Jan 25
AP STAT

1. Notes ch 7

2. return Exam/review at end.....
cant keep...need back

Chapter 7 Notes

Discrete and Continuous Random Variables.

Discrete Random Variable is COUNTABLE (whole number)-distinct.(gaps)

Probability Distribution is a TABLE of probabilities as associated with the random variables being studied.

We use the Distribution to answer probability questions. We can also make ***Probability HISTOGRAMS***

The Probability Distribution (conditions)

1. has to contain probabilities that are between zero and one, inclusive.
2. Probabilities sum to one.

EX:

Make a Distribution for the number of boys in a three child household.

x	0	1	2	3
p(x)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

← # Boys

GGG $\frac{1}{2}\frac{1}{2}\frac{1}{2}$

BGG $-\frac{1}{8}$
GBG
GGB

BBG
GBB
BGB

What is Probability of having at least 2 boys?

$P(X \geq 2)$ $P(2)+P(3)$ $\frac{3}{8}+\frac{1}{8}=\frac{4}{8}=\frac{1}{2}$

What is the Probability of having less than 2 boys?

$P(X<2)$ $P(0)+P(1)=\frac{1}{8}+\frac{3}{8}=\frac{1}{2}$

What is the expected value (mean) of number of boys per family?

$\mu=\sum x(Px)$

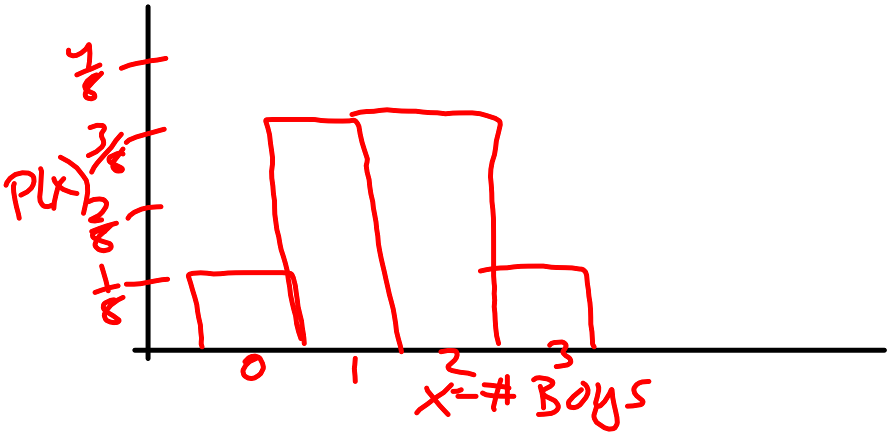
$\cancel{0(\tfrac{1}{8})}+1(\tfrac{3}{8})+2(\tfrac{3}{8})+3(\tfrac{1}{8})$

$\frac{3}{8}+\frac{6}{8}+\frac{3}{8}=\frac{12}{8}=1.5$ boys

Recall Expected value aka MEAN:

$$\boxed{\Sigma x P(x)}$$

Make a Probability Histogram for the 3 boys Distribution.



CONTINUOUS Random Variable

Now includes decimals- interval of values.

The distribution is now a *density curve* whose area *under* the curve is 1.
Note: for continuous you can ignore < vs ≤ (not in discrete though)

Normal Probability Distribution is the most common continuous curve.
(Emipirical Rule/ Z-scores) help you find the probability (area under curve) for a desired value.

Recall $\boxed{z = \frac{x-\mu}{\sigma}}$ Find the z score, find the area under the curve (chart)

While Normal is one type of continuous graph, others can be determined by the probability histogram.

7.2 More about mean/expected value.....and also introducing ....Variance/Standard Deviation

First:

Really quick **EXAMPLE**:

1. Verify if the following is a probability distribution:

Days of Rain (X)	0	1	2	3
Probabilities(P(x))	.216	.432	.288	.064

2. Draw a probability Histogram

3. What is the expected Value?

$0(.216) + 1(.432) + 2(.288) + 3(.064) = 1.2$

4. What is the variance and standard deviation?

Formula: $\sigma^2 = \Sigma(x-\mu)^2 P(x)$

$\sigma = \sqrt{(\text{variance})}$

$(0-1.2)^2(.216) + (1-1.2)^2(.432) + (2-1.2)^2(.288) + (3-1.2)^2(.064)$

$\sigma^2 = .72 \quad \sigma = .85$

EXAMPLE:

Apgar Scores

In 1952, Dr. Virginia Apgar suggested five criteria for measuring a baby's health at birth: skin color, heart rate, muscle tone, breathing, and response to stimulation. She developed a 0-1-2 scale to rate a newborn in each of the 5 criteria. A babies Apgar score is the SUM of the ratings in each of the 5 categories. This gives a whole number value between 0-10. These Apgar scores are still used today to evaluate the health of a newborn.

