Jan 16

Chapter 5: Discrete Probability Distributions

Section 5.1: Probability Distributions

(RV) **Random Variable** (X) : is a numerical measure of the outcome of a probability "experiment", so its value is determined by chance.

Capital X

Notation
X = name of RV
x = value X can take

Discrete Random Variable: either finite or countable number of values.

gaps!

go forever

Continuous Random Variable: has (measurable) infinitely many values.

no gaps!

Ex: Identify the random variable and sample space.

(a) Coin toss for Heads

X = # of Heads in 10 coin flips,

the sample space x = 3 (eg.) ($x = 0, 1, \dots, 10$).

(simple event) one outcome HTTHHTTTTT

(b) An experiment that measures the time between arrivals of cars at a drive-through.

continuous

X = time b/w car arrivals,

the sample space t $t > 0$.

0 → t

PROBABILITY DISTRIBUTIONS

Def A **probability distribution** is a description that gives the probability for each value of the random variable (X). It is often expressed in the format of a graph, table, or formula. (x)

REQUIREMENTS

1. $\sum P(x) = 1$ (where x assumes all possible values.)
2. $0 \le P(x) \le 1$ (for every individual value of x.)

Ex: Are the following a probability distribution? If not, state why.

A.

x	$P(x)$
0	0.16
1	0.18
2	0.22
3	0.10
4	0.30
5	0.01

2 ✓
1 ?
$\Sigma P(x) = 1$?
NO $\Sigma = 0.97$
$\Sigma = 0.97$

B. Not PD

x	$P(x)$
0	0.16
1	0.18
2	0.22
3	0.10
4	0.30
5	– 0.01

BAD

C. valid prob. dist.

x	$P(x)$
0	0.16
1	0.18
2	0.22
3	0.10
4	0.30
5	0.04

2 ✓
1 ? ✓
yes, satisfies both conditions!
$\Sigma = 1.0$

Jan 17

Ex: A couple plans to have four children. Let x be the number of boys the couple will have. Find the probability distribution for the number of boys.

X = # of boys out of 4 kids.

$x = 0, 1, 2, 3, 4$

$S = \{GGGG, BGGG, \dots\}$

$\#S = 16$

GGGG – $x = 0$

BGGG, GBGG, GGBG, GGGB } $x = 1$

BBGG, GGBB, GBBG, BGGB, BGBG, GBGB } $x = 2$

BBBG, BBGB, BGBB, GBBB } $x = 3$

BBBB $x = 4$

Probability Dist.

x	$P(x)$
0	1/16 = 0.0625
1	4/16 = 0.25
2	6/16 = 0.375
3	4/16 = 0.25
4	1/16 = 0.0625

$\text{Prob} = \dfrac{\#\text{ of comes}}{\#\text{ sample space}}$

Formulas are used when given a Prob. Dist.

MEAN VALUE

FORMULA: $\boxed{\mu = \Sigma\left[\, x \cdot P(x) \,\right]}$ Note: The Greek letter μ is read "mu"

VARIANCE

FORMULA: $\boxed{\sigma^2 = \Sigma\left[\, (x-\mu)^2 \cdot P(x) \,\right]}$ or $\sigma^2 = \Sigma\left[\, x^2 \cdot P(x) \,\right] - \mu^2$

STANDARD DEVIATION

FORMULA: $\boxed{\sigma = \sqrt{\Sigma\left[\, x^2 \cdot P(x) \,\right] - \mu^2}}$ Note: The Greek letter σ is read "sigma"

Round-Off Rule: Carry one more decimal place than is used for the random variable. → Stats Rule for Rounding

Ex: In a study of brand recognition, random groups of four people are interviewed. Let x be the number of people who recognize Jeff Bezos when shown a picture. The following probability distribution gives the likelihood of the random variable. (RV) X = # of people who recognize Jeff Bezos out of 4

values of RV X

x	$P(x)$	$x \cdot P(x)$	$x^2 \cdot P(x)$
0	0.01	$0 * 0.01 = 0$	$0^2 * 0.01 = 0$
1	0.10	$1 * 0.10 = 0.1$	$1^2 * 0.10 = 0.10$
2	0.24	$2 * 0.24 = 0.48$	$2^2 * 0.24 = 0.96$
3	0.30	$3 * 0.30 = 0.90$	$3^2 * 0.30 = 2.7$
4	0.35	$4 * 0.35 = 1.4$	$4^2 * 0.35 = 5.6$
	$\Sigma P(x) = 1$	$\underbrace{\Sigma\left[\, x \cdot P(x) \,\right]}_{\mu} = 2.88$	$\Sigma\left[\, x^2 \cdot P(x) \,\right] = 9.36$

(a) What is the probability that more than two people will recognize a picture of Jeff Bezos?

