Math 214 – Introductory Statistics **Summer 2008**
6-20-08 Class Notes

Sections 6.4, 6.5
6.4: 1, 4, 5, 9, 11, 15, 18, 29
6.5: 9, 11, 15, 21

<u>The Central Limit Theorem</u>

In the last section, we learned how to analyze the standard normal distribution. But we also noted that there are infinitely many normal distributions. What if we want to analyze a normal distribution that isn't standard? We standardize it. Recall, that for every data value x, there is a standard value z given by the formula $z = \dfrac{x-\mu}{\sigma}$.

Example 1: The length of human pregnancy is normally distributed with a mean of 266 days and a standard deviation of 16 days. What is the probability that a randomly selected pregnancy is between 260 and 271 days?

If we first convert the raw x-values 260 and 271 into their z-scores, we get

$$z_1 = \frac{260-266}{16} = -0.38 \text{ and } z_2 = \frac{271-266}{16} = 0.31.$$

So, $P(260 < x < 271) = P(-0.38 < z < 0.31) = .1480 + .1217 = .2697$.

Example 2: The length of a Grubber fish is normally distributed with a mean of 12 inches and a standard deviation of 1.4 inches. If you are only allowed to keep a caught Grubber fish if it is in the top 5% in terms of length, what is the smallest fish you can keep?

In the first example, we knew the x-values, found the z-scores and then the probability. In this example, we know the probability, find the z-score, and then the x-value.

If we want the top 5%, we can look up .4500 on Table E and find the z-score (1.64). Using the formula $z = \dfrac{x-\mu}{\sigma}$ (solving for x), we get $x = (1.64)(1.4) + 12 = 14.3$.

Now what if the data isn't even normally distributed? We have the most amazing result.

Theorem (Central Limit Theorem): Suppose we have a probability distribution (normal or not normal) with mean μ and standard deviation σ. Instead of looking at the entire population, we take samples of size n. As the sample size n increases without bound, the distribution of the sample means *will* be normal and its mean will be μ and its standard deviation will be $\sigma/\sqrt{n}$.

What this means is that if the data is already normally distributed, then the sample means will be normally distributed too, regardless of the sample size n. If the data is not normal, then the sample means will be normally distributed as long as the sample size is "large enough," let's say greater than or equal to 30. The larger the sample size, the more normal the distribution will be.

Example 3: (#10) The average teacher's salary in New Jersey is \$52,174. Suppose the distribution is normal with standard deviation of \$7,500. (a) What is the probability that a randomly selected teacher makes less than \$50,000 a year? (b) If we sample 100 teachers' salaries, what is the probability that the sample mean is less than \$50,000?

(a) $z = \dfrac{50000 - 52174}{7500} = -0.29$, which gives us a probability of .3859.

(b) $z = \dfrac{50000 - 52174}{\left(7500/\sqrt{100}\right)} = -2.90$, which gives a probability of .0019.

Example 4: (#8) A survey found that the American family generates an average of 17.2 pounds of glass garbage each year. Assume the standard deviation is 2.5 pounds. Find the probability that the mean of a sample of 55 families will be between 17 and 18 pounds.

$z_1 = \dfrac{17 - 17.2}{\left(2.5/\sqrt{55}\right)} = -0.59$ and $z_2 = \dfrac{18 - 17.2}{\left(2.5/\sqrt{55}\right)} = 2.37$. So the probability is

$.2224 + .4911 = .7135$.