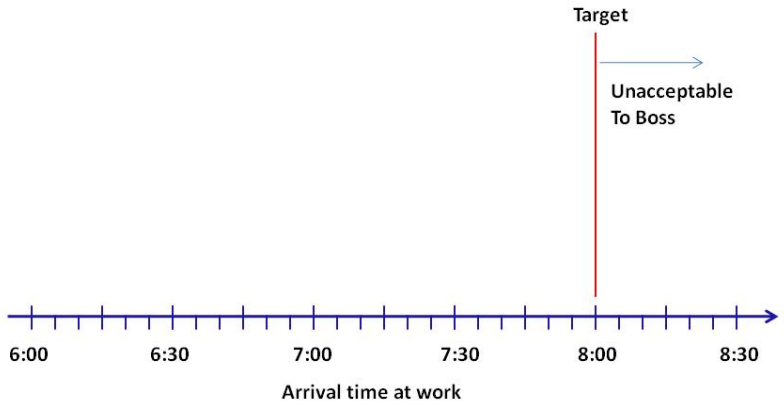
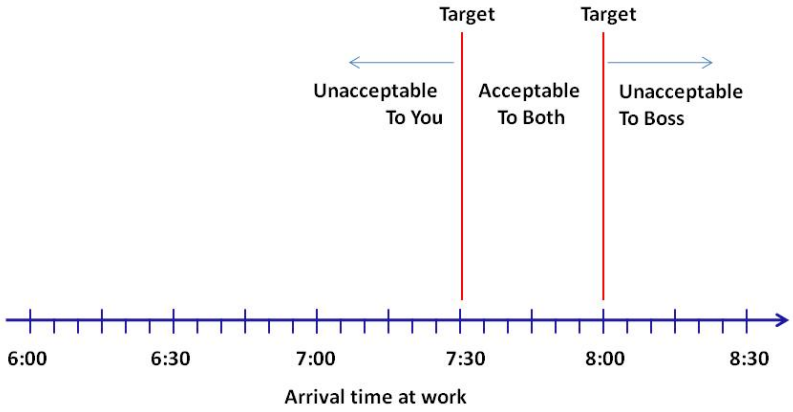
The Basics of Variation

Variation

A question that is often asked is, “How much variance is acceptable?” Taking the work arrival time example, your boss probably doesn’t care how long it takes you to get to work, as long as you are at your desk by 8:00 in the morning. On a graph, that might look like this:



As far as your boss is concerned, you can come in as early as you want as long as it’s by 8:00. You, on the other hand, would rather not be into work any earlier than 7:30 am. Adding that to our graph gives us two targets, yours and your boss’s:

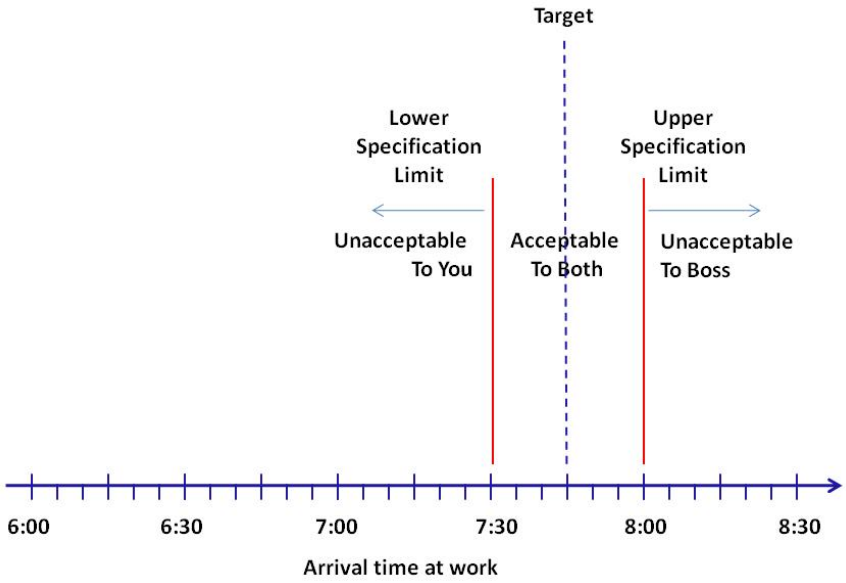


Targets and Specifications

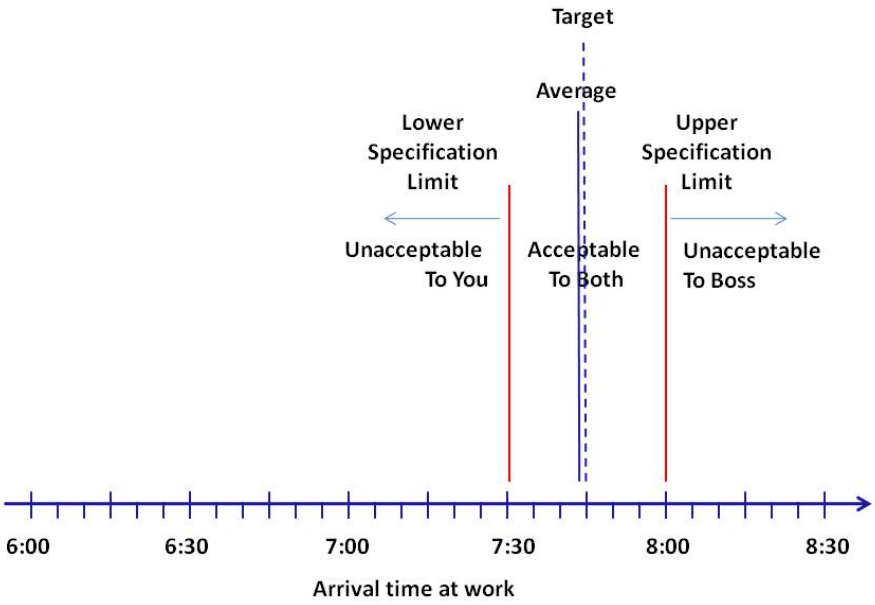
These two targets are called **specifications**. The one on the right, your boss’s target, is called an **upper specification limit** and the one on the left, your target, is called a **lower specification limit.**

We can see from the graph, then, that as long as you are at work between 7:30 and 8:00, it will be acceptable to both you and your boss.

In fact, let's change the names of the two targets to upper and lower specification limits. In addition, let's add a new target half-way between the two specification limits:

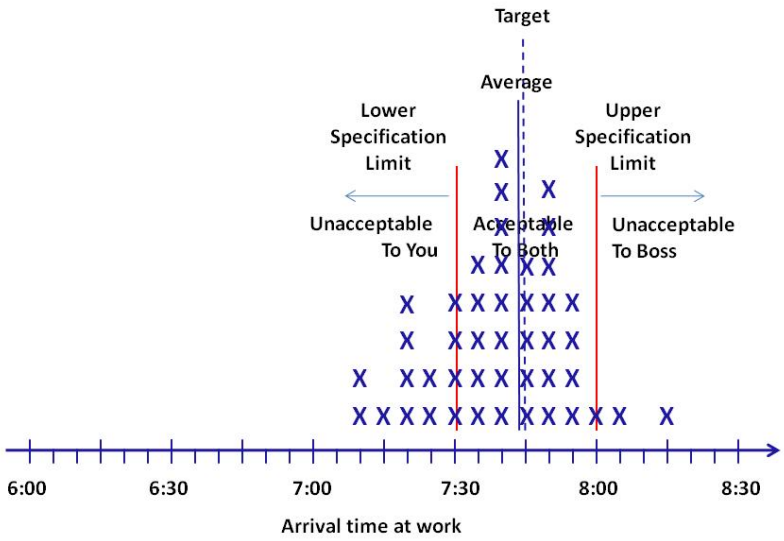


Now that we know what the difference is between unacceptable and acceptable arrival time performance, let's look at how you've done. Looking at your arrival time data for the past nine weeks, your average arrive time was calculated to be 7:39 am. This is shown on the graph below:



When Hitting the Target Isn't Good Enough

This is good news! You are meeting both your specification and your boss's specification and you are very close to the new target that was established. Or, are you? Over that nine week period you actually arrived late twice and got in before 7:30 nine times. Maybe there is something else we need to look at besides the average. Each "**X**" on the graph below represents your arrival time for a particular day.



This graph makes it clear that there is variation in your arrival time. If there were no variation in your arrival time, there would be 45 Xs stacked at one time on the timeline. It is also clear from the graph the average does not tell the whole story of how you are doing with respect to the specifications or target.

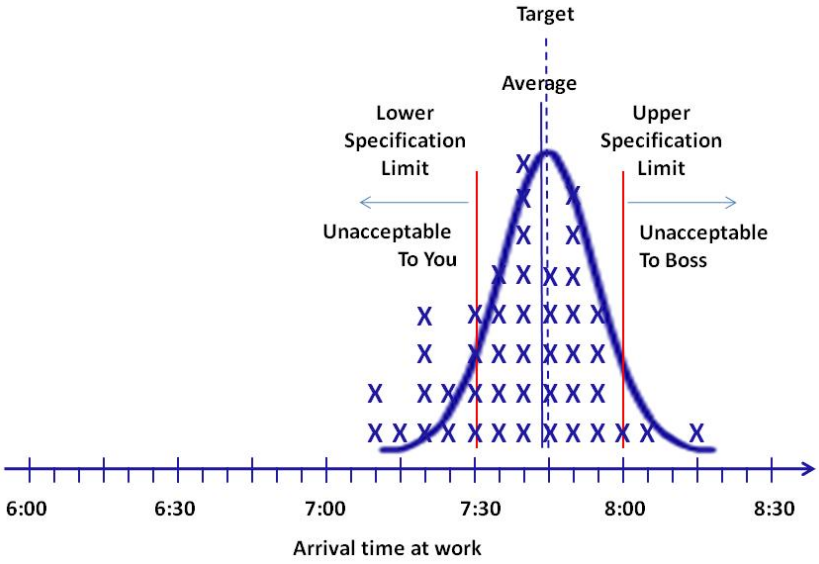
Even though the average of 7:39 am is in the acceptable range and very close to the target, there are 11 values outside the acceptable range. The average doesn't represent any single value in the sample of arrival times. In fact the average value of 7:39 am does not occur at all in the sample (there are 8 values at 7:40 am).