$x = 3, 4$ "more than 2"

$P(x=3 \text{ or } x=4) = P(x=3) + P(x=4) - \underbrace{P(x=3 \text{ and } x=4)}_{=0} = 0.3 + 0.35 = \boxed{0.65}$

(b) What is the probability that at most three people will recognize a picture of Jeff Bezos?

$x = \boxed{0\ 1\ 2\ 3}\ 4$ → "inequality" $x \leq 3$

Use the "1 minus trick" $P(\bar{A}) = 1 - P(A)$

$P(x \leq 3) = P(\bar{x=4}) = 1 - P(x=4) = 1 - 0.35 = \boxed{0.65}$

(c) Find the mean number of people who recognize a picture of Jeff Bezos.

$\mu = \Sigma\left[x \cdot P(x)\right] = 2.88 = \boxed{2.9 \text{ people recognize J. Bezos out of 4}}$

rounding rule use RV

(d) Find the standard deviation of the given probability distribution.

$\sigma = \sqrt{\underbrace{\Sigma\left[x^2 \cdot P(x)\right]}_{9.36} - \mu^2} = \sqrt{9.36 - (2.88)^2} = 1.032279 \ldots\ldots$

$\boxed{\sigma = 1.0 \text{ people recognize Bezos out of 4}}$

EXPECTED VALUE

Def The **expected value** (denoted E) of a discrete random variable represents the mean value of the outcomes.

FORMULA: $\boxed{E = \sum[x \cdot P(x)]}$ units of E are same as units of x (or X)

Assume: if win get \$1 back ("bet")

Ex: There is a game in Vegas where you can win \$4 or \$9 but it costs \$1 to play the game. The probability of winning \$4 is 0.3 and the probability of winning \$9 is 0.1. Find the expected value for this game.

x	$P(x)$
-1	0.6
4	0.3
9	0.1

$\sum P(x) = 1$

$0.6 = 1 - 0.3 - 0.1$

- X = value of winning or losing from playing game (once).
- $x = -\$1, \$+4, \$9$
- $E = \sum[x \cdot P(x)] = (-1)*0.6 + (4)*(0.3) + (9)*0.1$

$\boxed{E = \$1.5}$

Interpretation of μ: Over the long run (if we play this game MANY times) we **expect** the mean profit to be \$1.50.

- if E > 0, keep playing • if E < 0, don't even play • if E = 0, up to you!

Ex: When someone buys a life insurance policy, that policy will pay out a sum of money to a benefactor upon the death of the policyholder. Suppose a 25-year-old male buys a \$150,000 1-year term life insurance policy for \$250. The probability that the male will not survive the year is 0.0013.

The experiment has two possible outcomes: survive or die. Let the random variable X represent the money lost or gained by the life insurance company for the 25-year-old male after many years. What is the expected value for the company?

X = money lost or gained by insurance Co.

x =

	x	$P(x)$
survive	\$250	0.9987
die	-\$149,750	0.0013

P(die) = 0.0013

P(survive) = 1 − P(die) = 1 − 0.0013 = .9987

"1 minus trick"

$E = \sum[x \cdot P(x)] = (250)*(0.9987) + (-149750)*(0.0013) = \boxed{\$55}$

Over 1,000 policies, how much should they expect to make?

Insurance Co. expects to make only \$55 per policy.

- 1000 * 0.0013 = 1.3
- 1.3(−149750) + 998.7(250) = 55000 (total made) / 1000 policies

Ex: Find the expected value of the random variable. Round to the three decimal places.

A contractor is considering a sale that promises a profit of \$29,000 with a probability of 0.7 or a loss (due to bad weather, strikes, and such) of \$3,000 with a probability of 0.3. What is their expected profit?

A. \$19,400
B. \$21,200
C. \$20,300
D. \$22,400
E. \$26,000

Section 5.2: Binomial Probability Distributions

Binomial Probability Distribution

Def A **binomial probability distribution** results from a procedure that meets the given requirements.

