Chapter 2

2.1 Eleanor's standardized score, $z = \frac{680-500}{100} = 1.8$, is higher than Gerald's standardized score, $z = \frac{27-18}{6} = 1.5$.

2.2 The standardized batting averages (*z*-scores) for these three outstanding hitters are:

Player	*z*-score
Cobb	$z = \frac{.420-.266}{.0371} = 4.15$
Williams	$z = \frac{.406-.267}{.0326} = 4.26$
Brett	$z = \frac{.390-.261}{.0317} = 4.07$

All three hitters were at least 4 standard deviations above their peers, but Williams' *z*-score is the highest.

2.3 (a) Judy's bone density score is about one and a half standard deviations below the average score for all women her age. The fact that your standardized score is negative indicates that your bone density is below the average for your peer group. The magnitude of the standardized score tells us how many standard deviations you are below the average (about 1.5). (b) If we let σ denote the standard deviation of the bone density in Judy's reference population, then we can solve for σ in the equation $-1.45 = \frac{948-956}{\sigma}$. Thus, $\sigma \doteq 5.52$.

2.4 (a) Mary's *z*-score (0.5) indicates that her bone density score is about half a standard deviation above the average score for all women her age. Even though the two bone density scores are exactly the same, Mary is 10 years older so her *z*-score is higher than Judy's (−1.45). Judy's bones are healthier when comparisons are made to other women in their age groups. (b) If we let σ denote the standard deviation of the bone density in Mary's reference population, then we can solve for σ in the equation $0.5 = \frac{948-944}{\sigma}$. Thus, $\sigma \doteq 8$. There is more variability in the bone densities for older women, which is not surprising.

2.5 (a) A histogram is shown below. The distribution of unemployment rates is symmetric with a center around 5%, rates varying from 2.7% to 7.1%, and no gaps or outliers.

(b) The average unemployment rate is $\bar{x} = 4.896\%$ and the standard deviation of the rates is $s = 0.976\%$. The five-number summary is: 2.7%, 4.1%, 4.8%, 5.5%, 7.1%. The distribution is symmetric with a center at 4.896%, a range of 4.4%, and no gaps or outliers. (c) The unemployment rate for Illinois is the 84th percentile; Illinois has one of the higher unemployment rates in the country. More specifically, 84% of the 50 states have unemployment rates at or below the unemployment rate in Illinois (5.8%). (d) Minnesota's unemployment rate (4.3%) is at the 30th percentile and the *z*-score for Minnesota is $z = -0.61$. (e) The intervals, percents guaranteed by Chebyshev's inequality, observed counts, and observed percents are shown in the table below.

k	Interval	% guaranteed by Chebyshev	Number of values in interval	Percent of values in interval
1	3.920−5.872	At least 0%	35	70%
2	2.944−6.848	At least 75%	47	94%
3	1.968−7.824	At least 89%	50	100%
4	0.992−8.800	At least 93.75%	50	100%
5	0.016−9.776	At least 96%	50	100%

As usual, Chebychev's inequality is very conservative; the observed percents for each interval are higher than the guaranteed percents.

2.6 (a) The rate of unemployment in Illinois increased 28.89% from December 2000 (4.5%) to May 2005 (5.8%). (b) The *z*-score $z = \frac{4.5 - 3.47}{1} = 1.03$ in December 2000 is higher than the *z*-score $z = \frac{5.8 - 4.896}{0.976} = 0.9262$ in May 2005. Even though the unemployment rate in Illinois increased substantially, the *z*-score decreased slightly. (c) The unemployment rate for Illinois in December 2000 is the 86th percentile. $\left(\frac{42+1}{50} = 0.86\right)$ Since the unemployment rate for Illinois in May 2005 is the 84th percentile, we know that Illinois dropped one spot $\left(\frac{1}{50} = 0.02\right)$ on the ordered list of unemployment rates for the 50 states.

2.7 (a) In the national group, about 94.8% of the test takers scored below 65. Scott's percentiles, 94.8th among the national group and 68th, indicate that he did better among all test takers than he did among the 50 boys at his school. (b) Scott's *z*-scores are

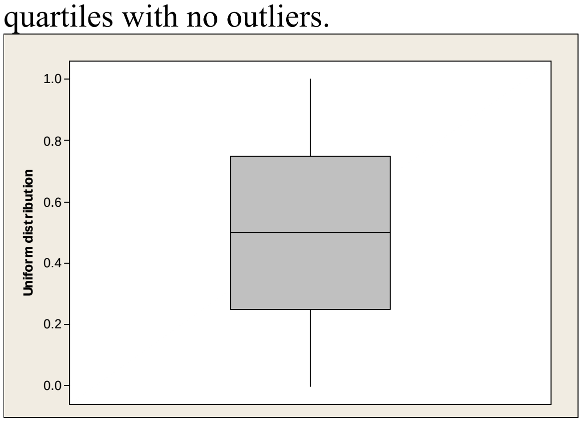
$z = \frac{64 - 46.9}{10.9} \doteq 1.57$ among the national group and $z = \frac{64 - 58.2}{9.4} \doteq 0.62$ among the 50 boys at his school. (c) The boys at Scott's school did very well on the PSAT. Scott's score was relatively better when compared to the national group than to his peers at school. Only 5.2% of the test takers nationally scored 65 or higher, yet about 23.47% scored 65 or higher at Scott's school. (d) Nationally, at least 89% of the scores are between 20 and 79.6, so at most 11% score a perfect 80. At Scott's school, at least 89% of the scores are between 30 and 80, so at most 11% score 29 or less.

2.8 Larry's wife should gently break the news that being in the 90th percentile is not good news in this situation. About 90% of men similar to Larry have identical or lower blood pressures. The doctor was suggesting that Larry take action to lower his blood pressure.

2.9 Sketches will vary. Use them to confirm that the students understand the meaning of (a) symmetric and bimodal and (b) skewed to the left.

2.10 (a) The area under the curve is a rectangle with height 1 and width 1. Thus, the total area under the curve is 1×1 = 1. (b) The area under the uniform distribution between 0.8 and 1 is 0.2×1 = 0.2, so 20% of the observations lie above 0.8. (c) The area under the uniform distribution between 0 and 0.6 is 0.6×1 = 0.6, so 60% of the observations lie below 0.6. (d) The area under the uniform distribution between 0.25 and 0.75 is 0.5×1 = 0.5, so 50% of the observations lie between 0.25 and 0.75. (e) The mean or "balance point" of the uniform distribution is 0.5.

2.11 A boxplot for the uniform distribution is shown below. It has equal distances between the quartiles with no outliers.



2.12 (a) Mean C, median B; (b) mean A, median A; (c) mean A, median B.

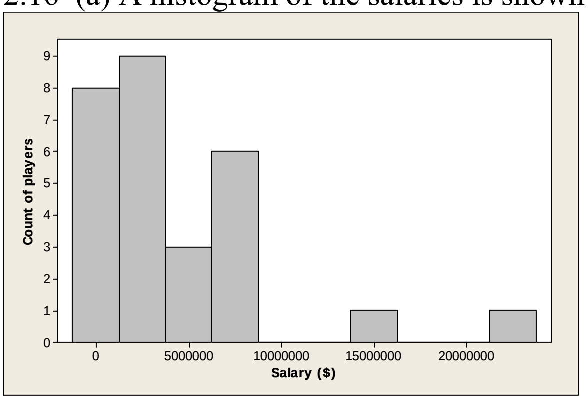
2.13 (a) The curve satisfies the two conditions of a density curve: curve is on or above horizontal axis, and the total area under the curve = area of triangle + area of 2 rectangles = $\frac{1}{2} \times 0.4 \times 1 + 0.4 \times 1 + 0.4 \times 1 = 0.2 + 0.4 + 0.4 = 1$. (b) The area under the curve between 0.6 and 0.8 is 0.2×1 = 0.2. (c) The area under the curve between 0 and 0.4 is

$\frac{1}{2}\times 0.4\times 1+0.4\times 1=0.2+0.4=0.6$. (d) The area under the curve between 0 and 0.2 is $\frac{1}{2}\times 0.2\times 0.5+0.2\times 1.5=0.05+0.3=0.35$. (e) The area between 0 and 0.2 is 0.35. The area between 0 and 0.4 is 0.6. Therefore the "equal areas point" must be between 0.2 and 0.4.

2.14 (a) The distribution should look like a uniform distribution, with height 1/6 or about 16.67%, depending on whether relative frequency or percent is used. If frequency is used, then each of the 6 bars should have a height of about 20. (b) This distribution is similar because each of the bars has the same height. This feature is a distinguishing characteristic of uniform distributions. However, the two distributions are different because in this case we have only 6 possible outcomes {1, 2, 3, 4, 5, 6}. In Exercise 2.10 there are an infinite number of possible outcomes in the interval from 0 to 1.

2.15 The *z*-scores are $z_w = \dfrac{72-64}{2.7} \doteq 2.96$ for women and $z_m = \dfrac{72-69.3}{2.8} \doteq 0.96$ for men. The *z*-scores tell us that 6 feet is quite tall for a woman, but not at all extraordinary for a man.

