$\frac{5!}{1!(5-1)!}(0.5)^1(0.5)^{5-1}$
2	$P(X = 2)$	0.3456 =	$\frac{5!}{2!(5-2)!}(0.5)^2(0.5)^{5-2}$
3	$P(X = 3)$	0.2304 =	$\frac{5!}{3!(5-3)!}(0.5)^3(0.5)^{5-3}$
4	$P(X = 4)$	0.0768 =	$\frac{5!}{4!(5-4)!}(0.5)^4(0.5)^{5-4}$
5	$P(X = 5)$	0.0102 =	$\frac{5!}{5!(5-5)!}(0.5)^5(0.5)^{5-5}$

Binomial Probability Distribution house sale example

n = 5, P= .4

No. House sales	P(X)
0	0.078
1	0.259
2	0.346
3	0.23
4	0.077
5	0.0102



The binomial distribution is used to find the probability of a **specific or cumulative number of successes in n trials.**
Let's look at the cumulative probability: P (x < 2 houses), P(x $\geq$ 3)

NUMBER OF HEADS (r)		PROBABILITY = $\frac{5!}{r!(5-r)!}(0.5)^r(0.5)^{5-r}$
0	P(X = 0)	0.0778 = $\frac{5!}{0!(5-0)!}(0.5)^0(0.5)^{5-0}$
1	P(X = 1)	0.2592 = $\frac{5!}{1!(5-1)!}(0.5)^1(0.5)^{5-1}$
2	P(X = 2)	0.3456 = $\frac{5!}{2!(5-2)!}(0.5)^2(0.5)^{5-2}$
3	P(X = 3)	0.2304 = $\frac{5!}{3!(5-3)!}(0.5)^3(0.5)^{5-3}$
4	P(X = 4)	0.0768 = $\frac{5!}{4!(5-4)!}(0.5)^4(0.5)^{5-4}$
5	P(X = 5)	0.0102 = $\frac{5!}{5!(5-5)!}(0.5)^5(0.5)^{5-5}$

The binomial distribution is used to find the probability of a **specific or cumulative number of successes in *n* trials.**
Let's look at the cumulative probability: P (x < 2 houses), P(x ≥ 3)

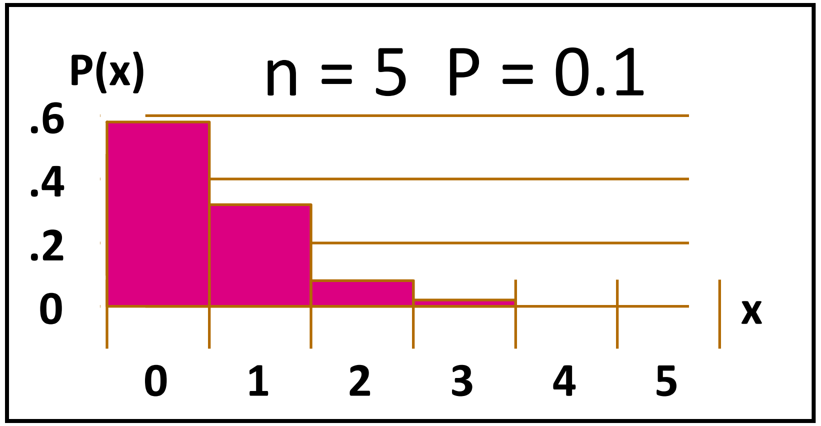
P (x < 2 houses) = P(0 house) + P(1 house) = 0.0778 + 0.2592 = .337 or 33.7%

P(x ≥ 3 houses) = P(3 houses) + P(4 houses) + P(5 houses) = 0.2304 + 0.0768 + 0.0102 = 0.3174

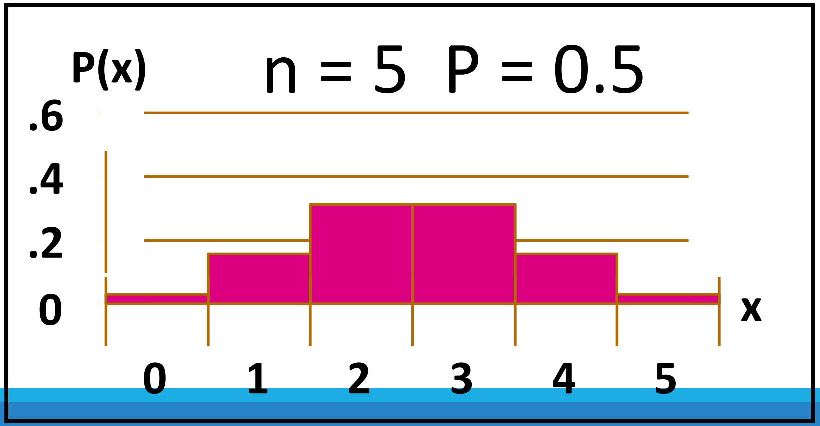
Shape of Binomial Distribution

The shape of the binomial distribution depends on the values of P and n

- Here, n = 5 and P = 0.1



- Here, n = 5 and P = 0.5



Binomial Distribution shapes

When P = .5 the shape of the distribution is ***perfectly symmetrical*** and resembles a bell-shaped (normal distribution)

When P = .2 the distribution is ***skewed right***. This skewness increases as P becomes smaller.

When P = .8, the distribution is ***skewed left.*** As P comes closer to 1, the amount of skewness increases.

Using Binomial Tables instead of to calculating Binomial probabilites

N	x	…	p=.20	p=.25	p=.30	p=.35	p=.40	p=.45	p=.50
10	0	…	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	…	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098
	2	…	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439
	3	…	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172
	4	…	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051
	5	…	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461
	6	…	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051
	7	…	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172
	8	…	0.0001	0.0004	0.0014	0.0043	0.0106	**0.0229**	0.0439
	9	…	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098
	10	…	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010

Examples:

n = 10, x = 3, P = 0.35: $P(x = 3 | n = 10, p = 0.35) = .2522$

n = 10, x = 8, P = 0.45: $P(x = 8 | n = 10, p = 0.45) = .0229$

Solving Problems with Binomial Tables

MSA Electronics is experimenting with the manufacture of a new USB-stick and is looking into the

- Every hour a random sample of 5 USB-sticks is taken
- The probability of one USB-stick being defective is 0.15
- What is the probability of finding 3, 4, or 5 defective USB-sticks ?
  P(x = 3), P(x = 4), P(x= 5)

$n = 5$, $p = 0.15$, and $r = 3, 4,$ or 5

Solving Problems with Binomial Tables

TABLE 2.9 (partial) – Table for Binomial Distribution, n= 5,

		P		
n	***r***	**0.05**	**0.10**	**0.15**
5	0	0.7738	0.5905	0.4437
	1	0.2036	0.3281	0.3915
	2	0.0214	0.0729	0.1382
	3	0.0011	0.0081	0.0244
	4	0.0000	0.0005	0.0022
	5	0.0000	0.0000	0.0001