Computer Science

CSC380: Principles of Data Science

Statistics 2

Xinchen Yu

Outline

- Interval estimation
- Hypothesis testing

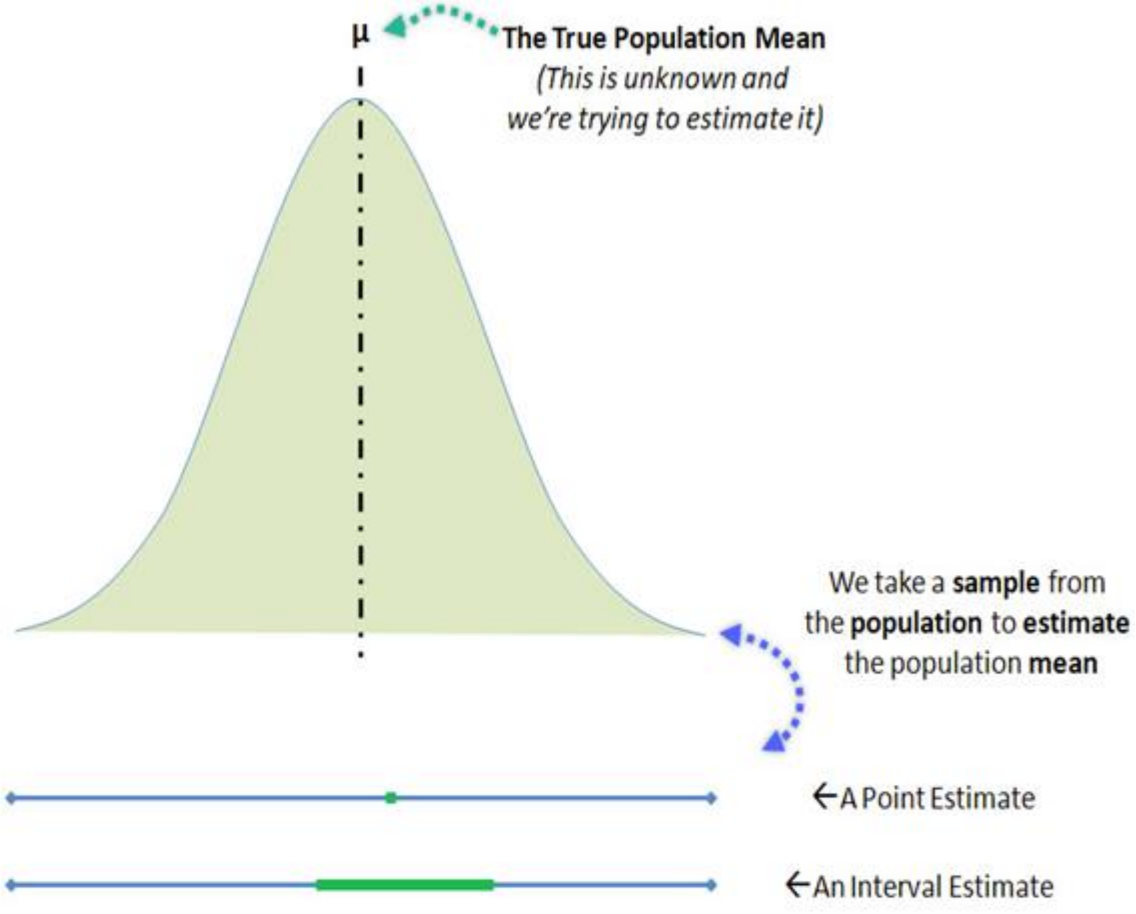
Interval estimation

Motivation

- Point estimation:
  - "Given the data, I estimate the bias of the coin to be 0.73"
  - "Given the data, I estimate the mean height of UA students to be 172cm"

- In many applications, we'd like to make statements with uncertainty quantifications
  - "Given the data, I estimate the bias of the coin to be 0.73 $\pm$ 0.05"
  - "Given the data, I estimate the mean height of UA students to be 172 $\pm$ 2cm"

- This is called *interval estimation*



Interval Estimation: basic setup

$$\theta \;\to X_1, \dots, X_n \;\to I_n = [\hat{\theta}_n \pm b_n]$$

data generation process — Confidence Interval (CI) for θ

Examples

Coin toss: $\theta = p,\ \ X_1, \dots, X_n \sim \mathrm{Bernoulli}(p)$

Student height: $\theta = \mu, X_1, \dots, X_n \sim N(\mu, 8^2)$

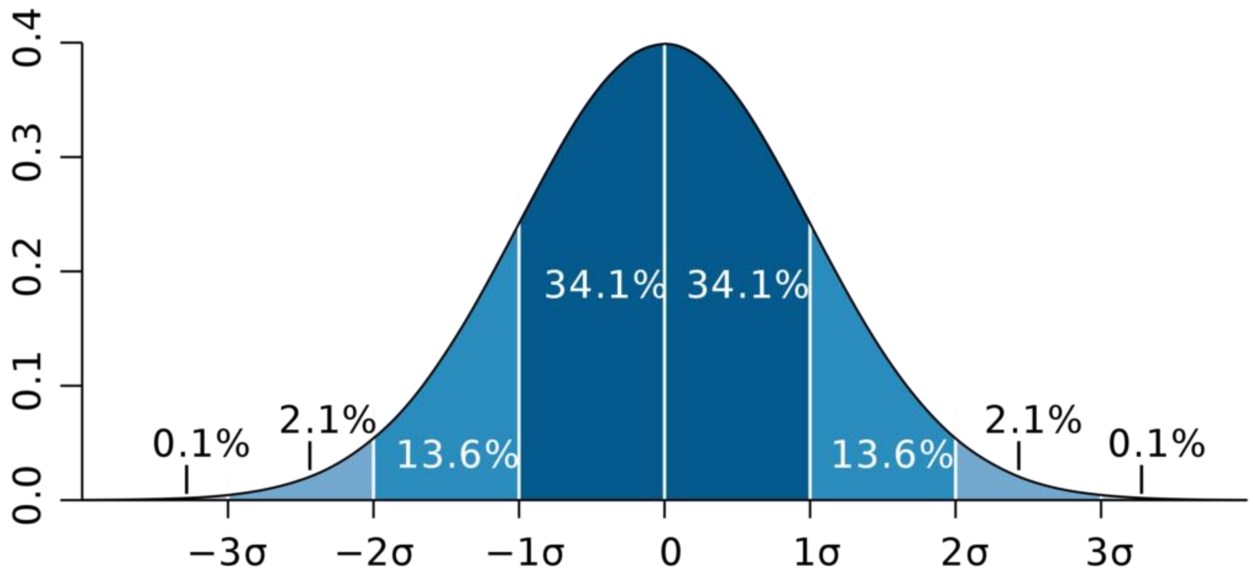
Goal: construct I_n using data, such that with 95% confidence (say),

$$\theta \in I_n$$

We will mostly focus on estimating θ = population mean, and will take $\hat{\theta}_n$ = sample mean.