2.16 (a) A histogram of the salaries is shown below.



(b) Numerical summaries are provided below.

```
Variable   N    Mean   StDev  Minimum      Q1   Median       Q3   Maximum
Salaries  28 4410897 4837406   316000  775000  2875000  7250000  22000000
```

The distribution of salaries is skewed to the right with a median of \$2,875,000. There are two major gaps, one from \$8.5 million to \$14.5 million and another one from \$14.5 million to \$22 million. The salaries are spread from \$316,000 to \$22 million. The \$22 million salary for Manny Ramirez is an outlier. (c) David McCarty's salary of \$550,000 gives him a *z*-score of $z = \dfrac{550000-4410897}{4837406} \doteq -0.80$ and places him at about the 14th percentile. (d) Matt Mantei's salary of \$750,000 places him at the 25th percentile and Matt Clement's salary of \$6.5 million places him at the 75th percentile. (e) These percentiles do not match those calculated in part (b) because the software uses a slightly different method for calculating quartiles.

2.17 Between 2004 and 2005, McCarty's salary increased by \$50,000 (10%), while Damon's increased by \$250,000 (3.125%). The *z*-score for McCarty decreased from $z = \dfrac{500000-4243283.33}{5324827.26} \doteq -0.70$ in 2004 to −0.80 in 2005 while the *z*-score for Damon increased from $z = \frac{8000000 - 4243283.33}{5324827.26} \doteq 0.71$ in 2004 to 0.79 in 2005. Damon's salary percentile increased from the 87th (26 out of 30) in 2004 to the 93rd (26 out of 28) in 2005, while McCarty's decreased from the 20th (6 out of 30)in 2004 to the 14th (4 out of 28) in 2005.

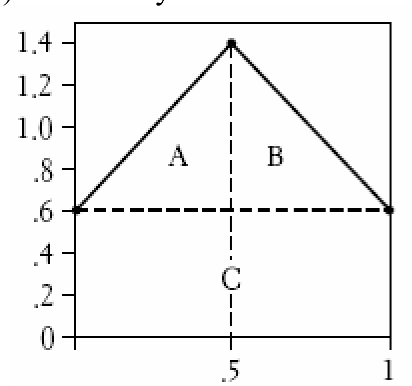
2.18 (a) The intervals, percents guaranteed by Chebyshev's inequality, observed counts, and observed percents are shown in the table below.

k	Interval	% guaranteed by Chebyshev	Number of values in interval	Percent of values in interval
1	73.93−86.07	At least 0%	18	72%
2	67.86−92.14	At least 75%	23	92%
3	61.79−98.21	At least 89%	25	100%
4	55.72−104.28	At least 93.75%	25	100%
5	49.65−110.35	At least 96%	25	100%

As usual, Chebyshev's inequality is very conservative; the observed percents for each interval are higher than the guaranteed percents. (b) Each student's z-score and percentile will stay the same because all of the scores are simply being shifted up by 4 points, $z = \frac{(x+4)-(\bar{x}+4)}{s} = \frac{x-\bar{x}}{s}$. (c) Each student's z-score and percentile will stay the same because all of the scores are being multiplied by the same positive constant, $z = \frac{1.06x - 1.06\bar{x}}{1.06s} = \frac{x-\bar{x}}{s}$. (d) This final plan is recommended because it allows the teacher to set the mean (84) and standard deviation (4) without changing the overall position of the students.

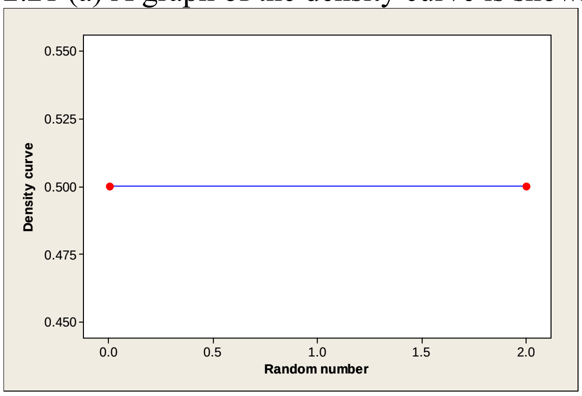
2.19 (a) Erik had a relatively good race compared the other athletes who completed the state meet, but had a poor race by his own standards. (b) Erica was only a bit slower than usual by her own standards, but she was relatively slow compared to the other swimmers at the state meet.

2.20 (a) The density curve is shown below.



The area under the density curve is equal to the area of A + B + C = $\frac{1}{2}\times 0.5\times 0.8 + \frac{1}{2}\times 0.5\times 0.8 + 1\times 0.6 = 1$. (b) The median is at $x = 0.5$, and the quartiles are at approximately $x = 0.3$ and $x = 0.7$. (c) The first line segment has an equation of $y = 0.6 + 1.6x$. Thus, the height of the density curve at 0.3 is $0.6 + 1.6\times 0.3 = 1.08$. The total area under the density curve between 0 and 0.3 is $\frac{1}{2}\times 0.3\times 0.48+0.3\times 0.6=0.252$. Thus, 25.2% of the observations lie below 0.3. (d) Using symmetry of the density curve, the area between 0.3 and 0.7 is $1 - 2\times 0.252 = 0.496$. Therefore, 49.6% of the observations lie between 0.3 and 0.7.

2.21 (a) A graph of the density curve is shown below.

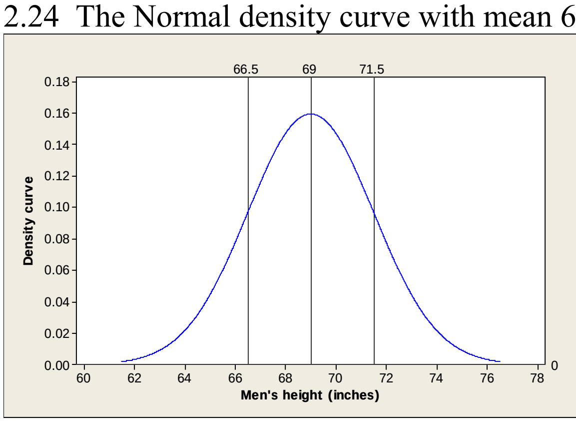


(b) The proportion of outcomes less than 1 is $1\times\frac{1}{2}=\frac{1}{2}$. (c) Using the symmetry of the distribution, it is easy to see that median = mean = 1, $Q_1 = 0.5$, $Q_3 = 1.5$. (d) The proportion of outcomes that lie between 0.5 and 1.3 is $0.8\times\frac{1}{2}=0.4$.

2.22 (a) Outcomes from 18 to 32 are likely, with outcomes near 25 being more likely. The most likely outcome is 25. (d) The distribution should be roughly symmetric with a single peak around 25 and a standard deviation of about 3.54. There should be no gaps or outliers. The normal density curve should fit this distribution well.

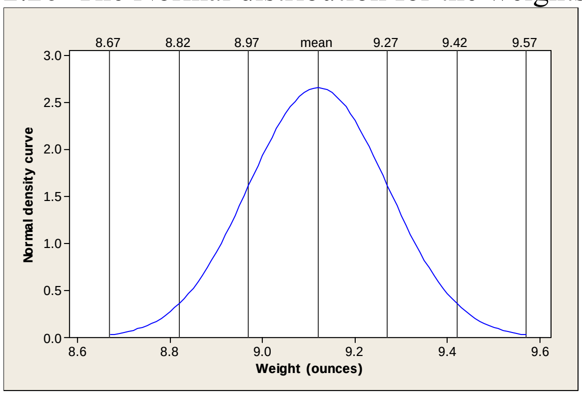
2.23 The standard deviation is approximately 0.2 for the tall, more concentrated one and 0.5 for the short, less concentrated one.

2.24 The Normal density curve with mean 69 and standard deviation 2.5 is shown below.



2.25 (a) Approximately 2.5% of men are taller than 74 inches, which is 2 standard deviations above the mean. (b) Approximately 95% of men have heights between 69−5=64 inches and 69+5=74 inches. (c) Approximately 16% of men are shorter than 66.5 inches, because 66.5 is one standard deviation below the mean. (d) The value 71.5 is one standard deviation above the mean. Thus, the area to the left of 71.5 is the 0.68 + 0.16 = 0.84. In other words, 71.5 is the 84th percentile of adult male American heights.

2.26 The Normal distribution for the weights of 9-ounce bags of potato chips is shown below.



The interval containing weights within 1 standard deviation of the mean goes from 8.97 to 9.27. The interval containing weights within 2 standard deviations of the mean goes from 8.82 to 9.42. The interval containing weights within 3 standard deviations of the mean goes from 8.67 to 9.57. (b) A bag weighing 8.97 ounces, 1 standard deviation below the mean, is at the 16th percentile. (c) We need the area under a Normal curve from 3 standard deviations below the mean to 1 standard above the mean. Using the 68−95−99.7 Rule, the area is equal to $0.68+\frac{1}{2}(0.95-0.68)+\frac{1}{2}(0.997-0.95)=0.8385$, so about 84% of 9-ounce bags of these potato chips weigh between 8.67 ounces and 9.27 ounces.