What Apgar scores are typical? To find out researchers recorded the Apgar scores of over 2 million newborn babies in a single year. Imagine selecting one of the newborns at random. (that is the chance process). Define the random variable X=Apgar score of a randomly selected baby one minute after birth. The table gives the probability distribution for X.

Value	0	1	2	3	4	5	6	7	8	9	10
Probability	0.001	0.006	0.007	0.008	0.012	0.020	0.038	0.099	0.319	0.437	0.053

A) Show the probability distribution for X is legitimate.

B) Doctors decided that an Apgar score of 7 or higher indicated a healthy baby. What is the probability that a randomly selected baby is healthy?

$P(X \geq 7) = \cancel{.5}\ .908$

Normal Probability (Continuous Random Variable)

Example:

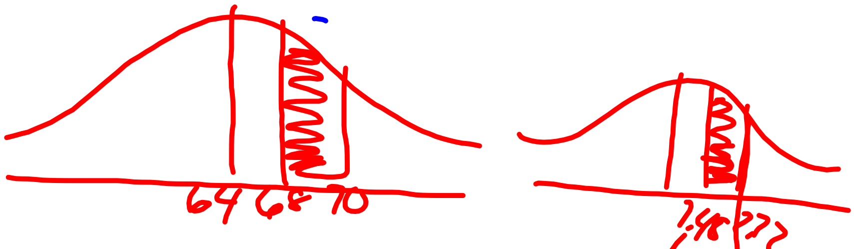
The heights of young woman closely follow the Normal Distribution with a mean μ= 64 inches and a standard deviation σ=2.7 inches. This is a distribution for a large set of data. Now choose one young woman at random. What is the probability that the randomly chosen young woman has height between 68 and 70 inches tall?

P(68≤Y≤70)

If using standardized Z scores, find the z score for each height.

$z=\frac{x-\mu}{SD}=\frac{68-64}{2.7}=1.48$ $\quad z=\frac{70-64}{2.7}=2.22$

Notation and graph:



Read table: .9868 − .9306 = .0562

A = .9306 A = .9868

normalcdf(1.48, 2.22, 0, 1)

On Calculator: normalcdf(68, 70, 64, 2.7)

Other Normal Examples:

1. The lengths of pregnancies of humans are normally distributed with a mean of 268 days and a standard deviation of 15 days. A baby is premature if it is born 3 weeks early. What percentage of babies are born prematurely?

2. Refer to example 1. What is the probability a pregnancy lasts more than 300 days?

3. IQ scores are normally distributed with a mean of 100 and standard deviation of 15. Find the x value (the IQ) that corresponds to a z score of -1.645.

4. Refer to 3. What is the IQ score that corresponds to the upper 10%?

5. Assume blood pressure readings are normally distributed with a mean of 116 and standard deviation of 4.8. If 36 people are randomly selected, what is the probability that their mean blood pressure will be less than 118?

6. The average number of pounds of red meat a person eats each ear is 196 with a standard deviation of 22. A sample of 50 people are selected. Find the probability that the mean of the sample will be greater than 200lbs.

Law of Large Numbers....quick revisit.....

The sample mean(*x bar*) is rarely the same as the population mean (*mu*).

The more we measure the more accurate we will be toward the pop mean.

This applies to probability as well. The more times you perform an "experiment" the closer you get to the ***theoretical probability*** (aka classical)

The more random the data, the more trials you need (there is no set number for "large numbers)

Rules for Means and Variances-Transforming and Combining Variables

- For random variables X and Y, you can add their **means** to get the Mean Sum. (this is also expected value *E*).
- You can add variances **ONLY of the events are INDEPENDENT**. Standard deviations **DO NOT ADD.** For independent events, you must *combine your variances FIRST*, then square root for standard deviation.

and

- For a random variable X, if **mean** needed to be adjusted by constants a and b, then Mean $X_{a+bx} = a + b\mu x$

**The transformation does NOT change the shape of the distribution

- For **standard deviation**, transformed $\boldsymbol{\sigma_y = |b|\sigma_x}$ (absolute value bc standard deviation cant be negative)

EXs.

