STA258H5
University of Toronto Mississauga

Al Nosedal, Asal Aslemand and Omid Jazi

Winter 2023

Statistical Power

Statistical Power

The statistical power of a test is its ability to detect an effect if it exists in reality.

It is the probability of correctly rejecting H_0 when H_0 is false in reality.

$$\text{Power} = P(\text{reject } H_0 \mid H_0 \text{ false})$$

$[0 < \text{power} < 1]$

Power close to 1 (high power)
Test is good at detecting effects

Power close to 0 (low power)
Tes is not reliable (i.e. we expect the test will not reject H_0 when H_0 is false)

Power is affected by

- The effect
  (larger differences between reality and the null are easier to detect)
- Sample size
  (larger samples increase power)
- Significance level
  (as α increases, easier to reject H_0)
- Variability in data
  (lower variability, higher power)

Type I and II Errors

It is possible to make an incorrect conclusion on a hypothesis test.

Type I: Incorrectly reject H_0 when H_0 is true in reality

Type II: Incorrectly fail to reject H_0 when H_0 is false in reality

Conclusion of hyp. test		Reality: H_0 True	Reality: H_0 False
	reject H_0	Type I (α)	No error ✓
	Fail to reject H_0	No error ✓	Type II (β)

Note:
Type I errors are generally considered worse

β = Probability of a Type II error

power $= 1 - \beta = 1 - P(\text{Type II})$

Power

The probability that a fixed level α significance test will reject H_0 when a particular alternative value of the parameter is true is called the **power** of the test against that alternative.

Sweetening colas: power

Specific $H_a: \mu = 1.1$

The cola maker of our example determines that a sweetness loss is too large to accept if the mean response for all tasters is $\mu = 1.1$. Will a 5% significance test of the hypotheses

$\alpha = 0.05$

$H_0 : \mu = 0$

$H_a : \mu > 0$

$n = 10$

based on a sample of 10 tasters usually detect a change this great?

We want the power of the test against the alternative $\mu = 1.1$. This is the probability that the test rejects H_0 when $\mu = 1.1$ is true. Assume $\sigma = 1$.

power to detect Specific $H_a: \mu = 1.1$

$\sigma = 1$ (known)

Example (Slide 4)

$H_0: \mu = 0$ (μ_0)

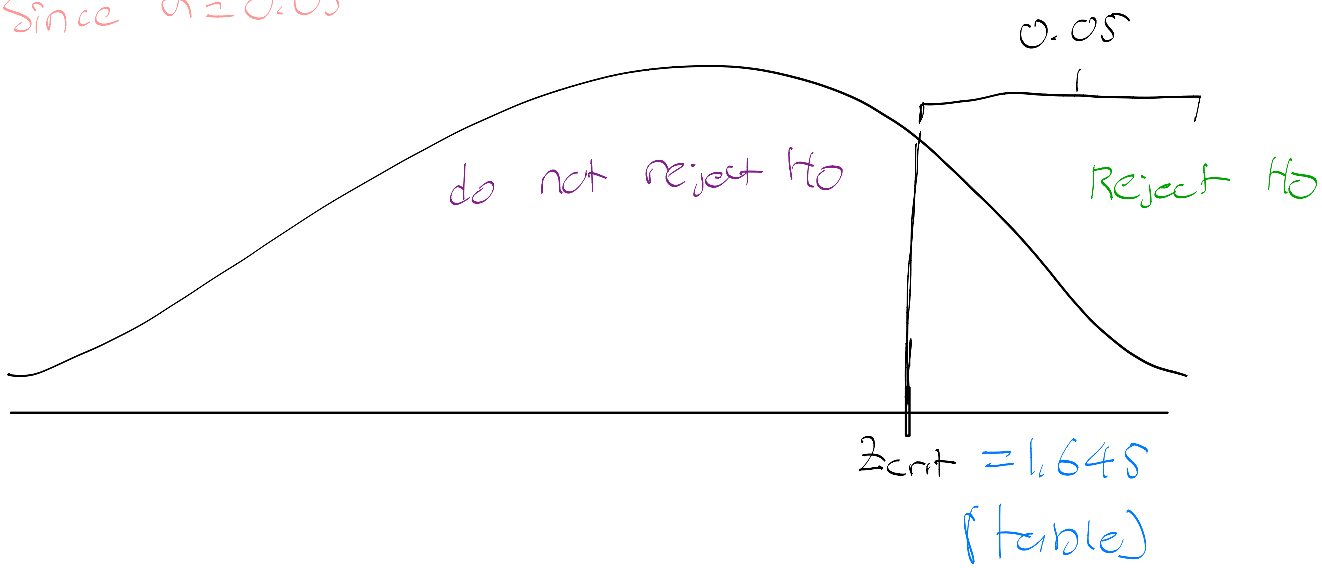
$H_a: \mu > 0$ (μ_0) (one-sided, $\mu > 0$ right tailed)

$n = 10$, $\sigma = 1$, $\alpha = 0.05$

Calculate the power of reject H_0 and detecting specific $H_a: \mu = 1.1$

1/ Find $\underbrace{\text{Rejection Region}}_{\text{where } H_0 \text{ gets rejected}}$ and critical z value (z_{crit})

Since $\alpha = 0.05$



we reject H_0 if $z^* > 1.645$

2/ Find equivalent $\bar{X}_{crit}$ which is a threshold for rejecting $H_0: \mu = 0$

Since σ known

$$Z_{crit} = \frac{\bar{X}_{crit} - \mu_0}{\sigma/\sqrt{n}}$$

$$1.645 = \frac{\bar{X}_{crit} - 0}{1/\sqrt{10}}$$

$$\bar{X}_{crit} \approx 0.520$$

$$\begin{pmatrix}\text{reject } H_0: \mu = 0 \\ \text{when } z^{*} > 1.645\end{pmatrix} \Leftrightarrow \begin{pmatrix}\text{reject } H_0: \mu = 0 \\ \text{when } \bar{x}_{crit} > 0.520\end{pmatrix}$$

