**Transcripts for X-bar Chart Video**

SPEAKER: This lecture was organized by Dr. Alemi. This set of slides we'll walk you through the development of an x-bar control chart. Suppose that, over several months, we track the satisfaction within our unit. In each time period, we ask four patients to rate us. The question is whether the unit has improved over time.

How would you answer this question? Look at the data. There are wide variations in the data. Could these variations be due to chance?

You begin with checking assumptions of x-bar control charts. The first assumption is that we are measuring a variable that has a continuous interval scale. The variable being averaged must be a continuous variable on an interval scale where the differences between the scores are meaningful. An ordinal scale cannot be averaged. Satisfaction rating and health status ratings are generally assumed to be interval scales. Cost and time are obviously continuous interval scales.

The second assumption is that the patients are independent from each other. The observations over each time period are not affected by the previous observations. In our example, the satisfaction ratings in time period 2 should not be affected by ratings in the first time period. This assumption will be violated in an example where the same patient is rating the unit in every time period. It is likely that the patient's first impression affects subsequent evaluations. The assumption seems reasonable when different patients are rating in different time periods.

The third assumption is that the data have a normal distribution. If we were to stack all the ratings, most will fall under average rating, some on each side. A normal distribution such as [INAUDIBLE] will peak on the average, slowly decline on both sides of the average, and the shape of the curve will be symmetrical. In a normal distribution, the average and the median are on the same spot.

The law of large numbers says that no matter what the distribution of a variable is, the average of the variable will tend to become normal. As the number of cases for calculation of the average increases, the average is more likely to be exactly normal. A minimum of four cases is needed for applying the law of large numbers. The law of large numbers tells us that the average of any distribution, no matter how strange the original distribution is, has a normal distribution.

So we can also look at the data and try to eyeball it to see if it is normal or not. Using Excel, we created a histogram of the data. Select Data Analysis, then select Histogram to reach this option.

Here is how the data looked. An eyeball test suggests that this data may be near normal. Most of the data is in the center. The data is spread symmetrically around the peak of the center.

The last assumption is that the variance is constant. The variable variance will be observed if the patient's severity of illness changes over time. This assumption can be verified on the control chart. It states that deviations from the average should not consistently increase or decrease over time, so plotting the deviations from the average can help us understand if this assumption is met.

The next step, after checking the assumptions, is to calculate and plot the average for each time period. This x pi plot already tells you a lot about changes in the average rating, but it doesn't tell the full story. It will tell us more if we add to it the upper and lower control limits. These are the limits between which one expects 95 percent of the data.

Next, calculate and plot the grand average. The grand average is the average of all ratings across all time periods. Do not calculate this by averaging the mean of each time period. The correct way to do this is to sum all the ratings for all time periods and divide the sum by the number of ratings. In the example provided here, it makes no difference how you calculate the grand average. But in many situations, where the number of cases in each time period is changing, it does make a difference.

With the central tendency line in, we have a visual line to compare the data to. It already tells us a lot. It gives us a sense of which time periods are closer to our central tendency line.

To calculate control limits, we need to first calculate standard deviation. Excel as a function for calculating standard deviation. Here, we are showing a formula for doing so, in case you want to do it by yourself.

First, sum the square residuals. Then, this sum is divided by the number of cases minus 1. The square root of the calculated value is the standard deviation. Estimate the standard deviations for observations within each time period by dividing the standard deviations for all of the observations by the square root of the number of cases in the time period.

Calculate the upper control limit for each time period as the grand average plus 1.96 times the standard deviation. 1.96 sets the limit so that 95% of data fall within the two limits. Calculate the lower control limit for each time period as the grand average minus 1.96 times the standard deviation of that period.

Let us look at these calculations inside Excel. We begin with the calculation of standard deviation. You can use the standard deviation function in Excel to calculate the standard deviation of all observations. The function is equals STDEV, and the arguments of the functions are the range of the data. In this example, it's taking data from the range B2 through E2.

The standard deviation is then divided by the square root of the count of the cases in your time period. Taking something to the power of 0.5 is the same as taking the square root. This provides the estimate for the standard deviation now for that time period.

The t-statistic is 1.96. That guarantees that 95% of data fall within the two control limits. The grand average is calculated using the average function within Excel from the entire set of the data.

Now we are ready to plot the data. In x-bar control chart, the x-axis is time, the y-axis is observed average for each time period. Observations are shown by single markers. Control limits are shown as lines with no markers, and preferably in red. Note that the control limits are straight lines in this example, because in every time period, we sampled four cases. If this were not the case, the control limits would be tighter when the sample size was larger, and wider otherwise.

Any point that falls outside the control limit is not due to chance. The first time period is above the limit. Patients rated as higher than the historical trend in the first time period. Note also that the second time period is lower than the lower control limit. Therefore, patients rated our services worse in this time period, and a change in the ratings were not due to event chances. This change marks a real change in satisfaction with our services. We would not have known this until we created the control chart.

Once you have created the control chart, tell your story. Distribute your control chart. Include in the report how you check the assumptions of the control chart. Include a summary of key findings, and hypothesize the likely factors that contributed to points outside the control limit.

X-bar chart tells if observations are within historical patterns.