SPC

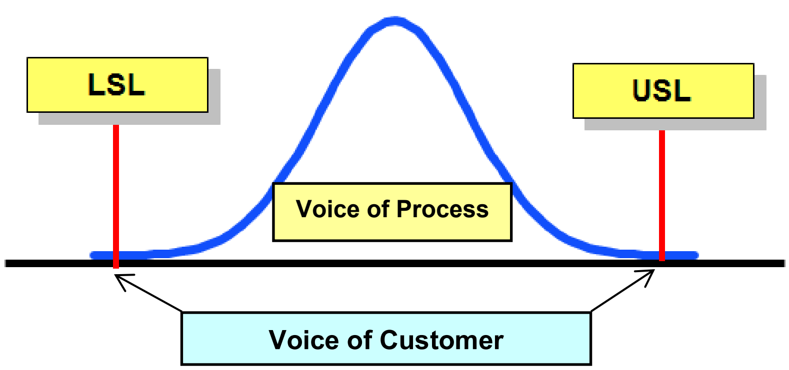
LESSON: Capability Analysis

- A **Capability Analysis** gathers and analyzes data to see how well a process meets **customer specifications**
- As we gather data for our processes, we must continue to ask: “Is our process ***capable*** of meeting these specifications and to what extent?”

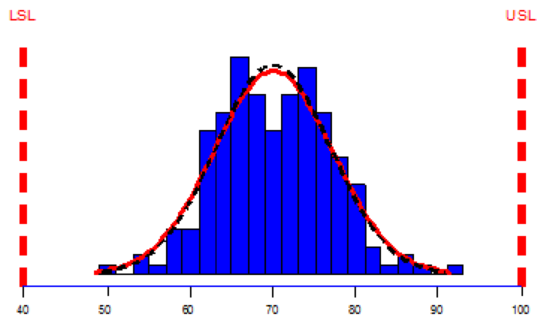
Here are some examples of **specifications** for some of my Six Sigma class projects:

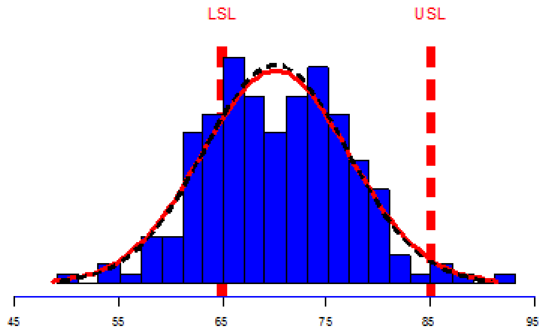
- Time between ordering and receiving food at Chauncey’s is at most **10 minutes**
- **Electricity use** in Percopo Hall is at most **x kilowatt hours** between 2 a.m. and 6 a.m.
- **Honeysuckle invasive plant pulling time** is at most **4 minutes** for a “medium-sized” (circumference at base ~5 to 7 inches) plant using The Honeysuckle Popper (tool for pulling honeysuckle plants)
- **Weight of hand towel waste** in Crapo bathrooms is at most ***y* grams**
- Who typically **sets the specification limits?**
- How do we define a “defect” in a process?
- The output of a capability study is a **Capability Ratio**. There are several different capability ratios, the most common being Cp, Cpk, Pp, Ppk, and Cpm. The difference between them is how they estimate the process standard deviation σ. In general, a capability ratio is defined as:

$$\textbf{Capability Ratio} = \frac{\text{Voice of Customer}}{\text{Voice of Process}}$$



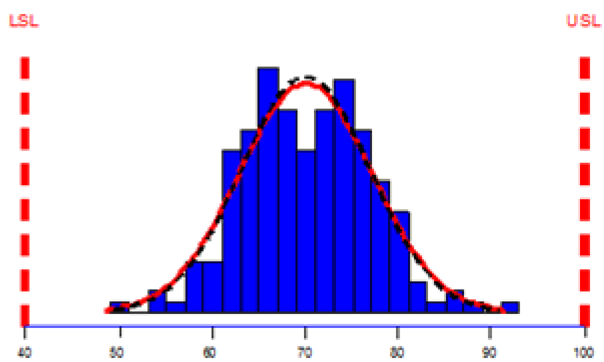
Example 1. Which of these two processes, whose histograms are in blue, **produces fewer defects?** Which of these two has a **higher capability ratio?** [Higher capability ratio ➜ Better Process = Less Defectives]





- Two important aspects of the distribution of a process: center and spread
- If specifications are too tight and a process is unable to meet them, what can we do?

Let's give the **capability ratio** a statistical definition with the appropriate notation. So, using USL, LSL, and σ (as an estimate of the process standard deviation), create a formula for the **capability index C_p**.

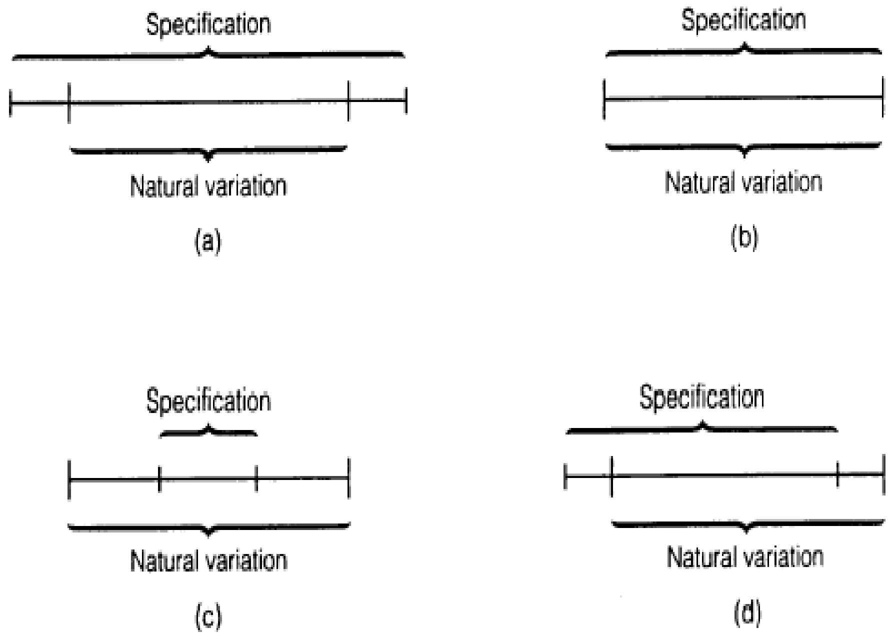


$$C_p = \textbf{Capability Index} = \frac{\text{Voice of Customer}}{\text{Voice of Process}} = \frac{\text{Specification Spread}}{\text{Process Spread}} =$$

Measure of "Voice of Process": If the process is in-control, the process standard deviation σ can be estimated using the pooled standard deviation s_p, the mean standard deviation $\bar{s}$, or the mean range R (the one we've been using):

$$\hat{\sigma} = \frac{s_p}{c_4} \text{ or } \hat{\sigma} = \frac{\bar{s}}{c_4} \text{ or } \hat{\boldsymbol{\sigma}} = \frac{\bar{\boldsymbol{R}}}{\boldsymbol{d_2}}$$

What is a "good" value for C_p?



