Inspire…Educate…Transform.

Inspire…Educate…Transform.

Stat Skills

Confidence Intervals, t-Distribution, Hypothesis Testing, t-Tests

Dr. Anand Jayaraman

Dec 25, 2016

Thanks to Dr.Sridhar Pappu for the material

Merry Christmas

CSE 7315c

Review

- Probability Distributions
  - Bernoulli, Geometric, Binomial, Poisson, Exponential
- Gaussian Distribution
  - Areas under the curve
  - Reading the normal distribution table
  - Approximating B(n,p) with Normal distribution
  - Continuity correction
- Central Limit Theorem

Activity – R

According to the US Bureau of the Census, about 75% of the commuters in the United States drive to work alone. Suppose 150 US commuters are randomly sampled.

- What is the probability that fewer than 105 commuters drive to work alone?
- What is the probability that between 110 and 120 (inclusive) commuters drive to work alone?
- What is the probability that more than 95 commuters drive to work alone?

Answers: 0.0657, 0.6485, 0.9993

- Expected Mean = 0.75*150=112.5
- Variance = n*p*q= 150*0.75*0.25

1. Area under the curve from –Infinity to 104.5 (continuity correction)

```
> pnorm(104.5,112.5,sqrt(150*0.75*0.25))
[1] 0.06571401
```

2. Area under the curve between 120.5 and 109.5

```
> pnorm(120.5,150*0.75,sqrt(150*0.75*0.25)) -  pnorm(109.5,150*0.75,sqrt(150*0.75*0.25))
[1] 0.6484822
```

3. Area under the curve above 95.5

```
> 1-pnorm(95.5,112.5,sqrt(150*0.75*0.25))
[1] 0.999326
```

Activity – R

According to National Center for Health Statistics of the US, the distribution of serum cholesterol levels for 20-74 year old males has a mean of 211mg/dl with a standard deviation of 46mg/dl.

- What is the probability that the serum cholesterol level of a male is >230mg/dl?
- What is the probability that the average serum cholesterol level of a random sample of 25 males will be >230mg/dl?

Answer: 34.0%, 1.9%


TOTAL RECALL
Spigot

" I got the instructions from my Statistics Professor. He was 80% confident that the true location of the restaurant was in this neighborhood."

CONFIDENCE INTERVALS

When we use samples to provide population estimates, we cannot be CERTAIN that they will be accurate. There is an amount of uncertainty, which needs to be calculated.

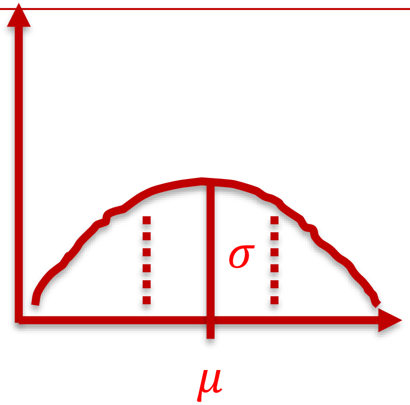
Publish Date	Source	Polling Organisation	NDA	UPA	Other
12 May 2014	[177]	CNN-IBN – CSDS – Lokniti	**276 (±6)**	97 (±5)	148 (±23)
	[177][178]	India Today – Cicero	**272 (±11)**	115 (±5)	156 (±6)
	[177][179]	News 24 – Chanakya	**340 (±14)**	70 (±9)	133 (±11)
	[177]	Times Now – ORG	**249**	148	146
	[177][180]	ABP News – Nielsen	**274**	97	165
	[177]	India TV – CVoter	**289**	101	148
14 May 2014	[181][182]	NDTV – Hansa Research	**279**	103	161
12 May 2014	[177]	Poll of Polls	283	105	149
16 May 2014		Actual Results[2]	**336**	**58**	**149**

Source: http://en.wikipedia.org/wiki/Indian_general_election,_2014
Last accessed: March 27, 2015

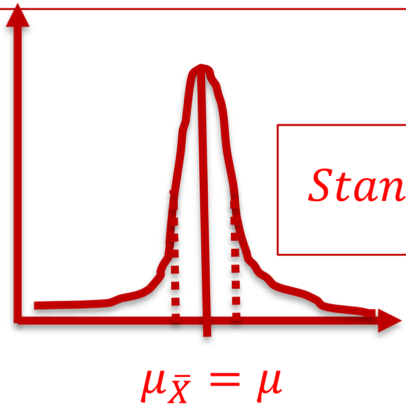
Polling Organisation	NDA	UPA	Other
CNN-IBN – CSDS – Lokniti	**276 (±6)**	97 (±5)	148 (±23)
India Today – Cicero	**272 (±11)**	115 (±5)	156 (±6)
News 24 – Chanakya	**340 (±14)**	70 (±9)	133 (±11)

Incorrect way to present data as it gives the feeling that the population parameter will lie within these ranges.

$Population\ distribution$



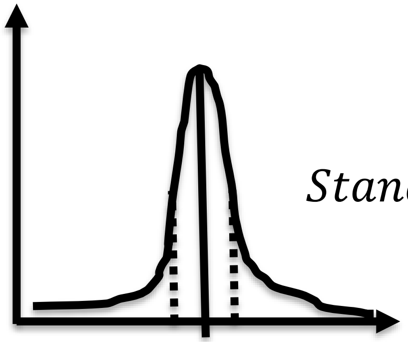
$Sampling\ distribution\ of\ sample\ means$



$$Standard\ Error\ of\ the\ Mean = \frac{\sigma}{\sqrt{n}}$$

Standard Error (SE) is the same as Standard Deviation of the sampling distribution and a sample with 1 SE may or may not include the population parameter.

$$Sampling\ distribution\ of\ sample\ means$$



$$Standard\ Error\ of\ the\ Mean = \frac{\sigma}{\sqrt{n}}$$

$$\mu_{\bar{X}} = \mu$$

We have seen that ~ 95% of the samples will have a mean value within the interval +/- 2 SE of the population mean *(recall the Empirical Rule for Normal Distribution)*.

Alternatively, 95% of such intervals include the population mean.

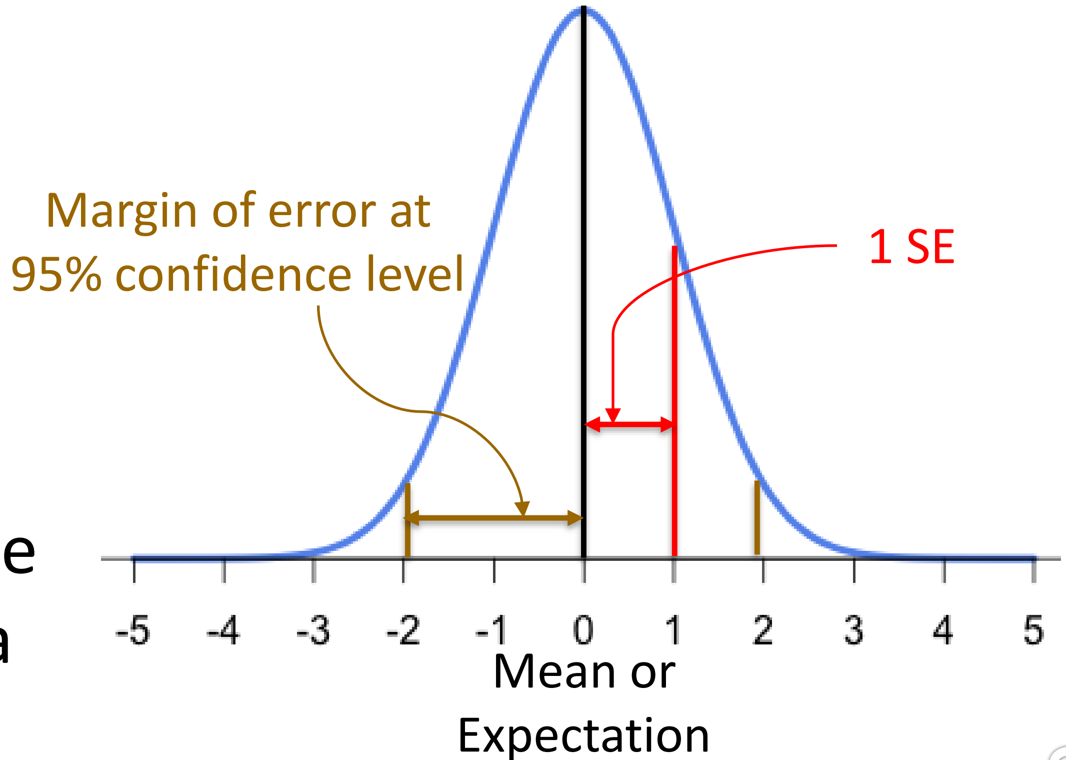
Here, 95% is the Confidence Level and the interval is called the Confidence Interval.

SE, Margin of Error, Confidence Interval and Sample Size

$$SE = \frac{\sigma}{\sqrt{n}}$$

$$Margin\ of\ Error = z * SE$$

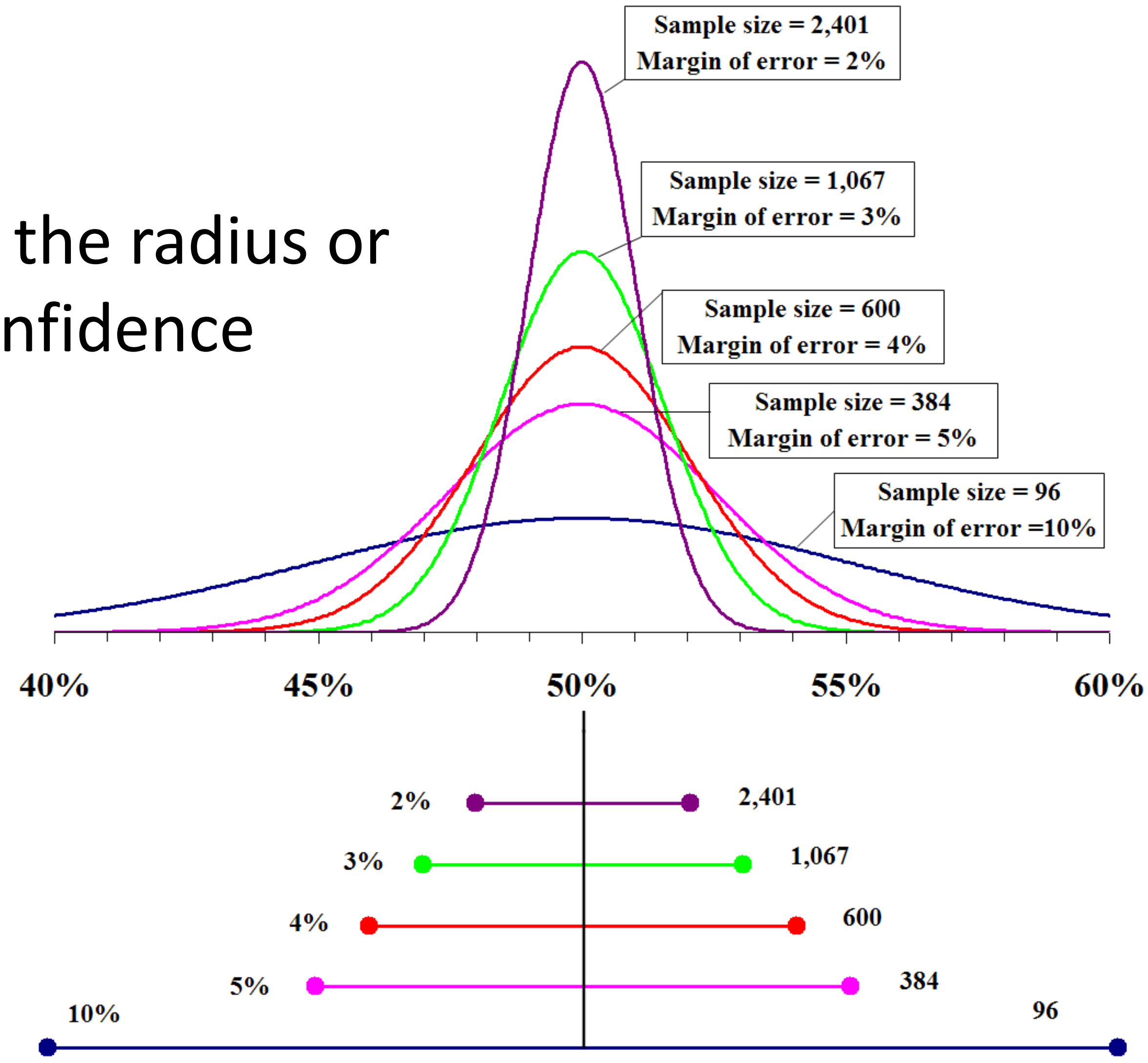
Margin of error is the **maximum expected difference** between the true population parameter and a sample estimate of that parameter.