3/ Calculate the power

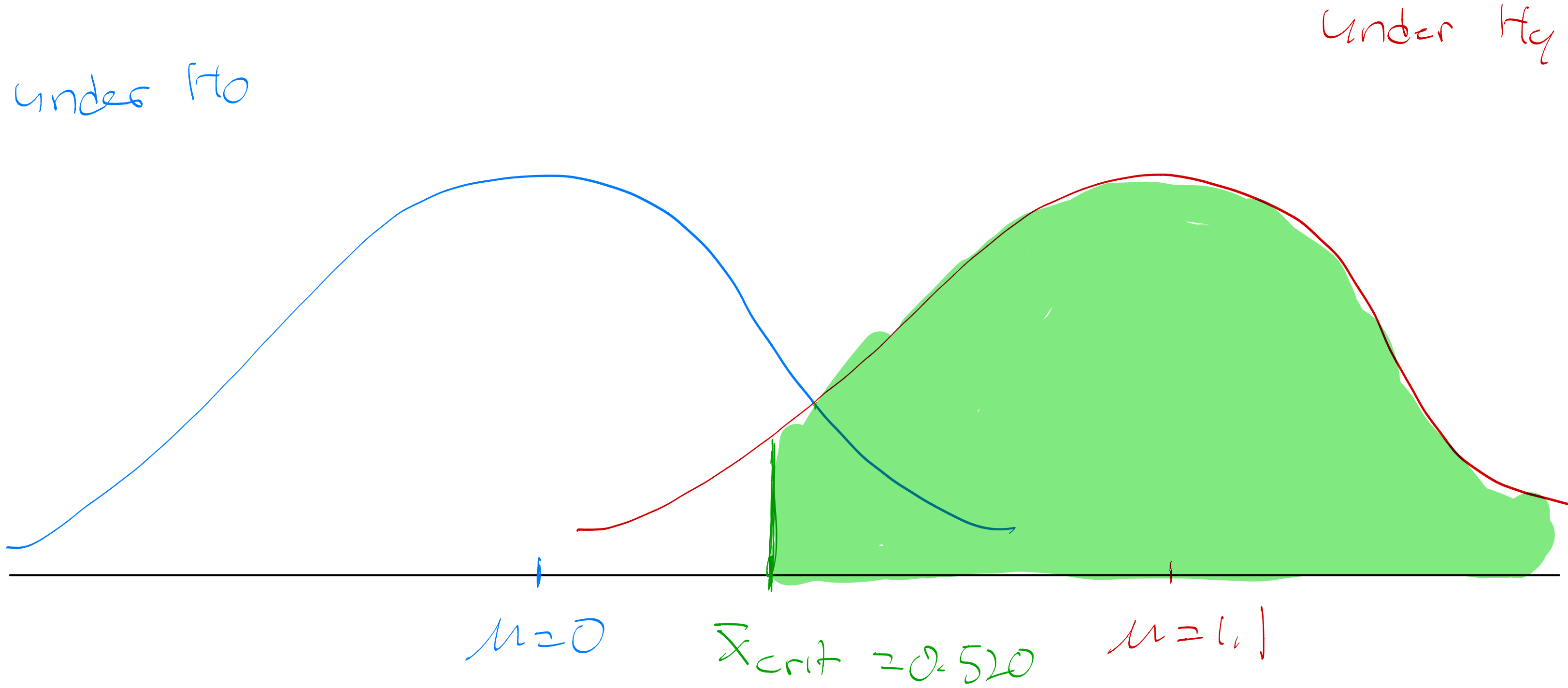
(reject $H_0: \mu = 0$ when $\mu = 1.1$ in reality)

Specific H_a

↳ i.e How sensitive is the test at detecting the difference between $\underbrace{\mu \text{ under } H_0}_{\mu = 0}$ and $\underbrace{\mu \text{ in reality}}_{H_a: \mu = 1.1}$

Use $\bar{x}_{crit}$ (since it is boundary condition)

$= 0.520$



$$\text{Power} = P\left(Z > \frac{\bar{x}_{crit} - \mu}{\sigma/\sqrt{n}} \;\middle|\; \mu = 1.1 \right)$$