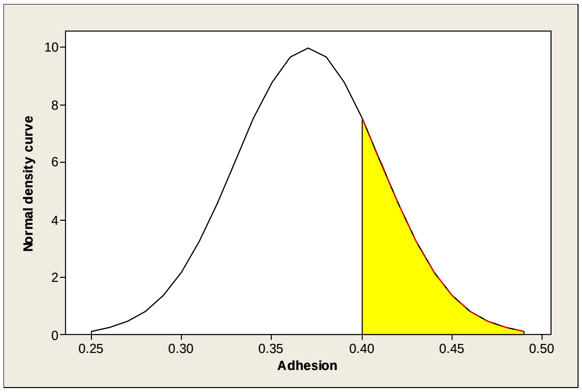
2.27 Answers will vary, but the observed percents should be close to 68%, 95%, and 99.7%.

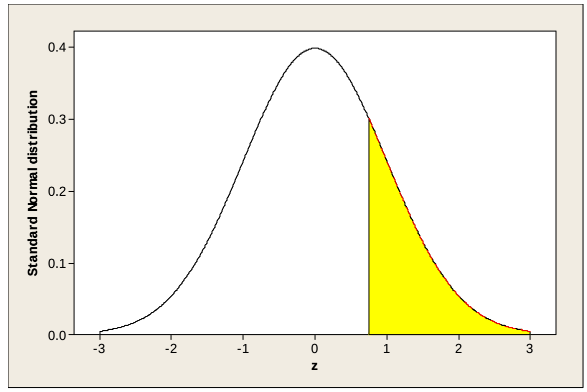
2.28 Answers will differ slightly from 68%, 95%, and 99.7% because of natural variation from trial to trial.

2.29 (a) 0.9978 (b) 1 − 0.9978 = 0.0022 (c) 1 – 0.0485 = 0.9515 (d) 0.9978 – 0.0485 = 0.9493

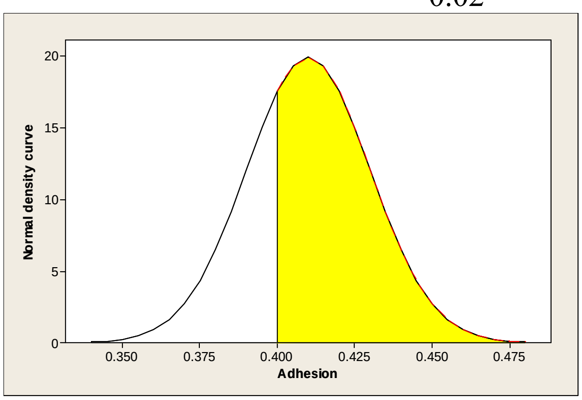
2.30 (a) 0.0069 (b) 1 − 0.9931 = 0.0069 (c) 0.9931 − 0.8133 = 0.1798 (d) 0.1020 − 0.0016 = 0.1004

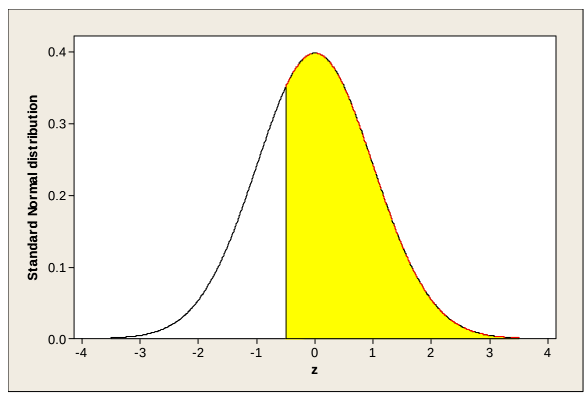
2.31 (a) We want to find the area under the N(0.37, 0.04) distribution to the right of 0.4. The graphs below show that this area is equivalent to the area under the N(0, 1) distribution to the right of $z=\frac{0.4-0.37}{0.04}=0.75$.





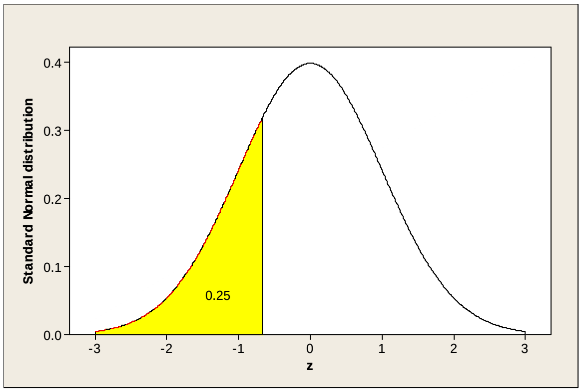
Using Table A, the proportion of adhesions higher than 0.40 is 1 − 0.7734 = 0.2266. (b) We want to find the area under the N(0.37, 0.04) distribution between 0.4 and 0.5. This area is equivalent to the area under the N(0, 1) distribution between $z = \dfrac{0.4 - 0.37}{0.04} = 0.75$ and $z = \dfrac{0.5 - 0.37}{0.04} = 3.25$. (Note: New graphs are not shown, because they are almost identical to the graphs above. The shaded region should end at 0.5 for the graph on the left and 3.25 for the graph on the right.) Using Table A, the proportion of adhesions between 0.4 and 0.5 is 0.9994 − 0.7734 = 0.2260. (c) Now, we want to find the area under the N(0.41, 0.02) distribution to the right of 0.4. The graphs below show that this area is equivalent to the area under the N(0, 1) distribution to the right of $z = \dfrac{0.4 - 0.41}{0.02} = -0.5$.

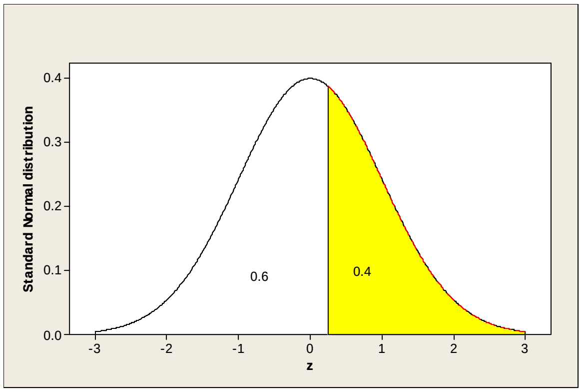




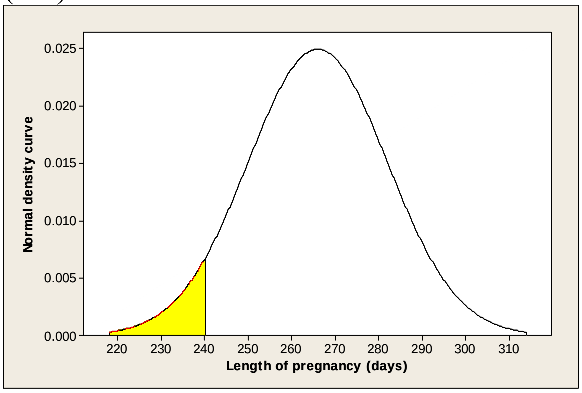
Using Table A, the proportion of adhesions higher than 0.40 is 1 − 0.3085 = 0.6915. The area under the N(0.41, 0.02) distribution between 0.4 and 0.5 is equivalent to the area under the N(0, 1) distribution between $z = \dfrac{0.4 - 0.41}{0.02} = -0.5$ and $z = \dfrac{0.5 - 0.41}{0.02} = 4.5$. Using Table A, the proportion of adhesions between 0.4 and 0.5 is 1 − 0.3085 = 0.6915. The proportions are the same because the upper end of the interval is so far out in the right tail.

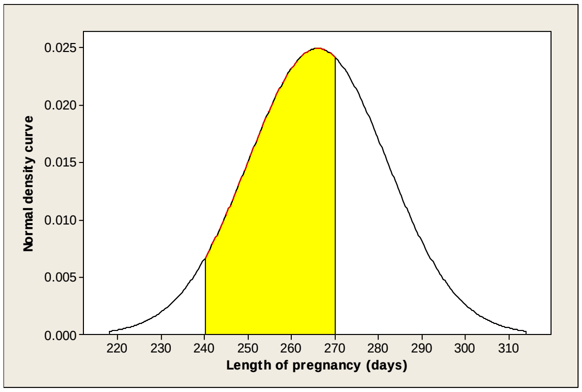
2.32 (a) The closest value in Table A is −0.67. The 25th percentile of the N(0, 1) distribution is −0.67449. (b) The closest value in Table A is 0.25. The 60th percentile of the N(0, 1), distribution is 0.253347. See the graphs below.



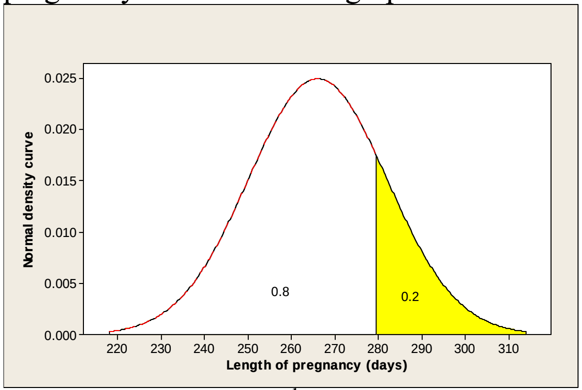


2.33 (a) The proportion of pregnancies lasting less than 240 days is shown in the graph below (left).





