STAT 513 fa 2020 Lec 05 slides

Some large-sample tests and sample size calculations

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Central limit theorem and corollary using Slutsky's theorem

Let $X_1, \ldots, X_n$ be a rs from a dist. with mean μ and variance $\sigma^2 < \infty$. Then

1. $$\sqrt{n}(\bar{X}_n - \mu)/\sigma \rightarrow \text{Normal}(0,1) \text{ in distribution}$$
   as $n \rightarrow \infty$.

2. Moreover, if S_n is a consistent estimator of σ, then
   $$\sqrt{n}(\bar{X}_n - \mu)/S_n \rightarrow \text{Normal}(0,1) \text{ in distribution}$$
   as $n \rightarrow \infty$.

Large-sample tests about the mean:

Let $X_1, \dots, X_n$ be iid with mean μ and variance $\sigma^2 < \infty$.

For some μ_0, define the test statistic

$$T_n = \sqrt{n}(\bar{X}_n - \mu_0)/S_n.$$

Then we have the following:

H_1	Rej. H_0 at α iff	Approx. power function $\gamma(\mu)$	Approx. p-value
$\mu > \mu_0$	$T_n > z_\alpha$	$1 - \Phi(z_\alpha - \sqrt{n}(\mu - \mu_0)/\sigma)$	$1 - \Phi(T_n)$
$\mu < \mu_0$	$T_n < -z_\alpha$	$\Phi(-z_\alpha - \sqrt{n}(\mu - \mu_0)/\sigma)$	$\Phi(T_n)$
$\mu \neq \mu_0$	$\lvert T_n \rvert > z_{\alpha/2}$	$1 - [\Phi(z_{\alpha/2} - \sqrt{n}(\mu - \mu_0)/\sigma)$ $-\Phi(-z_{\alpha/2} - \sqrt{n}(\mu - \mu_0)/\sigma)]$	$2[1 - \Phi(\lvert T_n \rvert)]$

Large-sample tests about the proportion: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Bernoulli}(p)$.

For some $p_0 \in (0, 1)$, define the test statistic

$$Z_n = \frac{\sqrt{n}(\hat{p}_n - p_0)}{\sqrt{p_0(1 - p_0)}}, \quad \text{where} \quad \hat{p}_n = \bar{X}_n.$$

Then we have the following:

H_1	Rej. H_0 at α iff	Approx. power function $\gamma(p)$	Approx. p-value
$p > p_0$	$Z_n > z_\alpha$	$1 - \Phi((\sigma_0/\sigma)(z_\alpha - \sqrt{n}(p - p_0)/\sigma_0))$	$1 - \Phi(Z_n)$
$p < p_0$	$Z_n < -z_\alpha$	$\Phi((\sigma_0/\sigma)(-z_\alpha - \sqrt{n}(p - p_0)/\sigma_0))$	$\Phi(Z_n)$
$p \neq p_0$	$\lvert Z_n \rvert > z_{\alpha/2}$	$1 - [\Phi((\sigma_0/\sigma)(z_{\alpha/2} - \sqrt{n}(p - p_0)/\sigma)) - \Phi((\sigma_0/\sigma)(-z_{\alpha/2} - \sqrt{n}(p - p_0)/\sigma_0))]$	$2[1 - \Phi(\lvert Z_n \rvert)]$

where $\sigma_0^2 = p_0(1 - p_0)$ and $\sigma^2 = p(1 - p)$.

Exercise: Derive the approximate power function for the large-sample test of

$$H_0\colon p \leq p_0 \text{ vs } H_1\colon p > p_0$$

on the previous slide.

Large-sample tests for comparing two means:

Let $X_{k1}, \dots, X_{kn_k}$ be iid w/ mean μ_k and variance $\sigma_k^2 < \infty$, $k = 1, 2$. For some δ_0, define

$$T = \frac{\bar{X}_1 - \bar{X}_2 - \delta_0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}.$$

Then we have the following (supposing, say $\min\{n_1, n_2\} \geq 30$):

H_1	Rej. H_0 at α iff	Approx. power $\gamma(\delta)$, $\delta = \mu_1 - \mu_2$	Approx. p-val
$\mu_1 - \mu_2 > \delta_0$	$T > z_\alpha$	$1 - \Phi\left(z_\alpha - \frac{\delta - \delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}\right)$	$1 - \Phi(T)$
$\mu_1 - \mu_2 < \delta_0$	$T < -z_\alpha$	$\Phi\left(-z_\alpha - \frac{\delta - \delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}\right)$	$\Phi(T)$
$\mu_1 - \mu_2 \neq \delta_0$	$\lvert T\rvert > z_{\alpha/2}$	$1 - \left[\Phi\left(z_{\alpha/2} - \frac{\delta - \delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta - \delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}\right)\right]$	$2[1 - \Phi(\lvert T\rvert)]$

How to choose n?

