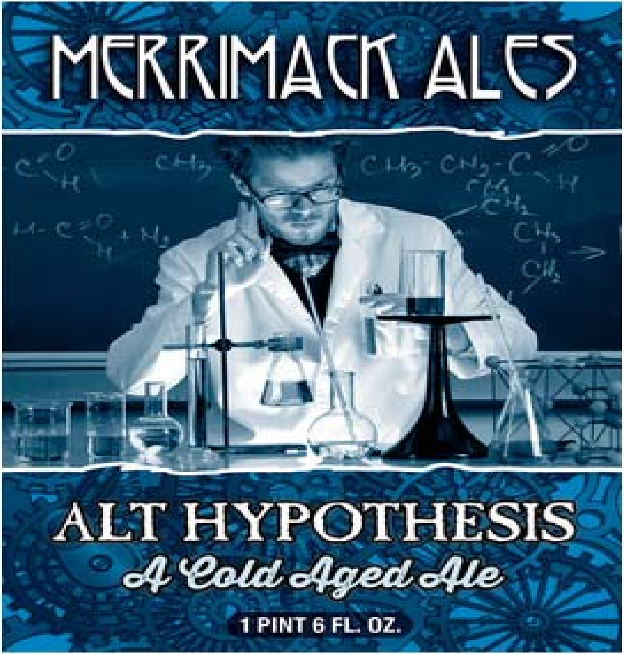
MERRIMACK ALES
ALT HYPOTHESIS
A Cold Aged Ale
1 PINT 6 FL. OZ.

Fixed Significance Level Tests

We have conducted hypothesis tests by computing a p-value to measure the evidence against H_0 and in favor of H_a. When we wanted to conduct the tests at a fixed level of significance, α, we first computed the p-value and rejected H_0 if and only if the p-value was less than α.

Fixed Significance Level Tests

We have conducted hypothesis tests by computing a p-value to measure the evidence against H_0 and in favor of H_a. When we wanted to conduct the tests at a fixed level of significance, α, we first computed the p-value and rejected H_0 if and only if the p-value was less than α.

An alternative way to conduct a test at a fixed level of significance is to determine which values of the test statistic will lead to rejection of H_0 in favor of H_a. Here are the steps involved, illustrated using the LDL reduction example:

$\approx$

Fixed Significance Level Tests

1. **Specify hypotheses to be tested.**

$$\begin{aligned} H_0 : \quad \mu &= 0 \\ H_a : \quad \mu &> 0 \end{aligned}$$

(i.e. $\mu_0 = 0$)

Fixed Significance Level Tests

1. **Specify hypotheses to be tested.**

$$\begin{aligned} H_0 &: \quad \mu \quad = \quad 0 \\ H_a &: \quad \mu \quad > \quad 0 \end{aligned}$$

   (i.e. $\mu_0 = 0$)

2. **Set the significance level α.** Usual choices are 0.10, 0.01 or 0.05. We'll choose the latter.

$\approx$

Fixed Significance Level Tests

3. **Specify the (standardized) test statistic and it's distribution under $\mathbf{H_0}$.** The standardized test statistic is

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}} = \frac{\overline{y} - 0}{15.45/\sqrt{10}},$$

and under H_0 it has a $t_{n-1} = t_9$ distribution.

$\approx$

4. **Find the critical region of the test.** The critical region of the test is the set of values of the (standardized) test statistic for which H_0 will be rejected in favor of H_a. Here, H_a tells us that the critical region has the form

$$[t_{n-1,1-\alpha}, \infty) = [t_{9,0.95}, \infty) = [1.8331, \infty),$$

meaning H_0 will be rejected if and only if the observed value of t is greater than or equal to 1.8331.

$\approx$

Fixed Significance Level Tests

5. **Perform the test.** For the LDL example, the observed value of t is

$$t^* = \frac{21.19 - 0}{15.45/\sqrt{10}} = 4.337,$$

which falls in the critical region, so H_0 is rejected in favor of H_a.

$\approx$

Power of a Test

In a fixed significance level test, power is the proportion of all samples for which H_0 will be rejected in favor of H_a. Power will vary for different values of the parameter being tested, so it is written as a function of that parameter.

