Concise Lecture Notes Series

5

TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI

Fun with

Introductory Probability

Dr. Ralph deLaubenfels

TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI

Teacher-Scholar Institute LLC

Columbus, Ohio

2014

I. PRESENTATION OF DATA

1.1. The **population,** or **universe** is the set of things being studied.

1.2. A **sample** is a subset of the population. This is usually what we pick to study, like the subjects of an IRS audit.

The big question of statistics is: what can we deduce about the population, just by looking at a sample?

1.3. The **variable** is the characteristic of the population being studied.

1.4. A variable is **quantitative** if it can be measured with a number; otherwise, it's **qualitative.**

1.5. A quantitative variable is **discrete** if all possible values are separated.

1.6. A quantitative variable is **continuous** if it can assume any value in an interval or intervals.

1.7. Examples. Height and weight are continuous. The population of the United States is discrete.

1.8. A **frequency distribution** is a listing of all possible values, or ranges of values, along with their frequencies.

1.9. Frequency distribution paraphernalia: The ranges of values $[a, b]$ appearing in the leftmost column are called *classes* or *class intervals.* The *lower limit* is a, the *upper limit* is b. The *class mark* is $\frac{1}{2}(a+b)$. The *class size* is the difference between consecutive upper limits. A *class boundary* is the average of the upper limit of a given class and the lower limit of the next class. The *relative frequency* is the frequency divided by the total number of data. The *percent frequency* is the relative frequency multiplied by 100, with a percent symbol tacked on.

1.10. A **cumulative frequency distribution** lists the number of values below each class boundary.

1.11. A **histogram** is a set of rectangles representing a frequency distribution; the bases go between consecutive class boundaries, with the height equal to the frequency of the class interval.

1.12. An **ogive,** or cumulative frequency polygon, represents a cumulative frequency distribution. Graph ordered pairs (x, y) where x is a class boundary and y is the corresponding cumulative frequency, then draw straight lines connecting those ordered pairs.

1.13. Examples.

1. Put the following numbers into a frequency distribution, with classes 1–5, 6–10, 11–15 and 16–20. Then construct a cumulative frequency distribution, histogram and ogive, for the same data.

$$6, 7.5, 10, 19, 1, 12, 14, 19, 20, 10.$$

Solution.

class	frequency
$1-5$	1
$6-10$	4
$11-15$	2
$16-20$	3

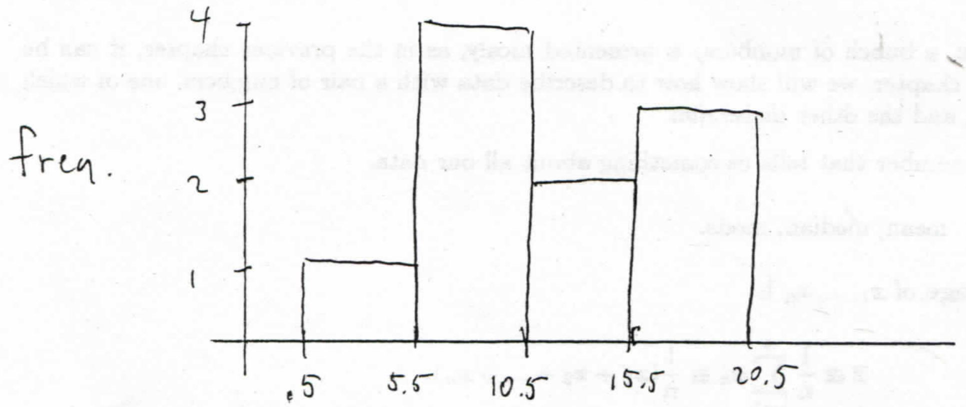
is the frequency table.

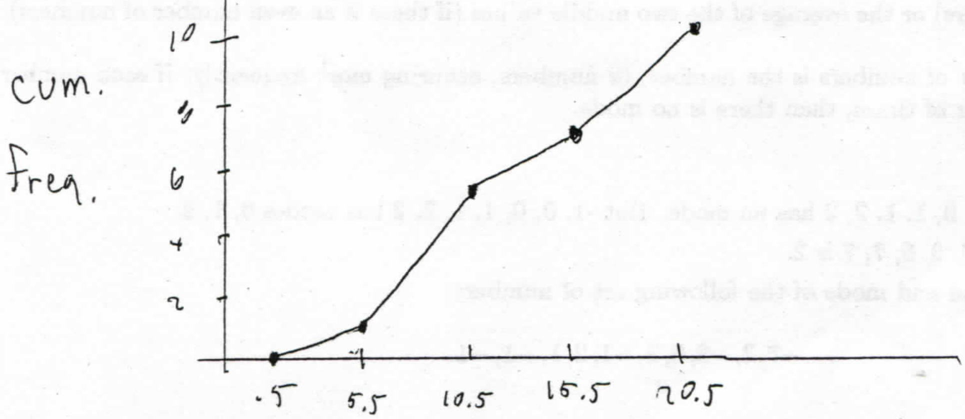
The class boundaries are $.5, 5.5 (= \frac{1}{2}(5+6)), 10.5, 15.5$ and 20.5. Here is the cumulative frequency distribution.

class	cumulative frequency
$< .5$	0
< 5.5	1
< 10.5	5
< 15.5	7
< 20.5	10

Note that you must make a completely separate table for the cumulative frequency; don't tack it onto the frequency table.

Here is the histogram (only for the frequency table) and the ogive (only for the cumulative frequency table).





2. In the previous example, what is the class size? Expand the frequency table to include upper and lower limits, class marks, relative frequency and percent frequency.

Solution. The class size is 5 (6 - 1). Here is a fancier frequency distribution.

class	frequency	upper limit	lower limit	class mark	relative frequency	percent frequency
1 – 5	1	5	1	3	.1	10
6 – 10	4	10	6	8	.4	40
11 – 15	2	15	11	13	.2	20
16 – 20	3	20	16	18	.3	30
TOTAL	10				1	100

II. CENTRAL TENDENCY AND DISPERSION

Even when data (that is, a bunch of numbers) is presented nicely, as in the previous chapter, it can be hard to work with. In this chapter, we will show how to describe data with a pair of numbers, one of which measures *central tendency*, and the other *dispersion*.

First, we'd like a single number that tells us something about all our data.

2.1. Central tendency: mean, median, mode.

2.2. The **mean,** or **average** of $x_1, ..., x_n$ is

$$\overline{x} \equiv \frac{1}{n}\sum_{k=1}^{n} x_k \equiv \frac{1}{n}(x_1 + x_2 + ... + x_n).$$

2.3. The **median** of a set of numbers is, after they've been arranged in order, the middle value (if there is an odd number of numbers) or the average of the two middle values (if there is an even number of numbers).

2.4. The **mode** of a set of numbers is the number, or numbers, occuring most frequently. If each number appears the same number of times, then there is no mode.

2.5. Examples.

1. The set of numbers 0, 0, 1, 1, 2, 2 has no mode. But -1, 0, 0, 1, 1, 2, 2 has modes 0, 1, 2.
2. The median of -1, 0, 0, 2, 5, 7, 7 is 2.
3. Find the mean, median and mode of the following set of numbers:

$$-7, 2, -3, 0, 2, -1, 2, 1, -1, -1.$$

Solution. The mean is

$$\overline{x} = \frac{1}{10}(-7 + 2 - 3 + 0 + 2 - 1 + 2 + 1 - 1 - 1) = -.6.$$

For the median and mode, put the numbers in order:

$$-7, -3, -1, -1, -1, 0, 1, 2, 2, 2.$$

The two middle numbers are -1 and 0, so the median is $\frac{1}{2}(-1 + 0) = -.5$.
The modes are -1 and 2.

4. Suppose a car is tested on a quarter-mile track. On its first five runs, its average time is 12 seconds. How quickly should it cover the quarter-mile on its sixth run, if it wants its average time over the first six runs to be 11.5 seconds?

Solution. For $k = 1, 2, ..., 6$, let t_k be the time (in seconds) of the k^{th} run. We want t_6. We know that

$$\frac{1}{5}(t_1 + t_2 + t_3 + t_4 + t_5) = 12, \quad \text{and} \quad \frac{1}{6}(t_1 + t_2 + t_3 + t_4 + t_5 + t_6) = 11.5;$$

by the first equation, $(t_1 + ... + t_5) = 5 \cdot 12$, so we can put this into the second equation, to get

$$\frac{1}{6}((5 \cdot 12) + t_6) = 11.5;$$

solving for t_6 gives us $t_6 = 6(11.5) - 5 \cdot 12 = 9$ (seconds).

Note that this is the time taken for the first five runs ($5 \cdot 12 = 60$ secondds) subtracted from the total time you want to take for all six runs ($6(11.5) = 69$ seconds).

Second, we'd like to measure how "spread-out" our data is.

2.6. Dispersion: range, mean deviation, standard deviation, variance.

2.7. The **range** of some data is the largest value minus the smallest value.

2.8. The **mean deviation** of $x_1, ..., x_n$ is

$$\frac{1}{n}\sum_{k=1}^{n}|x_k - \overline{x}| \equiv \frac{1}{n}(|x_1 - \overline{x}| + |x_2 - \overline{x}| + ... + |x_n - \overline{x}|).$$

2.9. The **variance** of $x_1, ..., x_n$ is

$$s^2 \equiv \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2 \equiv \frac{1}{n}\left[(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + ... + (x_n - \overline{x})^2\right].$$

2.10. The **standard deviation** is

$$s \equiv \sqrt{s^2}.$$

2.11. Examples. Find the range, mean deviation, standard deviation, and variance, of the following set of numbers: 5, -1, 0, 0, 5, 3.

Solution. The range is $5 - (-1) = 6$.

For deviations, we need the mean first, even though it's not asked for.

$$\overline{x} = \frac{1}{6}(5 - 1 + 0 + 0 + 5 + 3) = 2.$$

The mean deviation is

$$\frac{1}{6}(|5-2| + |-1-2| + |0-2| + |0-2| + |5-2| + |3-2|) = \frac{1}{6}(3+3+2+2+3+1) = \frac{7}{3}.$$

The variance is

$$\begin{aligned} s^2 &= \frac{1}{6}((5-2)^2 + (-1-2)^2 + (0-2)^2 + (0-2)^2 + (5-2)^2 + (3-2)^2) \\ &= \frac{1}{6}(3^2 + 3^2 + 2^2 + 2^2 + 3^2 + 1^2) = 6. \end{aligned}$$

The standard deviation is

$$s = \sqrt{6}.$$

2.12. The **computing formula**, for the variance, is

$$s^2 = \frac{1}{n}\left[\sum_{k=1}^{n}x_k^2 - \frac{1}{n}\left(\sum_{k=1}^{n}x_k\right)^2\right] = \frac{1}{n}\left[(x_1^2 + x_2^2 + ... + x_n^2) - \frac{1}{n}(x_1 + x_2 + ... + x_n)^2\right].$$

2.13. Mean and variance for frequency distribution: If, for $1 \le i \le k$, y_i occurs with frequency f_i, then $n = \sum_{i=1}^{k} f_i$,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} y_i f_i, \quad s^2 = \frac{1}{n}\sum_{i=1}^{k}(y_i - \overline{x})^2 f_i,$$

and the computing formula for s^2 is now

$$s^2 = \frac{1}{n}\left[\sum_{i=1}^{k} y_i^2 f_i - \frac{1}{n}\left(\sum_{i=1}^{k} y_i f_i\right)^2\right].$$

The mean and the computing formula for variance can be smoothly calculated by adding two columns to the frequency table: $y_i f_i$ and $y_i^2 f_i$.

2.14. Example. Find the mean, mode, median, variance and standard deviation of the following data.

number	frequency
−2	10
0	10
1	2
2	4
5	4

Solution. The mode is the value (or values) with the largest frequency. Thus the modes are −2 and 0. For the median, imagine all thirty numbers lined up in order: −2 ten times, then 0 ten times, 1 twice, 2 four times, 5 four times. The middle values are the 15^{th} and 16^{th}; these are both 0, so the median is 0.

Now let's expand the table:

number, y_i	frequency, f_i	$y_i f_i$	$y_i^2 f_i$
−2	10	−20	40
0	10	0	0
1	2	2	2
2	4	8	16
5	4	20	100
sum	30	10	158

From our enlarged table, we can read off the mean $\overline{x} = \frac{10}{30} = \frac{1}{3}$, the variance $s^2 = \frac{1}{30}\left(158 - \frac{1}{30}(10)^2\right)$, and the standard deviation

$$s = \sqrt{\frac{1}{30}\left(158 - \frac{1}{30}(10)^2\right)}.$$

2.15. Mean and variance for frequency distribution with class intervals: This is the same as 2.13, with y_i chosen to be the class mark of the i^{th} interval.

2.16. Example. Find the mean and variance of the following data.

class	frequency
1 − 3	6
4 − 6	2
7 − 9	2

Solution. Expand the table:

class	frequency, f_i	class mark, y_i	$y_i f_i$	$y_i^2 f_i$
$1-3$	6	2	12	24
$4-6$	2	5	10	50
$7-9$	2	8	16	128
total	10		38	202

Now stare at the final row, to get $\overline{x} = \frac{38}{10}, s^2 = \frac{1}{10}\left(202 - \frac{1}{10}(38)^2\right)$

2.17. Terminology remark. A distinction is sometimes made between *population* mean, variance, standard deviation, and *sample* mean, variance, standard deviation. There is even different terminology: $\overline{x}$ for sample mean, μ for population mean, s^2 for sample variance, σ^2 for population variance. The formulas for sample mean and population mean are the same as Definition 2.2; the formula for population variance is the same as Definition 2.9; however, the formula for sample variance of $x_1, ..., x_n$ is

$$\frac{1}{n-1}\sum_{k=1}^{n}(x_k - \overline{x})^2 \equiv \frac{1}{n-1}\left[(x_1-\overline{x})^2 + (x_2-\overline{x})^2 + ... + (x_n-\overline{x})^2\right].$$

III. PROBABILITY

3.1. A **sample space** is the set of all possible outcomes of an experiment.

3.2. An **event** is a subset of the sample space. All subsets we consider will be events.

3.3. A **sure event** contains all the outcomes in the sample space. An **impossible event** contains no outcomes.

3.4. Example. A poker hand is 5 cards, dealt from a deck of 52 cards. The sample space here is sets of 5 (cards), chosen from a set of 52. One example of an event would be getting 4 aces; more precisely, this is sets of 5 cards that include all 4 aces.

3.5. If A is an event, then we write $P(A)$ for the probability of A.

3.6. If A is sure, then $P(A) = 1$. If A is impossible, then $P(A) = 0$. In general, for any event A,

$$0 \leq P(A) \leq 1.$$

3.7. Probability $\sim$ relative frequency: When all outcomes are equally likely, then, for any event A,

$$P(A) = \frac{(\text{ number of outcomes in } A)}{(\text{ number of possible outcomes })}.$$

3.8. We say that things are chosen **at random** from a population if any member of the population is equally likely to be chosen.

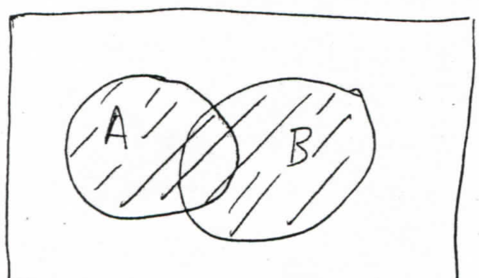
3.9. Example. Suppose a barrel contains ten wolverines, three of whom are rabid. You reach in and pull out a wolverine at random (you wear gloves, for safety). What is the probability that the wolverine you chose is rabid?

Solution. $\frac{3}{10}$.

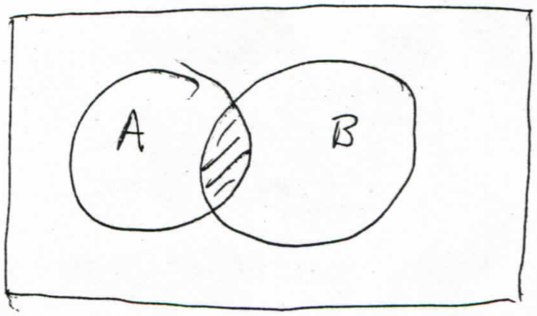
3.10. Addition Law:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$

It is worthwhile to think of probability as area. In the following picture, the area of the piece representing an event is thought of as the probability of that event.



A or B



A and B

3.11. Example. Suppose the probability that I will fish is .5, the probability that I will cut bait is .6, and the probability that I will both fish and cut bait is .2. What is the probability that I will either fish or cut bait?

Solution. Let A be "I will fish", B be "I will cut bait". Then $P(A) = .5, P(B) = .6, P(A \text{ and } B) = .2$, and we want

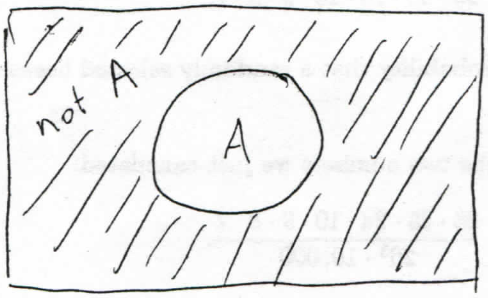
$$P(A \text{ or } B) = .6 + .5 - .2 = .9.$$

3.12. Two events A and B are **mutually exclusive** if $P(A \text{ and } B) = 0$.
Then the addition law becomes simpler:

$$P(A \text{ or } B) = P(A) + P(B) \quad \text{if } A \text{ and } B \text{ are mutually exclusive .} \tag{3.13}$$

3.14. Law of the Complement:

$$P(\text{ not } A) = 1 - P(A).$$



3.15. Examples.

1. Suppose that the probability that a wolverine is hungry is .7, the probability that it is thirsty is .75, and the probability that it is both hungry and thirsty is .6.

a. Find the probability that a wolverine is neither hungry nor thirsty.

b. Find the probability that a wolverine is either not hungry or not thirsty.

Solution. Let A be "wolverine is hungry," B be "wolverine is thirsty"; $P(A) = .7, P(B) = .75, P(A \text{ and } B) = .6$, and we want $P(\text{not}(A \text{ or } B))$.

a. By the Addition Law,

$$P(A \text{ or } B) = .7 + .75 - .6 = .85,$$

so by the Law of the Complement,

$$P(\text{ not}(A \text{ or } B)\ ) = 1 - .85 = .15.$$

b. Here we want

$$P((\text{ not } A) \text{ or } (\text{ not } B)) = P(\text{ not } (A \text{ and } B)) = 1 - .6 = .4,$$

by the Law of the Complement.

2. Suppose I am sitting at home all day waiting for phone calls. The probability that I will get no phone calls is .2, and the probability that I will get exactly one phone call is .5. What is the probability that I will get at least two phone calls?

Solution. Let A be "getting no phone calls," B be "getting exactly one phone call," C be "getting at least two phone calls." Then $C = \text{not } (A \text{ or } B)$, so

$$P(C) = 1 - P(A \text{ or } B) = 1 - (P(A) + P(B)) = 1 - (.2 + .5) = .3,$$

since A and B are mutually exclusive.

3.16. Basic Counting Principle: If an experiment has r outcomes, and another experiment has k outcomes, regardless of the outcome of the first experiment, then the combined experiment has rk outcomes.

3.17. Examples.

1. How many license plates, consisting of three letters (A, B, C, ..., Z), followed by four numbers $(0, 1, 2, ..., 9)$, can be formed?

Solution. For each of the 3 letters, 26 choices; for each of the 4 numbers, 10 choices. Multiply them together:

$$26^3 \cdot 10,000.$$

2. Same as 1, except that no number or letter may be repeated.

Solution. First letter: 26 choices; second letter: 25 choices; third letter: 24 choices; first number: 10 choices; second number: 9 choices; third number: 8 choices; fourth number: 7 choices. Multiply:

$$26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 \cdot 7.$$

3. For license plates as in 1, what is the probability that a randomly selected license plate will have no letter or number repeated?

Solution. By 3.7, we take the ratio of the two numbers we just calculated:

$$\frac{26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 \cdot 7}{26^3 \cdot 10,000}.$$

3.18. A **permutation** (or ordered arrangement) of n things, taken r at a time, means an ordered set of size r, chosen from a set of size n. The number of such permutations, written ${}_nP_r$, is given by

$${}_nP_r = \frac{n!}{(n-r)!},$$

where $k!$ ("k factorial") is $k(k-1)(k-2)\cdots 2 \cdot 1$, for k an integer greater than one, and $0! \equiv 1$.

The permutation counting, ${}_nP_r$, is a special case of the basic counting principle: imagine choosing the elements in your ordered set of size r; you have n choices for the first element, $(n-1)$ choices for the second, $(n-2)$ choices for the third, ..., $(n-r+1)$ choices for the r^{th}. The basic counting principle says you multiply them together

$${}_nP_r = n(n-1)(n-2)\cdots(n-r+1),$$

which equals $\frac{n!}{(n-r)!}$.

3.19. A **combination** of n things, taken r at a time, means a subset of size r, chosen from a set of size n. The number of such combinations, which is written $\binom{n}{r}$, is given by

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}.$$

This can be understood by taking the permutation counting, and correcting it to adjust to the fact that we no longer care about order. In the permutation counting, each set of size r got counted $r!$ times, since there are $r!$ ways to arrange r things.

3.20. Examples.

1. There are 15 wolverines in a barrel, 10 of whom are rabid. You choose 6 of them at random. What is the probability that exactly 4 of those 6 is rabid?

Solution. For the sample space, we have combinations of 15 things, taken 6 at a time. So $\binom{15}{6}$ goes in the denominator, in 3.7. For the numerator in 3.7, first you have combinations of 10 (rabid) things, taken 4 at a time; there are $\binom{10}{4}$ of those. Then you must also take combinations of 5 (non-rabid) things, taken 2 at a time; there are $\binom{5}{2}$ of those; the remaining 2 must not be rabid, or you will have more than 4 rabid wolverines. The Basic Counting Principle says you multiply $\binom{10}{4}$ by $\binom{5}{2}$, for the number of desired outcomes. Finally, 3.7 gives us

$$\frac{\binom{10}{4}\binom{5}{2}}{\binom{15}{6}} = \frac{\frac{10!}{4!6!}\frac{5!}{2!3!}}{\frac{15!}{6!9!}} = \frac{10!5!6!9!}{4!6!2!3!15!}.$$

2. Suppose there are four political factions, Monarchists, Anarchists, Deists, and Exorcists; call them M, A, D and E, for short. There are 10 Ms, 20 As, 10 Ds and 60 Es in a room. If I choose a committee of 10 at random from these people, what is the probability that they will be represented proportionately, that is, 1 M, 2As, 1 D and 6 Es?

Solution. For our sample space, we have combinations of 100 things, taken 10 at a time; there are $\binom{100}{10}$ of those.

For our desired outcome, we have four experiments: choosing 1 M out of 10, choosing 2 As out of 20, choosing 1 D out of 10 and choosing 6 Es out of 60. The number of outcomes in each experiment is, respectively, $\binom{10}{1}, \binom{20}{2}, \binom{10}{1}$ and $\binom{60}{6}$. To count the outcomes in the combined experiment, you multiply those 4 numbers together (Basic Counting Principle), to get (from 3.7 again)

$$\frac{\binom{10}{1}\binom{20}{2}\binom{10}{1}\binom{60}{6}}{\binom{100}{10}}$$

3. If a committee of 9 is chosen, from a group of 25 men and 15 women, what is the probability that it will contain exactly 5 women? At least 5 women?