***Cp* values:**

***Cp* > 1** means that the **process variation is less than the specification**, however, defects might be made if the process is not centered on the target value.

***Cp* = 1** means that the **process is just meeting specifications**. A minimum of 0.3% defects will be made and more if the process is not centered.

***Cp* < 1** means the **process variation exceeds specification**, and a significant number of defects are being made.

Example 2. Process Capability: The Goodman Tire and Rubber Company periodically tests its tires for tread wear under simulated road conditions. Twenty subgroups, each containing **three radial tires**, were chosen from different shifts over several days of operation, see Lesson15_Treadwear_WaitTimes_DATA. The specifications for a tire are 30 ±10 hundredths of an inch. How capable is the process given the data collected below?

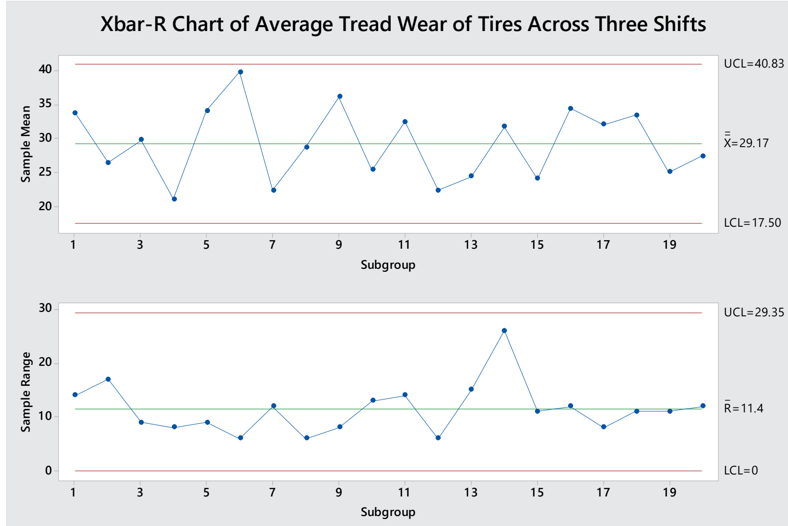
Subgroup	Shift 1 Tread Wear	Shift 2 Tread Wear	Shift 3 Tread Wear
1	31	42	28
2	26	18	35
3	25	30	34
4	17	25	21
5	38	29	35
…	…	…	…
18	40	29	31
19	18	29	28
20	22	34	26

(a) Is this a “good” subgrouping plan? Why or why not?

(b) Before checking capability, what should we first consider?

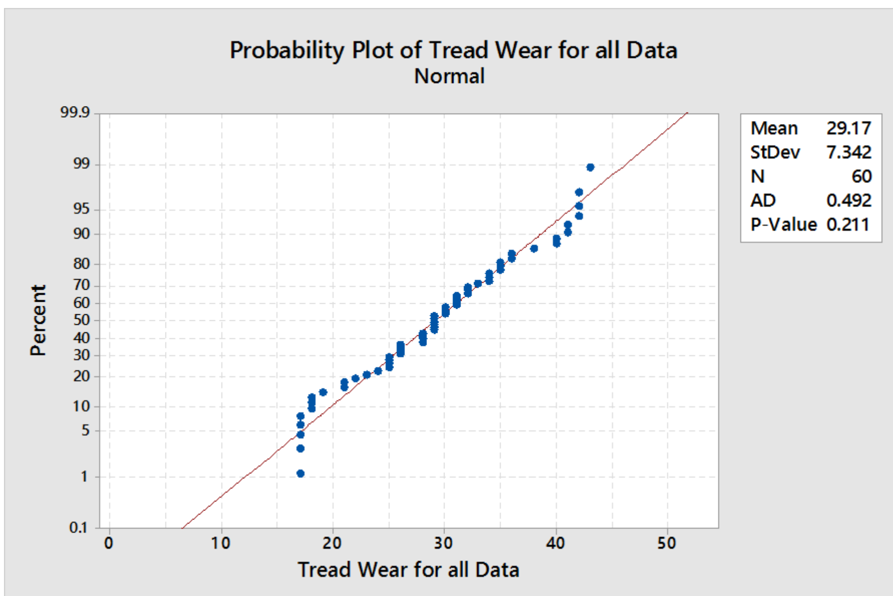
(c) The “correct” control charts (not necessarily subgrouping plan) to use are Xbar-R – why?

(d) Let X be the tread wear for a tire. What are the mean and standard deviation of X as estimated from the Xbar-R charts?