Margin of error is meaningful only when stated in conjunction with a probability (confidence level).

SE, Margin of Error, Confidence Interval and Sample Size

Margin of error is the radius or half-width of a confidence interval.



Source: https://en.wikipedia.org/wiki/Margin_of_error
Last accessed: June 18, 2015

SE, Margin of Error, Confidence Interval and Sample Size

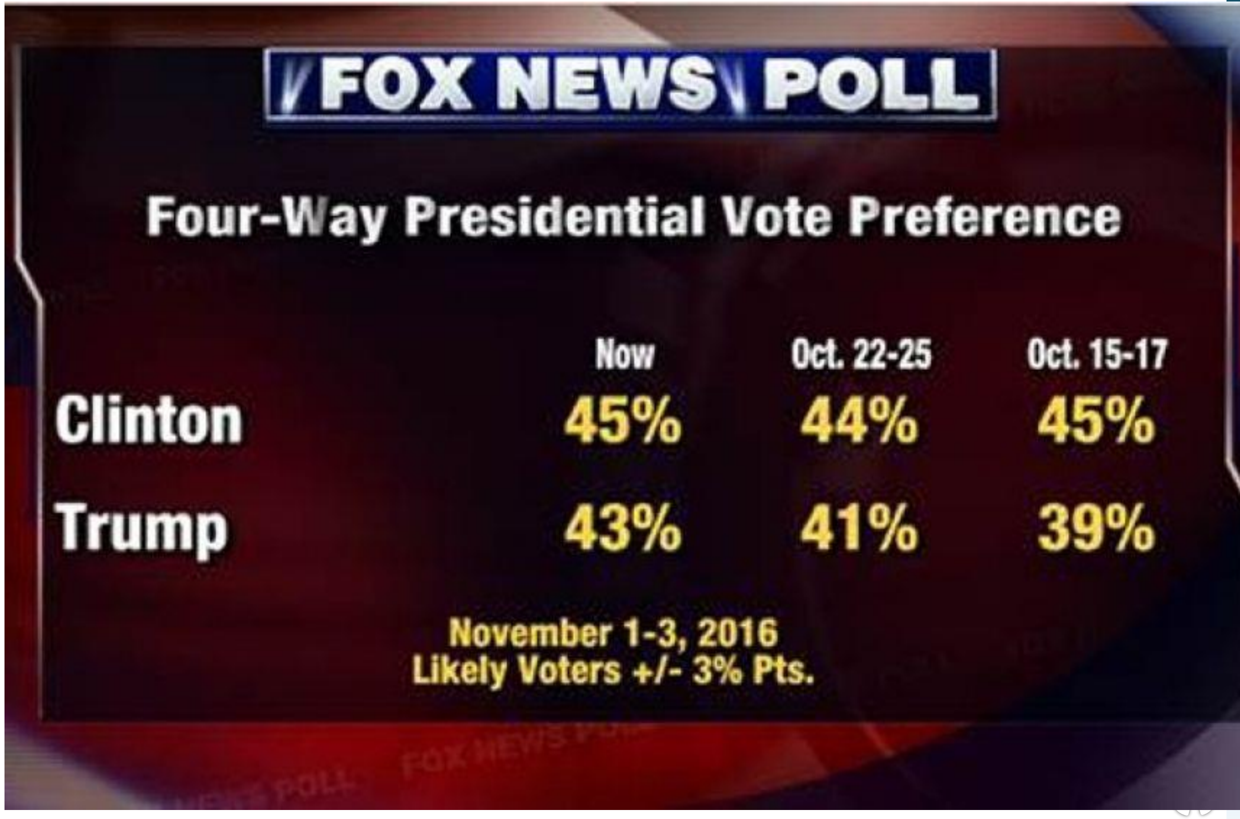
Just like Mean, Proportion is another common parameter of interest in many problems.

Expectation of a sample proportion = p

SE of a sample proportion = $\sqrt{\frac{pq}{n}}$

SE, Margin of Error, Confidence Interval and Sample Size

In a poll by FOX News conducted between November 1 –3 2016, a survey of 1107 randomly sampled likely voters predicted that 45% would vote for Hillary Clinton.



What is the margin of error at 95% confidence level (z = 1.96)?

Check qnorm(0.975, 0, 1). Why 0.975?

$$\text{Margin of error} = 1.96 * \sqrt{\frac{0.45 * 0.55}{1107}} \approx 2.93\%$$

SE, Margin of Error, Confidence Interval and Sample Size

If the desired margin of error at 95% confidence level is 1%, what should be the sample size?

$$0.01 = 1.96 * \sqrt{\frac{0.45 * 0.55}{n}}$$

$$\therefore n = \left(\frac{1.96}{0.01} * \sqrt{0.45 * 0.55}\right)^2 = 9508$$

Confidence Intervals

A survey was taken of US companies that do business with firms in India. One of the survey questions was: Approximately how many years has your company been trading with firms in India?

A random sample of 44 responses to this question yielded a mean of 10.455 years. Suppose the population standard deviation for this question is 7.7 years. Using this information, construct a 90% confidence interval for the mean number of years that a company has been trading in India for the population of US companies trading with firms in India.

Confidence Intervals

- $n = 44$
- $\bar{x} = 10.455$
- $\sigma = 7.7$

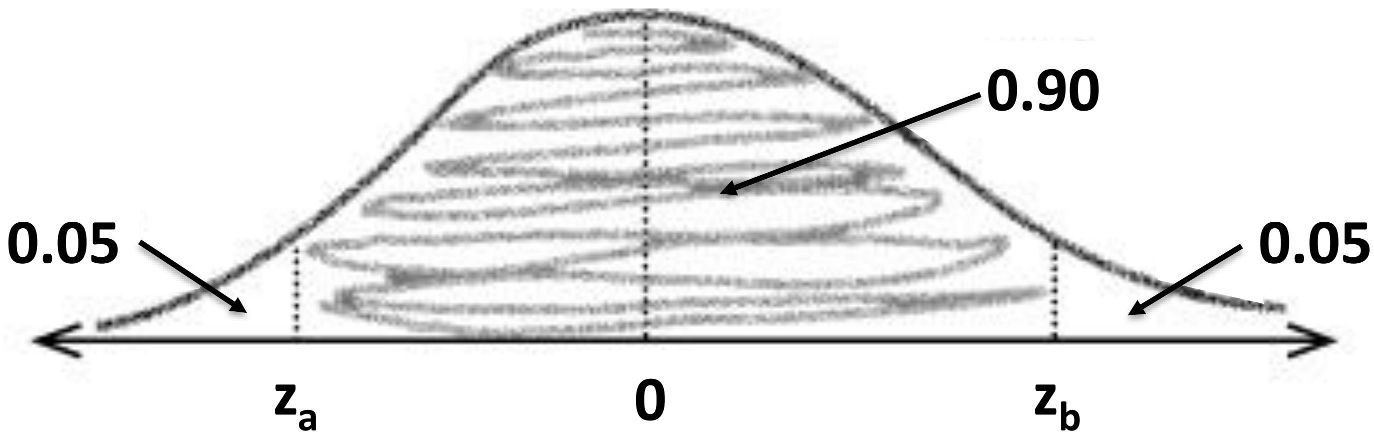
$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ or Margin of error = $z * \frac{\sigma}{\sqrt{n}}$

$\therefore$ Confidence Interval for the Population Mean is

Sample Mean $\pm$ Margin of Error

Confidence Intervals

Find z_a and z_b where $P(z_a < Z < z_b) = 0.90$



$P(Z < z_a) = 0.05$ and $P(Z > z_b) = 0.05$

Confidence Intervals

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

From probability tables using interpolation, we get $z_a = -1.645$ and $z_b = 1.645$.

Check qnorm(0.05, 0, 1) and qnorm(0.95, 0, 1) in R.

Confidence Intervals

$$Margin\ of\ error\ at\ 90\%\ Confidence\ Level = 1.645 * \frac{7.7}{\sqrt{44}} = 1.91$$

Recall Confidence Interval for the Population Mean is Sample Mean $\pm$ Margin of Error

$$\bar{X} - 1.91 < \mu < \bar{X} + 1.91$$

Since the sample mean is 10.455 years, we get the confidence interval for 90% as 8.545 < μ < 12.365.

The analyst is 90% confident that if a census of all US companies trading with firms in India were taken at the time of the survey, the actual population mean number of trading years of such firms would be between 8.545 and 12.365 years.

Shortcuts for Calculating Confidence Intervals

Population Parameter	Population Distribution	Conditions	Confidence Interval
μ	Normal	You know σ^2 n is large or small $\bar{X}$ is the sample mean	$(\bar{X} - z\frac{\sigma}{\sqrt{n}}, \bar{X} + z\frac{\sigma}{\sqrt{n}})$
μ	Non-normal	You know σ^2 n is large (> 30) $\bar{X}$ is the sample mean	$(\bar{X} - z\frac{\sigma}{\sqrt{n}}, \bar{X} + z\frac{\sigma}{\sqrt{n}})$
μ	Normal or Non-normal	You don't know σ^2 n is large (> 30) $\bar{X}$ is the sample mean s^2 is the sample variance	$(\bar{X} - z\frac{s}{\sqrt{n}}, \bar{X} + z\frac{s}{\sqrt{n}})$
p	Binomial	n is large p_s is the sample proportion q_s is $1 - p_s$	$(p_s - z\sqrt{\frac{p_s q_s}{n}}, p_s + z\sqrt{\frac{p_s q_s}{n}})$

Shortcuts for Calculating Confidence Intervals

Level of Confidence	Value of z
90%	1.64
95%	1.96
99%	2.58

You took a sample of 50 Gems and found that in the sample, the proportion of red Gems is 0.25. Construct a 99% confidence interval for the proportion of red Gems in the population.

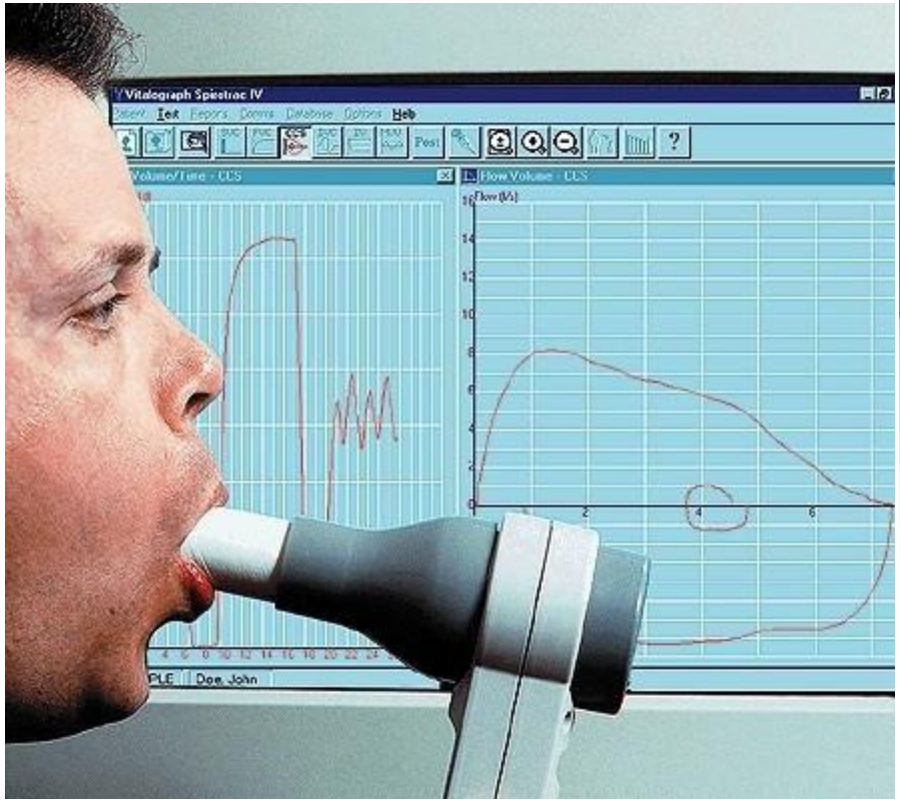
$$0.25 - 2.58 * \sqrt{\frac{0.25 * 0.75}{50}} < p < 0.25 + 2.58 * \sqrt{\frac{0.25 * 0.75}{50}}$$

$$0.09 < p < 0.41$$

Shortcuts for Calculating Confidence Intervals

Level of confidence	Value of z
90%	1.64
95%	1.96
99%	2.58

The lung function in 57 people is tested using FEV1 (Forced Expiratory Volume in 1 Second) measurements. The mean FEV1 value for this sample is 4.062 litres and standard deviation, *s* is 0.67 litres. Construct the 95% Confidence Interval.