How to choose b_n? **uncertainty of our estimate**

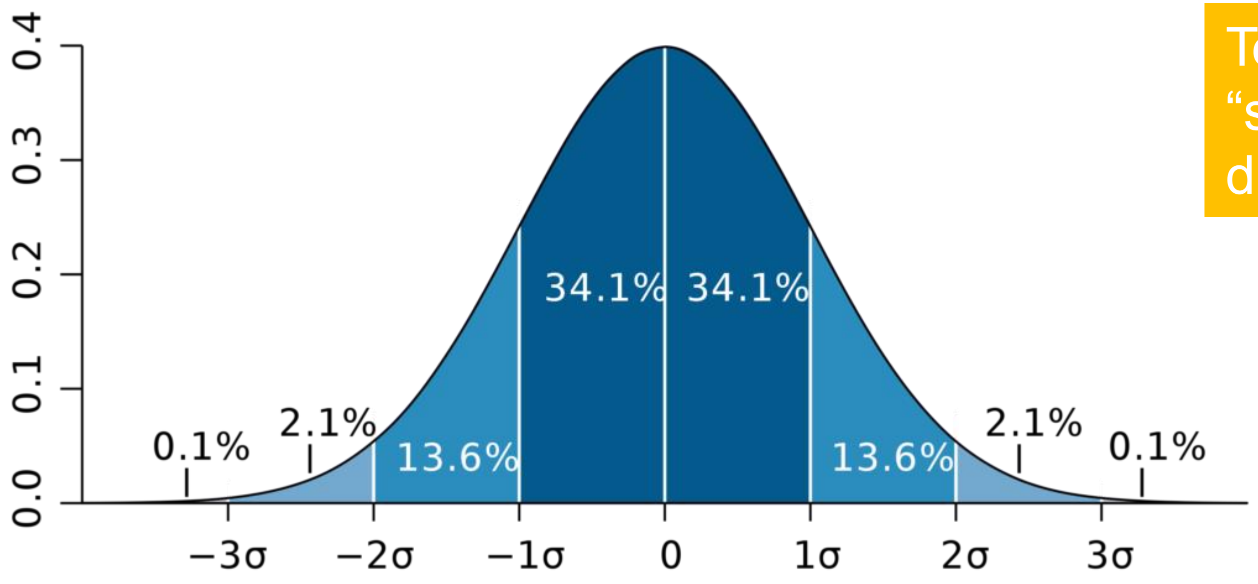
Recall: Normal distribution



For $X \sim N(\mu, \sigma^2)$, we can transform it into $X - \mu \sim N(0, \sigma^2)$

- the area under a normal distribution curve (PDF) represents probability.
- the total area under the curve is equal to 100%.
- the area within a certain range of values corresponds to the probability of a random variable falling within that range.

Recall: Normal distribution



Terminology: "standard" normal distribution := $N(0,1)$

Fact If $X \sim N(\mu, \sigma^2)$ or $X - \mu \sim N(0, \sigma^2)$, then

$$P(-1.96\sigma \leq X - \mu \leq 1.96\sigma) = 0.95$$

In words, with 95% confidence, X falls within 1.96 standard deviation of μ

$$P(X - 1.96\sigma \leq \mu \leq X + 1.96\sigma) = 0.95$$

i.e, with 95% confidence, μ falls within 1.96 standard deviation of X

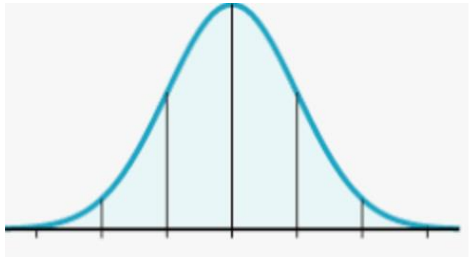
$[X \ - 1.96\sigma, X + 1.96\sigma]$ **is a 95% confidence interval for μ**

Constructing confidence interval

- **We know if $X \sim N(\mu, \sigma^2)$, then $[X - 1.96\sigma, X + 1.96\sigma]$ is a 95% CI for μ**

- **Fact**: Let $X_1, \ldots, X_n$ be iid with mean μ and variance σ^2. Then for large n, the sample mean $\bar{X}_n$ roughly follow a normal distribution:

$$\bar{X}_n \approx N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$$



Corollary with 95% confidence, μ lies within $1.96\frac{\sigma}{\sqrt{n}}$ of $\bar{X}_n$

Our confidence interval for μ: $I_n = [\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}]$

Example: UA student height

Example Assume that UA students' heights (in centimeters) follow $N(\mu, 8^2)$, and we observe 4 students' heights:

163, 171, 179, 167

Find a 95% confidence interval for μ

Solution

our CI for μ: $I_n = [\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}]$

Sample mean $= 170$; population stddev $\sigma = 8$; sample size n=4

Plugging in all values, $I_n = [170 \pm 7.84] = [162.1, 177.8]$

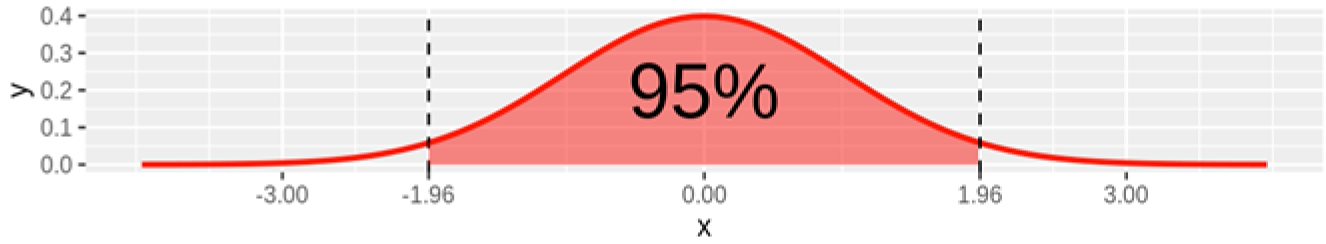
Confidence intervals: extensions

Given if $X \sim N(\mu, \sigma^2)$ or $X - \mu \sim N(0, \sigma^2)$, then

$$P(-1.96\sigma \leq X - \mu \leq 1.96\sigma) = 0.95$$

Where does the 1.96 come from?

st.norm.ppf(0.975) gives 1.96



Φ: **standard normal CDF**

Fact If $X \sim N(\mu, \sigma^2)$, then

$$P(-k\,\sigma \leq X - \mu \leq k\,\sigma) = 2\Phi(k) - 1 = p$$

$$2\Phi(k) - 1 = 0.95 \Rightarrow \; k = \Phi^{-1}\left(\frac{0.95+1}{2}\right) = \Phi^{-1}(0.975) = 1.96$$

k: $\left(\frac{1+p}{2}\right)$-quantile of *the standard normal distribution*

CI for μ: $I_n = [\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}]$

Confidence intervals: extensions

- What if we'd like to find 99% confidence interval? 99.9%? 90%?

