SUMMARY: Confidence Interval for the Difference Between Two Population Proportions

A confidence interval for the difference $(p_1 - p_2)$ between two population proportions is

$$(\hat{p}_1 - \hat{p}_2) \pm z(se), \text{ where } se = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}.$$

The z-score depends on the confidence level, such as $z = 1.96$ for 95% confidence. To use this method, you need

- A categorical response variable for two groups
- Independent random samples for the two groups, either from random sampling or a randomized experiment
- Large enough sample sizes n_1 and n_2 so that, in each sample, there are at least 10 successes and at least 10 failures. The confidence interval for a single proportion required at least 15 successes and 15 failures. Here, the method works well with slightly smaller samples, at least 10 of each type in each group.

SUMMARY: Interpreting a Confidence Interval for a Difference of Proportions

- Check whether 0 falls in the confidence interval. If so, it is plausible (but not necessary) that the population proportions are equal.
- If all values in the confidence interval for $(p_1 - p_2)$ are positive, you can infer that $(p_1 - p_2) > 0$, or $p_1 > p_2$. The interval shows just how much larger p_1 might be. If all values in the confidence interval are negative, you can infer that $(p_1 - p_2) < 0$, or $p_1 < p_2$.
- The magnitude of values in the confidence interval tells you how *large* any true difference is. If all values in the confidence interval are near 0, the true difference may be relatively small in practical terms.

SUMMARY: Two-Sided Significance Test for Comparing Two Population Proportions

1. **Assumptions**
   - A categorical response variable for two groups
   - Independent random samples, either from random sampling or a randomized experiment
   - n_1 and n_2 are large enough that there are at least five successes and five failures in each group if using a two-sided alternative
2. **Hypotheses**

   Null H_0: $p_1 = p_2$ (that is, $p_1 - p_2 = 0$)

   Alternative H_a: $p_1 \neq p_2$ (one-sided H_a also possible; see after Example 5)
3. **Test Statistic**

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{se_0} \text{ with } se_0 = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)},$$

   where $\hat{p}$ is the pooled estimate.
4. **P-value**

   P-value = Two-tail probability from standard normal distribution (Table A) of values even more extreme than observed z test statistic presuming the null hypothesis is true
5. **Conclusion**

   Smaller P-values give stronger evidence against H_0 and supporting H_a. Interpret the P-value in context. If a decision is needed, reject H_0 if P-value $\leq$ significance level (such as 0.05).

SUMMARY: Confidence Interval for Difference Between Population Means

For two samples with sizes n_1 and n_2 and standard deviations s_1 and s_2, a 95% confidence interval for the difference $(\mu_1 - \mu_2)$ between the population means is

$$(\bar{x}_1 - \bar{x}_2) \pm t_{.025}(se), \text{ with } se = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Software provides $t_{.025}$, the t-score with right-tail probability 0.025 (total probability = 0.95 between $-t_{.025}$ and $t_{.025}$).

This method assumes:

- Independent random samples from the two groups, either from random sampling or a randomized experiment.
- An approximately normal population distribution for each group. (This is mainly important for small sample sizes, and even then the method is robust to violations of this assumption.)

SUMMARY: Two-Sided Significance Test for Comparing Two Population Means

1. **Assumptions**
   - A quantitative response variable for two groups
   - Independent random samples, either from random sampling or a randomized experiment
   - Approximately normal population distribution for each group. (This is mainly important for small sample sizes, and even then the two-sided test is robust to violations of this assumption.)
2. **Hypotheses**

   H_0: $\mu_1 = \mu_2$

   H_a: $\mu_1 \neq \mu_2$ (one-sided H_a: $\mu_1 > \mu_2$ or H_a: $\mu_1 < \mu_2$ also possible)
3. **Test Statistic**

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{se} \text{ where } se = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$
4. **P-value**

   P-value = Two-tail probability from t distribution of values even more extreme than observed t test statistic, presuming the null hypothesis is true with df given by software.
5. **Conclusion**

   Smaller P-values give stronger evidence against H_0 and supporting H_a. Interpret the P-value in context, and if a decision is needed, reject H_0 if P-value $\leq$ significance level (such as 0.05).

SUMMARY: Comparing Population Means, Assuming Equal Population Standard Deviations

Using the pooled standard deviation estimate s of $\sigma = \sigma_1 = \sigma_2$, the standard error of $(\bar{x}_1 - \bar{x}_2)$ simplifies to

$$se = \sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}} = s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

Otherwise, inference formulas look the same as those that do not assume $\sigma_1 = \sigma_2$:

- A 95% confidence interval for $(\mu_1 - \mu_2)$ is $(\bar{x}_1 - \bar{x}_2) \pm t_{.025}(se)$.
- The test statistic for H_0: $\mu_1 = \mu_2$ is $t = (\bar{x}_1 - \bar{x}_2)/se$.

These methods have $df = n_1 + n_2 - 2$. To use them, you assume

- Independent random samples from the two groups, either from random sampling or a randomized experiment
- An approximately normal population distribution for each group (This is mainly important for small sample sizes, and even then the confidence interval and two-sided test are usually robust to violations of this assumption.)
- $\sigma_1 = \sigma_2$ (In practice, this type of inference is not usually relied on if one sample standard deviation is more than double the other one.)

For Dependent Samples, Mean of Differences = Difference of Means

For dependent samples, the difference $(\bar{x}_1 - \bar{x}_2)$ between the means of the two samples equals the mean $\bar{x}_d$ of the difference scores for the matched pairs.

SUMMARY: Comparing Means of Dependent Samples

To compare means with dependent samples, construct confidence intervals and significance tests using the single sample of difference scores,

$$d = \text{observation in Sample 1} - \text{observation in Sample 2}.$$

The 95% confidence interval $\bar{x}_d \pm t_{.025}(se)$ and the test statistic $t = (\bar{x}_d - 0)/se$ are the same as for a single sample. The assumptions are also the same: A random sample or a randomized experiment and a normal population distribution of difference scores.

SUMMARY: McNemar Test Comparing Proportions From Dependent Samples

Hypotheses: H_0: $p_1 = p_2$, H_a can be two-sided or one-sided.

Test Statistic: For the two counts for the frequency of yes on one response and no on the other, the z test statistic equals their difference divided by the square root of their sum. The sum of the counts should be at least 30, but in practice the two-sided test works well even if this is not true.

P-value: For H_a: $p_1 \neq p_2$, two-tail probability of z test statistic values more extreme than observed z, using standard normal distribution.