Baby and Bathwater:

One brand of bathtub comes with a dial to set the water temperature. When "babysafe" setting is selected and the tub is filled, the temperature X of the water follows a Normal distribution with mean of 34^{o}C and standard deviation of 2^{o}C. Define the random variable Y to be water temp in degrees F. (F=9/5C +32)

A) Find the mean and standard deviation for Y.

B) According to Babys R Us, the temperature of babies bathwater should be between 90^{o}F and 100^{o}F. Find the probability that the water temp on a randomly selected day when the "babysafe" setting is used meets Babys R Us recommendations. Be sure to sketch graph also.

notation $P(90 \le Y \le 100)$

<u>Another example:</u>

Pete and Erin are trip planners. Let X be the number of passengers on Pete's trip and Y be the number of passengers on Erin's trip. C is the amount of money pete collects on a randomly selected day. We have obtained the following information:

$\mu_x = 3.75$ $\mu_y = 3.10$ $\mu_c = 562.50$

$\sigma_x = 1.090$ $\sigma_y = 0.943$ $\sigma_c = 163.50$

P ER P

A) Erin charges $175 per passenger for her trip. Let G= the amount of money that she collects on a randomly selected day. Find the mean and standard deciation of G.

B) Find the mean and standard deviation of the total amount that Pete and Erin collect on a randomly chosen day.

1/29/18

AP STAT

1. Check HWs
2. Book assignment (calendar-skip #33on 491 replace with #40 pg 499)/New practice packet.-work on for class tomorrow...will have class time too.

Upcoming: SHORT QUIZ (ch 7) WEDNESDAY

CHAPTER 8 Binomial and Geometric Distributions

Gretchen is a 60% free throw shooter. In a season, she shoots an average of 75 free throws. What is the probability that Gretchen (a) makes exactly 50 out of 75 free throws, (b) makes more than 50 free throws, (c) makes more than 40 free throws, (d) makes between 45 and 55 free throws. ***By the end of this chapter you will be able to answer these and more!***

Binomial Distributions

BINS

- Each observation falls into two categories: success or failure
- Each observation is INDEPENDENT.
- There is a fixed number (n) of observations.
- The probability for SUCCESS is the SAME for each observation.

If it falls into this category, the random variable is now a ***Binomial*** Random Variable and the probability distribution is now a ***binomial*** distribution.

FORMULA:

Binomial Coefficient: $_nC_x$ alternate notation: $\binom{n}{x}$

Probability : $_nC_x\, p^x\underbrace{(1-p)}_{q}{}^{n-x}$ $\binom{n}{x}p^x q^{n-x}$

n= number of trials

p= probability success

x= random variable

A **Binomial Experiment** is an experiment consisting of many many INDEPENDENT trials of the same chance process and it is recorded the Number of times a particular outcome occurs.

Examples:

1. A certain surgical procedure has an 85% chance for success. A doctor peerforms the procedure on 8 people. The random variable represents the number of successful surgeries.

A) Is this Binomial...Check BINS

$n = 8$
$p = .85$

B) What values could the random variable be?

$X = 0, 1, 2, 3, 4, 5, 6, 7, 8$

C) What is the probability that 5 will be successful?

$P(x=5) = {}_8C_5 (.85)^5 (.15)^3 = .0839$

2. A card is selected from a deck of cards and replaced. the experiment is repeated 5 times. Find the probability of selecting 3 clubs.

N= 5 p= .25 x = 3

Using Formula:
$P(x=3) = {}_5C_3 (.25)^3 (.75)^2 = .0879$

Using the Calc:
binompdf(n, p, x)
binompdf(5, .25, 3) =

3. A survey indicates that 41 percent of woman in the US consider reading as their favorite leisure activity. you randomly select 4 US woman and ask them if reading is theri favorite leisure activity.

$n = 4$
$p = .41$

A) Find the probability that 2 say yes.

$P(x=2) = \text{binompdf}(4, .41, 2)$
$= .3511$

B) Find the probability that at least 2 say yes.

$P(x \geq 2) = P(2) + P(3) + P(4)$
$.3511 + .1626 + .0282 = .5420$

C) Find the probability that fewer than 2 say yes.