$\approx$

Power of a Test

Example: A random sample of size n is taken from a $N(\mu, 25)$ population. We want to test $H_0 : \mu = 10$ versus $H_a : \mu < 10$ at the 0.05 level of significance using a fixed significance level test.

Power of a Test

Example: A random sample of size n is taken from a $N(\mu, 25)$ population. We want to test $H_0 : \mu = 10$ versus $H_a : \mu < 10$ at the 0.05 level of significance using a fixed significance level test.

The test statistic is $\overline{y}$, which under H_0 has distribution $N(10, 25/n)$, where n is the sample size.

Power of a Test

Example: A random sample of size n is taken from a $N(\mu, 25)$ population. We want to test $H_0 : \mu = 10$ versus $H_a : \mu < 10$ at the 0.05 level of significance using a fixed significance level test.

The test statistic is $\overline{y}$, which under H_0 has distribution $N(10, 25/n)$, where n is the sample size.

The form of H_a tells us that small values of $\overline{y}$ should lead to rejection of H_0, which means small values of the standardized statistic $z = (\overline{y} - 10)/(5/\sqrt{n})$ should lead to rejection.

Power of a Test

Example: A random sample of size n is taken from a $N(\mu, 25)$ population. We want to test $H_0 : \mu = 10$ versus $H_a : \mu < 10$ at the 0.05 level of significance using a fixed significance level test.

The test statistic is $\overline{y}$, which under H_0 has distribution $N(10, 25/n)$, where n is the sample size.

The form of H_a tells us that small values of $\overline{y}$ should lead to rejection of H_0, which means small values of the standardized statistic $z = (\overline{y} - 10)/(5/\sqrt{n})$ should lead to rejection.

Since $Pr(z \leq z_{0.05}) = Pr(z \leq -1.645) = 0.05$, the rejection region is defined by $(\overline{y} - 10)/(5/\sqrt{n}) \leq -1.645$, which after some algebra, becomes $\overline{y} \leq 10 - (1.645)(5)/\sqrt{n}$.

$\approx$

Power of a Test

To compute the power of this test, we need to evaluate the proportion of all samples which lead to rejection when the true population mean is μ.

Power of a Test

To compute the power of this test, we need to evaluate the proportion of all samples which lead to rejection when the true population mean is μ.

We'll write this as

$$\Pi(\mu) = Pr_{\mu}(\text{reject } H_0) = Pr_{\mu}(\bar{y} \leq 10 - (1.645)(5)/\sqrt{n}),$$

for all values $\mu < 10$.

$\approx$

Power of a Test

We evaluate by standardizing:

$$\begin{aligned}\Pi(\mu) &= Pr_{\mu}\left(\frac{\overline{y}-\mu}{5/\sqrt{n}} \leq \frac{10-(1.645)(5)/\sqrt{n}-\mu}{5/\sqrt{n}}\right) \\ &= Pr(z \leq \sqrt{n}(10-\mu)/5 - 1.645).\end{aligned}$$

Power of a Test

We evaluate by standardizing:

$$\begin{aligned}\Pi(\mu) &= Pr_{\mu}\left(\frac{\overline{y}-\mu}{5/\sqrt{n}} \leq \frac{10-(1.645)(5)/\sqrt{n}-\mu}{5/\sqrt{n}}\right)\\ &= Pr(z \leq \sqrt{n}(10-\mu)/5 - 1.645).\end{aligned}$$

For any value of μ, $\Pi(\mu)$ can be computed using online applets, statistical software, a calculator or a table of the normal distribution.

$\approx$

Power of a Test

For instance, if $n = 16$ and $\mu = 7$, we get

$$\begin{aligned}
\Pi(\mu) &= Pr(z \leq \sqrt{n}(10-\mu)/5 - 1.645) \\
&= Pr(z \leq \sqrt{16}(10-7)/5 - 1.645) \\
&= Pr(z \leq 0.755) = 0.775
\end{aligned}$$

Power of a Test

For instance, if $n = 16$ and $\mu = 7$, we get

$$\begin{aligned}\Pi(\mu) &= Pr(z \leq \sqrt{n}(10-\mu)/5 - 1.645) \\ &= Pr(z \leq \sqrt{16}(10-7)/5 - 1.645) \\ &= Pr(z \leq 0.755) = 0.775\end{aligned}$$

Here is a plot of the power functions, $\Pi(\mu)$ for $n = 16$ and $n = 32$:

$\approx$

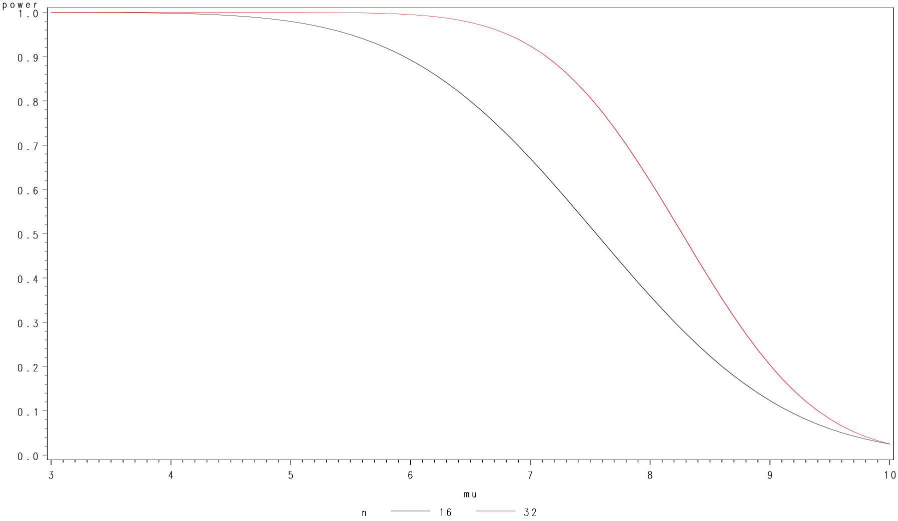

Figure: 2: Power of the one-sided test.

Power of a Test

Consider again the random sample of size n is taken from a $N(\mu, 25)$ population. We will find a formula for the rejection region of a fixed significance level test of $H_0 : \mu = 10$ versus $H_a : \mu \neq 10$ at the 0.05 level of significance. The rejection region is defined by the absolute value of the standardized test statistic exceeding a specified value.

$\approx$

Power of a Test

To define the rejection region, We need to find a value A so that $Pr(|z| \geq A) = 0.05$, where $z = (\bar{y} - 10)/(5/\sqrt{n})$.

Power of a Test

To define the rejection region, We need to find a value A so that $Pr(|z| \geq A) = 0.05$, where $z = (\bar{y} - 10)/(5/\sqrt{n})$.

Since $z \sim N(0, 1)$, $A = z_{0.975} = 1.96$.

Power of a Test

To define the rejection region, We need to find a value A so that $Pr(|z| \geq A) = 0.05$, where $z = (\overline{y} - 10)/(5/\sqrt{n})$.

Since $z \sim N(0, 1)$, $A = z_{0.975} = 1.96$.

Therefore, the rejection region is defined by $|(\overline{y} - 10)/(5/\sqrt{n})| \geq 1.96$, which, after some algebra, specifies rejection if and only if $\overline{y} \geq 10 + (1.96)(5)/\sqrt{n}$ or $\overline{y} \leq 10 - (1.96)(5)/\sqrt{n}$.

$\approx$

We will now find a formula for the power function of this test. Remember the specifics:

Power of a Test

We will now find a formula for the power function of this test. Remember the specifics:

- $\overline{y}$ is the mean of a sample of size n taken from a $N(\mu, 25)$ population.

Power of a Test

We will now find a formula for the power function of this test. Remember the specifics:

- $\overline{y}$ is the mean of a sample of size n taken from a $N(\mu, 25)$ population.
- A level 0.05 test of $H_0 : \mu = 10$ versus $H_a : \mu \neq 10$ rejects H_0 if and only if

$$\overline{y} \geq 10 + (1.96)(5)/\sqrt{n} \text{ or } \overline{y} \leq 10 - (1.96)(5)/\sqrt{n}.$$

$\approx$

Power of a Test

The power function is

$$
\begin{aligned}
\Pi(\mu) &= Pr_\mu(\text{reject } H_0) \\
&= Pr_\mu(\overline{y} \leq 10 - (1.96)(5)/\sqrt{n}) + Pr_\mu(\overline{y} \geq 10 + (1.96)(5)/\sqrt{n}) \\
&= Pr(z \leq \sqrt{n}(10 - \mu)/5 - 1.96) \\
&+ Pr(z \geq \sqrt{n}(10 - \mu)/5 + 1.96)
\end{aligned}
$$