1. The procedure has a *fixed number of trials*.
2. The trials must be *independent*.
   (i.e. the outcome of an individual trial does not affect the probabilities in other trials.)
3. Each trial must have all outcomes classified into *two categories*.
   (often referred to as success and failure)
4. The probability of a success remains *constant* in all trials.

Notation

Symbol	*Represents*
n	fixed number of trials
p	probability of success
q	probability of failure
x	specific number of success in n trials
$P(x)$	probability of getting exactly x successes among the n trials

Ex: According to Wikipedia, 19% of Mexican residents are vegetarians. If we randomly survey 20 Mexican residents, what's the probability 3 of the Mexicans are vegetarians? Fill in the values given the information presented.

Two Possible Outcomes: ______________________________ or ______________________________.

$n =$ ____________ $p =$ ____________ $q =$ ____________ $x =$ ____________

Two Methods to Find Probability of a Specific Value

Graphing Calculator for ${}_nC_x$
math $\Rightarrow$ PROB $\Rightarrow$ 3: ${}_nC_x$

1. Formula: $P(x) = {}_nC_x \cdot p^x \cdot q^{n-x}$
2. Use Graphing Calculator (TI-83 or 84)

   Instructions: (a) $2^{nd} \Rightarrow$ VARS $\Rightarrow$ DISTR

   (b) binompdf (n, p, x)

Ex: Use the problem from above (regarding vegetarianism in Mexico) to set up and evaluate using both methods.

Ex: We survey 5 PCC students and ask "Is this your first year here?" Assume that 20% of all PCC students are in their first year.

(a) What is the probability that all 5 students are new?

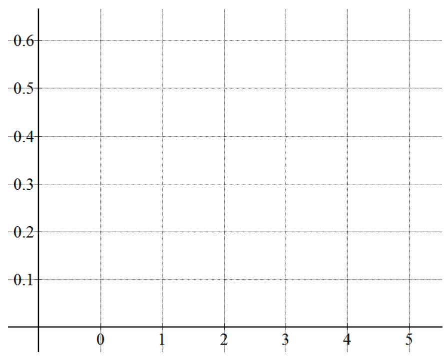
(b) What is the probability that 2 or 3 will be new?

(c) What is the probability that at least one will be new?

(d) Create a probability distribution table for this exercise.

x	$P(x)$

(e)



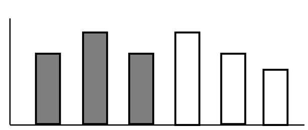
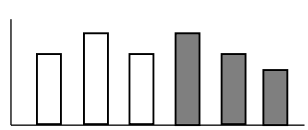
*If we surveyed 200 students, and I ask the probability that between 113 and 187 students are new, how would you find your answer?

Finding a Cumulative Probability

Use Graphing Calculator (TI-83 or 84)

Instructions: (a) 2nd $\Rightarrow$ VARS $\Rightarrow$ DISTR

(b) $P(x \leq \#) = \text{binomcdf}(n, p, \#)$

"no more than" or "at most #" or "less than or equal to #" 	"fewer than #" or "less than #" 	*Calculator
"at least" or "no less than" or "greater than or equal to" 	"more than" or "greater than" 	*Calculator

Ex: A basketball player makes 75% of the free throws he tries. If the player attempts 10 free throws in a game, find the probability that:

(a) the player will make at most six free throws.

(b) the player will make at least eight free throws.

Ex: According to 2017 Washington Post article, approximately 53% of all U.S. households are wireless-only households (no landline). In a random sample of 20 households, what is the probability that…

(a) fewer than 6 are wireless only?

(c) more than 13 are wireless only?

Using Mean and Standard Deviation for Critical Thinking

Mean Value

Formula: $\mu = np$

Variance

Formula: $\sigma^2 = npq$

Standard Deviation

Formula: $\sigma = \sqrt{npq}$

Round-Off Rule: Round to the nearest tenth.

Ex: According to the U.S. Office of Adolescent Health, nearly 90% of adult smokers in America started smoking before turning 18 years old.

(a) If 300 adult smokers are randomly selected, how many would we expect to have started smoking before turning 18 years old?

(b) Would it be unusual (significantly high/low) to observe 240 smokers who started smoking before turning 18 years old in a random sample of 300 adult smokers? What may this suggest about the population that was observed?