1. Suppose a clinically accepted value for mean systolic blood pressure in females, aged 65-74 is 133 mmHg and the standard deviation is 20 mmHg.

   a) If a 70-year-old- woman is selected at random from the population, what is the probability that her systolic blood pressure is equal to or less than 120 mmHg?
   $P[X \le 120 \mid \mu=133, \sigma=20] = P[Z \le (120-133)/20] = P[Z \le -0.65] = 0.2578$
   b) The systolic blood pressure of a 65-year-old woman selected at random from the population was 163 mmHg. How many standard deviations above the mean is this value? (HINT: Z values are in standard deviation units.)
   Z=(163-133)/20 = 1.5 (standard deviations above the mean)
   c) Between what two blood pressure readings will 95% of all systolic blood pressures for 65-74-year-old women lie?
   $(\mu - 1.96\sigma, \mu + 1.96\sigma)$ **encompasses 95% of the population (93.8, 172.2)**
   d) What proportion will lie inside the range 123 to 133 mmHg?
   $P[123<X<133 \mid \mu=133, \sigma=20] = P[(123-133)/20<Z<(133-133)/20] = P[-0.5<Z<0]$
   $= P[Z < 0] - P[Z < -0.5] = .5000 - 0.3085 = 0.1915.$
   e) What proportion will lie outside the range 73 to 193 mmHg?
   $P[X<73 \text{ or } X>193 \mid \mu=133, \sigma=20] = 1-P[73<X<193 \mid \mu=133, \sigma=20]$
   $=1-P[(73-133)/20 < Z < (193-133)/20] = 1-P[-3<Z<3]= 1 - .9974 = 0.0026$

To calculate the above results, what assumption must be made about systolic blood pressures in this population? **That the distribution of systolic blood pressures in females, aged 65-74, was normally distributed.**

2. Refer to problem 1 above. As part of a screening project, the systolic blood pressures of 100 Seattle 65-74-year-old women were taken. The mean systolic blood pressure of this sample was 130 mmHg.

   a) What is the standard error of the mean systolic blood pressure in this case?
   Standard error of the mean is $\sigma/\sqrt{n}$ = 20/10=2.
   b) What is the probability of obtaining a sample mean systolic blood pressure less than or equal to 130 mmHg from a population with true mean 133 mmHg?
   $P[\bar{X} \le 130 \mid \mu=133, \sigma=20, n=100]= P[(\bar{X}- \mu)/(\sigma/\sqrt{n})] = P[Z \le -1.5]=0.0668.$
   c) Between what two mean systolic blood pressure values would we expect 95% of all sample means (of size n=100) to lie?
   $(\mu \pm 1.96\, \sigma/\sqrt{n})$ = (129.08,136.92)
   d) Based on the above information and computations, what conclusions might be drawn concerning the systolic blood pressure of Seattle women. What must you assume about the distribution of systolic blood pressures to make the answers to parts b and c valid?
   Their average blood pressure is somewhat low in comparison with the population mean (for a sample size of 100) but not extremely unlikely.

3. If a random variable, Z, has a standard normal distribution N(0,1), find the following:

   a) $P[Z \le 2]$ **= 0.9772**
   b) $P[Z \le -1]$ **= 0.1587**
   c) $P[Z > 1.645]$ **= (.0505 + .0495)/2 = .0500.**
   d) $P[0.4 < Z \le 1]$ **= 0.1859**
   e) $P[Z \le -1.96 \text{ or } Z \ge 1.96]$ **= .0500**

4. If a random variable, X, has a normal distribution with mean 1 and variance 4, find the following:

   a) $P[X \le 2]$ =**$P[Z \le (2-1)/2]=P[Z \le 0.50] = 0.6915$**
   b) $P[1 \le X < 3]$ = **$P[(1-1)/2 \le (X-\mu)/\sigma < (3-1)/2]=P[0 \le Z < 1]=0.3413$**

5. Suppose it is known that the hourly wages of a certain type of hospital employee are distributed with a mean of \$10.00 and a standard deviation of \$1.00. A random sample of 25 is selected from this population.

   a) What is the standard error of the mean in this case?
   Standard error of the mean is $\sigma/\sqrt{n}$ = \$1.00/5 = \$0.20
   b) What is the difference in interpretation between the standard deviation and the standard error in this case? (Express in words)
   The standard deviation, σ, is a population parameter that characterizes the variation in the population for all values. The standard error, σ/sqrt(n) (n is the sample size), characterizes the variation in the sample means. The former is a measure of variation for the population distribution; the latter is a measure of precision of sample means.
   c) What is the probability that the mean hourly wage for the sample will be between \$9.90 and \$10.10? **$P[-0.5 < Z < 0.5]=0.3829$.**
   d) What must you assume about the distribution of hourly wages to make the answer to part c. valid?
   Since the sample size is reasonably large, we can rely on the central limit theorem and assume the sample means are approximately normally distributed. We also require that the observations are independent, which they are by definition of a "random sample".