$$= P\left(Z > \frac{0.520 - 1.1}{1/\sqrt{10}} \right)$$

$$= P(Z > -1.83)$$

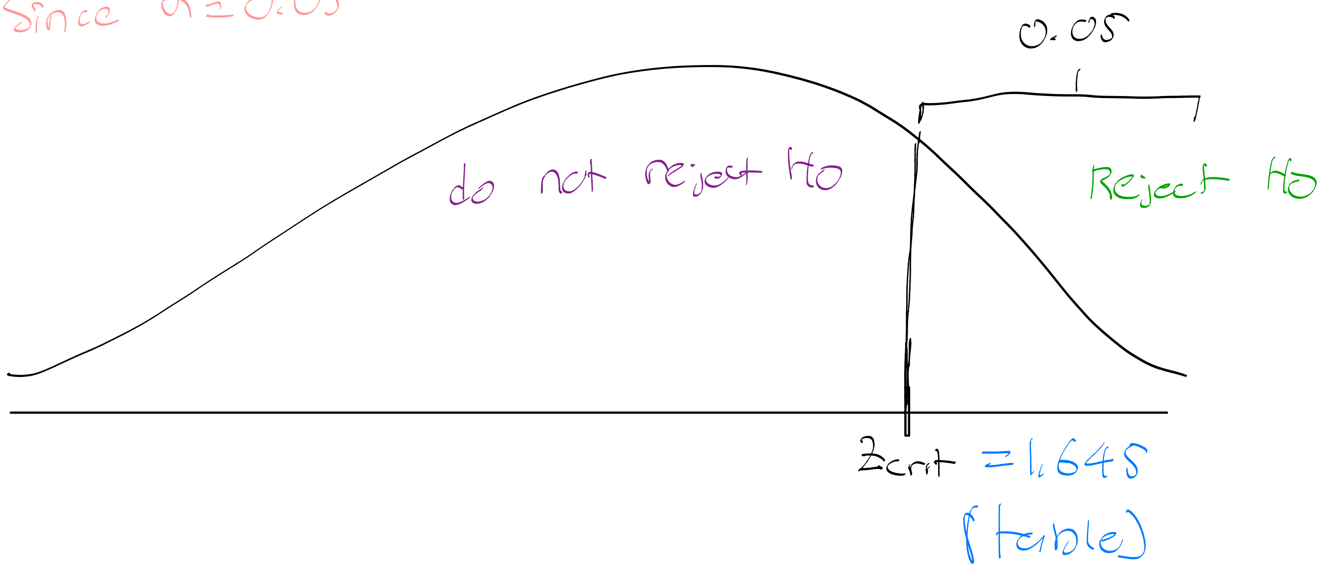
$$= 1 - 0.0336$$

$$= 0.9664$$

Examine approach involving Type II error

For the same example (slide 4), calculate the probability of a type II error for the same situation. (i.e in reality $H_a: \mu = 1.1$)

Since $\alpha = 0.05$



we reject Ho if $z^* > 1.645$

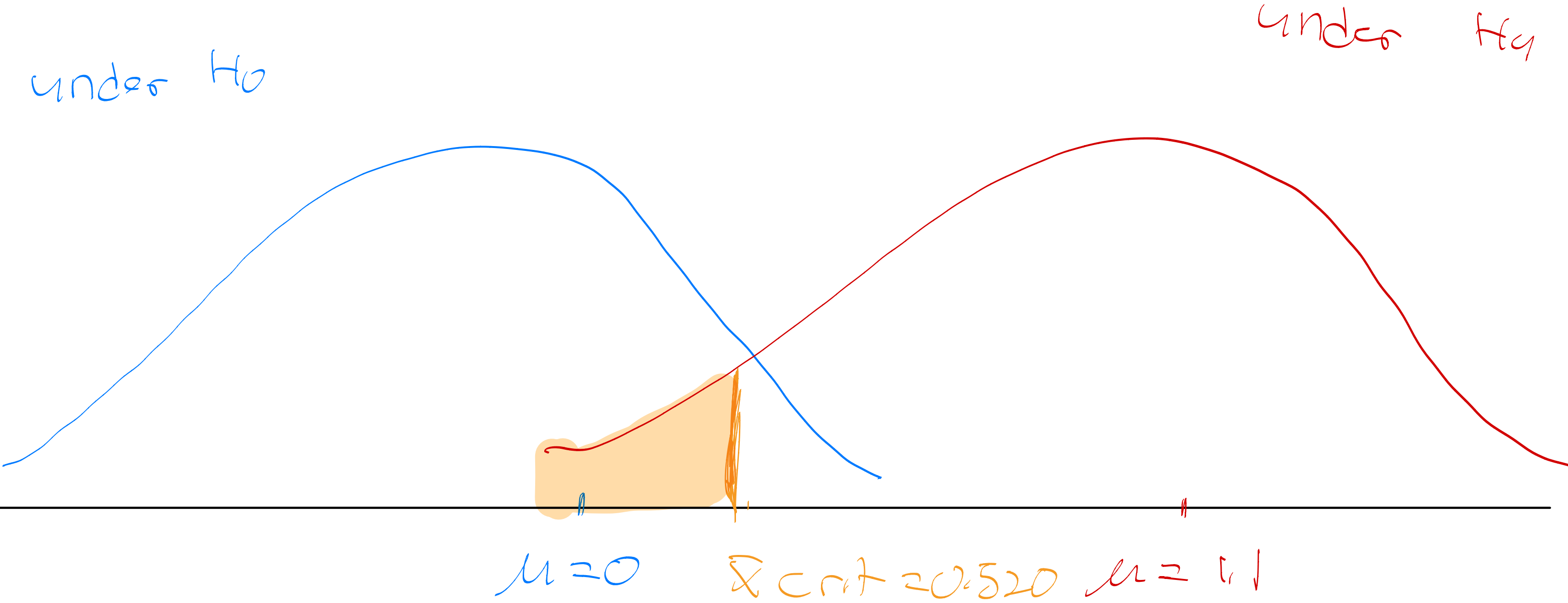
Since σ known

$$z_{crit} = \frac{\bar{x}_{crit} - \mu_0}{\sigma/\sqrt{n}}$$

$$1.645 = \frac{\bar{x}_{crit} - 0}{1/\sqrt{10}}$$

$$\bar{x}_{crit} = 0.520$$

$\left(\begin{array}{l}\text{not reject } H_0: \mu = 0 \\ \text{when } z^* < 1.645\end{array}\right) \iff \left(\begin{array}{l}\text{not reject } H_0: \mu = 0 \\ \text{when } \bar{x}_{crit} < 0.520\end{array}\right)$



$\beta = P(\text{type II error})$

$= P(\underbrace{\text{not reject } H_0 : \mu = 0}_{\bar{X}_{crit} < 0.520} \mid \mu = 1.1)$

$= P(\bar{X} < 0.520 \mid \mu = 1.1)$

$$= P\left(Z < \frac{0.520 - 1.1}{1/\sqrt{10}} \right)$$

$= P(Z < -1.83)$

$= 0.0336 = \beta = P(\text{type II})$

Let's use β to find the power

$$\text{power} = 1 - \beta$$
$$= 1 - 0.0336$$
$$= 0.9664$$ (consistent to earlier value)

$$\text{power} = 1 - P(\text{Type II})$$

$$P(\text{Type I}) = 1 - \text{power}$$

Step 1.

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}$.
We know that $\sigma = 1$, so the z test rejects H_0 at the $\alpha = 0.05$ level when

$$z = \frac{\bar{x} - 0}{1/\sqrt{10}} \geq 1.645$$

This is the same as

$$\bar{x} \geq 0 + 1.645 \frac{1}{\sqrt{10}}$$

or

$$\text{Reject } H_0 \text{ when } \bar{x} \geq 0.520$$

Step 2.

