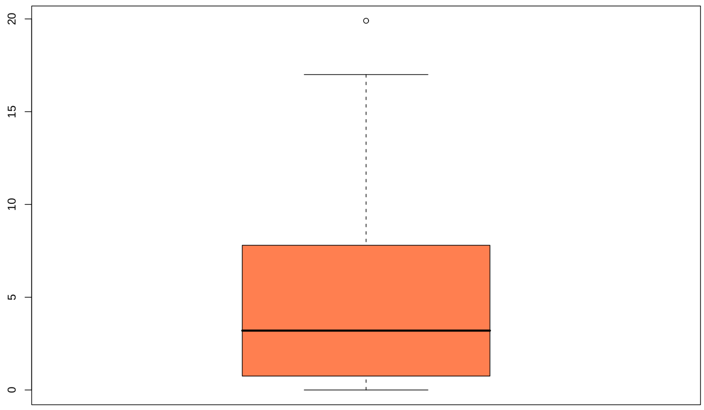
Homework 2 solutions - 85 total points

STAT 201-502

Lecture 2 & 3 material

1 Table 1.6 on page 33 of the text book gives CO_2 emmissions per person from countries with populations over 20 million.

a. Find the 5-number summary for this data set and use it to make a boxplot. Does the distribution look right-skewed?



5 number summary is Min=0, Q_1=.75, M=3.2, Q_3=7.8, Max=17 with one extreme outlier=19.9.

b. Are there any suspected outlier? If so, make sure to account for them on your boxplot. IQR=7.8-.75=7.05 so suspected outliers if $X < Q_1 - 1.5(IQR) = -9.83$ *or* $X > Q_3 + 1.5(IQR) = 18.38$. There is one outlier, USA emmissions per person is 19.9 which is greater than 18.38.

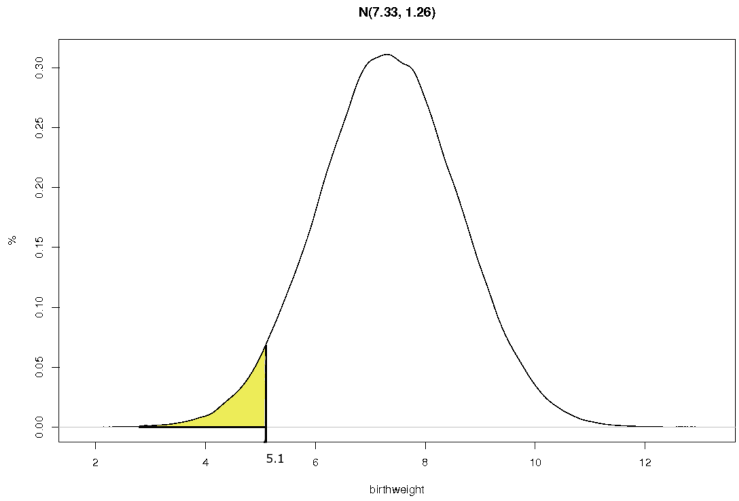
c. Find the mean CO_2 emmisions per person. Explain why the mean and median differ so greatly for this distribution. $\bar{X} = \frac{2.3+3.9+17+0.2+...+0.5}{48} = 4.61$ The mean is greater than the median because this distribution is right-skewed meaning it has more larger values.

2 The CDC reports that in 2003 the weight of a baby at birth is normally distributed with mean 7.33 pounds and standard deviation of 1.26 pounds.

a. One baby is selected at random and it's weight is 9.1 pounds. What is the Z-score for this observation?
$Z = \frac{9.1-7.33}{1.26} = 1.41$

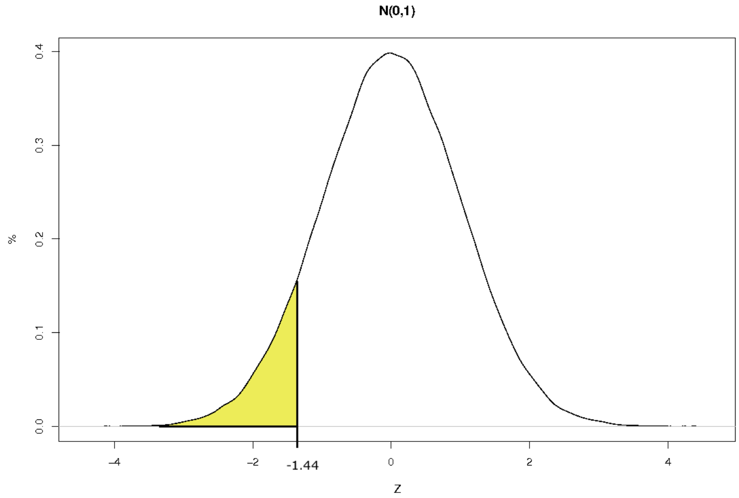
b. What does this Z-score tell us about this particular baby's relationship to the mean?
This baby's weight (9.1 pounds) is 1.41 standard deviations above the mean.

c. Low Birth Weight is considered a newborn baby weighing less than 5.51 pounds. Sketch the normal curve for the distribution of birth weight. Shade the area representing the proportion of babies that have Low Birth Weight.