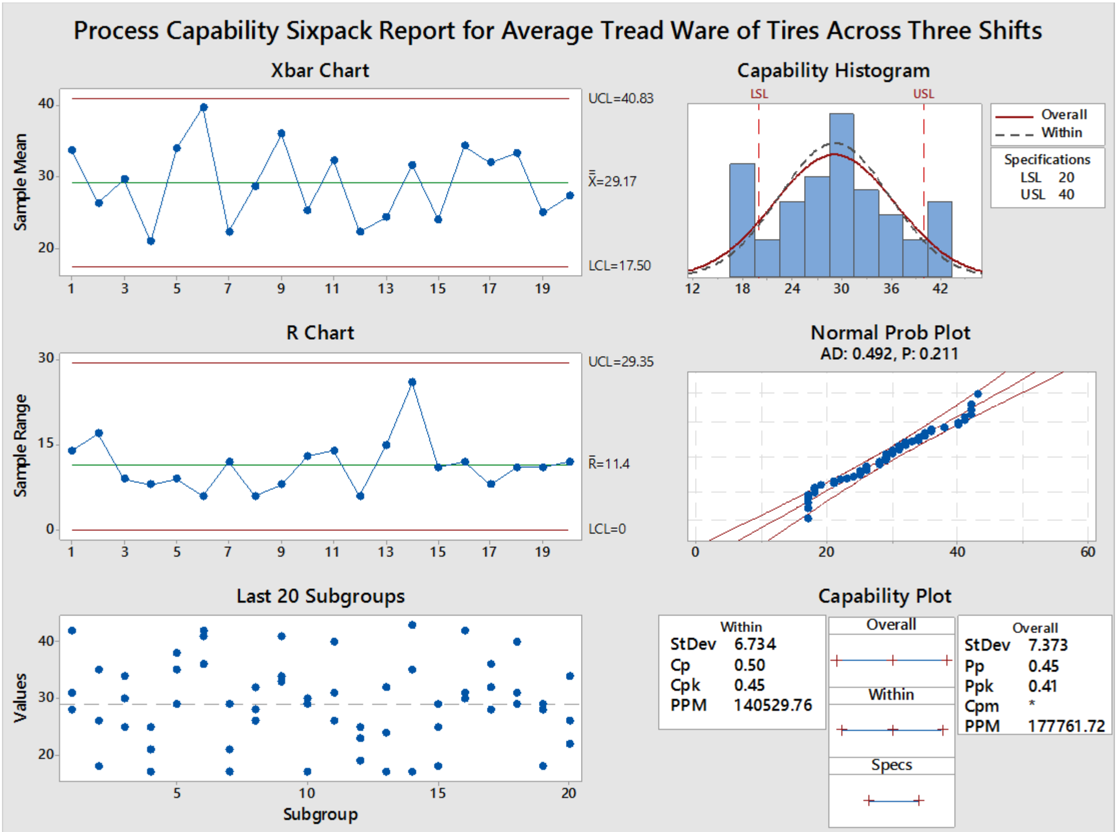
(e) Are the individual data points from a normally distributed population? Is there dependency in the data?

H_0: Data is from a normally distributed population **H_a:** Data is NOT from a normally distributed population

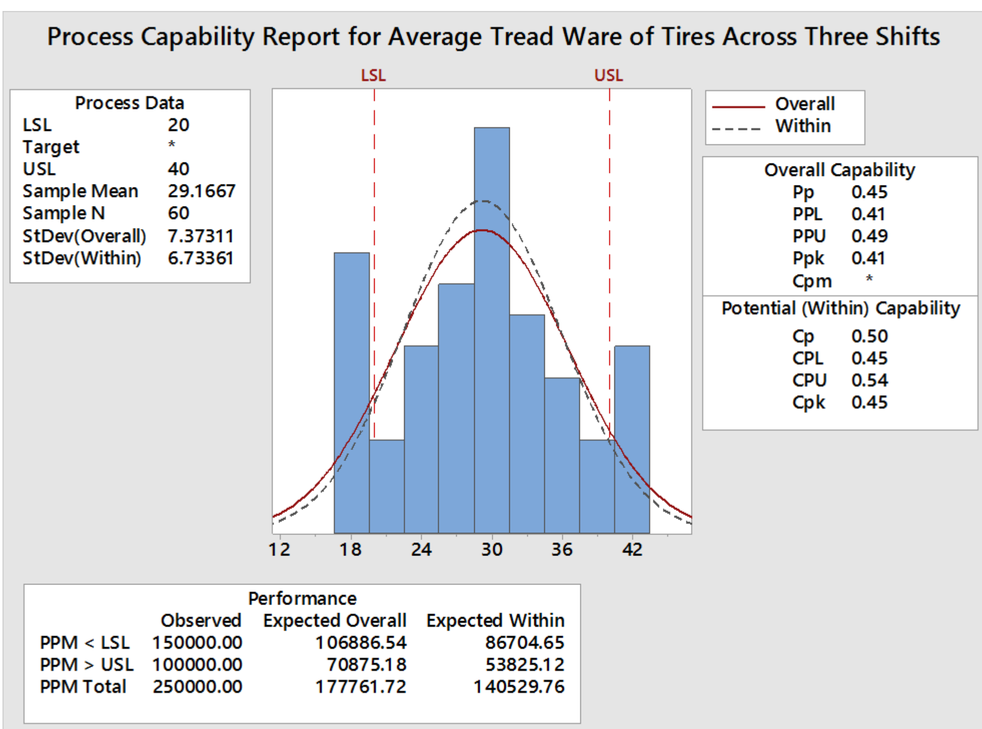


(f) Compute Cp by-hand: Cp =

(g) Compute Cp using Minitab's Process Capability Sixpack:
Stat > **Quality Tools** > **Capability Sixpack** > **Normal**

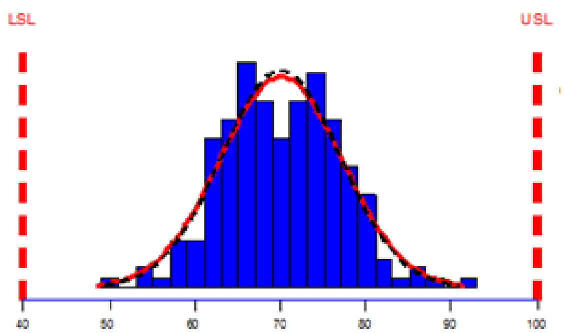


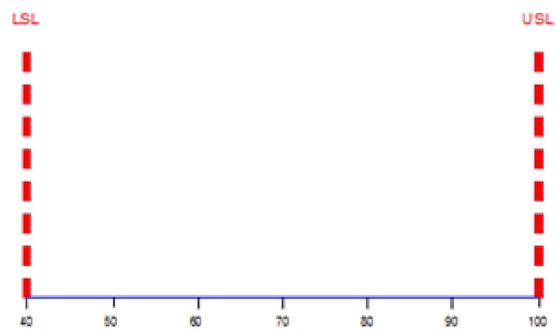
Compute Cp using Minitab's Capability Analysis: **Stat** > **Quality Tools** > **Capability Analysis** > **Normal**



Problem: The index C_p does not take what into account?

Draw a picture of a process with the same LSL and USL as the left plot and has the exact same capability index C_p, but the customer would NOT appreciate the resulting product.





