The class today contains new material, and continues where video 13 ended. When we do not have large amounts of data, we take small samples from two distributions [*we will look at two cases*]:

Two histograms overlapping each other When there is no difference	Two histograms apart from each other, with little overlap. When an actual, large difference exists

Our end goal: determine if $\mu_A \neq \mu_B$ in other words, did an improvement occur? Skipping ahead, this is what we want to achieve: a confidence interval for $\mu_B - \mu_A$

Where **a difference exists** (e.g. $\mu_B = 10$ and $\mu_A = 8$): $1.5 \leq \mu_B - \mu_A \leq 2.4$ (for example)

Where there is **no difference** (e.g. $\mu_B = 10$ and $\mu_A = 10$): $-0.8 \leq \mu_B - \mu_A \leq 1.6$ (for example)

For no difference, the confidence interval will span zero.
Note, the signs just flip if you switch A and B around.

This is one of the final equations we will end up with:

$$(\bar{x}_B - \bar{x}_A) - c_n\sqrt{\sigma^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)} < \mu_B - \mu_A < (\bar{x}_B - \bar{x}_A) + c_n\sqrt{\sigma^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}$$

1. Assume data for case A is normally distributed: $A \sim \mathcal{N}(\mu_A, \sigma_A^2)$
2. Assume data for case B is normally distributed: $B \sim \mathcal{N}(\mu_B, \sigma_B^2)$
3. Assume data for sample A and sample B have $\sigma_A = \sigma_B = \sigma$
4. From CLT (assumes data in A and data in B are independent)
   - $\mathcal{V}\{\bar{x}_A\} = \dfrac{\sigma_A^2}{n_A}$
   - $\mathcal{V}\{\bar{x}_B\} = \dfrac{\sigma_B^2}{n_B}$

Let's start with a derivation to get there.

We are going to create a z-value, and unpack it into a confidence interval.

5. Assume: $\bar{x}_A$ and $\bar{x}_B$ are independent [*likely true in many cases*]
   - $\mathcal{V}\left\{\bar{x}_B - \bar{x}_A\right\} = \dfrac{\sigma^2}{n_A} + \dfrac{\sigma^2}{n_B} = \sigma^2\left(\dfrac{1}{n_A} + \dfrac{1}{n_B}\right)$
6. Create a z-value:
   - $z = \dfrac{(\text{variable "}x\text{"}) - (\text{"location"})}{\text{"spread"}} = \dfrac{(\bar{x}_B - \bar{x}_A) - (\mu_B - \mu_A)}{\sqrt{\sigma^2\left(\dfrac{1}{n_A} + \dfrac{1}{n_B}\right)}}$
7. Create a confidence interval for z

$$(\bar{x}_B - \bar{x}_A) - c_n\sqrt{\sigma^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)} < \mu_B - \mu_A < (\bar{x}_B - \bar{x}_A) + c_n\sqrt{\sigma^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}$$

Let's look at how to use **step 6**, the z-value (it can be confusing!):

- Make the assumption that there really is no difference: $\mu_B = \mu_A$ in other words: $\mu_B - \mu_A = 0$
- You might recognize this as a null hypothesis, which you have learned about in a prior course.
- Consider two cases

What would typical values for z be if $\mu_B = \mu_A$ is **true** and you measured samples of data:	What would typical values for z be if $\mu_B = \mu_A$ is **false** and you measured samples of data:
A z-value would be close to zero (above and below zero) [*Hint*: Remember the video about the feedback controller?]	Your z-values would be large magnitude (positive or negative)

- Let's use some numbers: assume that a $z = 2$ was calculated from samples of data from A and B. The probability of getting a value of $z = 2$ from minus infinity up to +2 is 97.7%; so the probability of a value of 2 or greater is 100 - 97.7 = 2.3%
- That value of 2.3% is a clear signal our assumption of $\mu_B = \mu_A$ was wrong. We have very low probability of being correct. Conversely, we are almost certain that the true average of A (μ_A) and the true average of B (μ_B) are different.
- Let's look at the opposite case: assume a change has happened, so assuming $\mu_B = \mu_A$ is wrong. Take samples, and imagine you get a z-value of 2.0. It shows that assuming no change was a bad assumption, because that z-value has a low probability of occurring: 2.3%. That confirms to us that an actual change has occurred between system A and B.
- There is only a 2.3% risk that you are wrong in saying that system A and B are different, and 97.7% chance that you are correct in concluding they are different.

Now, as I said this can be confusing initially. Risk and probabilities can be opposites of each other. This can be confusing, so let's look at quantifying this as a confidence interval (far more intuitive to engineers).

In step 7 we expanded the z value between lower and upper critical values, $\pm c_n$(we saw this process last week): $-c_n \leq z \leq +c_n$

The question is what values to use. Let's look at the example from the video, the feedback controllers. Sub in in these numbers:

- $\bar{x}_A = 79.9$and $\bar{x}_B = 82.9$and $\bar{x}_B - \bar{x}_A = 3.04$
- $\sigma = 6.61$(found by using all the 300 data points, called an external estimate of spread)
- $n_A = n_B = 10$
- $c_n = 1.96$ for a 95% confidence interval, read from tables, or use `qnorm(0.025)` or `qnorm(0.975)`
- So the lower bound for the interval is (82.9 - 79.9) - 1.96 * sqrt(6.61^2*2/10) = -2.8
- The upper bound for the interval is (82.9 - 79.9) + 1.96 * sqrt(6.61^2*2/10) = +8.8
- Interpretation of the interval:

  The interval spans zero, so there is a clear possibility that the effect from the new controller is negligible in the long term. There might be a benefit to be had, because the asymmetry of the range, but if we have to pay a substantial amount of money for the controller, I would not buy it. If we were getting the controller for free, at least we would be no worse off (maybe, just maybe) a little better off.