Two questions for determining sample size

1. How small a deviation from H_0 is it of interest to detect?
2. With what probability do we wish to detect it?

Exercise: You wish to test

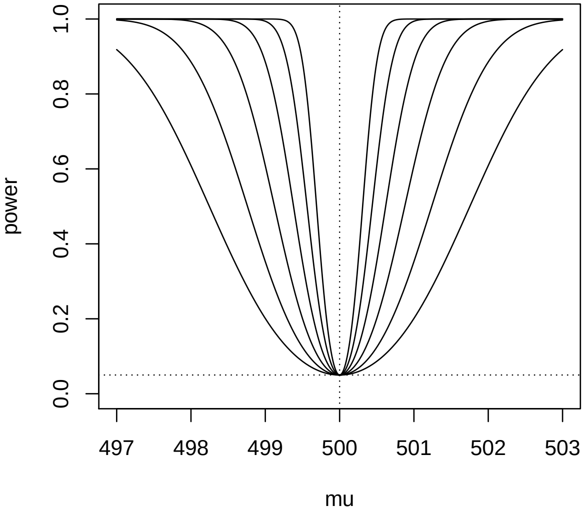
$$H_0\colon \mu = 500\,\text{mL versus } H_1\colon \mu \neq 500\,\text{mL},$$

where μ is the mean amount of a drink in bottles labeled as containing $500\,\text{mL}$.

Assume $\sigma = 2\,\text{mL}$. Test based on vols $X_1, \dots, X_n$ of n randomly selected bottles:

$$\text{Reject } H_0 \text{ iff } |\sqrt{n}(\bar{X}_n - 500)/2| > z_{0.025}.$$

1. Plot power curves of the test for the sample sizes $n = 5, 10, 20, 40, 80, 160$ across $\mu \in (497, 503)$, assuming that the volumes are Normally distributed.
2. You wish to detect a deviation as small as $1\,\text{mL}$ of the mean from $500\,\text{mL}$ with probability at least 0.80. Which of the sample sizes ensure this?
3. What is the smallest sample size among $n = 5, 10, 20, 40, 80, 160$ under which the power is at least 0.80 for values of μ at least $1\,\text{mL}$ away from $500\,\text{mL}$?


power
1.0
0.8
0.6
0.4
0.2
0.0
497
498
499
500
501
502
503
mu

We must specify two values when making sample size determinations:

1. δ^* is the smallest deviation from H_0 we wish to detect.
2. γ^* is the smallest probability with which we wish to detect it.

The quantity δ^* expresses: how untrue is the null?
The quantity γ^* is the desired power when the null is untrue by the amount δ^*

Exercise (cont): You now wish to test the one-sided set of hypotheses

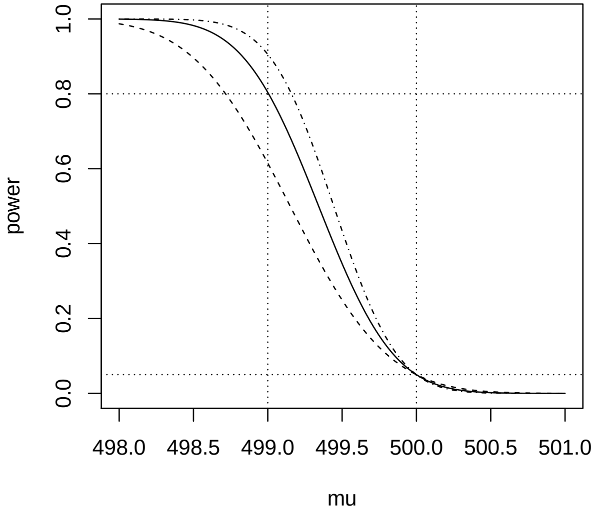
$$H_0\colon \mu \geq 500\,\text{mL versus } H_1\colon \mu < 500\,\text{mL},$$

where μ is the mean amount of a drink in bottles labeled as containing $500\,\text{mL}$.