So, we need ***another*** capability index that accounts for the ***centering*** of the distribution of the process.

$$C_{pk} = \min\{C_{pl}, C_{pu}\}, \text{ where}$$

$C_{pl} =$ $\qquad\qquad$ $C_{pu} =$

Example 3. Suppose your company makes McDonald's fries and the specification limits for salt content is between 4.1% and 5.0% for the customer to be happy. You have compiled actual salt content data for the last several weeks and have determined that salt content is independent and normally distributed with mean $\hat{\mu}$ = 4.4% and standard deviation $\hat{\sigma}$ = 0.2%.

(a) Determine the capability of the salt process to meet customer specifications by calculating C_{pk} and then C_{pk}. Draw a picture of the situation and include the Voice of Customer (VOC) and Voice of Process (VOP) spreads.

(b) How could we increase the value of C_{pk}? C_{pk}?

(c) Can C_{pk} and C_{pk} be the same value? What does C_{pk} equaling C_{pk} mean?

(d) Will C_{pk} always be less than or equal to C_{pk}?

(e) Is it possible for C_p to be negative? If so, when would this happen? Can C_{pk} be negative? If so, when would this happen?

Capability Analysis ASSUMPTIONS

1. What must be true about our process before we can perform a capability analysis?

2. What other assumption must we check that I've clearly alluded to in the graphics that I've created this lecture?

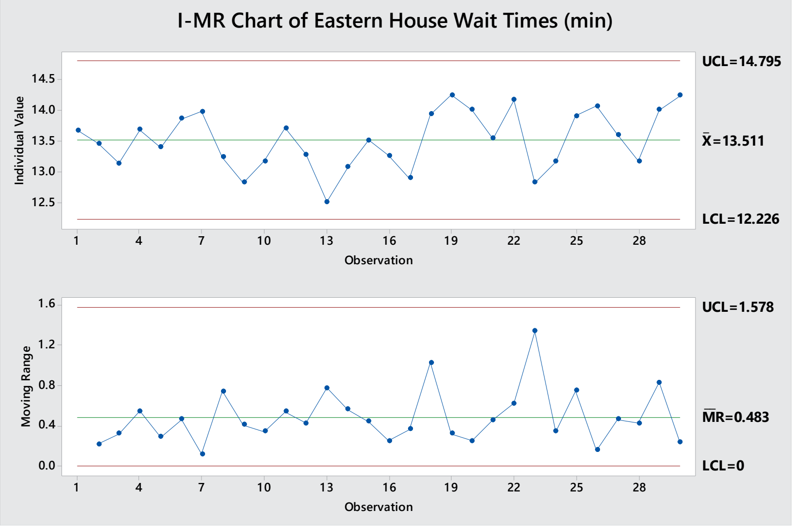
3. Other assumption we should check?

- **If these assumptions are not met, we can still perform a capability study, but the indices will likely be meaningless.**

Example 4. I have been able to collect data for **waiting times** at the Eastern House Restaurant for mealtimes starting at 7 p.m. The manager wanted to improve customer satisfaction related to customer waiting for their meals. The **lower specification of 10 minutes** was established to increase beverage sales. The manager found if the wait time was less than 10 minutes people wouldn't order a second drink. The **upper specification was 16 minutes** because manager started to receive complaints when wait time went over 16 minutes. If the manager could decrease the variation in wait time, then she could better predict how long people would have to wait for tables and how many customers she could serve in one evening.

Let's first make a control chart of the wait time per table data. What control chart do we want to build?

(a) I turned on all 8 "out of control rules" (one point beyond 3 sigma control limits, 6 in a row trending up or down, 2 of 3 within 1 sigma control limit on same side of chart, etc.) Is the process "stable" (i.e. in-control)?

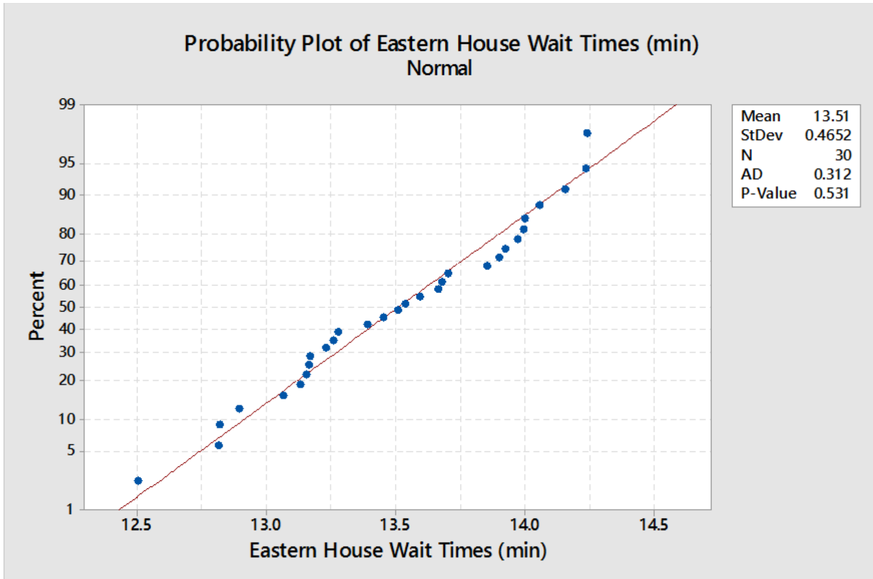


(b) Is the data normally distributed?

