6.1

Variables

$$\triangle \longrightarrow x \longrightarrow \theta \longrightarrow \mu$$

Random Variable (X)

A variable whose value is a probability

Discrete RVs

The probabilities that a customer selects 1, 2, 3, 4 or 5 items is:

X	1	2	3	4	5
P(X)	.32	.12	.23	.18	.15

= 1.00 ✓

X = the number of items bought

$$P(X > 3.5) = P(X = 4 \text{ or } 5) = .33$$

$$P(1.0 < X < 3.0) = P(X = 2) = .12$$

$$P(X < 5) = 1 - P(X = 5) = .85$$

Mean of a RV (Weighted Average)

$$\boxed{E(X) = \mu_X = \sum x_i p_i}$$

$$\mu_X = 1(.32) + 2(.12) + 3(.23) + 4(.18) + 5(.15)$$

$$= \boxed{2.72 \text{ items}}$$

Don't Round

Variance of a RV

$$\boxed{Var(X) = \sigma_X^2 = \sum (x_i - \mu_X)^2 p_i}$$

$$\sigma^2 = (1-2.72)^2(.32) + (2-2.72)^2(.12) + (3-2.72)^2(.23)$$
$$+ (4-2.72)^2(.18) + (5-2.72)^2(.15) = 2.1016$$

$$\sigma = \sqrt{2.1016} = 1.4496$$

Ex Game of Chance

If a player rolls 2 dice and gets a sum of 2 or 12, s/he wins \$20. If the person gets a sum of 7, s/he wins \$5. The cost to play is \$3. Find the expected payout.

$X = \text{Payout}$

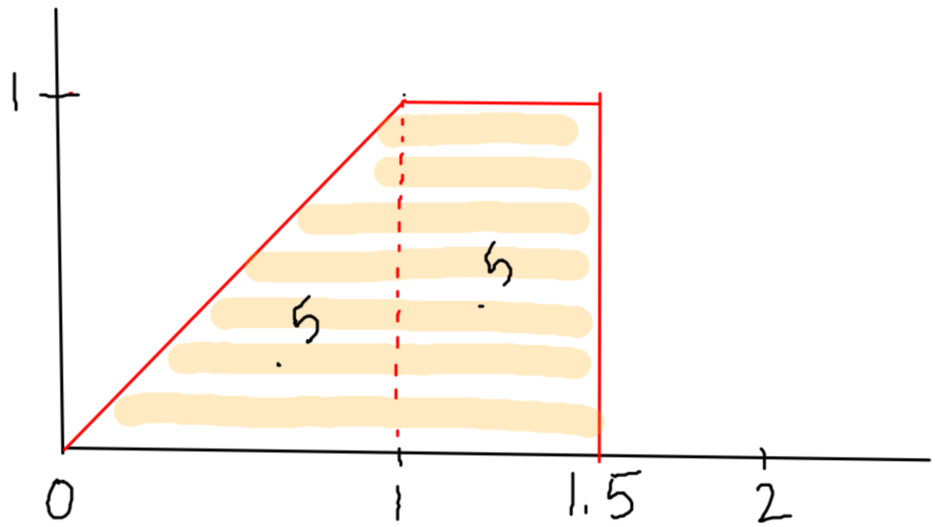
X	\$20	\$5	\$0
$P(x)$	$2/36$	$6/36$	$28/36$

$$\mu_X = 20\left(2/36\right) + 5\left(6/36\right) + 0\left(28/36\right) = \$1.94$$

Is the game fair?

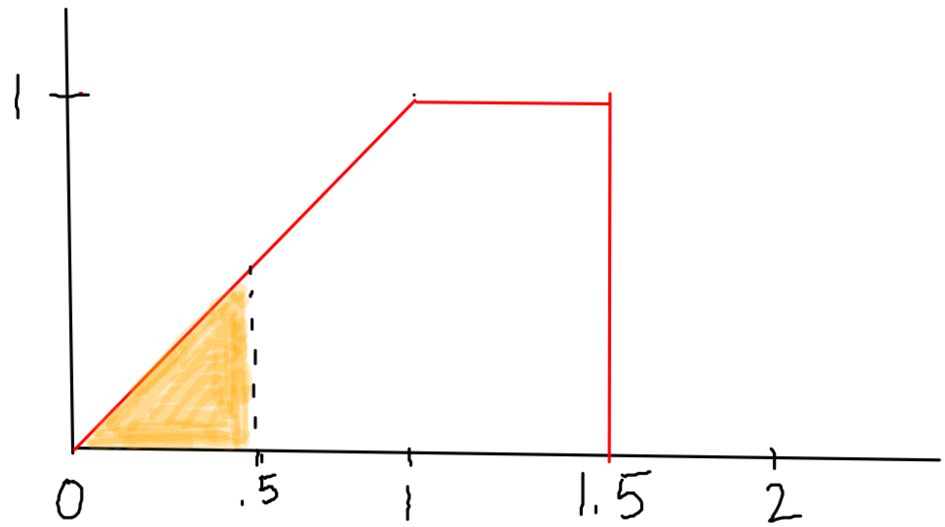
Continuous RVs (Heights, Time...)

