SPC

LESSON: Xbar Charts - Homework

Homework 3 NAME:____

Topics: Xbar control chart, Type I & II errors

Solve the following problems and show your work.

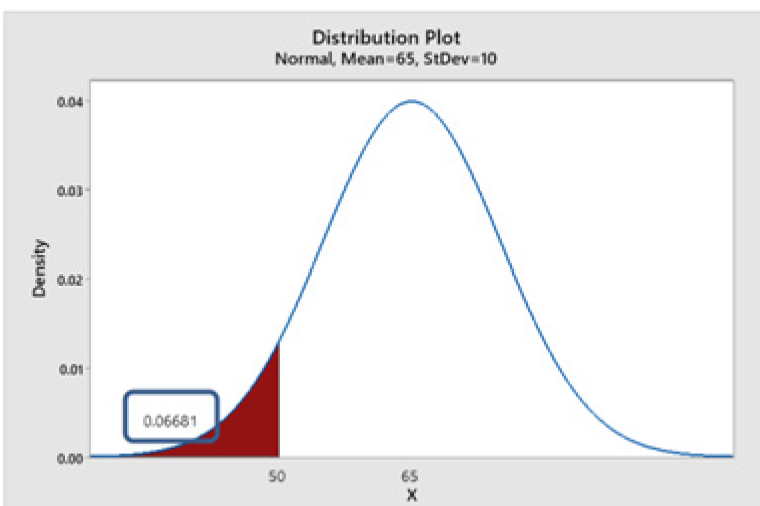
Problem 0. Warm-up Practice Problems from Lesson 7: Distribution of Xbar

Situation 1: Let X represent the time it takes a randomly selected student in the ONL class to finish this homework set. We'll assume that X ~ Normal (μ = 65 minutes, σ = 10 minutes; that is, the distribution of homework times is normally distributed with a mean of μ = 65 minutes and a standard deviation of σ = 10 minutes. Determine the following probabilities.

(a) What's the probability that we randomly choose a student (put names in a hat and drawn one) and the student finishes this homework set in less than 50 minutes? Show your work. If you are using Minitab, you can just sketch a picture of the distribution and shade the area corresponding to "less than 50 minutes."

Note: If it's not possible to determine this probability from what we've learned in class, just say "not possible."

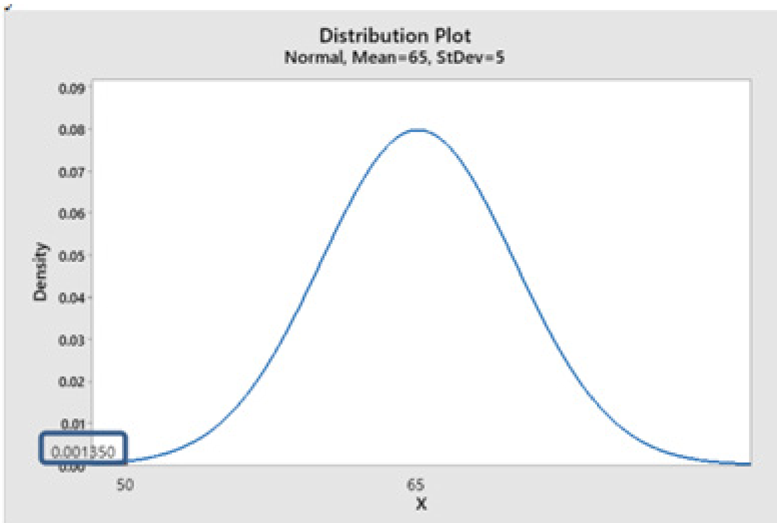
Solution: Since X is **normally distributed**, we can compute the value for just 1 randomly selected student. We know the distribution and its parameters (mean, std dev). Using Minitab's Probability Distribution Plot, we get **0.06681**.



(b) What's the probability that we randomly select 4 students and their average time to finish the homework set is less than 50 minutes? Show your work, which may just be sketching a distribution and shading it.

Note: If it's not possible to determine this probability from what we've learned in class, just say "not possible."

