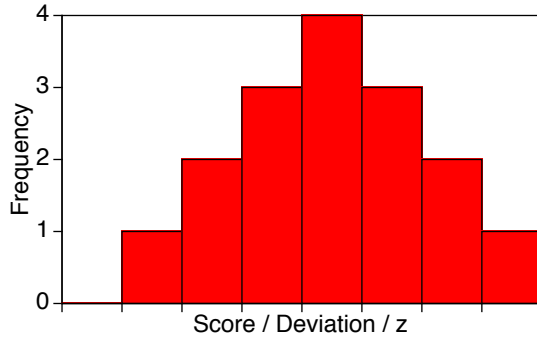# Chapter 6-The Normal Distribution

6.1 Distribution of original values:

For the first distribution the abscissa would take on the values of:
$$1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7$$
For the second distribution the values would be:
$$-3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3$$
For the third distribution the values would be:
$$-1.90 \quad -1.27 \quad -0.63 \quad 0 \quad 0.63 \quad 1.27 \quad 1.90$$

> In these calculations I used the parameters as given, rather than the statistics calculated on the sample.

6.3 Psychology 1 exam grades:

$$z = \frac{X - \mu}{\sigma} = \frac{165 - 195}{30} = -1.0 \quad z = \frac{X - \mu}{\sigma} = \frac{225 - 195}{30} = 1.0$$

a) The percentage between 165 and 225 is the percentage between $z = -1.0$ and $z = 1.0$. This is twice the area between $z = 0$ and $z = 1$. The probability is $2 \times 0.3413 = .6826$.

b) The percentage below 195 is just the percentage below $z = 0 = .500$.

c) The percentage below $z = 1$ is the percentage in the larger portion $= .8413$.

6.5 Guessing on the Psychology 1 exam:

    a) We know the mean and standard deviation if the students guess; they are 75 and 7.5, respectively. We also know that a z score of 1.28 cuts off the upper 10%. We simply need to convert $z = 1.28$ to a raw score.

$$1.28 = \frac{X - 75}{7.5} \quad 1.28 * 7.5 + 75 = X$$

$$X = 9.6 + 75 = 84.6$$

    b) For the top 25% of the students the logic is the same except that $z = 0.675$.

$$.675 - \frac{X - 75}{7.5} \quad .675 * 7.5 + 75 = X$$

$$5.0625 + 75 = X = 80.0625$$

    c) For the bottom 5% the cutoff will be $z = -1.645$.

$$-1.645 = \frac{X - 75}{30} \quad -1.645 * 30 + 75 = X$$

$$X = 75 - 49.35 = 25.65$$

    d) I would conclude that students were not just guessing, and could make use of test-taking skills that they had acquired over the years.

> There is a possibility that students will fail to see the difference between Exercises 6.3 and 6.4 on the one hand, and 6.5 on the other. In the first two we are talking about performance on the test if students take it normally. There the mean is 195. In Exercise 6.5 we are talking about performance if the students just guessed purely at random without seeing the questions, but only the answers. Here the mean is 75, with a standard deviation of 7.5. These parameters are given by the binomial distribution with $N = 300, p = .25$, and $q = .75$, though the students would certainly not be expected to know this.

6.7 Reading scores for fourth and ninth grade children:

    a)
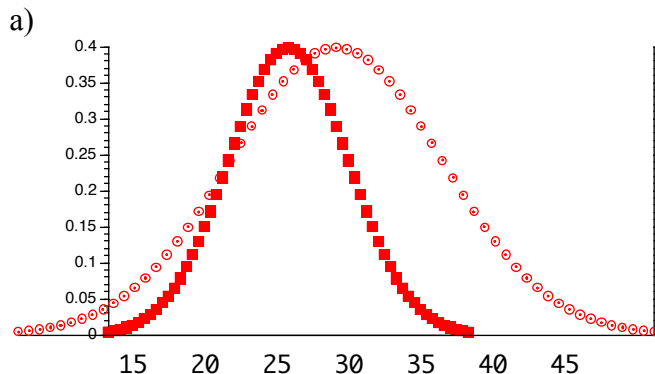
b) To do better than the average 9th grade student, the 4th grader would have to have a score of 30 or higher.

$$z = \frac{X - \mu}{\sigma} = \frac{30 - 25}{5} = 1$$

The probability that a fourth grader would exceed a score of 30 is the probability of a z greater than $1.00 = .1587$.

c) The probability of a 9th grader doing worse than the average 4th grader is the probability of a 9th grader getting a score below 25, which is the probability of being more than half a standard deviation below the mean, which is .3085.