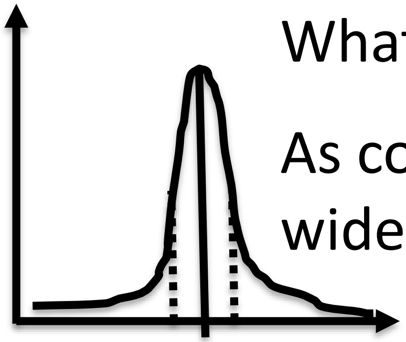
FEV1 values of 57 male medical students

Level of confidence	Value of z
90%	1.64
95%	1.96
99%	2.58

2.85	2.85	2.98	3.04	3.10	3.10	3.19	3.20	3.30	3.39
3.42	3.48	3.50	3.54	3.54	3.57	3.60	3.60	3.69	3.70
3.70	3.75	3.78	3.83	3.90	3.96	4.05	4.08	4.10	4.14
4.14	4.16	4.20	4.20	4.30	4.30	4.32	4.44	4.47	4.47
4.47	4.50	4.50	4.56	4.68	4.70	4.71	4.78	4.80	4.80
4.90	5.00	5.10	5.10	5.20	5.30	5.43			

$$95\%\ CI: \left(4.062 - 1.96 * \frac{0.67}{\sqrt{57}}, 4.062 + 1.96 * \frac{0.67}{\sqrt{57}}\right)$$
$$= (3.89, 4.23)$$

Attention Check



What happens to confidence interval as confidence level changes?

As confidence level increases, the confidence interval becomes wider and *vice-versa*.

What happens to the confidence interval as sample size changes?

As sample size increases, the confidence interval becomes narrower.

Remember $(\bar{X} - z\frac{\sigma}{\sqrt{n}}, \bar{X} + z\frac{\sigma}{\sqrt{n}})$.

Confidence Intervals for a Sample Median

Confidence limits are given by actual values in the sample using the following formulae:

Lower 95% Confidence Limit: $\frac{n}{2} - 1.96 * \frac{\sqrt{n}}{2}$ ranked value.

Upper 95% Confidence Limit: $1 + \frac{n}{2} + 1.96 * \frac{\sqrt{n}}{2}$ ranked value.

Confidence Intervals for a Sample Median

2.85	2.85	2.98	3.04	3.10	3.10	3.19	3.20	3.30	3.39
3.42	3.48	3.50	3.54	3.54	3.57	3.60	3.60	3.69	3.70
3.70	3.75	3.78	3.83	3.90	3.96	4.05	4.08	4.10	4.14
4.14	4.16	4.20	4.20	4.30	4.30	4.32	4.44	4.47	4.47
4.47	4.50	4.50	4.56	4.68	4.70	4.71	4.78	4.80	4.80
4.90	5.00	5.10	5.10	5.20	5.30	5.43			

Lower 95% CL

Median

Upper 95% CL

Lower 95% Confidence Limit: $\frac{57}{2} - 1.96 * \frac{\sqrt{57}}{2} = 21.10$ ranked value. 21st ranked value is 3.70.

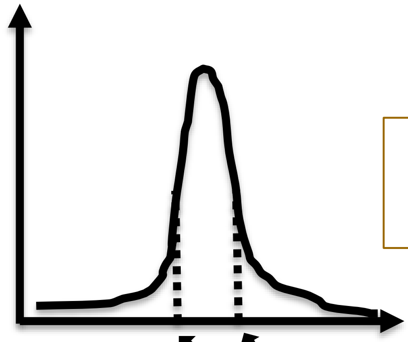
Upper 95% Confidence Limit: $1 + \frac{57}{2} + 1.96 * \frac{\sqrt{57}}{2} = 36.90$ ranked value. 37th ranked value is 4.32.

95% CI: (3.70,4.32) *Recall 95% CI using Mean: (3.89,4.23)*

Confidence Intervals for a Sample Median

- Lack of distributional assumptions makes it difficult to obtain an exact CI for the median.
- CI are not necessarily symmetric around the sample estimate.

The Summary of CI



Confidence Interval = Sample statistic $\pm$ Margin of Error

$$\left(\bar{X} - z\frac{\sigma}{\sqrt{n}}, \bar{X} + z\frac{\sigma}{\sqrt{n}}\right)$$

Margin of error = z * Standard Error *(Recall the standardization formula)*

$$\frac{\sigma}{\sqrt{n}}$$

Depends on the confidence level

Probability density.
Area under the curve between the limits.
Probability that a certain % of samples will contain the population mean within this interval.

Standard deviation of the population: Measure of deviation from the mean.

A short detour – Variance Formula

- Population Parameter

$$\mu = \frac{\Sigma x}{N}$$

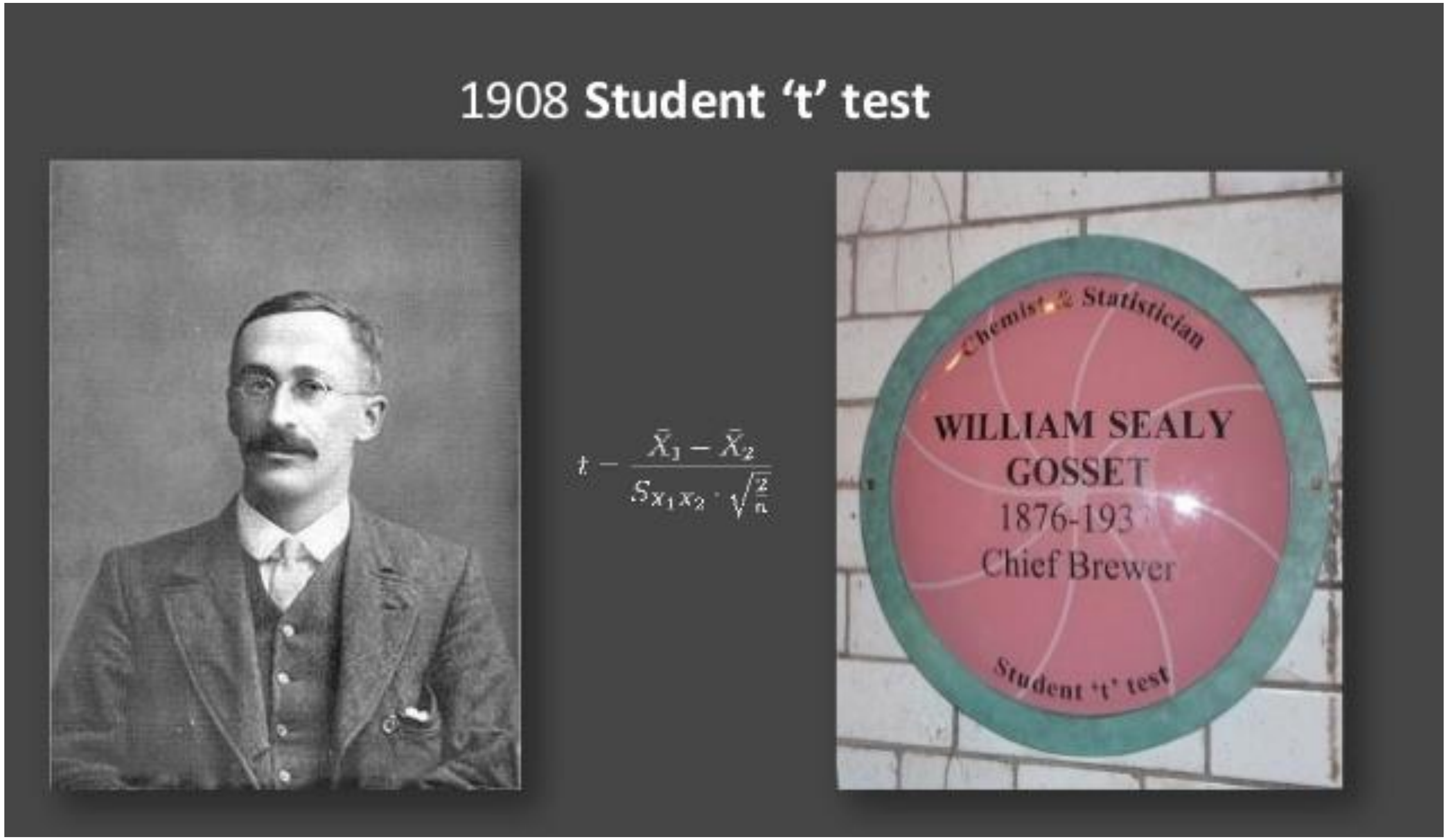
$$\text{Variance } \sigma^2 = \frac{\Sigma(x-\mu)^2}{N}$$

- Sample Statistic

$$\bar{x} = \frac{\Sigma x}{n}$$

$$\text{Variance } s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}$$

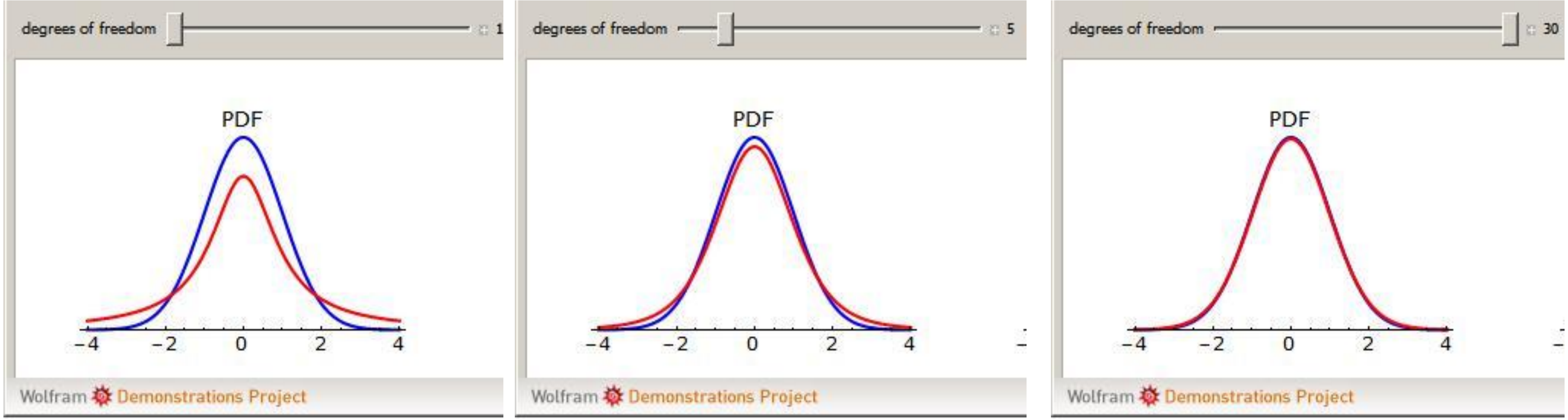
t-Distribution



Ref: http://image.slidesharecdn.com/2013-ingenious-ireland-theingeniousirishiet-slideshow-130524065705-phpapp01/95/2013-ingeniousirelandthe-ingenious-irishietslideshow-43-638.jpg?cb=1369825611
Last accessed: October 31, 2015

t-Distribution

If the sample size is small (<30), the variance of the population is not adequately captured by the variance of the sample. Instead of z-distribution, t-distribution is used. It is also the appropriate distribution to be used when population variance is **not known**, irrespective of sample size.



Ref: http://demonstrations.wolfram.com/ComparingNormalAndStudentsTDistributions/
Last accessed: October 31, 2015

t-Distribution

$$t\ statistic\ (or\ t\ score), t = \frac{(\bar{x} - \mu)}{\frac{s}{\sqrt{n}}}$$

Degrees of freedom, ν: # of independent observations for a source of variation minus the number of independent parameters estimated in computing the variation.*

When estimating mean or proportion from a single sample, the # of independent observations is equal to *n-1*.

* Roger E. Kirk, *Experimental Design: Procedures for the Behavioral Sciences*. Belmont, California: Brooks/Cole, 1968.

