AP Statistics – 6.1	Name: KEY
Goal: Understanding Discrete and Continuous Random Variables	Date:

I. Warm Up CYU (page 349): *Discrete RV's*

- Always state what your random variable is: **X =** THE NUMBER OF CARS SOLD DURING THE 1ST HOUR OF BUSINESS ON A RANDOMLY SELECTED FRIDAY

- Calculate the mean of X by hand and interpret in context

$E(x) = \mu_x = 0(.3) + 1(.4) + 2(.2) + 3(.1) = 0 + .4 + .4 + .3 = \boxed{1.1}$

NOTE: $E(x) = \sum x_i p_i$

* THE LONG-RUN AVERAGE, OVER MANY FRIDAY MORNINGS, WILL BE ABOUT 1.1 CARS SOLD.

- Calculate the standard deviation of X by hand and interpret in context

$VAR(x) = \sigma_x^2 = \sum (x_i - \mu_x)^2 \cdot p_i = (0-1.1)^2(.3) + (1-1.1)^2(.4) + (2-1.1)^2(.2) + (3-1.1)^2(.1) = .363 + .004 + .162 + .361 = \boxed{.89}$

$SD(x) = \sigma_x = \sqrt{.89} = .943$

* ON AVERAGE, THE NUMBER OF CARS SOLD ON A RANDOMLY SELECTED FRIDAY WILL DIFFER FROM THE MEAN (1.1) BY ABOUT .94 CARS SOLD

II. Discrete Random Variables

Example: *"National Hockey League's Goals"*

In 2010, there were 1319 games played in the regular NHL season. Imagine selecting one of these games at random and then randomly selecting one of the two teams that played in the game. Define the random variable X = number of goals scored by a randomly selected team in a randomly selected game. **The table below gives the probability distribution of X:**

x_i	Goals:	0	1	2	3	4	5	6	7	8	9
p_i	Probability:	0.061	0.154	0.228	0.229	0.173	0.094	0.041	0.015	0.004	0.001

(a) Show that the probability distribution for X is legitimate.

① ALL THE PROBABILITIES ARE BETWEEN 0 AND 1

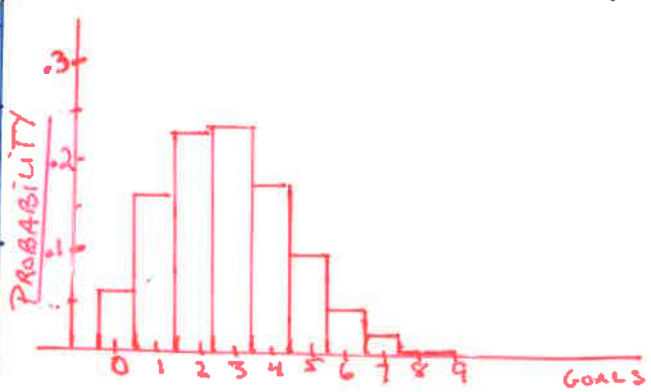
② THE SUM OF THE PROBABILITIES IS 1. $(\sum p_i = 1)$

(b) Make a histogram of the probability distribution. Describe what you see. Remember CUSS and BS

Calc
L1 = x_i's
L2 = p_i's

STAT PLOT
XLIST: L1
FREQ: L2

WINDOW
XMIN -1, XMAX 10, XSCL 1
YMIN -.1, YMAX .3, YSCL=.1



* The Histogram is skewed to the right meaning the number of goals are relatively low scoring.
* The center is between 2 and 3 goals
* With the majority of goals ranging from 0 to 7.

(c) What is the probability that the number of goals scored by a randomly selected team in a randomly selected game is at least 6?

write prob. stmt

$P(X \geq 6) = P(x=6) + P(x=7) + P(x=8) + P(x=9) = .041 + .015 + .004 + .001 = \boxed{.061}$

Random Variables denoted with capital letters.

(d) We defined the random variable X to be the number of goals scored by a randomly selected team in a randomly selected game. **Compute the mean of the random variable X and interpret this value in context.**

$\Rightarrow E(x) = \mu_x = \sum x_i p_i = 0(.61) + 1(.154) \ldots \; 9(.001) \Rightarrow \boxed{\mu_x = 2.851}$

* THE MEAN NUMBER OF GOALS FOR A RANDOMLY SELECTED TEAM IN A RANDOMLY SELECTED GAME IS 2.851. THAT IS ... IF YOU WERE TO REPEAT THE RANDOM SAMPLING PROCESS OVER AND OVER AGAIN, THE MEAN # OF GOALS SCORED WOULD BE ABOUT 2.851 IN THE LONG RUN.

