Math 4720: Statistical Methods

7^{th} Week Summary (2/27/25)

- In practice, we don't know σ, so we substitute s for σ. The resulting statistic does not have a normal distribution.
- Draw a random sample of size n from a normal distribution with mean μ and standard deviation σ. The **one-sample t statistic**: $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ has the **t distribution** with $n-1$ degrees of freedom.

  The p_value for a test of $H_0 : \mu = \mu_0$ against

  $H_a : \mu > \mu_0$ is p-value $= P(T \geq t)$

  $H_a : \mu < \mu_0$ is p-value $= P(T \leq t)$

  $H_a : \mu \neq \mu_0$ is p-value $= 2P(|T| \geq |t|)$

  These p-values are exact if the population distribution is Normal; they are approximately correct for large n in other cases.
- the t distribution with $n-1$ degrees of freedom approaches the standard normal distribution $N(0,1)$ as n increases.
- Confidence interval for μ is $\bar{x} \pm t_{\alpha/2}\dfrac{s}{\sqrt{n}}$ with $df = n-1$, if

  data is coming from normal population having unknown mean μ and unknown σ, or

  large random sample from a non-normal population having unknown mean μ and unknown σ.

- **Comparing the means of two populations** ....
- **Independent samples** : A so-called two-sample problem can arise from a *randomized comparative experiment* that randomly divides subjects into two groups and exposes each group to a different treatment.

  To perform inference about $\mu_1 - \mu_2$, the difference between the means of the two populations, we start from $\bar{x}_1 - \bar{x}_2$, the difference between the means of the two samples.

Two Numerical Variables – Population Standard Deviations unknown

	Independent Samples	
	$\sigma_1 = \sigma_2$	$\sigma_1 \neq \sigma_2$
Parameters of Interest:	Mean Difference, $\mu_1 - \mu_2$	Mean Difference, $\mu_1 - \mu_2$
Confidence Interval Formula:	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ $df = n_1 + n_2 - 2$, where $s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$ $df = \frac{(n_1-1)(n_2-1)}{(1-c)^2(n_1-1)+c^2(n_2-1)}$, where $c = \frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$
Name of Hypothesis Test, H_0	Two-sample T Test, $H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 - \mu_2 = 0$	Two-sample T Test, $H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 - \mu_2 = 0$
Test Statistic Formula:	$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$
P-value:	$H_a: \mu_1 \neq \mu_2$, p-value = $2P(T \geq \lvert t \rvert)$ $H_a: \mu_1 > \mu_2$, p-value = $P(T \geq t)$ $H_a: \mu_1 < \mu_2$, p-value = $P(T \leq t)$	