Solution: The average of normally distributed random variables is normal. We don't even need the CLT to know this. We need to adjust the standard deviation to be in sync with the sample size n = 4. We use Minitab's Probability Distribution Plot with the same mean as part (a): 65 minutes, but with a reduced standard deviation: 10 / sqrt(4) = 5. The answer is: **0.00135**



Situation 2: Let Y represent the time it takes a randomly selected student in the F2F class to finish this homework set. We'll assume that the distribution of Y is highly skewed (i.e. long right tail) with mean μ = 65 minutes, σ = 10 minutes. Determine the following probabilities.

(c) What's the probability that we randomly choose a student (put names in a hat and drawn one) and the student finishes this homework set in less than 50 minutes? Show your work. If you are using Minitab, you can just sketch a picture of the distribution and shade the area corresponding to "less than 50 minutes."

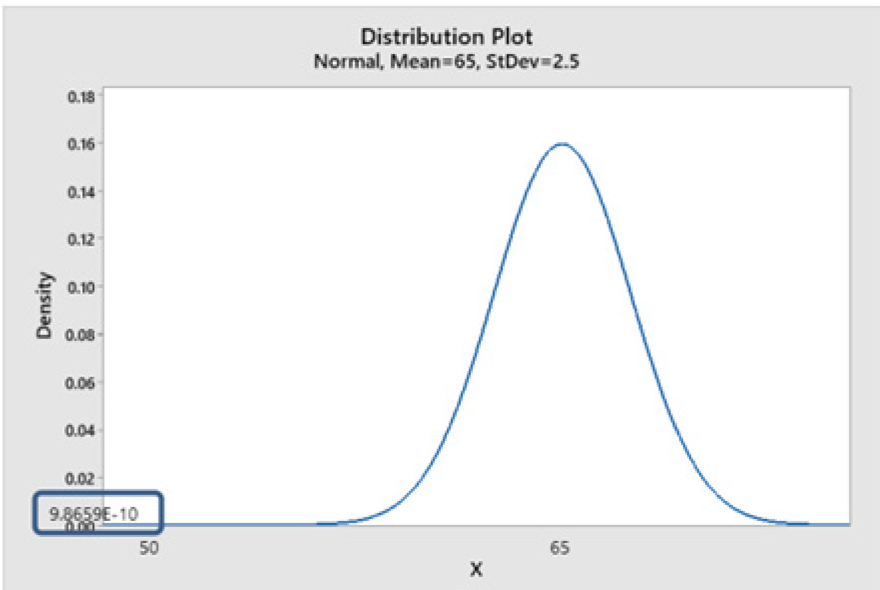
Note: If it's not possible to determine this probability from what we've learned in class, just say "not possible."

Solution: We don't know Y's distribution even though we know its mean and standard deviation. Without knowing its distribution, we can't compute a probability.

(d) What's the probability that we randomly select n = 16 students and their average time to finish the homework set is less than 50 minutes? From what we talked about in class, we will assume that n = 16 is a large enough sample size for the Central Limit Theorem to hold true. Show your work, which may just be sketching a distribution and shading it.

Note: If it's not possible to determine this probability from what we've learned in class, just say "not possible."

Solution: Even though you are told in your first Engineering Stats class that you need a sample size of n > 30 to take advantage of the CLT, you've seen in this class that in many cases the n = 16 is definitely enough. Because of the CLT, we know the **average of non-normal distribution** for a large enough n (in this case 16) is **normal**. We can compute this probability, but we do need to be careful to shrink the standard deviation by the square root of 16. The new standard deviation will be: 10/sqrt(16) = 2.5 minutes.



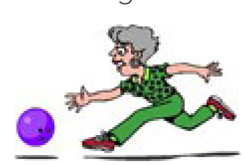
1. A person's **bowling score** is the quality characteristic of interest. Granny's average score (for the population of all games that she bowled in the last year) is **μ = 110** pins with a standard deviation of **σ = 4** pins.

To determine if Granny's bowling scores are currently in control, **k = 30 samples each of subgroup size n = 4**, will be recorded over the next month. (Example: Granny bowls next Friday night and her 4 bowling scores are 105, 110, 112, and 120. This is one subgroup of size 4 has an average of $\overline{X}$ = 111.75 pins and range R = 15 pins. A total of 30 such bowling trips or outings will be recorded.)