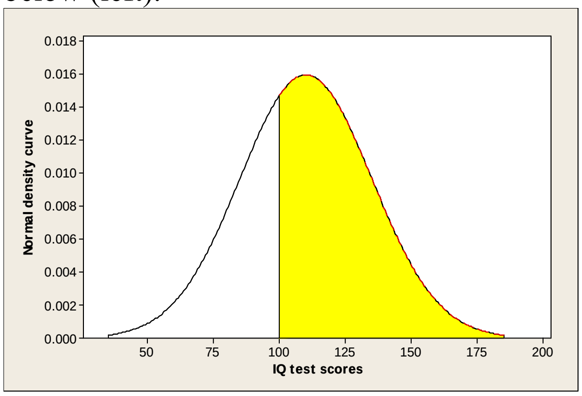
The shaded area is equivalent to the area under the N(0, 1) distribution to the left of $z = \frac{240-266}{16} \doteq -1.63$, which is 0.0516 or about 5.2%. (b) The proportion of pregnancies lasting between 240 and 270 days is shown in the graph above (right). The shaded area is equivalent to the area under the N(0, 1) distribution between $z = -1.63$ and $z = \frac{270-266}{16} = 0.25$, which is $0.5987 - 0.0516 = 0.5471$ or about 55%. (c) The 80th percentile for the length of human pregnancy is shown in the graph below.

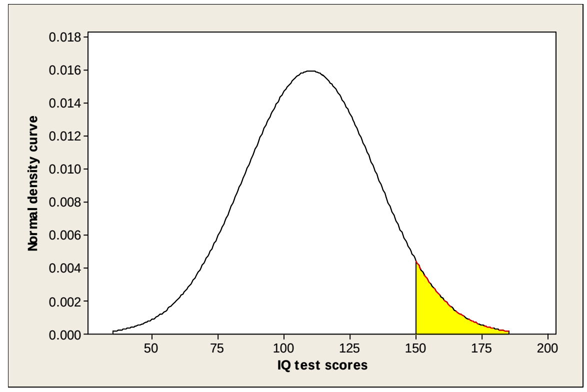


Using Table A, the 80th percentile for the standard Normal distribution is 0.84. Therefore, the 80th percentile for the length of human pregnancy can be found by solving the equation

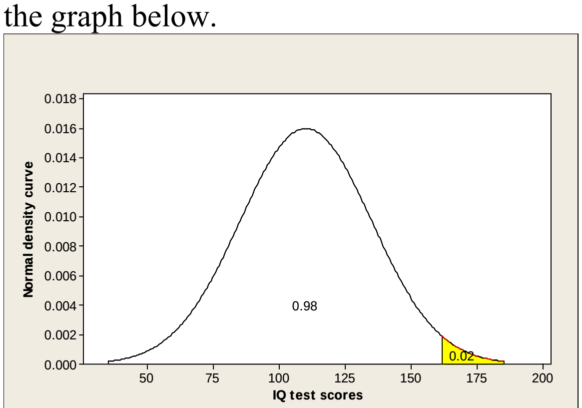
$0.84 = \dfrac{x - 266}{16}$ for x. Thus, $x = 0.84 \times 16 + 266 = 279.44$. The longest 20% of pregnancies last approximately 279 or more days.

2.34 (a) The proportion of people aged 20 to 34 with IQ scores above 100 is shown in the graph below (left).





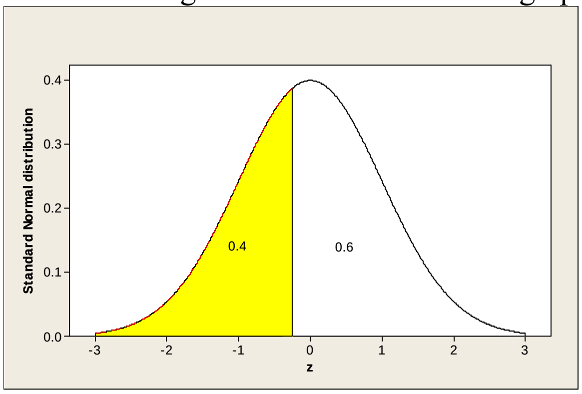
The shaded area is equivalent to the area under the N(0, 1) distribution to the right of $z = \dfrac{100 - 110}{25} = -0.4$, which is 1 − 0.3446 = 0.6554 or about 65.54%. (b) The proportion of people aged 20 to 34 with IQ scores above 150 is shown in the graph above (right). The shaded area is equivalent to the area under the N(0, 1) distribution to the right of $z = \dfrac{150 - 110}{25} = 1.6$, which is 1 − 0.9452 = 0.0548 or about 5.5%. (c) The 98^{th} percentile of the IQ scores is shown in the graph below.

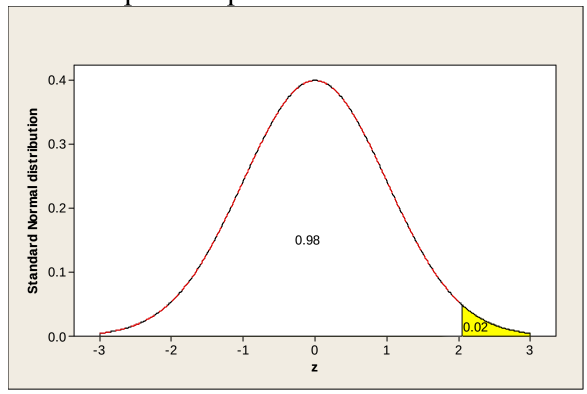


Using Table A, the 98^{th} percentile for the standard Normal distribution is closest to 2.05. Therefore, the 80^{th} percentile for the IQ scores can be found by solving the equation $2.05 = \dfrac{x - 110}{25}$ for x. Thus, $x = 2.05 \times 25 + 110 = 161.25$. In order to qualify for MENSA membership a person must score 162 or higher.

2.35 (a) The quartiles of a standard Normal distribution are at ± 0.675. (b) Quartiles are 0.675 standard deviations above and below the mean. The quartiles for the lengths of human pregnancies are 266 ± 0.675(16) or 255.2 days and 276.8 days.

2.36 Use the given information and the graphs below to set up two equations in two unknowns.



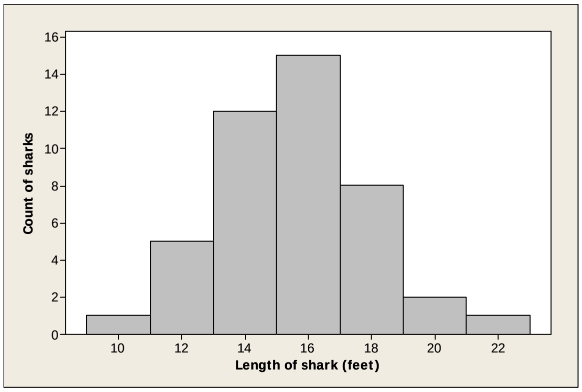


The two equations are $-0.25 = \frac{1-\mu}{\sigma}$ and $2.05 = \frac{2-\mu}{\sigma}$. Multiplying both sides of the equations by σ and subtracting yields $-2.3\sigma = -1$ or $\sigma = \frac{1}{2.3} \doteq 0.4348$ minutes. Substituting this value back into the first equation we obtain $-0.25 = \frac{1-\mu}{0.4348}$ or $\mu = 1 + 0.25 \times 0.4348 \doteq 1.1087$ minutes.

2.37 Small and large percent returns do not fit a Normal distribution. At the low end, the percent returns are smaller than expected, and at the high end the percent returns are slightly larger than expected for a Normal distribution.

2.38 The shape of the quantile plot suggests that the data are right-skewed. This can be seen in the flat section in the lower left—these numbers were less spread out than they should be for Normal data—and the three apparent outliers that deviate from the line in the upper right; these were much larger than they would be for a Normal distribution.

2.39 (a) *Who?* The individuals are great white sharks. *What?* The quantitative variable of interest is the length of the sharks, measured in feet. *Why?* Researchers are interested in the size of great white sharks. *When, where, how, and by whom?* These questions are impossible to answer based on the information provided. *Graphs:* A histogram and stemplot are provided below.