It's worth noting here that the specification limits have nothing to do with the actual arrival times. No matter what your average arrival time is, the specification limits stay the same.

> **MoreSteam Note:** Specification limits are independent of any process characterization and are often called the Voice of the Customer.

Spread and Location, Indicators of Variation

The pattern of the Xs in the graph above is called a **distribution**. The Xs are distributed in a pattern according to the number of days a particular arrival time was logged. The **shape** of this distribution will be of interest in more in-depth studies. Suffice it for now to say that the shape of the distribution of many process characteristics is **bell-shaped** and is symmetrical around the average:

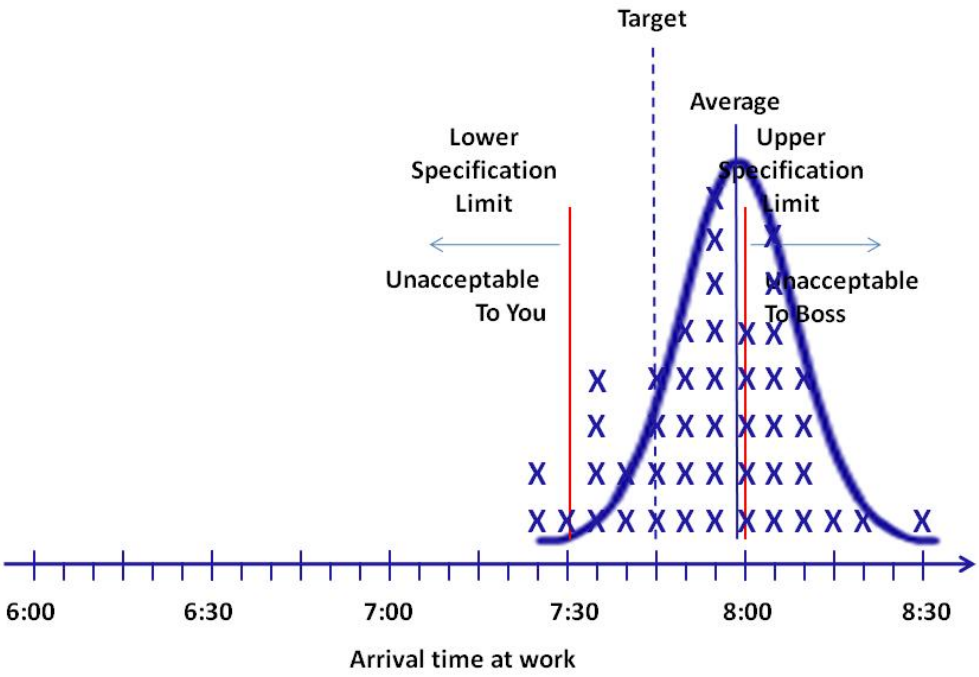


Looking at the bell shape, an indication of how much variation there is in your arrival time is the width of the wide end of the bell. The wider the bell, the more variation there is. Another way to look at it is, the further the lowest and highest values are from each other, the more variation there is. This variation characteristic of a process is called **spread**.

Looking at the spread of a distribution as an indication of variation is very intuitive.

There is another characteristic that may not be as intuitive because, unlike the spread which is solely a process characteristic, this new characteristic is based on the process and the specifications. It's the difference between the target (from the specifications) and the average (from the process) and is called the **location** of the distribution.

If the distribution of your process characteristic is bell-shaped, then a location of zero (target = average) gives the best chance of meeting the specifications. When this occurs, the process is said to be **centered**, i.e. centered between the specification limits. If the location is large, i.e. the process distribution is shifted far from the target, much of the process output will be outside the specification limits:



Note that in this situation there are 16 days where the arrival time is outside the specification limits. This is compared to 11 days outside the specification limits when the distribution of arrival times was closer to being centered.

MoreSteam Note: A bell-shaped process distribution will have the least number of data points outside the specification limits when the distribution is centered between the specification limits.

Summary

There are two things to look at when considering the variation in the distribution of a process characteristic. 1) the **spread** of the distribution and 2) the **location** of the distribution. The spread is purely a process characteristic and the location is a combination characteristic being the difference between the process distribution average and the target (half the difference of the specification limits).

The location can be easily quantified in a simple difference calculation as shown. The spread can be easily quantified by subtracting the lowest value of the distribution from the highest. This is also called the **range**. **Standard deviation** is another way to quantify spread and is covered in a more advanced lesson.

If the range or spread can be reduced, the distribution will become more **centered** and more, if not all, of the data in the distribution will fall within the specification limits. This is called **variation reduction** and is a key thrust of Lean Six Sigma.