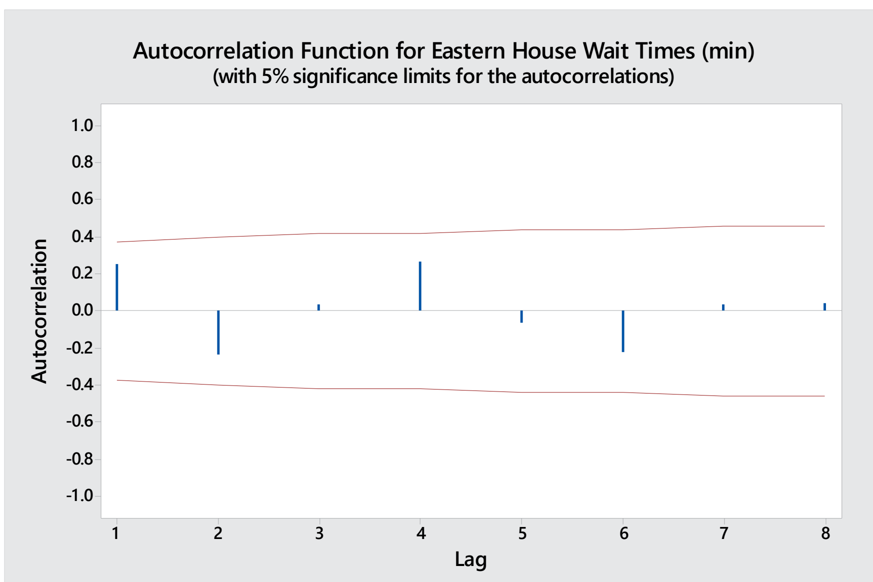
You are performing the following hypothesis test:

H_0: Data is from a normally distributed process

H_a: Data is not from a normally distributed process

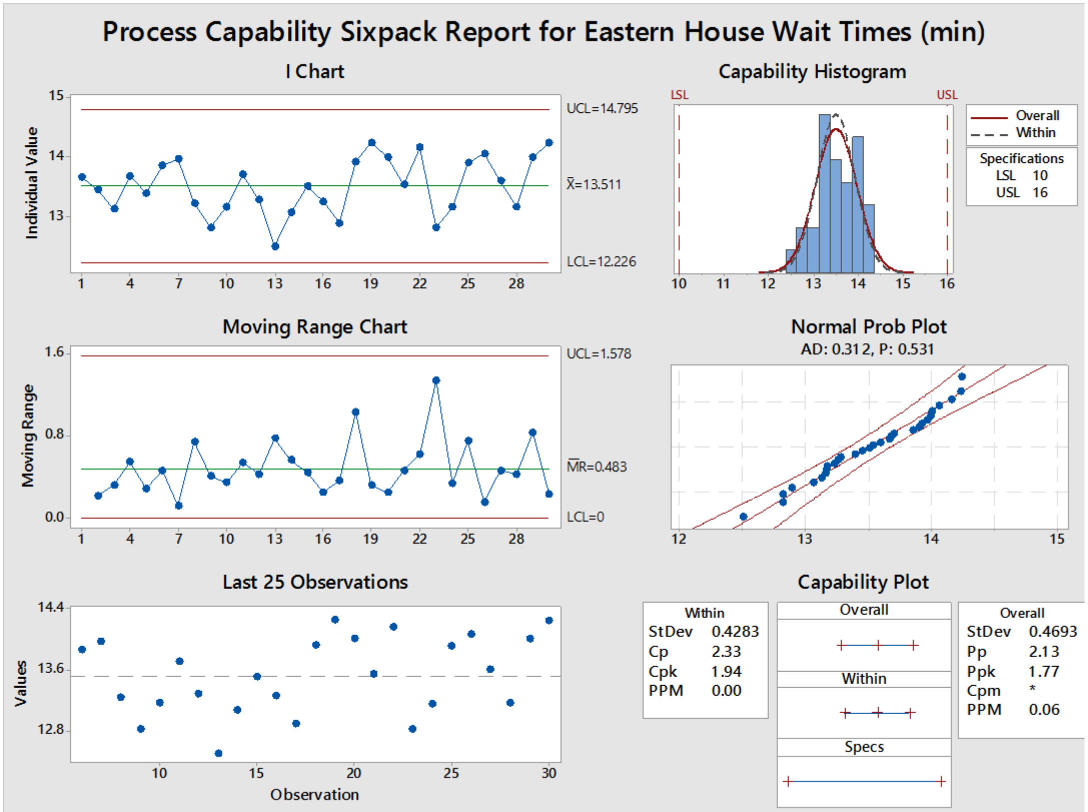


(c) Autocorrelation plot:

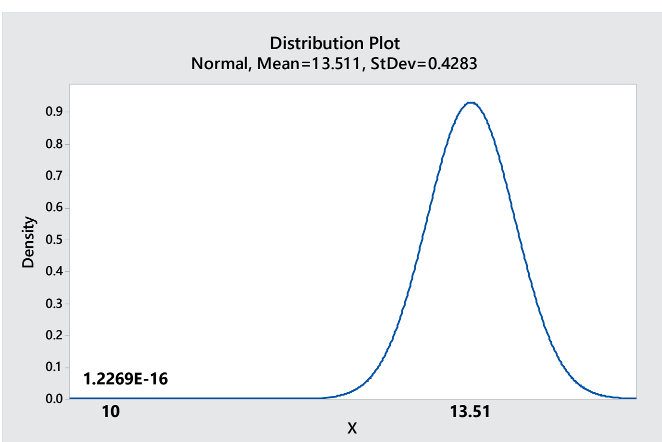


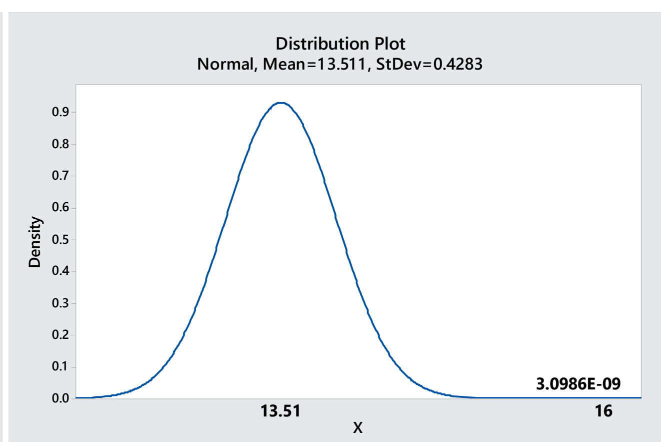
(d) Have we met the assumption to run a capability study with normality?