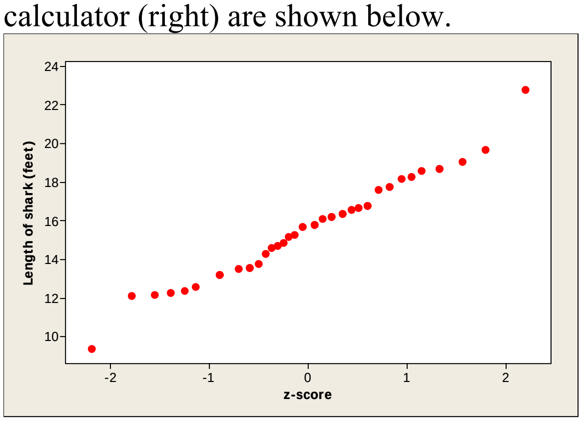


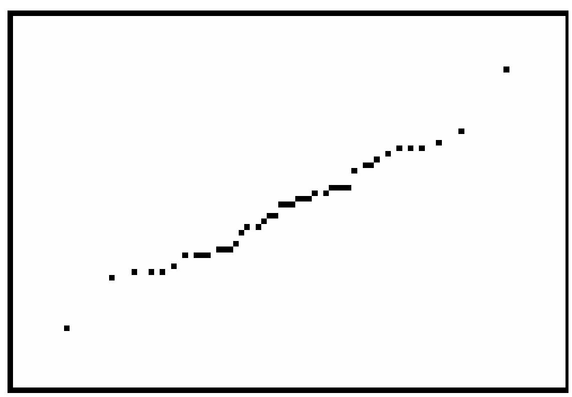

```
Stem-and-leaf of shlength  N  = 44
Leaf Unit = 0.10

 1   9   4
 1   10
 1   11
 6   12  12346
 14  13  22225668
 18  14  3679
(6)  15  237788
 20  16  122446788
 11  17  688
 8   18  23677
 3   19  17
 1   20
 1   21
 1   22  8
```

Numerical Summaries: Descriptive statistics are provided below.

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
shlength	44	15.586	2.550	9.400	13.525	15.750	17.400	22.800

Interpretation: The distribution of shark lengths is roughly symmetric with a peak at 16 and a spread from 9.4 feet to 22.8 feet.
(b) The mean is 15.586 and the median is 15.75. These two measures of center are very close to one another, as expected for a symmetric distribution. (c) Yes, the distribution is approximately normal—68.2% of the lengths fall within one standard deviation of the mean, 95.5% of the lengths fall within two standard deviations of the mean, and 100% of the lengths fall within 3 standard deviations of the mean. (d) Normal probability plots from Minitab (left) and a TI calculator (right) are shown below.



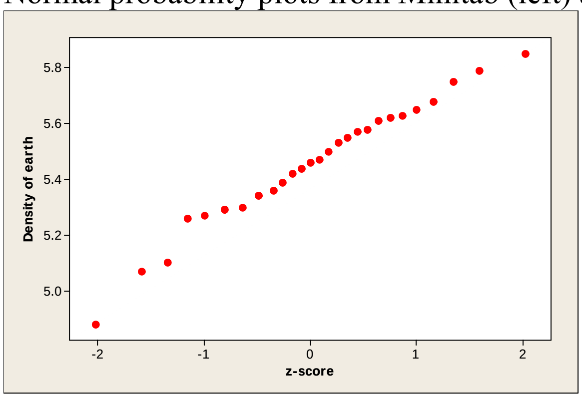


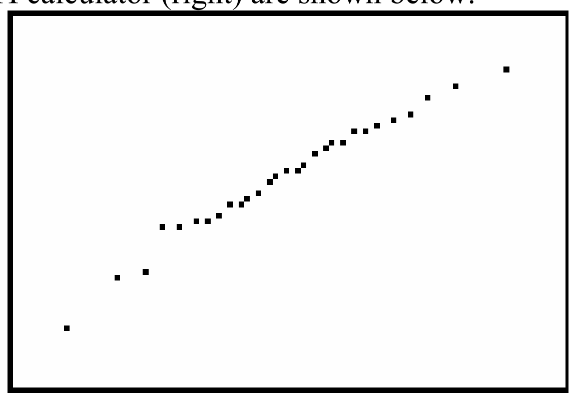
Except for one small shark and one large shark, the plot is fairly linear, indicating that the Normal distribution is appropriate. (e) The graphical displays in (a), comparison of two measures of center in (b), check of the 68−95−99.7 rule in (c), and Normal probability plot in (d) indicate that shark lengths are approximately Normal.

2.40 (a) A stemplot is shown below. The distribution is roughly symmetric.

```
Stem-and-leaf of density  N  = 29
Leaf Unit = 0.010

 1   48  8
 1   49
 2   50  7
 3   51  0
 7   52  6799
 12  53  04469
(4)  54  2467
 13  55  03578
 8   56  12358
 3   57  59
 1   58  5
```

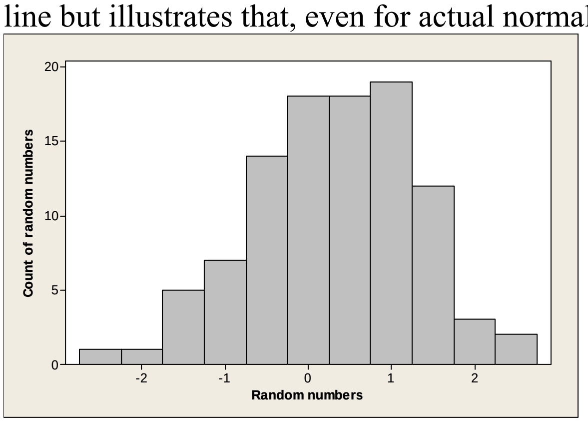
(b) The mean is $\bar{x} = 5.4479$ and the standard deviation is $s = 0.2209$. The densities follow the 68−95−99.7 rule closely—75.86% (22 out of 29) of the densities fall within one standard deviation of the mean, 96.55% (28 out of 29) of the densities fall within two standard deviations of the mean, and 100% of the densities fall within 3 standard deviations of the mean. (c) Normal probability plots from Minitab (left) and a TI calculator (right) are shown below.

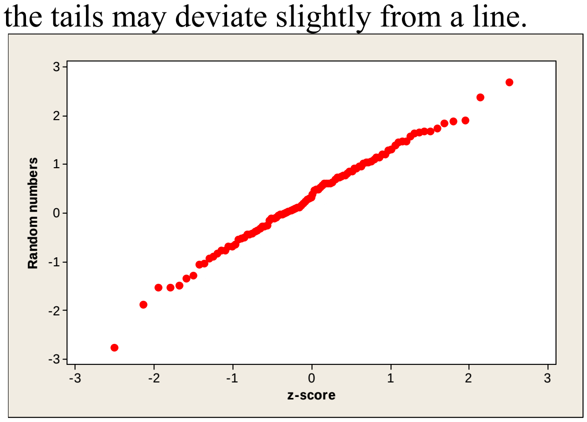




Yes, the Normal probability plot is roughly linear, indicating that the densities are approximately Normal.

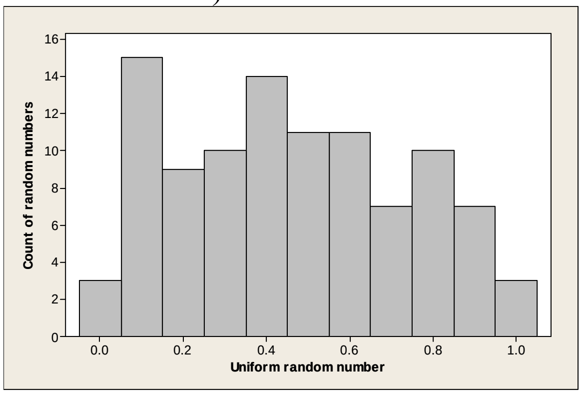
2.41 (a) A histogram from one sample is shown below. Histograms will vary slightly but should suggest a bell curve. (b) The Normal probability plot below shows something fairly close to a line but illustrates that, even for actual normal data, the tails may deviate slightly from a line.

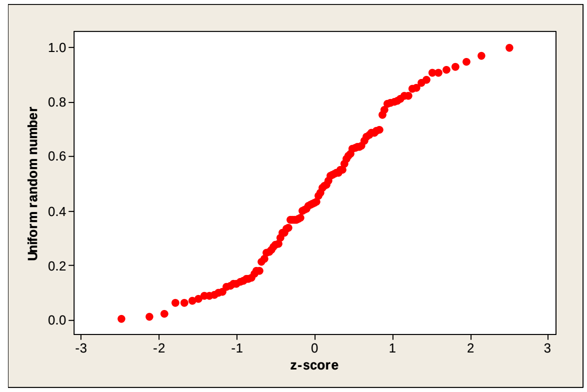




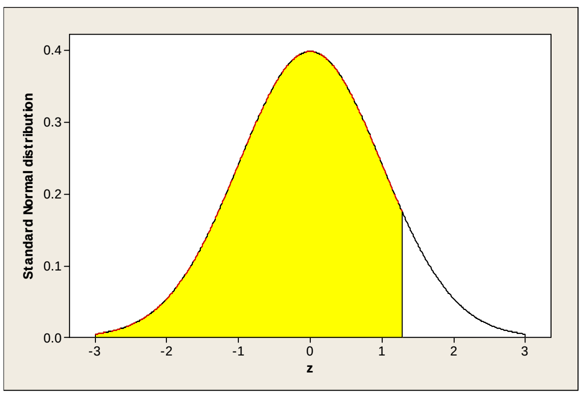
2.42 (a) A histogram from one sample is shown below. Histograms will vary slightly but should suggest the density curve of Figure 2.8 (but with more variation than students might expect).