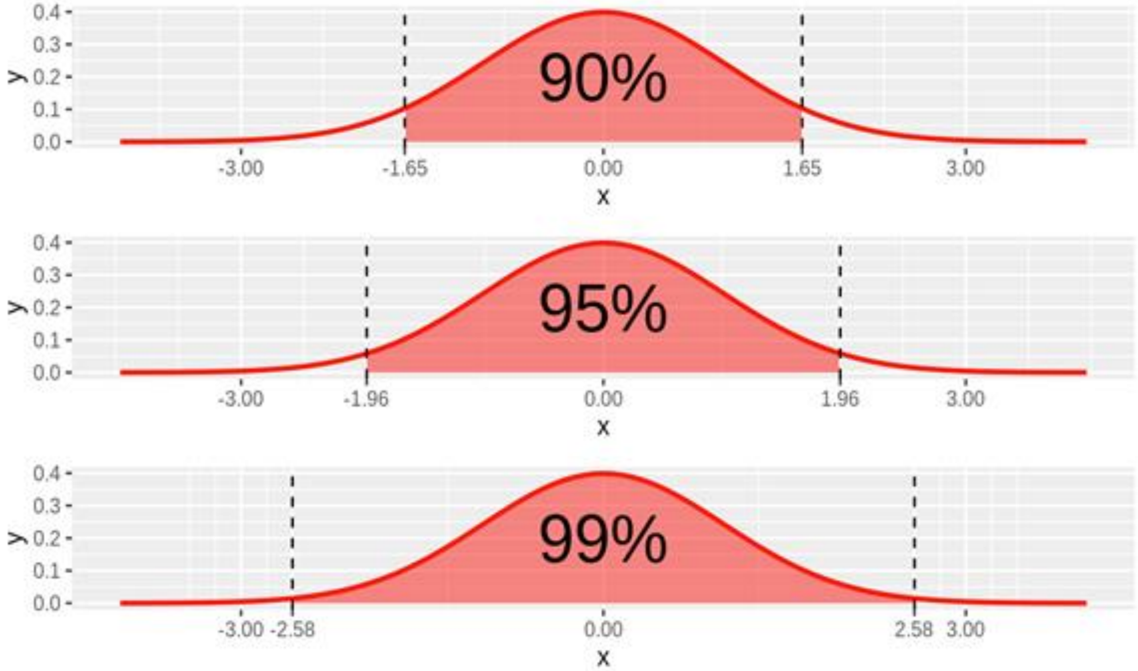
Fact If $X \sim N(\mu, \sigma^2)$, then

$$P(-k\,\sigma \le X - \mu \le k\,\sigma) = 2\Phi(k) - 1 = p$$

Our p confidence interval for μ:

$$I_n = [\bar{X}_n \pm \Phi^{-1}\left(\frac{p+1}{2}\right)\frac{\sigma}{\sqrt{n}}] = [\bar{X}_n \pm k\,\frac{\sigma}{\sqrt{n}}]$$

p	$\mathrm{k} = \Phi^{-1}\left(\frac{p+1}{2}\right)$
0.95	1.96
0.99	2.58
0.999	3.29



Example: UA student height

Example Assume that UA students' heights (in centimeters) follow $N(\mu, 8^2)$, and we observe 4 students' heights:

163, 171, 179, 167

Find 99%, 99.9% confidence intervals for μ

p	$\Phi^{-1}\left(\frac{p+1}{2}\right)$
0.95	1.96
0.99	2.58
0.999	3.29

Solution

our p-CI for μ: $I_n = [\bar{X}_n \pm \Phi^{-1}\left(\frac{p+1}{2}\right)\frac{\sigma}{\sqrt{n}}]$

$p = 0.99 \Rightarrow [159.7, 180.3]$

$p = 0.999 \Rightarrow [156.9, 183.1]$

Confidence interval: observations

$$p\text{-CI for } \mu\text{: } I_n = [\bar{X}_n \pm \Phi^{-1}\left(\frac{p+1}{2}\right)\frac{\sigma}{\sqrt{n}}]$$

$p = 0.95 \Rightarrow [162.1, 177.8]$
$p = 0.99 \Rightarrow [159.7, 180.3]$
$p = 0.999 \Rightarrow [156.9, 183.1]$

The center is always at $\bar{X}_n$

The width of the interval depends on:

- Sample size n: width smaller when n larger
- Confidence level p: width larger when p closer to 1
- Population stddev σ: width larger when σ large (more noise)

What if σ is unknown?

- We will address this soon..

Is confidence = probability?

Example Assume that UA students' heights (in centimeters) follow $N(\mu, 8^2)$, and we observe 4 students' heights:

163, 171, 179, 167

we found that a 95% CI for μ is $[162.1, 177.8]$

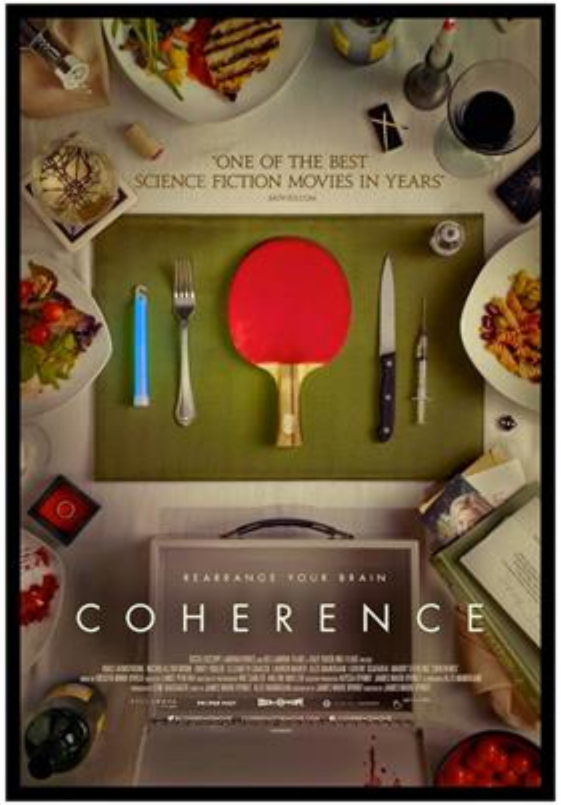
Can we say "with probability 95%, the population mean height μ lies in interval $[162.1, 177.8]$"?

No! This is a common misinterpretation

- μ is deterministic, and $[162.1, 177.8]$ is deterministic,
- Proposition $\mu \in [162.1, 177.8]$ is either true or false!

Then, what does "95% probability" mean?

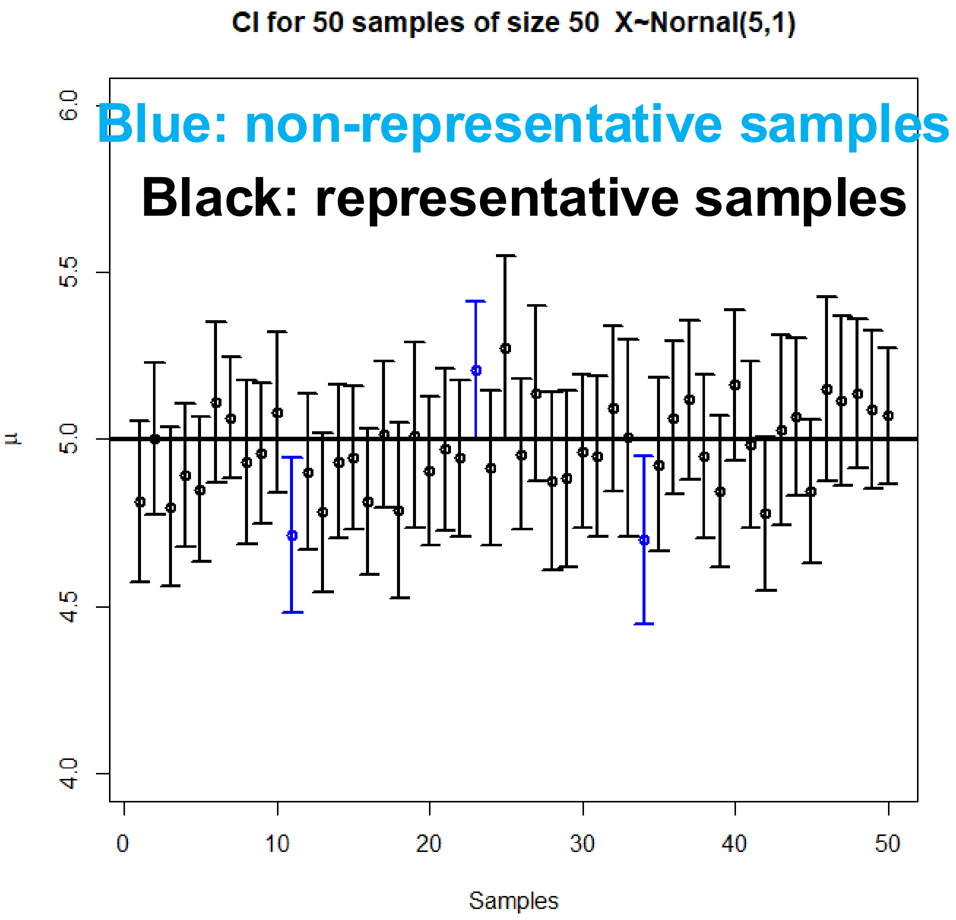
Interpreting CI (think of parallel universe…)