Solution. The sample space is combinations of 40 things (all right, people) taken 9 at a time; so our denominator is $\binom{40}{9}$. For the numerator, we need combinations of 15 women taken 5 at a time ($\binom{15}{5}$ of those) and combinations of 25 men taken 4 (9 minus 5) at a time ($\binom{25}{4}$). Basic Counting tells us to multiply those last two countings together, to count what we want, so 3.7 gives us

$$\frac{\binom{15}{5}\binom{25}{4}}{\binom{40}{9}}$$

for the answer to the first question. For the second question, the numerator now needs a large sum: one term for 5 women, one for 6, etc.:

$$\frac{\binom{15}{5}\binom{25}{4} + \binom{15}{6}\binom{25}{3} + \binom{15}{7}\binom{25}{2} + \binom{15}{8}\binom{25}{1} + \binom{15}{9}\binom{25}{0}}{\binom{40}{9}}$$

4. If a President, Vice-President, and Premier are chosen at random from 25 men and 15 women, what is the probability that all will be women, if anyone may hold any number of offices?

Solution. For the sample space, we have 40 choices for Pres., 40 choices for *V.P.*, and 40 choices for Premier; Basic Counting counts $40 \times 40 \times 40 \equiv 40^3$ possibilities. For the desired event, replace "40" with "15," to get the fraction

$$\frac{15^3}{40^3}.$$

5. Same as 4, except each person may hold at most one office.

Solution. Now our sample space choices are 40 for Pres., 39 for V.P., 38 for Premier, to make a denominator of $40 \times 39 \times 38$. Identically, for the numerator, $15 \times 14 \times 13$, so the answer is

$$\frac{15 \times 14 \times 13}{40 \times 39 \times 38}.$$

This problem could also be done with combinations: think of the sample space as choosing 3 people out of 40 available, the desired space as choosing 3 women out of 15 available, to get

$$\frac{\binom{15}{3}}{\binom{40}{3}}.$$

I leave it to you to verify that the two different-looking answers we got are actually the same number; otherwise we're in trouble.

6. At a party of 10, each person shakes hands with everyone else. How many handshakes will there be?

Solution. This is the same as counting the number of pairs of people that can be chosen from 10; in other words, combinations of 2 people taken 10 at a time: $\binom{10}{2} = 45$.

7. Again, 10 people at a party. What is the probability that 2 people have the same birthday? (Assume no one is born on February 29 of a leap year.)

Solution. The sample space here is combinations of 10 (possibly repeated) days taken out of 365 available; 365^{10} possibilities. It's easier to look at the complement: the event of no two people having the same birthday. Using Basic Counting as in 3.17.2 gives us $365 \times 364 \times ... \times 356$ ways to do that. So the probability of the *complement* is $\frac{365 \times 364 \times ... \times 356}{365^{10}}$, thus the probability we want is

$$1 - \frac{365 \times 364 \times ... \times 356}{365^{10}} \sim .12.$$

IV. CONDITIONAL PROBABILITY

4.1. Conditional Probability: $P(A|B)$ means the (conditional) probability of A, given B. This means the probability of A when you know that B has occurred.

4.2. Example. Let $A \equiv$ you study, $B \equiv$ you pass the class. You must believe that

$$P(B|\text{ not } A) < P(B) < P(B|A).$$

Here is a formula for conditional probability:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \tag{4.3}$$

when $P(B) = 0$.

Note that a formula for $P(A \text{ and } B)$ follows quickly:

$$P(A \text{ and } B) = P(A|B)P(B). \tag{4.4}$$

4.5. Example. Suppose 95 percent of rabid wolverines are excitable, 60 percent of non-rabid wolverines are excitable and 1 percent of wolverines are rabid. Find the probability that a wolverine selected at random is

a. excitable and rabid;

b. excitable and not rabid;

c. excitable.

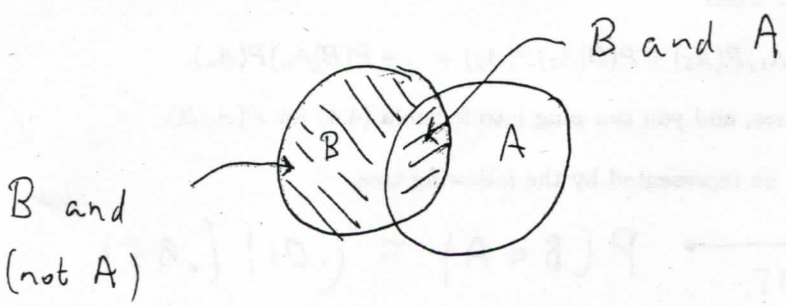
Solution. Let A be the event of the wolverine being rabid, B the event of the wolverine being excitable. Then

$$.95 = P(B|A), \quad .6 = P(B|\text{ not } A), \quad .01 = P(A).$$

a. $P(B \text{ and } A) = P(B|A)P(A) = (.95)(.01) = .0095.$

b. $P(B \text{ and } (\text{ not } A)) = P(B|\text{ not } A)P(\text{ not } A) = (.6)(1-.01) = .594.$

c. $P(B) = P(B \text{ and } A) + P(B \text{ and } (\text{ not } A)) = .0095 + .594 = .6035$, by 3.13.



4.6. Two events A and B are **independent** if $P(A|B) = P(A)$.

Then we have a simpler formula for $P(A \text{ and } B)$:

$$P(A \text{ and } B) = P(A)P(B) \text{ if } A \text{ and } B \text{ are independent .} \tag{4.7}$$

Note that this implies that $P(B|A) = P(B)$.

4.8. Example. Suppose the probability of getting dandruff is .1 and the probability of getting a hangnail is .2.

a. If dandruff and hangnails are independent, find the probability of getting both dandruff and a hangnail.

b. If dandruff and hangnails are independent, find the probability of getting either dandruff or a hangnail.

c. If dandruff and hangnails are independent, find the probability of getting neither dandruff nor a hangnail.

d. If the probability of getting either dandruff or a hangnail is .12, find the probability of getting both dandruff and a hangnail.

Solution. Let A be the event of getting dandruff, B the event of getting a hangnail. $P(A) = .1$ and $P(B) = .2$.

a. By independence, $P(A \text{ and } B) = P(A)P(B) = (.1)(.2) = .02$.

b. By the Addition Law,

$$P(A \text{ or } B) = .1 + .2 - .02 = .28.$$

c. By the Law of the Complement,

$$P(\text{ not } (A \text{ or } B)) = 1 - .28 = .72.$$

OR, by independence,

$$P((\text{ not } A) \text{ and } (\text{ not } B)) = P(\text{ not } A)P(\text{ not } B) = (1 - .1)(1 - .2) = .72.$$

d. Now $P(A \text{ or } B) = .12$. By the Addition Law again,

$$P(A \text{ and } B) = .1 + .2 - .12 = .18.$$

NOTE: in d, A and B are not independent.

4.9. Bayes' theorem: Suppose $A_1, A_2, ..., A_n$ are mutually exclusive and $P(A_1 \text{ or } A_2 \text{ or } ... \text{ or } A_n) = 1$. Then for any event B, $1 \le i \le n$,

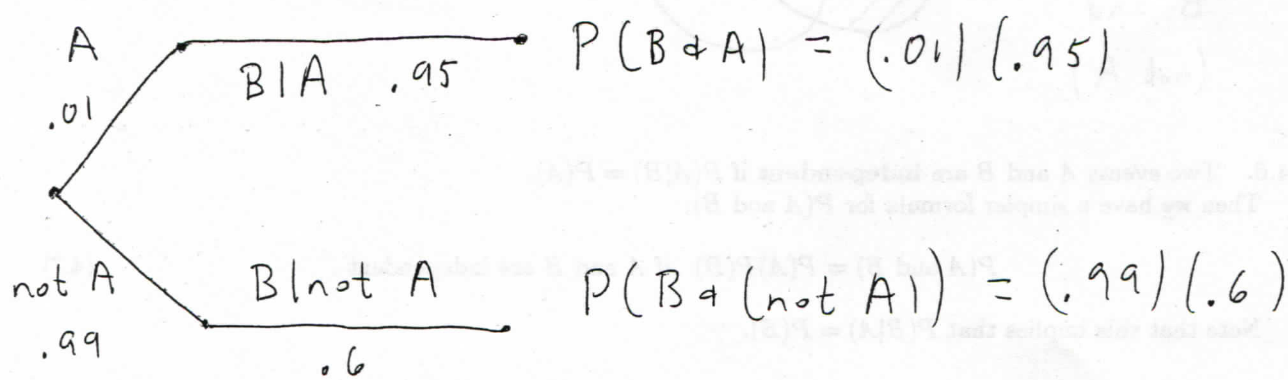
$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + ... + P(B|A_n)P(A_n)}.$$

But don't use this formula directly. I recommend that you draw a tree. The first set of branches represents $A_1, ..., A_n$, then each A_i branch leads to a $B|A_i$ branch. At the end of the i^{th} branch, you obtain $P(B \text{ and } A_i)$, by multiplying $P(A_i)$ and $P(B|A_i)$. Then

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + ... + P(B|A_n)P(A_n),$$

the sum of the ends of those branches, and you can plug into formula (4.4) for $P(A_i|B)$.

For instance, Example 4.5 could be represented by the following tree.



Bayes' theorem (or a tree representing it) is of particular interest in situations where you're given $P(B|A)$ and you want $P(A|B)$.

B may be thought of as a test for A; for example, A could be the event of having the flu, and B could be the event of having a fever. A is what we care about, but it's hard to determine directly; bacteria are so small. So we test for A with B; that is, if you have a fever then I'd like to deduce that you have the flu. The probability that this deduction is true is precisely $P(A|B)$.

4.10. Examples.

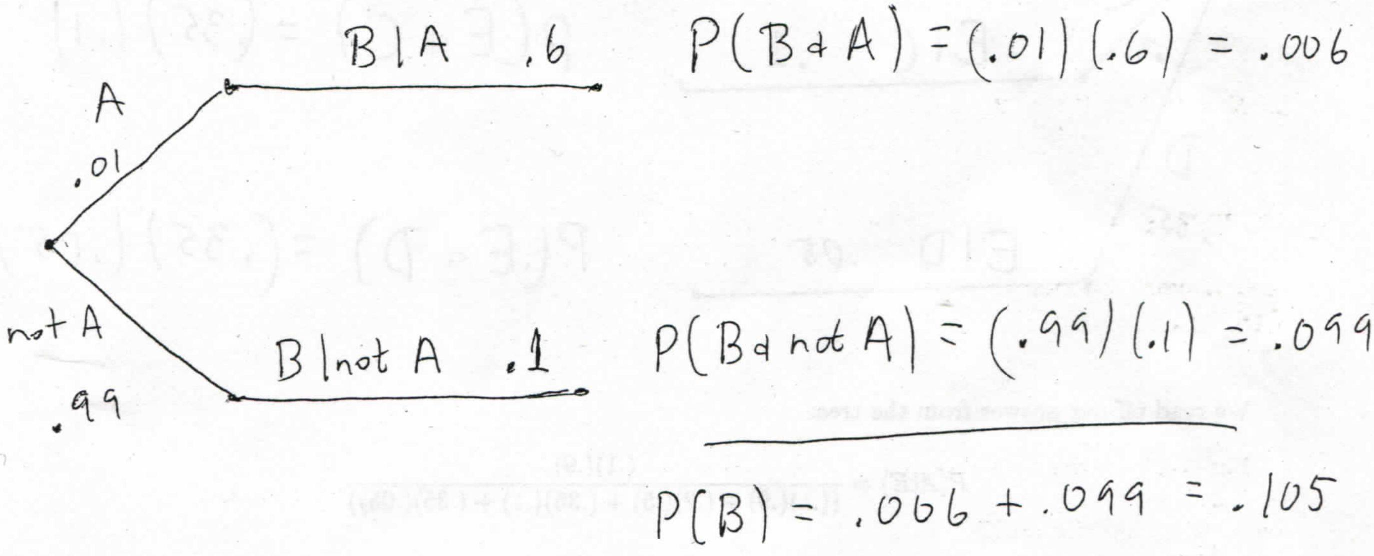
1. Suppose that if you are a psychopath there is a sixty percent chance you are a loner, if you are not a psychopath there is a ten percent chance you are a loner, and one percent of people are psychopaths.

a. What percent of the population are loners?

b. Loners are looking pretty suspicious to me, so every time I see a loner, I will accuse him/her of being a psychopath (this is my test for being a psychopath). What percent of the time will I be right?

Solution. Let A be the event of being a psychopath, B the event of being a loner. We're told that $P(B|A) = .6, P(B|\text{ not } A) = .1$ and $P(A) = .01$.

In (a) we're asked for $P(B)$; in (b) we're asked for $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$. We need a tree that tells us everything.



From the tree, we can read off our answers:

a. $P(B) = (.01)(.6) + (.99)(.1) = .105$.

b. $P(A|B) = \frac{(.01)(.6)}{((.01)(.6)+(.99)(.1))} = \frac{.006}{.105} = \frac{2}{35}$.

2. Suppose there are four colleges, call them A, B, C, D, that students from Brutopia attend. 10 percent of students attend College A, 20 percent attend College B, 35 percent attend College C, and 35 percent attend College D. 90 percent of students who attend College A are successful, 50 percent of students who attend College B are successful, 10 percent of students who attend College C are successful, and 5 percent of students who attend College D are successful.

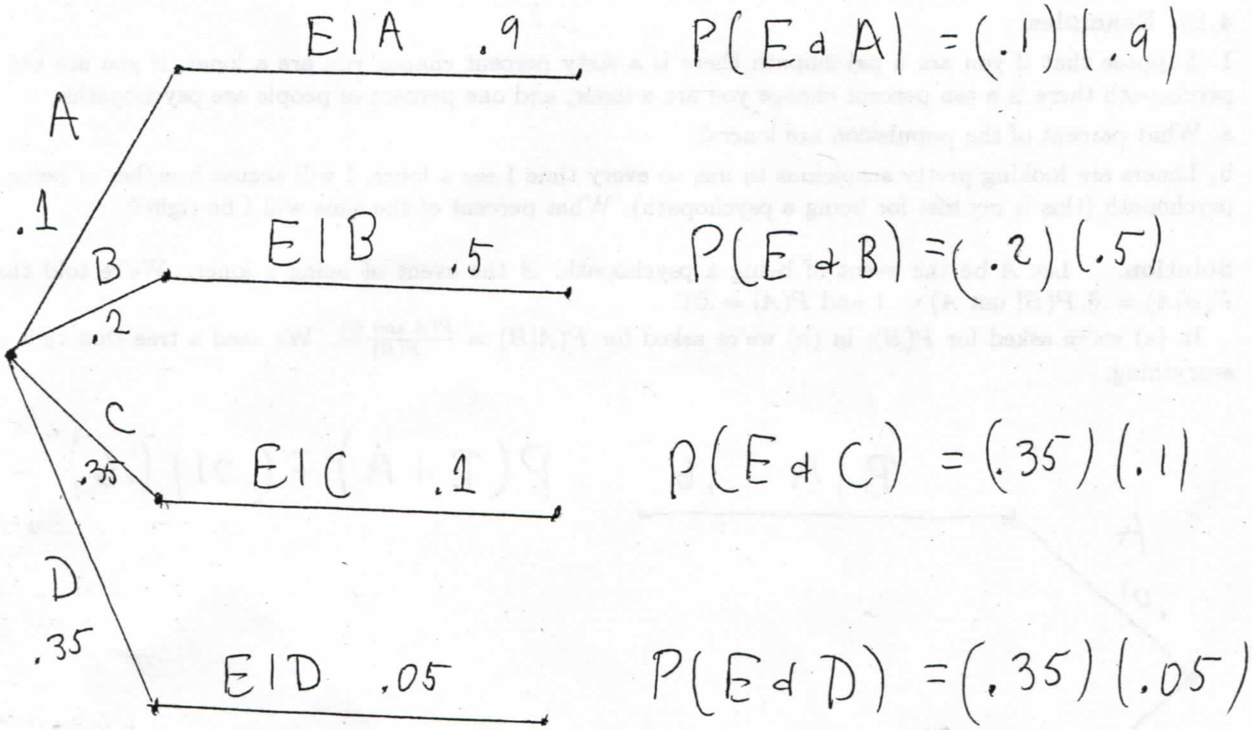
If a former student in Brutopia is successful, what is the probability that he/she attended College A?

Solution. For short, let's write A for the event "attends college A," likewise for B, C and D. The other event of interest is being successful; let's call that E.

We have $P(A) = .1, P(B) = .2, P(C) = .35, P(D) = .35, P(E|A) = .9, P(E|B) = .5, P(E|C) = .1$ and $P(E|D) = .05$.

We want $P(A|E) = \frac{P(A \text{ and } E)}{P(E)}$.

Now our tree has four branches:



We read off our answer from the tree:

$$P(A|E) = \frac{(.1)(.9)}{((.1)(.9) + (.2)(.5) + (.35)(.1) + (.35)(.05))}.$$

V. RANDOM VARIABLES

We have probabilities defined for events, that is, subsets of a sample space. These can be mysterious or irrelevant objects. For example, if you're flipping a coin, and getting a dollar for every head, and losing 2 dollars for every tail, the sample space is $\{H, T\}$, and our probabilities are $P(H)$ and $P(T)$ ($T \equiv$ tails, $H \equiv$ heads). But what you care about is *money*, or, more precisely, the probability of making various amounts of money.

Defining a random variable will allow us to focus on what we care about (such as money).

5.1. A **random variable** is a function, X, from a sample space into the real numbers; that is, it assigns a numerical value to each outcome in a sample space.

In the example before this definition, we would define $X(H) \equiv 1$ (in words, "X of H equals one"), $X(T) \equiv -2$.

5.2. A random variable is **discrete** if all possible values are separated.

5.3. A random variable X is **continuous** if it can assume any value in an interval or intervals, & $P(X = c) = 0$, for any number c.

5.4. For a discrete random variable, X, we construct a **probability function:** list all values X can take, along with their probabilities, $p(x_k) \equiv P(X = x_k)$.

5.5. Example. Suppose there are three grocery stores in town, call them A, B and C. Store A is 1.5 miles from my home, B is 4 miles from my home, and I live in C. Every day I go to one and only one of those three stores; I go to A 10 percent of the time, B 60 percent of the time, and C 30 percent of the time. Let X be the distance I travel to and from a grocery store each day (round trip). Write down the probability function for X.

Solution. As a random variable, $X(A) \equiv 1.5, X(B) \equiv 4$ and $X(C) \equiv 0$. This is not the probability function. Here is the probability function for X.

x	0	1.5	4
$p(x)$	.3	.1	.6

All we care about are values of X, and how likely each of those values is to occur; the identity of the grocery stores is of no interest.

5.6. Properties of probability function: Suppose $x_1, ..., x_n$ are the possible values of the discrete random variable X, with probabilities $p(x_1), ..., p(x_n)$. Then

(1) $0 \leq p(x_k) \leq 1$, for $1 \leq k \leq n$; and

(2) $p(x_1) + ... + p(x_n) = 1$.

5.7. The **mean, or expected value** of the discrete random variable X, with possible values $x_1, ..., x_n$ is

$$E(X) = \mu \equiv x_1 p(x_1) + ... + x_n p(x_n).$$

5.8. The **variance** of X is

$$\sigma^2 = (x_1 - \mu)^2 p(x_1) + ... + (x_n - \mu)^2 p(x_n).$$

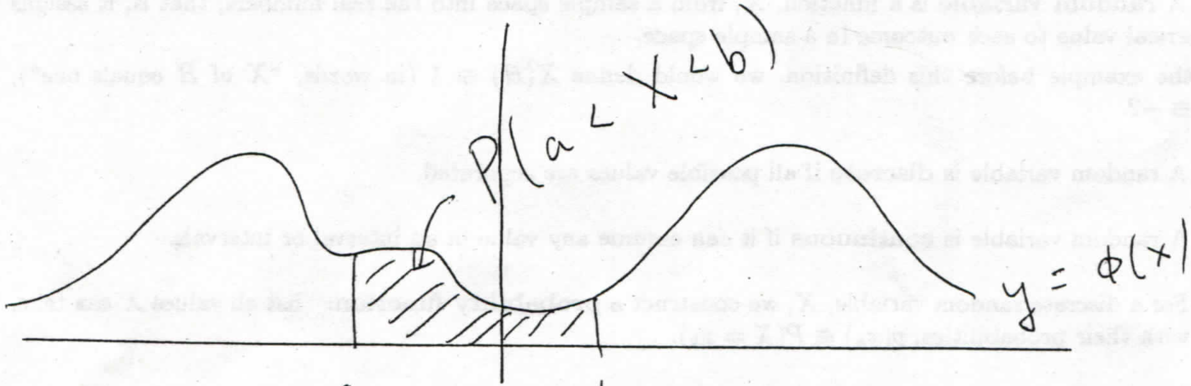
5.9. The **standard deviation** of X is

$$\sigma \equiv \sqrt{\sigma^2}.$$

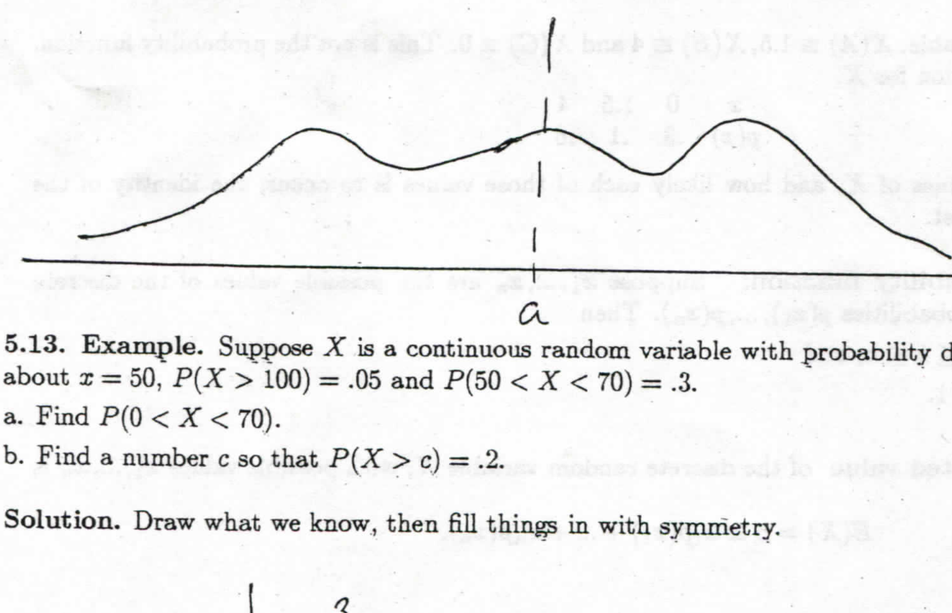
5.10. Example. For the random variable in Example 5.5,

$$\mu = 0(.3) + 1.5(.1) + 4(.6) = 2.55,$$
$$\sigma^2 = (0-2.55)^2(.3) + (1.5-2.55)^2(.1) + (6-2.55)^2(.6),$$
$$\sigma = \sqrt{(0-2.55)^2(.3) + (1.5-2.55)^2(.1) + (6-2.55)^2(.6)}.$$

5.11. Probability density function, for a *continuous* random variable X: a function ϕ such that, for any a, b, $P(a < X < b)$ equals the area between $y = \phi(x)$, $x = a, x = b$ and the x axis.



5.12. A probability density function is symmetric about $x = a$ if the graph to the right of $x = a$ is the mirror image of the graph to the left.

