MATH 243, LECTURE 13

1. The Big Picture: Statistical Estimation

With our understanding of probability - plus one huge fact we will soon cover - we can now go back to the important central question of with what certainty we may draw conclusions from random surveys and samples of data.

We use the term *population* refers to the total population we wish to study and the *sample* refers to a particular (usually much smaller) group of the population we have information on.

Suppose we wish to measure the mean amount people spend purchasing a car in the U.S. We call this a *parameter*. It is almost impossible to calculate this number, as will be the case with most parameters we study. Instead we may interview 1000 people, and get a *statistic* based on that 1000 people. Eventually, we would like some idea how well our statistic approximates our parameter.

One of the most common parameters we want to measure is the mean of an entire population, which we now denote by μ. The associated statistic, which is the mean of some sample, will be denoted our usual $\overline{x}$.

Definition 1.
- *A parameter is some feature of your population that you are interested in.*
- *A statistic is some feature of your sample with which you would like to approximate the parameter.*

One of the basic facts of probability and statistics is that if one takes larger and larger samples from a population, a statistic will approach its corresponding parameter (so for example $\overline{x}$ will approach μ). When formalized, this is called the "Law of Large Numbers." But this does not help us in our ultimate goal of extrapolating data It says for example that we approach the right answer if we can survey larger and larger random samples of car buyers. But if we can only sample 1000 people, then what we really want to know is how close we are to the actual parameter "in all likelihood."

2. Sampling distributions

A basic practice in statistics is to take a distribution and *sample* from it.

Definition 2. *Given a population, the sampling distribution for samples of size n is the probability distribution of some parameter as we take values n times.*

We go thoroughly through an example looking at sampling distributions and comment thoroughly as we go along.

Consider the collection of numbers

$$\mathbf{S} = \{2, 2, 5, 6, 7\}.$$

What is the mean? Ans: 4.4.

We can consider all samples from our collection of size 1. There are 5 of them and again, the mean of the samples is again 4.4. The standard deviation is 2.302.

We can now consider all samples from $\mathbf{S}$ of size 2, and take the mean of each sample. Our samples are

$$\{\{2,2\},\{2,5\},\{2,6\},\{2,7\},\{2,5\},\{2,6\},$$
$$\{2,7\},\{5,6\},\{5,7\},\{6,7\}\}.$$

Our means are

$$\{2, 3.5, 4, 4.5, 3.5, 4, 4.5, 5.5, 6, 6.5\}$$

Notice that the spread is smaller, even though there are more numbers. The mean is still 4.4, and the standard deviation is now 1.329.

The probability distribution of these numbers is an example of *sampling distribution*. In this case it is for the mean of a sample of 2 from the set **S**. There are 7 possible outcomes which *don't* have equal probabilities.

X	2	3.5	4	4.5	5.5	6	6.5
Probability	.1	.2	.2	.2	.1	.1	.1

Now consider all samples of size 3, and their means. There are again 10 such samples (you might recall this from our digression on binomial coefficients), which we won't list. The means for these 10 samples are

$$\{3, 3.333, 3.667, 4.333, 4.667, 4.333, 4.667, 5, 5, 6\}$$

The standard deviation is now only .886.

Our sampling distribution for samples of 3 out of our original group S is

X	3	3.333	3.667	4.333	4.667	5	6
Probability	.1	.1	.1	.2	.2	.2	.1

Example 3. *Calculate the sampling distribution for the mean of samples of two out of the data set* $\{-1, 0, 0, 2, 3, 6\}$. *Graph both the original distribution and the sampling distribution.*

The process of taking a sampling distribution is not an easy one to understand the first time you see it; go through this example and the book carefully.

There are two important cases of sampling distributions. The first is one we have seen in a different language.

Theorem 4. *The sampling distribution measuring number of heads in n coin flips is called the* binomial distribution, *with values giving by Pascal's triangle. As n gets large, the sampling distribution for the number of heads approaches a normal distribution,* $N(\frac{n}{2}, \frac{\sqrt{n}}{2})$.

Example 5. *What is the probability, approximately, of getting more than* 550 *heads when flipping a coin* 1000 *times?*

The next theorem brings us closer to the techniques we will ultimately use to do statistics.

Theorem 6. *The sampling distribution for the mean of n measurements of some data which is known to be following a distribution of $N(\mu, \sigma)$ is itself normal, given by $N(\mu, \sigma/\sqrt{n})$.*

Example 7. *For men's heights, distributed according to* $N(70, 4)$, *what is the probability that nine men chosen at random have an average height between* 70 *and* 73*?*

Compare this with our previous answer as to the probability of finding a single man with such height.

