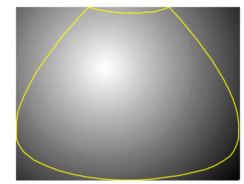
Hypothesis testing

Statistical logic

Hypothesis testing

A game of chance

- I claim that in a bowl there are 10 red and 90 blue marbles. You are allowed to draw 10 marbles. What will you conclude if:
  - 2 red and 8 blue marbles are drawn
  - 5 red and 5 blue marbles are drawn
  - 9 red and 1 blue marbles are drawn



Hypothesis testing

It is simple logic

- In the previous example, one expects to draw significantly more blue marbles if the statement is true
- A similar logic applies to formal hypothesis testing as well
- The conclusion we reach may or may not be correct

Hypothesis testing

The power of rejection

- If a pharmaceutical company claims that their product relieves any headache always within 15 minutes, how can you test its truth?
  - Take one next time you have a headache and it works
  - Take one for 99 times and it works
  - Take one the 100th time and it does not

Terminology

Errors, we have them

- α probability of making Type I error, level of significance
- $1-\alpha$ Level of confidence
- β probability of making Type II error, it is a function not a single number
- $1-\beta$ Power of a test

- Although α and β are related, their relation is not an additive one. They do not add to any particular value.

The hypotheses

What we test

- Hypothesis testing requires a pair of hypotheses:
  - A null hypothesis, H_0, the opposite of what we wish to prove
  - An alternate hypothesis, H_a, the statement we wish to prove

Possible errors

Not all conclusions are correct

- The U.S. criminal justice system is based on the premise that the individual is "*innocent until proven guilty*."

		True State	
		Innocent	*Guilty*
Verdict	*Not guilty*	*Correct*	*Error*
	Guilty	*Error*	*Correct*

Possible errors

A statistical view

- Type I and Type II errors

		Null hypothesis is	
		True	*False*
Test result	*Not reject*	*Correct*	*Type II*
	Reject	*Type I*	*Correct*

The hypotheses

How we test

- The process
  - Considering the desired conclusion, state the alternate hypothesis as the statement to prove
  - State its exact opposite as the null hypothesis to test
  - Test the null hypothesis with the evidence from the sample
  - If the evidence supports, reject it and consequently accept the alternate hypothesis

An example

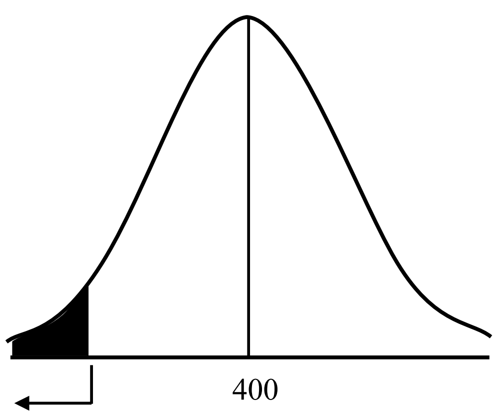
State the hypotheses

- H_o: $\mu \geq 400$ lbs
  H_a: $\mu < 400$ lbs
  We would like to prove that the population mean is less than 400 lbs
- H_o: $\pi \leq 0.40$
  H_a: $\pi > 0.40$
  We would like to prove that the population proportion is greater than 40%

One tailed test

Trying to prove “less than”

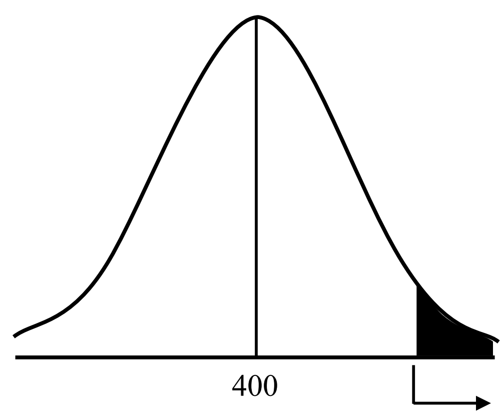
H_o: $\mu \geq 400$ lbs
H_a: $\mu < 400$ lbs