Multiple different universes…

Caveat: interpreting confidence intervals

Recommended point of view:



universe 1: get sample 1, and confidence interval 1

universe 2: get sample 2, and confidence interval 2

……

universe 50: get confidence interval 50

True: With probability 0.95 *over the draw of a sample*, $[\bar{X}_n \pm 1.96 \frac{\sigma}{\sqrt{n}}]$ contains μ

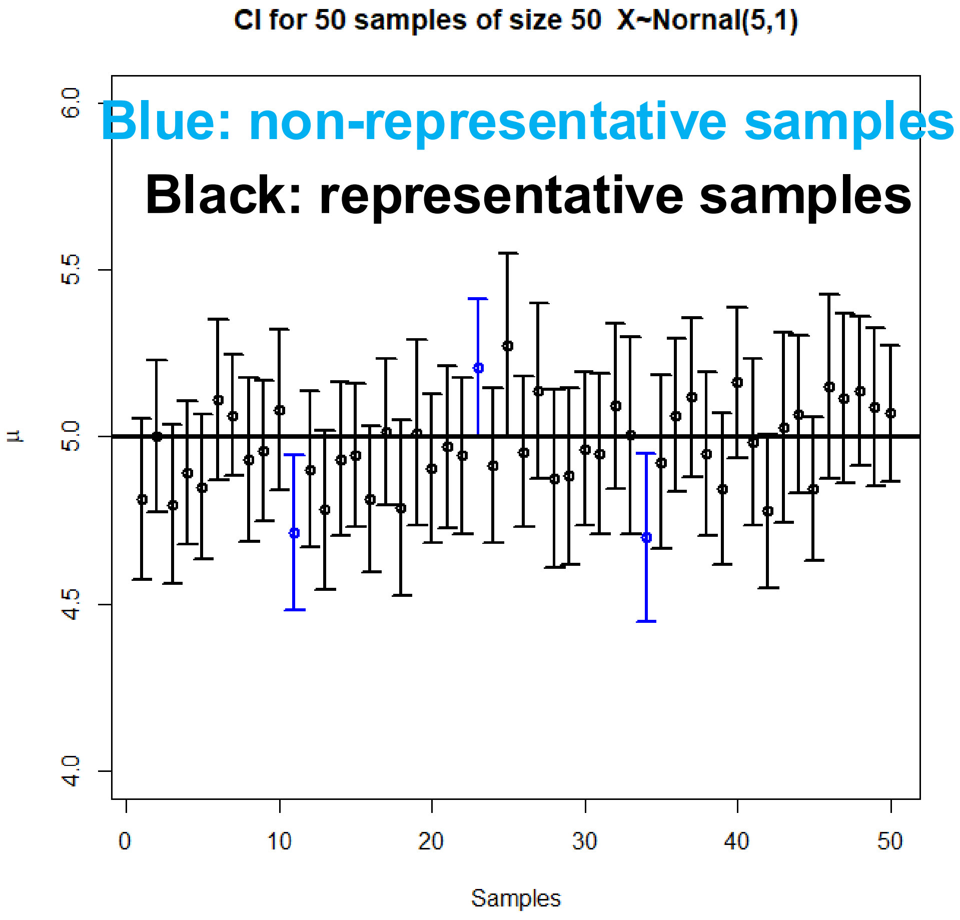
Confidence interval: interpretation

Example Assume that UA students' heights (in centimeters) follow $N(\mu, 8^2)$, and we observe 4 students' heights:

163, 171, 179, 167

True: With probability 0.95 *over the draw of a sample*, $[\bar{X}_n \pm 1.96 \frac{\sigma}{\sqrt{n}}]$ contains μ

50 draws of samples
$\Rightarrow$ 50 CIs
$\Rightarrow$ expect $50 \times 95\% = 47.5$ CI's to contain μ



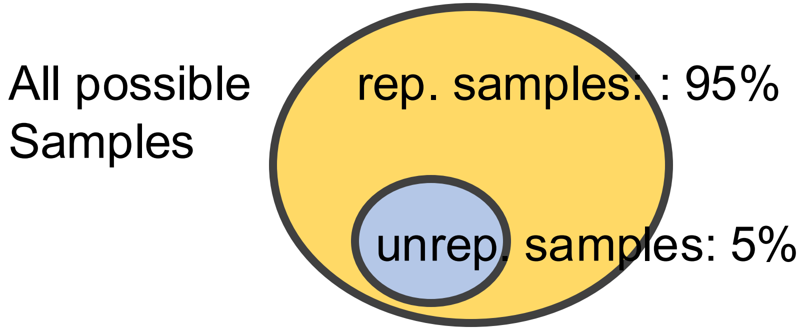
Confidence interval: interpretation

Example Assume that UA students' heights (in centimeters) follow $N(\mu, 8^2)$, and we observe 4 students' heights:

163, 171, 179, 167

True: With probability 0.95 *over the draw of a sample*, $[\bar{X}_n \pm 7.84]$ contains μ

As long as we are not extremely unlucky / our sample is mildly representative, my CI contains μ



Confidence interval for mean with unknown variance

Example Assume that UA students' weights (in kgs) follow $N(\mu, \sigma^2)$, and we observe 4 students' weights:

60, 65, 70, 75

Find a 95% confidence interval for μ

Note The CI construction before $[\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}]$ no longer works, since σ is *unknown*

How to fix this?

The student-t distribution

- $[\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}]$ no longer works: σ is unknown

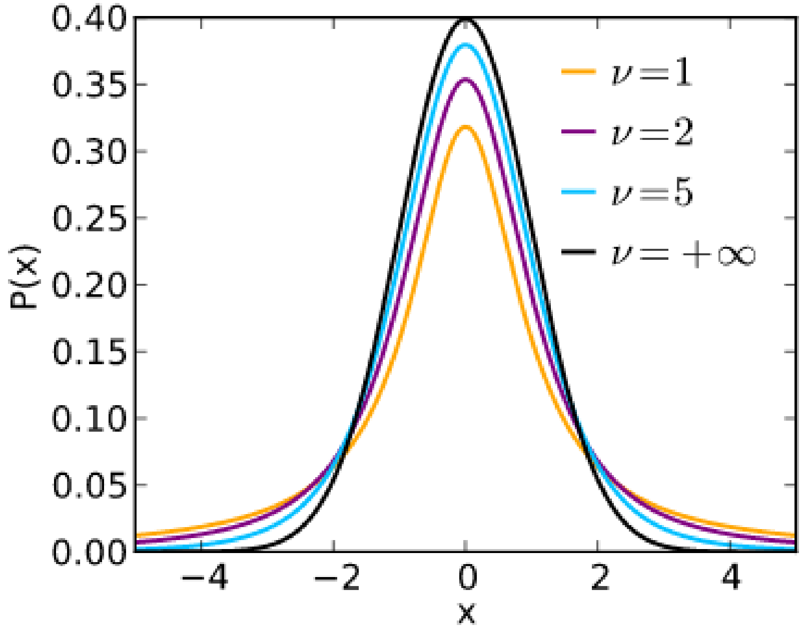
Fact $X_1, \dots, X_n$ is an iid sample with unknown μ & σ^2.