2.1. **Facts about sampling distributions.** We see in these examples that the sampling distribution of size n always has the same mean as the original list and has a standard deviation which is decreasing as the size gets larger.

Theorem 8. *Let X be a population with mean μ and standard deviation σ.*

Consider the sampling distribution of means of samples of size n from our population.

- *The mean of the sampling distribution is again μ.*
- *The standard distribution of the sampling distribution is $\sigma/\sqrt{n}$.*
- *If our population is $N(\mu, \sigma)$ then our sampling distribution is $N(\mu, \sigma/\sqrt{n})$.*

These features of sampling distributions, which are true in general, will be critical in making statistical inferences.

3. The Central Limit Theorem

Theorem 9. *The sampling distribution of means of random samples of size n from a population with mean μ and standard deviation σ is approximately*

$$N(\mu, \sigma/\sqrt{n})$$

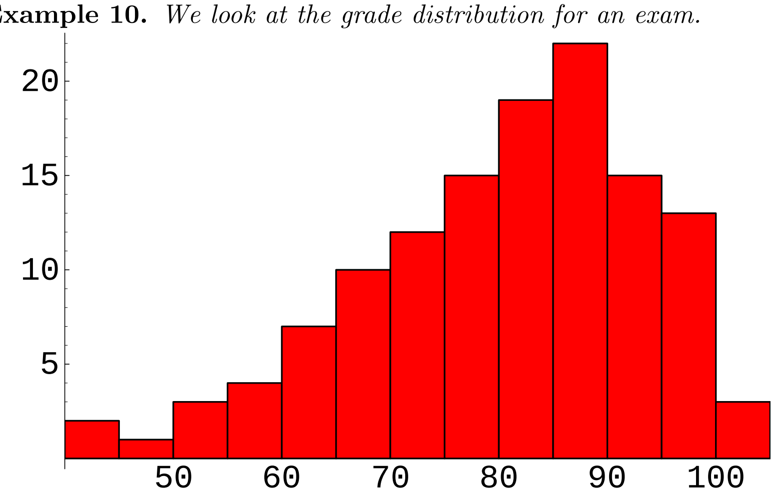
when n is large.

This theorem is true no matter what the original distribution of our population is! (unlike the previous theorem)

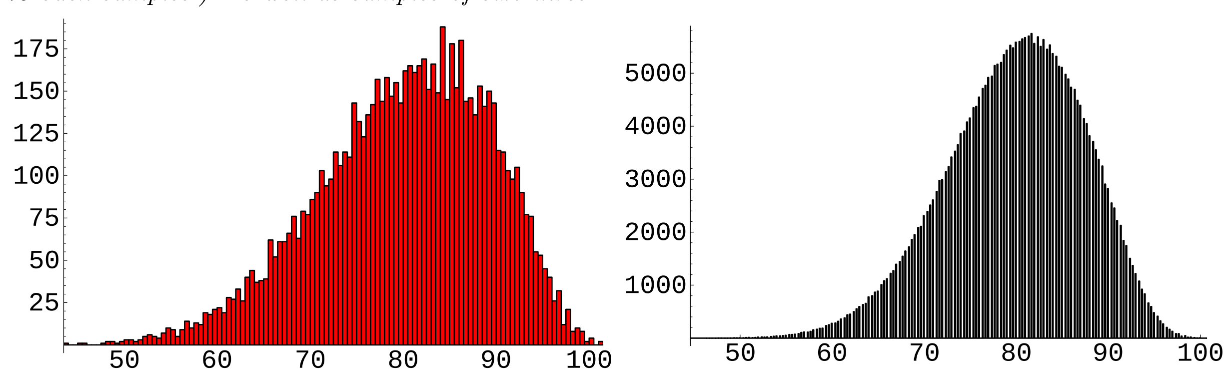
This theorem is key to the kingdom of statistics. Think of what is happening to the standard deviation - getting smaller. What does that tell us? Knowing the deviation of the sampling distribution and the fact that it is approximately normal, we know how likely it is that a sample is within that deviation (or some multiple) from the mean. So we can understand our basic question: how far our sample mean probably is from the real mean.

Before we do precisely these kinds of computations, let's see the Central Limit Theorem in action.

Example 10. *We look at the grade distribution for an exam.*



Now we take means of all samples of size 2, and look at the histogram of those numbers. (There are 7875 such samples.) As well as samples of size three.



Example 11. *Suppose that the average price of a new car purchase is \$24145 with a standard deviation of \$3615. Suppose you take a survey of 1000 car purchases. What is the probability that the average over your survey is over \$25000?*