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 1.1$ is true.

To calculate this probability, standardize $\bar{x}$ using $\mu = 1.1$.

$$\begin{aligned} \text{power } &= P(\bar{x} \geq 0.520 \text{ when } \mu = 1.1) \\ &= P\left(\frac{\bar{x}-1.1}{1/\sqrt{10}} \geq \frac{0.520-1.1}{1/\sqrt{10}}\right) \\ &= P(Z \geq -1.83) = 1 - 0.0336 = 0.9664 \end{aligned}$$

Type I and Type II errors

Type I and Type II Errors

If we reject H_0 when in fact H_0 is true, this is a **Type I error.**
If we fail to reject H_0 when in fact H_a is true, this is a **Type II error.**
The **significance level** α of any fixed level test is the probability of a Type I error.
The **power** of a test against any alternative is 1 minus the probability of a Type II error for that alternative.
The probability of making a Type II error is denoted by β.

Decision Errors in Tests

Type I Error
H_0 is true, but sampling variation in the data leads you to reject H_0, you've made a Type I error.
When H_0 is true, a Type I error occurs if H_0 is rejected.
Type II Error
H_0 is false, but sampling variation in the data does not lead you to reject H_0, you've made a Type II error.
When H_0 is false, a Type II error occurs if H_0 is NOT rejected.

Decision Errors in Tests

		TRUTH	
		Null Hypothesis True	Null Hypothesis False
Decision	Reject Null Hypothesis	Type I Error	Correct
	Fail to Reject Null Hypothesis	Correct	Type II Error

Example of Decisions Errors in Tests

In medical disease testing, the null hypothesis is usually the assumption that a person is healthy. The alternative is that the person has the disease we are testing for. H_0 : Healthy versus H_a: Infected

- Type I error: Reject Ho when it is true.

A Type I error is a false positive: A healthy person is diagnosed with the disease. That is, a person must go under further test.

- Type II error: Fail to reject Ho (?Accept Ho?) when it is false.

A Type II error is a false negative in which an infected person is diagnosed as disease-free. That is, a sick person gets untreated.
For example: If a new treatment is being tested for a disease (e.g., epilepsy), a Type I error will lead to future patients getting a useless treatment; a Type II error means a useful treatment will remain undiscovered.

What type of error could we making in our example?

In 1980s, it was generally believed that congenital abnormalities affect 5% of the nation's children. Some people believe that the increase in the number of chemicals in the environment in recent years has led in the incidence of abnormalities. A recent study examined 384 children and found that 46 of them showed signs of abnormality. Is this strong evidence that the risk has increased?
$H_0 : p = 0.05$ versus $H_a : p > 0.05$ $Z \approx 6.28$, $p - value < 0.0001$ (which is less than $\alpha = 0.05$). We reject H_0 and conclude H_a. This means we could be making a Type I error. We decided that the true percentage is more than 5% based on our data (as evidence against H_0), however, it could be that the hypothesized value of 5% is true (e.g., $H_0 : p = 0.05$ could be true).

What type of error could we making in our example?

According to Access and Support to Education and Training Survey (2008), of 4,756 adult Canadians, 1,581 indicated that they worked at a job or business at anytime (between July 2007 and June 2008), regardless of the number of hours per week . Is there evidence to suggest that the true proportion p is greater than 0.50?

$H_0 : p = 0.50$ vs $H_a : p > 0.50$

What type of error could we making in our example?

```
prop.test(x=1581, n = 4756, p=0.50,
alternative = "greater", correct = FALSE);

##
##  1-sample proportions test without continuity
##  correction
##
## data:  1581 out of 4756, null probability 0.5
## X-squared = 534.24, df = 1, p-value = 1
## alternative hypothesis: true p is greater than 0.5
## 95 percent confidence interval:
##  0.3212845 1.0000000
## sample estimates:
##         p
## 0.3324222
```

What type of error could we making in our example? (cont.)

$P - value > \alpha = 0.05$; we Fail to Reject Ho.

This means we could be making a Type II error. We indicated that there is no evidence to conclude that the true proportion of adult Canadians who worked at a job or business at anytime (between July 2007 and June 2008), regardless of the number of hours per week was more than 0.50 - this conclusion implies that $H_0 : p = 0.50$ is plausible, but we could be wrong.

Example. Sweetening colas: error probabilities

When we select the significance level α of a test, we are setting the probability of a Type I error. Calculating the probability of a Type II error is just like calculating the power, except that we find the probability of the wrong decision (failing to reject H_0) rather than the probability of the right decision (rejecting).

$$H_0 : \mu = 0$$

$$H_a : \mu > 0$$

Example. Sweetening colas: error probabilities

The example shows that the z test rejects the null hypothesis at level $\alpha = 0.05$ when the mean sweetness loss assigned by 10 tasters satisfies $\bar{x} \geq 0.520$. The two error probabilities are

$$\begin{aligned} \text{P(Type I error)} &= \text{P(reject } H_0 \text{ when } \mu = 0) \\ &= P(\bar{x} \geq 0.520 \text{ when } \mu = 0) \\ &= P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \geq \frac{0.520-0}{1/\sqrt{10}}\right) \\ &= P(Z \geq 1.6443) = 0.05. \end{aligned}$$

Example. Sweetening colas: error probabilities

$$
\begin{aligned}
\text{P(Type II error)} &= \text{P(fail to reject } H_0 \text{ when } \mu = 1.1) \\
&= P(\bar{x} < 0.520 \text{ when } \mu = 1.1) \\
&= P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < \frac{0.520-1.1}{1/\sqrt{10}}\right) \\
&= P(Z < -1.8341) = 0.0336.
\end{aligned}
$$