Let $sample\ stddev$: $\hat{\sigma}_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}$. Then, approximately:

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\hat{\sigma}_n} \sim \text{student-t}(n-1)$$

t-statistic

degree of freedom

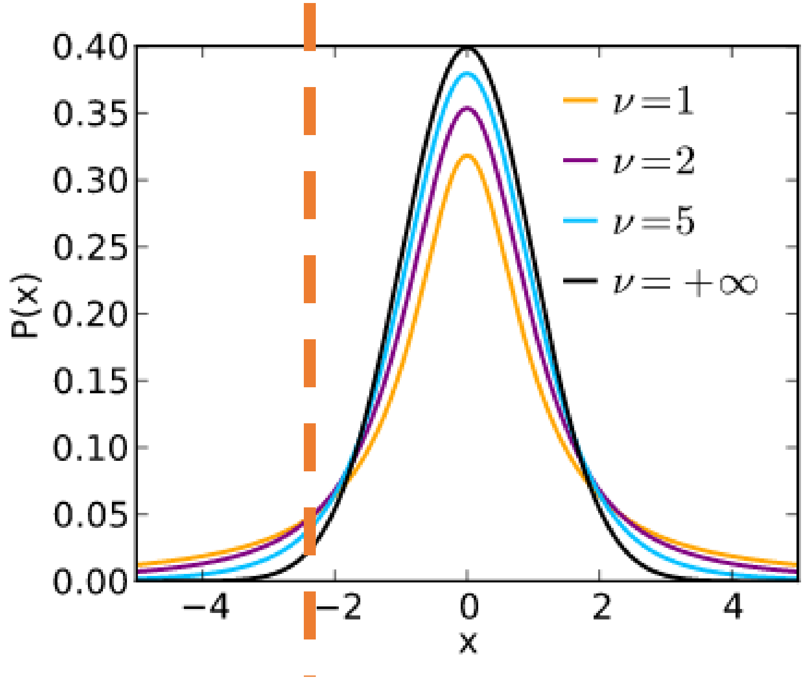


student-t(ν) is a family of distributions

The student-t distribution

student-t(ν) distribution family

- goes to Gaussian when ν is large
- generally has heavier tail than Gaussian




```
import scipy.stats as st

st.t.ppf(0.975,df=3)
=> 3.18

st.t.ppf(0.975,df=5)
=> 2.57

st.t.ppf(0.975,df=10)
=> 2.23

st.t.ppf(0.975,df=100)
⇒ 1.98
```

Recall:
st.norm.ppf(0.975) gives 1.96

CI: $\left[\bar{X}_n - w\frac{\hat{\sigma}_n}{\sqrt{n}}, \bar{X}_n + w\frac{\hat{\sigma}_n}{\sqrt{n}}\right]$, w: $\left(\frac{1+p}{2}\right)$−quantile of the $t(n-1)$ distribution

Confidence interval for mean with unknown variance

Example Assume that UA students' weights (in kgs) follow $N(\mu, \sigma^2)$, and we observe 4 students' weights:

$$60, 65, 70, 75$$

st.t.ppf(0.975,df=3)
=> 3.18

Find a 95% confidence interval for μ

Solution With 95% confidence, **= 6.45**

$$\Rightarrow \mu \in \left[\bar{X}_4 - 3.18\frac{\hat{\sigma}_4}{\sqrt{4}}, \bar{X}_4 + 3.18\frac{\hat{\sigma}_4}{\sqrt{4}}\right]$$

= 67.5

Plugging data,

our CI is $[67.5 - 10.3, 67.5 + 10.3] = [57.2, 77.8]$ **Our confidence interval**

Confidence interval for mean with unknown variance

General result given a sample $X_1, \ldots, X_n$ drawn from a distribution with mean μ, a p-confidence interval (e.g. p=95%) is

$$\left[\bar{X}_n - w\frac{\hat{\sigma}_n}{\sqrt{n}}, \bar{X}_n + w\frac{\hat{\sigma}_n}{\sqrt{n}}\right],$$

where w is the $\left(\frac{1+p}{2}\right)$-quantile of the $t(n-1)$ distribution

Example p=0.95, n=4 $\Rightarrow w = 3.18$

```
st.t.ppf(0.975,df=3)
=> 3.18
```

p=0.99, n=4 $\Rightarrow w = 5.84$

p=0.99, n=9 $\Rightarrow w = 3.35$

Confidence interval: closing remarks

How to construct confidence intervals for μ?

- When σ is known
  - $\boldsymbol{CI}$: $\left[\bar{X}_n - k\frac{\sigma}{\sqrt{n}}, \bar{X}_n + k\frac{\sigma}{\sqrt{n}}\right]$, k: $\left(\frac{1+p}{2}\right)$−quantile of the standard normal distribution

    st.norm.ppf((1+p)/2)

- When σ is unknown
  - $\boldsymbol{CI}$: $\left[\bar{X}_n - w\frac{\hat{\sigma}_n}{\sqrt{n}}, \bar{X}_n + w\frac{\hat{\sigma}_n}{\sqrt{n}}\right]$, w: $\left(\frac{1+p}{2}\right)$−quantile of the $t(n-1)$ distribution

    st.t.ppf((1+p)/2,df=n-1)

Hypothesis testing

Announcements

- Fill out SCS at scsonline.ucatt.arizona.edu
- If 80% of class complete the survey, one of the lowest quizzes grade will be replaced with full points (1.5/1.5 pts)

- My office hour next Thursday will change to Monday, Dec 15

- A note on final project
  - Please use the following to print the output of your best system:
    - **from** sklearn.metrics **import** classification_report
    - print(classification_report(y_true, y_pred))

Hypothesis

- Statements about parameter / property θ of a distribution / population

Examples

- Average GPA < 2.8
- Probability of head of a coin > 0.6
- People eat more on weekends than weekdays

Simple vs. composite hypotheses

$\theta = 3.2$ (simple) , $\theta \in \{3.2, 4\}$ (composite), $\theta \in [3.2, 4]$ (composite)

One-sided vs. two-sided

$\theta > 3.2$ (one-sided) , $\theta < 1.5 \text{ or } \theta > 3.2$ (two-sided), $\theta \neq 2$ (two-sided)

Hypothesis testing

Hypothesis testing: choosing from two hypotheses:

- Null hypothesis H_0
  - Status quo, assumption believed to be true
  - Coin in my pocket, probability of head p = 0.5

- Alternative hypothesis H_1: Complement of H_0
  - Novel finding after research
  - Coin has probability of head $p \neq 0.5$

Hypothesis testing

- How to test?
- Design experiment, collect data, check:

If data shows strong evidence against H_0:
Reject H_0 (in favor of H_1)

Else
Do not reject H_0 **Note: does not necessarily mean "accept H_0 "**

- Analogy with the legal principle:
  - Presume innocent (H_0) until proven guilty (H_1) with strong evidence against innocence