A probability density function is comprised of 2 straight line segments. The first begins at $(0,0)$ and goes to $(1,1)$. The second goes from $(1,1)$ to $(1.5,1)$.



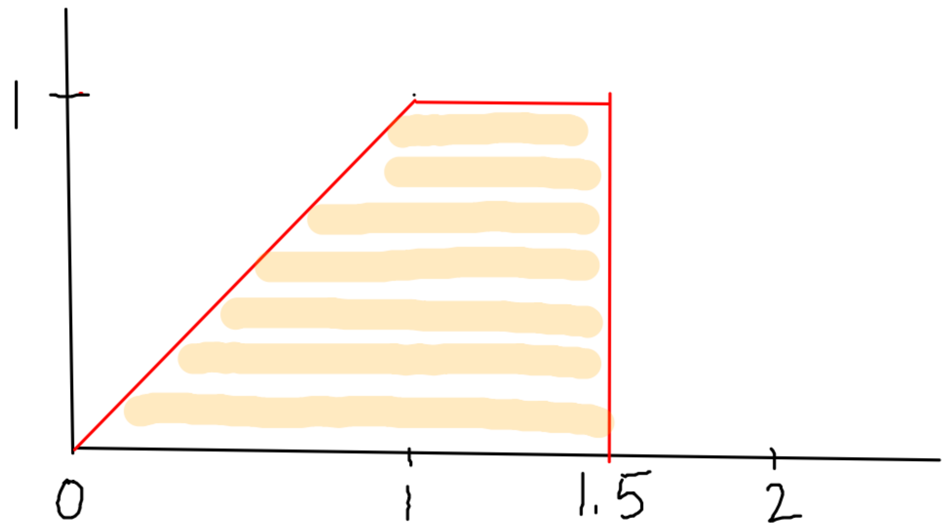
1) Verify this is a legitimate prob distribution

$A = \triangle + \square = \frac{1}{2}(1)(1) + (.5)(1) = 1$ ✓



2) Find $P(0 < X < .5) = \frac{1}{2}(.5)(.5) = .125$

3) Find $P(X > .5) = 1 - .125 = .875$



4) Find $P(X=1) = 0$

No Area Above A Point!

6.2

Data Set	$\bar{x}$		s	
{2, 5, 8}	5		3	
+4 {6, 9, 12}	9	+4	3	NC
x2 {4, 10, 16}	10	x2	6	x2

Rules For Random Variables

1) $Y = bX + a$

$$\mu_Y = b\mu_X + a$$

$$\sigma_Y = b\sigma_X$$

$$\sigma^2_Y = b^2\sigma^2_X$$

$\underline{Ex}$ Let $y = .9x - .2$

If $M_x = 3$, find M_y:

$$M_y = .9(M_x) - .2$$

$$M_y = .9(3) - .2$$

$$M_y = 2.5$$

Ex Assume $\sigma_x = 20$; Find a and b such that $Y = bX + a$ has a standard deviation of 1.

$$\sigma_y = 1$$

$$b(\sigma_x) = 1$$

$$20b = 1$$

$$b = \frac{1}{20}$$

$$a = \{\text{real}\}$$

2) $M_{X \pm Y} = M_X \pm M_Y$

X = SAT Math Score ($M_X = 625$)

Y = SAT Verbal Score ($M_Y = 590$)

$M_{X+Y} = 625 + 590 = 1215$

$M_{X-Y} = 625 - 590 = 35$

3) $\sigma^2_{x \pm y} = \sigma^2_x + \sigma^2_y$ <u>But</u> $\sigma_{x+y} \neq \sigma_x + \sigma_y$

<u>Ex</u> <u>If</u> $\sigma_x = 10$ and $\sigma_y = 8$, find σ_{x+y}

Find $\sigma^2_{x+y} = \sigma^2_x + \sigma^2_y$

$= 10^2 + 8^2$

$= 164$

$\sigma_{x+y} = \sqrt{164} = 12.8$

Ex 2002 AP Exam

Relay Team	Mean (Mile)	Stand Dev
1	4.9 min	.15 min
2	4.7	.16
3	4.5	.14
4	4.8	.15

1) What is the team's total mean time?

