Test 2 will be on October 20th, and cover the material on lessons 12,15-19, 21-24. We have covered all of this material in the lecture. What follows are highlights of the material and is not intended to be comprehensive.
In particular, you should know about all of the following

boxplots The boxplot is a way of displaying 5 numbers the min, the max, the median and the 25th and 75th percentiles, Q_1 and Q_3.

percentile ranks and z scores The percentile rank of a data point tells you how much of the data is less than that value as a percentage. It is found with

$$\text{percentile rank of} x = \frac{\#\{\text{data points } < x\} + 1/2\#\{\text{data points } = x\}}{\text{number of data points}}.$$

The z score of a data point measures how many sample standard deviations it is from the sample mean. It allows us to gauge spread by the number of standard deviations, as opposed to the percentile rank which gauges spread by percentages. The z scores of x_i is found by

$$z - \text{score} = \frac{x_i - \bar{x}}{s}$$

Probability definition The language of probability includes the words *outcome*, *event*, *outcome space*, *equally likely outcomes*. The definition of a probability for equally likely outcomes is

$$P(E) = \frac{\#\{\text{outcomes in } E\}}{\#\{\text{outcomes in total}\}}.$$

Counting We spent a lot of time working on counting. In particular we learned a few useful formulas. None more so than the formula for counting

$$\#\text{ of choices } = n_1 \cdot n_2 \cdots n_j,$$

where we have a sequence of stages where the ith stage has n_i choices. This counting formula is visualized in a tree diagram.

From this formula, we easily get one for permutations of n distinct objects:

$${}_nP_r = n \cdot (n-1) \cdots (n-r+1) = \frac{n!}{(n-r)!}.$$

Another formula applies when the objects are not distinct, and we permute all of them.

The number of combinations differs from the number of permutations as we do not count order. This number is

$${}_nC_r = \binom{n}{r} = \frac{n!}{(n-r)!r!}.$$

That is there are $r!$ permutations counted for each combination of size r

Odds A brief discussion on the relationship of odds to probabilities was given. To tie it to intuition, the mathematical expectation of several games was found.

Random Variables A random variable is defined as a numeric summary of an outcome, such as the number of heads in 3 coin tosses. For random variables, as with data sets, we can describe them in various levels of accuracty:

- The range is the set of all possible values.
- The distribution of X is a specification of the probability that X takes a value for each value in the range. This can be specified with a table or a function. For example. If X is 1 for heads and 0 for tails, then the distribution of X is $P(X = 1) = 1/2$, $P(X = 0) = 1/2$.

  If Y is the number of heads in 2 coin tosses, then a table may be used

k	$P(X = k)$
0	1/4
1	1/2
2	1/4.

- The mean of X, μ or $E(X)$, describes the center of a distribution as does the sample mean of a data set. It is found by the formula

  $$\mu = \sum kP(X = k).$$

- The variance of X measures spread as the sample variance does for a sample. It is found by

  $$\sigma^2 = \sum (k - \mu)^2 P(X = k) \text{ or } \sum k^2 P(X = k) - \left[\sum kP(X = k)\right]^2.$$

The Binomial distribution If we have several trials where a trial is either a success with probability p or a failure with probability $q = 1 - p$ then the number of successes in n trials is a random variable. It is a Binomial random variable and has distribution

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

and mean $\mu = np$ and variance $\sigma^2 = np(1 - p)$.

The mean is very intuitive. For example, it tells you how many heads to expect in 200 coin tosses ($np = 200(1/2) = 100$), or the number of doubles in 60 rolls of a pair of dice ($np = 60(1/6) = 10$).

What follows are typical problems. Be very surprised if they should show up on the exam

1. Draw a boxplot by hand of data with summary

```
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   6     315     684     933    1320    4940
```

Is this data set skewed?

2. For the data set

   `55  4 34 26 31 12 21 44 30  6 13`

   with mean 25.09 and standard deviation 15.92 find the percentile rank of 21 and its z score.

3. List all possible outcomes of tossing a coin 4 times. Circle the times where the number of heads is 3 and find the probability of that assuming equally likely events.

4. If there are 30 M&M's in a bag: 10 red, 5 blue, 5 brown and 10 green. What is the probability of picking one at random and having it be green or blue?

5. How many license plates of the form 3 letters then 3 numbers if we assume that there are no Q's for letters and no 0's for numbers.

6. A new car can have stick or automatic; be in red, blue or yellow, and have a sun roof or not. What are all the possible different cars (write them all down using some abbreviations).

7. Simplify as much as possible leaving a fraction written in terms of factors. For example $4!/2! = 4 \cdot 3$. (n is an unknown.)

   $${}_5P_3 \quad \binom{8}{4} \quad \binom{50}{4} \quad \binom{n}{2}.$$

8. A class of 8 students arrives 1 by 1. How many different orders are there for them to arrive?

9. A game is played by making 3 letter words out of 7 letter blocks. All 7 letters are distinct. How many possible words can be made?

10. A team of 12 players must select a captain and co-captain. How many possibilities are there?

11. A team of 12 players must select 5 players to start a game. How many possibilities are there.

12. Write our all possible color combinations of 3 colors when you can choose from the 5 colors Red, Blue, Green, Yellow, Orange.

13. A coin is tossed 3 times. Let Y be the number of tails. Find the range and distribution of Y.

14. After years of being a waitress it is decided that the chance a person leaves a tip of a certain percent has this probability

percent	5	10	15	20
	.25	.25	.40	.10

Find the expected tip percentage.

15. A pair of dice is rolled 5 times.

- What is the probability the sum of the first roll is a 4 or less? (see p 337)
- Let X be the number of times in the 5 rolls the sum is 4 or less. Fill in this table to find the distribution of X

k	$P(X = k)$
0	0.3349
1	
2	0.2009
3	
4	
5	

- What is the expected value of X? What is the standard deviation?