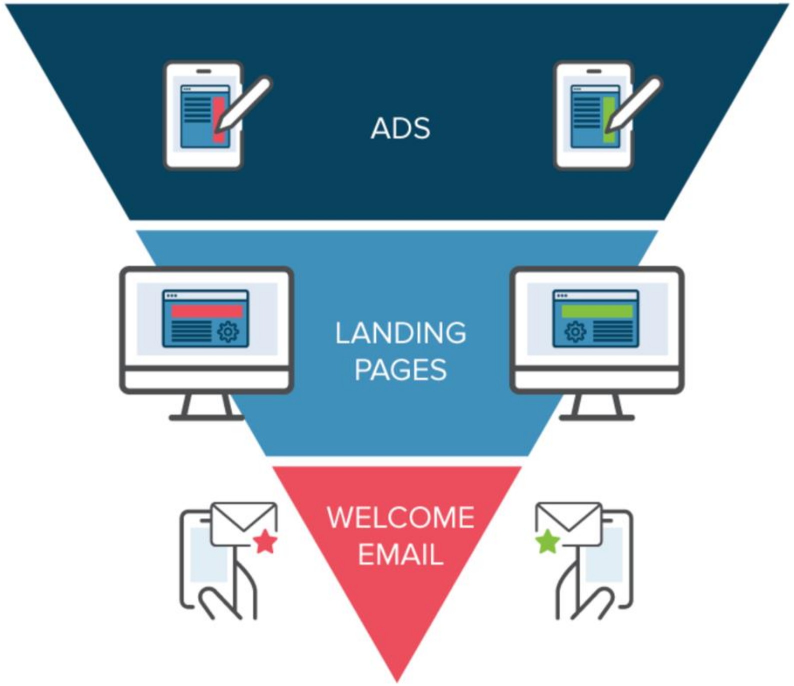
Application: A/B Testing

These days, Internet companies run A/B testing extensively

Try out an alternative of user interface (UI) on **<u>randomly chosen subset of users</u>** to collect their feedback (e.g. rating)

- E.g, choosing b/w **list view** vs grid view

How do we know if the new UI is better than older one? (i.e., statistically significant)



(from optimizely.com)

Application: A/B Testing

Evaluator:	1	2	3	4	5	6
Old UI	5	2	2	5	4	2
New UI	4	4	1	3	3	5

Compute the score differences:

Evaluator:	1	2	3	4	5	6
Score difference X	-1	+2	-1	-2	-1	+3

Can view X's as drawn from some distribution with unknown mean μ

"Does new UI improve over old UI?" is now a hypothesis testing problem:

$$H_0\colon \mu \leq 0, \qquad H_1\colon \mu > 0$$

we can perform e.g. t-test based on data (we will see)

Is the true μ equal to 168?

Example Assume that UA students' heights (in centimeters) follow $N(\mu, 8^2)$, test the hypothesis

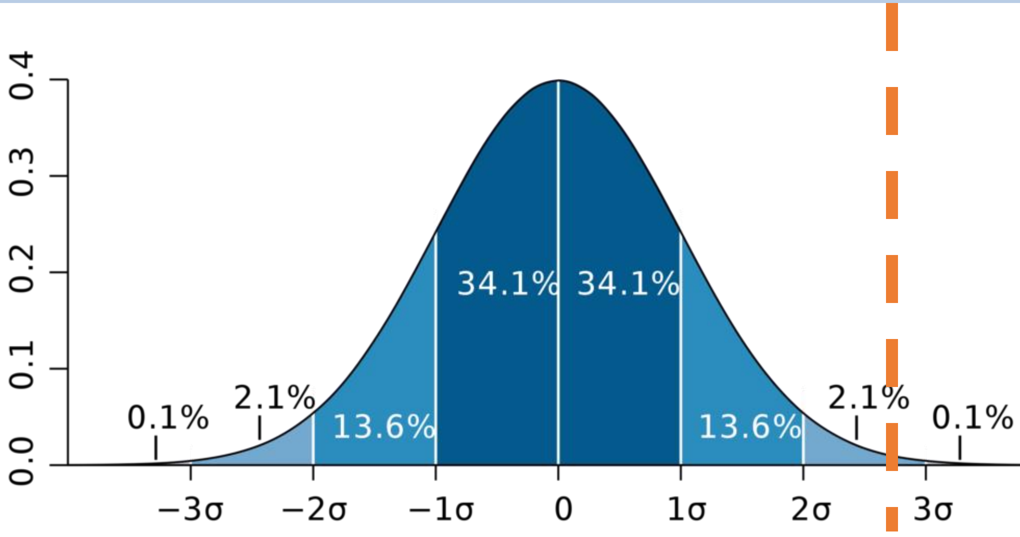
$$H_0\colon \mu = 168, \qquad\qquad H_1\colon \mu \neq 168$$

Suppose we observe 4 students' heights: 173, 181, 189, 177

- We want to know if the data provides evidence against this claim H_0.
- **Fact**: $Z = \frac{\sqrt{n}(\bar{X}_n - 168)}{8} \sim N(0,1)$

Is the true μ equal to 168?

- **Fact**: $Z = \frac{\sqrt{n}(\bar{X}_n - 168)}{8} \sim N(0,1)$
- If H_0 is true:
  - Z close to 0 happen frequently
  - Z values moderately far from 0 (like -2.1, +2.5) happen rarely (only 5% of the time beyond ±1.96)

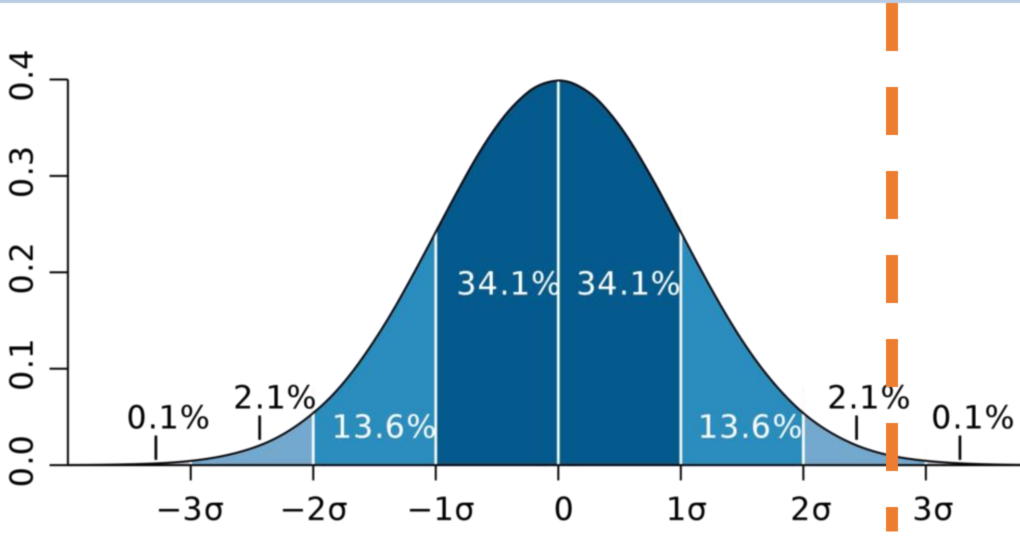


- Let's say we observe $Z = 3$
  - If H_0 is true, getting |Z| ≥ 1.96 has a probability of about 5%.
  - Surprising! this should rarely happen under H_0

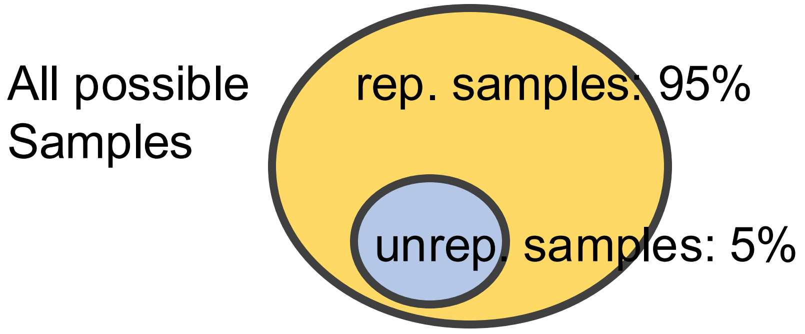
st.norm.ppf(0.975) gives 1.96

Is the true μ equal to 168?