Example. Sweetening colas: error probabilities

$$\begin{aligned}\text{P(Type II error)} &= \text{P(fail to reject } H_0 \text{ when } \mu = 0.52) \\ &= P(\bar{x} < 0.520 \text{ when } \mu = 0.52) \\ &= P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < \frac{0.520-0.52}{1/\sqrt{10}}\right) \\ &= P(Z < 0) = 0.5.\end{aligned}$$

R Code

```
# Step 1. Computing power;

# vec.mu = vector with several different values of mu;
vec.mu = seq(0,3,by=0.01);
# n = sample size;
n = 10;
# sigma = population std. dev.;
sigma = 1;
# vec.z = vector of z-scores;
vec.z = sqrt(n)*(0.52-vec.mu)/sigma;
#beta = type II error;
beta = pnorm(vec.z);
power = 1 - beta;
```

R Code

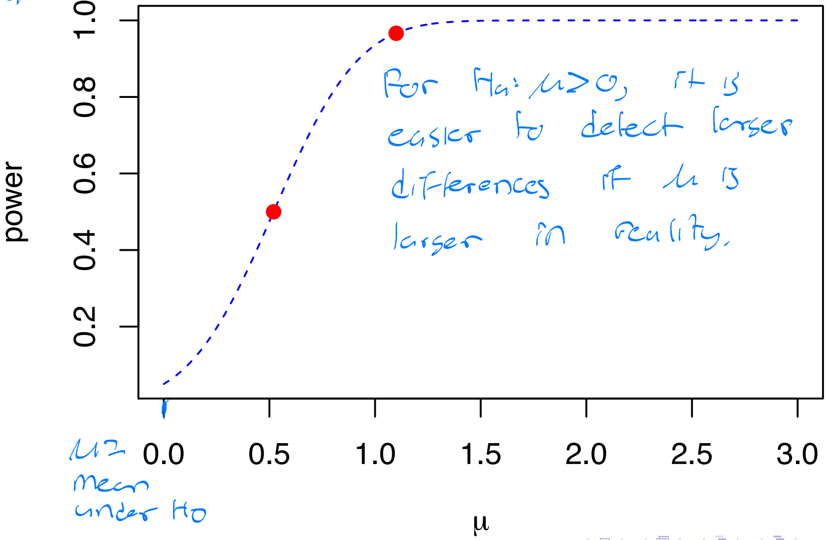
```
# Step 2. Graphing power;

plot(vec.mu, power,type="l",lty=2,
col="blue", xlab=expression(mu));
points(0.52,0.5,pch=19,col="red");
points(1.1,1-0.0336,pch=19,col="red");
```

Power curve

$H_0: \mu = 0$

$H_a: \mu > 0$

non-decreasing (as μ under H_a ↑)



For $H_a: \mu > 0$, it is easier to detect larger differences if μ is larger in reality.

$\mu =$ mean under H_0

Cola bottles: power

Bottles of a popular cola are supposed to contain 300 millimeters (ml) of cola. There is some variation from bottle to bottle because the filling machinery is not perfectly precise. The distribution of contents is Normal with standard deviation $\sigma = 3$ ml. Will inspecting 6 bottles discover underfilling? The hypotheses are

$$H_0 : \mu = 300$$

$$H_a : \mu < 300$$

Cola bottles: power

A 5% significance test rejects H_0 if $z_* \leq -1.645$, where the test statistic z_* is

$$z_* = \frac{\bar{x} - 300}{3/\sqrt{6}}$$

Power calculations help us see how large a shortfall in the bottle contents the test can be expected to detect. Find the power of this test against the alternative $\mu = 299$.

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}$.
We know that $\sigma = 1$, so the z test rejects H_0 at the $\alpha = 0.05$ level when

$$z = \frac{\bar{x} - 300}{3/\sqrt{6}} < -1.645$$

This is the same as

$$\bar{x} < 300 - 1.645 \frac{3}{\sqrt{6}}$$

or

$$\text{Reject } H_0 \text{ when } \bar{x} < 297.985$$

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 299$ is true.

To calculate this probability, standardize $\bar{x}$ using $\mu = 299$.

$$\begin{aligned} \text{power } &= P(\bar{x} < 297.985 \text{ when } \mu = 299) \\ &= P\left(Z < \frac{297.985 - 299}{3/\sqrt{6}}\right) \\ &= P(Z < -0.83) = 0.2033 \end{aligned}$$

Using Power to Determine Sample Size

Example

Suppose an experimenter wishes to test

$$H_0 : \mu = 100$$

power = 0.60

$$H_a : \mu > 100$$

at the $\alpha = 0.05$ level of significance and wants $1 - \beta$ to equal 0.60 when $\mu = 103$. What is the smallest (i.e., cheapest) sample size that will achieve that objective? Assume that the variable being measured is Normally distributed with $\sigma = 14$.

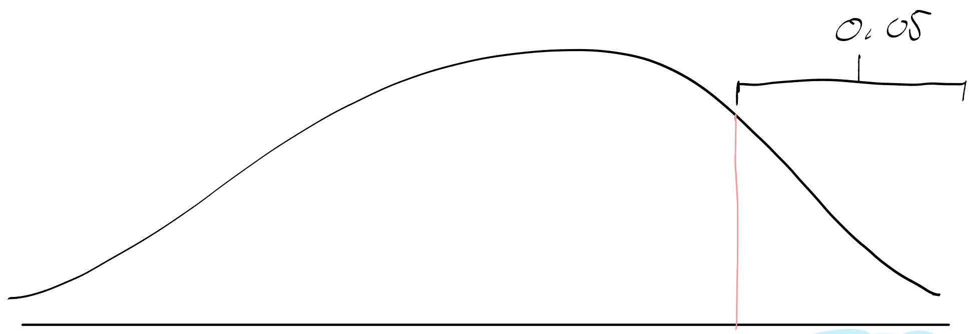
Example (slide 28)

$H_0: \mu = 100$
$H_a: \mu > 100$

Assume $\sigma = 14$. Use $\alpha = 0.05$.

Find smallest n to detect $H_a: \mu = 103$ such that power $= 1 - \beta = 0.60$.