Stat > **Quality Tools** > **Capability Sixpack** > **Normal**

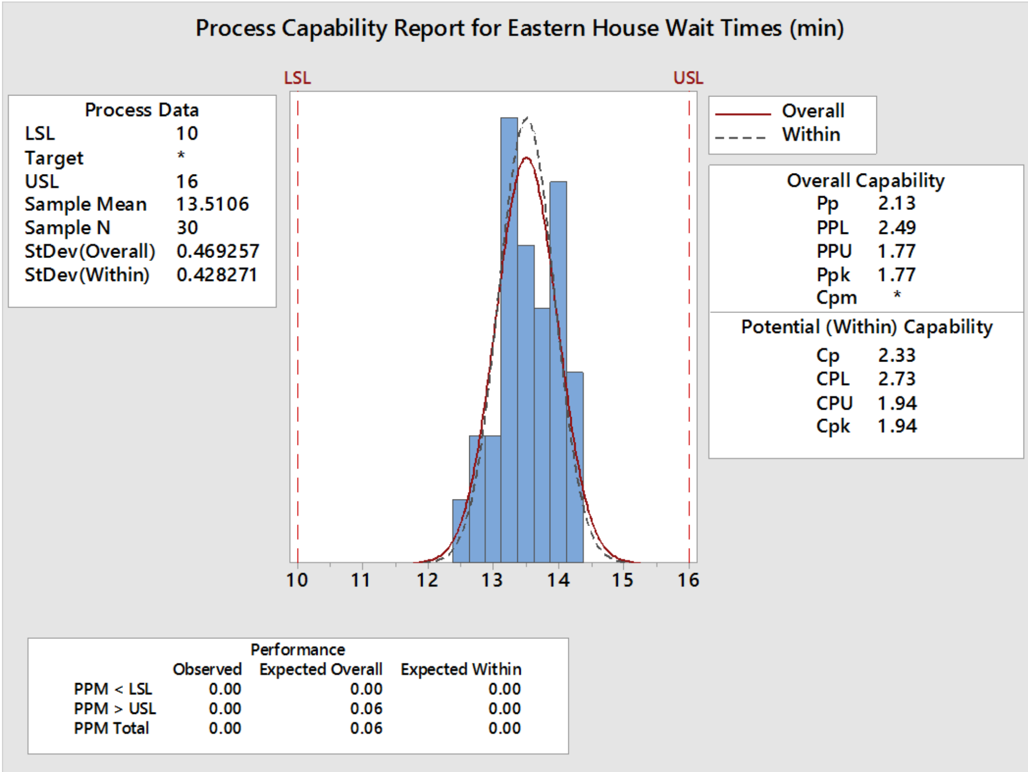


PPM below LSL and above USL:





Stat > Quality Tools > Capability Analysis > Normal



- P_p and P_{pk} are quite like C_p and C_{pk}. I think you may be able to tell from the above graphic.
- P_p and Ppk are computed using the "overall" standard deviation of the process (black solid line) and C_p and C_{pk} are computed using the "within" standard deviation of the process (red dotted line).

Thus, the **capability ratios P_p and P_{pk}** have the **same formulas as C_p and C_{pk}**, ***except the process standard deviation is computed differently***. That is, σ is computed using "long term" variation of the process, and not the "short term" variation as in C_p and C_{pk}.

$$\boldsymbol{P_p} = \textbf{Capability Index} = \frac{\text{Voice of Customer}}{\text{Voice of Process}} = \frac{(\text{USL}-\text{LSL})}{6\hat{\sigma}}$$

$$\boldsymbol{P_{pk}} = \textbf{min}\{\boldsymbol{P_{pl}}, \boldsymbol{P_{pu}}\}\text{, where}$$

$$P_{pl} = \frac{(\hat{\mu}-\text{LSL})}{3\hat{\sigma}} \qquad P_{pu} = \frac{(\text{USL}-\hat{\mu})}{3\hat{\sigma}}$$

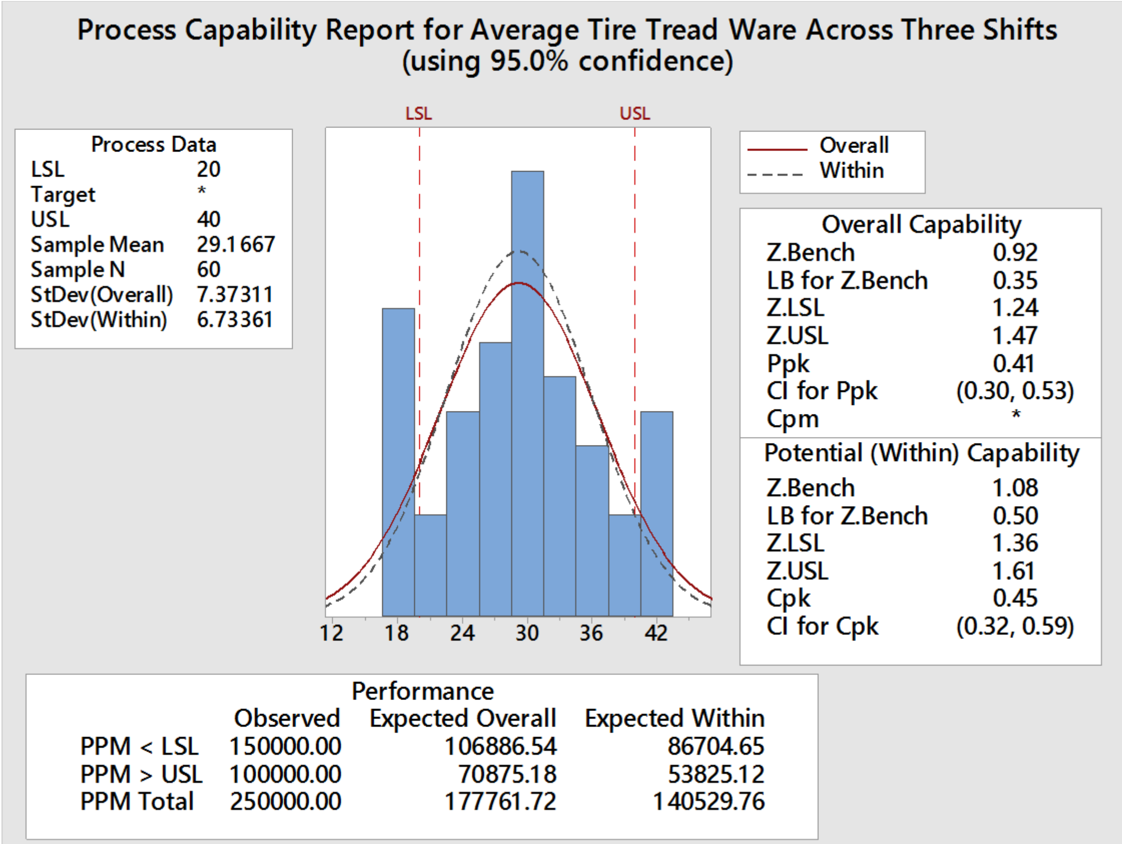
In the above formulas, σ is the variation for all the data. This is the formula that you are used to compute standard deviation in Engineering Stats I:

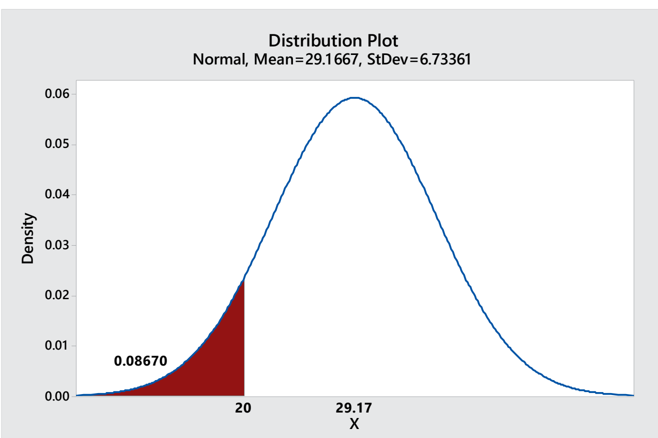
$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{k}(x_i-\hat{\mu})^2}{k-1}} \quad \textbf{overall}\text{ standard deviation}$$

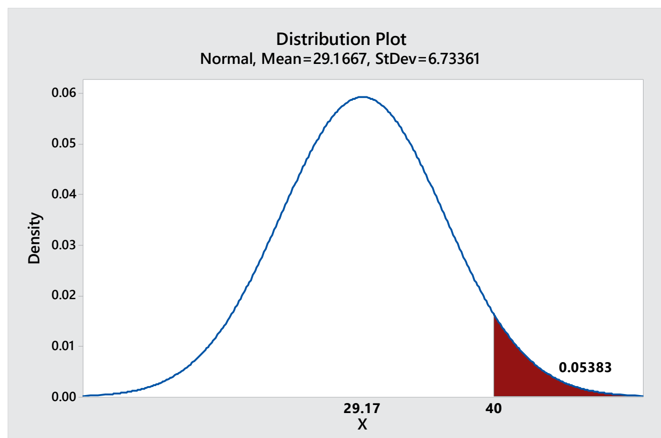
where *k* is the number of trials.

When the within and overall standard deviations are similar, then C_{pk} and P_{pk} are similar, too. Within and overall standard deviation tend to be the same when the process is in-control.

Going full circle: Capability Indices and Sigma Levels







Sigma Levels: Left tail: $\frac{20-29.1667}{6.73361} \cong -1.361$

Right tail: $\frac{40-29.1667}{6.73361} \cong 1.609$

PPM: ~86,705

PPM: ~53,825

Long Term Sigma:

Long Term Sigma

1.5	3	66,807
1.4	2.9	80,757
1.3	2.8	96,800
1.2	2.7	115,070
1.1	2.6	135,666

Long Term Sigma Level	Short Term Sigma Level	Defects Per Million
1.9	3.4	28,717
1.8	3.3	35,930
1.7	3.2	44,565
1.6	3.1	54,799
1.5	3	66,807

Mayberry, Ppk: Episode 1 (published on Minitab Blog, 8/27/12)
Making intuitive sense of capability output

In this episode, Sheriff Andy Taylor uses cars and parking spaces to **explain process capability** to Deputy Barney Fife.

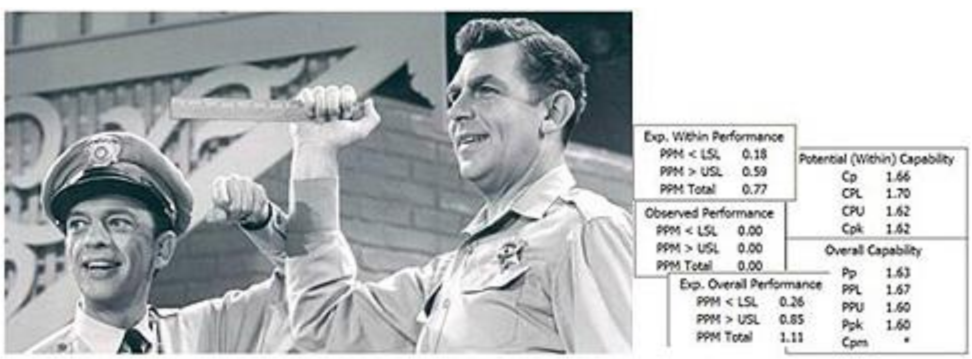
Sheriff Taylor: What's wrong, Barney?

Barney: Someone spilled alphabet soup all over my capability output! Wait till I catch the practical joker who did this! They'll be sorry!!

Sheriff Taylor: That's not alphabet soup, Barney. Those are capability indices.

Barney: I don't see any indices, Andy. I just see a big mess of letters: Cp, Pp, Cpk, Ppk, Cpm, PPM… LMAO!

Sheriff Taylor: LMAO is not part of capability output, Barney. At any rate, ***the indices help you evaluate whether your process is meeting your customer requirements.***



Barney: Well, my customers don't require random letters of the alphabet. I wish they did. It'd be a lot easier to satisfy 'em.

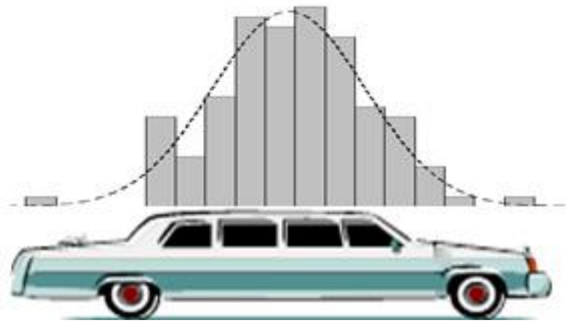
Sheriff Taylor: OK, let's try to break this down for you, Barney. Should we start with Pp?

Barney: No thanks, I just went a couple minutes ago… ha, ha!

Sheriff Taylor (sighing): Hmm… this might be harder than I thought. OK, let's try an analogy. Think of the **variation of your process as a car.**

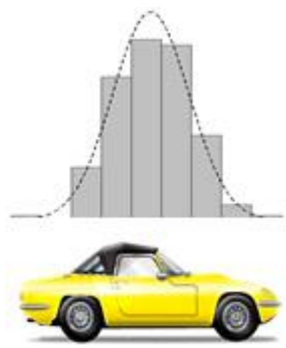
Barney: What kind of car?