- We observe $Z = 3$
  - If H_0 is true, getting $|Z| \geq 1.96$ has a probability of about 5%.
  - Surprising! this should rarely happen under H_0



- Two explanations:
  - (a) H_0 is TRUE, but I got unlucky
    - my sample happened to be extreme that produce such large Z values
  - (b) H_0 IS FALSE
    - My sample is actually representative of the true population
    - null hypothesis wrong: 168 is not true mean



Is the true μ equal to 168?

- Two explanations:
  - (a) H_0 is TRUE, but I got unlucky
    - my sample happened to be extreme that produce such large Z values
    - requires believing a rare event (2% probability) occurred

Type-I error: reject but H_0 is true

  - (b) H_0 IS FALSE
    - null hypothesis wrong: 168 is not true mean
    - doesn't require believing in rare events

- Need to decide: How rare does the data need to be (under H_0) before I'll reject H_0?

This is where we choose **α (significance level)**.

Is the true μ equal to 168?

Example Assume that UA students' heights (in centimeters) follow $N(\mu, 8^2)$, test the hypothesis

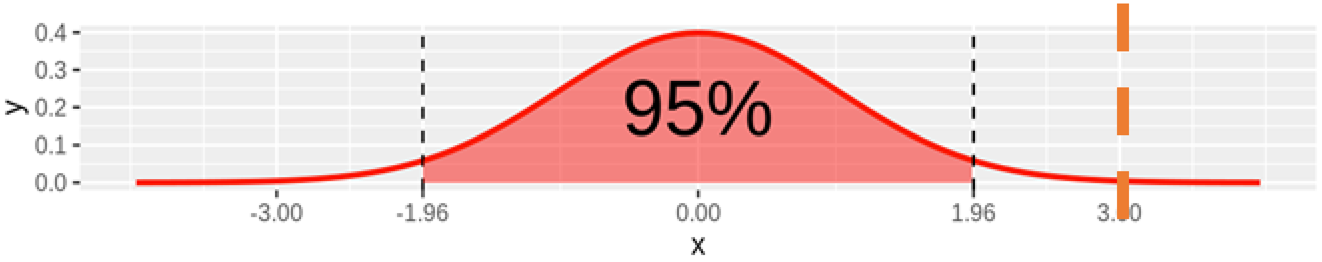
$$H_0\colon \mu = 168, \qquad H_1\colon \mu \neq 168$$

Suppose we observe 4 students' heights: 173, 181, 189, 177

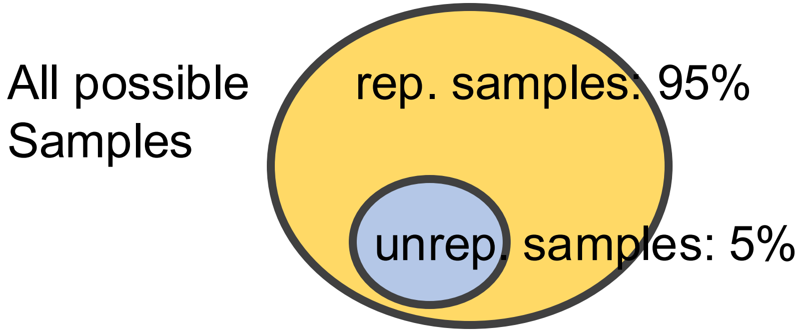
Solution:

How to choose c for other α?

- We choose $\alpha = 0.05$, $Z = \frac{\sqrt{n}(\bar{X}_n - 168)}{8} = 3, P_{H_0}(|Z| \geq 1.96) \leq 0.05$
- Reject H_0: my data would occur with probability ≤ 5% under H_0



Hypothesis testing



- How to choose c?
- Significance level α:

$$P_{H_0}(|Z| \geq c) \leq \alpha$$

Type-I error: we reject H_0 (due to $\bar{X}_n$ far from 168), but H_0 is true

- Usually α is small, e.g. 0.05
- I.e., stay with the null hypothesis as long as our sample is 95%-representative

Smaller α => more inclined to stay with H_0 => Need stronger evidence to reject H_0

Hypothesis testing

Choose c such that

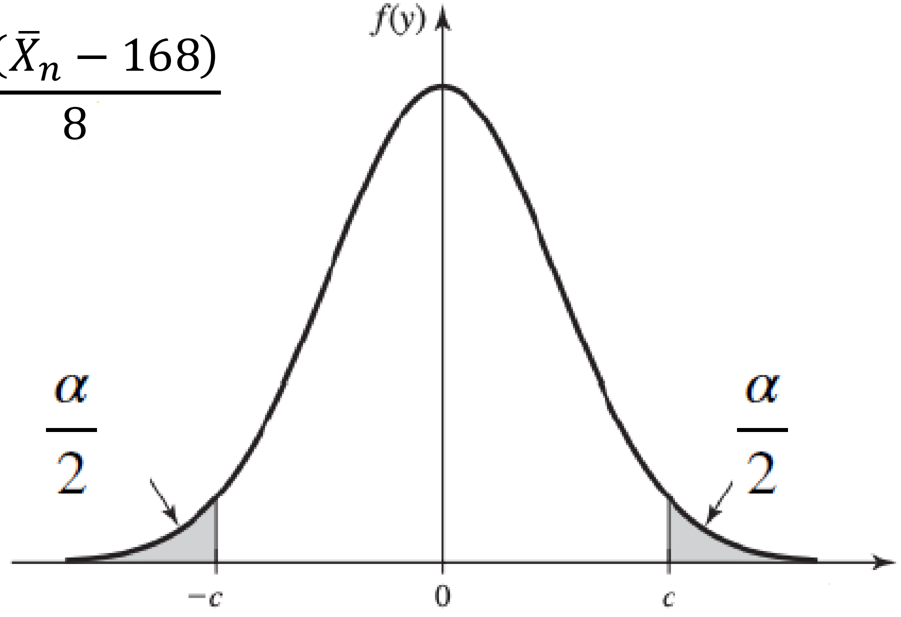
$$P_{H_0}(|Z| \geq c) = \alpha = 5\%$$

Reject H_0 if $|Z| \geq c$, i.e.

Z falls in the shaded region

Let's find the value of c..

$$Z = \frac{\sqrt{n}(\bar{X}_n - 168)}{8}$$



PDF of Z

Hypothesis testing

- under H_0, by central limit theorem:

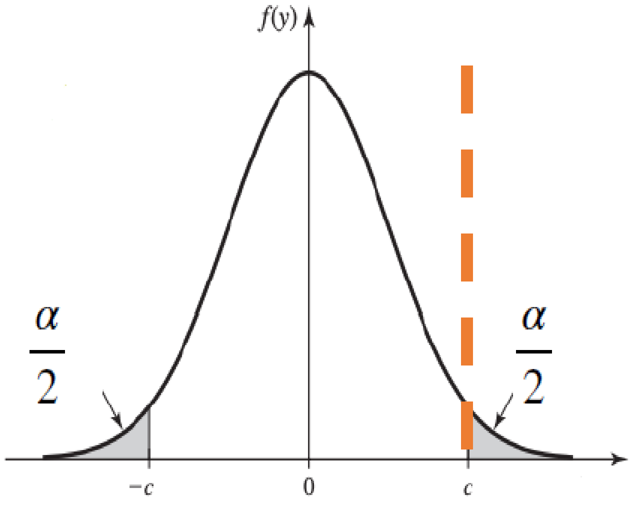
$$Z = \frac{\sqrt{n}(\bar{X}_n - 168)}{8} \sim N(0,1)$$

z-statistic: a statistic that is supposed to follow $N(0,1)$

Z is a valid z-statistic

c is such that $P_{Z\sim N(0,1)}(|Z| \geq c) = \alpha$

$$\Rightarrow c = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$



PDF of Z

$$\alpha = 0.05 \Rightarrow c = \Phi^{-1}(0.975) = \text{st.norm.ppf}(0.975) = 1.96$$

z-test

Example Assume that UA students' heights (in centimeters) follow $N(\mu, 8^2)$, test the hypothesis

$$H_0\colon \mu = 168, \qquad\qquad H_1\colon \mu \neq 168$$

Suppose we observe 4 students' heights: 173, 181, 189, 177

We reject if $Z = \frac{\sqrt{n}}{8}|\bar{X}_n - 168| \geq \Phi^{-1}(0.975)$ This is called a z-test