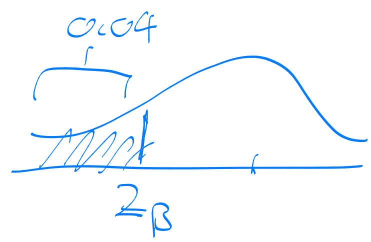
Use $\alpha = 0.05$ to find threshold z_α^* $\quad H_a: >$



$z_\alpha^* = 1.645$ (table)

Determine β and calculate z_β^* threshold

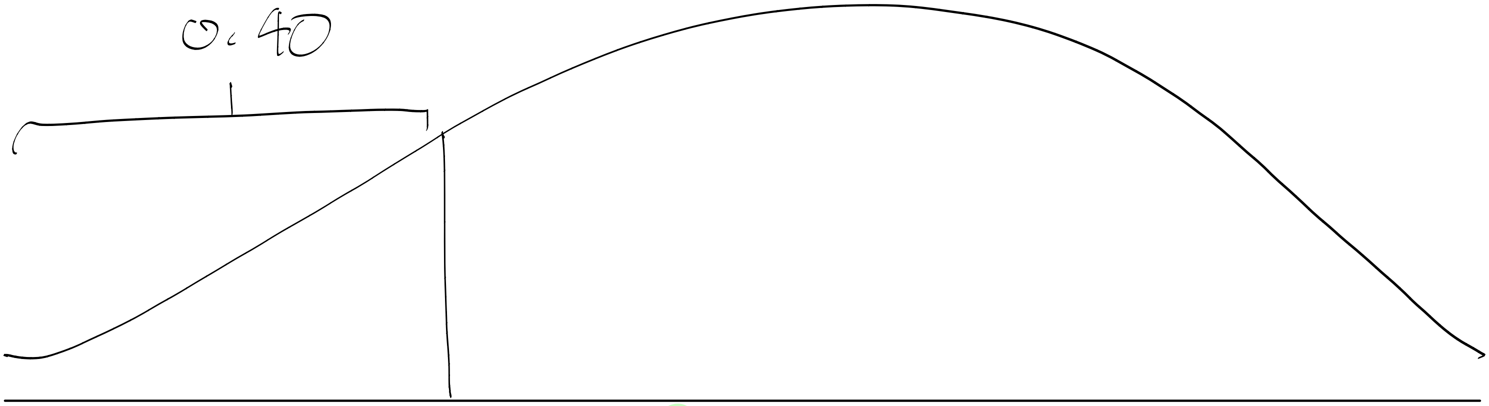
$$1 - \beta = 0.60$$
$$\beta = 0.40$$



$\beta = P(\text{Type II})$

$= P(\underbrace{\text{Not reject } H_0}_{z < z_\beta^*} \mid H_0 \text{ is false}, \mu = 103)$

(find z_β^*)



$z_\beta \approx -0.25$ (table)

how many st. devs away

z_β informs us how far the critical threshold is for rejecting H_0 from the mean of the alternative.

Use z_α, z_β and the form of test stat to find n

(test stat when σ known: $z^* = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$)

Under H_0:

$$\mu_c = \mu_0 + z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$$

under H_a

$$\mu_c = \mu_A + z_\beta \cdot \frac{\sigma}{\sqrt{n}}$$

where μ_c is the critical value of $\bar{x}$ which is a threshold to reject (or not reject) the test

Equate the 2 expressions

$$\mu_c = \mu_0$$

$$\mu_0 + Z_\alpha \cdot \frac{\sigma}{\sqrt{n}} = \mu_A + Z_\beta \cdot \frac{\sigma}{\sqrt{n}}$$

$$Z_\alpha \frac{\sigma}{\sqrt{n}} - Z_\beta \frac{\sigma}{\sqrt{n}} = \mu_A - \mu_0$$

mean under Ha

mean under Ho

$$Z_\alpha - Z_\beta = \frac{\mu_A - \mu_0}{\sigma/\sqrt{n}}$$

$$1.645 - (-0.25) = \frac{103 - 100}{14/\sqrt{n}}$$

$$\sqrt{n} = 8.843$$

$$n = 78.205$$

$$n = 79$$

$$1.645 - (-0.25) = \frac{103 - 100}{14/\sqrt{n}}$$

$$1.645 + 0.25 = \frac{103 - 100}{\left(\frac{14}{\sqrt{n}}\right)}$$

$$\frac{14}{\sqrt{n}} = \frac{103 - 100}{1.645 + 0.25}$$

$$\frac{\sqrt{n}}{14} = \frac{1.645 + 0.25}{103 - 100}$$

$$\sqrt{n} = \frac{14(1.645 + 0.25)}{103 - 100}$$

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}_*$.
By definition,

$$\begin{aligned}\alpha &= \text{P(we reject } H_0 \text{ given } H_0 \text{ is true)}\\ &= P(\bar{X} > \bar{x}_* | \mu = 100)\\ &= P\left(\frac{\bar{X}-100}{14/\sqrt{n}} > \frac{\bar{x}_*-100}{14/\sqrt{n}}\right)\\ &= P\left(Z > \frac{\bar{x}_*-100}{14/\sqrt{n}}\right) = 0.05\end{aligned}$$

But $P(Z > 1.645) = 0.05$, so

$$\bar{x}_* = 100 + 1.645\frac{14}{\sqrt{n}}$$

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 103$ is true.
To calculate this probability, standardize $\bar{x}$ using $\mu = 103$.

$$\text{power} = 1 - \beta = P(\bar{X} > \bar{x}_* \text{ when } \mu = 103)$$
$$= P\left(\frac{\bar{X}-103}{14/\sqrt{n}} > \frac{\bar{x}_*-103}{14/\sqrt{n}}\right) = 0.60$$

From our Table, $P(Z > -0.25) = 0.5987 \approx 0.60$, so

$$\frac{\bar{x}_* - 103}{14/\sqrt{n}} = -0.25$$

which implies that $\bar{x}_* = 103 - 0.25\left(\frac{14}{\sqrt{n}}\right)$

Step 3. Solving for n

It follows from Steps 1 and 2 that

$$100 + 1.645\left(\frac{14}{\sqrt{n}}\right) = 103 - 0.25\left(\frac{14}{\sqrt{n}}\right)$$

(Solving for n)

$$n = \left[\frac{(1.645 + 0.25)(14)}{(103 - 100)}\right]^2 \approx 78.2045$$

Therefore, a minimum of 79 observations must be taken to guarantee that the hypothesis test will have the desired precision.