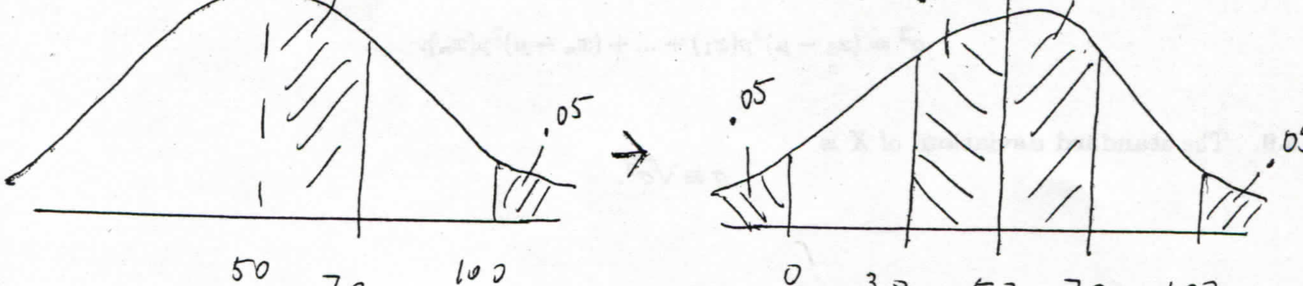


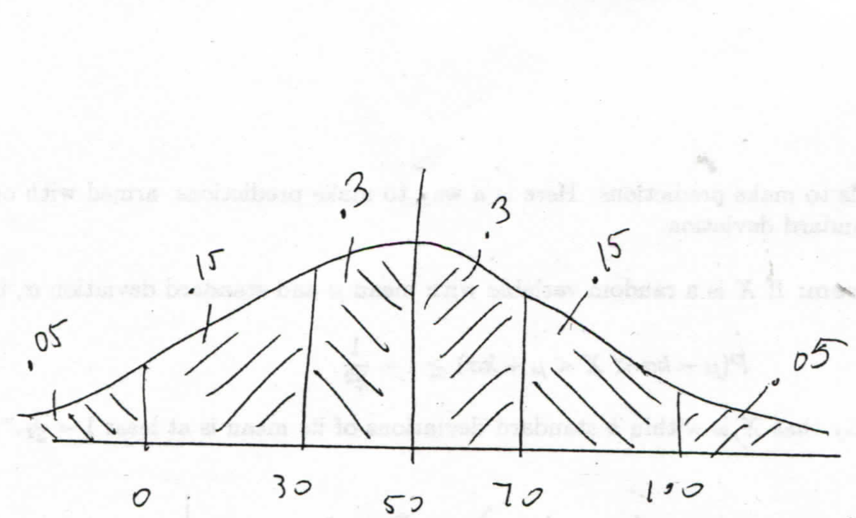
5.13. Example. Suppose X is a continuous random variable with probability density function symmetric about $x = 50$, $P(X > 100) = .05$ and $P(50 < X < 70) = .3$.

a. Find $P(0 < X < 70)$.

b. Find a number c so that $P(X > c) = .2$.

Solution. Draw what we know, then fill things in with symmetry.





$(.15 = .5 - (.05 + .3)$; OR $.15 = \frac{1}{2}(1 - (.05 + .3 + .3 + .05)))$

a. Now it is clear that $P(0 < X < 70) = .15 + .3 + .3 = .75$.

b. Well, *guess*, starting on the right: $P(X > 100) = .05 = .2$; $P(X > 70) = .15 + .05 = .2$, YES, so choose $c = 70$.

Example 5.13b is an example of *percentile*.

5.14. For any number α between 0 and 100, X a continuous random variable, the **αth percentile** of X is a number c so that $P(X < c) = \alpha\%$.

For example, in Example 5.13, 70 is the 20th percentile of X and 100 is the 5th percentile of X.

Without calculus, we can't do a lot with probability density functions; that shaded area in 5.11, representing

$$P(a < X < b),$$

the probability that X is between a and b, is the integral, from a to b, of the probability density function ϕ. But I can't talk about that, since calculus is not a prerequisite for this class.

In practice, we will use tables to get these probabilities.

Remark 5.15. For any random variable, we still have a notion of *mean* and *standard deviation*, as we defined for discrete random variables in 5.7 and 5.9. Without calculus, we can't write down the definition of mean or standard deviation for a continuous random variable. But the idea is the same: two numbers, one of which describes central tendency, the other dispersion.

Example 5.16. Suppose a chasm is 36 inches wide, and your long jump has a mean of 37 inches. Is it safe for you to try to jump across the chasm?

The correct answer is "it depends." First it depends on what you mean by "safe." Let's say by "safe" you mean a probability of less than 1% that you will plummet into the chasm; that is, if X is your long jump, you want

$$P(X < 36) < .01.$$

The answer to the question now depends on the *standard deviation* of your long jump X. If the standard deviation of X is large, that is, your long jumps vary a lot, then that average jump of 37 inches isn't so reassuring; you might very easily jump shorter than the necessary 36 inches.

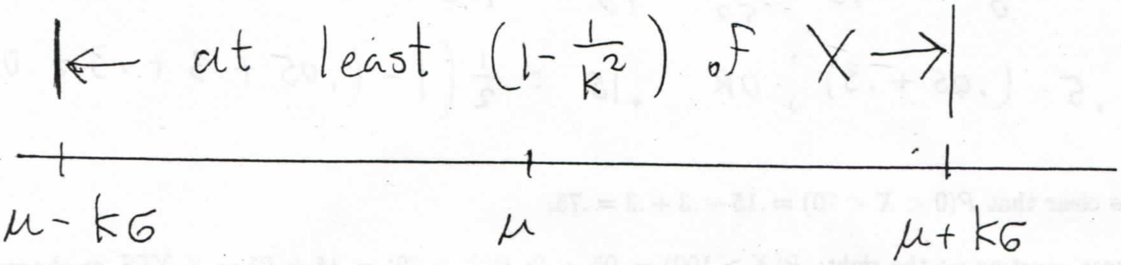
Chebyshev's theorem below will give us a precise way of answering questions like this, if we have both the mean and standard deviation of the relevant random variable.

The goal of probability is to make predictions. Here is a way to make predictions, armed with only two numbers, the mean and standard deviation.

5.17. Chebyshev's theorem: If X is a random variable with mean μ and standard deviation σ, then for any number $k \geq 1$,

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}.$$

In words: "the probability that X is within k standard deviations of its mean is at least $1 - \frac{1}{k^2}$."
Here is a picture:

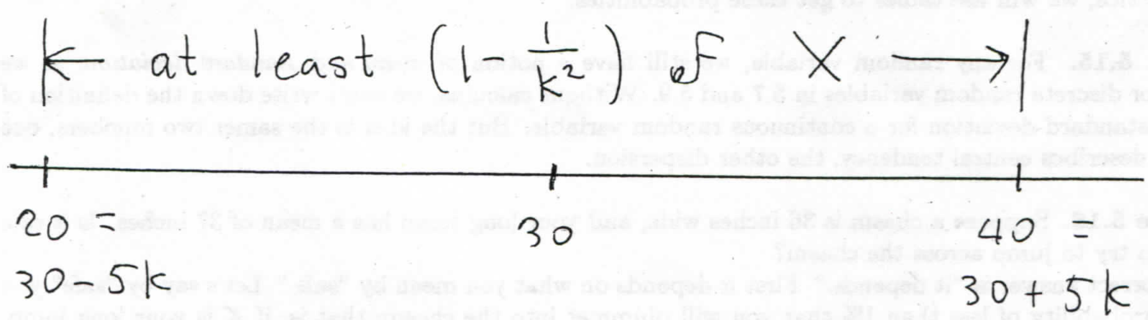


If you have no information other than the mean and standard deviation of a random variable, then you should use Chebyshev's theorem to estimate probabilities. Later, we will see how to make more precise predictions when we have more information about the random variable.

5.18. Examples:

1. Suppose that wolverine weight has a mean of thirty pounds and a standard deviation of five pounds. Estimate the probability that a randomly selected wolverine will weigh between twenty and forty pounds.

Solution. Let X be wolverine weight; here $\mu = 30, \sigma = 5$. We want to estimate $P(20 < X < 40)$.
To use Chebyshev's theorem, we need $40 = \mu + k\sigma = 30 + 5k$, and $20 = \mu - k\sigma = 30 - 5k$.



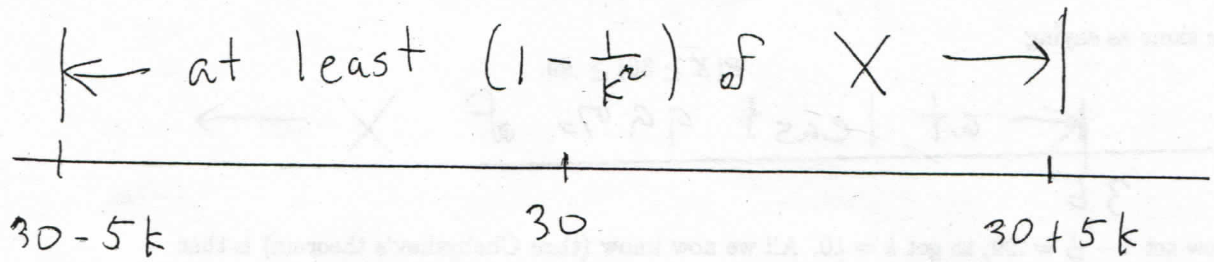
In both equations, we get $k = 2$. Thus, by Chebyshev's theorem,

$$P(20 < X < 40) = P(\mu - 2\sigma < X < \mu + 2\sigma) \geq 1 - \frac{1}{2^2} = \frac{3}{4};$$

the probability is at least $\frac{3}{4}$.

2. You want to construct a wolverine carrying crate. It will carry only one wolverine at a time. For wolverines as in 5.18(1), how strong should this crate be—that is, how much weight should it be able to carry without breaking—if you want it to carry at least 90 percent of wolverines without breaking?

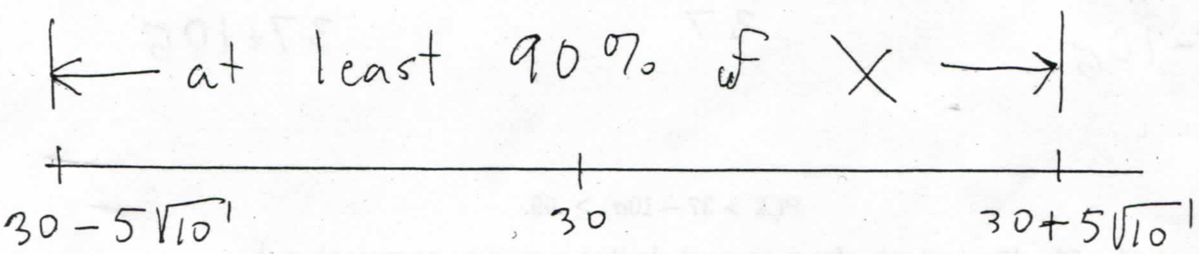
Solution. Let X, μ, σ be as in number 1. To use Chebyshev's theorem, set $.90 = 1 - \frac{1}{k^2}$, and solve for k: $\frac{1}{k^2} = 1 - .90 = .1 = \frac{1}{10}$, so $k^2 = 10, k = \sqrt{10}$.



By Chebyshev's theorem,

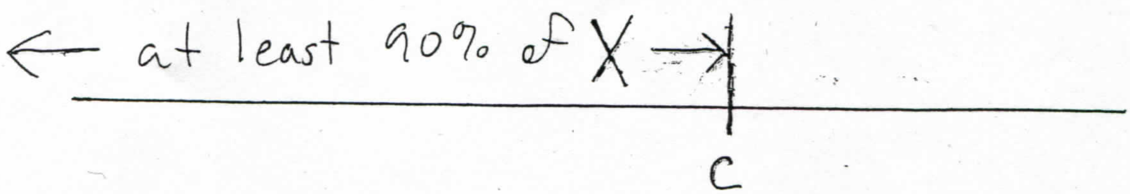
$$P(30 - 5\sqrt{10} < X < 30 + 5\sqrt{10}) = P(\mu - \sqrt{10}\sigma < X < \mu + \sqrt{10}\sigma) \geq 1 - \frac{1}{(\sqrt{10})^2} = .9.$$

In other words, at least 90 percent of wolverine weights are in the interval $(30 - 5\sqrt{10}, 30 + 5\sqrt{10})$.



Let c be the crate strength we're looking for. Recall that X is wolverine weight. In order that the crate not break, we want $X < c$; since this is supposed to happen at least 90 percent of the time, we want

$$P(X < c) \geq 90\%$$



Comparing those last two pictures tells us to choose

$$c \equiv \text{ weight of crate } = 30 + 5\sqrt{10} \text{ pounds .}$$

3. In order that the answer to the question in Example 5.16 be "yes," how small must the standard deviation of your long jump be?

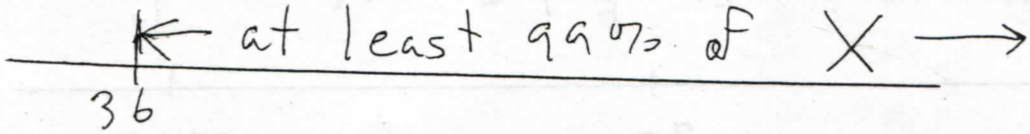
Solution. Let σ be the unknown standard deviation.

As in the previous example, we could set $1 - \frac{1}{k^2} = .01$, and solve for k. This does not give us much information; telling us where one percent of the jumps are is not telling us much. We can get much more information by noting that

$$P(X < 36) < .01$$

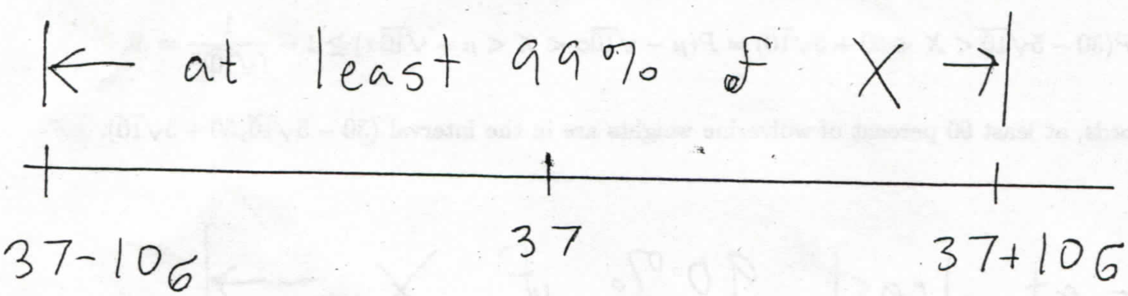
is the same as saying

$$P(X \geq 36) \geq .99.$$



Now set $1 - \frac{1}{k^2} = .99$, to get $k = 10$. All we now know (thru Chebyshev's theorem) is that

$$P(37 - 10\sigma < X < 37 + 10\sigma) \geq .99;$$



thus

$$P(X > 37 - 10\sigma) \geq .99.$$

Thus we set $36 = 37 - 10\sigma$, and solve for σ, to conclude that σ must be no more than $\frac{1}{10}$.

VI. NORMAL RANDOM VARIABLE

6.1. A **normal** random variable is a continuous random variable with probability density function

$$\phi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

This is symmetric about $x = \mu$, with mean μ and standard deviation σ. e is a famous irrational number that we won't tell you about, unless you ask. Even if you ask, it's not clear what will happen.

This probability density function is often called a *bell-shaped curve* or *bell curve.*

A random variable that is *normally distributed*, or *has a normal distribution* is a normal random variable.

A normal random variable with mean 0 and standard deviation 1 is called Z, the *standard normal.*

6.2. Important facts about Z**:** the probability density function is symmetric about $x = 0$. Aside from that, it's hard to summarize or outline how to use the Z-table. For $0 \leq z \leq 3$, our Z-table gives $P(Z \geq z)$. z is a point on the x axis, while the probabilities $P(Z \geq z)$ are areas between the probability density function and the x axis, as with any probability density function.

Using the Z-table (see after homework answers) and symmetry, you can look up any Z probability.

6.3. Examples.

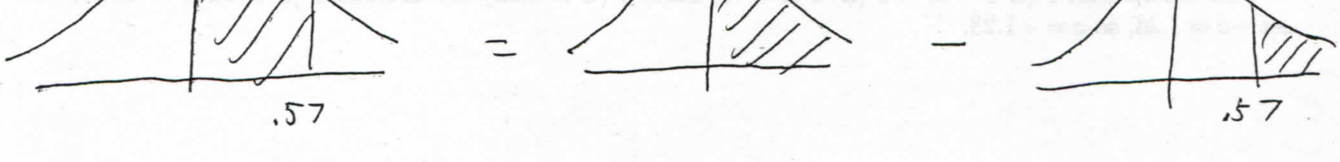
1. Find each of the following.
   (a) $P(Z > 1.5)$.
   (b) $P(Z < -1.32)$.
   (c) $P(0 < Z < .57)$.
   (d) $P(-2 < Z < 2.5)$.
   (e) $P(1.5 < Z < 2)$.

Solution. In all these, use the Z-table and symmetry.

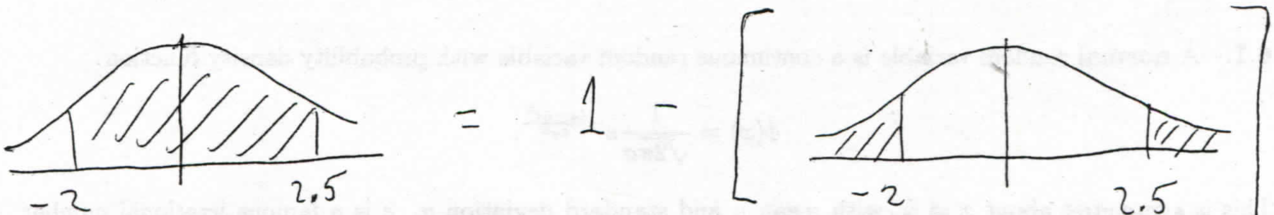
a. $P(Z > 1.5) = .0668$.

b. $P(Z < -1.32) = P(Z > 1.32) = .0934$.

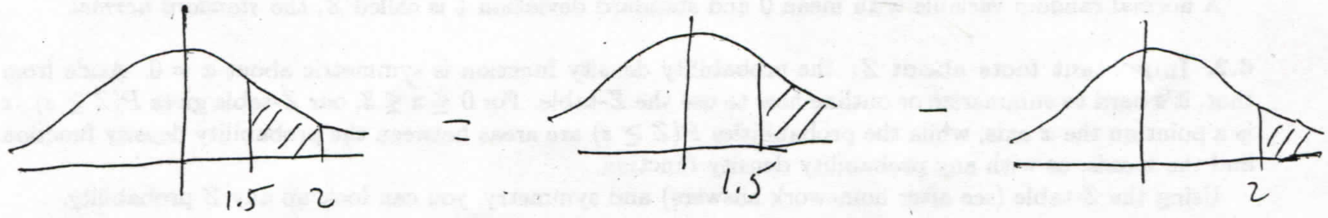
c. $P(0 < Z < .57) = P(Z > 0) - P(Z \geq .57) = .5 - .2843$.



d. $P(-2 < Z < 2.5) = 1 - [P(Z \le -2) + P(Z \ge 2.5)] = 1 - P(Z \ge 2) - P(Z \ge 2.5) = 1 - .0228 - .0062.$

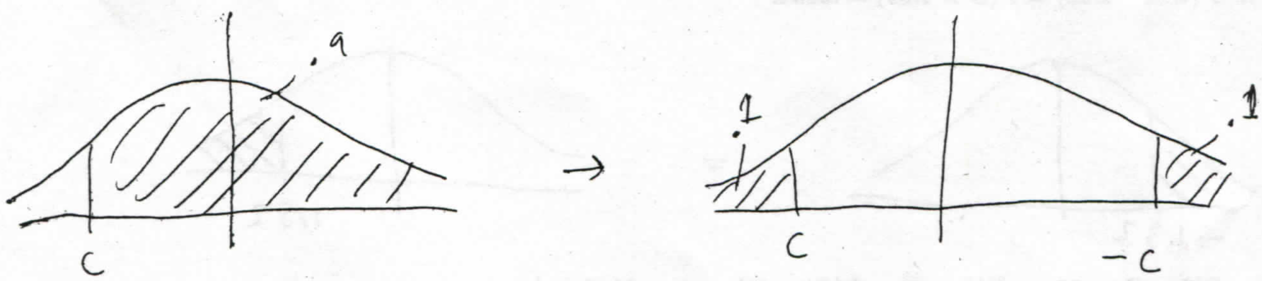


e. $P(1.5 < Z < 2) = P(Z > 1.5) - P(Z \ge 2) = .0668 - .0228.$



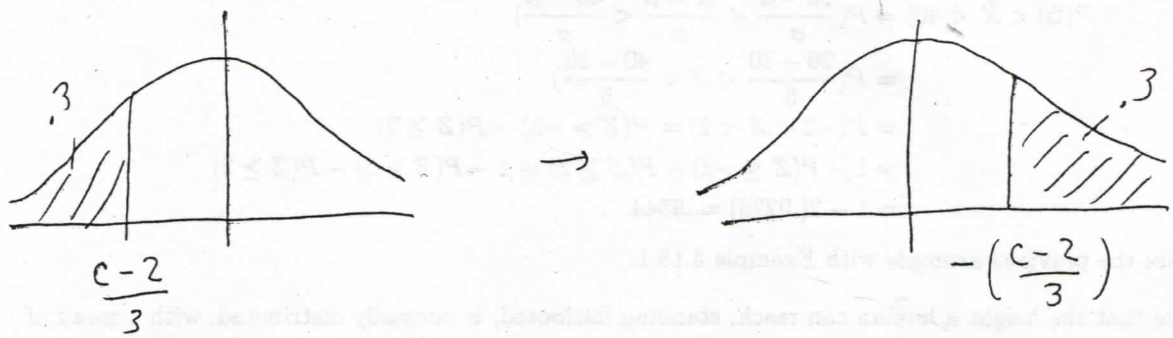
2. For each of the following, find an approximate value of c that satisfies the assertion.
   (a) $P(Z > c) = .01.$
   (b) $P(Z > c) = .9.$
   (c) $P(Z < \frac{c-2}{3}) = .3.$
   (d) $P(Z < \frac{c-19}{\frac{5}{8}}) = .79.$

Solution.
a. Looking in the middle of the table for a shaded area close to .01, I see $P(Z > 2.32) = .0102, P(Z > 2.33) = .0099$. So let $c = 2.33$.
b. Here is the picture.



This means that $P(Z > -c) = P(Z < c) = .1$. Since $P(Z > 1.28) = .1003$ and $P(Z > 1.29) = .0985$, we set $-c = 1.28$, so $c = -1.28$.

c. Here is the picture.

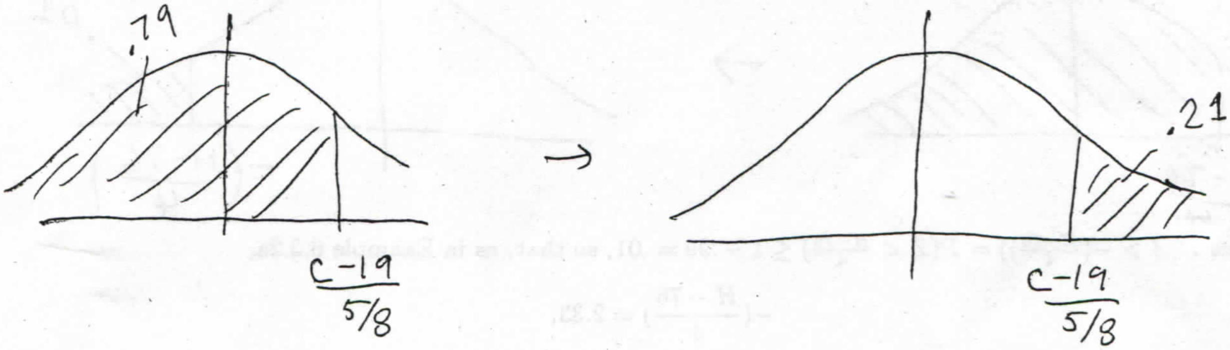


We have $P(Z > -(\frac{c-2}{3})) = P(Z < \frac{c-2}{3}) = .3$; since $P(Z > .52) = .3015, P(Z > .53) = .2981$, we let

$$-(\frac{c-2}{3}) = .52$$

and solve for $c = 2 - 3(.52)$.

d. Now we have this picture.



Thus $P(Z > \frac{c-19}{\frac{5}{8}}) = 1 - .79 = .21$; since $P(Z > .80) = .2119$ and $P(Z > .81) = .2090$, set

$$\frac{c-19}{\frac{5}{8}} = .81,$$

and solve for $c = 19 + .81(\frac{5}{8})$.