The upper and lower control limits (UCL and LCL) for the $\overline{X}$ control chart for her average bowling scores are Solution:

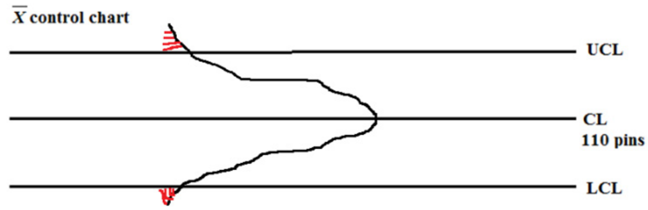
$$3\sigma \text{ upper limit: } \mathbf{110 + 3 \cdot \frac{4}{\sqrt{4}} \cong 116 \text{ pins}}$$

$$3\sigma \text{ lower limit: } \mathbf{110 - 3 \cdot \frac{4}{\sqrt{4}} \cong 104 \text{ pins}}$$



(a) Draw a normal curve for X̅ that is "in control" and centered about its mean in the control chart below. Then shade in the area of the normal curve that corresponds to the probability of a Type I error (**a**) and calculate this probability correct to 4 decimal places.

Solution:



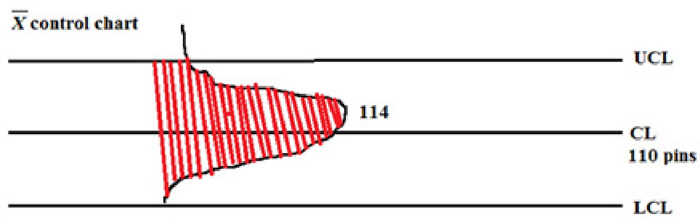
If the bowling scores are in-control, then the center of the X̅ curve is "on target" at 110. The probability of a Type I error (aka "false alarm") is getting a point beyond the 3 **σ** limits when X̅ is on target. This is:

$$P(\bar{X} > 116) + P(\bar{X} < 104) = P\left(\frac{\bar{X} - 110}{4/\sqrt{4}} > \frac{116 - 110}{4/\sqrt{4}}\right) + P\left(\frac{\bar{X} - 110}{4/\sqrt{4}} < \frac{104 - 110}{4/\sqrt{4}}\right)$$

$$= P(Z > 3) + P(Z < -3) \cong 2 \cdot 0.00135 = \mathbf{0.0027}$$

(b) If Granny's bowling improves and her true mean average shifts to 114 pins, draw a normal X̅ that is "out of control" with a mean of 114 pins on the control chart below. Determine the probability (correct to 4 decimal places) of not detecting this shift on Granny's first trip bowling after the shift has occurred and shade in this area on the X̅ chart above. This is the probability of a Type II Error (β).

Solution:



The X̅ curve has shifted and is centered at 114 – what's the probability of getting a point within the control limits (given the shift has occurred)? I could have used Minitab's Probability Distribution Plot to determine this value.

$$\beta = P(104 < \bar{X} < 116) = P\left(\frac{104 - 114}{4/\sqrt{4}} < \frac{\bar{X} - 114}{4/\sqrt{4}} < \frac{116 - 114}{4/\sqrt{4}}\right) = P(-5 < Z < 1) \cong 0.8413 - 0 = \mathbf{0.8413}$$

(c) Determine the probability (correct to 4 decimal places) of **detecting** this shift on Granny's first trip bowling after the shift has occurred. This value is called **power**. Power is the complement of Type II Error.

Solution: 1 – 0.8413 = **0.1587**

(d) [+0.5] If her process mean shifts to 115 pins, what is the probability (correct to 4 decimals places) of detecting this shift on her first bowling trip after the shift has occurred? That is, what's the **power** associated with this shift?