(b) The Normal probability plot below shows that, compared to a normal distribution, the uniform distribution does not extend as low or as high (not surprising, since all observations are between 0 and 1).

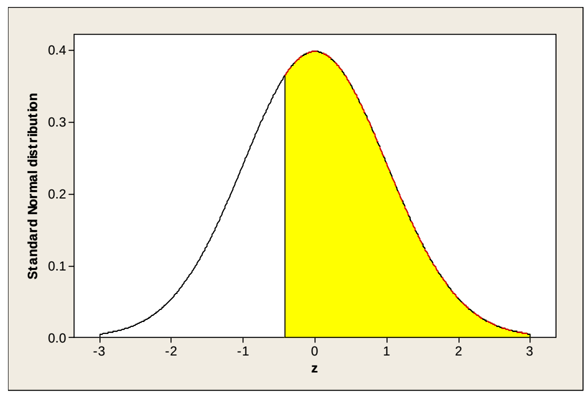




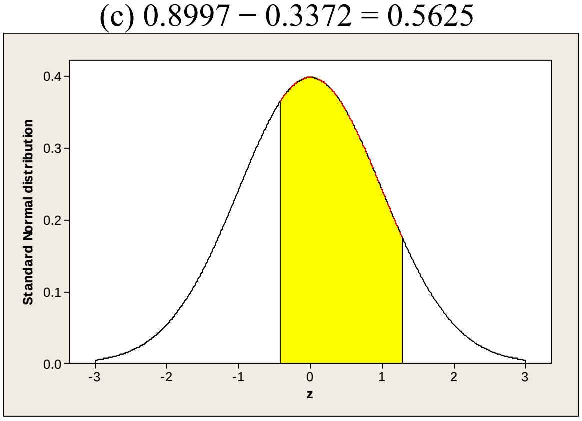
2.43 (a) 0.8997



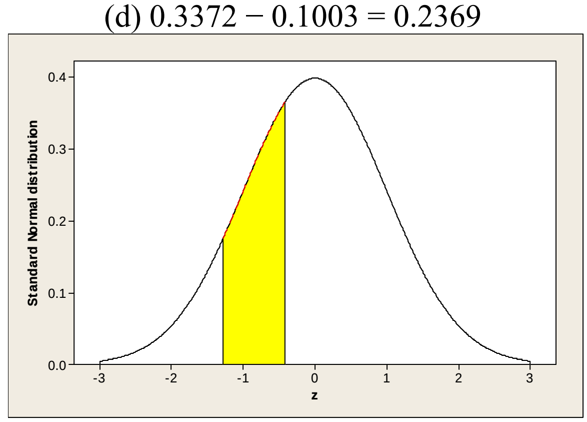
(b) $1 - 0.3372 = 0.6628$



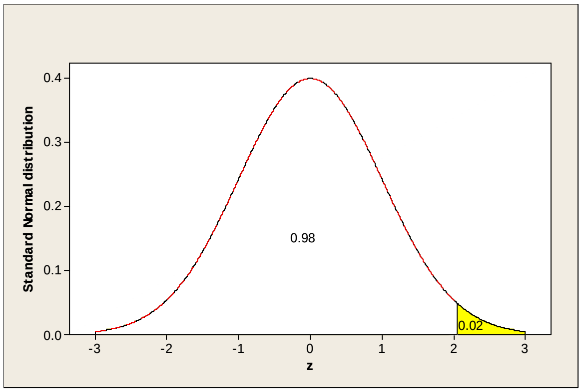
(c) $0.8997 - 0.3372 = 0.5625$

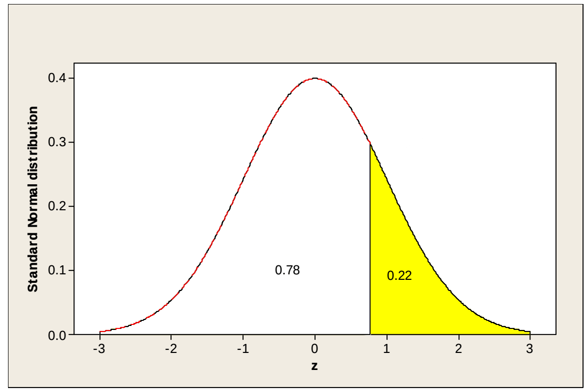


(d) $0.3372 - 0.1003 = 0.2369$

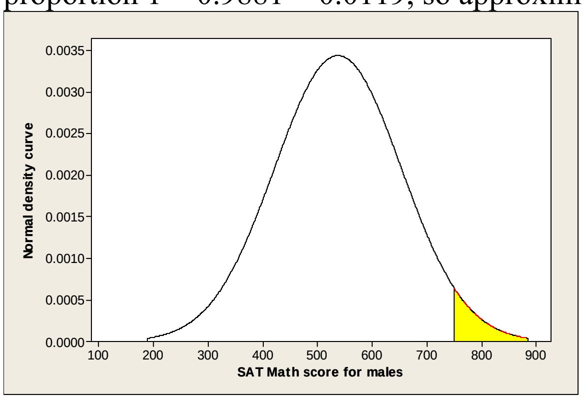


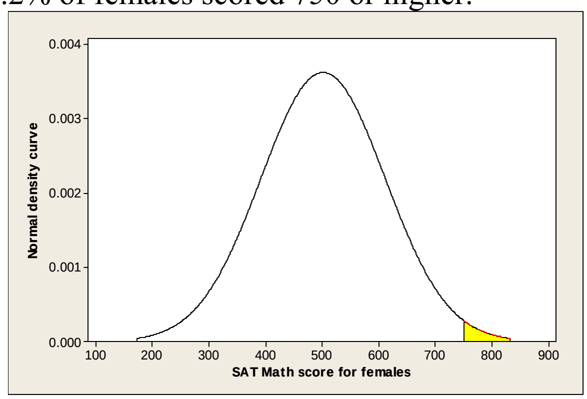
2.44 (a) Using Table A, the closest value to the 98th percentile is 2.05. (b) Using Table A, the closest value to the 78th percentile is 0.77.



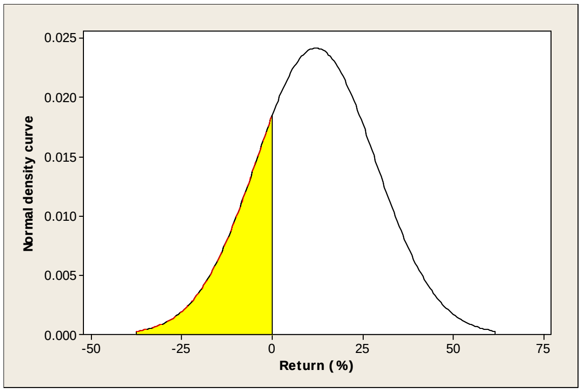


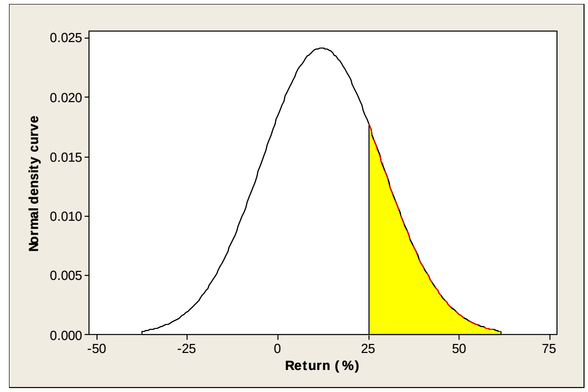
2.45 **(a)** To find the shaded area below for men, standardize the score of 750 to obtain the z-score of $z = \dfrac{750-537}{116} \doteq 1.84$. Table A gives the proportion 1 − 0.9671 = 0.0329, so approximately 3.3% of males scored 750 or higher. (b) For women, the shaded area below corresponds to getting a standardized score greater than $z = \dfrac{750-501}{110} \doteq 2.26$. Table A gives the proportion 1 − 0.9881 = 0.0119, so approximately 1.2% of females scored 750 or higher.





2.46 (a) According to the 68−95−99.7 rule, the middle 95% of all yearly returns are between 12 − 2×16.5 = −21% and 12 + 2×16.5 = 45%. (b) To find the shaded area below zero, indicated on the figure below, standardize 0 to obtain the z-score of $z = \dfrac{0-12}{16.5} \doteq -0.73$. Table A gives the proportion 0.2327 (software gives 0.233529). (c) To find the shaded area above 25%, indicated on the figure below, standardize 25 to obtain the z-score of $z = \dfrac{25-12}{16.5} \doteq 0.79$. Table A gives the proportion 1 − 0.7852 = 0.2148 (software gives 0.215384).





2.47 (a) Using Table A, the closest values to the deciles are ±1.28. (b) The deciles for the heights of young women are 64.5 ± 1.28×2.5 or 61.3 inches and 67.7 inches.