6.4. How to calculate normal probabilities: If X is normal, with mean μ and standard deviation σ, then

$$Z = \frac{X - \mu}{\sigma}; \tag{6.5}$$

use the Z table.

6.6. Examples.

1. Suppose that wolverine weight is normally distributed, with a mean of thirty pounds and a standard deviation of five pounds. What is the probability that a randomly selected wolverine will weigh between twenty and forty pounds?

Solution. Let X be wolverine weight. We're told that X is normal, with $\mu = 30, \sigma = 5$, and we want $P(20 < X < 40)$. We change it to Z, with (6.5):

$$\begin{aligned} P(20 < X < 40) &= P(\frac{20-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{40-\mu}{\sigma}) \\ &= P(\frac{20-30}{5} < Z < \frac{40-30}{5}) \\ &= P(-2 < Z < 2) = P(Z > -2) - P(Z \geq 2) \\ &= 1 - P(Z \leq -2) - P(Z \geq 2) = 1 - P(Z \geq 2) - P(Z \geq 2) \\ &= 1 - 2(.0228) = .9544. \end{aligned}$$

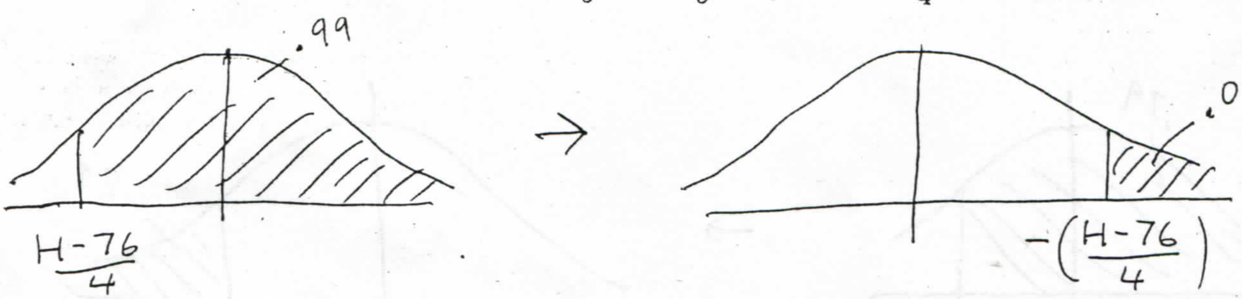
Compare the previous example with Example 5.18.1.

2. Suppose that the height a human can reach, standing flatfooted, is normally distributed, with a mean of 76 inches and a standard deviation of 4 inches. How high should a shelf be, so that at least 99 percent of people can reach it, standing flatfooted?

Solution. Let X be the height a human can reach, standing flatfooted. We're told that X is normal, with $\mu = 76, \sigma = 4$ (inches). Let H be the height of the desired shelf.

We want $P(X > H) \geq .99$. We must change to Z, with (6.5):

$$.99 \leq P(X > H) = P(\frac{X-\mu}{\sigma} > \frac{H-\mu}{\sigma}) = P(Z > \frac{H-76}{4}),$$



thus $P(Z > -(\frac{H-76}{4})) = P(Z < \frac{H-76}{4}) \leq 1 - .99 = .01$, so that, as in Example 6.3.2a,

$$-(\frac{H-76}{4}) = 2.33,$$

and we can solve for $H = 76 - 4(2.33) = 66.68$ inches.

3. Same as number 2, except do not assume that the height a human can reach, standing flatfooted, is normal.

Solution. Let X, μ, σ be as in number 2; now we cannot assume that X is normal, so we must use Chebyshev's theorem (5.17).

Set $1 - \frac{1}{k^2} = .99$; this implies that $k = 10$. Chebyshev's theorem says that

$$P(36 < X < 116) = P(\mu - 10\sigma, \mu + 10\sigma) \geq 1 - \frac{1}{10^2} = .99.$$

Since we still want $P(X > H) \geq .99$, we choose $H = 36$ inches.

NOTE that our answer in number 3 is much more extreme than the answer in number 2. That's because we have more information in number 2.

---

4. Suppose human I.Q. (Intelligence Quotient) is normally distributed and has a mean of 100 and a standard deviation of 15. Find the 99th percentile of human I.Q.

Solution. We want c so that $.99 = P(X < c) = P(Z < \frac{c-100}{15})$, so as in Examples 6.3.2, $c = 100 + 15(2.33) \sim 135$.

VII. BINOMIAL RANDOM VARIABLE AND FRIENDS

7.1. Binomial random variable: an experiment, with only two possible outcomes (success or failure), is repeated n times; at each repetition, the probability of success is a constant p.
$X \equiv$ the number of successes.

7.2. Probability function for binomial:

$$b(x,n,p) \equiv P(X=x) = \binom{n}{x} p^x q^{n-x} \ (x = 0, 1, ..., n),$$

where $q \equiv 1-p$.

This can be seen by a counting argument. The probability of a particular arrangement of x successes, e.g., x successes followed by $(n-x)$ failures, is $p^x q^{n-x}$. Then there are $\binom{n}{x}$ ways of arranging those x successes; think of it as choosing x places among n possible places, where success will occur.

7.3. Examples.

1. The archetypical binomial is coin flipping. Suppose you have a coin weighted in such a way that on each flip, the probability that it will come up heads is .3. If you flip it 27 times, what is the probability of getting exactly 5 heads?

Solution. The experiment here is to flip a coin. There are 2 possible outcomes: heads or tails (we assume the coin doesn't land on an edge). We define success to be a head. Thus our binomial random variable is the number of heads in 27 flips. The constant probability of success, at each repetition of our experiment, is the probability of getting a head.

Here $n = 27, p = .3, x = 5$, so we want

$$b(5,27,.3) = \binom{27}{5}(.3)^5(.7)^{22} = \frac{27!}{5!22!}(.3)^5(.7)^{22}.$$

Many problems can be modelled as binomial, by being open minded about how you define "success." Here's a particularly unpleasant way to do this.

2. Suppose that four percent of wolverines in the Yukon are rabid. If you select 30 wolverines from the Yukon at random, what is the probability that at least 28 of them are rabid?

Solution. Our experiment is to pick a wolverine from the Yukon. Define success to be getting a rabid one. Then the number of rabid wolverines, in selecting 30 wolverines, is a binomial random variable, with $n = 30, p = .04$.

We want

$$P(X \geq 28) = b(28,30,.04) + b(29,30,.04) + b(30,30,.04)$$
$$= \binom{30}{28}(.04)^{28}(.96)^2 + \binom{30}{29}(.04)^{29}(.96)^1 + \binom{30}{30}(.04)^{30}(.96)^0.$$

In practice, we don't like numbers like this; they take a lot of work to calculate. See 7.6, for the methods I want you to use, to calculate binomial probabilities.

7.4. Mean and variance of binomial:

$$\mu = np, \ \sigma^2 = npq.$$

7.5. The **Poisson** random variable, with parameter λ, is discrete, with probability function

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}, \quad k=0,1,2,\ldots.$$

λ is the mean of X.

7.6. How to calculate binomial probabilities.

(1) If possible, use the binomial table.

(2) If the binomial tables are impossible, and either np or nq is less than or equal to five, use the *Poisson approximation*: use the Poisson table, with parameter $\lambda \equiv np$.

(3) If the binomial tables are impossible, and both np and nq are greater than five, use the *normal approximation*:

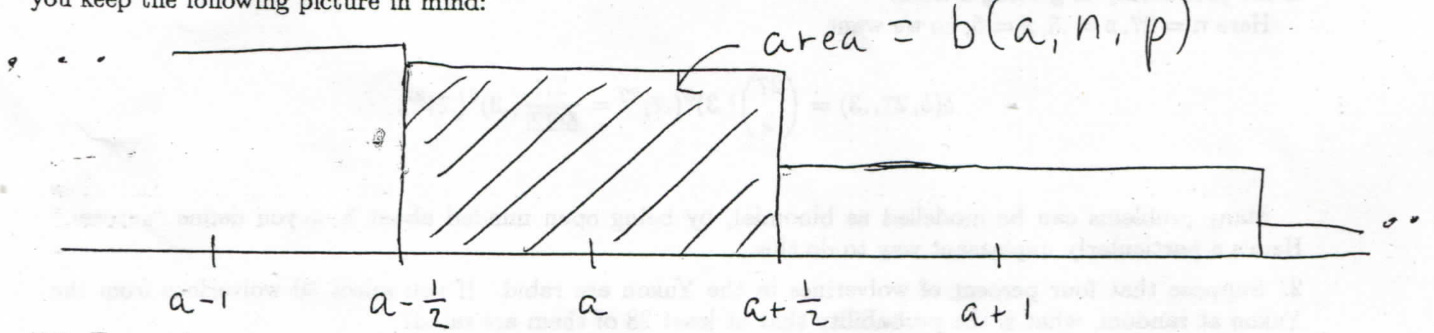
$$b(a,n,p)\equiv P(X=a)\sim P\left(\frac{a-\frac{1}{2}-\mu}{\sigma}<Z<\frac{a+\frac{1}{2}-\mu}{\sigma}\right);$$

$$P(X\geq b)\sim P\left(Z>\frac{b-\frac{1}{2}-\mu}{\sigma}\right);$$

$$P(X\leq a)\sim P\left(Z<\frac{a+\frac{1}{2}-\mu}{\sigma}\right).$$

The $\pm\frac{1}{2}$ in the normal approximation of binomial is called a *continuity correction*; it is necessary because we are approximating a discrete random variable with a continuous random variable.

I recommend that, rather than memorize those three formulas in the normal approximation of binomial, you keep the following picture in mind:



7.7. Examples.

1. Suppose that, every time the clown throws a pie, there is a .6 chance that he/she will hit someone. If he/she throws seven pies, what is the probability that he/she will hit no more than one person?

Solution. Our experiment is the clown throwing a pie. Define success to be hitting someone. Let X be the number of times that someone is hit, in seven throws. This is binomial, with $n=7, p=.6$. We want

$$P(X\leq 1)=P(X=0)+P(X=1)=.002+.017=.019,$$

from the binomial table.

2. Suppose that .1 percent of cherry pies contain a cherry pit. If you eat 1,000 pies, what is the probability that exactly one of them will contain a cherry pit?

Solution. The experiment will be to eat a cherry pie. Define success to be getting a cherry pit in that pie. If X is defined to be the number of pies that contain a cherry pit, out of the 1,000 that you eat, then X is binomial, with $n=1{,}000, p=.001$. We want $\sim P(X=1)$.

Clearly we can't use the binomial table. We must choose which approximation to use. Since $np = 1 \le 5$, we use the Poisson approximation, with $\lambda = 1$, to get, from the Poisson tables,

$$P(X = 1) \sim .3679.$$

3. Suppose that 20 percent of dates are no fun. If you go on 49 dates, what is the probability that

a. exactly 12 of them will be no fun;

b. at least 12 of them will be no fun?

Solution. The experiment will be to go on a date. Define success to be having no fun. Let X be the number of dates, out of 49, on which you have no fun. Then X is a binomial random variable, with $n = 49, p = .2$.

We can't use the binomial table, because n is too large. To choose our approximation, look at $np = 9.8 > 5$, also $nq = 49(.8) > 5$; so we must use the normal approximation.

First we'll need $\mu = np = 9.8$; also

$$\sigma = \sqrt{npq} = \sqrt{49(.2)(.8)} = \sqrt{49\frac{16}{100}} = 7(\frac{4}{10}) = 2.8.$$

a.

$$P(X = 12) \sim P(\frac{12 - .5 - 9.8}{2.8} < Z < \frac{12 + .5 - 9.8}{2.8})$$
$$\sim P(.61 < Z < .96) = P(Z > .61) - P(Z \ge .96) = .2709 - .1685.$$

b.

$$P(X \ge 12) \sim P(Z > \frac{12 - .5 - 9.8}{2.8}) \sim P(Z > .61) = .2709.$$

4. Suppose that a medicine is declared to be "effective," if it works at least 95 percent of the time, when tested. Suppose that, in fact, the medicine works 90 percent of the time. An omniscient presence would thus assert that the medicine is ineffective. But I don't know this 90 percent success rate; all I can do is test the success rate. Find the probability that it will be (mistakenly) declared "effective," if it is tested

a. once;

b. 8 times;

c. 64 times;

d. 256 times.

Solution. My experiment will be to give someone this medicine. Success will mean that the medicine works. If I test the medicine n times, then $X \equiv$ "the number of successes" is binomial, with $p = .9$. We want $P(X \ge n(.95))$.

a. Here $n = 1$. $P(X \ge .95) = P(X = 1) = .9$ (from (7.2)).

b. $n = 8$, so we want $P(X \ge 8(.95)) = P(X \ge 7.6) = P(X = 8) = .430$ (from binomial tables).

c. $n = 64$, so we want $P(X \ge 64(.95)) = P(X \ge 60.8) = P(X \ge 61)$. n is too large for binomial tables, $np = 64(.9) = 57.6 > 5, nq = 64(.1) = 6.4 > 5$, so we use a normal approximation.

$\mu = np = 57.6, \sigma = \sqrt{npq} = \sqrt{64(.9)(.1)} = \sqrt{64(\frac{9}{100})} = 8(\frac{3}{10}) = 2.4$, so we have

$$P(X \geq 61) \sim P(Z > \frac{61 - .5 - 57.6}{2.4}) \sim P(Z > 1.21) = .1131.$$

d. $n = 256$, so we want $P(X \geq 256(.95)) = P(X \geq 243.2) = P(X \geq 244)$. As in c, this needs a normal approximation. $\mu = 256(.9) = 230.4$, $\sigma = \sqrt{256(.9)(.1)} = 16(\frac{3}{10}) = 4.8$, so

$$P(X \geq 244) \sim P(Z > \frac{244 - .5 - 230.4}{4.8}) \sim P(Z > 2.73) = .0032.$$

The point being made here is: larger samples are more likely to be accurate.

5. Suppose there are 100 fish in a pond. You want to know how many of them have Fish Ick, a fish disease. To estimate this, you do the following 5 times: grab a fish at random, check to see if it has Ick, then throw it back in.

If in fact 20 of the 100 fish has Ick, what is the probability that more than 2 of the 5 you checked has Ick?

Solution. The experiment here is to grab a fish. We'll call it successful if the fish has Ick. Let X be the number of fish with Ick, from the 5 you grabbed. X is binomial, with $n = 5, p = \frac{20}{100} = .2$. We want

$$P(X > 2) = P(X = 3) + P(X = 4) + P(X = 5) = .051 + .006 + 0 = .057$$

(from the binomial tables).

If, after checking each fish, you had not replaced it in the pond, you could not calculate this probability as a binomial, because p (in this case, the percentage of fish in the pond with Ick) changes each time you remove a fish from the pond.

6. Same as number 5, except that you feed each fish to your cat, after you've checked to see if it (the fish, not the cat) has Ick. HUMANE NOTE: cats can't catch Fish Ick.

Solution. As I commented just before this problem, this is no longer binomial. You must go back to a combinations counting argument, and the relative frequency definition of probability, as in 3.7 and 3.19, to calculate the probabilities.

Our sample space would be sets of 5 fish, chosen from the 100 in the pond; there are $\binom{100}{5}$ such sets. There are $\binom{20}{3}$ sets of 3 fish with Ick chosen from the 20 fish with Ick, thus

$$P(\text{3 fish with Ick}) = \frac{\binom{20}{3}\binom{80}{2}}{\binom{100}{5}}.$$

Similarly,

$$P(\text{4 fish with Ick}) = \frac{\binom{20}{4}\binom{80}{1}}{\binom{100}{5}}, \quad P(\text{5 fish with Ick}) = \frac{\binom{20}{5}\binom{80}{0}}{\binom{100}{5}},$$

so that

$$P(\text{ more than 2 fish with Ick}) = \frac{\binom{20}{3}\binom{80}{2} + \binom{20}{4}\binom{80}{1} + \binom{20}{5}\binom{80}{0}}{\binom{100}{5}}.$$

This is an example of a *hypergeometric* random variable.

7.8. Hypergeometric random variable: there are N things, a of which are good. Choose n of the N things at random, without replacement.

$X \equiv$ number of good things chosen.

7.9. Probability function for hypergeometric:

$$P(X = x) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}.$$

This can be seen by a combinations counting argument, exactly as in the last example.

Note that, if this were with replacement, we would have binomial $b(x, n, p)$, with $p = \frac{a}{N}$.

See 3.20.1 for another example of a hypergeometric problem.

7.10. Multinomial random variable: an experiment, with r possible outcomes, is repeated n times; at each repetition, the probability of the k^{th} outcome is a constant p_k $(1 \le k \le r)$. Then the probability of having x_k of the k^{th} outcome $(1 \le k \le r)$ is

$$\frac{n!}{x_1!x_2! \cdots x_r!} p_1^{x_1} p_2^{x_2} \cdots p_r^{x_r} \quad (x_k \ge 0, x_1 + x_2 + ... + x_r = n).$$

To be more precise, this is actually a multivariable version of what we called a random variable. Note that, when $r = 2$, this is a binomial random variable $(p \equiv p_1, p_2 = 1 - p \equiv q, x \equiv x_1, x_2 = n - x)$.

7.11. Example. Suppose a grocery store owner knows that 50 percent of customers will buy only fruit, 30 percent will buy only nuts, and 20 percent will buy nothing. If 12 people enter his/her store, what is the probability that 6 of them will buy fruit, 4 of them will buy nuts, and 2 of them will buy nothing?

Solution. This is multinomial, with $r = 3, p_1 = .5, p_2 = .3, p_3 = .2, n = 12, x_1 = 6, x_2 = 4, x_3 = 2$. So our probability is

$$\frac{12!}{6!4!2!}(.5)^6(.3)^4(.2)^2.$$

VIII. SAMPLING DISTRIBUTION OF THE MEAN

8.1. Sampling distribution of the mean, or (for short) **Sample mean, $\overline{X}$:** if X is a random variable, then $\overline{X}$ is the average of X over a sample of size n. Note that $\overline{X}$ is another random variable.

Write $\mu_{\overline{X}}$ for the mean of $\overline{X}$, $\sigma^2_{\overline{X}}$ for the variance of $\overline{X}$, also called the *standard error of the mean.*

8.2. Formulas for $\mu_{\overline{X}}, \sigma^2_{\overline{X}}$: Suppose X is a random variable with mean μ, standard deviation σ, and a sample of size n is taken from a population of size N.

(1) *Sampling with replacement:* $\mu_{\overline{X}} = \mu$, $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$.

(2) *Sampling without replacement:* $\mu_{\overline{X}} = \mu$, $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}(\frac{N-n}{N-1})$.
NOTE: for N large, $(\frac{N-n}{N-1}) \sim 1$.

(3) *N unknown or very large:* $\mu_{\overline{X}} = \mu$, $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$.

8.3. Example. Suppose the sugar content of pieces of a particular brand of candy has a mean of 1 gram and a standard deviation of .2 grams. I take samples of 80 pieces of candy. Find the mean and standard deviation of the average sugar content in these samples, if

a. I sample with replacement from a shipment of 1,000 pieces;

b. I sample without replacement from a shipment of 1,000 pieces;

c. I sample from a very large population.

Solution. Let X be the sugar content of a randomly chosen piece of that candy. We have the mean and standard deviation of $X: \mu = 1, \sigma = .2$. We want $\mu_{\overline{X}}$ and $\sigma_{\overline{X}}$, for a sample size $n = 80$.

a. $\mu_{\overline{X}} = 1, \sigma_{\overline{X}} = \frac{.2}{\sqrt{80}}$.

b. $\mu_{\overline{X}} = 1, \sigma_{\overline{X}} = \frac{.2}{\sqrt{80}}\sqrt{\frac{1{,}000-80}{1{,}000-1}} = \frac{.2}{\sqrt{80}}\sqrt{\frac{920}{999}}$.

c. $\mu_{\overline{X}} = 1, \sigma_{\overline{X}} = \frac{.2}{\sqrt{80}}$.

8.4. $\overline{X}$, when X is normal: If X is a normal random variable, with mean μ and standard deviation σ, then $\overline{X}$ is also normal, with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$; that is,

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}, \quad \text{if } X \text{ is normal.} \qquad (8.5)$$

8.6. Example. Suppose that garbage production in each house in a week is normally distributed, with a mean of 1.5 cans and a standard deviation of .4 cans.

a. What is the probability that a randomly chosen house will produce more than 1.6 cans of garbage in a week?

b. What is the probability that 25 randomly chosen houses will produce an average of more than 1.6 cans of garbage in a week?

c. What is the probability that 100 randomly chosen houses will produce an average of more than 1.6 cans of garbage in a week?

Solution. Let X be the garbage production in a randomly chosen house in a week. We are told that X is normal, $\mu = 1.5, \sigma = .4$.

a. We want $P(X > 1.6)$. Use (6.5):

$$P(X > 1.6) = P(\frac{X-\mu}{\sigma} > \frac{1.6-\mu}{\sigma}) = P(Z > \frac{1.6-1.5}{.4}) = P(Z > .25) = .4013.$$

b. We want $P(\overline{X} > 1.6)$, when $n = 25$. Use (8.5):

$$P(\overline{X} > 1.6) = P(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{1.6-\mu}{\frac{\sigma}{\sqrt{n}}}) = P(Z > \frac{1.6-1.5}{\frac{.4}{\sqrt{25}}}) = P(Z > 1.25) = .1056.$$

c. Same as b, except $n = 100$:

$$P(\overline{X} > 1.6) = P(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{1.6-\mu}{\frac{\sigma}{\sqrt{n}}}) = P(Z > \frac{1.6-1.5}{\frac{.4}{\sqrt{100}}}) = P(Z > 2.5) = .0062.$$

Note that

(answer to c) < (answer to b) < (answer to a).

Sampling and taking averages makes strange behaviour less likely. The larger the sample, the less likely strangeness is. This is why large samples are better; in using the sample mean $\overline{X}$ to estimate the population mean μ, you are more likely to be accurate, when you use a larger sample.

8.7. Central limit theorem: If X is any random variable, and n is large enough, then

$$Z \sim \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}, \tag{8.8}$$

where μ is the mean of X, σ is the standard deviation of X, and n is the sample size for $\overline{X}$.

"Large enough" usually means more than 30; in general, the larger n is, the better the approximation.

This enables you to use the Z tables to approximate probabilities involving averages or sums of *any* random variable X.

Note that (8.5) is an equality, for X normal; the central limit theorem says that (8.5) is *approximately* true, for any X.

8.9. Examples.

1. Suppose that the number of headaches you get in a day has a mean of 5 and a standard deviation of 2. What is the probability that, in 64 days, you will have an average of more than 5.5 headaches per day?

Solution. Let X be the number of headaches you get in a day. It has a mean $\mu = 5$, standard deviation $\sigma = 2$. What is being asked for is $P(\overline{X} > 5.5)$, where the sample size n is 64.

We use the central limit theorem:

$$P(\overline{X} > 5.5) = P(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{5.5-\mu}{\frac{\sigma}{\sqrt{n}}})$$

$$\sim P(Z > \frac{5.5-5}{\frac{2}{\sqrt{64}}}) = P(Z > 2) = .0228.$$

2. Suppose that the number of calories I consume in a meal has a mean of 500 and a standard deviation of 120. Suppose further that I will be overweight if I consume a total of more than 47,000 calories in 100 meals. What is the probability that, after 100 meals, I will be overweight?

Solution. Let X be the number of calories that I consume in a meal. X has a mean $\mu = 500, \sigma = 120$. We want $P(\overline{X} > \frac{47{,}000}{100})$. Use the central limit theorem again:

$$P(\overline{X} > 470) = P(\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{470 - \mu}{\frac{\sigma}{\sqrt{n}}})$$
$$\sim P(Z > \frac{470 - 500}{\frac{120}{\sqrt{100}}}) = P(Z > -2.5)$$
$$= 1 - P(Z \leq -2.5) = 1 - P(Z \geq 2.5) = 1 - .0062 = .9938.$$

3. Suppose, in addition to the information in number 2, that the number of calories I consume in a meal is normally distributed. Let's define "gross overeating" for me to be eating more than 800 calories in a meal. What is the probability that I will grossly overeat at least twice, in 100 meals?

