Probability Distributions and Statistics

- Distributions of Random Variables (II.C.1, 2, 3)
- Expected Value (II.C.2, 4)
- Variance and Standard Deviation (II.C.3, 5)
- The Binomial Distribution (II.C.6)
- The Normal Distribution (II.C.7)
- Applications of the Normal Distribution (II.C.7)

Distributions of Random Variables

- A *random variable* is a rule that describes the outcome of an experiment.
  - For our purposes, a random variable is just a number.
- In both of the next two cases, the random variable X is a *finite discrete random variable* because it assumes only finitely many values.

Example: Let the random variable X denote the number of boys in a 4-child family.

a) Describe the uniform sample space.

b) Describe all possible values of X.

c) What event E corresponds to $X = 1$?

Example: Three cards are drawn from a deck of cards. Let the random variable X denote the number of hearts. Describe all possible values of X.

- An *infinite discrete random variable* assumes infinitely many values that may be arranged in a sequence.

Example: A single die is rolled. Let the random variable X denote the number of rolls required in order to observe a 6. Describe all possible values of X.

- A *continuous random variable* assumes values that comprise an interval of real numbers.

- Examples include:
  - The weight, in ounces, of a newborn baby
  - The time, measured in minutes, that a light bulb will work

Example: Choose whether the random variable is finite discrete, infinite discrete, or continuous. Describe all possible values of the random variable.

a) The time, measured in minutes, students spend studying for a final examination over one weekend

b) The number of dominoes obtained with 2 dots on one side, when drawing six dominoes

c) The number of draws needed, without replacement, until the queen of spades is drawn

d) The number of draws needed, with replacement, until the queen of spades is drawn

e) The volume of water in a swimming pool, measured in cubic feet

Example: Let X be the number of boys in a 4-child family. Complete the probability distribution table.

X					
Probability					

Example continued:

a) Graph the data from the probability distribution table in a *histogram*.

b) Shade the part of the histogram associated with $P(1 < X \leq 3)$.

c) $P(X > 2) =$

Histograms and Measures of Central Tendency/Dispersion

- The *mean*, or *average*, is where the histogram "balances."
- The *mode* is the tallest rectangle.
  - This is the most frequently occurring value.
- The *median* is where the area is cut in half.
  - This is the middle value.
- The *range* is the number of rectangles.
  - Some may have a height of 0.

Example: What are the mean, median, mode and range of 1, 1, 2, 2, 3?

Mean =

Median =

Mode =

Range =

Example: What are the mean, median, mode and range of 1, 1, 2, 100?

Mean =

Median =

Mode =

Range =

Example: Find the mean, median and mode of the following test scores: 77, 98, 87, 84, 66, 94, 77, 72, 85, 64

Example: A class of 100 students took a 5-question quiz with the following results:

Number of Questions Correct	0	1	2	3	4	5
Number of Students	2	15	16	22	25	20

a) What is the *mean* number of correct answers?

b) What was the score that happened the most often?

c) What was the score in the middle?

d) What is the *range* of values?

Expected Value

- The *expected value* for the finite discrete random variable X in a probability distribution is:

$$E(X) = x_1p_1 + x_2p_2 + \ldots + x_np_n$$

- Expected value is used for life insurance premiums, lotteries, raffles, expected winnings, expected profits, etc.

- Organize the values and probabilities in a table to find the expected value more easily.

Example: What is the expected number of boys from the earlier example?

X					
Probability					

Example: What is the expected number of questions answered correctly from the earlier example?

Number of Questions Correct	0	1	2	3	4	5
Number of Students	2	15	16	22	25	20

Example: A set of 3 animals is to be sold from a group of 5 cows and 8 horses. Find the expected number of horses in the set.

Example: A car worth $10,000 is being raffled. 3000 tickets are sold for $10 each. What are the expected net winnings for a person buying 1 ticket?

Example: Suppose the lottery jackpot is $17 million and is won by matching 6 of the numbers 1 through 50 in any order. Find the expected winnings.

Numbers Matched	**Payoff (in $)**	**Probability**
0	0	
1	0	
2	0	
3	4	
4	120	
5	1900	
6	17,000,000	

Example: A musician wants to insure her violin for $100,000 in the event of theft or damage. Using insurance tables, the likelihood that her violin will remain safe is 94%. What is the minimum amount that she can expect to pay for her premium? (The minimum premium occurs when the insurance company's expected profit is zero.)

Odds

- If $P(E)$ is the probability of event E occurring, then
  - the odds *in favor* of E are calculated by:

$$\frac{P(E)}{1-P(E)} = \frac{P(E)}{P(E^C)}, \quad P(E) \neq 1$$

  - the odds *against* E occurring are calculated by:

$$\frac{1-P(E)}{P(E)} = \frac{P(E^C)}{P(E)}, \quad P(E) \neq 0$$

- Express the odds as a ratio of whole numbers if possible.
  - a to b or $a{:}b$

- If the odds in favor of an event E occurring are a to b, then

$$P(E) = \frac{a}{a+b}$$

Example: The odds of rain tomorrow are 2 to 7. What is the probability of rain tomorrow?

Example: If the chance that the Aggies will win a football game is 80%, what are the odds in favor of the Aggies?

Variance

- *Variance* is a measure of distance from the mean.
- Suppose a random variable has expected value $E(X) = \mu$ and probability distribution as follows:

$\boldsymbol{X}$	x_1	x_2	x_3	$\ldots$	x_n
$\boldsymbol{P(X = x)}$	p_1	p_2	p_3	$\ldots$	p_n

Then the variance of the random variable X is

$$\mathrm{Var}(X) = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \ldots + p_n(x_n - \mu)^2$$

- The units are the square of the units of the random variable.

Standard Deviation

- The *standard deviation* describes how far away from the mean one can expect the outcomes to be.
- It has the same units as the random variable.
- Given by: $\sigma = \sqrt{\mathrm{Var}(X)}$

Example: Find the mean, median, and standard deviation of each of the following data sets.

Set A	**Set B**	**Set C**
20	22	40
20	16	5
20	24	12
20	17	46
20	21	–3