(e) We defined the random variable X to be the number of goals scored by a randomly selected team in a randomly selected game. **Compute and interpret the standard deviation of the random variable X.**

$\Rightarrow VAR(x) = \sigma_x^2 = \sum (x_i - \mu_x)^2 \cdot p_i = (0-2.851)^2 \cdot (.061) + (1-2.851)^2 \cdot (.154) + \ldots$

$(9-2.851)^2(.001)$

$\sigma_x = \sqrt{2.66} = 1.63$

$\sigma_x^2 = 2.68$

* ON AVERAGE, A RANDOMLY SELECTED TEAM'S NUMBER OF GOALS IN A RANDOMLY SELECTED GAME WILL DIFFER FROM THE MEAN (2.851) BY ABOUT 1.63 GOALS.

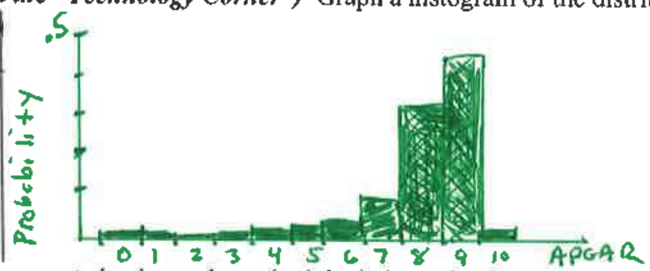
III. Technology Corner (page 348)

Example: ***"Apgar Scores"***

a) Define the Random Variable X = THE APGAR SCORE OF A RANDOMLY SELECTED BABY

b) *(step 2 from the "Technology Corner")* Graph a histogram of the distribution. Sketch it. Describe the shape.

CALC
L1 = x_i L2 = p_i
STAT PLOT
XLIST = L1
FREQ = L2
SET WINDOW!!
XMIN −1 YMIN −.1
XMAX 11 YMAX .3
XSCL 1 YSCL .1



THE HISTOGRAM OF APGAR SCORES IS HEAVILY SKEWED TO THE LEFT TELLING US THAT A RANDOMLY SELECTED NEWBORN TENDS TO HAVE A SCORE AROUND 9 WITH MOST SCORES FALLING BETWEEN 7 AND 9.

c) Calculate the expected value and standard deviation using lists and 1-VarStats – show your work

$E(x) = \mu_x$ $VAR(x) = \sigma_x^2$

L1 x_i	L2 p_i	L3 $x_i \cdot p_i$	L4 $(x_i - \mu_x)^2 \cdot p_i$
0	.001	0	.066
1	.006	.006	.305
2	.007	.014	.263
3	.008	.024	.210
4	.012	.048	.204
5	.020	.1	.196
6	.038	.228	.172
7	.099	.693	.125
8	.319	2.552	.005
9	.437	3.933	.332
10	.053	.53	.186
Totals	1.00	8.128	2.066

THIS IS A LEGIT PROB. DISTRIB.

1-VAR STAT x =

- Show your work to calculate the expected value and standard deviation

$E(x) = \mu_x = \sum x_i p_i = 8.128$

The Sum of $x_i p_i$'s in L3

$VAR(x) = \sigma_x^2 = \sum (x_i - \mu_x) \cdot p_i$

$= 2.066$

The sum of L4

$SD(x) = \sigma_x = \sqrt{2.066} = 1.437$

d) *(step 3 from the "Technology Corner")* Now check your above work following this step to find the expected value and standard deviation

1VAR STATS

LIST: $\boxed{L1}$ x_i's

FREQ LIST: $\boxed{L2}$ p_i's

$\bar{x} = 8.128 \leftarrow$ This is
$\sum x = 8.128 \leftarrow$ μ_x or $E(x)$
$S_x = \cancel{\;} \leftarrow$ Not a sample
$\sigma_x = 1.437 \leftarrow SD(x)$

FOR RV'S → NEVER USE $\bar{x}$!!!