Solution. First, let's calculate the probability that I will grossly overeat in any randomly chosen meal. Since X is normal, by (6.5), this is

$$P(X > 800) = P(\frac{X - \mu}{\sigma} > \frac{800 - \mu}{\sigma}) = P(Z > \frac{800 - 500}{120}) = P(Z > 2.5) = .0062.$$

What we've got now is a binomial. Define the experiment to be eating a meal; this experiment is repeated 100 times. Define "success" to be grossly overeating. Let Y be the number of meals, out of those 100, that you overeat. This is binomial, with $n = 100, p = .0062$. We want $P(Y \geq 2)$.

Clearly we can't use the binomial tables. To choose our approximation, look at $np = .62$. Since this is less than or equal to 5, we use the Poisson approximation, with $\lambda = .6$ (to use the Poisson table, it is necessary to round off $np = .62$). Looking at the Poisson table, we now get

$$P(Y \geq 2) = 1 - P(Y < 2) = 1 - (P(Y = 0) + P(Y = 1)) = 1 - (.5488 + .3293).$$

HOMEWORK 0.

1. Simplify each of the following.

$$a.\frac{3/4}{1/8}\quad b.\frac{.001}{.02}\quad c.\sqrt{(.9)(.1)}\quad d.\frac{.21-.19}{.004}\quad e.\sqrt{(400)(.8)(.2)}\quad f.\frac{10/3}{5/9}$$

$$g.\sqrt{36(.9)(.1)}\quad h.\sqrt{(.005)(.5)}\quad i.(3\times 10^{-5})(6\times 10^{2})\quad j.4^3\quad k.19^0\quad \ell.25^1$$

2. Write each of the following as decimals: a. $\frac{1}{3}$ b. $\frac{2}{5}$ c. $\frac{3}{1000}$ d. $\frac{5}{8}$ e. 150 % f. 75 % g. $66\frac{2}{3}$ % h. $19\frac{1}{5}$ %.

3. Write each of the following as fractions: a. .01 b. .025 c. $.\overline{3}$ d. 75 % e. $66\frac{2}{3}$ %

4. Write each of the following as %s: a. $\frac{1}{2}$ b. $\frac{5}{4}$ c. .9 d. .865 e. $\frac{2}{3}$

5. Solve each of the following, for whatever unknowns are sitting around.
a. $1-\frac{1}{k^2}=.99$
b. $\frac{c-5}{3}=2.33$
c. $\frac{.25}{\sqrt{n}}=.001$.
d. $\frac{1}{9}(x+8(11.5))=13$.
e. $a+b=1, a+3b=4$.

Definition For $n=1,2,3,...$, define $n!$, pronounced n *factorial*, to be

$$n!\equiv n(n-1)(n-2)\cdots 2\cdot 1.$$

$0!\equiv 1$.

6. Simplify each of the following: a. 3! b. 6! c. $\frac{7!}{5!}$ d. $\frac{4!2!}{6!}$ e. $\frac{0!6!}{6!}$

HOMEWORK NUMBER ONE

1. Which of the following variables are qualitative and which are quantitative? For those that are quantitative, identify which are continuous and which are discrete.

a. The smell of a plant
b. A person's religion
c. A monthly bill (in dollars)
d. The verdict of a jury
e. The concentration (mg) of cholesterol in your blood
f. The life of an insect (in days)
g. The number of students in the room
h. How a student feels
i. A student's height
j. The number of students that belong to a fraternity
k. A student's salary
l. The speed of a car
m. The length of time one has to wait for a bus to arrive
n. The amount of non-alcoholic beer in a mug
o. The number of redwood trees on an acre of land
p. The prize (in dollars) won in a lottery
q. A person's pulse rate

2. Arrange the following data into a frequency table using the classes 2.00–3.99, 4.00–5.99, 6.00–7.99, and 8.00–9.99.

5.00	6.18	7.63	8.41	6.48	6.89
2.85	8.31	5.81	2.67	9.45	8.88
5.68	7.38	8.69	4.85	2.78	9.99
5.87	3.76	7.68	5.87	5.00	8.76

3. The following table gives the number of ounces per week of hamburger eaten, among fifty teenagers.

Number of ounces	Number of teen-agers
0–9	3
10–19	7
20–29	17
30–39	12
40–49	10
50 and over	1

Answer the following questions, where possible. Find how many teen-agers ate (in a week):

a. less than 20 ounces

b. less than 22 ounces
c. less than 29 ounces
d. more than 20 ounces
e. more than 22 ounces
f. more than 29 ounces.

4. The following table gives the distribution, in thousandths of a millimeter, of the lengths of 100 paramecia.

length	Number of paramecia
1400–1499	8
1500–1599	20
1600–1699	12
1700–1799	35
1800–1899	18
1900–1999	7

Find the following.
a. The class marks
b. The class boundaries
c. The class boundary below which there are 75 percent of the paramecia.
d. The class boundary above which there are 60 percent of the paramecia.

5. Complete the following table by adding in the relative frequency and the percentage frequency.

class	frequency
12.1–16.0	8
16.1–20.0	14
20.1–28.0	20
28.1–32.0	23
32.1–40.0	18
40.1–48.0	7

6. Draw the histogram, cumulative frequency distribution and ogive, for the following frequency distribution of the lives of 400 insects.

Life of insect (in days)	Number of insects
600–699	85
700–799	77
800–899	124
900–999	78
1000–1099	36

7. Here is the cumulative frequency distribution of the distance (in feet) of the jumps of 120 frogs.

distance (in feet)	Cumulative frequency
less than 2.45	0
less than 2.95	6
less than 3.45	26
less than 3.95	49
less than 4.45	74
less than 4.95	102
less than 5.45	120

Find how many frogs jumped:

a. less than 3.95 feet

b. not less than 4.45 feet

c. 3.45 feet or more

d. less than 4.45 feet but 3.45 feet or more.

8. Here is a frequency distribution.

class	frequency
48–50	12
51–53	32
54–56	50
57–59	85
60–62	15
63–65	6

a. Obtain the cumulative frequency distribution.

b. Draw the histogram (of the frequency distribution).

c. Draw the ogive (of the cumulative frequency distribution).

HOMEWORK NUMBER TWO

1. The heights (in inches) of five people were 56, 78, 66, 55, and 80. Find the average height of the people.

2. Suppose there are ten numbers, the lowest is 40 and the highest is 80. Answer T(true) or F(false) to each of the following assertions.

a. The mean is $(80+40)/2$, or 60.
b. The mean is at least 40.
c. The mean is at most 80.

What is the largest that the mean could be? What is the smallest?

3. The mean weight of 15 rats in one cage was 20 pounds, of 8 rats in a second cage was 15 pounds, and of 17 rats in a third cage was 10 pounds. Find the mean weight of all the rats in the three cages.

4.

a. Suppose the mean of the first twenty numbers is 99.8. What would the next number have to be for the mean of the first twenty-one numbers to be 100?
b. Suppose the mean of the first 1000 numbers is 99.8. What would the next number have to be for the mean of the first 1001 numbers to be 100?

5. Suppose you got an average of 86 percent on the first five tests.
a. What should your percent on the next test have to be, to raise your average to 90, if all tests count equally?
b. Same question as (a), except that the next test counts twice as much as each of the first five.

6. The number of crickets on the hearth during nine days were 410, 250, 27, 69, 1, 568, 398, 1000 and 387. Find the median number of crickets.

7. Find the mean, median and mode of the following numbers.

2.4 2.2 2.0 2.3 2.1 2.0 2.1 2.1 2.4 2.4

8. Find the median and mode of the set of data described by the following frequency distribution.

datum	frequency
1	6
2	8
3	1
4	2
5	8
6	2
7	1

9. If the range of scores is 62 and the lowest score is 25, what is the highest score?

10. The following is the rainfall, in inches, each month for six months, on a Caribbean island: 19.8, 18.9, 18.0, 18.8, 18.4, and 19.5.

Find the number of months when the rainfall is within one mean deviation of the mean rainfall.

11. Compute the mean deviation, variance and standard deviation of: 1, 0, 0, −1, 5.

12. I measure some people, and conclude that their average height is 68 inches, with a standard deviation of 12 inches. Later I notice that I had been holding the bottom of my tape measure ten inches above the ground; thus I must add ten inches to everyone's height. What is the mean, standard deviation and variance of the corrected heights?

13. Among thirty loudspeakers, fifteen produced 150 decibels each, ten produced 170 each, and five produced 200 each.

a. Find the mean number of decibels produced.
b. Find the standard deviation.
c. Find the median number of decibels produced.

14. Find the mean and variance of the following data. Use the computing formula for variance.

number	frequency
145	9
150	8
160	11
170	2

15. Find the mean and standard deviation of the following data. Use the computing formula for the standard deviation.

class	frequency
20–24	1
25–29	9
30–34	8
35–39	5
40–44	2
45–49	1

16. Find the mean, median, mode and variance of the following data. Use the computing formula for the variance.

number	frequency
−1	8
0	7
1	3
3	2

HOMEWORK NUMBER THREE

1. If three children are born, draw a sample space that represents all possible arrangements of sons and daughters. If, at each birth, the probability of a son equals the probability of a daughter, find the probabilities of the following events.

a. There are exactly two daughters.
b. There are at least two daughters.
c. There are no more than two daughters.

2. Your small child has a bucket containing plastic letters, one of each letter of the English alphabet (A–Z). If he reaches in and grabs one at random, find the probability that the letter chosen

a. is a vowel
b. is a consonant
c. precedes the letter *p*
d. precedes the letter *p* and is a vowel
e. follows the letter *p* or is a vowel.

3. A barrel contains ten monkeys of which four are known to be hairless. If one monkey is picked at random, find the probability that it is:

a. hairless.
b. hairy (not hairless).

4. Four hundred people participate in a lottery, where each person is given a different number, between 1 and 400. One number is called at random. Find the probability that the number called:

a. is 333
b. has the same three digits
c. begins with 3 and ends in 9.

5. Two die are rolled. Draw the sample space and calculate the probabilities of each of the following events.

a. The second die has a different number than the first die.
b. The same number is on each die.
c. The sum of the numbers on the dice is 7.
d. The number on the first die is twice the number on the second die.
e. The sum of the numbers on the dice is less than 5.
f. The number on the first die is greater than 5.
g. The same odd number is on each die.

6. If $P(A) = 0.3$, $P(B) = 0.5$, and $P(A \text{ and } B) = 0.15$, find $P(A \text{ or } B)$.

7. If $P(A) = 0.6$, $P(\text{not } B) = 0.3$, and $P(A \text{ and } B) = 0.5$, find $P(A \text{ or } B)$.

8. Suppose the probability that an arachnid chosen at random is hairy or gruesome is 0.3. What is the probability that the arachnid is neither hairy nor gruesome?

9. The probability that a grass seed is crabgrass is 0.6; the probability that it will grow is 0.5; and the probability that it is crabgrass or will grow is 0.75. Find the probability that a randomly picked grass seed:

a. is crab grass and will grow
b. is not crab grass nor will it grow
c. either is not crab grass or will not grow.

10. A businessperson is making two investments. If the probability that he/she will make money on one or the other of the investments is 0.65, that he/she will make money on both investments is 0.2, and that he/she will not make money on the first investment is 0.6, find the probability that:

a. he/she will make money on the first investment
b. he/she will make money on the second investment.

11. The probability that I will cough is 0.7, the probability that I will sneeze is 0.6, and the probability that I will cough or sneeze is 0.9. Find the probability that I will:

a. cough and sneeze
b. neither cough nor sneeze
c. either not cough or not sneeze.

12. The probabilities that you will make, when shooting a tiddlywink, 0, 10, 20, 30, or 40 points, are .15, .25, .3, .1, and .2, respectively. Find the probability that, when shooting a tiddlywink:

a. you will make more than 20 points
b. you will make at most 20 points
c. you will make at least 20 points.

13. A six-letter word is formed with the letters A, B, G, I, N and Q with no repetitions. Find the probability that:

a. the word ends in a vowel.
b. the word ends with Q.

14. A six-letter word is formed from A, B, C, D, E, F, G, H, I, J (any letter might be repeated any number of times). Find the probability that:

a. the word ends in a vowel.
b. no letter is repeated.

15. Five mice, call them 1, 2, 3, 4 and 5, are running a maze. If they finish at different times and at random, find the probability that:

a. Mouse 1 will arrive first, Mouse 2 second and Mouse 3 third.
b. Each of Mice 1, 2 and 3 will arrive first, second or third.
c. Mouse 1 will arrive first and Mouse 2 second.
d. Mouse 1 will arrive first and Mouse 2 last.

16. Three of the ten frogs in a shoebox are enchanted royalty. If I pick six of the ten frogs at random, what is the probability that I will pick all three enchanted ones?

17. If a committee of five is selected from a group of eight tall people, three short people, and seven average-height people, find the probability that the committee has:

a. two tall people, two short people, and one average-height person
b. no tall people
c. at least one tall person.

18. A classroom contains nine students, Hilda, Bertha, Mortimer, Ignatz, Eustacias, Ignatius, Stanislav, Hortense and Matilda. If five students are picked at random, find the probability that Ignatz and Hortense are among the five selected.

19. A box of precious stones contains 10 diamonds, 15 rubies and 12 emeralds. I choose 6 stones from the box.

a. What is the probability that my choice consists of 2 diamonds, 2 rubies and 2 emeralds?
b. What is the probability that my choice includes precisely 4 diamonds?
c. What is the probability that my choice includes no diamonds?
d. What is the probability that my choice consists entirely of diamonds?

20. There are 50 books being given away: 10 math books, 20 history books and 20 English books. You choose 5 of the books at random. Find the probability that

a. all of them are math books;
b. one is a math book, 2 are history and 2 are english;
c. exactly one is a math book;
d. none of them are math books;
e. the number of math books equals the number of history books.

HOMEWORK NUMBER FOUR

1. If A and B are two events, find $P(A|B)$ if:

a. $P(B) = 0.4$, $P(A \text{ and } B) = 0.1$
b. $P(\text{not } B) = 0.6$, $P(A \text{ and } B) = 0.2$
c. $P(B) = 0.7$, $P(A \text{ and } B) = 0.4$.

2. Suppose 60 percent of people are heavy drinkers and 10 percent of the heavy drinkers have liver trouble. Find the percentage of people who are heavy drinkers and have liver trouble.

3. The probability that an arachnid is hairy is 0.3. Given that it is hairy, the probability that it is not gruesome is 0.2. Find the probability that a randomly chosen arachnid will be hairy and not gruesome.

4. The probability that it will rain is 0.3. If it rains, the probability is 0.4 that I will stay dry, and if it doesn't rain, the probability is 0.9 that I will stay dry. Find the probability that:

a. it will rain and I will stay dry
b. it will not rain and I will stay dry
c. I will stay dry.

5. The probability that a person picked at random is dysfunctional is 0.05. If two people are picked at random, assuming independence, find the probability that:

a. they are both dysfunctional
b. neither is dysfunctional
c. at least one is dysfunctional.

6. Suppose a student has to reach three different phone numbers in order to register for classes.

If the phone numbers are independent of each other and the probabilities of reaching them are 0.6, 0.4, and 0.2, find the probability that:

a. the student will register
b. the student will not register.

7. Suppose a sleeper is protected by six mosquito nets, operating independently. The probabilities of a mosquito getting through each of the six nets are 0.3, 0.2, 0.2, 0.2, 0.1, and 0.1, respectively.

Find the probability that a randomly chosen mosquito will get through all six nets.

8. Suppose a bridge will fall down if a certain piece is defective. This piece is checked by four people. The first person has a 1 percent chance of not noticing it's defective, when it is;

the second person a 3 percent chance, the third person a .5 percent chance, and the fourth person a 1 percent chance. If their checks are independent, and the piece is defective, what is the probability that none of them will notice it's defective?

9. Suppose you're in space, being bombarded by cosmic rays. You put up shields that operate independently. When a cosmic ray hits one of these shields, it has a 3 percent chance of getting through. How many shields should you put up, if you want fewer than .00001 percent of cosmic rays to get through all the shields?

10. Suppose that, on any day, the probability that you will get a phone call is .4 and the probability that you will get a knock on the front door is .3.

(a) If phone calls and front door knocks are independent, what is the probability that, on any day, you will get both a phone call and a knock on the front door?

(b) If phone calls and front door knocks are independent, what is the probability that, on any day, you will get either a phone call or a knock on the front door?

(c) If phone calls and front door knocks are independent, what is the probability that, on any day, you will get neither a phone call nor a knock on the front door?

(d) If, on any day, the probability that you will get both a phone call and a knock on the front door is .25, what is the probability that, on any day, you will get either a phone call or a knock on the front door?

(e) If, on any day, the probability that you will get both a phone call and a knock on the front door is .25, what is the probability that, on any day, you will get neither a phone call nor a knock on the front door?

11. You are served food that you must eat. The probability that your food is rotten is .05. If your food is rotten, the probability of getting sick is .8. If your food is not rotten, the probability of getting sick is .1.

(a) Find the probability that your food is rotten and you will get sick.

(b) Suppose you get sick. What is the probability that your food was rotten?

12. Suppose that one percent of the population has the strain of virus known as the Scumvirus floating around in their blood. Those with the Scumvirus have a .7 probability of growing warts on their foreheads. Those without the Scumvirus have a .1 chance of growing warts on their foreheads.

(a) Find the probability that you will grow warts on your forehead.

(b) Suppose you grow warts on your forehead. What is the probability that the Scumvirus is in your blood?

13. Suppose the probability that you will study is .7. If you study, the probability of passing is .9. If you don't study, the probability of passing is .2. If you passed, find the probability that you studied.

14. Three barrels of monkeys contain monkeys that can't see, monkeys that can't hear and monkeys that can't talk, as in the table below.

Barrel	can't see	can't hear	can't talk
1	8	2	6
2	4	5	3
3	2	7	3

Assume that each monkey is lacking in one and only one of the skills of seeing, hearing and talking. A barrel is picked at random and a monkey drawn from it at random.

a. Find the probability that the monkey can't talk.
b. Given that the monkey can't talk, what is the probability that it came from barrel 2?

15. Suppose you have three scouts, Rufus, Hieronymus, and Gladys. There is a .1 chance that the scoutmaster will choose Rufus to pitch the tent, a .3 chance that he will choose Hieronymus, and a .6 chance that he will choose Gladys. 20 percent of tents pitched by Rufus fall down, 10 percent of tents pitched by Hieronymus fall down, and 5 percent of tents pitched by Gladys fall down. If the tent falls down, what is the probability it was pitched by Rufus?

16. Suppose your young son has four toys to choose from, call them A, B, C and D. There is a 5 percent chance that he will choose toy A, a 30 percent chance that he will choose toy B, a 25 percent chance that he will choose toy C and a 40 percent chance that he will choose toy D. Toy A breaks the window 50 percent of the times it's used, toy B 20 percent, toy C 20 percent, and toy D 10 percent.

He chooses a toy and uses it.
a. What is the probability that the window will break?
b. If the window breaks, what is the probability that your young son chose toy A?

HOMEWORK NUMBER FIVE

1. Find the expected value and variance of the random variable X with the following probability function.

x	-5	-1	0	4
$p(x)$	0.05	0.15	0.45	0.35

2. A random variable X has the following probability function.

x	-15	5	10	25	30	40
$p(x)$	0.10	0.25	0.10	0.25	0.15	0.15

Find the following.

a. $P(X \geq 25)$
b. $P(5 \leq X < 12)$
c. $P(X < 30)$
d. $P(X \leq 10)$
e. $P(X \geq 30)$
f. $P(X \geq 8)$
g. $P(5 < X < 30)$
h. All numbers c so that $P(X > c) = .30$.
i. All numbers c so that $P(X \leq c) > .45$.

3. Suppose I play a game where I roll a fair die, and get, for any even number, half as many dollars as that number, while for any odd number I lose two dollars. Find the probability function of my winnings. Find my expected winnings.

4. Suppose that a serving of Captain Canine (CC) cereal has 1,000 calories, a serving of Green Forest (GF) cereal has 600 calories and a serving of Wholesome Homeopathy (WH) cereal has 300 calories.

Assume these are my only food possibilities in the morning. On any morning, the probability that I will eat CC is .4, the probability that I will eat GF is .3 and the probability that I will eat WH is .2. Find the expected number of calories that I will consume in a morning.

5. Suppose it costs \$2 for me to find out if it's raining in Fairbanks. I win \$5 if it's raining and lose \$3 if it's not raining. If the probability of rain in Fairbanks is 0.6, find my expected net winnings, including the phone call.

6. Hortense, Bertha, Mortimer and Snerd weigh 200, 170, 150, and 120 pounds, respectively. A person in the group is picked. Hortense, Bertha and Mortimer are all equally likely to be picked; however, Snerd is so obnoxious that he is twice as likely to be picked as any one of the other three. Find the following.

a. The expected weight of the person picked
b. The variance of the weight

7. Suppose the only possible values for a random variable X are -1, 1, and 2. If

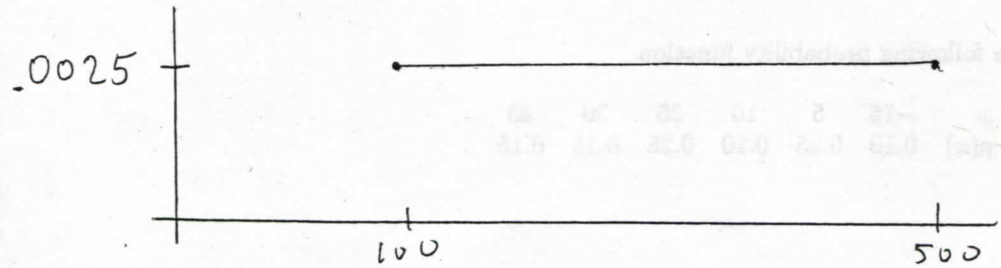
$$P(X = -1) = 0.3$$

and the expected value of X is .5, find the probability function for X.

8. Suppose a random variable X assumes only the values -1, 0, and 2. If the expected value of X is 0, and the variance is 1, find the probability function for X.

9. If X is a continuous random variable with the probability density curve below, find the following.

a. $P(X < 200)$.
b. $P(100 < X < 300)$.
c. $P(X > 200)$.
d. A value c so that $P(200 < X < c) = 0.60$
e. A value c so that $P(X > c) = 0.15$
f. A value c so that $P(X < c) = 0.20$.



10. Consider a continuous random variable with a probability density function that is symmetric about the vertical axis at 0. If $P(-2 < X < 2) = 0.3$ and $P(X > 3) = 0.25$, compute the following probabilities.

a. $P(0 < X < 3)$
b. $P(-3 < X < -2)$
c. $P(-2 < X < 3)$
d. $P(-2 < X < 0)$
e. $P(X < -2)$
f. $P(X < 2)$

11. Consider a continuous random variable X with a probability density function that is symmetric about the vertical axis $x = 50$. Given that $P(X > 60) = 0.1$ and $P(40 < X < 45) = 0.22$, compute the following probabilities.

a. $P(55 < X < 60)$
b. $P(45 < X < 55)$
c. $P(X < 40)$
d. $P(40 < X < 55)$
e. $P(45 < X < 50)$
f. $P(X > 45)$

12. Let X be the random variable in number 11.

a. Find c so that $P(X > c) = .32$.
b. Find c so that $P(50 - c < X < 50 + c) = .8$.
c. Find c so that $P(40 < X < c) = .58$.
d. What is the 90th percentile of X?

13. Suppose my frog's long distance jumps have a mean of 2 feet and a standard deviation of 6 inches. He is six inches from a 3 foot wide stream. What is the probability that, if he jumps perpendicularly to the stream, he will land in the stream?