Properties of *t*-Distribution

- Mean of the distribution = 0
- Variance = $\frac{\nu}{\nu-2}$, where $\nu > 2$
- Variance is always greater than 1, although it is close to 1 when there are many degrees of freedom (sample size is large)
- With infinite degrees of freedom, *t* distribution is the same as the standard normal distribution

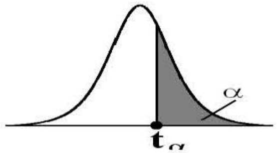
Confidence Interval to Estimate μ

- Population standard deviation UNKNOWN and the population normally distributed.
- $\bar{x} - t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$
  - Sample mean, standard deviation and size can be calculated from the data; *t* value can be read from the table or obtained from software.
  - α is the area in the tail of the distribution. For 90% Confidence Level, α=0.10. In a Confidence Interval, this area is symmetrically distributed between the 2 tails ($\alpha/2$ in each tail).

t-table

Percentage Points of the *t* Distribution; $t_{\nu,\alpha}$

$P(T > t_{\nu,\alpha}) = \alpha$



	α													
ν	**0.40**	**0.30**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.02**	**0.015**	**0.01**	**0.0075**	**0.005**	**0.0025**	**0.0005**
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706	15.895	21.205	31.821	42.434	63.657	127.322	636.590
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303	4.849	5.643	6.965	8.073	9.925	14.089	31.598
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182	3.482	3.896	4.541	5.047	5.841	7.453	12.924
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776	2.999	3.298	3.747	4.088	4.604	5.598	8.610
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571	2.757	3.003	3.365	3.634	4.032	4.773	6.869
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447	2.612	2.829	3.143	3.372	3.707	4.317	5.959
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306	2.449	2.634	2.896	3.085	3.355	3.833	5.041
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262	2.398	2.574	2.821	2.998	3.250	3.690	4.781
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228	2.359	2.527	2.764	2.932	3.169	3.581	4.587
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201	2.328	2.491	2.718	2.879	3.106	3.497	4.437
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179	2.303	2.461	2.681	2.836	3.055	3.428	4.318
13	0.259	0.538	0.870	1.079	1.350	1.771	2.160	2.282	2.436	2.650	2.801	3.012	3.372	4.221
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145	2.264	2.415	2.624	2.771	2.977	3.326	4.140
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131	2.249	2.397	2.602	2.746	2.947	3.286	4.073
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120	2.235	2.382	2.583	2.724	2.921	3.252	4.015
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110	2.224	2.368	2.567	2.706	2.898	3.222	3.965
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101	2.214	2.356	2.552	2.689	2.878	3.197	3.922
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093	2.205	2.346	2.539	2.674	2.861	3.174	3.883
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086	2.197	2.336	2.528	2.661	2.845	3.153	3.850
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080	2.189	2.328	2.518	2.649	2.831	3.135	3.819
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074	2.183	2.320	2.508	2.639	2.819	3.119	3.792
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069	2.177	2.313	2.500	2.629	2.807	3.104	3.768
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064	2.172	2.307	2.492	2.620	2.797	3.091	3.745
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060	2.167	2.301	2.485	2.612	2.787	3.078	3.725
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052	2.158	2.291	2.473	2.598	2.771	3.057	3.690
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048	2.154	2.286	2.467	2.592	2.763	3.047	3.674
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045	2.150	2.282	2.462	2.586	2.756	3.038	3.659
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042	2.147	2.278	2.457	2.581	2.750	3.030	3.646
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021	2.123	2.250	2.423	2.542	2.704	2.971	3.551
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000	2.099	2.223	2.390	2.504	2.660	2.915	3.460
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980	2.076	2.196	2.358	2.468	2.617	2.860	3.373
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960	2.054	2.170	2.326	2.432	2.576	2.807	3.291

t-Distribution - Example

The labeled potency of a tablet dosage form is 100 mg. As per the quality control specifications, 10 tablets are randomly assayed.

A researcher wants to estimate the interval for the true mean of the batch of tablets with 95% confidence. Assume the potency is normally distributed.

Data are as follows (in mg):

98.6	102.1	100.7	102.0	97.0
103.4	98.9	101.6	102.9	105.2

t-Distribution - Example

Mean, $\bar{x}$ = 101.24 mg

Standard deviation, s = 2.48

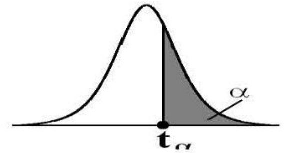
n = 10

$\nu = 10 - 1 = 9$

At 95% level, $\alpha = 0.05$, and $\therefore, \frac{\alpha}{2} = 0.025$

t-table

Percentage Points of the *t* Distribution; $t_{\nu,\alpha}$
$P(T > t_{\nu,\alpha}) = \alpha$



ν	α = 0.40	0.30	0.20	0.15	0.10	0.05	0.025	0.02	0.015	0.01	0.0075	0.005	0.0025	0.0005
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706	15.895	21.205	31.821	42.434	63.657	127.322	636.590
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303	4.849	5.643	6.965	8.073	9.925	14.089	31.598
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182	3.482	3.896	4.541	5.047	5.841	7.453	12.924
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776	2.999	3.298	3.747	4.088	4.604	5.598	8.610
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571	2.757	3.003	3.365	3.634	4.032	4.773	6.869
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447	2.612	2.829	3.143	3.372	3.707	4.317	5.959
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306	2.449	2.634	2.896	3.085	3.355	3.833	5.041
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262	2.398	2.574	2.821	2.998	3.250	3.690	4.781
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228	2.359	2.527	2.764	2.932	3.169	3.581	4.587
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201	2.328	2.491	2.718	2.879	3.106	3.497	4.437
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179	2.303	2.461	2.681	2.836	3.055	3.428	4.318
13	0.259	0.538	0.870	1.079	1.350	1.771	2.160	2.282	2.436	2.650	2.801	3.012	3.372	4.221
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145	2.264	2.415	2.624	2.771	2.977	3.326	4.140
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131	2.249	2.397	2.602	2.746	2.947	3.286	4.073
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120	2.235	2.382	2.583	2.724	2.921	3.252	4.015
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110	2.224	2.368	2.567	2.706	2.898	3.222	3.965
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101	2.214	2.356	2.552	2.689	2.878	3.197	3.922
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093	2.205	2.346	2.539	2.674	2.861	3.174	3.883
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086	2.197	2.336	2.528	2.661	2.845	3.153	3.850
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080	2.189	2.328	2.518	2.649	2.831	3.135	3.819
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074	2.183	2.320	2.508	2.639	2.819	3.119	3.792
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069	2.177	2.313	2.500	2.629	2.807	3.104	3.768
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064	2.172	2.307	2.492	2.620	2.797	3.091	3.745
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060	2.167	2.301	2.485	2.612	2.787	3.078	3.725
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052	2.158	2.291	2.473	2.598	2.771	3.057	3.690
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048	2.154	2.286	2.467	2.592	2.763	3.047	3.674
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045	2.150	2.282	2.462	2.586	2.756	3.038	3.659
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042	2.147	2.278	2.457	2.581	2.750	3.030	3.646
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021	2.123	2.250	2.423	2.542	2.704	2.971	3.551
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000	2.099	2.223	2.390	2.504	2.660	2.915	3.460
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980	2.076	2.196	2.358	2.468	2.617	2.860	3.373
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960	2.054	2.170	2.326	2.432	2.576	2.807	3.291

$t_{9,0.025} = 2.262$

t-Distribution - Example

Mean, $\bar{x}$ = 101.24 mg, Standard deviation, s = 2.48

n = 10, $\nu = 10 - 1 = 9$

$$\bar{x} - t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

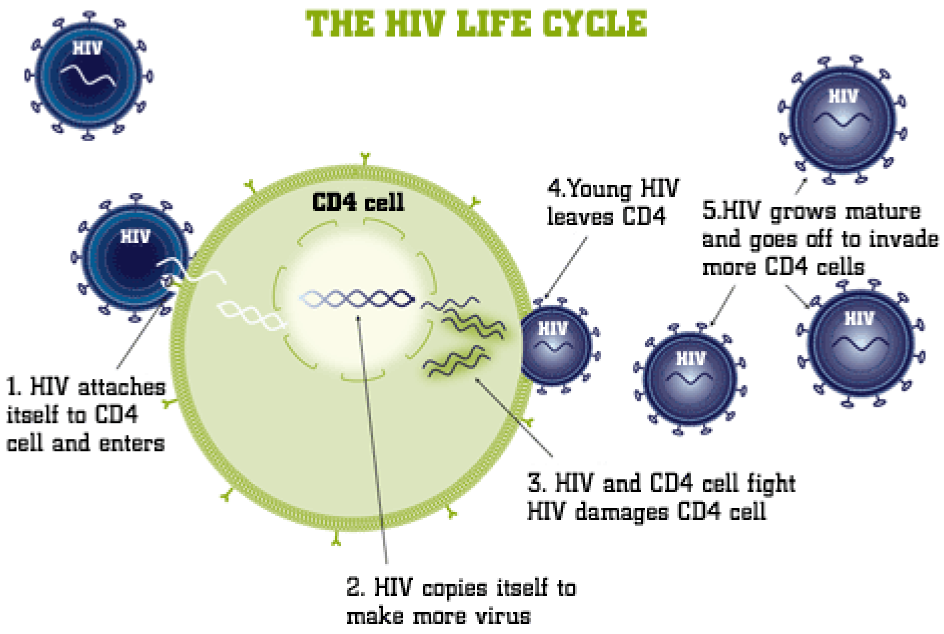
$$101.24 - 2.262 * \frac{2.48}{\sqrt{10}} \leq \mu \leq 101.24 + 2.262 * \frac{2.48}{\sqrt{10}}$$

$99.47 \leq \mu \leq 103.01$

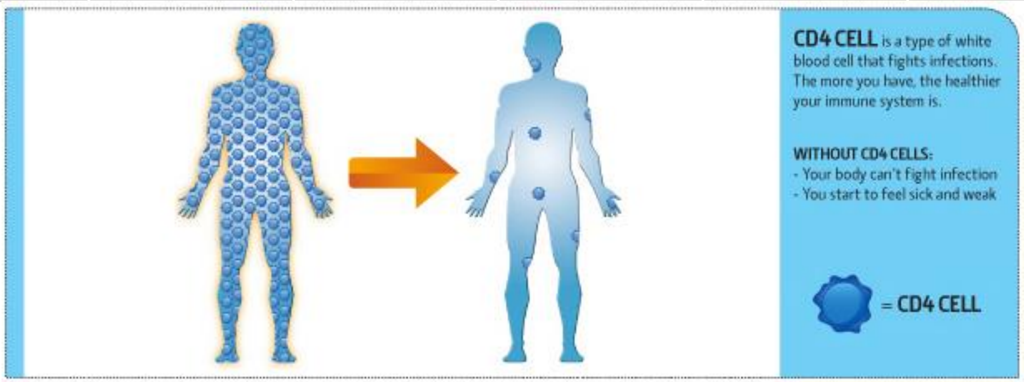
The batch mean is 101.24 mg with an error of +/-1.77 mg. The researcher is 95% confident that the average potency of the batch of tablets is between 99.47 mg and 103.01 mg.

t-Distribution - Example

A researcher wants to examine CD4 counts for HIV+ patients at her clinic. She randomly selects a sample of 25 HIV+ patients and measures their CD4 levels. Calculate a 95% CI for population mean given the following sample results:



Variable	*n*	$\bar{x}$	SE of Mean	*s*	Min	Q1	Median	Q3	Max
CD4 (cells/μl)	25	321.4	14.8	73.8	208.0	261.5	325.0	394.0	449.0



CSE 7315c

t-Distribution - Example

Variable	n	$\overline{x}$	SE of Mean	s	Min	Q1	Median	Q3	Max
CD4 (cells/μl)	25	321.4	14.8	73.8	208.0	261.5	325.0	394.0	449.0

$$CI(0.05):\ \bar{x} - t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

$$321.4 - t_{25-1,\frac{0.05}{2}} \frac{73.8}{\sqrt{25}} \le \mu \le 321.4 + t_{25-1,\frac{0.05}{2}} \frac{73.8}{\sqrt{25}}$$