Example

A vending machine advertises that it dispenses 225 ml cups of coffe ($\sigma = 7$ ml). You believe the mean volume of coffee per cup is something less than 225 ml. You plan to sample 40 cups of coffee from this machine to test your hypothesis.

a. If the true mean volume of coffee per cup is 223 ml, what is the power of your test at $\alpha = 0.05$? HW?

b. How many coffee cups should you sample if you want to raise the power in part (a) to 0.80?

Step 1. Write the rule for rejecting H_0 in terms of $\bar{x}_*$.

By definition,

$$
\begin{aligned}
\alpha &= \text{P(we reject } H_0 \text{ given } H_0 \text{ is true)} \\
&= P(\bar{X} < \bar{x}_* | \mu = 225) \\
&= P\left(\frac{\bar{X}-225}{7/\sqrt{n}} < \frac{\bar{x}_*-225}{7/\sqrt{n}}\right) \\
&= P\left(Z < \frac{\bar{x}_*-225}{7/\sqrt{n}}\right) = 0.05
\end{aligned}
$$

But $P(Z < -1.645) = 0.05$, so

$$\bar{x}_* = 225 - 1.645\frac{7}{\sqrt{n}}$$

Solution b)

Step 2. The power is the probability of this event under the condition that the alternative $\mu = 223$ is true.

To calculate this probability, standardize $\bar{x}$ using $\mu = 223$.

$$\begin{aligned} \text{power } &= 1 - \beta = P(\bar{X} < \bar{x}_* \text{ when } \mu = 103) \\ &= P\left(\frac{\bar{X}-223}{7/\sqrt{n}} < \frac{\bar{x}_*-223}{7/\sqrt{n}}\right) = 0.80 \end{aligned}$$

From our Table, $P(Z < 0.84) = 0.7995 \approx 0.80$, so

$$\frac{\bar{x}_* - 223}{7/\sqrt{n}} = 0.84$$

which implies that $\bar{x}_* = 223 + 0.84\left(\frac{7}{\sqrt{n}}\right)$

Solution b)

Step 3. Solving for n

It follows from Steps 1 and 2 that

$$225 - 1.645\left(\frac{7}{\sqrt{n}}\right) = 223 + 0.84\left(\frac{7}{\sqrt{n}}\right)$$

(Solving for n)

$$n = \left[\frac{(1.645 + 0.84)(7)}{(225 - 223)}\right]^2 \approx 75.6465$$

Therefore, a minimum of 76 observations must be taken to guarantee that the hypothesis test will have the desired precision.

Another example

A newsletter reports that 90% of adults drink milk. The researchers are interested in investigating if less than 90% of adults drink milk (at $\alpha = 0.05$). They collect a random sample of 200 adults in a certain region.
a. Calculate power of the test if the percentage of adults who drink milk is really 85%.
b. Calculate beta if the percentage of adults who drink milk is really 85%.
c. How many adults should you sample if you want to raise the power in part (a) to 0.80?

a. Calculate power of the test if the percentage of adults who drink milk is really 85%.

$$\begin{aligned} \alpha = 0.05 &= P(\text{reject } H_0 | H_0 \text{ is true }) \\ P(Z < -1.645) &= 0.05 \text{ (this is the rejection region)} \\ & \text{Z critical value is -1.645} \end{aligned}$$

$$
\begin{aligned}
\text{Power} &= P(\text{reject } H_0 | H_0 \text{ is false }) \\
&= P\left(\frac{\hat{p} - 0.90}{\sqrt{\frac{0.90(1-0.90)}{200}}} | p = 0.85 \right) \\
&= P(\hat{p} < 0.8651 | p = 0.85) \\
&= P\left(Z < \frac{0.8651 - 0.85}{\sqrt{\frac{0.85(1-0.85)}{200}}} \right) \\
&= P(Z < 0.5980) \\
&\approx 0.7250
\end{aligned}
$$

```r
pnorm(0.5980, mean =0, sd = 1);

## [1] 0.72508
```

b. Calculate beta if the percentage of adults who drink milk is really 85%.

$$\begin{aligned}\beta &= 1 - \text{power} \\ &= 1 - 0.73 \\ &= 0.27\end{aligned}$$

c. How many adults should you sample if you want to raise the power in part (a) to 0.80? Recall part a. Calculate power of the test if the percentage of adults who drink milk is really 85%.

Step 1: Write the rule for rejecting Ho in term of $\hat{p}^*$

$$
\begin{aligned}
\alpha &= P(\text{reject } H_0 | H_0 \text{ is true }) \\
&= P(\hat{p} < \hat{p}^* | p = 0.90) \\
&= P\left(\frac{\hat{p} - 0.90}{\sqrt{\frac{0.90(1-0.90)}{n}}} < \frac{\hat{p}^* - 0.90}{\sqrt{\frac{0.90(1-0.90)}{n}}} \right) \\
&= P\left(Z < \frac{\hat{p}^* - 0.90}{\sqrt{\frac{0.90(1-0.90)}{n}}} \right) \\
&= 0.05
\end{aligned}
$$

But, $P(Z < -1.645) = 0.05$

So, $\hat{p}^* = 0.90 - 1.645\sqrt{\frac{0.90(1-0.90)}{n}}$

Step 2: The power is the probability of this event under the condition that the alternative $p = 0.85$ is true Standardize $\hat{p}$ using $p = 0.85$

$$P\left(\frac{\hat{p} - 0.85}{\sqrt{\frac{0.85(1-0.85)}{n}}} < \frac{\hat{p}^* - 0.85}{\sqrt{\frac{0.85(1-0.85)}{n}}}\right) = 0.80$$