14. Suppose that the number of days after the end of summer that a mosquito lives in your yard has a mean of 30 days, and a standard deviation of 2 days. If you know there's a mosquito in your yard at the end of summer, how many days should you wait to go into your yard, if you want to be 99 percent certain that the mosquito is dead?

15. Suppose some data have a mean of 18 and a standard deviation of 4.
(a) What fraction of measurements will lie between 6 and 30?

(b) What fraction of measurements will lie between 0 and 36?
(c) Find an interval that will contain at least 75 percent of the data.
(d) Find an interval that will contain at least 95 percent of the data.

16. Suppose some data has a mean of 5. Determine how small the standard deviation should be, so that 99 percent of the data will lie between 4 and 6.

17. Suppose that the SAT of high school graduates has a mean of 500 and a standard deviation of 30. A college wants to choose a minimum SAT score for students entering; that is, a student enters if and only if their SAT score is above this minimum.

Find what this minimum score should be, so that at least 75 percent of students will enter.

18. Suppose that thirty percent of the population is lying. A lie detector test is given to randomly chosen people. A liar has a sixty percent chance of failing the test, and a non-liar has a twenty percent chance of failing.

If someone fails the test, what is the probability that he/she is lying?

19. Suppose the height a randomly chosen cow can jump has a mean of 4 inches and a standard deviation of .6 inches. How high should I build a fence, so that fewer than 4% of cows can jump over it?

HOMEWORK NUMBER SIX

1. If Z is a standard normal variable, find the following probabilities.

a. $P(0 \leq Z < 1.5)$
b. $P(Z \geq 1.5)$
c. $P(Z > -1.5)$
d. $P(1 < Z < 1.5)$
e. $P(-1 < Z < 1.5)$

2. Suppose Z is a standard normal variable. Find the approximate value of c so that:

a. $P(0 < Z \leq c) = 0.2$.
b. $P(Z \leq c) = 0.87$
c. $P(Z \leq c) = .1$
d. $P(c \leq Z < 0) = .3$.
e. $P(Z \geq c) = .1$.
f. $P(Z \geq c) = .8$.
g. $P(-2 \leq Z \leq c) = 0.055$
h. $P(-2 \leq Z \leq c) = 0.49$
i. $P(-2 \leq Z \leq c) = 0.9$
j. $P(Z < \frac{c-5}{7}) = .8$.
k. $P(Z < \frac{c+4}{5}) = .1$
l. $P(Z > \frac{c-1}{.2}) = .99$.

3. The diameter of the pegs, in a child's toy work bench, has a normal distribution with a mean diameter equal to 2 inches and a standard deviation equal to 0.05 inches. Find approximately the diameter of a circular hole through which at least 97 percent of the pegs will pass.

4. The chocolate content in a certain type of candy has a normal distribution with a mean content of 1.8 grams and a standard deviation of 0.2 grams. Find the probability that the chocolate content of a randomly picked piece of this type of candy will be:

a. less than 1.45 grams
b. between 1.45 and 1.65 grams
c. between 1.95 and 2.15 grams
d. more than 2.15 grams.

5.

(a) The demand for non-alcoholic beer at a store, during any week, has a mean of 5000 quarts and a standard deviation of 300 quarts. How much non-alcoholic beer should the store have in stock per week so that they will run out of non-alcoholic beer no more than 1 percent of the time?

(b) Same question as (a), except assume the demand is normally distributed.

6. Suppose the amount of water consumed in a certain house on any day is normally distributed with a mean of 5000 liters and a standard deviation of 300 liters.

a. If the house has 5300 liters of water on a given day, what is the probability that it has more than enough?
b. If the house has 5300 liters of water on each of two days, what is the approximate probability that it has more than enough both days?
c. How much water should the house have ready for consumption per day so as not to run short more than 10 percent of the time?

7. Do problems 13 and 14, of homework five, under the additional assumption that the random variables are normally distributed.

8. Suppose the mean annual income of a worker is 23,000 dollars, with a standard deviation of 1,400 dollars. I define poverty in terms of income, that is, a worker is defined to be in poverty if and only if his annual income is below a certain maximum. If I want to define things so that no more than four percent of workers are in poverty, what should this maximum annual income be?

9. Same as number 8, except now assume that annual income is normally distributed.

10. With income as in number 9, what is the probability that two workers picked at random will each earn less than 22,500?

11. Suppose that sixty percent of the population eats well. Further suppose that those who eat well have a seventy percent chance of longevity, while those who don't eat well have a twenty percent chance of longevity. If someone has longevity, what is the probability that he/she eats well?

12. Suppose human weight is normally distributed, with a mean of 150 pounds and a standard deviation of 30 pounds. What is the 95th percentile of human weight?

HOMEWORK NUMBER SEVEN

On all binomial problems, approximate with an appropriate table.

1. A military dictator is concerned about his food being poisoned. He has eight food testers that operate independently. The probability that a food taster will get sick when tasting poisoned food is 0.90. Find the probability that poisoned food will:

a. not make any of the tasters sick
b. make at most four tasters sick.

2. When AJAX seat covers are built, the probability that they can survive pizza stains is 0.8. If eight AJAX seat covers are stained with pizza, find the probability that:

a. none will survive
b. exactly four will survive
c. all will survive.

3. Suppose that, if a certain coin is tossed ten times, the probability of having four heads is 0.111. What is the probability of having five heads in eight tosses?

4. Suppose ten percent of the corn crop fails to grow after being planted. If 100 corn seeds are planted, what is the expected number of successful corn plants? What is the variance of the number of corn plants grown?

5. Suppose that 80 percent of the animals you catch on a fishing line are fish. What is the probability that, when you catch 100 animals on this fishing line, the number of fish you catch will be:

a. less than seventy
b. greater than eighty-six
c. between seventy-six and eighty-six, inclusive?

6. Suppose the media asserts that 90 percent of people watch TV. This assertion will be tested by interviewing 100 people; if 82 or more of those 100 watch TV, the assertion will be declared true. Find the probability that the assertion will be declared false when it is actually true.

7. Suppose 10 percent of students want to wear uniforms to class. Find the probability that on a committee of 64 students selected at random, exactly 6 will favor wearing uniforms to class.

8. A photo-copy shop tests its machines by running 400 papers and checking to see if they copy correctly. If more than 225 are perfect, the machine is declared OK. What is the probability that a machine will be declared OK, when actually only half of papers run through it are perfect?

9. The number of flies caught in a fly trap in an hour has a Poisson distribution with a mean of 2 flies. Find the probability that in an hour there will be:

a. no flies caught
b. at least 4 flies caught.

10. Suppose the number of children showing up at an institution between 9:00 and 9:30 AM has a Poisson distribution with a mean number of children equal to 2.5. Find the probability that between 9:00 and 9:30 AM there will be:

a. no children showing up
b. between 2 and 6 children, inclusive, showing up.

11. Suppose that only .6 percent of train trips include a derailment. If I take 400 train trips, what is the probability that at most four of those trips will include a derailment?

12. Suppose the probability that a drug will have side effects, when taken, is .004. If I take it 350 times, what is the probability that I will have no side effects?

13. Suppose the amount of non-alcoholic beer poured into your mug is normally distributed, with a mean of one pint and a standard deviation of .5 cups. How large a mug should you have, so that the non-alcoholic beer poured into it will overflow no more than two percent of the time?

14. Suppose there are ten monkeys in a barrel, three of whom are bald. Five times, I choose a monkey at random, make note of whether or not he is hairy, then throw him (gently) back. What is the probability that I will choose a bald monkey precisely twice?

15. Suppose there are ten monkeys in a barrel, three of whom are bald. I pull out five of them. What is the probability that precisely two of them are bald?

16. There are nine students in a room, four of whom are nervous. If I choose three different students, what is the probability that at least one of them is nervous?

17. Suppose 20 percent of the population is Monarchist. If I choose 8 people at random, what is the probability that at least two of them are Monarchist?

18. Suppose 20 percent of the fifty people in a room is Monarchist. If I choose 8 of these people at random, what is the probability that at least two of them are Monarchist?

19. Suppose 20 percent of Amazonian frogs are green, 50 percent are red, 10 percent are blue and 20 percent are yellow. If I choose 10 Amazonian frogs at random, what is the probability that

(a) 1 of them is green, 5 of them are red, 2 are blue and 2 are yellow;

(b) none of them are green;

(c) 5 of them are green;

(d) all of them are green or red;

(e) all of them are green, red or blue.

20. Suppose 20 percent of 20 Amazonian frogs in a bucket are green, 50 percent are red, 10 percent are blue and 20 percent are yellow. If I choose 10 (different) Amazonian frogs at random from the bucket, what is the probability that

(a) 1 of them is green, 5 of them are red, 2 are blue and 2 are yellow;

(b) none of them are green;

(c) 5 of them are green;

(d) all of them are green or red;

(e) all of them are green, red or blue.

21. Suppose there are 50 cookies in a box: 10 macaroons, 20 oatmeals, 12 snickerdoodles, 8 chocolate chips. You choose 16 of them at random, for your guests to consume. Find the probability that

(a) exactly 10 of the cookies you chose were snickerdoodles and chocolate chips;

(b) you chose no macaroons;

(c) you chose only macaroons;

(d) you chose only macaroons and oatmeals;

(e) you chose exactly 4 macaroons.

22. Suppose that 20 percent of cookies are macaroons, 50 percent are oatmeals, 25 percent are snickerdoodles and 5 percent are chocolate chips. You choose 9 of them at random, for your guests. Find the probability that

(a) exactly 5 of the cookies you chose were oatmeals and chocolate chips, and 2 of them were snickerdoodles;

(b) you chose no macaroons, 5 oatmeals, and 3 snickerdoodles;

(c) you chose no macaroons;

(d) you chose only macaroons and oatmeals;

(e) you chose exactly 4 macaroons.

23. Suppose that 95 percent of people with an internal infection have high blood pressure, 10 percent of people without an internal infection have high blood pressure, and 20 percent of the population have an internal infection. If someone has high blood pressure, what is the probability that he/she has an internal infection?

24. Suppose that the amount of coffee poured by a waiter has a mean of .8 cups and a standard deviation of .3 cups. How large a mug should you use to receive coffee, if you want the coffee to overflow no more than 1 percent of the time?

25. Suppose the probability is .8 that a person picked at random will support a constitutional amendment requiring an annual balanced budget. If 9 individuals are interviewed and they respond independently, what is the probability that at least 7 of them will support the amendment?

26. Suppose the amount of non-alcoholic beer poured into your mug is normally distributed, with a mean of one pint and a standard deviation of .5 cups. If your mug holds 3.2 cups, what is the probability that non-alcoholic beer poured into your mug will overflow?

27. Suppose the amount of non-alcoholic beer poured into your mug is normally distributed, with a mean of one pint and a standard deviation of .5 cups. If non-alcoholic beer is poured into 100 mugs, each of which holds 3.2 cups, what is the probability that more than one mug will overflow?

HOMEWORK NUMBER EIGHT

On all binomial problems, approximate with an appropriate table. On all sample mean problems, approximate if necessary with an appropriate table.

1. Suppose twelve items are picked at random from a population having eighty items. If the population variance of a variable X is 60, find the variance of the distribution of $\overline{X}$ if the sampling is:

a. with replacement

b. without replacement.

2. A population consisting of 100 items has a mean of 45 and a standard deviation of 5. If a sample of 16 items is drawn without replacement from this population, what is the mean and the standard deviation of the distribution of the sample mean?

3. Suppose the protein content of fish has a distribution with a mean of 2 grams and a standard deviation of .3 grams. If a sample of sixty-four fish is picked at random, without replacement, find the standard error of the mean protein content if the population:

a. consists of 400 fish

b. can be considered extremely large.

4. For fish as in number 3, what is the probability that a sample of 900 fish there will be a mean protein content exceeding 2.02 grams?

5. Suppose I sell "foot-long hot dogs"; their length is normally distributed with a mean length of 12 inches and a standard deviation of .2 inch.

(a) What is the probability that a randomly chosen foot-long hot dog is less than 11.97 inches or greater than 12.04 inches?

(b) A shipment of 100 hot dogs will be considered unfit for consumption if the mean length is less than 11.97 inches or greater than 12.04 inches. Find the probability that the shipment is fit for consumption.

6. Suppose the number of zithers per week produced by a randomly chosen zither manufacturer has a mean of 100 zithers with a standard deviation of 8 zithers. If 64 of the zither manufacturers are picked at random, find the probability that the mean number of zithers produced in a week will:

a. exceed 102.5 zithers

b. be less than 98.5 zithers

c. be between 98 and 102 zithers.

7. A truck can carry at most 36,450 pounds of dogs without suffering. If it's carrying 900 dogs and the distribution of the weight of a dog has a mean of 40 pounds and a standard deviation of 6 pounds, find the probability that the truck will suffer.

8. Suppose that the amount of coffee drunk by a randomly chosen person in a day has a distribution with a mean of one pint, with a standard deviation of .4 cups.

(a) Find the probability that a randomly chosen person will drink more than 3 cups or less than 1 cup in a day.

(b) Find the probability that the average amount of coffee drunk by 64 people in a day will be more than 2.1 cups or less than 1.9 cups.

(c) If 64 people are at a restaurant in a day, and the restaurant has only 68 pints, what is the probability that there will not be enough coffee?

9. In cockroach extermination, the cumulative amount of poison picked up by the cockroach is what kills the cockroach. Suppose some poison is laid down, at a spot that a particular cockroach walks over 400 times in a week. Suppose the amount of poison that he picks up each time he walks over this spot is a distribution with a mean of .1 mg and a standard deviation of .05 mg. If it takes an accumulation of 38 mg of poison to kill the cockroach, what is the probability that he will die in a week?

10. Suppose that the amount of sunlight a person absorbs on the beach in an afternoon of sunbathing has a mean of ten skin cancers (this unit is not real) and a standard deviation of two skin cancers. If 36 people are sunbathing on the beach, find the probability that the total amount of skin cancers absorbed by these people in an afternoon will exceed 369 skin cancers.

11. Now suppose that the amount of sunlight absorbed, from the previous problem, is normally distributed. What is the probability that a randomly chosen individual will absorb at least 10.4 skin cancers (in scientific terminology, this is "medium to well-done")?

12. Suppose 60 percent of students want to wear uniforms to class. Find the probability that on a committee of ten students selected at random:

a. all will favor wearing uniforms to class.
b. precisely half will favor wearing uniforms to class.

13. The probability that a student will pass is 5/8. If he passes, he will go home for the weekend with a probability of 5/6; however, if he fails, there is a probability of 1/3 that he will go home for the weekend.

a. Find the probability that the student will go home for the weekend.
b. Given that he does go home for the weekend, find the probability that he passed.

14. A photo-copy shop tests its machines by running ten papers and checking to see if they copy correctly. If fewer than three are imperfect, the machine is declared OK. What is

the probability that a machine will be declared OK, when actually only half of papers run through it are perfect?

15. When AJAX seat covers are built, the probability that they can survive pizza stains is 0.98. If 100 AJAX seat covers are stained with pizza, find the probability that all of them will survive.

16. The diameter of a lead shot produced by a machine has a mean of 2 inches and a standard deviation of .05 inches. Find what the diameter of a circular hole should be, so that no more than 3 percent of the lead shots will pass through.

17. Suppose the probability is .8 that a person picked at random will support a constitutional amendment requiring an annual balanced budget. If 900 individuals are interviewed and they respond independently, what is the probability that at least 700 of them will support the amendment?

18. Suppose 3 out of the 10 AJAX seat covers in a store have pizza stains. If I choose 6 of them at random, what is the probability that at least 2 will have pizza stains?

19. For dogs as in problem 5, suppose now that a shipment of 100 hot dogs is considered unfit if it contains more than two that are less than 11.6 inches. Find the probability that a shipment of 100 will be considered fit.

20. For dogs as in problem 7, suppose now that a truck can carry at most one dog that weighs over 55 pounds without suffering, and dog weight is normally distributed. If it's carrying 400 dogs, find the probability that the truck will suffer.

21. Suppose a random variable X has a standard deviation of 5 and an unknown mean μ. I will use the sample mean, $\overline{X}$, to estimate μ. If my sample size is 6400, what is the probability that $\overline{X}$ is within .1 of μ?

22. Suppose X has a standard deviation of 2 and an unknown mean μ. If, in a sample of 100, $\overline{X}$ is 10, what is the probability that μ is greater than 10.5?

23. Suppose X has a standard deviation of 2 and an unknown mean μ. If, in a sample of 100, $\overline{X}$ is 10, find an interval $[a, b]$ that contains μ with a probability of 99 percent.

24. In problem 8, if 64 people eat at a restaurant in a day, how much coffee should the restaurant have available each day, if they want to have enough coffee 99.5 percent of the days?

25. Suppose the vertical dimension of a brick is normally distributed, with a mean of 3 inches and a standard deviation of 1 inch. If a building is built in such a way that the height is 1600 bricks, how high should a telephone wire be so that the probability that the building will hit the telephone wire is less than .2 percent?

26. With sunlight as in problems 10 and 11, define a person to be "half-baked" if they've absorbed more than ten skin cancers in their afternoon of sunbathing. What is the probability that more than 12 people, out of those 36 sunbathers, will be half-baked?

HOMEWORK NUMBER NINE

On all binomial and sample mean problems, approximate with an appropriate table.

1. Suppose that the length your hair grows in a day has a mean of .1 inch and a standard deviation of .04 inches. What is the probability that, in 100 days, your hair will grow more than 9 inches?

2. Find the mean, median and mode of the following: $5, 3, 3, 1, -1, 1, 2, 5$.

3. A barrel contains ten wolverines, four of which are rabid. If I pull out three of them (without replacement), what is the probability that none of them are rabid?

4. Suppose 25 percent of unhealthy people have bad diets, 5 percent of healthy people have bad diets, and 10 percent of people are unhealthy.

   (a) What is the probability that a randomly chosen person has a bad diet?

   (b) If you have a bad diet, what is the probability that you are healthy?

5. Construct a frequency table with classes 0–9, 10–19, 20–29 and 30–39, a histogram, a cumulative frequency table, and an ogive, for the following data:

$$9, 11, 8, 1, 37, 30, 32, 31, 32, 8.$$

6. Suppose the diameter of the human head has a mean of six inches and a standard deviation of .3 inches.

   (a) Determine how large the diameter of a hat should be, so that at least 99 percent of human heads will fit inside it, without stretching the hat.

   (b) Same as a, except assume human head diameter is normally distributed.

7. The probability that you will trip while walking to class is .1, the probability that you will fall is .08, and the probability that you will neither trip nor fall is .85. What is the probability that you will both trip and fall?

8. How many possible license plates are there, consisting of three letters (A through Z), followed by four numbers (0 through 9)? Both letters and numbers could appear more than once.

9. Suppose that, if a dangerous person shows up, there's an 80 percent chance that my dogs will bark. On the other hand, when someone who is not dangerous shows up, there's a 60 percent chance that my dogs will bark. Suppose that ten percent of people who show up are dangerous. If someone shows up and my dogs bark, what is the probability that the person who showed up is dangerous?

10. Suppose you have a ten percent chance of sweating 2 quarts, a fifty percent chance of sweating 1 quart, and a forty percent chance of not sweating at all. What is your expected amount of sweating?

11. Suppose that, every time I teach, there is a .03 chance that I will run out of chalk. If I give forty lectures, what is the probability that I will run out of chalk at least once?

12. Suppose that ten percent of students get As. In a class of 400 students, what is the probability that more than fifty students will get As?

13. Suppose human weight has a mean of 168 pounds and a standard deviation of 20 pounds. An elevator will collapse if it carries more than $17,000$ pounds. If this elevator carries 100 people, what is the probability that it will collapse?

14. Suppose the number of clams you will eat at a meal has the following probability function.

x	0	1	2	3	4	5	6
$p(x)$	.2	.3	.1	.1	.15	.1	.05

(a) What is the probability that you will eat no more than one clam at a meal?

(b) What is the expected number of clams that you will eat at a meal?

(c) What is the standard deviation of the number of clams that you will eat at a meal?

15. Suppose that ten percent of wolverines are rabid. If you bring 100 wolverines to this country, what is the probability that more than 15 of them will be rabid?

16. Suppose that, every time you walk to class, there is a .01 probability that you will forget your pencil. If you walk to class 100 times, what is the probability that you will forget your pencil on exactly three of those trips?

17. Suppose that the amount of hamburger a student consumes at a party has a distribution with a mean of 2 pounds and a standard deviation of 1.2 pounds. If there are 36 students at a party, and 90 pounds are available, what is the probability that there will not be enough hamburger?

18. Find the mean, median, mode and variance of the data in the following frequency distribution.

number	frequency
−1	10
0	5
1	2
2	3

19. Suppose you have a 70 percent chance of passing the first test, a 60 percent chance of passing the second test, and a 40 percent chance of passing the third test.

(a) If your scores on the three tests are independent of each other, what is the probability of passing all three tests?

(b) If you are certain to pass either the first or second test, what is the probability that you will pass both the first and second test? (Do not assume independence.)

20. Suppose that if you have malaria there is a .95 chance you will tremble. On the other hand, 10 percent of people who don't have malaria tremble. Suppose one percent of the population has malaria.

(a) What percent of the population trembles?

(b) If I am trembling, what is the probability that I have malaria?

21. Suppose an elevator can carry only one person. There is a ten percent chance that it will carry a 100 pound person, a fifty percent chance that it will carry a 120 pound person, a twenty percent chance that it will carry a 180 pound person, a ten percent chance that it will carry a 250 pound person, and a ten percent chance that it will carry no one. Find the expected weight that it will carry, and the variance of the weight that it will carry.

22. Suppose that, every time you turn on your TV, there is a thirty percent chance that it will be during a commercial. If you turn on your TV five times at random, what is the probability that at least two of those times will be during a commercial?

23. Suppose that the number of decibels required to wake up a sleeping person is normally distributed, with a mean of 60 decibels and a standard deviation of 1.5 decibels. Suppose it's Halloween, so you naturally want to wake up 99 percent of the sleeping population. How many decibels of noise should you make?

24. Suppose that annual snowfall in Athens is a random variable with a mean of 20 inches, and a standard deviation of 10 inches.

(a) Find the probability that a total of at least 2,200 inches will fall in 100 years.

(b) Find the probability that the average annual snowfall over 100 years will be less than 21 inches.

(c) If annual snowfall is normally distributed, find the probability that annual snowfall will be less than 21 inches.

25. Suppose you have an average time of 11 seconds, on your first five runs of the hundred-yard dash. What should your time be on the sixth run, if you want to have an average time of 10.8 seconds on your first six runs of the hundred-yard dash?

26. Suppose that the number of decibels required to wake up a sleeping person has a mean of 60 decibels and a standard deviation of 1.5 decibels. Suppose it's Halloween, so you naturally want to wake up 99 percent of the sleeping population. Determine how many decibels of noise you should make.

27. Suppose that 90 percent of students who study are happy, 15 percent of students who don't study are happy, and 60 percent of students study.

(a) What percent of students are happy?

(b) If someone is happy, what is the probability that he/she studied?

28. Suppose that ten percent of apples are mealy. If you choose eight apples at random, what is the probability that no more than two of them are mealy?

29. Suppose that, every time you pick up your pencil, there is a .00001 chance that it will explode. If you pick it up 20,000 times, what is the probability that it will explode more than once?

30. Suppose that daily rainfall is a random variable with a mean of .1 inches and a standard deviation of .05 inches. You leave a cylindrical glass 11 inches tall outdoors in the rain for 100 days. What is the probability that it will overflow with water, if the only source of water is rain and no evaporation occurs?

31. Suppose I give 10 percent As, 20 percent Bs, 40 percent Cs, 20 percent Ds and 10 percent Fs. In a class of 10 students, what is the probability that there will be

(a) 1 A;

(b) 1 A, 2 Bs, 4 Cs, 2 Ds, and 1F;

(c) 5 As and 5 Bs;

(d) 5 As and Bs.