$P(x < 2) = P(0) + P(1)$
$P(x \leq 1) = 1 - .5420 = .4580$
binomcdf(4, .41, 1)

pdf vs cdf

pdf: probability distribution function

Use for equal to value (=)

cdf: cumulative distribution function

Use for less than or equal to (≤)

If greater than situation, use complement

****Input examples:*

for P(x<17) means 16 or less so imput 16 in the binomcdf

for P(x≤ 17) use 17

for P(x> 17) means 18 or more so 1- binomcdf (with 17)- this knocks out 17 or less

for P(x ≥ 17) means 17 or more so 1- binomcdf(with 16) this knocks out 16 or less

Other Formulas:

Mean for binomial : $\mu = np$

Standard deviation: $\sigma=\sqrt{npq}$ *where* $q=1-p$

Worth noting: as n gets increasingly large in a binomial distribution, it begins to resemble a Normal distribution. (so can be used for excessively large n values. however with todays technology, its not entirely necessary as most technologies can handle binomials with large n.

8.2 Geometric Distributions

ASSESSMENT
CH 7-8
FRIDAY

Probability to FIRST success...What?

Do the experiment ***UNTIL first success***.

EX: The probability of rolling a 3 on one die is 16.6%.

Let our success be defined as rolling a 3 and lets roll a die until we get a three. This is a ***GEOMETRIC*** setting.

Question: What is the probability that we roll our first 3 on the second roll?

You can simulate it by perfoming this experiment many many times.....we won't.

Mnemonic: **BITS**

Binary- success or fail for each observation

Independent- observations are independent.

Trials- # trials for first success

Success- probability for success is the same for every observation.

$P(x=2) = (\quad)$

Before we roll the dice, let's figure out the proability.

Formula: $P(x=n) = (1-p)^{n-1}p$

$P(x=2) = (.834)^1(.166)$
$= .1384$

EX: In monopoly, one way to get out of jail is to roll doubles. How likely is it that someone would roll doubles on their first, second or third attempt. Let's use out calculator for this one. *(use geometpdf (n,p))*

$\frac{1}{6} \cdot \frac{1}{6}$ $\frac{1}{36}$ $\cdot \frac{6}{36}$ $\frac{1}{6}$

Find:

P(1)= geompdf ((1/6), 1) = .1667

P(2)= geompdf(1/6, 2) = .1389

P(3)= geompdf(1/6, 3) = .1157

Add

Add these probabilities together. P(1 or 2 or 3)= .4213

$\leq$

EX. From experience, a salesman knows that he will make a sale on any given telephone call is 0.23. Find the probability that on any given day, the first sale will occur on the forth or fifth call.

P(4)= geompdf(.23, 4) = .1050

P(5)= geompdf(.23, 5) = .0809

P(4th or 5th)= .1050 + .0809 = .1859

1, 2, 3

What is the probability that it will happen in less than 4 calls?

(*this is where you use the geometcdf. geometcdf is LESS THAN in calc.)*

P(x<4) = geomcdf(.23, 4) = .6484 .1050

What is the proability it occurs in ***MORE*** than 4 calls?

P(x > 4) = 1 − geomcdf(.23, 3)

<- 4 5 ->

1 2 3 4

~~.2466~~

.2707

More Formulas:

Mean for geometric distribution: $\mu = 1/p$

Variance: $(1-p)/p^2 \Rightarrow SD \sqrt{}$

EX: What are the mean and standard deviation for the previous problem?

$p = .23 \quad \mu = \frac{1}{.23} = 4.3 \quad \sqrt{\frac{(1-.23)}{.23^2}} = 3.81$

EX:Suppose your teacher was going to play a game which will determine how many HW problems you have to complete. You are to keep guessing birthday days of week for her list of birthdays she has. How ever many guesses it takes you to get a correct day will equal the number of HW problems assigned. If first guess is right, then that is one HW problem, if it takes 3 guess for first one right then 3 HW problems, etc....the birthdays are all independent. It's a list of DIFFERENT birthdays so if guessed wrong it goes to the next birthday on the list etc.

A)Is this binomial or geometric? G

B)What is the probability that it will take 10 guesses?

$P(X=10)$ geompdf$(1/7, 10) = .0357$

C)What is the probability that it will take less than 10 guesses.

$P(X<10) =$ geomcdf$(1/7, 9)$

D) What are the mean and standard deviation?

$\mu = \frac{1}{1/7} = 7$

$SD = \sqrt{\frac{(1-1/7)}{1/7^2}} = 6.4$

Feb 6

AP STAT

1. Check/rev HW

2. Practice packet

Upcoming:

Project due WED 2/7- this is tomorrow

Assessment friday 2/9- ch 7-8

Project Due WED 2/7

HW- see calendar