Power of a Test

The power function is

$$
\begin{aligned}
\Pi(\mu) &= Pr_\mu(\text{reject } H_0) \\
&= Pr_\mu(\bar{y} \le 10 - (1.96)(5)/\sqrt{n}) + Pr_\mu(\bar{y} \ge 10 + (1.96)(5)/\sqrt{n}) \\
&= Pr(z \le \sqrt{n}(10-\mu)/5 - 1.96) \\
&+ Pr(z \ge \sqrt{n}(10-\mu)/5 + 1.96)
\end{aligned}
$$

Here's a plot for $n = 16$ and 32:

$\approx$

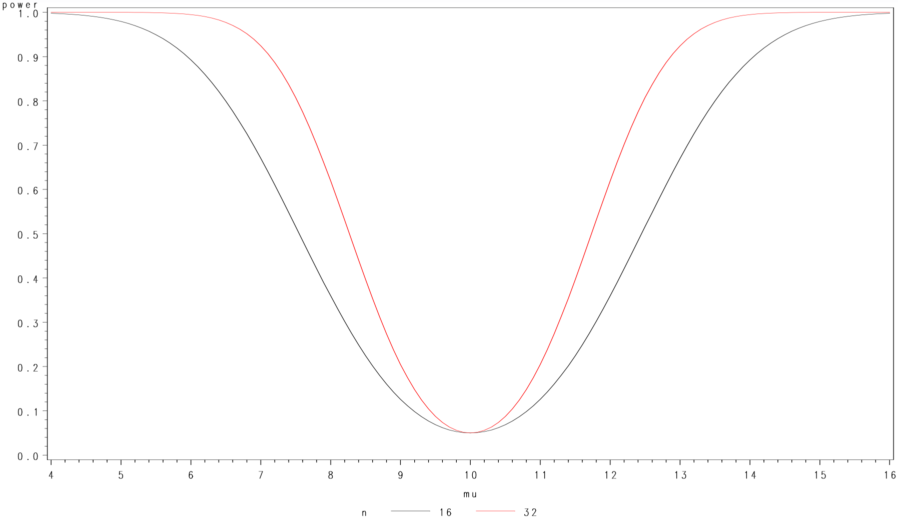

Figure: 2: *Power of the two-sided test.*

Power of a Test

You have seen that the power curve depends on the sample size, n. This means the power function can also be used to specify a sample size. If the researcher specifies a significance level for the test and a desired power at a specified value of the parameter being tested, then using the formulas given, (s)he can find what value of n is needed.

Power of a Test

You have seen that the power curve depends on the sample size, n. This means the power function can also be used to specify a sample size. If the researcher specifies a significance level for the test and a desired power at a specified value of the parameter being tested, then using the formulas given, (s)he can find what value of n is needed.

In our example of a test for a population mean μ, the researcher would specify a significance level α and a desired power $\Pi(\mu_0)$ at the particular mean value μ_0. A minimum sample size n to satisfy these requirements would then be obtained.

$\approx$

The relation between hypothesis tests and confidence intervals

There is an interesting parallel between many hypothesis tests and confidence intervals. Specifically, to each fixed-level hypothesis test we have considered, there corresponds a confidence interval that can be used to conduct the test.

$\approx$

The relation between hypothesis tests and confidence intervals

As an example, consider the following two inference procedures for a population mean μ, conducted on the same set of data:

- A two-sided level α test of $H_0 : \mu = \mu_0$ versus $H_a : \mu \neq \mu_0$.
- A level $L = 1 - \alpha$ confidence interval for μ.

$\approx$

The relation between hypothesis tests and confidence intervals

The relation can be summarized in two ways:

The relation between hypothesis tests and confidence intervals

The relation can be summarized in two ways:

The test will reject H_0 in favor of H_a if and only if μ_0 lies outside the confidence interval, so we can use the confidence interval to conduct the test: just construct the interval and if μ_0 is outside the interval, reject H_0 in favor of H_a.

The relation between hypothesis tests and confidence intervals

The relation can be summarized in two ways:

The test will reject H_0 in favor of H_a if and only if μ_0 lies outside the confidence interval, so we can use the confidence interval to conduct the test: just construct the interval and if μ_0 is outside the interval, reject H_0 in favor of H_a.