Assume $\sigma = 2\,\text{mL}$. Test based on vols $X_1, \dots, X_n$ of n randomly selected bottles:

$$\text{Reject } H_0 \text{ iff } \sqrt{n}(\bar{X}_n - 500)/2 < -z_{0.05}.$$

1. Sketch the shape of the power curve.
2. Find the smallest sample size under which the test will have a power of at least 0.80 to detect a deviation as small as $1\,\text{mL}$ from the null.
3. Plot the power curve of the test under this sample size. Also include power curves for the test under a sample with 10 fewer and 10 more observations.


power
1.0
0.8
0.6
0.4
0.2
0.0
498.0
498.5
499.0
499.5
500.0
500.5
501.0
mu

Sample sizes for tests about a mean:

Let $X_1, \dots, X_n$ be iid with mean μ and variance $\sigma^2 < \infty$, and for some μ_0, let

$$T_n = \sqrt{n}(\bar{X}_n - \mu_0)/S_n.$$

Sample size needed to detect dev. from null as small as δ^* with power at least γ^*.

H_0	H_1	Rej. at α iff	choose n^* as smallest integer greater than
$\mu \leq \mu_0$	$\mu > \mu_0$	$T_n > z_\alpha$	$\sigma^2(z_\alpha + z_{\beta^*})^2/(\delta^*)^2$
$\mu \geq \mu_0$	$\mu < \mu_0$	$T_n < -z_\alpha$	$\sigma^2(z_\alpha + z_{\beta^*})^2/(\delta^*)^2$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert T_n \rvert > z_{\alpha/2}$	$\sigma^2(z_{\alpha/2} + z_{\beta^*})^2/(\delta^*)^2$

In the table $\beta^* = 1 - \gamma^*$.

Sample sizes for tests about a proportion:

Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Bernoulli}(p)$, and for some p_0 define

$$Z_n = \sqrt{n}(\hat{p} - p_0)/\sqrt{p_0(1-p_0)}, \quad \text{where } \hat{p}_n = \bar{X}_n.$$

Sample size needed to detect dev. from null as small as δ^* with power at least γ^*.

H_1	Rej. at α iff	choose n^* as smallest integer greater than
$p > p_0$	$Z_n > z_\alpha$	$\left[\sqrt{(p_0+\delta^*)(1-(p_0+\delta^*))}z_{\beta^*} + \sqrt{p_0(1-p_0)}z_\alpha\right]^2/(\delta^*)^2$
$p < p_0$	$Z_n < -z_\alpha$	$\left[\sqrt{(p_0-\delta^*)(1-(p_0-\delta^*))}z_{\beta^*} + \sqrt{p_0(1-p_0)}z_\alpha\right]^2/(\delta^*)^2$
$p \neq p_0$	$\lvert Z_n\rvert > z_{\alpha/2}$	$\max\left\{\left[\sqrt{(p_0+\delta^*)(1-(p_0+\delta^*))}z_{\beta^*} + \sqrt{p_0(1-p_0)}z_{\alpha/2}\right]^2, \left[\sqrt{(p_0-\delta^*)(1-(p_0-\delta^*))}z_{\beta^*} + \sqrt{p_0(1-p_0)}z_{\alpha/2}\right]^2\right\}/(\delta^*)^2$

In the table $\beta^* = 1 - \gamma^*$.

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Bernoulli}(p)$ and suppose

- Fausto will test H_0: $p \leq 1/4$ vs H_1: $p > 1/4$ with

$$\text{Reject } H_0 \text{ iff } \frac{\hat{p}_n - 1/4}{\sqrt{(1/4)(1-1/4)/n}} > z_\alpha$$

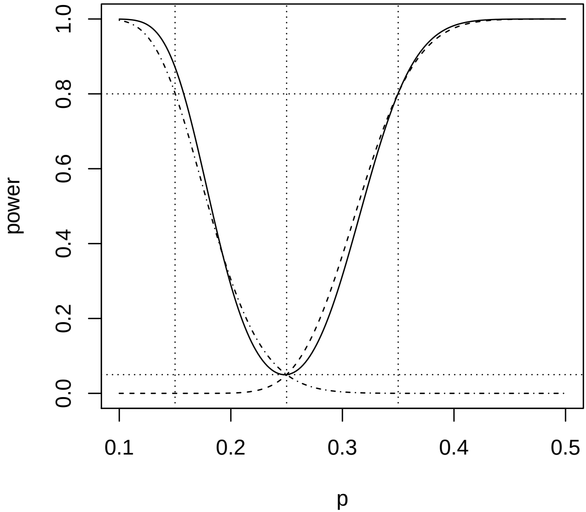
- Inés will test H_0: $p \geq 1/4$ vs H_1: $p < 1/4$ with

$$\text{Reject } H_0 \text{ iff } \frac{\hat{p}_n - 1/4}{\sqrt{(1/4)(1-1/4)/n}} < -z_\alpha$$