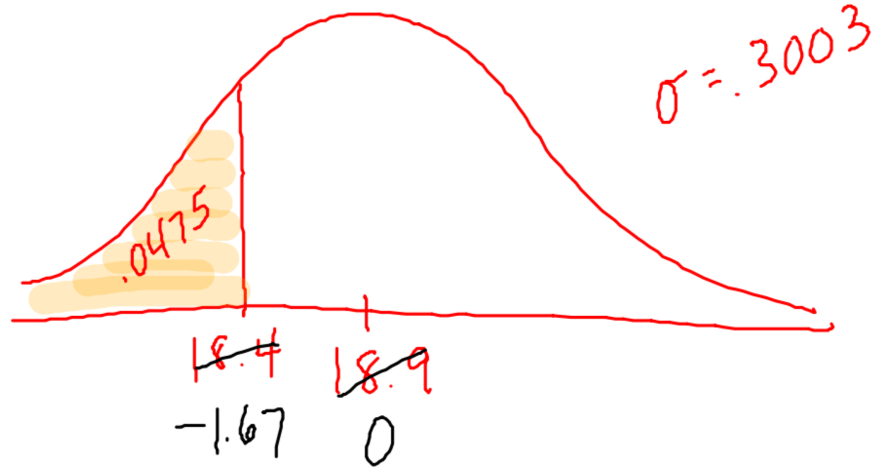
$$\mu_T = 4.9 + 4.7 + 4.5 + 4.8 = 18.9 \text{ min}$$

2) What is the stand dev of the team?

$$\sigma^2_T = (.15)^2 + (.16)^2 + (.14)^2 + (.15)^2 = .0902$$

$$\sigma_T = \sqrt{.0902} \approx .3003 \text{ min}$$

3) Suppose the team's best time to date is 18.4 min. If the distributim is aprox normal, what is the probability that the team will beat its own best time in the next race?



normalcdf (0, 18.4, 18.9, .3003) ≈ .0479

Sec 6.3

Randomly guess answers → 40 SAT questions

1) 2 Outcomes (Success / Failure)

2) P(success) = .2 (Constant)

3) Guesses are independent

Binomial Setting

4) X = # correct guesses out of 40

X = # successes from a fixed number of observations (n)

Notation

$$B(n, p) = B(40, .2)$$

# of observations

probability of success

1) What is the probability that $X = 10$

a) Use binomial formula

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(X=10) = \binom{40}{10} (.2)^{10} (.8)^{30}$$

$$P(X=10) = .1075$$

b) Calculator Function

$\boxed{\text{DISTR}} \rightarrow \text{binompdf}(\underset{n}{40}, \underset{p}{.2}, \underset{k}{10}) = .1075$

2) Find $P(X=0)$

a) Formula: $P(X=0) = \binom{40}{0}(.2)^0(.8)^{40} = .0001$

b) Calculator: $\text{binompdf}(40, .2, 0) = .0001$

3) Find $P(X \leq 10)$

a) Formula Inefficient

X	0	1	2	...	10
P(X)					

★ b) Use Calculator

binomcdf (40, .2, 10) = .8392

$\uparrow$

$\leq K$

4) Find $P(X > 10)$

a) Formula Inefficient

b) Use Calculator / Complement Rule

$P(X > 10) = 1 - \text{binomcdf}(40, .2, 10)$

$= 1 - .8392$

$= .1608$

If $B(n,p)$ then

$$\mu_x = np = (40)(.2) = 8$$

$$\sigma^2_x = np(1-p) = (40)(.2)(.8) = 6.4$$

$$\sigma_x = \sqrt{n(p)(1-p)} = \sqrt{6.4} = 2.5298$$

If $B(n, p)$ and $np \geq 10$, $n(1-p) \geq 10$

then normal distribution rules can be used...

Ex Attitudes toward shopping (P. 396)

$B(2500, .6)$; Find $P(X \geq 1520)$

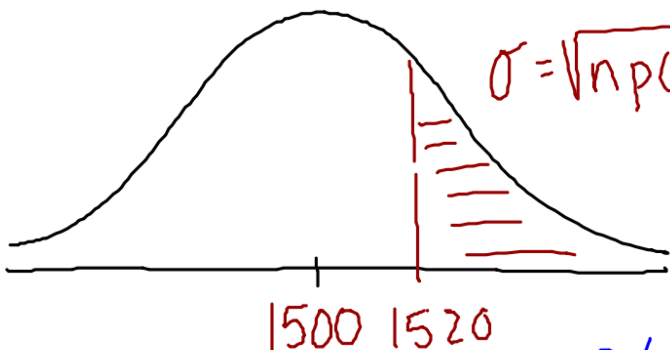
1) $1 - \text{binomcdf}(2500, .6, 1519) \approx .2131$

2) Since $np = (2500)(.6) = 1500 > 10$

and $n(1-p) = (2500)(.4) = 1000 > 10$

then normal distribution rules apply:

$\sigma = \sqrt{np(1-p)} = \sqrt{(2500)(.6)(.4)} = 24.49$



$np = (2500)(.6)$

$P\left(z > \frac{1520-1500}{24.49} > .82\right) \approx .2061$

normalcdf $(1520, 1000, 1500, 24.49) \approx .2071$

Geometric Setting

- 2 Outcomes (Success / Failure)
- P(success) is constant
- observations independent

X = # trials required before you get your first success

Geometric Formulas

$$P\begin{pmatrix}\text{1st Success On}\\ \text{nth Trial}\end{pmatrix} = P(X=n) = p\,q^{n-1}$$

$q = (1-p)$

As the # trials increase

the probability for the first success decreases...