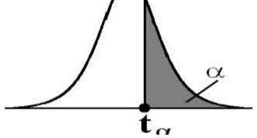
$$t_{24,0.025} = 2.064$$

$$\therefore CI(0.05): [290.85, 351.95]$$

t-table

Percentage Points of the *t* Distribution; $t_{\nu,\alpha}$

$P(T > t_{\nu,\alpha}) = \alpha$



ν	α 0.40	0.30	0.20	0.15	0.10	0.05	0.025	0.02	0.015	0.01	0.0075	0.005	0.0025	0.0005
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706	15.895	21.205	31.821	42.434	63.657	127.322	636.590
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303	4.849	5.643	6.965	8.073	9.925	14.089	31.598
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182	3.482	3.896	4.541	5.047	5.841	7.453	12.924
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776	2.999	3.298	3.747	4.088	4.604	5.598	8.610
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571	2.757	3.003	3.365	3.634	4.032	4.773	6.869
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447	2.612	2.829	3.143	3.372	3.707	4.317	5.959
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306	2.449	2.634	2.896	3.085	3.355	3.833	5.041
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262	2.398	2.574	2.821	2.998	3.250	3.690	4.781
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228	2.359	2.527	2.764	2.932	3.169	3.581	4.587
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201	2.328	2.491	2.718	2.879	3.106	3.497	4.437
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179	2.303	2.461	2.681	2.836	3.055	3.428	4.318
13	0.259	0.538	0.870	1.079	1.350	1.771	2.160	2.282	2.436	2.650	2.801	3.012	3.372	4.221
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145	2.264	2.415	2.624	2.771	2.977	3.326	4.140
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131	2.249	2.397	2.602	2.746	2.947	3.286	4.073
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120	2.235	2.382	2.583	2.724	2.921	3.252	4.015
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110	2.224	2.368	2.567	2.706	2.898	3.222	3.965
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101	2.214	2.356	2.552	2.689	2.878	3.197	3.922
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093	2.205	2.346	2.539	2.674	2.861	3.174	3.883
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086	2.197	2.336	2.528	2.661	2.845	3.153	3.850
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080	2.189	2.328	2.518	2.649	2.831	3.135	3.819
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074	2.183	2.320	2.508	2.639	2.819	3.119	3.792
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069	2.177	2.313	2.500	2.629	2.807	3.104	3.768
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064	2.172	2.307	2.492	2.620	2.797	3.091	3.745
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060	2.167	2.301	2.485	2.612	2.787	3.078	3.725
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052	2.158	2.291	2.473	2.598	2.771	3.057	3.690
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048	2.154	2.286	2.467	2.592	2.763	3.047	3.674
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045	2.150	2.282	2.462	2.586	2.756	3.038	3.659
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042	2.147	2.278	2.457	2.581	2.750	3.030	3.646
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021	2.123	2.250	2.423	2.542	2.704	2.971	3.551
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000	2.099	2.223	2.390	2.504	2.660	2.915	3.460
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980	2.076	2.196	2.358	2.468	2.617	2.860	3.373
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960	2.054	2.170	2.326	2.432	2.576	2.807	3.291

t-Distribution - Example

What does $CI(0.05)$: $[290.85, 351.95]$ mean in the business context given that US Government classifies AIDS under three official categories of CD4 counts:

- Asymptomatic: $\geq$ 500 cells/μl
- AIDS related complex (ARC): 200-499 cells/μl
- AIDS (Stage 3 infection): < 200 cells/μl

Interview Question

If you toss a coin 20 times and get 15 heads, would you say the coin is biased?

Let us apply our learning thus far…

Q. What distribution is it?

A. Binomial; X~B(20,0.5) assuming the coin is fair.

Q. What is the expectation?

A. $np = 10$

Q. What is the standard deviation?

A. $\sqrt{npq} = \sqrt{5} = 2.236$

Q. How many standard deviations away from the mean is 15?

A. $\frac{15-10}{2.236} = 2.236$

Q. What is the probability of getting 15 or more heads?

A. $P(X \geq 15) = P(X = 15) + P(X = 16) + P(X = 17) + P(X = 18) + P(X = 19) + P(X = 20) = 0.021$

Q. Can it be approximated to a normal distribution?

A. $np = 10 \; and \; nq = 10$. Since both are greater than 5, it can be approximated to X~N(10,5)

Q. What is the probability of getting 15 or more heads?

A. $\mathrm{P}(\mathrm{X} > 14.5) = 1 - P(X < 14.5)$

Q. What is the z-score?

"I've narrowed it to two hypotheses: it grew or we shrunk."

HYPOTHESIS TESTS

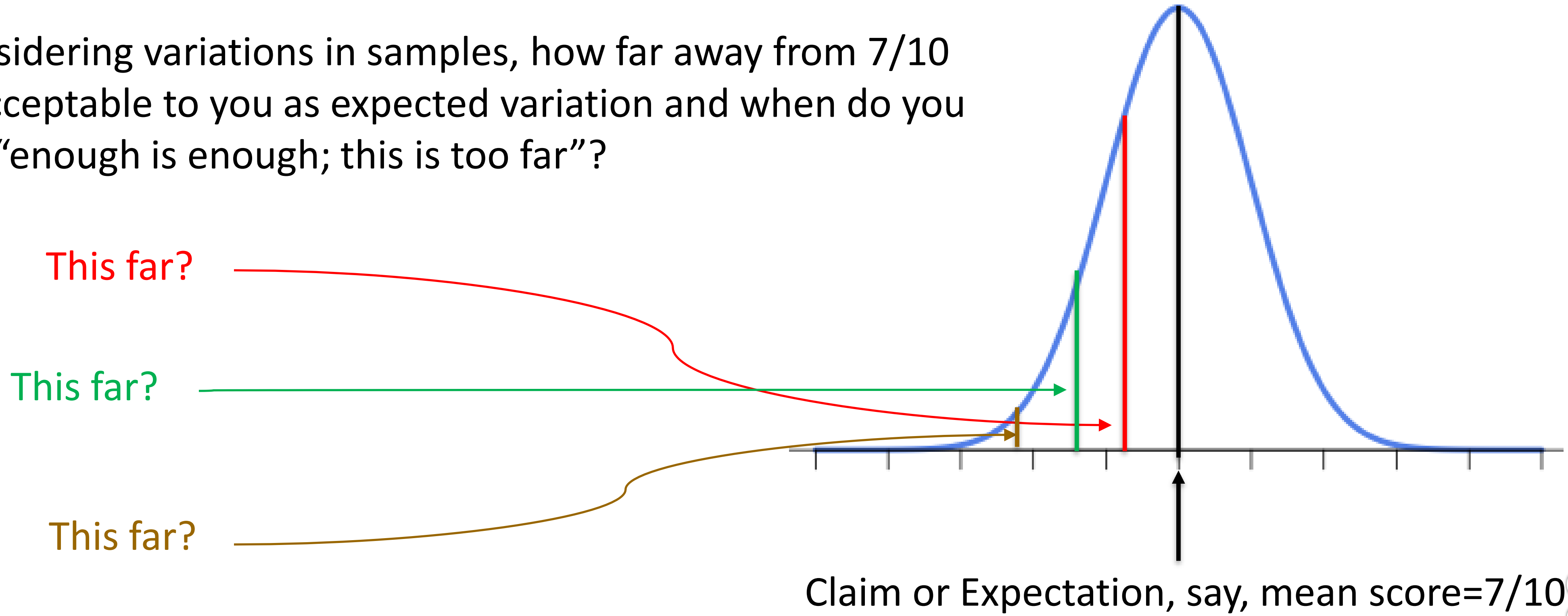
- Hypothesis tests give a way of using samples to test whether or not statistical claims are likely to be true or not.

Hypothesis Testing

- A school principal claims that the students from her school have an average score of 7/10 in a English Proficiency test.
- You doubt that claim and take a random sample of 40 students and you find a mean score of 5.5/10, with a sample standard deviation of 1. Can you reject the principal's claim?

Hypothesis Testing Process

Considering variations in samples, how far away from 7/10 is acceptable to you as expected variation and when do you say "enough is enough; this is too far"?



Step 1: Decide on the hypothesis

Average score on the test is 7/10.

This is called Null Hypothesis and is represented by H_0.

In this case, H_0: $\mu = 0.7$

If Null Hypothesis is rejected based on evidence, an Alternate Hypothesis, H_1, needs to be accepted. **We always start with the assumption that Null Hypothesis is true.**

In this case, H_1: $\mu < 0.7$

Examples of Hypotheses

- Two hypotheses in competition:
  - H_0: The NULL hypothesis, usually the most conservative.
  - H_1 or H_A: The ALTERNATIVE hypothesis, the one we are actually interested in.
- Examples of NULL Hypothesis:
  - The coin is fair
  - The new drug is no better (or worse) than the placebo
- Examples of ALTERNATIVE hypothesis:
  - The coin is biased (either towards heads or tails)
  - The coin is biased towards heads
  - The coin has a probability 0.6 of landing on tails
  - The drug is better than the placebo

Step 2: Choose your statistic

Sample size = 40

Normal distribution is a good approximation

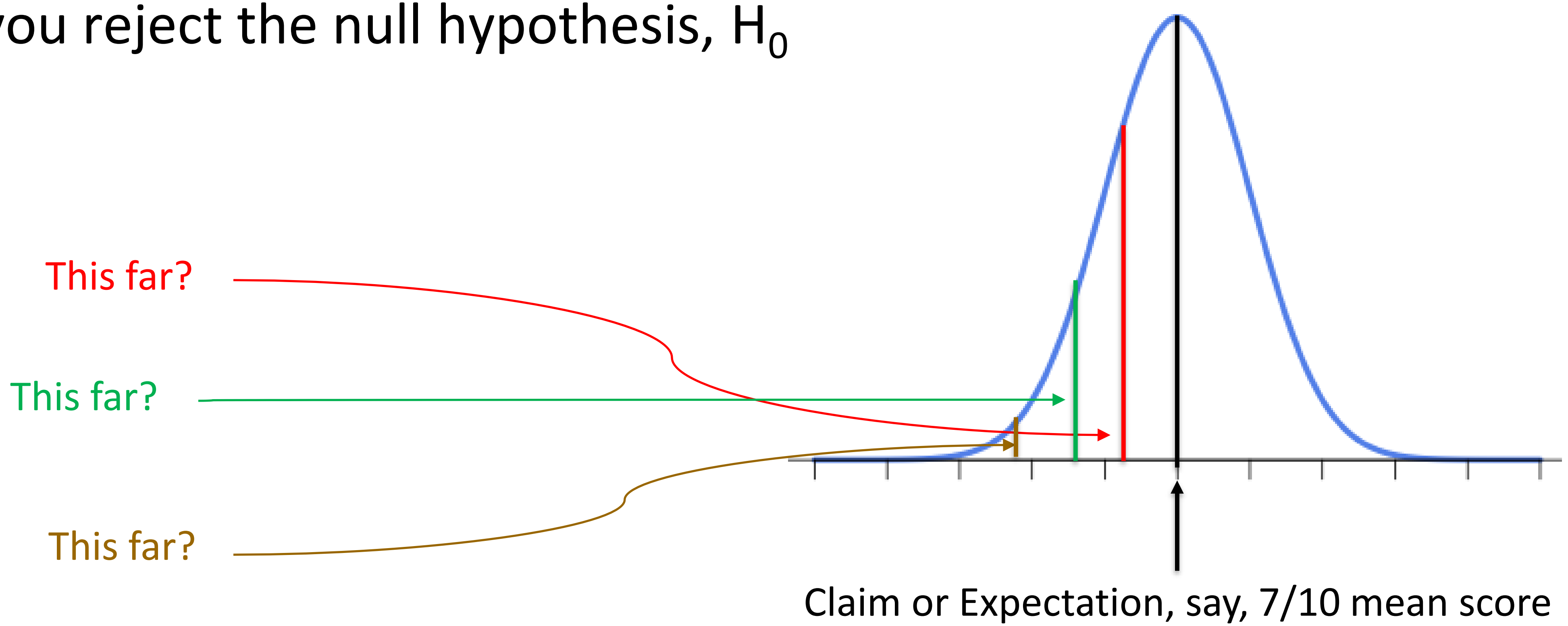
Std Err = $\frac{s}{\sqrt{n}}$ = $\frac{1.0}{\sqrt{40}}$ = 0.158

X ~ N(0.7, 0.158^2) = N(0.7, 0.025)

$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ = $\frac{0.55-0.7}{0.158}$ = -0.94

Step 3: Specify the Significance Level

First, we must decide on the Significance Level, α. It is a measure of how unlikely you want the results of the sample to be before you reject the null hypothesis, H_0



Step 4: Determine the critical region

If X represents the sample mean score, the critical region is defined as $P(X < c) < \alpha$ where $\alpha = 5\%$.

Recall that in a 95% CI, there is a 5% chance that the sample will not contain the population mean. Hence if the sample falls in the critical region, the null hypothesis that 0.7 is the mean score is rejected.