III. Technology Corner (page 348) **Example: *"Apgar Scores"***

a) Define the Random Variable

X = The Apgar Score of a randomly selected baby

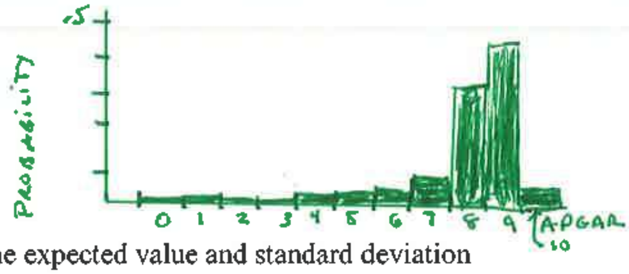
b) *(step 2 from the "Technology Corner")* Graph a histogram of the distribution. Sketch it. Describe the shape.

CALC

$L1 = x_i$ $L2 = p_i$

STATPLOT XLIST = L1
FREQ = L2

Xmin −1 Ymin −.1
Xmax 11 Ymax .5
Xscl 1 Yscl .1



THE HISTOGRAM OF APGAR SCORES IS HEAVILY SKEWED TO THE LEFT TELLING US THAT A RANDOMLY SELECTED NEW BORN TENDS TO HAVE A SCORE OF 9 WITH MOST SCORES FALLING BETWEEN 7 TO 9.

c) Calculate the expected value and standard deviation

- Create lists for each column given in the table. Make sure you understand each column
- Fill in the table. You can round columns 4 and 5 (1 decimal) and column 6 (3 decimals).
- Use 1-Var Stats to get each column total

X $E(x) = \mu_x$ $VAR(x) = \sigma_x^2$

L1 x_i	L2 p_i	L3 $x_i \cdot p_i$	L4 $x_i - \mu_x$	L5 $(x_i - \mu_x)^2$	L6 $(x_i - \mu_x)^2 \cdot p_i$
0	.001	0	−8.1	66.0	.066
1	.006	.006	−7.1	50.8	.305
2	.007	.014	−6.1	37.6	.263
3	.008	.024	−5.1	26.3	.210
4	.012	.048	−4.1	17.0	.204
5	.020	.1	−3.1	9.8	.196
6	.038	.228	−2.1	4.5	.172
7	.099	.693	−1.1	1.3	.125
8	.319	2.552	−0.1	.2	.005
9	.437	3.933	.87	.8	.332
10	.053	.53	1.87	3.5	.186
Totals	1.00	8.128	−34.4	217.63	2.066

THIS IS A LEGIT PROBABILTY DISRIBUTION

- Show your work to calculate the expected value and standard deviation

$E(x) = \mu_x = \Sigma x_i p_i = 8.128$ SEE COLUMN 2

$VAR(x) = \sigma_x^2 = 2.066 \longrightarrow SD(x) = \sigma_x = \sqrt{2.066} = 1.437$

d) *(step 3 from the "Technology Corner")* Now check your above work following this step to find the expected value and standard deviation

1 VAR STATS
LIST: $\boxed{L1}$
FREQLIST $\boxed{L2}$

$\bar{x} = 8.128$ ← This μ_x OR $E(x)$ NEVER $\bar{x}$!

$\Sigma x = 8.128$

$S_x = \cancel{\ }$ NOT A SAMPLE

$\sigma_x = 1.437$ ← SD(X)

IV Vocabulary you need to understand for this chapter. Here is space for you to take your own notes:

- **Random Variable** (p313) TAKES NUMERICAL VALUES THAT DESCRIBE THE OUTCOMES OF SOME CHANCE PROCESS.
  * RV'S ARE DENOTED WITH CAPITAL LETTERS (X)

- **Probability Distribution** OF A RANDOM VARIABLE GIVES ITS POSSIBLE VALUES AND THEIR PROBABILITIES. THIS IS A DISTRIBUTION LIKE IN EARLIER CHAPTER. Remember to graph and DESRIBE WITH "CUSS and BS".

- **Discrete Random Variable and Their Probability Distribution** (p313)

  DISCRETE RV's (X) takes a fixed set of possible Values.
  * GAPS MAY OCCUR
  * PROBABILITY DISTRIBUTION:

X	all possible outcomes
Value	x_1 to x_i
Probabilities	p_1 to p_i

- *The probabilities must satisfy 2 requirements*
  1) p_i's are betweel 0 and 1
  2) The som of probabilities is 1 $\Sigma p_i = 1$

  To find the probability of an EVENT, add the p_i's for the x_i's IN THE EVENT.

