**2-sample methods**

*Difference of 2 (independent) means*:

Some archaeologists theorize that ancient Egyptians interbred with several different immigrant populations over thousands of years. To see if there is any indication of changes in body structure that might have resulted, in a random sample they measured 30 skulls of male Egyptians dated from 4000 BCE and 30 others dated from 200 BCE.

- Is there sufficient evidence that the mean breadth of males' skulls increased (as theorized by archaeologists) over this period? Conduct hypothesis test (all 5 steps)
  1. Hypotheses, assumptions if requested
  $$H_0: \mu_{200} = \mu_{4000}\ (or\ \mu_1 = \mu_2)\ vs. H_a: \mu_{200} > \mu_{4000}\ (or\ \mu_1 > \mu_2)$$
  Assumptions: Independence (is random), Randomization (yes), normality $n = 30$ so yes
  2. State test statistic, $df$, $pvalue$: t = 3.5797, df = 54.973, p-value = 0.000364
  3. Results: reject null if $pvalue \leq \alpha(0.05)$
  $pvalue = 0.000364 \leq \alpha(0.05) \therefore H_0$ is rejected
  4. There is sufficient (significant) evidence that Egyptian male skulls breadths have increased from 4000 BCE to 200 BCE
  5. Error: since we rejected null, we could have made a Type I error. We think there is an increase in skull breadths when there was not
- Estimate the true difference of means with 95% confidence and interpret.
  CI: 1.878030, 6.655303. We are 95% confident the true difference in mean skull breadths is between 1.88 and 6.66 mm. The skulls from 200 BCE are between 1.88 and 6.66 mm larger than the skulls from 4000 BC

Welch Two Sample t-test

data:  breadth by era
t = 3.5797, df = 54.973, p-value = 0.000364
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 2.27257     Inf
sample estimates:
 mean in group 200BCE mean in group 4000BCE
           135.6333            131.3667

Welch Two Sample t-test

data:  breadth by era
t = 3.5797, df = 54.973, p-value = 0.000728
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.878030 6.655303

sample estimates:
mean in group 200BCE mean in group 4000BCE
135.6333 131.3667

When doing a one tailed test (either upper or lower), if you need a CI, a separate analysis will be done and provided for you

*Difference of 2 independent proportions*

Sludge is a dried product remaining from processed sewage and is often used as a fertilizer on crops; there could be dangerous concentrations of nickel in the crops. A new method of processing sewage has been developed and a randomized experiment conducted to evaluate its effectiveness in removing heavy metals. Sewage of a known concentration of nickel is treated using both old and new methods and applied to 100 tomato plants that were randomly assigned to pots containing sewage sludge processed by one of the two methods and the nickel was measured in the tomatoes.

- Is there sufficient evidence that the concentration of nickel from the new treatment is less than the old treatment?
  1. Hypotheses, assumptions
  $$H_0: \pi_{new} = \pi_{old}\ (or\ \pi_1 = \pi_2)\ \ vs. H_a: \pi_{new}\ < \pi_{old}\ (or\ \pi_1 < \pi_2)$$
  Assumptions: independence (random), random (yes), normality: usually we want $n_i \geq 60$ . Here we have $n_i = 50$ ➔ it will be ok for this example
  2. State test statistic, df, pvalue: t = -1.1489, df = 92.56, p-value = 0.1268
  3. Results: reject null if $pvalue \leq \alpha(0.05)$
  $pvalue = 0.1268 \nleq \alpha(0.05) \therefore H_0$ is NOT rejected
  4. We did not reject the null, so there is no evidence that the new method is better than the old method. There is no significant difference between new and old method.
  5. Error: since the null was not rejected, we could have a Type II error. We think there is no difference in the amount of metals removed from new or old treatment, but the new method could be better
- Estimate the true difference of proportions with 95% confidence and interpret
  CI: -0.2182892, 0.0582892 Since the CI contains 0, we say that there is no significant difference between the two groups.

Welch Two Sample t-test

data: sludge.new and sludge.old
t = -1.1489, df = 92.56, p-value = 0.1268
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
-Inf 0.03569674
sample estimates:

mean of x mean of y
    0.10      0.18

Welch Two Sample t-test

data:  sludge.new and sludge.old
t = -1.1489, df = 92.56, p-value = 0.2536
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.2182892  0.0582892 ➔ (mu1-mu2=0)
sample estimates:
mean of x mean of y
    0.10      0.18