From R, $P(Z < 0.8416212) = 0.80$

```
qnorm(0.80, mean=0,sd=1);

## [1] 0.8416212
```

So, $\frac{\hat{p}^* - 0.85}{\sqrt{\frac{0.85(1-0.85)}{n}}} = 0.8416$

In terms of $\hat{p}^*$: $\hat{p}^* = 0.85 + 0.8416\sqrt{\frac{0.85(1-0.85)}{n}}$

Solving for n, it follows from Steps 1 and 2 that

$$
\begin{aligned}
0.90 - 1.645\sqrt{\frac{0.90(1-0.90)}{n}} &= 0.85 + 0.8416\sqrt{\frac{0.85(1-0.85)}{n}} \\
0.90 - 0.85 &= 1.645\sqrt{\frac{0.90(1-0.90)}{n}} \\
&+ 0.8416\sqrt{\frac{0.85(1-0.85)}{n}} \\
0.05 &= \frac{1}{\sqrt{n}}(0.4935 + 0.3005) \\
0.05\sqrt{n} &= 0.7940 \\
n &= 252.1744
\end{aligned}
$$

Final answer: $n = 253$.

Another example

A newsletter reports that 90% of adults drink milk. The researchers are interested in investigating if less than 90% of adults drink milk (at $\alpha = 0.05$). They collect a **random sample of 100 adults** in a certain region.

a. Calculate power of the test if the percentage of adults who drink milk is really 85%.

a. Calculate power of the test if the percentage of adults who drink milk is really 85%.

$$\begin{aligned} \alpha = 0.05 &= P(\text{reject } H_0 | H_0 \text{ is true }) \\ P(Z < -1.645) &= 0.05 \text{ (this is the rejection region)} \\ & \quad \text{Z critical value is -1.645} \end{aligned}$$

$$
\begin{aligned}
\text{Power} &= P(\text{reject } H_0 | H_0 \text{ is false }) \\
&= P\left(\frac{\hat{p} - 0.90}{\sqrt{\frac{0.90(1-0.90)}{100}}} < -1.645 | p = 0.85 \right) \\
&= P(\hat{p} < 0.85065 | p = 0.85) \\
&= P\left(Z < \frac{0.85065 - 0.85}{\sqrt{\frac{0.85(1-0.85)}{100}}} \right) \\
&= P(Z < 0.0182) \\
&\approx 0.5072
\end{aligned}
$$

```
pnorm(0.0182, mean =0, sd = 1);

## [1] 0.5072603
```

Another example

A newsletter reports that 90% of adults drink milk. The researchers are interested in investigating if less than 90% of adults drink milk (at $\alpha = 0.05$). They collect a **random sample of 50 adults** in a certain region.

a. Calculate power of the test if the percentage of adults who drink milk is really 85%.

a. Calculate power of the test if the percentage of adults who drink milk is really 85%.

$$\begin{aligned} \alpha = 0.05 &= P(\text{reject } H_0 | H_0 \text{ is true }) \\ P(Z < -1.645) &= 0.05 \text{ (this is the rejection region)} \\ &\quad \text{Z critical value is -1.645} \end{aligned}$$

$$
\begin{aligned}
\text{Power} &= P(\text{reject } H_0 | H_0 \text{ is false }) \\
&= P\left(\frac{\hat{p} - 0.90}{\sqrt{\frac{0.90(1-0.90)}{50}}} < -1.645 | p = 0.85 \right) \\
&= P(\hat{p} < 0.8302 | p = 0.85) \\
&= P\left(Z < \frac{0.0.8302 - 0.85}{\sqrt{\frac{0.85(1-0.85)}{50}}} \right) \\
&= P(Z < -0.3921) \\
&\approx 0.3475
\end{aligned}
$$

```
pnorm(-0.3921, mean =0, sd = 1);

## [1] 0.3474922
```

If we keep α at the same size, larger sample sizes increase the power of test because sampling variability (sampling distributions) are much narrower.

Power Formulas

If $H_a : \mu > \mu_0$, then

$$Power(\mu_a) = \pi(\mu_a) = P\left[Z \geq Z_\alpha + \frac{\sqrt{n}(\mu_0 - \mu_a)}{\sigma}\right]$$

where $\mu_a > \mu_0$. Need to know

- Standard deviation, σ.
- Significance level, α.
- Effect size you want to detect, $\mu_0 - \mu_a$.

Power Formulas

If $H_a : \mu < \mu_0$, then

$$Power(\mu_a) = \pi(\mu_a) = P\left[Z \leq -Z_\alpha + \frac{\sqrt{n}(\mu_0 - \mu_a)}{\sigma}\right]$$

where $\mu_a < \mu_0$. Need to know

- Standard deviation, σ.
- Significance level, α.
- Effect size you want to detect, $\mu_0 - \mu_a$.

Power Formulas

If $H_a : \mu \neq \mu_0$, then

$$Power(\mu_a) = \pi(\mu_a)$$

$$= 1 - P\left[-Z_{\alpha/2} + \frac{\sqrt{n}(\mu_0 - \mu_a)}{\sigma} \leq Z \leq Z_{\alpha/2} + \frac{\sqrt{n}(\mu_0 - \mu_a)}{\sigma}\right]$$

where $\mu_a \neq \mu_0$. Need to know

- Standard deviation, σ.
- Significance level, α.
- Effect size you want to detect, $\mu_0 - \mu_a$.

Calculations of power (or of error probabilities) are useful for planning studies because we can make these calculations before we have any data. Once we actually have data, it is more common to report a P-value rather than a reject-or-not decision at a fixed significance level α. The P-value measures the strength of the evidence provided by the data against H_0. It leaves any action or decision based on that evidence up to each individual. Different people may require different strengths of evidence.