Solution: The $\overline{X}$ curve has shifted and is centered at 115 – what's the probability of getting a point outside of the control limits (given the shift has occurred)?

$$\text{Power} = 1 - \beta = 1 - P(104 < \bar{X} < 116) = 1 - P\left(\frac{104-115}{4/\sqrt{4}} < \frac{\bar{X}-114}{4/\sqrt{4}} < \frac{116-115}{4/\sqrt{4}}\right)$$
$$= 1 - P(-5.5 < Z < 0.5) \cong 1 - (0.6915 - 0) = \mathbf{0.3085}$$

(e) True or False. As the process mean shifts toward the upper control limit, the probability of a Type II Error (β) increases and the power decreases.

True **False**

(f) True or False. An **Operating Characteristic** (OC) curve is a graph showing the probability of committing a Type II Error as the process mean shifts away from its center.

True False

(g) Increase the sample size from n = 4 to n = 9. Compute the new UCL and LCL for Granny's bowling averages.

Solution: Using definitions of UCL and LCL, we have:

3σ upper limit: $\mathbf{110 + 3 \cdot \frac{4}{\sqrt{9}}} \cong$ **114 pins**

3σ lower limit: $\mathbf{110 - 3 \cdot \frac{4}{\sqrt{9}}} \cong$ **106 pins**

(h) Determine the probability (correct to 4 decimal places) of a Type I error by increasing the sample size from n = 4 to n = 9. Recall, for Type I error, the process mean does not shift.

Solution: Don't forget to **change your UCL and LCL** as calculated in part (g).

The probability of a Type I error (aka "false alarm" by Rule 1) is getting a point beyond the 3 **σ** limits when $\overline{X}$ is on target. This is:

$$P(\bar{X} > 114) + P(\bar{X} < 106) = P\left(Z > \frac{114-110}{4/\sqrt{9}}\right) + P\left(Z < \frac{106-110}{4/\sqrt{9}}\right)$$
$$= P(Z > 3) + P(Z < -3) \cong 2 \cdot 0.00135 = \mathbf{0.0026} \text{ or } \mathbf{0.0027}$$

(i) Compute the probability (correct to 4 decimal places) of a Type II error for a shift in her mean to 112 pins for n = 16. Show your work for full credit. Don't forget to update your LCL and UCL for a sample size of n = 16.

Solution: Don't forget to change your UCL and LCL for a sample size of n = 16.

$$3\sigma \text{ upper limit: } \mathbf{110 + 3 \cdot \frac{4}{\sqrt{16}} \cong 113 \text{ pins};}\ 3\sigma \text{ lower limit: } \mathbf{110 - 3 \cdot \frac{4}{\sqrt{16}} \cong 107 \text{ pins}}$$

$$\beta = P(107 < \bar{X} < 113) = P\left(\frac{107-112}{4/\sqrt{16}} < Z < \frac{113-112}{4/\sqrt{16}}\right) = P(-5 < Z < 1) \cong \mathbf{0.8413}$$

(j) Compute the probability (correct to 4 decimal places) of correctly detecting this shift (from part k) within the next 4 sample means (of size n = 16).

This means that you either:

- Detect the shift in the first observation (sample mean) after the shift.
  This is **POWER: 0.1587**.
- You don't detect the shift in the first observation and you detect it on the second observation (sample mean).
  P(Type II error on 1st observation) + P(detect on 2nd observation) = **0.8413 * 0.1587**
- You don't detect the shift on the first observation, you don't detect the shift on the second observation, and you detect the shift on the third observation (sample mean).
  0.8413^2 * 0.1587
- You don't … on the first, you don't … on the second, you don't … on the third, you DO on the fourth.
  0.8413^3 * 0.1587

You should be summing these four probabilities to obtain your answer. Hint: To move onto detecting it on the second means that you did not detect it on the first. So, for the 2nd bullet point, you're actually computing:

P(shift detected on 2nd sample mean | shift not detected on 1st sample mean) - P(shift not detected on 1st sample mean).

Since detecting the shift on the 1st and 2nd can be assumed to be independent, this simplifies to:

P(shift detected on 2nd sample mean) - P(shift not detected on 1st sample mean).

P(detect the shift BY the 4th observation) = $\sum_{i=1}^{4} 0.1587\, 0.8413^{i-1} = 0.4990394170$

0.4990