- Germán will test H_0: $p = 1/4$ vs H_1: $p \neq 1/4$ with

$$\text{Reject } H_0 \text{ iff } \left| \frac{\hat{p}_n - 1/4}{\sqrt{(1/4)(1-1/4)/n}} \right| > z_{\alpha/2}$$

1. If each researcher wishes to detect a deviation from the null as small as 0.10 with probability at least 0.80, what sample size should each use?
2. Using these sample sizes, plot the power curves for the three researchers' tests.


power
0.0
0.2
0.4
0.6
0.8
1.0
0.1
0.2
0.3
0.4
0.5
p

Sample sizes for tests comparing two means:

Let $X_{k1}, \dots, X_{kn_k}$ be iid w/ mean μ_k and variance $\sigma_k^2 < \infty$, $k = 1, 2$. Define

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (0)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}.$$

We tabulate below the smallest values of n such that for all n_1 and n_2 satisfying

$$n_1 \geq n \cdot \sigma_2(\sigma_1 + \sigma_2)^{-1} \quad \text{and} \quad n_2 \geq n \cdot \sigma_2(\sigma_1 + \sigma_2)^{-1}$$

the power to detect a deviation from the null as small as $\delta^* > 0$ will be at least γ^*.

H_1	Rej. at α iff	choose n greater than or equal to
$\mu_1 - \mu_2 > 0$	$T > z_\alpha$	$[(z_{\beta^*} + z_\alpha)(\sigma_1 + \sigma_2)]^2/(\delta^*)^2$
$\mu_1 - \mu_2 < 0$	$T < -z_\alpha$	$[(z_{\beta^*} + z_\alpha)(\sigma_1 + \sigma_2)]^2/(\delta^*)^2$
$\mu_1 - \mu_2 \neq 0$	$\lvert T \rvert > z_{\alpha/2}$	$[(z_{\beta^*} + z_{\alpha/2})(\sigma_1 + \sigma_2)]^2/(\delta^*)^2$

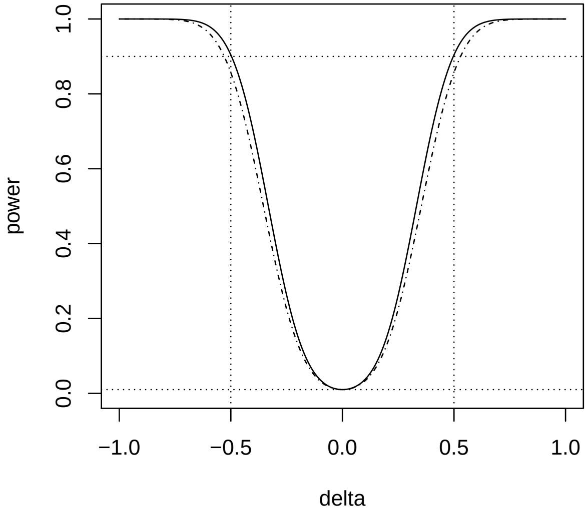
In the table $\beta^* = 1 - \gamma^*$.

Exercise: A researcher wishes to test H_0: $\mu_1 - \mu_2 = 0$ vs H_1: $\mu_1 - \mu_2 \neq 0$. A pilot study resulted in $\hat{\sigma}_1^2 = 1.22$ and $\hat{\sigma}_2^2 = 0.26$.

If a difference as large as 0.50 units exists, the researcher would like to detect it with probability no smaller than 0.90. The researcher will use test

$$\text{Reject } H_0 \text{ iff } \left| \frac{\bar{X}_1 - \bar{X}_2 - (0)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \right| > 2.575829.$$

1. Recommend sample sizes n_1 and n_2 to the researcher.
2. Plot the power curve of the test under these sample sizes, assuming $\sigma_1^2 = 1.22$ and $\sigma_2^2 = 0.26$.
3. Add to the plot the power curve resulting from using the same total sample size, but when samples of equal size are drawn from the two populations. Explain how the power curves are different and why.



power
delta
0.0
0.2
0.4
0.6
0.8
1.0
−1.0
−0.5
0.0
0.5
1.0