6. Suppose it is known that the hourly wages of a certain type of hospital employee are distributed with a mean of \$12.00 and a standard deviation of \$2.00. A random sample of 16 is selected from this population.

   a) What is the standard error of the mean in this case?
   Standard error of the mean is $\sigma/\sqrt{n}$ = \$2.00/4 = \$0.50
   b) What is the difference in interpretation between the standard deviation and the standard error in this case? (Express in words) **See interpretation in problem 5(b).**

7. **Diabetes.** A number of clinical characteristics were ascertained in a large group of subjects with insulin-dependent diabetes mellitus (IDDM). Suppose the distribution of percentage of ideal body weight in this group of patients is normal with a mean of 110 and a standard deviation of 13.
   X is $N(\mu=110,\sigma=13)$
   a) What percentage of subjects with IDDM are above their ideal body weight (i.e., above 100% ideal body weight)?
   $P[X \geq 100 \mid \mu=110, \sigma=13] = P[Z \geq (100-110)/13] = P[Z \geq -0.77] = 0.7791$.
   b) What percentage of subjects are overweight (defined as 10% or more above ideal body weight)?
   $P[X \geq 110 \mid \mu=110, \sigma=13] = P[Z \geq (110-110)/13] = P[Z \geq 0]=0.5000$
   c) What percentage of subjects are obese (defined as 20% or more above ideal body weight)?
   $P[X \geq 120 \mid \mu=110, \sigma=13] = P[Z \geq (120-110)/13] = P[Z \geq 0.77] = 0.2209$.
   d) What percentage of subjects are underweight (defined as 10% or more below ideal body weight)?
   $P[X \leq 90 \mid \mu=110, \sigma=13] = P[Z \leq (90-110)/13] = P[Z \leq -1.54] = 0.0618$
   e) What percentage of subjects have normal body weight (within 10% of ideal body weight)?
   $P[90 < X < 110 \mid \mu=110, \sigma=13] = P[-1.54 < Z < 0] = 0.5000 - 0.0618 = 0.4382$.

8. Suppose the average length of stay in a chronic disease hospital of a certain type of patient is 60 days with a standard deviation of 15 days. Find the probability that a randomly selected patient from this group will have a length of stay:

   a) less than 30 days
   $P[X < 30 \mid \mu=60, \sigma=15] = P[Z < (30-60)/15] = P[Z < -2] = 0.0228$
   b) greater than 50 days
   $P[X > 50 \mid \mu=60, \sigma=15] = P[Z > -0.67] = 0.7486$
   c) between 30 and 60 days
   $P[30 < X < 60] = P[-2 < Z < 0] = P[0 < Z < 2] = 0.9772 – 0.5000 = 0.4772$

d) What number of days is at the 80th percentile?
$P[Z < z_{0.80}] = 0.80$. Now look at probability that cuts off 80 percent of the distribution to the left and then read off the value for the standard normal value. We see the value 0.845 cuts off 80 percent of the standard normal distribution. $P[Z < 0.845] = 0.80$. Now convert the value from the standard (Z-scale) scale to the scale of the length of stay, X, with mean $\mu=60$, and standard deviation $\sigma=15$, as $X = \mu + \sigma\ z_{0.80} = 60 + 15(0.845) = 72.68$.

e) What must you assume about the distribution of length of stay to make the answers in a-d valid?
That, X, length of stay, is normally distributed

9. A patient checks her diastolic blood pressure at home and finds her average blood pressure for a two-week period to be 84 mmHg. Assume her blood pressure to be normally distributed with a standard deviation of 5 mmHg. A nurse checks her diastolic blood pressure in the office and finds a value of 110 mmHg. The office reading is apparently

a) not atypical of her distribution of blood pressures
b) consistent with normotensive diastolic blood pressure
c) below the 95th percentile for this patient
d) extremely high for this patient

$P[X > 110 \mid \mu=84, \sigma=5] = P[Z > (110-84)/5] = P[Z > 5.2] < 0.0000001$ (very close to zero)

10. **Nutrition**. Suppose that total carbohydrate intake in 12- to 14-year old boys is normally distributed with a mean of 124 g/1000 cal and a standard deviation of 20 g/1000 cal.

a) What percentage of boys in this age range will have carbohydrate above 140 g/1000 cal?
$P[X \geq 140 \mid \mu=124, \sigma=20] = P[Z \geq (140-124)/20] = P[Z \geq 0.8] = 0.2119$.

b) What percentage of boys in this age range will have carbohydrate below 90 g/1000 cal?
$P[X < 90 \mid \mu=124, \sigma=20] = P[Z < (90-124)/20] = P[Z < -1.7] = 0.0446$.

11. **Hypertension.** The Second Task Force Report on Blood Pressure Control in Children (*Pediatrics*, *79(1)*, 1-25, 1987) reported blood pressure norms for children by age and sex groups. The mean +/– standard deviation for 17-year old boys for diastolic blood pressure is 63.7 +/– 11.4 mm Hg, based on these data.

a) One approach of defining elevated diastolic blood pressure is to use 90 mm Hg – the standard for elevated diastolic blood pressure for adults – as the threshold. What percentage of 17-year old boys would be found to have elevated diastolic blood pressure by this standard?
$P[X > 90 \mid \mu=63.7, \sigma=11.4] = P[Z > (90-63.7)/11.4] = P[Z > 2.31] = 0.0104$.

b) Suppose there are 2000 17-year old boys sampled, of whom 25 had elevated diastolic blood pressure by the criterion used in 11(a). Is this an unusually high number of 17-year old boys with elevated systolic blood pressure? Approximately how many of the 17-year old boys would you expect to have elevated diastolic blood pressure?
No. This does not appear to be unusually high for a sample of 2000 17-year old boys. In a sample of N=2000 and a probability of success, p=.0104 (i.e., probability of a boy in the population having elevated diastolic blood pressure), we would expect N*p = 2000*.0104 = 20.8 or approximately 21 boys to have elevated diastolic blood pressure.