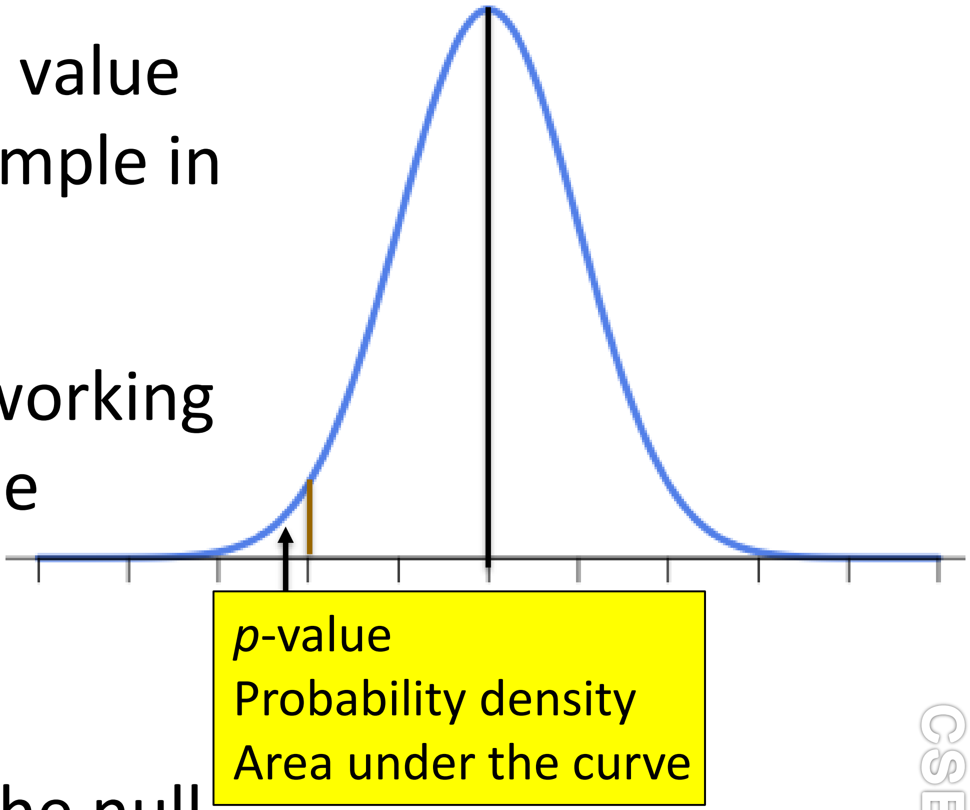
That is the reason 5% or 0.05 is called the Significance Level. In a 99% CI, 0.01 is the Significance Level.

Step 5: Find the *p*-value

p-value is the probability of getting a value up to and including the one in the sample in the direction of the critical region.

It is a way of taking the sample and working out whether the result falls within the critical region of the hypothesis test.

Essentially, this is the value used to determine whether or not to reject the null hypothesis.



Step 5: Find the *p*-value

In our sample, we found a mean score of 5.5/10. This means our *p*-value is $P(X \leq 0.55)$, where X is the distribution of the mean scores in the sample.

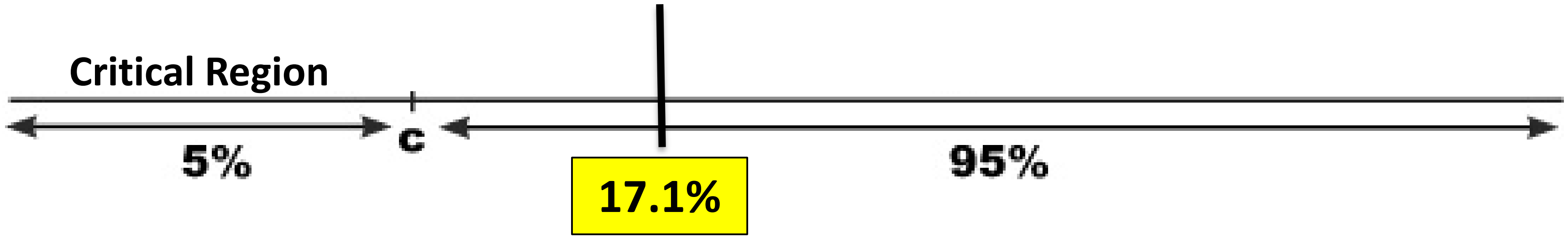
If $P(X \leq 0.55) < 0.05$ (Significance Level), it indicates that 0.55 is inside the critical region, and hence H_0 can be rejected.

Given that $Z = -0.94$, $P(X \leq 0.55) = 0.171$

```
> pnorm(0.55,0.7,1/sqrt(40))
[1] 0.1713909
```

So there is a 17% probability of find a mean score of 5.5/10 or less.

Step 6: Is the sample result in the critical region?



Step 7: Make your decision

There isn't sufficient evidence to reject the null hypothesis and so, the claims of the principal are accepted.

Would your conclusion be any different if the same average score of 5.5/10 was found from a sample of size 400?

What are the null and alternate hypotheses?

H_0: $\mu = 0.7$

H_1: $\mu < 0.7$

What is the test statistic?

$$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{0.55-0.7}{\frac{1}{\sqrt{400}}} = -3$$

p-value = P(Z < -3.0) = 0.00135

CSE 7315c

What is your decision?

Since the *p*-value (0.00135) is less than the Significance Level of 0.05, the null hypothesis can be rejected.

Attention Check

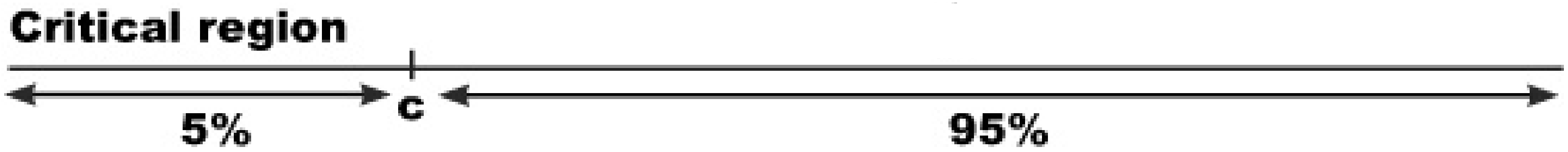
In hypothesis testing, do you assume the null hypothesis to be true or false?

True.

If there is sufficient evidence against the null hypothesis, do you accept it or reject it?

Reject it.

Attention Check



If the *p*-value is less than 0.05 for the above significance level, will you accept or reject the null hypothesis?

Reject it.

Do you need weaker evidence or stronger to reject the null hypothesis if you were testing at the 1% significance level instead of the 5% significance level?

Stronger.

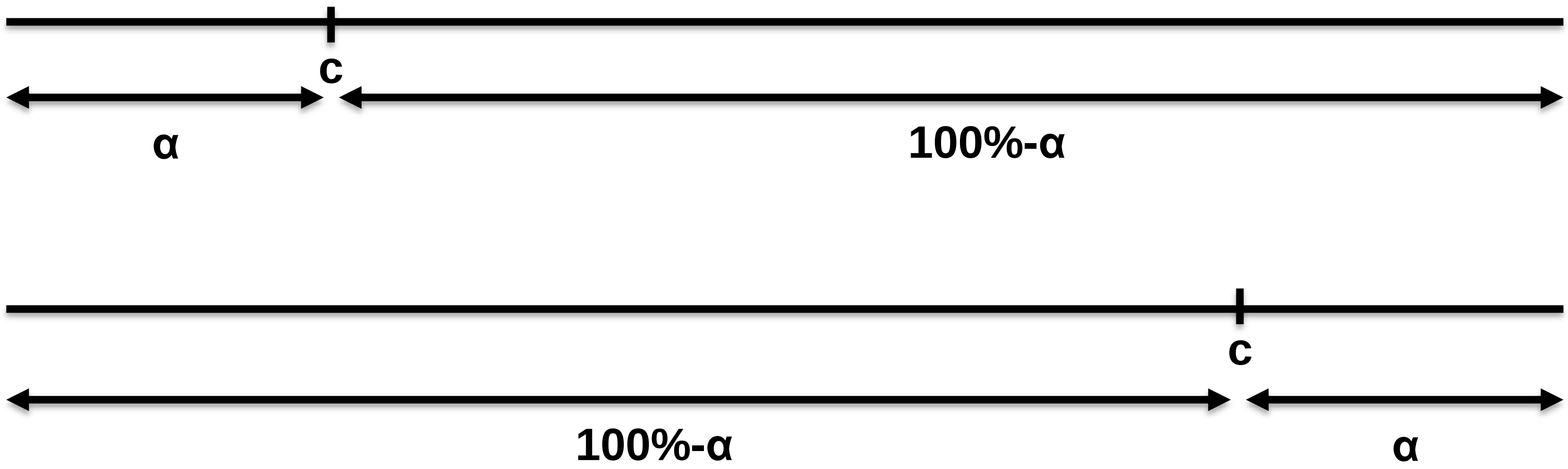
Critical Region Up Close

One-tailed tests

The position of the tail is dependent on H_1.

If H_1 includes a < sign, then the **lower tail** is used.

If H_1 includes a > sign, then the **upper tail** is used.

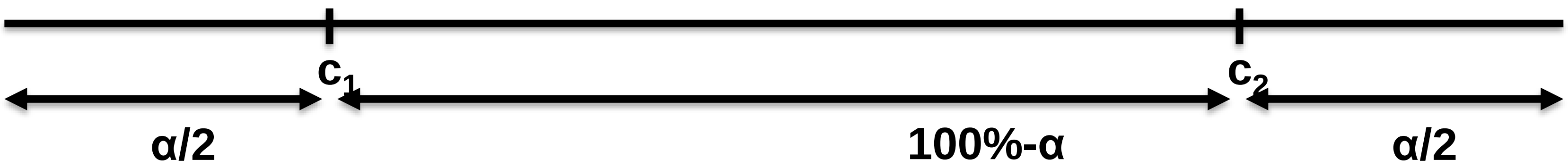


Critical Region Up Close

Two-tailed tests

Critical region is split over both ends. Both ends contain α/2, making a total of α.

If H_1 includes a $\neq$ sign, then the two-tailed test is used as we then look for a change in parameter, rather than an increase or a decrease.



Critical Region Up Close

For each of the scenarios below, identify what type of test you would require.

- Average test score problem as discussed till now.

  One-tailed/Lower-tailed

- If we were checking whether the average is significantly different from 7/10, i.e., H_1: $\mu \neq 0.7$.

  Two-tailed test

- The coin is biased.

  Two-tailed test

- The coin is biased towards heads with probability 0.8.

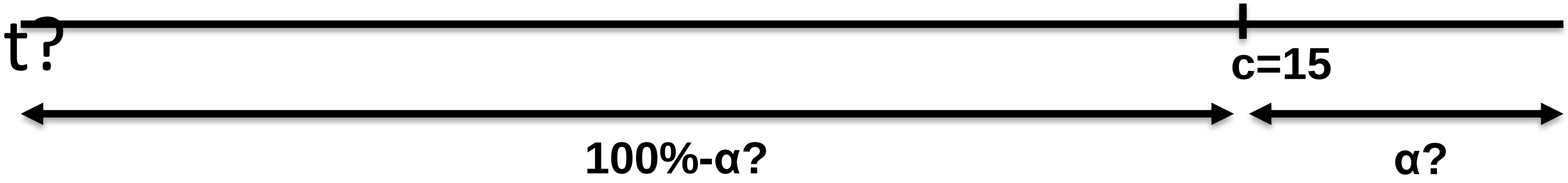
  One-tailed/Upper-tailed

The Missing Link in the Interview

Q. What is the probability of getting 15 or more heads?

A. $P(X \geq 15) = P(X = 15) + P(X = 16) + P(X = 17) + P(X = 18) + P(X = 19) + P(X = 20) = 0.021$

What can you now say about the coin being biased or not?



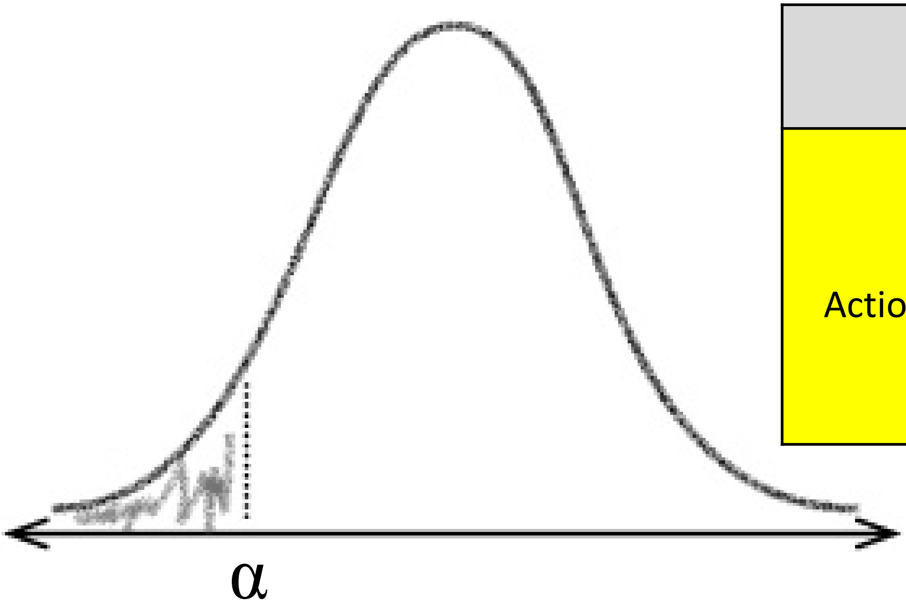
The hypothesis test doesn't answer the question whether the coin is biased or not; it only states whether the evidence is enough to reject the null hypothesis or not ***at the chosen significance level***.