Sheriff Taylor: Depends. If your **process has a lot of variation**, you might think of it as a **long, stretched-out limousine**, like this:



Barney: Oh boy, that's what I call a lottttt-a-car!

Sheriff Taylor: Well, Barney, it's a lot of process variation. But suppose **your process doesn't produce much variation**. In that case it might be more like a **small sports car**, like this:

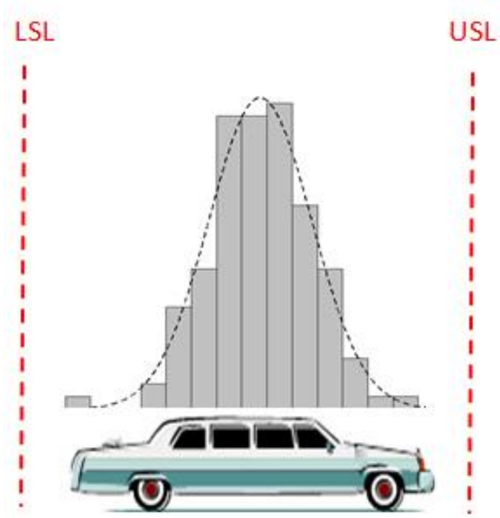


Barney: Boy, would I love to drive a zippy little process like that!

Sheriff Taylor: Wouldn't we all. Now, you need a parking space to park your car in, don't you Barney?

Barney: You're darn tootin' I do! I'd like a space right on Main Street Mayberry.

Sheriff Taylor: Well, your parking space is determined by your customer requirements—in other words, the **specification limits for your process**. Those **limits are the parking space** that your car—your process—needs to fit into:



Barney: Well, duh, I get that. You ain't talkin' to a rock, ya know. But what does that have to do with all those darn p's in my output?

Sheriff Taylor: Well, Barney, Pp, for example, is just the **ratio** of the **parking space to the length of the car.**

Barney (excited): I get it! So Pp just tells me how many cars I can get into my parking space!

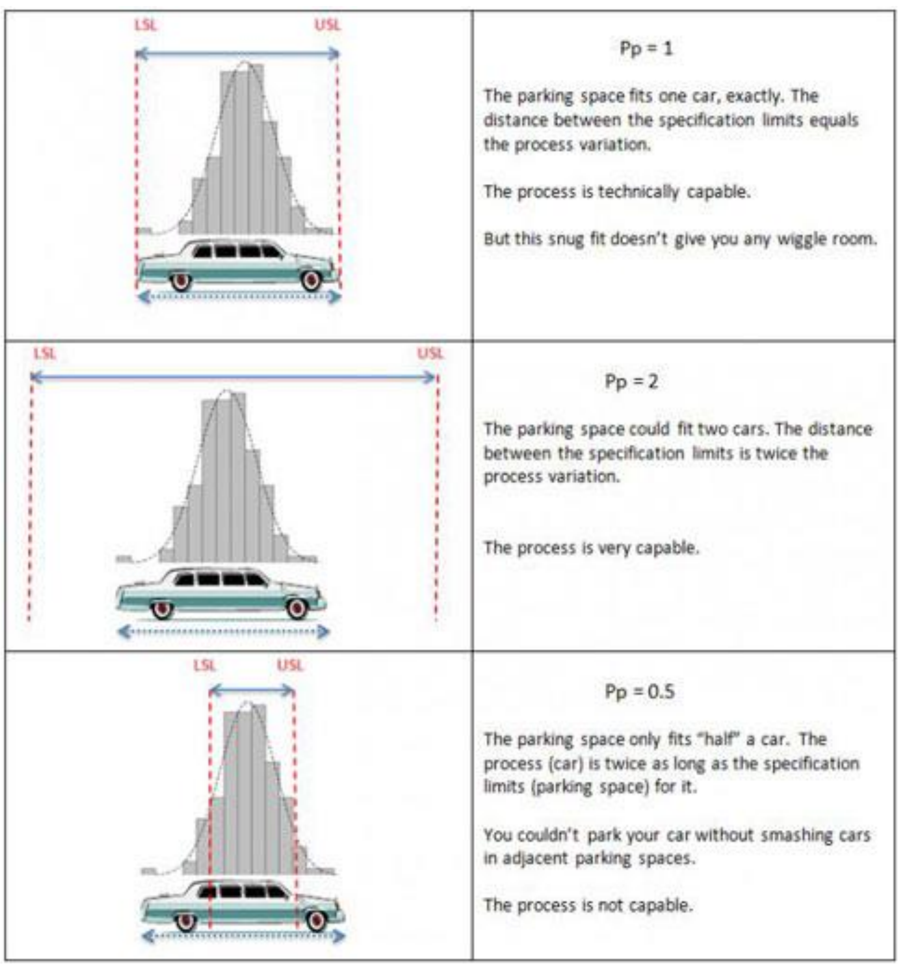
Sheriff Taylor: You got it, Barney. It's **how many times your process spread fits into your specification spread.** So the smaller the car, or the larger the parking space, in relation to each other, the higher the value of Pp and the more capable the process is.

Barney: Eureka! So all those p's stands for parking!

Sheriff Taylor: Well... I guess you could say that, Barney. But don't forget about the c's in the output, too, like C_p.

Barney: Just when I was getting a handle on things, you got to spring something new on me!

Sheriff Taylor: Don't worry, Barney. If you **understand Pp, you understand Cp. They're the same ratio—the specification spread to the process spread**. The only difference is that **Cp is calculated using the short-term variation** of the process—the **variation within subgroups**—while **Pp is calculated using overall variation** of the process. **Cp tells you the short-term** (or potential) capability, and **Pp tells you the long-term, actual capability.**



Barney (bored, playing with gun): Are ya done with your anomaly?

Sheriff Taylor: “Analogy,” Barney. Yes, I think that’s probably enough for one day. Why do you ask?

Barney: Because I got real work to do!

Sheriff Taylor: Real work?

Barney: Writing parking tickets, Andy! While we’ve been yapping, every Tom, Dick, and Harriet in Mayberry’s been out there parking illegally on Main Street!