32. Suppose, in a class of 50 students, there are 10 percent As, 20 percent Bs, 40 percent Cs, 20 percent Ds and 10 percent Fs. If I choose 10 of those students at random, what is the probability that, among them, there will be

(a) 1 A;

(b) 1 A, 2 Bs, 4 Cs, 2 Ds, and 1F;

(c) 5 As and 5 Bs;

(d) 5 As and Bs.

33. Suppose the amount of garbage produced by a dorm in a day has a mean of 100 quarts and a standard deviation of 8 quarts. If the dorm has only one garbage can, and all garbage goes into this garbage can, what size should the garbage can be, if you want it to overflow no more than 4% of days?

34. Suppose the height above the ground of my eyes, when I jump, has a mean of 90 inches and a standard deviation of 4 inches. The bottom of a window is 80 inches above the

ground and the top of the window is 100 inches above the ground. If I jump from the ground directly below the window, what is the probability that I will be able to look through the window at the top of my jump?

35. Same as number 34, except now assume the height of your jumps is normally distributed.

ANSWERS: HOMEWORK NUMBER ZERO

1. a. 6 b. $\frac{1}{20}$ c. .3 d. 5 e. 8
f. 6 g. 1.8 h. .05 i. 18×10^{-3} (or 1.8×10^{-2}; that's called scientific notation)
j. 64 k. 1 ℓ. 25

2. a. $.\overline{3}$ b. .4 c. .003 d. .625 e. 1.5
f. .75 g. $.\overline{6}$ h. .192

3. a. $\frac{1}{100}$ b. $\frac{1}{40}$ c. $\frac{1}{3}$ d. $\frac{3}{4}$ e. $\frac{2}{3}$

4. a. 50 % b. 125 % c. 90 % d. 86.5 % e. $66\frac{2}{3}$ %

5. a. $k = 10$ b. $c = 11.99$ c. $n = 62500$ d. $x = 25$ e. $b = \frac{3}{2}, a = -\frac{1}{2}$

6. a. 6 b. 720 c. 42 (equals 7×6) d. $\frac{1}{15}$ (cancel $\frac{4!}{6!} = \frac{1}{6 \times 5}$) e. 1

ANSWERS: HOMEWORK NUMBER ONE

1.

a. The smell of a plant: qualitative
b. A person's religion: qualitative
c. A monthly bill (in dollars): quantitative, discrete
d. The verdict of a jury: qualitative
e. The concentration (mg) of cholesterol in your blood: quantitative, continuous
f. The life of an insect (in days): quantitative, continuous
g. The number of students in the room: quantitative, discrete
h. How a student feels: qualitative
i. A student's height: quantitative, continuous
j. The number of students that belong to a fraternity: quantitative, discrete
k. A student's salary: quantitative, discrete
l. The speed of a car: quantitative, continuous
m. The length of time one has to wait for a bus to arrive: quantitative, continuous
n. The amount of non-alcoholic beer in a mug: quantitative, continuous
o. The number of redwood trees on an acre of land: quantitative, discrete
p. The prize (in dollars) won in a lottery: quantitative, discrete
q. A person's pulse rate: quantitative, discrete.

2.

class	frequency
2.00–3.99	4
4.00–5.99	7
6.00–7.99	6
8.00–9.99	7

3.

a. 10.
b. impossible to tell
c. impossible to tell, because "less than 29" does not include 29, and we don't know how many of the 17 values between 20 and 29 are equal to 29
d. impossible to tell, since we don't know how many of the 17 values between 20 and 29 are equal to 20
e. impossible to tell
f. 23

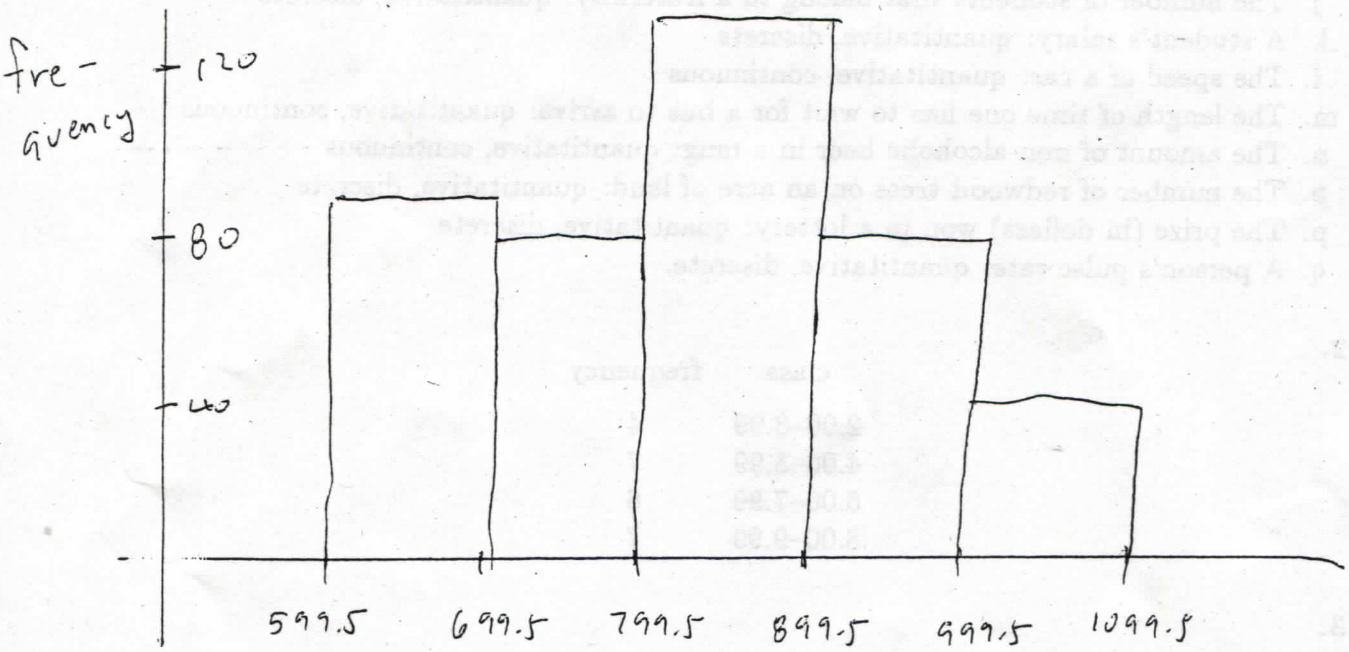
4.

a. 1449.5, 1549.5, 1649.5, 1749.5, 1849.5, 1949.5
b. 1399.5, 1499.5, 1599.5, 1699.5, 1799.5, 1899.5, 1999.5.
c. 1799.5.
d. 1699.5.

5.

class	frequency	relative frequency	percentage frequency
12.1–16.0	8	.089	8.9
16.1–20.0	14	.156	15.6
20.1–28.0	20	.222	22.2
28.1–32.0	23	.256	25.6
32.1–40.0	18	.2	20
40.1–48.0	7	.078	7.8
sum	90	1	100

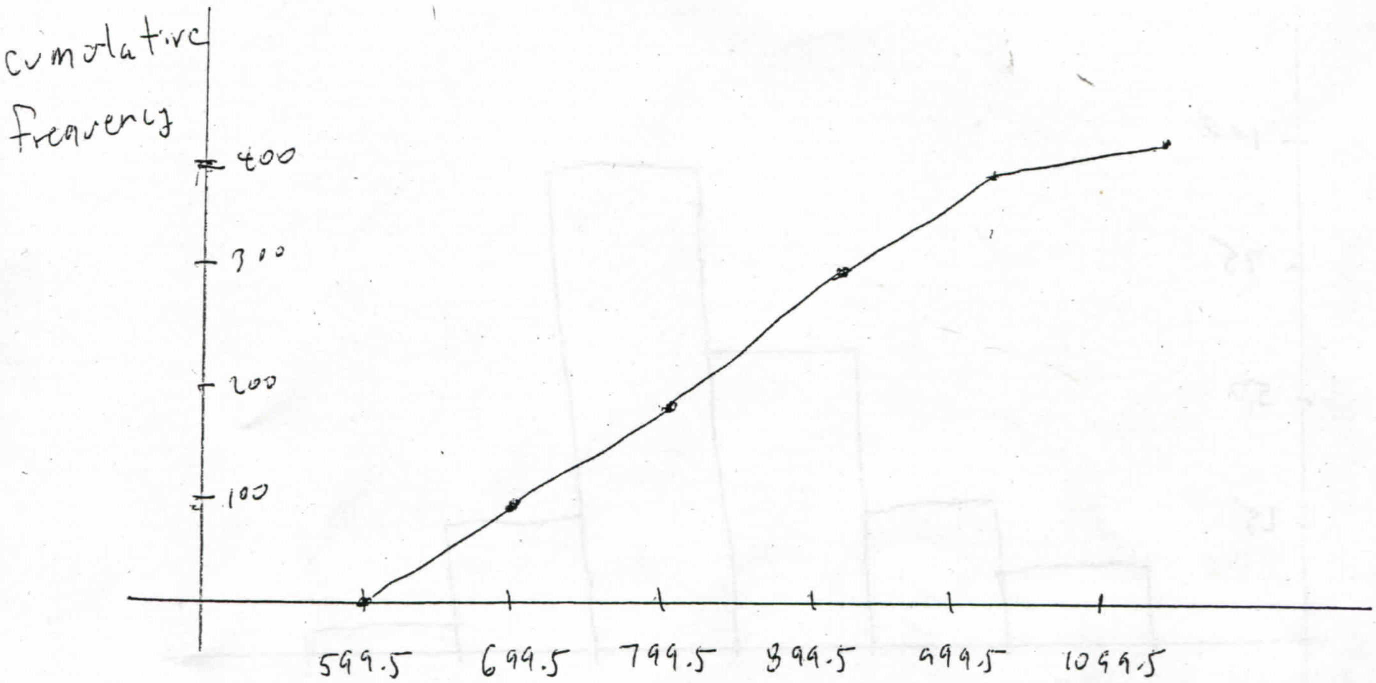
6. The class boundaries are 599.5, 699.5, 799.5, 899.5, 999.5 and 1099.5.
histogram:



cumulative frequency distribution:

Life of insect (in days)	Cumulative number of insects
< 599.5	0
< 699.5	85
< 799.5	162
< 899.5	286
< 999.5	364
< 1099.5	400

ogive:



7.

a. 49.

b. 46.

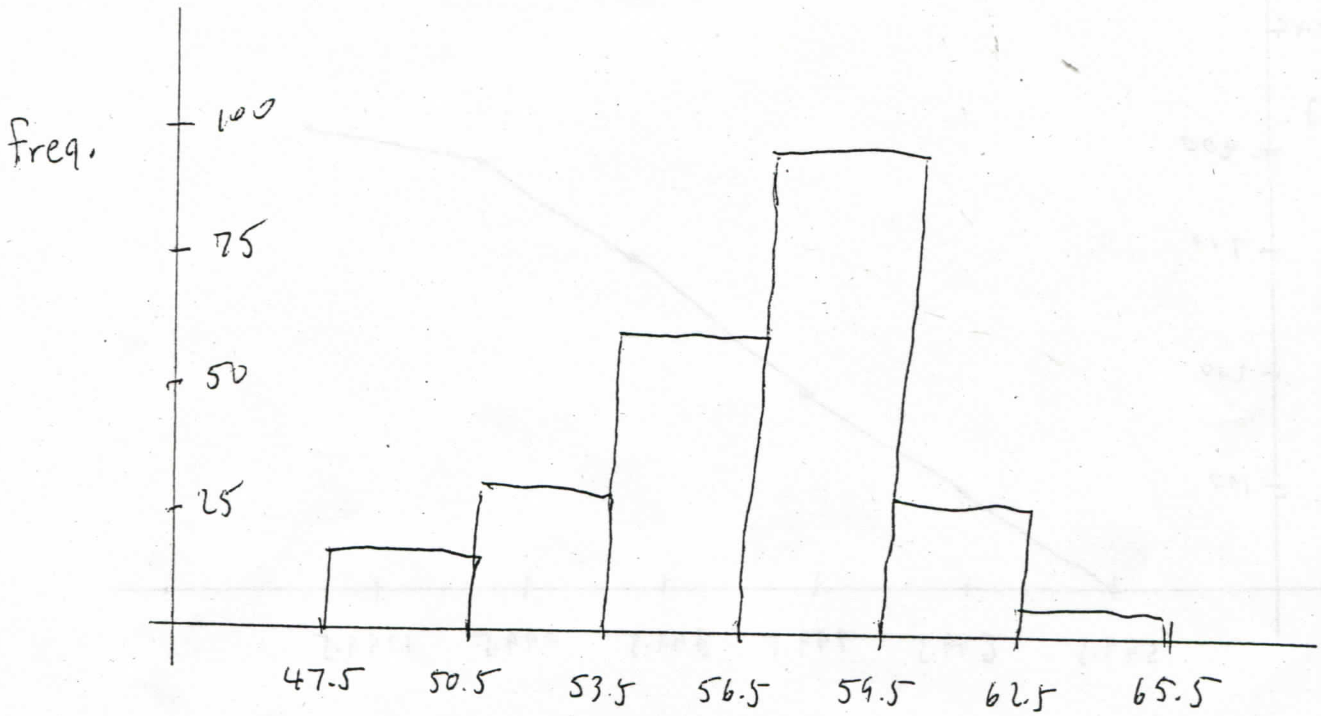
c. 94.

d. 48.

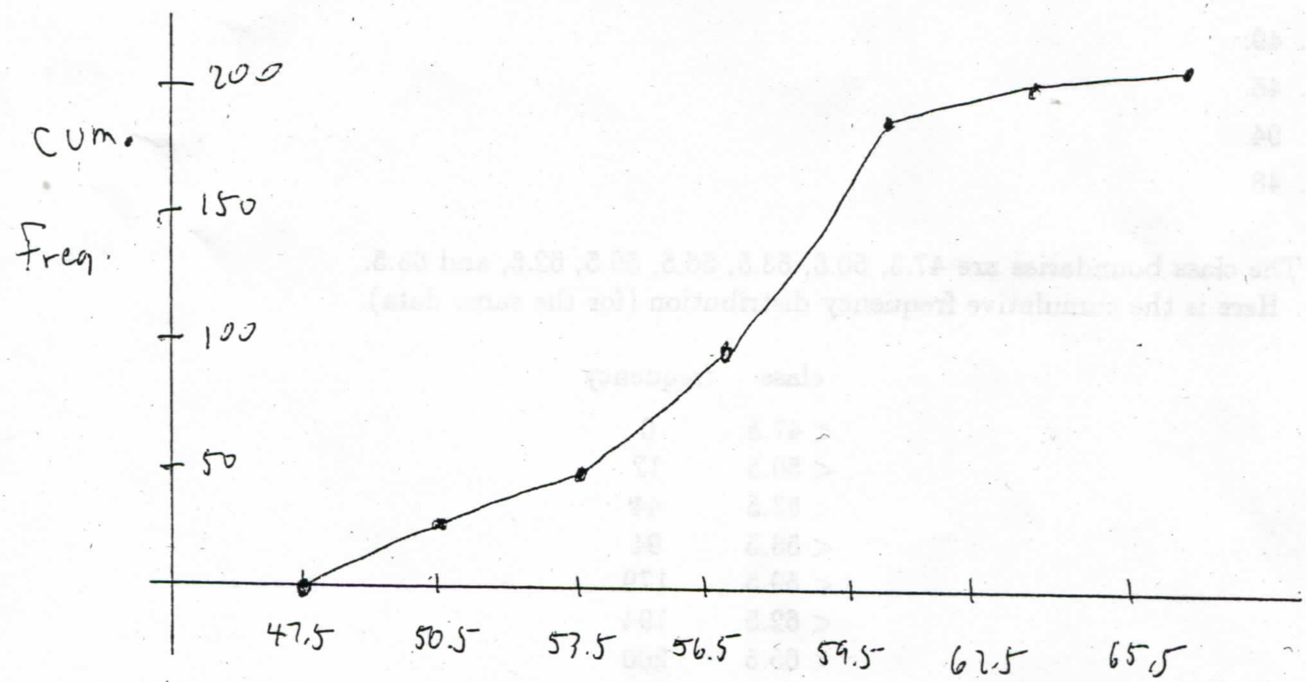
8. The class boundaries are 47.5, 50.5, 53.5, 56.5, 59.5, 62.5, and 65.5.
(a). Here is the cumulative frequency distribution (for the same data).

class	frequency
< 47.5	0
< 50.5	12
< 53.5	44
< 56.5	94
< 59.5	179
< 62.5	194
< 65.5	200

histogram:



ogive:



NOTE: the ogive is for the *cumulative* frequency distribution and the histogram is for the frequency distribution

ANSWERS: HOMEWORK NUMBER TWO

1. 67

2. a. F b. T c. T; highest the mean could be is 76; lowest is 44.

3. 14.75 pounds.

4. a. 104 b. 300

5. a. 110 percent b. 100 percent

6. 387

7. median is 2.15; modes are 2.1 and 2.4; mean is 2.2.

8. modes are 2 and 5, median is 2.5.

9. 87

10. 3

11. mean deviation is 1.6; variance is 4.4; standard deviation is $\sqrt{4.4}$.

12. mean is 78, variance is 144, standard deviation is 12

13. a. 165 decibels b. standard deviation is

$$\sqrt{\frac{1}{30}\left[15(15)^2 + 10(5)^2 + 5(35)^2\right]}$$

c. 160

14. mean is $\frac{1}{30}4605$, variance is

$$\frac{1}{30}\left[708,625 - \frac{1}{30}(4605)^2\right].$$

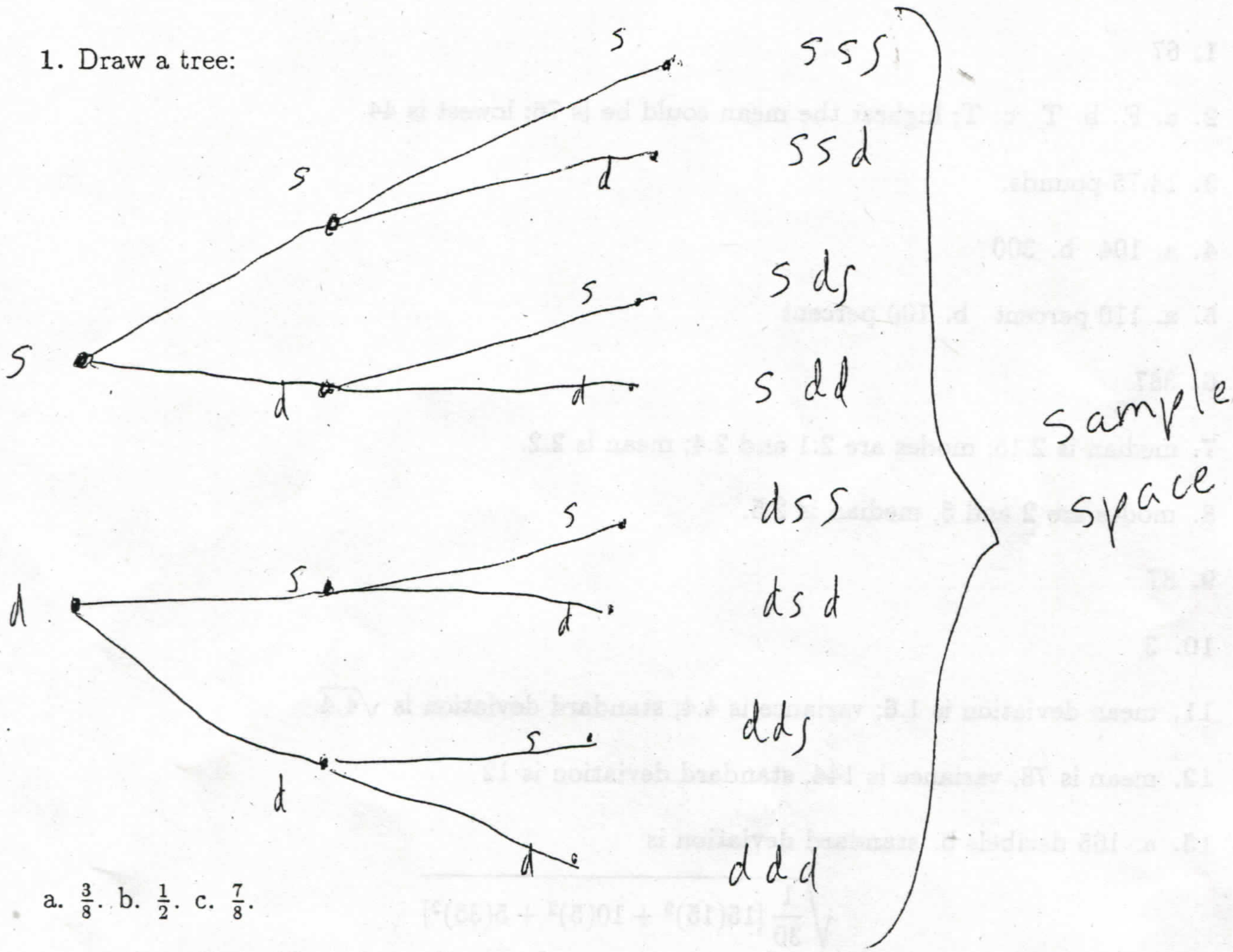
15. mean is $\frac{1}{26}(837)$, the standard deviation is

$$\sqrt{\frac{1}{26}\left[27,819 - \frac{1}{26}(837)^2\right]}$$

16. mean is $\frac{1}{20}$, median is 0, mode is -1, variance is $\frac{1}{20}\left(29 - \frac{1}{20}\right)$.

ANSWERS: HOMEWORK NUMBER THREE

1. Draw a tree:



a. $\frac{3}{8}$. b. $\frac{1}{2}$. c. $\frac{7}{8}$.

2. a. $\frac{5}{26}$ b. $\frac{21}{26}$ c. $\frac{15}{26}$ d. $\frac{2}{13}$ e. $\frac{7}{13}$

3. a. $\frac{4}{10}$ b. $\frac{6}{10}$.

4. a. $\frac{1}{400}$ b. $\frac{3}{400}$ c. $\frac{11}{400}$

5. a. $\frac{5}{6}$ b. $\frac{1}{6}$ c. $\frac{1}{6}$ d. $\frac{1}{12}$ e. $\frac{1}{6}$ f. $\frac{1}{6}$ g. $\frac{1}{12}$

6. .65

7. .8

8. .7

9. a. .35 b. .25 c. .65

10. a. .4 b. .45

11. a. .4 b. .1 c. .6

12. a. .3 b. .7 c. .6

13. a. $\frac{1}{3}$ b. $\frac{1}{6}$

14. a. $\frac{3}{10}$ b. $\frac{10\cdot9\cdot8\cdot7\cdot6\cdot5}{10^6}$

15. a. $\frac{1}{60}$ b. $\frac{1}{10}$ c. $\frac{1}{20}$ d. $\frac{1}{20}$

16. $\frac{\binom{7}{3}}{\binom{10}{6}} = \frac{7!6!4!}{4!3!10!}$

For the remaining answers, I will leave it to you to change $\binom{n}{k}$ to $\frac{n!}{k!(n-k)!}$, as I did in number 16.

17. a. $\frac{\binom{8}{2}\binom{3}{2}\binom{7}{1}}{\binom{18}{5}}$ b. $\frac{\binom{10}{5}}{\binom{18}{5}}$ c. $1 - \frac{\binom{10}{5}}{\binom{18}{5}}$

18. $\frac{\binom{7}{3}}{\binom{9}{5}}$.