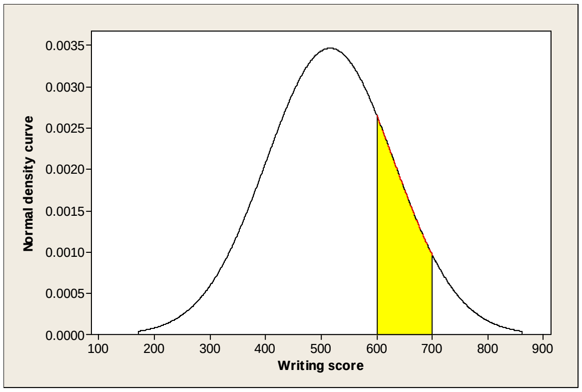
2.48 The quartiles for a standard Normal distribution are ±0.6745. For a $N(\mu,\sigma)$ distribution, $Q_1 = \mu - 0.6745\sigma$, $Q_3 = \mu + 0.6745\sigma$, and $IQR = 1.349\sigma$. Therefore, $1.5 \times IQR = 2.0235\sigma$, and the suspected outliers are below $Q_1 - 1.5 \times IQR = \mu - 2.698\sigma$ or above $Q_3 + 1.5 \times IQR = \mu + 2.698\sigma$. The proportion outside of this range is approximately the same as the area under the standard Normal distribution outside of the range from −2.7 to 2.7, which is 2 × 0.0035 = 0.007 or 0.70%.

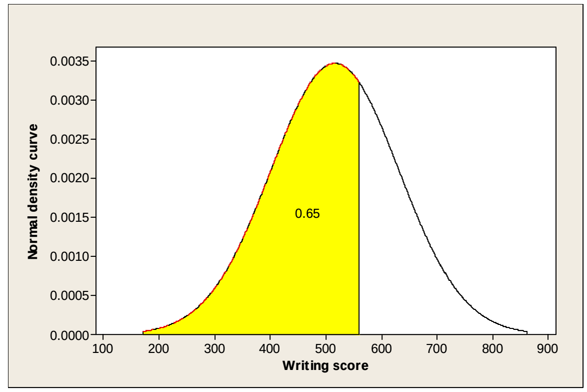
2.49 The plot is nearly linear. Because heart rate is measured in whole numbers, there is a slight “step” appearance to the graph.

2.50 Women’s weights are skewed to the right: This makes the mean higher than the median, and it is also revealed in the differences $M - Q_1 = 133.2 - 118.3 = 14.9$ pounds and $Q_3 - M = 157.3 - 133.2 = 24.1$ pounds.

CASE CLOSED!

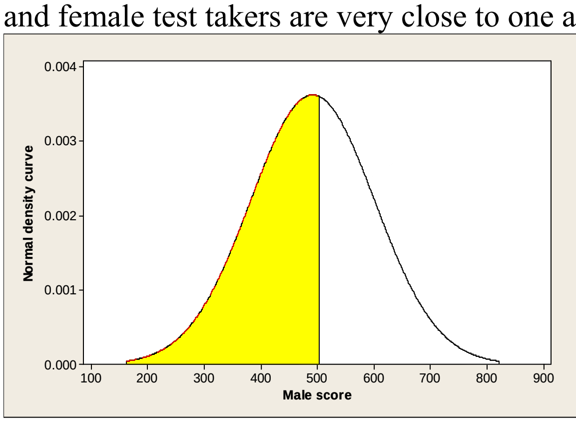
1. (a) The proportion of students who earned between 600 and 700 on the Writing section is shown below (left). Standardizing both scores yields *z*-scores of $z = \dfrac{600-516}{115} \doteq 0.73$ and $z = \dfrac{700-516}{115} = 1.6$. Table A gives the proportion 0.9452 − 0.7673 = 0.1779 or about 18%.

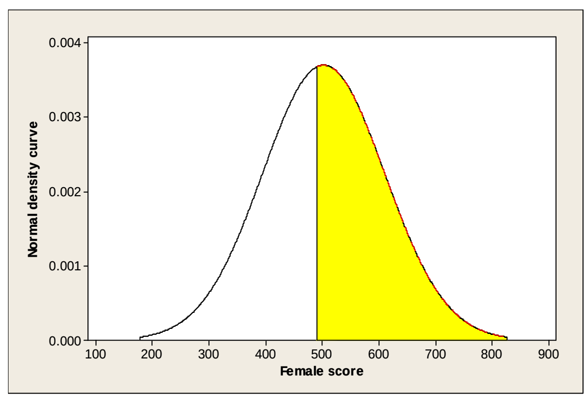




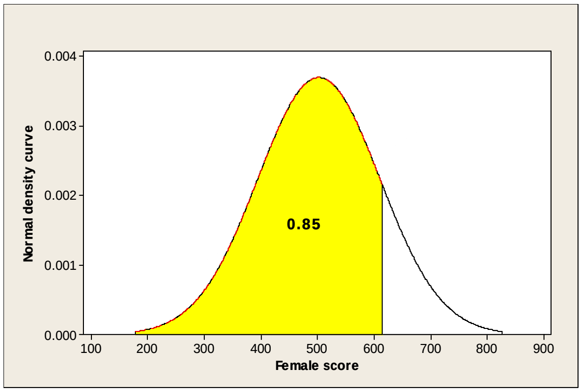
(b) The 65th percentile is shown above (right). Using Table A, the 65th percentile of a standard Normal distribution is closest to 0.39, so the 65th percentile for Writing score is $516 + 0.39 \times 115 = 560.85$.

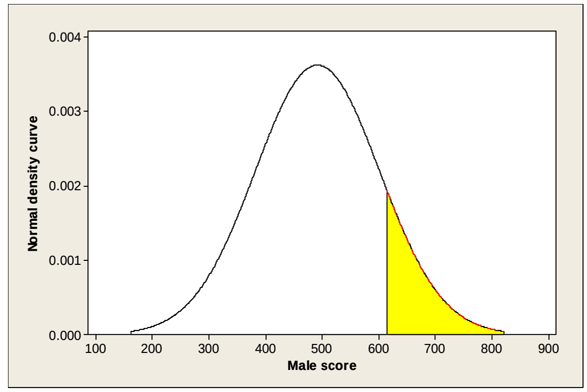
2. (a) The proportion of male test takers who earned scores below 502 is shown below (left). Standardizing the score yields a *z*-score of $z = \dfrac{502-491}{110} = 0.10$. Table A gives the proportion 0.5398 or about 54%. (b) The proportion of female test takers who earned scores above 491 is shown below (right). Standardizing the score yields a *z*-score of $z = \dfrac{491-502}{108} \doteq -0.10$. Table A gives the proportion $1 - 0.4602 = 0.5398$ or about 54%. (Minitab gives 0.5406.) The probabilities in (a) and (b) are almost exactly the same because the standard deviations for male and female test takers are very close to one another.



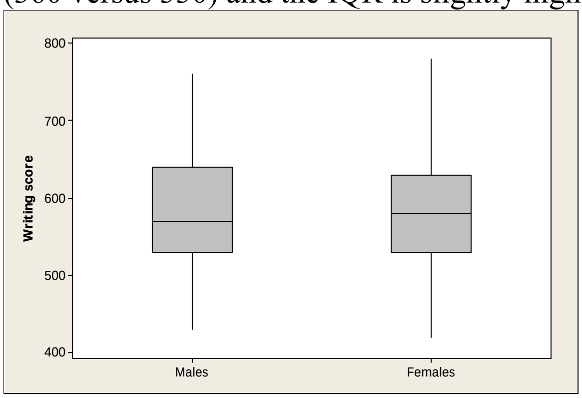


(c) The 85th percentile for the female test takers is shown below (left). Using Table A, the 85th percentile of the standard Normal distribution is closest to 1.04, so the 85th percentile for the female test takers is $502 + 1.04 \times 108 \doteq 614$. The proportion of male test takers who score above 614 is shown below (right). Standardizing the score yields a *z*-score of $z = \dfrac{614-491}{110} \doteq 1.12$. Table A gives the proportion $1 - 0.8686 = 0.1314$ or about 13%.



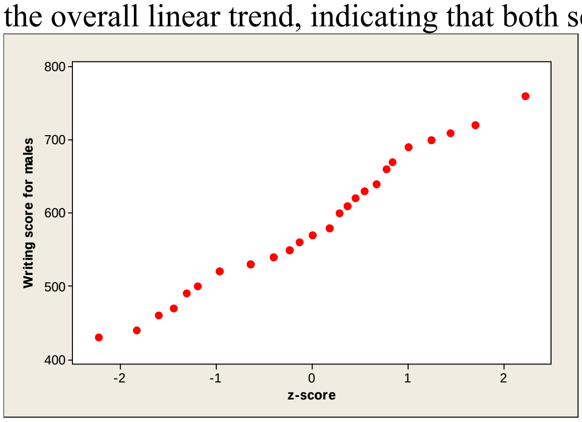


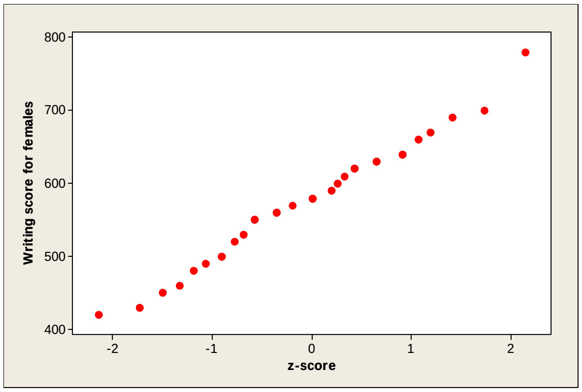
3. (a) The boxplot below shows that the distributions of scores for males and females are very similar. Both distributions are roughly symmetric with no outliers. The median for the females (580) is slightly higher than the median for the males (570). The range is higher for females (360 versus 330) and the IQR is slightly higher for males (110 versus 100).