- **Expected Value of a Discrete Random Variable** (p345)

  The EXPECTED VALUE is the mean of a RV.

$$\boxed{\mu_X = E(x) = \Sigma x_i p_i} \, *$$

CONTEXT: IF YOU WERE TO REPEAT THE RANDOM SAMPLING PROCESS OVER AND OVER AGAIN, THE MEAN VALUE would be about ______ in the long run.

(random variable have population means. NEVER use $\bar{X}$. No!!

- **Standard Deviation of a Discrete Random Variable**(p347)

  To find the SD(x), you must first find the variance:

$$\boxed{\sigma_X^2 = VAR(x) = \Sigma (x_i - \mu_X)^2 \cdot p_i} \, **$$

CONTEXT OF SD

ON AVERAGE, A RANDOMLY SELECTED PROCESS WILL DIFFER FROM THE MEAN BY ABOUT ______.

* The formulas are on YOUR AP GREEN SHEET

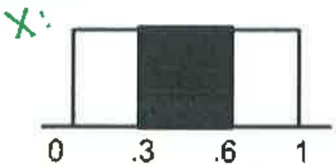
IV. Continuous Random Variables

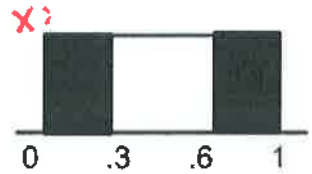
Continuous Probability Distribution: **described by the area under a density curve**

A continuous probability distribution differs from a discrete probability distribution in several ways.

- The probability that a continuous random variable will assume an exact value is **zero**.
- As a result, a continuous probability distribution cannot be expressed in tabular form.
- Instead, an equation or formula is used to describe a continuous probability distribution.
- We assign probabilities to **intervals** of outcomes rather than to individual outcomes.

EXAMPLE: Find the areas (shaded region) for the following density curves (*uniform distribution*).





$P(.3 \le X \le .6) = .3$

area = $l \cdot w$ =

$1 \cdot .3 = .3$

$1 - P(.3 \le X \le .6) =$

$1 - .3 = .70$

Your Notes: **Continuous Random Variables** (p350) SEE ABOVE NOTES.

IN MANY CASES ... DISCRETE RANDOM VARIABLES ARISE FROM COUNTING Something

CONTINUOUS RV'S ARISE FROM MEASURING SOMETHING

NOTE: $P(-1 \le X \le 2)$ is really the same as $P(-1 < X < 2)$

SINCE THERE IS NO AREA DIRECTLY ABOVE -1 OR 2.

***Example*: "Weights of Three-Year-Old Females"**

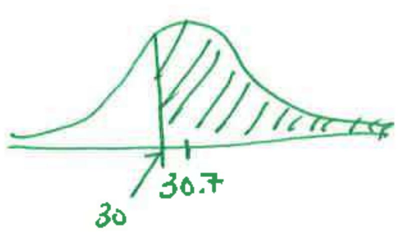
The weights of three-year-old females closely follow a Normal distribution with a mean of $\mu = 30.7$ pounds and a standard deviation of 3.6 pounds. Randomly choose one three-year-old female and call her weight X. Find the probability that the randomly selected three-year-old female weighs at least 30 pounds.

① DEFINE RV: X = THE WEIGHT OF A RANDOMLY CHOSEN 3 YEAR OLD FEMALE

② STATE DISTRIBUTION: $N(30.7, 3.6)$

③ STATE PROBABILITY OF INTEREST: $P(X \geq 30)$

④ DRAW A PICTURE AND STANDARDIZE THE WEIGHT TO FIND PROBABILITY



$Z = \frac{30 - 30.7}{3.6} = -.1944$

$P(Z \geq -.1944) = .5790$

normalcdf(-.1944, E99, 0, 1)

DO NOT NEED TO GIVE THIS.

⑤ CONCLUDE:

THERE IS ABOUT A 58% CHANCE THAT THE RANDOMLY SELECTED 3 YEAR OLD FEMALE WILL AT LEAST 30 pounds