19. a. $\frac{\binom{10}{2}\binom{15}{2}\binom{12}{2}}{\binom{37}{6}}$ b. $\frac{\binom{10}{4}\binom{27}{2}}{\binom{37}{6}}$ c. $\frac{\binom{27}{6}}{\binom{37}{6}}$ d. $\frac{\binom{10}{6}}{\binom{37}{6}}$

20. a. $\frac{\binom{10}{5}}{\binom{50}{5}}$ b. $\frac{\binom{10}{1}\binom{20}{2}\binom{20}{2}}{\binom{50}{5}}$. c. $\frac{\binom{10}{1}\binom{40}{4}}{\binom{50}{5}}$. d. $\frac{\binom{40}{5}}{\binom{50}{5}}$.

$$\text{e.} \quad \frac{\binom{20}{5} + \binom{10}{1}\binom{20}{1}\binom{20}{3} + \binom{10}{2}\binom{20}{2}\binom{20}{1}}{\binom{50}{5}}.$$

ANSWERS: HOMEWORK NUMBER FOUR

1.a. $\frac{1}{4}$ b. $\frac{1}{2}$ c. $\frac{4}{7}$

2. .06

3. .06

4. a. .12 b. .63 c. .75

5. a. $(.05)^2$ b. $(.95)^2$ c. $1-(.95)^2$

6. a. $(.6)(.4)(.2) = .048$ b. $1 - .048$

7. $(.3)(.2)(.2)(.2)(.1)(.1) = 2.4 \cdot 10^{-5}$

8. $1.5 \cdot 10^{-8}$

9. 5

10. a. .12 b. .58 c. .42 d. .45 e. .55

11. a. .04 b. $\frac{8}{27}$

12. a. .106 b. $\frac{7}{106}$

13. $\frac{21}{23}$

14. a. $\frac{7}{24}$ b. $\frac{2}{7}$

15. $\frac{(.1)(.2)}{(.1)(.2)+(.3)(.1)+(.6)(.05)} = \frac{1}{4}$

16. a. .175 b. $\frac{1}{7}$

ANSWERS: HOMEWORK NUMBER FIVE

1. The expected value $\mu = 1$. The variance is

$$\sigma^2 = (-5-1)^2(.05) + (-1-1)^2(.15) + (0-1)^2(.45) + (4-1)^2(.35).$$

2. a. .55 b. .35 c. .7 d. .45 e. .3 f. .65 g. .35
h. $25 \leq c < 30$ i. $c \geq 25$

3. 0;

x	1	2	3	−2
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$

4. 640

5. −20 cents

6. a. 152 pounds b. Var(X) equals

$$(200-152)^2(\frac{1}{5}) + (170-152)^2(\frac{1}{5}) + (150-152)^2(\frac{1}{5}) + (120-152)^2(\frac{2}{5}).$$

7.

x	−1	1	2
$p(x)$	.3	.6	.1

8.

x	−1	0	2
$p(x)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

9. a. .25 b. .5 c. .75 d. $c = 440$ e. $c = 440$ f. $c = 180$

10. a. .25 b. .1 c. .4 d. .15 e. .35 f. .65

11. a. .22 b. .36 c. .1 d. .58 e. .18 f. .68

12. a. $c = 55$ b. $c = 10$ c. $c = 55$ d. 60

13. At least $\frac{8}{9}$

14. 50 days

15. a. at least $\frac{8}{9}$ b. at least $\frac{77}{81}$ c. $[10, 26]$ d. $[18 - 8\sqrt{5}, 18 + 8\sqrt{5}]$

16. $\frac{1}{10}$

17. 440

18. $\frac{9}{16}$

19. 7 inches.

ANSWERS: HOMEWORK NUMBER SIX

1. a. .4332 b. .0668 c. .9332 d. .0919 e. .7745

2. a. .52 b. 1.13 c. -1.28 d. $-.84$ e. 1.28 f. $-.84$ g. -1.42
h. .03 i. 1.42 j. $5+7(.84)$ k. $-4-5(1.28)$ ℓ. $1-(.2)(2.33)$

3. 2.0945 inches

4. a. .0401 b. .1865 c. .1865 d. .0401

5. a. $5,000+300(10)=8,000$ quarts b. $5,000+300(2.33)=5,699$ quarts

6. a. .8413 b. $(.8413)^2$ c. 5384 liters

7. a. (no. 13, HW 5) .9974 b. (no. 14, HW 5) 34.66 days

8. $23,000-1,400(5)=16,000$ dollars

9. $23,000-1,400(1.75)=20,550$ dollars

10. $(.3594)^2$

11. $\frac{21}{25}$

12. $150+30(1.64)=199.2$ or $150+30(1.65)=199.5$.

ANSWERS: HOMEWORK NUMBER SEVEN

1. a. 0 (approximately) b. .005

2. a. 0 (approximately) b. .046 c. .168

3. .279

4. expected number is 90; variance is 9

5. a. .0043 b. .0516 c. .8192

6. .0023

7. .164

8. .0054

9. a. .1353 b. .1429

10. a. .0821 b. .6985

11. .9040

12. .2466

13. $2 + (.5)(2.06)$ cups

14. .309

15. $\frac{\binom{3}{2}\binom{7}{3}}{\binom{10}{5}} \sim .417$

16.

$$\frac{\binom{4}{1}\binom{5}{2} + \binom{4}{2}\binom{5}{1} + \binom{4}{3}\binom{5}{0}}{\binom{9}{3}} \quad \text{OR } 1 - \frac{\binom{5}{3}}{\binom{9}{3}}$$

17. .496

18. $1 - \left(\frac{\binom{40}{8} + \binom{10}{1}\binom{40}{7}}{\binom{50}{8}} \right)$

19. a. $\frac{10!}{1!5!2!2!}(.2)^1(.5)^5(.1)^2(.2)^2$ b. $b(0, 10, .2) = .107$ c. $b(5, 10, .2) = .026$
d. $b(10, 10, .7) = .028$ e. $b(10, 10, .8) = .107$

20. NOTE that there are 4 greens, 10 reds, 2 blues, 4 yellows.
a. $\frac{\binom{4}{1}\binom{10}{5}\binom{2}{2}\binom{4}{2}}{\binom{20}{10}}$ b. $\frac{\binom{16}{10}}{\binom{20}{10}}$ c. 0 d. $\frac{\binom{14}{10}}{\binom{20}{10}}$ e. $\frac{\binom{16}{10}}{\binom{20}{10}}$

21. a. $\frac{\binom{20}{10}\binom{30}{6}}{\binom{50}{16}}$ b. $\frac{\binom{40}{16}}{\binom{50}{16}}$ c. 0 d. $\frac{\binom{30}{16}}{\binom{50}{16}}$ e. $\frac{\binom{10}{4}\binom{40}{12}}{\binom{50}{16}}$

22. a. $\frac{9!}{2!2!5!}(.2)^2(.25)^2(.55)^5$ b. $\frac{9!}{5!3!1!}(.5)^5(.25)^3(.05)$ c. $b(0, 9, .2) = .134$
d. $b(9, 9, .7) = .040$ e. $b(4, 9, .2) = .066$

23. $\frac{19}{27}$

24. 3.8 cups

25. .738

26. .0082

27. .1912 (Poisson approximation of binomial, with $p = .008, n = 100$)

ANSWERS: HOMEWORK NUMBER EIGHT

1. a. 5 b. $5(\frac{68}{79})$

2. mean is 45; standard deviation is $\frac{5}{4}\sqrt{\frac{84}{99}}$

3. a. $\frac{.09}{64}\frac{336}{399}$ b. $\frac{.09}{64}$

4. .0228

5. a. .8611 b. .9104

6. a. .0062 b. .0668 c. .9544

7. .0062

8. a. at most .16 b. .0456 c. .0062

9. .9772

10. .2266

11. .4207

12. a. .006 b. .201

13. a. $\frac{31}{48}$ b. $\frac{25}{31}$

14. .055

15. .1353

16. $(2-(.05)(\frac{10}{\sqrt{3}})$ inches

17. .9564

18. $\frac{\binom{3}{2}\binom{7}{4}+\binom{7}{3}}{\binom{10}{6}}$

19. .5961
(This is $b(0,100,.023)+b(1,100,.023)+b(2,100,.023)$, Poisson approximation.)

20. .7127
(This is $1-(b(0,400,.0062)+b(1,400,.0062))$, Poisson approximation, $\lambda=2.5$.)

21. .8904

22. .0062

23. $[10-2(.258),10+2(.258)]$

24. $8(8+\frac{2.58}{5})$ pints

25. 4915.2 inches

26. .9664

ANSWERS: HOMEWORK NUMBER NINE

1. .9938

2. mean is $\frac{19}{8}$; median is 2.5; modes are 1, 3 and 5.

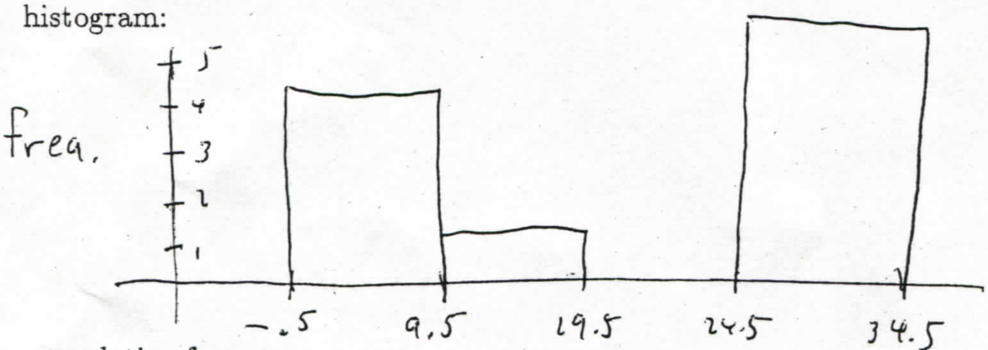
3. $\frac{\binom{6}{3}}{\binom{10}{3}}$

4. a. $(.9)(.05) + (.1)(.25) = .07$ b. $\frac{(.9)(.05)}{.07} = \frac{9}{14}$

5. frequency distribution:

number	frequency
0 − 9	4
10 − 19	1
20 − 29	0
30 − 39	5

histogram:

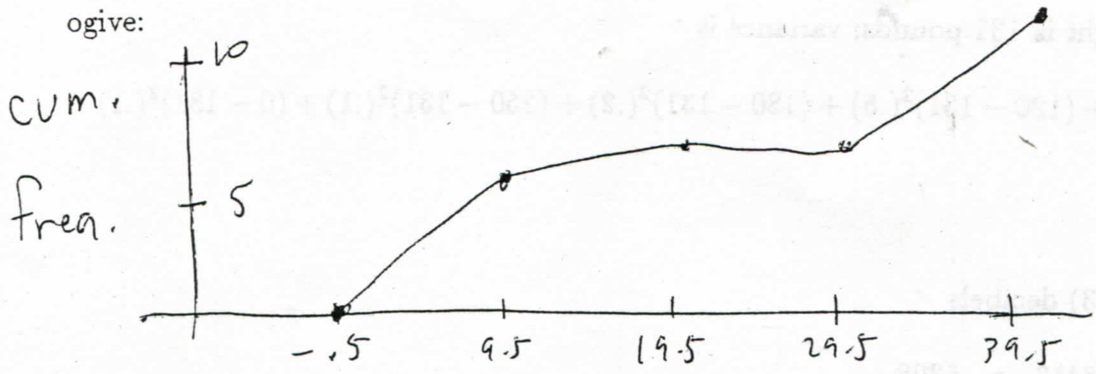


cumulative frequency:

number	cumulative frequency
$< -.5$	0
< 9.5	4
< 19.5	5
< 29.5	5
< 39.5	10

16

ogive:



6. a. $6 + 10(.3) = 9$ inches b. $6 + (2.33)(.3) = 6.699$ inches

7. .03

8. $26^3 \cdot 10^4$

9. $\frac{(.8)(.1)}{((.8)(.1)+(.6)(.9))} = \frac{4}{31}$

10. .7 quarts

11. .6988

12. .0401

13. .1587

14. a. .5 b. 2.2 c.

$$\sqrt{(-2.2)^2(.2) + (-1.2)^2(.3) + (-.2)^2(.1) + (.8)^2(.1) + (1.8)^2(.15) + (2.8)^2(.1) + (3.8)^2(.05)}$$

15. .0336

16. .0613

17. .0062

18. mean is $-\frac{1}{10}$; mode is -1; median is $-.5$; variance is $\frac{1}{20}\left[24 - \frac{1}{20}(-2)^2\right]$

19. a. $(.7)(.6)(.4) = .168$ b. .3

20. a. $(.95)(.01) + (.99)(.1)$ b.

$$\frac{(.95)(.01)}{(.95)(.01) + (.99)(.1)}$$

21. expected weight is 131 pounds; variance is

$$(100-131)^2(.1)+(120-131)^2(.5)+(180-131)^2(.2)+(250-131)^2(.1)+(0-131)^2(.1)$$

22. .472

23. $60+(1.5)(2.33)$ decibels

24. a. .0228 b. .8413 c. .5398

25. 9.8 seconds

26. $60+(1.5)(10)=75$ decibels

27. a. $(.6)(.9)+(.4)(.15)=.6$ b. $\frac{(.6)(.9)}{.6}=.9$

28. .962

29. .0176

30. .0228

31. a. $b(1,10,.1)=.387$ b. $\frac{10!}{1!2!4!2!1!}(.1)(.2)^2(.4)^4(.2)^2(.1)$ c. $\frac{10!}{5!5!}(.1)^5(.2)^5$
d. $b(5,10,.3)=.103$

32. NOTE: there are 5 As, 10 Bs, 20 Cs, 10Ds and 5 Fs in the class of 50.

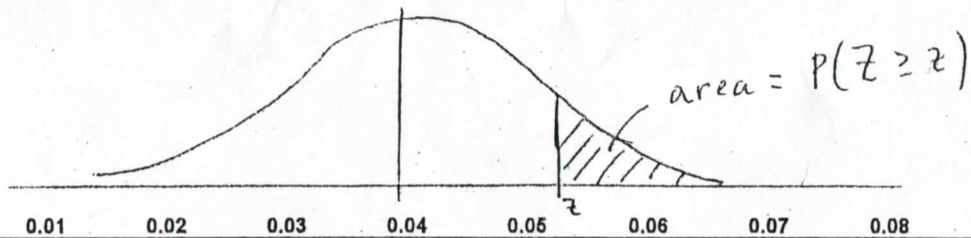
$$a. \frac{\binom{5}{1}\binom{45}{9}}{\binom{50}{10}} \quad b. \frac{\binom{5}{1}\binom{10}{2}\binom{20}{4}\binom{10}{2}\binom{5}{1}}{\binom{50}{10}} \quad c. \frac{\binom{5}{5}\binom{10}{5}}{\binom{50}{10}}=\frac{\binom{10}{5}}{\binom{50}{10}} \quad d. \frac{\binom{15}{5}\binom{35}{5}}{\binom{50}{10}}$$

33. 140 quarts.

34. At least .84.

35. .9876. (NOTICE the improved certainty over no. 34.)

The Standard Normal Distribution (Areas in the Right Tail)



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010

Binomial Probabilities, Denoted b(x; n, p)---Page 1 of 2

n	x	p = 0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
2	0	0.902	0.810	0.640	0.490	0.360	0.250	0.160	0.090	0.040	0.010	0.002
	1	0.095	0.180	0.320	0.420	0.480	0.500	0.480	0.420	0.320	0.180	0.095
	2	0.002	0.010	0.040	0.090	0.160	0.250	0.360	0.490	0.640	0.810	0.902
3	0	0.857	0.729	0.512	0.343	0.216	0.125	0.064	0.027	0.008	0.001	
	1	0.135	0.243	0.384	0.441	0.432	0.375	0.288	0.189	0.096	0.027	0.007
	2	0.007	0.027	0.096	0.189	0.288	0.375	0.432	0.441	0.384	0.243	0.135
	3		0.001	0.008	0.027	0.064	0.125	0.216	0.343	0.512	0.729	0.857
4	0	0.815	0.656	0.410	0.240	0.130	0.062	0.026	0.008	0.002		
	1	0.171	0.292	0.410	0.412	0.346	0.250	0.154	0.076	0.026	0.004	
	2	0.014	0.049	0.154	0.265	0.346	0.375	0.346	0.265	0.154	0.049	0.014
	3		0.004	0.026	0.076	0.154	0.250	0.346	0.412	0.410	0.292	0.171
	4			0.002	0.008	0.026	0.062	0.130	0.240	0.410	0.656	0.815
5	0	0.774	0.590	0.328	0.168	0.078	0.031	0.010	0.002			
	1	0.204	0.328	0.410	0.360	0.259	0.156	0.077	0.028	0.006		
	2	0.021	0.073	0.205	0.309	0.346	0.312	0.230	0.132	0.051	0.008	0.001
	3	0.001	0.008	0.051	0.132	0.230	0.312	0.346	0.309	0.205	0.073	0.021
	4			0.006	0.028	0.077	0.156	0.259	0.360	0.410	0.328	0.204
	5				0.002	0.010	0.031	0.078	0.168	0.328	0.590	0.774
6	0	0.735	0.531	0.262	0.118	0.047	0.016	0.004	0.001			
	1	0.232	0.354	0.393	0.303	0.187	0.094	0.037	0.010	0.002		
	2	0.031	0.098	0.246	0.324	0.311	0.234	0.138	0.060	0.015	0.001	
	3	0.002	0.015	0.082	0.185	0.276	0.312	0.276	0.185	0.082	0.015	0.002
	4		0.001	0.015	0.060	0.138	0.234	0.311	0.324	0.246	0.098	0.031
	5			0.002	0.010	0.037	0.094	0.187	0.303	0.393	0.354	0.232
	6				0.001	0.004	0.016	0.047	0.118	0.262	0.531	0.735

Binomial Probabilities, Denoted b(x; n, p)---Page 2 of 2

n	x	p = 0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
7	0	0.698	0.478	0.210	0.082	0.028	0.008	0.002				
	1	0.257	0.372	0.367	0.247	0.131	0.055	0.017	0.004			
	2	0.041	0.124	0.275	0.318	0.261	0.164	0.077	0.025	0.004		
	3	0.004	0.023	0.115	0.227	0.290	0.273	0.194	0.097	0.029	0.003	
	4		0.003	0.029	0.097	0.194	0.273	0.290	0.227	0.115	0.023	0.004
	5			0.004	0.025	0.077	0.164	0.261	0.318	0.275	0.124	0.041
	6				0.004	0.017	0.055	0.131	0.247	0.367	0.372	0.257
	7					0.002	0.008	0.028	0.082	0.210	0.478	0.698
8	0	0.663	0.430	0.168	0.058	0.017	0.004	0.001				
	1	0.279	0.383	0.336	0.198	0.090	0.031	0.008	0.001			
	2	0.051	0.149	0.294	0.296	0.209	0.109	0.041	0.010	0.001		
	3	0.005	0.033	0.147	0.254	0.279	0.219	0.124	0.047	0.009		
	4		0.005	0.046	0.136	0.232	0.273	0.232	0.136	0.046	0.005	
	5			0.009	0.047	0.124	0.219	0.279	0.254	0.147	0.033	0.005
	6			0.001	0.010	0.041	0.109	0.209	0.296	0.294	0.149	0.051
	7				0.001	0.008	0.031	0.090	0.198	0.336	0.383	0.279
	8					0.001	0.004	0.017	0.058	0.168	0.430	0.663
9	0	0.630	0.387	0.134	0.040	0.010	0.002					
	1	0.299	0.387	0.302	0.156	0.060	0.018	0.004				
	2	0.063	0.172	0.302	0.267	0.161	0.070	0.021	0.004			
	3	0.008	0.045	0.176	0.267	0.251	0.164	0.074	0.021	0.003		
	4	0.001	0.007	0.066	0.172	0.251	0.246	0.167	0.074	0.017	0.001	
	5		0.001	0.017	0.074	0.167	0.246	0.251	0.172	0.066	0.007	0.001
	6			0.003	0.021	0.074	0.164	0.251	0.267	0.176	0.045	0.008
	7				0.004	0.021	0.070	0.161	0.267	0.302	0.172	0.063
	8					0.004	0.018	0.060	0.156	0.302	0.387	0.299
	9						0.002	0.010	0.040	0.134	0.387	0.630
10	0	0.599	0.349	0.107	0.028	0.006	0.001					
	1	0.315	0.387	0.268	0.121	0.040	0.010	0.002				
	2	0.075	0.194	0.302	0.233	0.121	0.044	0.011	0.001			
	3	0.010	0.057	0.201	0.267	0.215	0.117	0.042	0.009	0.001		
	4	0.001	0.011	0.088	0.200	0.251	0.205	0.111	0.037	0.006		
	5		0.001	0.026	0.103	0.201	0.246	0.201	0.103	0.026	0.001	
	6			0.006	0.037	0.111	0.205	0.251	0.200	0.088	0.011	0.001
	7			0.001	0.009	0.042	0.117	0.215	0.267	0.201	0.057	0.010
	8				0.001	0.011	0.044	0.121	0.233	0.302	0.194	0.075
	9					0.002	0.010	0.040	0.121	0.268	0.387	0.315
	10						0.001	0.006	0.028	0.107	0.349	0.599

Poisson Probabilities $p(x) = 1/x!\ e^{-\lambda} \lambda^x$

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066	0.3679
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659	0.3679
2	0.0045	0.0164	0.0333	0.0536	0.0758	0.0988	0.1216	0.1438	0.1647	0.1839
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494	0.0613
4		0.0001	0.0002	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111	0.0153
5				0.0001	0.0002	0.0003	0.0007	0.0012	0.0020	0.0031
6							0.0001	0.0002	0.0003	0.0005
7										0.0001

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	0.3329	0.3012	0.2725	0.2466	0.2231	0.2019	0.1827	0.1653	0.1496	0.1353
1	0.3662	0.3614	0.3543	0.3452	0.3347	0.3230	0.3106	0.2975	0.2841	0.2707
2	0.2014	0.2169	0.2303	0.2417	0.2510	0.2584	0.2640	0.2678	0.2700	0.2707
3	0.0738	0.0867	0.0998	0.1127	0.1255	0.1378	0.1496	0.1607	0.1710	0.1804
4	0.0203	0.0260	0.0324	0.0395	0.0471	0.0551	0.0636	0.0723	0.0812	0.0902
5	0.0045	0.0062	0.0084	0.0111	0.0141	0.0176	0.0215	0.0260	0.0309	0.0361
6	0.0008	0.0012	0.0018	0.0026	0.0035	0.0047	0.0061	0.0078	0.0098	0.0120
7	0.0001	0.0002	0.0003	0.0005	0.0008	0.0011	0.0015	0.0020	0.0027	0.0034
8			0.0001	0.0001	0.0001	0.0002	0.0003	0.0005	0.0006	0.0009
9							0.0001	0.0001	0.0001	0.0002

x	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	0.1225	0.1108	0.1003	0.0907	0.0821	0.0743	0.0672	0.0608	0.0550	0.0498
1	0.2572	0.2438	0.2306	0.2176	0.2052	0.1931	0.1815	0.1703	0.1596	0.1494
2	0.2700	0.2681	0.2652	0.2613	0.2565	0.2510	0.2450	0.2384	0.2314	0.2240
3	0.1890	0.1966	0.2033	0.2090	0.2138	0.2176	0.2205	0.2225	0.2237	0.2240
4	0.0992	0.1082	0.1169	0.1254	0.1336	0.1414	0.1488	0.1557	0.1622	0.1680
5	0.0416	0.0476	0.0538	0.0602	0.0668	0.0735	0.0804	0.0872	0.0940	0.1008
6	0.0146	0.0174	0.0206	0.0241	0.0278	0.0319	0.0362	0.0407	0.0455	0.0504
7	0.0044	0.0055	0.0068	0.0083	0.0099	0.0118	0.0139	0.0163	0.0188	0.0216
8	0.0011	0.0015	0.0019	0.0025	0.0031	0.0038	0.0046	0.0057	0.0068	0.0081
9	0.0003	0.0004	0.0005	0.0007	0.0009	0.0011	0.0014	0.0018	0.0022	0.0027
10	0.0001	0.0001	0.0001	0.0002	0.0002	0.0003	0.0004	0.0005	0.0006	0.0008
11						0.0001	0.0001	0.0001	0.0002	0.0002
12										0.0001

INDEX