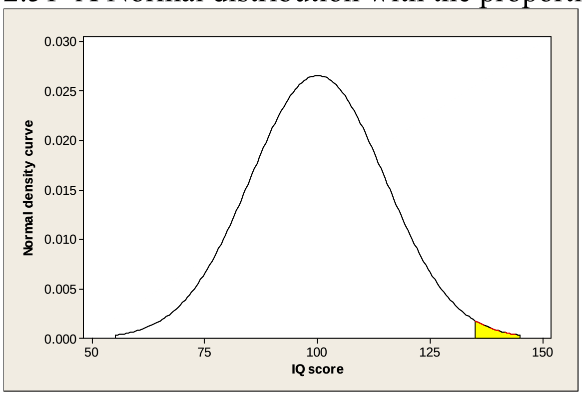
Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Males	48	584.6	80.1	430.0	530.0	570.0	640.0	760.0
Females	39	580.0	78.6	420.0	530.0	580.0	630.0	780.0

The mean for the males (584.6) is slightly higher than the mean for the females (580.0), but the overall performance for males and females is about the same at this school. (b) The students at this private school did much better than the overall national mean (516). There is also much less variability in the scores at this private school than the national scores. (c) Normal probability plots for the males and females are shown below. Both plots show only slight departures from the overall linear trend, indicating that both sets of scores are approximately Normal.



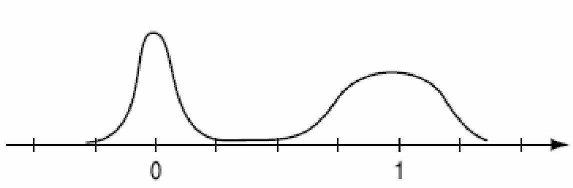


2.51 A Normal distribution with the proportion of "gifted" students is shown below.



A WISC score of 135 corresponds to a standardized score of $z = \frac{135-100}{15} \doteq 2.33$. Using Table A, the proportion of "gifted" students is 1 − 0.9901 = 0.0099 or .99%. Therefore, 0.0099×1300=12.87 or about 13 students in this school district are classified as gifted.

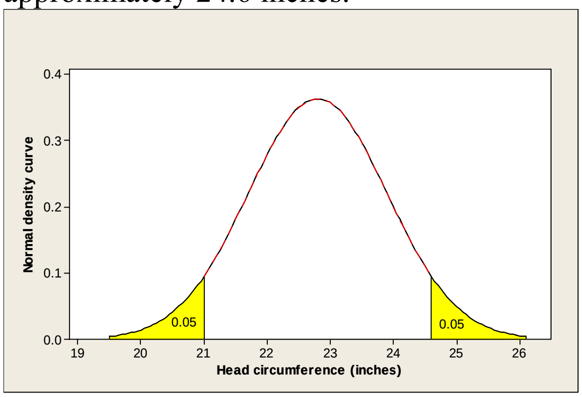
2.52 Sketches will vary, but should be some variation on the one shown below: The peak at 0 should be "tall and skinny," while near 1, the curve should be "short and fat."



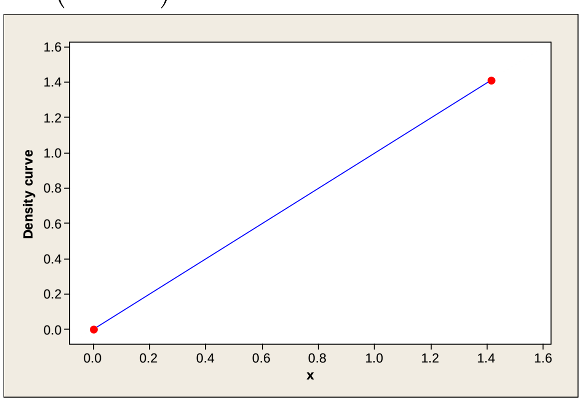
2.53 The percent of actual scores at or below 27 is $\frac{1052490}{1171460} \times 100 \doteq 89.84\%$. A score of 27 corresponds to a standard score of $z = \frac{27-20.9}{4.8} \doteq 1.27$. Table A indicates that 89.8% of scores in a Normal distribution would fall below this level. Based on these calculations, the Normal distribution does appear to describe the ACT scores well.

2.54 (a) Joey's scoring "in the 97th percentile" on the reading test means that Joey scored as well as or better than 97% of all students who took the reading test and scored worse than about 3%. His scoring in the 72nd percentile on the math portion of the test means that he scored as well as or better than 72% of all students who took the math test and worse than about 28%. That is, Joey did better on the reading test, relative to his peers, than he did on the math test. (b) If the test scores are Normal, then the *z*-scores would be 1.88 and for the 97th percentile and 0.58 for the 72nd percentile. However, nothing is stated about the distribution of the scores and we do not have the scores to assess normality.

2.55 The head sizes that need custom-made helmets are shown below. The 5th and 95th percentiles for the standard Normal distribution are ±1.645. Thus, the 5th and 95th percentiles for soldiers' head circumferences are $22.8 \pm 1.645 \times 1.1$. Custom-made helmets will be needed for soldiers with head circumferences less than approximately 21 inches or greater than approximately 24.6 inches.



2.56 (a) The density curve is shown below. The coordinates of the right endpoint of the segment are $\left(\sqrt{2},\sqrt{2}\right)$.



(b) To find the median *M*, set the area of the appropriate triangle $\left(\frac{1}{2} base \times height\right)$ equal to 0.5 and solve. That is, solve the equation $\frac{1}{2}M \times M = \frac{1}{2}$ for *M*. Thus, *M* = 1. The same approach yields $Q_1 = \frac{1}{\sqrt{2}} \doteq 0.707$ and $Q_3 = \sqrt{\frac{3}{2}} \doteq 1.225$. (c) The mean will be slightly below the median of 1 because the density curve is skewed left. (d) The proportion of observations below 0.5 is 0.5×0.5×0.5=0.125 or 12.5%. None (0%) of the observations are above 1.5.

2.57 (a) The mean $\bar{x} = \$17,776$ is greater than the median M = \$15,532. Meanwhile, $M - Q_1 = \$5,632$ and $Q_3 - M = \$6,968$, so Q_3 is further from the median than Q_1. Both of these comparisons result in what we would expect for right-skewed distributions. (b) From Table A, we estimate that the third quartiles of a Normal distribution would be 0.675 standard deviations above the mean, which would be \$17,776 + 0.675 × \$12,034 ≐ \$25,899. (Software gives 0.6745, which yields \$25,893.) As the exercise suggests, this quartile is larger than the actual value of Q_3.

2.58 (a) About 0.6% of healthy young adults have osteoporosis (the area below a standard *z*-score of −2.5 is 0.0062). (b) About 31% of this population of older women has osteoporosis: The BMD level that is 2.5 standard deviations below the young adult mean would standardize to −0.5 for these older women, and the area to the left of this standard *z*-score is 0.3085.

2.59 (a) Except for one unusually high value, these numbers are reasonably Normal because the other points fall close to a line. (b) The graph is almost a perfectly straight line, indicating that the data are Normal. (c) The flat portion at the bottom and the bow upward indicate that the distribution of the data is right-skewed data set with several outliers. (d) The graph shows 3 clusters or mounds (one at each end and another in the middle) with a gap in the data towards the lower values. The flat sections in the lower left and upper right illustrate that the data have peaks at the extremes.

2.60 If the distribution is Normal, it must be symmetric about its mean—and in particular, the 10th and 90th percentiles must be equal distances below and above the mean—so the mean is 250 points. If 225 points below (above) the mean is the 10th (90th) percentile, this is 1.28 standard deviations below (above) the mean, so the distribution's standard deviation is $\frac{225}{1.28} \doteq 175.8$ points.

2.61 Use window of $X[55,145]_{15}$ and $Y[-0.008, 0.028]_{.01}$. (a) The calculator command shadeNorm(135,1E99,100,15) produces an area of 0.009815. About .99% of the students earn WISC scores above 135. (b) The calculator command shadeNorm(-1E99,75,100,15) produces an area of 0.04779. About 4.8% of the students earn WISC scores below 75. (c) shadeNorm(70,130,100,15) = 0.9545. Also, 1 – 2(shadeNorm(-1E99,70,100,15)) = 0.9545.

2.62 The calculator command normalcdf (−1E99, 27, 20.9, 4.8) produces an area of 0.89810596 or 89.81%, which agrees with the value obtained in Exercise 2.53.

2.63 The calculator commands invNorm(.05,22.8,1.1) = 20.99 and invNorm(.95,22.8,1.1) = 24.61 agree with the values obtained in Exercise 2.55.