Similarly, given the level α test of $H_0 : \mu = \mu_0$ versus $H_a : \mu \neq \mu_0$, a level $L = 1 - \alpha$ confidence interval for μ consists of all μ_0 values for which the test does not reject H_0 in favor of H_a (this is called "inverting the test").

$\approx$

The relation between hypothesis tests and confidence intervals

The following quote neatly summarizes the relation between hypothesis tests and confidence intervals:

"... a hypothesis test tells us whether the observed data are consistent with the null hypothesis, and a confidence interval tells us which hypotheses are consistent with the data."

–William C. Blackwelder

$\approx$

The relation between hypothesis tests and confidence intervals

The relation between hypothesis tests and confidence intervals enables us to do more.

The relation between hypothesis tests and confidence intervals

The relation between hypothesis tests and confidence intervals enables us to do more.

First, it gives us an effective small-sample alternative to the large-sample test comparing two population proportions.

$\approx$

The relation between hypothesis tests and confidence intervals

Specifically, to test $H_0 : p_1 - p_2 = \delta_0$ versus $H_a : p_1 - p_2 \neq \delta_0$ at the α level of significance, we can construct the level $L = 1 - \alpha$ approximate score (Agresti-Coull) confidence interval

$$\tilde{p}_1 - \tilde{p}_2 \pm z_{(1+L)/2}\sqrt{\frac{\tilde{p}_1(1-\tilde{p}_1)}{\tilde{n}_1} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{\tilde{n}_2}},$$

where

$$\tilde{n}_i = n_i + 0.5z^2_{(1+L)/2}, \ \tilde{p}_i = \frac{y_i + 0.25z^2_{(1+L)/2}}{\tilde{n}_i}, \ i = 1, 2,$$

and reject H_0 if and only if δ_0 is not in the interval.

$\approx$

The relation between hypothesis tests and confidence intervals

Second, we can create one-sided confidence intervals by inverting tests of one-sided alternative hypotheses.

The relation between hypothesis tests and confidence intervals

Second, we can create one-sided confidence intervals by inverting tests of one-sided alternative hypotheses.

For example, consider the test of $H_0 : \mu = \mu_0$ versus $H_a : \mu > \mu_0$.

The relation between hypothesis tests and confidence intervals

Second, we can create one-sided confidence intervals by inverting tests of one-sided alternative hypotheses.

For example, consider the test of $H_0 : \mu = \mu_0$ versus $H_a : \mu > \mu_0$.

Based on what we have done, a level α test rejects H_0 in favor of H_a if

$$t_{n-1,1-\alpha} \leq \frac{\overline{y} - \mu_0}{s/\sqrt{n}} < \infty.$$

$\approx$

The relation between hypothesis tests and confidence intervals

After a little algebra, these inequalities are equivalent to

$$-\infty < \mu_0 \leq \bar{y} - \frac{s}{\sqrt{n}} t_{n-1,1-\alpha},$$

giving a level $L = 1 - \alpha$ confidence interval

$$\left(-\infty,\ \bar{y} - \frac{s}{\sqrt{n}} t_{n-1,1-\alpha}\right]$$

$\approx$

Recap: Hypothesis Testing

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - The Scientific Hypothesis

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - *o* The Scientific Hypothesis
  - *o* The Statistical Model

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - *o* The Scientific Hypothesis
  - *o* The Statistical Model
  - *o* The Statistical Hypotheses

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - *o* The Scientific Hypothesis
  - *o* The Statistical Model
  - *o* The Statistical Hypotheses
  - *o* The Test Statistic

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - The Scientific Hypothesis
  - The Statistical Model
  - The Statistical Hypotheses
  - The Test Statistic
  - The p-value

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - *o* The Scientific Hypothesis
  - *o* The Statistical Model
  - *o* The Statistical Hypotheses
  - *o* The Test Statistic
  - *o* The p-value
- Types of Hypotheses

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - *o* The Scientific Hypothesis
  - *o* The Statistical Model
  - *o* The Statistical Hypotheses
  - *o* The Test Statistic
  - *o* The p-value
- Types of Hypotheses
- One and Two-Sided Tests