Errors

- Type I: We reject the NULL hypothesis incorrectly
- Type II: We "accept" it incorrectly

		State of Nature	
		Null true	Null false
Action	Fail to reject null (negative)	Correct decision True Negative Specificity P(*Accept* H_0 | H_0 True)	Type II error (β) False Negative P(*Accept* H_0 | H_0 False)
	Reject null (positive)	Type I error (α) False Positive P(Reject H_0 | H_0 True)	Correct decision (Power) True Positive Sensitivity/Recall P(Reject H_0 | H_0 False)

Probability of Getting Type I Error



		State of Nature	
		Null true	Null false
Action	Fail to reject null (negative)	Correct decision True Negative Specificity	Type II error (β) False Negative
	Reject null (positive)	Type I error (α) False Positive	Correct decision (Power) True Positive Sensitivity/Recall

P(Type I error) = α

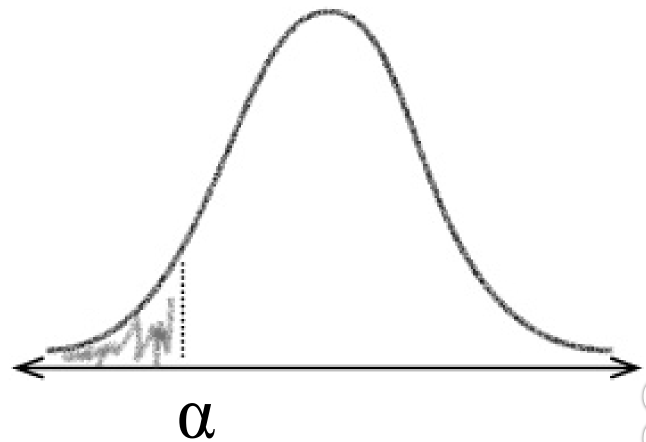
Probability of Getting Type II Error

P(Type II error) = β

To find β,

1. Check that you have a specific value for H_1.
2. Find the range of values outside the critical region of the test. If the test statistic has been standardized, it needs to be de-standardized for the purpose.
3. Find the probability of getting this range of values, assuming H_1 is true. In other words, find the probability of getting the range of values outside the critical region, but this time using the test statistic described by H_1 and not H_0.

		State of Nature	
		Null true	Null false
Action	Fail to reject null (negative)	Correct decision True Negative Specificity	Type II error (β) False Negative
	Reject null (positive)	Type I error (α) False Positive	Correct decision (Power) True Positive Sensitivity/Recall



A new miracle drug claims that it cures common cold and it has had a success rate of 90%. You conduct a random sample test with 100 patients and you find that 80 of them are cured.

At 5% significant level, do you reject or accept the claim by the drug company?

What are the null and alternate hypotheses?

H_0: $p = 0.9$

H_1: $p < 0.9$

What is the test statistic?

$X \sim B(100, 0.9)$

Since np>5 and nq>5, Normal distribution can be used instead.

$X \sim N(np, npq)$

$X \sim N(90, 9)$

What is the probability of 80 or fewer getting cured?

$$Z = \frac{80.5 - 90}{\sqrt{9}} = -3.17$$

p-value = P(Z < -3.17) = 0.0008

Probabilities of Errors in Our Example

P(Type I error) = 0.05

<u>To calculate P(Type II error)</u>

H_0: p = 0.9

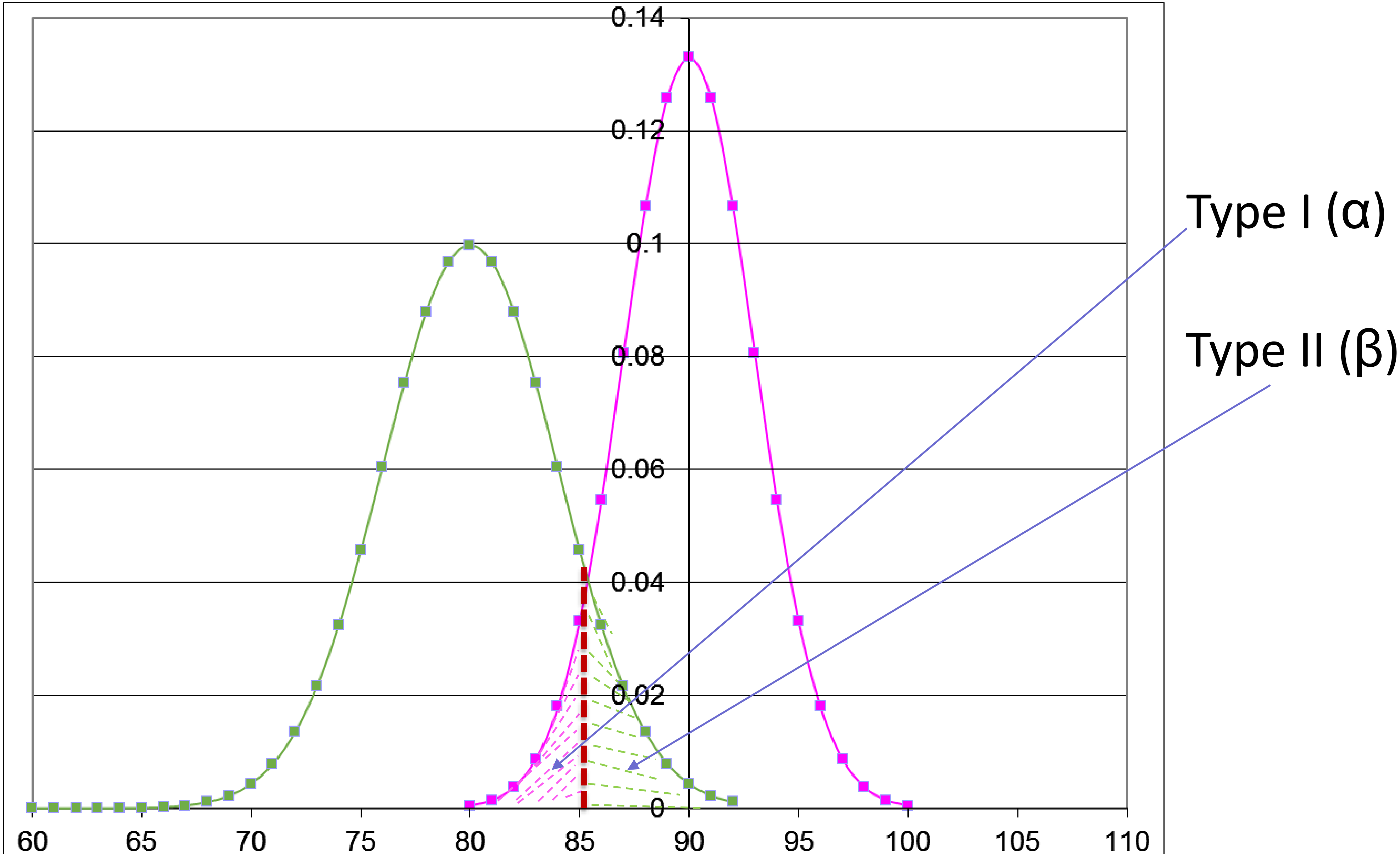
H_1: p = 0.8

P(Z < c) = 0.05 for 5% Significance Value. From probability tables, c = -1.64.

To de-standardize and find values outside the critical region,

$\frac{X-90}{\sqrt{9}} \geq -1.64; X \geq 85.08$, i.e., we would accept null hypothesis if 85.08 or more people had been cured.

Probabilities of Type I and Type II Errors



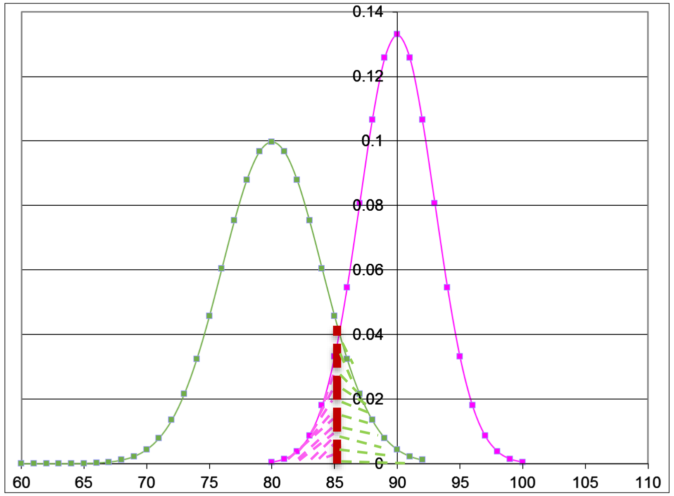
Probabilities of Errors in Our Example

Finally, we need to calculate $P(X \geq 85.08)$, assuming H_1 is true.

$X \sim N(np, npq)$ where n=100 and p=0.8. This gives $X \sim N(80, 16)$.

To calculate $P(X \geq 85.08)$ where $X \sim N(80, 16)$, we find

$$z = \frac{85.08 - 80}{\sqrt{16}} = 1.27$$



$P(Z \geq 1.27) = 1 - P(Z < 1.27) = 1\text{-}0.8980 = 0.102$

P(Type II error) = 0.102

The probability of accepting the null hypothesis that 90% are cured when its actually 80% is 10.2%.

Power of Hypothesis Test

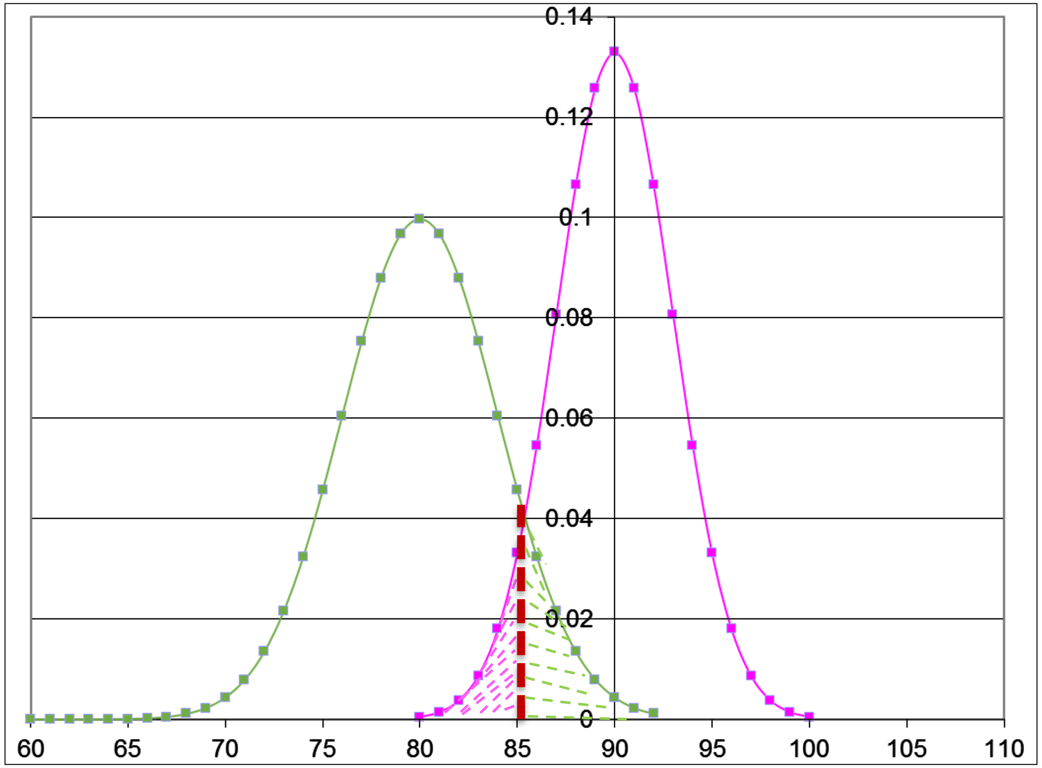
		State of Nature	
		Null true	Null false
Action	Fail to reject null (negative)	Correct decision True Negative Specificity P(*Accept* H_0 | H_0 True)	Type II error (β) False Negative P(*Accept* H_0 | H_0 False)
	Reject null (positive)	Type I error (α) False Positive P(Reject H_0 | H_0 True)	Correct decision (Power) True Positive Sensitivity/Recall P(Reject H_0 | H_0 False)

Power of a Hypothesis Test

We reject null hypothesis correctly when it is false.

It is actually the opposite of Type II error, and therefore,

Power = 1 – β = 1-0.102 = 0.898, i.e., the probability that we will make the correct decision in rejecting the null hypothesis is 89.8%.



Hypothesis Testing

A prisoner is on trial and you are on the jury. The jury's task is to assume that the accused is innocent, but if there is enough evidence, the jury needs to convict him.

In the trial, what is the null hypothesis?

The prisoner is innocent (or not guilty).

What is the alternate hypothesis?

The prisoner is guilty.

Hypothesis Testing

What are the possible ways of the jury coming to an incorrect verdict?