EXAMPLE 8.17 ROLL A DIE

The rule for calculating geometric probabilities can be used to construct a probability distribution table for X = number of rolls of a die until a 3 occurs:

X	1	2	3	4	5	6	7	...
$P(X)$	0.1667	0.1389	0.1157	0.0965	0.0804	0.0670	0.0558	...

1/6

Here's one way to find these probabilities with your calculator:

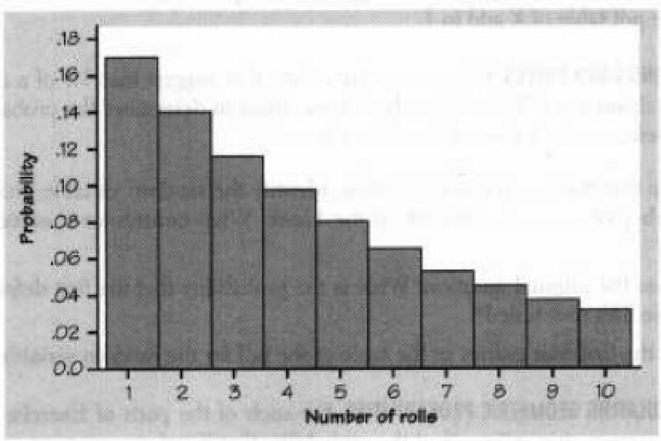
1. Enter the probability of success, 1/6. Press ENTER.
2. Enter *(5/6) and press ENTER.
3. Continue to press ENTER repeatedly.

```
1/6
          .1666666667
Ans*(5/6)
          .1388888889
          .1157407407
          .0964506173
          .0803755144
```

Verify that the entries in the second row are as shown:

X	1	2	3	4
$P(X)$	1/6	5/36	25/216	125/1296

Figure 8.4 is a graph of the distribution of X. As you might expect, the probability distribution histogram is strongly skewed to the right with a peak at the leftmost value, 1. It is easy to see why this must be so, since the height of each bar after the first is the height of the previous bar times the probability of failure $1 - p$. Since you're multiplying the height of each bar by a number less than 1, each new bar will be shorter than the previous bar, and hence the histogram will be right-skewed. Always.



skewed right always

FIGURE 8.4 Probability histogram for the geometric distribution.

P (More than n trials to get 1st success) $= P(X > n) = q^n$

$$\mu = \frac{1}{p}$$

Ex Roll a die...

1) Find P (Get a 3 on 4th Roll)

$$P(X=4) = \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{4-1} = .0965$$

$$P(X=4) = \text{geometpdf}(1/6, 4) = .0965$$

2) Find P (Get a 3 on one of 1st 4 rolls)

$P(X \le 4) = P(X=1) + P(X=2) + P(X=3) + P(X=4)$

$P(X \le 4) = \text{geometcdf}\left(\frac{1}{6}, 4\right) = .5177$

3) Find P(Takes more than 4 rolls to get a 3)

$P(X>4) = 1 - \text{geometcdf}\left(\frac{1}{6}, 4\right) = 1 - .5177 = .4823$

$P(X>4) = \left(\frac{5}{6}\right)^4 = .4823$

Success / Failure
P(success) = constant
Independence

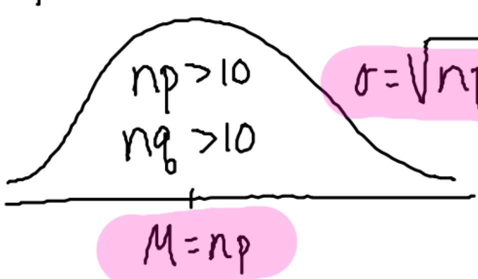
Binomial

Count # of successes from a fixed # trials:

$$P(X=k) = \binom{n}{k} p^k q^{n-k}$$

$P(X=k) = \text{binompdf}(n,p,k)$

$P(X \leq k) = \text{binomcdf}(n,p,k)$



$\sigma = \sqrt{npq}$

Geometric

Count # of trials to obtain 1st success:

$P(\text{success on } n\text{th trial}) = pq^{n-1}$

$P(X=n) = \text{geometpdf}(p,n)$

$P(X>n) = q^n$

$P(X \leq n) = \text{geomcdf}(p,n)$

$\mu = \frac{1}{p}$

Formulas Provided