1.96

From data, $Z = 3$, so we reject H_0.

z-test: general case

General fact Assume that we have a set of samples $X_1, \ldots, X_n$ that follow $N(\mu, \sigma^2)$, test the hypothesis

$$H_0\colon \mu = \mu_0, \qquad H_1\colon \mu \neq \mu_0$$

with significance level α

We can use the z-test:

$$\text{Reject if } |Z| \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right), \textit{ where } Z = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu_0)$$

rejection threshold r

Larger *n* $\Rightarrow$ more reject

Larger $\alpha \Rightarrow$ more reject

Larger $\sigma \Rightarrow$ less reject

Hypothesis testing

General fact Assume that we have a set of samples $X_1, \ldots, X_n$ that follow $N(\mu, \sigma^2)$, test the hypothesis

$$H_0\colon \mu = \mu_0, \qquad H_1\colon \mu \neq \mu_0$$

with significance level α

$$\text{z-test: Reject if } |Z| \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

rejection threshold r

Example $\sigma = 8$, $n = 4$, $\bar{X}_n = 180$, use z-test to test if $\mu = 168$

$$Z = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu_0) = 3$$

$\alpha = 0.05 \Rightarrow r = 1.96$ reject H_0

$\alpha = 0.01 \Rightarrow r = 2.58$ reject H_0

$\alpha = 0.001 \Rightarrow r = 3.29$ do not reject H_0

Other z-tests

- Other tests can be found using the same reasoning

Reject H_0 if:

- $H_0\colon \mu = \mu_0$, vs $H_1\colon \mu \neq \mu_0$

$|Z| \geq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ $\quad Z = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu_0)$

one-sided hypothesis testing problem

- $H_0\colon \mu \leq \mu_0$, vs $H_1\colon \mu > \mu_0$

$Z \geq \Phi^{-1}(1 - \alpha)$

- $H_0\colon \mu \geq \mu_0$, vs $H_1\colon \mu < \mu_0$

$Z \leq \Phi^{-1}(\alpha)$

All these are z-tests, since it uses the z-statistic $Z = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu_0)$

t-test

- Drawback of z-test: needs to know population stddev σ

Example Suppose the #of medical inpatient days in nursing homes follow a distribution with mean μ and variance σ^2. We'd like to perform hypothesis test between:

$$H_0\colon \mu = 200, \qquad H_1\colon \mu \neq 200$$

and we observe $n = 18$ samples with $\bar{X}_n = 182.17$ and $\hat{\sigma}_n = 17.72$

Should I reject H_0?

t-test

Example Suppose the #of medical inpatient days in nursing homes follow a distribution with mean μ and variance σ^2. We'd like to perform hypothesis test between:

$$H_0\colon \mu = 200, \qquad\qquad H_1\colon \mu \neq 200$$

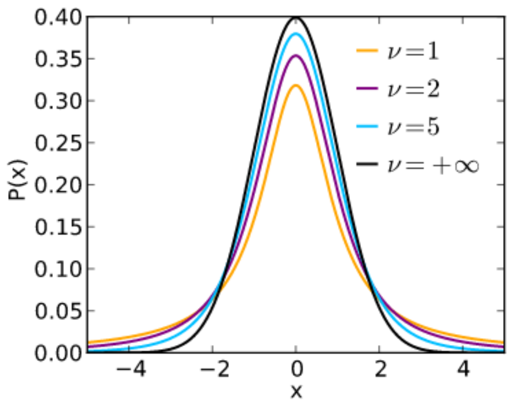
and we observe $n = 18$ samples with $\bar{X}_n = 182.17$ and $\hat{\sigma}_n = 17.72$

Approach When H_0 happens,

this is called a t-statistic, i.e, a statistic that follows t-distribution

$$T = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\hat{\sigma}_n} \sim t(n-1)$$

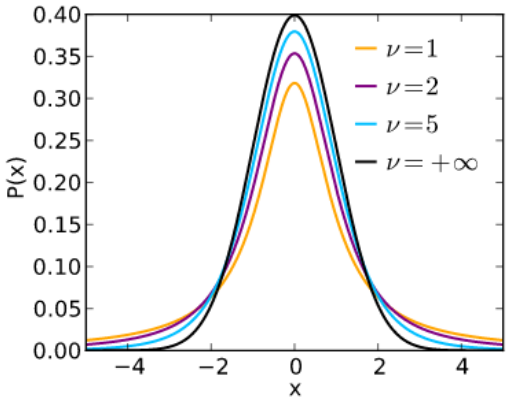
observed value $\frac{\sqrt{18}(182.17-200)}{17.72} = -1.018$



t-test

Approach We've seen that under H_0,

$$T = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\hat{\sigma}_n} \sim t(n-1)$$



Our test with significance α:

$$\text{reject when } |T| > F^{-1}\left(1 - \frac{\alpha}{2}\right)$$

F is now the CDF of the $t(n-1)$ distribution

(Note how similar this is to the z-test)

$|T| = 1.018$

$F^{-1}\left(1 - \frac{\alpha}{2}\right) = 2.11$

```
1 st.t.ppf(1-0.05/2, 17)
```

```
np.float64(2.1098155778331806)
```

thus, we do not reject H_0: $\mu = 200$

Hypothesis testing

- More specifically:
  - Design experiment
  - Design test statistic W (related to hypothesis)
    - T-statistic, Z-statistic, Chi-square statistic
  - Find distribution of W under H_0
  - Collect data $X_1, \dots, X_n$
  - Compute w, value of W applied on the data $X_1, \dots, X_n$
  - Define a rejection region R
  - Reject H_0 if $w \in R$, for "reasonable" rejection region R

Other t-tests

- Other tests can be found using the same reasoning

Reject H_0 if:

- H_0: $\mu = \mu_0$, vs H_1: $\mu \neq \mu_0$ — $|T| \geq F^{-1}\left(1 - \frac{\alpha}{2}\right)$ — $T = \frac{\sqrt{n}}{\hat{\sigma}_n}(\bar{X}_n - \mu_0)$

F: CDF of $t(n-1)$

- H_0: $\mu \leq \mu_0$, vs H_1: $\mu > \mu_0$ — $T \geq F^{-1}(1 - \alpha)$

- H_0: $\mu \geq \mu_0$, vs H_1: $\mu < \mu_0$ — $T \leq F^{-1}(\alpha)$

All these are called t-test, since it relies on computing T, a t-statistic

t-test: a practice final question

Example Metal fibers produced, length in millimeters; use t-test to test

$$H_0\colon \mu \leq 5.2, \qquad H_1\colon \mu > 5.2$$

n=15 fibers measured, $\bar{X}_n = 5.4$, $\hat{\sigma}_n = 0.4226$.

Shall we reject H_0 at significance 0.05?

Solution The t-test is "reject if $T \geq F^{-1}(1-\alpha)$"

t-statistic $T = \frac{\sqrt{n}}{\hat{\sigma}_n}(\bar{X}_n - \mu_0) = 1.83$

rejection threshold $F^{-1}(1-\alpha) =$ t.ppf(0.95, 14) = 1.76 we should reject