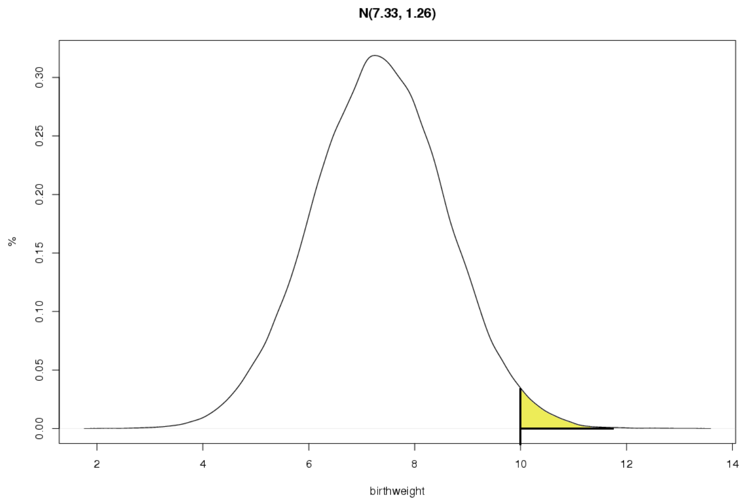
d. Now, find the Z-score for a baby weight of 5.51. What does this Z-score tell us about the relationship between an observed weight of 5.51 pounds and the mean?
$Z = \frac{5.51-7.33}{1.26} = -1.44$ This tells us that the weight 5.51 is 1.44 standard deviations below the mean.

e. Sketch the standard normal curve and shade the area representing the proportion of babies born with Low Birth Weight.



f. Using Table A find this proportion of babies born with Low Birth Weight.
Look for Z=-1.44 gives proportion .0749 = 7.59%

g. If I chose one baby at random, what's the probability that I would get a big, fat, healthy baby weighing more than 10 pounds?



First find the Z-score that corresponds to 10 pounds: $Z = \frac{10-7.33}{1.26} = 2.12$.
The corresponding probability in the table fo Z=2.12 is .9830. We want the proportion greater than 10 not less than 10 so the probability of me choosing a baby at random that is bigger than 10 pounds is 1-.9830=0.017=1.7%

3 Do textbook problems 1.86, 1.98, 1.99, 1.100, 1.101 and 1.106

1.86: Length$\sim$N(266, 16) which is symmetric and mound-shaped so we can use the Empirical Rule.
(a) 95% of values fall within $2s$ of the $\bar{x}$ so 95% of pregancies last between 234 and 298 days.
(b) The shortest 2.5% of pregnancies last less than 234 days. The longest 2.5% of pregnancies last more than 298 days.

For the next 5 problems we know SAT$\sim$ $N(\mu = 1026, \sigma = 209)$ and ACT$\sim$ $N(\mu = 20.8, \sigma = 4.8)$:

1.98: $Z_{SAT} = \frac{1318-\mu_{SAT}}{\sigma_{SAT}} = \frac{1318-1026}{209} = 1.40$ and $Z_{ACT} = \frac{27-\mu_{ACT}}{\sigma_{ACT}} = \frac{27-20.8}{4.8} = 1.29$.

Now we can compare Z_{SAT} to Z_{ACT} and we see that Z_{SAT}=1.40 is higher so Tonya scored higher than Jermaine.

1.99: $Z_{SAT} = \frac{670-\mu_{SAT}}{\sigma_{SAT}} = \frac{670-1026}{209} = -1.70$ and $Z_{ACT} = \frac{16-\mu_{ACT}}{\sigma_{ACT}} = \frac{16-20.8}{4.8} = -1.00$.

Now we can compare Z_{SAT} to Z_{ACT} and we see that Z_{SAT}=-1.00 is higher so Jacob scored higher than Emily.

1.100: $Z_{SAT} = \frac{1287-\mu_{SAT}}{\sigma_{SAT}} = \frac{1287-1026}{209} = 1.25$ A Z-score of 1.25 is equivalent to what ACT score?

$1.5 = \frac{ACT-\mu_{ACT}}{\sigma_{ACT}} = \frac{ACT-20.8}{4.8}$

Now, solve for ACT:
$4.8(1.25) = ACT - 20.8 \rightarrow 6 + 20.8 = ACT \rightarrow ACT = 26.8$.

1.101: Just like 1.100

$Z_m = \frac{28-20.8}{4.8} = 1.5$ and $1.5 = \frac{X_{SAT}-1026}{209} \rightarrow X_{SAT} = 1339.5$

1.106: The top 10% corresponds to the everything above the z-score where $P(Z < Z^*) = .90$. Using Table A, $Z^* = 1.29 = \frac{X_{SAT}-1026}{209} \approx 1295$. So, all SAT scores higher than 1295 represent the top 10% of all scores.