If the observed value falls in this range reject Ho

One tailed tests

Trying to prove “greater than”



H_o: $\mu \leq 400$ lbs
H_a: $\mu > 400$ lbs

If the observed value falls in this range reject Ho

One tailed tests

An example

- Before extending the warranty on picture tubes, a TV manufacturer wants to prove that the life expectancy of its tubes is greater than 5 years (60 months) with 95% level of confidence (5% level of significance)

One tailed test

An example continued

H_o: $\mu \leq 60$ months
H_a: $\mu > 60$ months

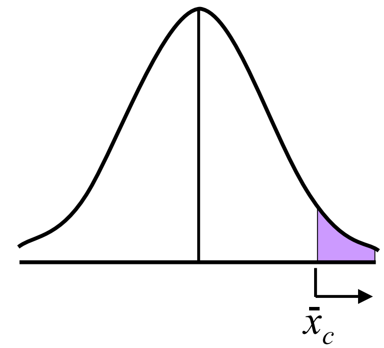
Given that

$\alpha = 0.05$

$z_\alpha = 1.645$

n = 64 tubes

$\sigma = 8$ months



Two-tailed tests

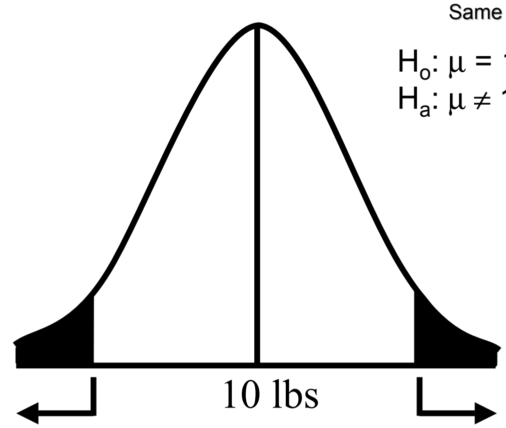
Test for equality

- Testing for the population parameter being in a range may be necessary
- This calls for a two tailed test with the same statistical logic

A mirror image

Same bird with two tails

H_o: $\mu = 10$ lbs
H_a: $\mu \neq 10$ lbs



If the observed value falls in this range reject Ho

If the observed value falls in this range reject Ho

An example

Two-tailed test

- Bags of potato expected to be 10 lbs
  - Too much is not desirable
  - Too little is not desirable
  - If the process deviates from this, the worker is instructed to stop the bagging machine and re-calibrate it

An example

Two tailed test

H_o: μ = 10 lbs

H_a: $\mu \neq$ 10 lbs ⬅ We want to prove this

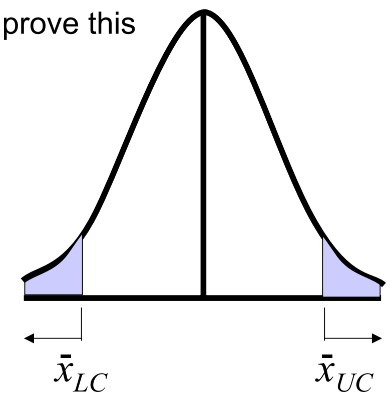
Given that

$\alpha = 0.05$

$z_\alpha = 1.96$

$n = 100$ bags

$\sigma = 1$ lbs



An example

Two-tailed test

$$\sigma_{\bar{x}} = \frac{1}{\sqrt{100}} = 0.1$$

$$\bar{x}_{UC} = 10 + (1.96)\,(0.1) = 10.2$$

$$\bar{x}_{LC} = 10 - (1.96)\,(0.1) = 9.8$$

Reject Ho if: $\bar{x} > \bar{x}_{UC}$

$\bar{x} < \bar{x}_{LC}$