  If we could, I would also ask to run a few more samples on the new controller, to get a clearer idea of the confidence level. It is in the vendor's favour to do this: more data points implies a narrower confidence interval, and maybe we will purchase it. Right now the statistics do not favour the vendor, but if they let us do a few more runs, they might end up making a sale.

- The z-value is 1.03 if we assume $\mu_B = \mu_A$and the area from negative infinity to this point corresponds to 84.8%, meaning there is a risk of 100-84.8 = 15.2% that we are wrong is saying the new controller is different [compare that to the dot plot result, of 11% risk].

Let's move onto a final point: in the above we use the variance from the 300 data points as a population variance. Now we are going to only use the 10+10 values. This is called an internal estimate of spread.

To calculate the overall variance, we "pool" the individual variances:

$$s_P^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A - 1 + n_B - 1}$$

It is a weighted sum of the variances. This is a common statistical technique to improve the variance estimate.

$$s_P^2 = \frac{9 \times 6.81^2 + 9 \times 6.70^2}{18} = 45.63$$

But, because we are estimating the variance from data (not using the population), the **z-value is now t-distributed with** $n_A - 1 + n_B - 1$ **degrees of freedom.**

$$(\bar{x}_B - \bar{x}_A) - c_t\sqrt{s_P^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)} < \mu_B - \mu_A < (\bar{x}_B - \bar{x}_A) + c_t\sqrt{s_P^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}$$

$$-c_t \leq z \leq +c_t$$

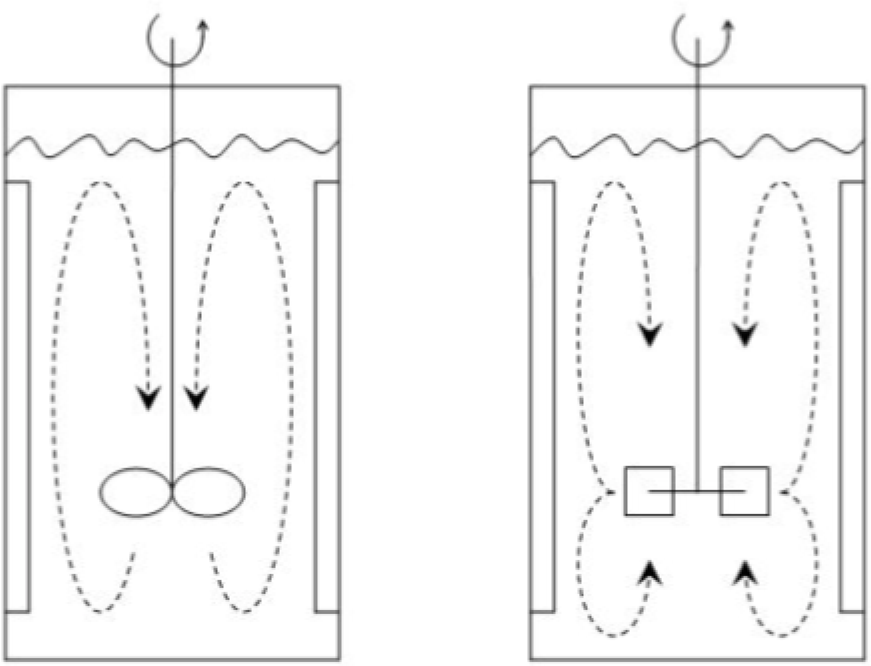
What is the lower bound value: (82.9 - 79.9) - 2.1 * sqrt(45.63*2/10) = -3.3

What is the upper bound value: (82.9 - 79.9) - 2.1 * sqrt(45.63*2/10) = +9.3

c_t = qt(0.975, 18) = 2.1

Notice how the confidence intervals have widened in this case (due to propagation of our error is estimating the variance).

We need more practice in interpreting and using the confidence interval. It is used the same way we learned about in class 03B.



Axial (left) and radial (right) impellers give different mixing times.
The objective is to have the *shortest* mixing times possible.

Try these cases:

- $43 \text{ min} < \mu_{\text{Axial}} - \mu_{\text{Radial}} < 95 \text{ min}$
- $-95 \text{ min} < \mu_{\text{Radial}} - \mu_{\text{Axial}} < -43 \text{ min}$
- $-12 \text{ min} < \mu_{\text{Axial}} - \mu_{\text{Radial}} < -7 \text{ min}$
- $-453 \text{ min} < \mu_{\text{Axial}} - \mu_{\text{Radial}} < 284 \text{ min}$
- $-21 \text{ min} < \mu_{\text{Axial}} - \mu_{\text{Radial}} < 187 \text{ min}$

- Radial
- Radial

- Axial
- Radial
- Radial

Be sure you can explain why in each case.