> It is important to get students used to the idea of sketching distributions. The sketches do not have to be exact (the one above isn't), but they help students visualize the problem.
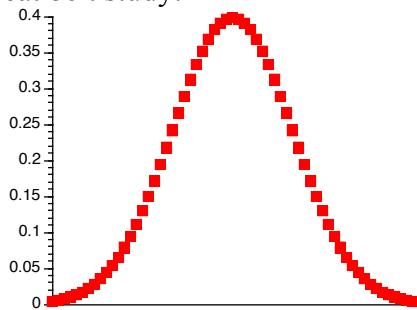
6.9 Diagnostically meaningful cutoffs:

$$1.28 = \frac{X - \mu}{\sigma} = \frac{X - 50}{10}$$

$$12.8 + 50 = X = 62.8$$

A T score of 62.8 is the score that cuts off the top 10% of the distribution, and is therefore a diagnostically meaningful cutoff.
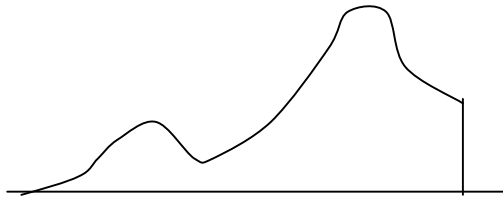
6.11 Seat belt study:

b) We need the probability of getting a 62 if the mean is 44 with a standard deviation of 7.

$$z = \frac{62 - 44}{7} = 2.57$$

The probability of $z \geq 2.57 = .0051$. This is such a small probability that we will probably conclude that the student just made up the data, rather than collecting them honestly.

6.13 Distribution of correct responses

   a) Distribution

b) The easiest way to find the cutoff for the lowest 10% is to simply take the sample data and count them, empirically finding the point with 10% of the scores below it.

This question could lead to two discussions. The first is whether the shape that I have so crudely drawn above is reasonable. I made it bimodal to suggest that there might be a distribution of scores for normal-language kids, and a second distribution of scores for kids with language problems. Your students might think that the distributions would be more overlapping, and I'm giving too much credit to normal-language kids. Such a discussion makes the question less abstract.

Second, you could ask what the sampling plan should be. Should we be making up a distribution only from using children with language problems, or using all children? That kind of question frequently arises in studies of psychiatric diagnosis.

6.15 *R* code for above plot

```
y <- dnorm(seq(0,5, .1), mean = mean(RTsec), sd = sd(RTsec))
x <- seq(0,5,.1)
hist(RTsec, breaks = 15, xlim = c(0,5),ylim = c(0,.8), freq = FALSE)
par(new = TRUE)
plot(y~x, xlim = c(0,5), ylim = c(0,.8), xlab = "", ylab = "", type = "l")
```

6.17 Identifying the highest 2% of Behavior Problem scores:

The upper 2% is cut off by $z = 2.05$

$$2.05 = \frac{X - 50}{10}$$

$2.05 * 10 + 50 = X = 70.5$

The critical cutoff is a score of 70.5.

6.19 The statisticians were upset because, by defining "overweight" as weighing more than 95% of peers (i.e. above the 95th percentile), the article seemed to be suggesting that

there were 22% of children in the top 5%. Moreover, the article says that in 1986 only 8% of children were in the top 15%. That is just silly—it is analogous to "all of the children are above average." I assume that they meant to say that 22% (etc.) were above what the 95th percentile was some years ago, but that is a different thing. Even if that is the case, the results still look too extreme to be likely.

6.21 Histogram of combined data on emotional reactivity
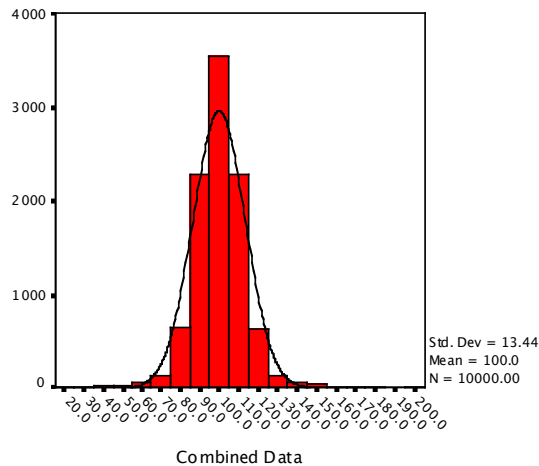
Notice that we have combined two normal distributions with the same mean, but the resulting distribution is not normal, as can be seen by comparing it to the superimposed normal curve. If the means were very different the distribution would become bimodal.

6.23 There are many out there. One is at http://davidmlane.com/hyperstat/z_table.html . An iPhone app named StatsMate will also work. Also try http://www.mathportal.org/calculators/statistics-calculator/normal-distribution-calculator.php