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - *o* The Scientific Hypothesis
  - *o* The Statistical Model
  - *o* The Statistical Hypotheses
  - *o* The Test Statistic
  - *o* The p-value
- Types of Hypotheses
- One and Two-Sided Tests
- The Philosophy of Hypothesis Testing

Recap: Hypothesis Testing

- The Components of a Statistical Hypothesis Testing Problem
  - *o* The Scientific Hypothesis
  - *o* The Statistical Model
  - *o* The Statistical Hypotheses
  - *o* The Test Statistic
  - *o* The p-value
- Types of Hypotheses
- One and Two-Sided Tests
- The Philosophy of Hypothesis Testing
- Statistical and Practical Significance

$\approx$

Recap: Hypothesis Testing

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - *o* 1-Sample Mean, Known Variance

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - *o* 1-Sample Mean, Known Variance
  - *o* 1 Sample Mean, Unknown Variance

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - *o* 1-Sample Mean, Known Variance
  - *o* 1 Sample Mean, Unknown Variance
  - *o* 1-Sample Proportion, Exact

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - 1-Sample Mean, Known Variance
  - 1 Sample Mean, Unknown Variance
  - 1-Sample Proportion, Exact
  - 1-Sample Proportion, Large Sample with Continuity Correction

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - *o* 1-Sample Mean, Known Variance
  - *o* 1 Sample Mean, Unknown Variance
  - *o* 1-Sample Proportion, Exact
  - *o* 1-Sample Proportion, Large Sample with Continuity Correction
  - *o* 2-Sample Mean, Paired Observations

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - *o* 1-Sample Mean, Known Variance
  - *o* 1 Sample Mean, Unknown Variance
  - *o* 1-Sample Proportion, Exact
  - *o* 1-Sample Proportion, Large Sample with Continuity Correction
  - *o* 2-Sample Mean, Paired Observations
  - *o* 2 Sample Mean, Pooled Variance

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - *o* 1-Sample Mean, Known Variance
  - *o* 1 Sample Mean, Unknown Variance
  - *o* 1-Sample Proportion, Exact
  - *o* 1-Sample Proportion, Large Sample with Continuity Correction
  - *o* 2-Sample Mean, Paired Observations
  - *o* 2 Sample Mean, Pooled Variance
  - *o* 2 Sample Mean, Separate Variance

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - *o* 1-Sample Mean, Known Variance
  - *o* 1 Sample Mean, Unknown Variance
  - *o* 1-Sample Proportion, Exact
  - *o* 1-Sample Proportion, Large Sample with Continuity Correction
  - *o* 2-Sample Mean, Paired Observations
  - *o* 2 Sample Mean, Pooled Variance
  - *o* 2 Sample Mean, Separate Variance
  - *o* 2-Sample Proportion, Large Sample

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - 1-Sample Mean, Known Variance
  - 1 Sample Mean, Unknown Variance
  - 1-Sample Proportion, Exact
  - 1-Sample Proportion, Large Sample with Continuity Correction
  - 2-Sample Mean, Paired Observations
  - 2 Sample Mean, Pooled Variance
  - 2 Sample Mean, Separate Variance
  - 2-Sample Proportion, Large Sample
- Fixed Significance Level Tests

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - 1-Sample Mean, Known Variance
  - 1 Sample Mean, Unknown Variance
  - 1-Sample Proportion, Exact
  - 1-Sample Proportion, Large Sample with Continuity Correction
  - 2-Sample Mean, Paired Observations
  - 2 Sample Mean, Pooled Variance
  - 2 Sample Mean, Separate Variance
  - 2-Sample Proportion, Large Sample
- Fixed Significance Level Tests
- Power

Recap: Hypothesis Testing

- Specific Hypothesis Testing Problems:
  - *o* 1-Sample Mean, Known Variance
  - *o* 1 Sample Mean, Unknown Variance
  - *o* 1-Sample Proportion, Exact
  - *o* 1-Sample Proportion, Large Sample with Continuity Correction
  - *o* 2-Sample Mean, Paired Observations
  - *o* 2 Sample Mean, Pooled Variance
  - *o* 2 Sample Mean, Separate Variance
  - *o* 2-Sample Proportion, Large Sample
- Fixed Significance Level Tests
- Power
- Hypothesis Tests and Confidence Intervals

$\approx$