If the prisoner is innocent, and the jury gives a ‘guilty’ verdict.
If the prisoner is guilty, and the jury gives an ‘innocent’ verdict.

Which one is Type I and which one Type II?

First one is Type I because null hypothesis actually was correct but rejected incorrectly.
Second one is Type II because null hypothesis was false but was accepted incorrectly.

What is the Power of the test?

Since it is opposite of Type II, it will be finding the prisoner guilty when the prisoner is actually guilty, i.e., rejecting the null hypothesis correctly.

Common Test Statistics for Inferential Techniques

Inferential techniques (Confidence Intervals and Hypothesis Testing) most commonly use 4 test statistics:

- z
- t
- χ^2 (Chi-squared)
- F

Closely related to Sampling Distribution of **Means** (z, t)

- Closely related to Sampling Distribution of **Variances** (χ^2, F)
- Derived from Normal Distribution

TWO-SAMPLE *t*-TEST FOR MEANS

CSE 7315c

- Do two samples come from the same population?
- If they come from different populations, what is the difference in the means of the two populations?

  - Does the average cost of a two-bedroom flat differ between Bengaluru and Hyderabad? What is the difference?
  - What is the difference in the strength of steel produced under two different temperatures?
  - Does the effectiveness of Head & Shoulders anti-dandruff shampoo differ from Pantene anti-dandruff shampoo?
  - What is the difference in the productivity of men and women on an assembly line under certain conditions?
  - Does an antibiotic affect the efficacy of another drug being taken by a patient?

Two-sample t-Test

- Paired Data
  - You have two sets of data, where there is a natural pairing in the elements. Eg: BloodPressure from 30 people – one from before a treatment and other from after treatment.
- Unpaired Data
  - Comparing apartment costs from two cities
  - Two data sets of different length
  - No Natural pairing

Two-Sample t-Test for Paired Data

When the effects of two alternative treatments is to be compared, sometimes it is possible to make comparisons in pairs, where, e.g., the pair can be the same person at two different occasions or matched pairs where they are alike in all respects.

To study if their means are the same – we can create a new data set from the difference of the individual data points.

$$X_{new} = X_1 - X_2$$

We can then look at how far away from zero is the mean $E(X_{new})$

$$t = \frac{\overline{X_{new}} - 0}{SE(\overline{X_{new}})}$$

Two-Sample t-Test for Paired Data

A Yoga guru suggests that meditation increases concentration. To test this hypothesis, you get 12 volunteers and get them to complete a puzzle and you measure the time taken for completing the puzzle. The next day, you put them through a 30 minute meditation routine and have them complete another puzzle of similar difficulty. The time taken for completion is measured again.

You want to test at 5% Significance Level (or 95% Confidence Level) if the time taken is shorter after meditation.

Two-Sample t-Test for Paired Data

Time to Solve the puzzle(min)			
Patient	After Yoga(A)	Before Yoga (B)	A-B
1	63	55	8
2	54	62	-8
3	79	108	-29
4	68	77	-9
5	87	83	4
6	84	78	6
7	92	79	13
8	57	94	-37
9	66	69	-3
10	53	66	-13
11	76	72	4
12	63	77	-14
TOTAL	**842**	**920**	**-78**
MEAN	**70.17**	**76.67**	**-6.5**

Mean of the differences, $\bar{d} = -6.5$

Standard Deviation of the differences, $s_d = 15.1$

Standard Error of the mean, $SE(\bar{d}) = \frac{s_d}{\sqrt{n}} = 4.37$

$$t = \frac{\bar{d}}{SE(\bar{d})} = \frac{-6.5}{4.37} = -1.487$$

- Number of degrees of freedom= 12-1 = 11
- One tail test or two tailed test?

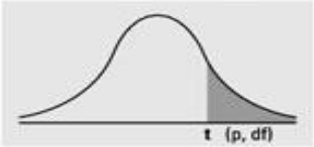
Two-Sample t-Test for Paired Data

$t = -1.487$

$t_{11,0.025} = 2.20099$

Comparing the absolute t-value we **cannot reject** the null hypothesis that the mean completion time is the same.

Numbers in each row of the table are values on a *t*-distribution with (*df*) degrees of freedom for selected right-tail (greater-than) probabilities (*p*).



df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Two-Sample t-Test for Paired Data

The 95% CI for the mean difference is given by $\bar{d} \pm t_{n-1,\frac{\alpha}{2}} * SE(\bar{d})$

$$-6.5 - 2.201 * 4.37 \leq D \leq -6.5 + 2.201 * 4.37$$

95% CI: (-16.1, 3.1).

As zero is included in the CI, we cannot reject the null hypothesis.

Business Decision (Yogic Decision?)

Although zero is included in CI, the range is very wide, which should lead the us to conduct a larger study to be sure.

Two sample t-Test : unpaired data

The Central Limit Theorem states that the difference in two sample means, $\overline{x_1} - \overline{x_2}$, is normally distributed for large sample sizes (both n_1 and $n_2 \geq 30$) whatever the population distribution.

Also, $\mu_{\overline{x_1}-\overline{x_2}} = \mu_1 - \mu_2$ *[Recall E(X-Y)=E(X)-E(Y)]*

and $\sigma_{\overline{x_1}-\overline{x_2}} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$ *[Recall Var(X-Y)=Var(X)+Var(Y)]*

$$z = \frac{observed\ difference - expected\ difference}{SE\ of\ the\ difference} = \frac{(\overline{x_1}-\overline{x_2})-(\mu_1-\mu_2)}{\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}}$$

This is the test statistic for a 2-sample z-test.

CSE 7315c

Two-Sample t-Test for Unpaired Data

$H_0\colon\ \mu_1 = \mu_2;\ H_1\colon\ \mu_1 \neq \mu_2$

Test statistic, $t = \frac{(\overline{x_1} - \overline{x_2})}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$

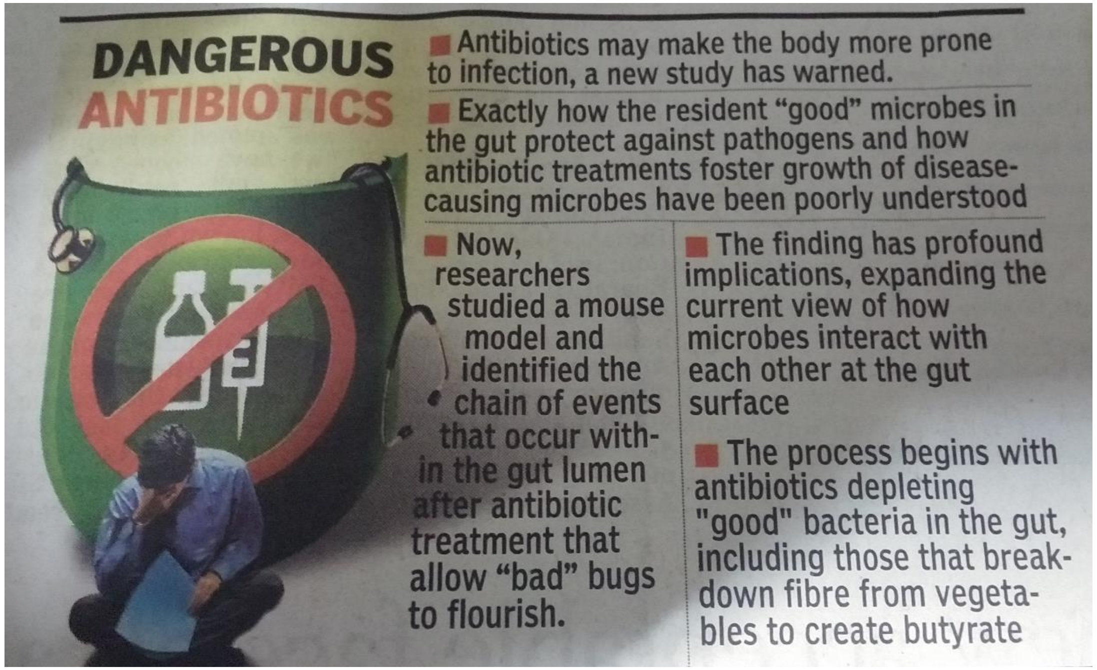
Assuming the two samples come from populations with the **same standard deviation** (Rule of thumb: The ratio between the higher *s* and the lower *s* is less than 2), pooled variance can be used to calculate SE.

$${s_p}^2 = \frac{(n_1 - 1){s_1}^2 + (n_2 - 1){s_2}^2}{(n_1 - 1) + (n_2 - 1)}$$

$t = \frac{(\overline{x_1} - \overline{x_2})}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ with $(n_1 + n_2 - 2)$ degrees of freedom.

Hypothesis Testing

Antibiotic rifampicin increases the amount of drug metabolizing enzyme present in the liver. This causes increase in the rate of elimination of a lot of other drugs.



DANGEROUS ANTIBIOTICS

- Antibiotics may make the body more prone to infection, a new study has warned.
- Exactly how the resident "good" microbes in the gut protect against pathogens and how antibiotic treatments foster growth of disease-causing microbes have been poorly understood
- Now, researchers studied a mouse model and identified the chain of events that occur within the gut lumen after antibiotic treatment that allow "bad" bugs to flourish.
- The finding has profound implications, expanding the current view of how microbes interact with each other at the gut surface
- The process begins with antibiotics depleting "good" bacteria in the gut, including those that break-down fibre from vegetables to create butyrate

Image Source: Deccan Chronicle, Hyderabad edition, May 04, 2016

An experiment was conducted to study whether rifampicin affects the metabolic removal of the anti-asthma drug theophylline. A high elimination rate would mean inadequate treatment of the patient's asthma.

Hypothesis Testing

Two groups of 15 subjects were pre-treated with oral rifampicin (600 mg daily for 10 days) and a placebo, respectively. All of them were then given intravenous injection of theophylline (3 mg/kg of body weight).

Drug content was then measured from the blood samples and efficiency of removal of theophylline reported as clearance (in ml/min/kg).

Hypothesis Testing

Clearance of theophylline (ml/min/kg)					
Control Subjects			Treated Subjects		
0.81	0.56	0.46	1.15	1.15	0.92
1.06	0.45	0.43	1.28	0.72	0.67
0.43	0.88	0.37	1.00	0.79	0.76
0.54	0.73	0.73	0.95	0.67	0.82
0.68	0.43	0.93	1.06	1.21	0.82

$n_2 = 15$
$\overline{x_2} = 0.633$
$s_2 = 0.216$
$s_2{}^2 = 0.0467$

$n_1 = 15$
$\overline{x_1} = 0.931$
$s_1 = 0.202$
$s_1{}^2 = 0.0408$

Hypothesis Testing

What is the null hypothesis?

$H_0: \mu_1 - \mu_2 = 0$ (Rifampicin does not cause a change in theophylline clearance)

What is the alternative hypothesis?

$H_1: \mu_1 - \mu_2 \neq 0$

Is it a one-tailed test or a two-tailed test?

Two-tailed

What could be a possible hypothesis for a one-tailed test?

Rifampicin decreases theophylline clearance.

Hypothesis Testing

At α = 0.05, determine if there is a significant difference between the two groups.

$$s_p^{\,2} = \frac{(n_1-1){s_1}^2+(n_2-1){s_2}^2}{(n_1-1)+(n_2-1)};\ t = \frac{(\overline{x_1}-\overline{x_2})}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \text{ with } (n_1+n_2-2) \text{ df.}$$

$$s_p^{\,2} = \frac{(15-1)*0.0408+(15-1)*0.0467}{(15-1)+(15-1)} = 0.04375;\ s_p = 0.209$$

$$t = \frac{0.931-0.633}{0.209*\sqrt{\frac{1}{15}+\frac{1}{15}}} = 3.91$$

You can find the p-value for this t-score or knowing that the t-score is way more than the critical value for 28 df (~ 2) at this significance level, you see that it is in the critical region in the right tail.

Hypothesis Testing

Will you reject the null hypothesis or fail to do so?

Reject. That means rifampicin does affect theophylline clearance.

Does it increase or decrease theophylline clearance and by how much?

As the treated patients showed a higher clearance (0.931 ml/min/kg) compared to the control group (0.633 ml/min/kg), rifampicin increases clearance by about 0.298 ml/min/kg).

R code: t.test(data1, data2, alternative="two.sided")

Confidence Intervals

$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ Rewriting:

$$(\overline{x_1} - \overline{x_2}) - t s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2 \leq (\overline{x_1} - \overline{x_2}) + t s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$0.298 - 2.048 * 0.0763 \leq \mu_1 - \mu_2 \leq 0.298 + 2.048 * 0.0763$$

95% CI: (0.142, 0.454)

Note zero difference is unlikely as at 95% Confidence Level, the difference ranges between 0.142 and 0.454 ml/min/kg, with a point estimate for the difference in mean clearance being 0.298 ml/min/kg.

Two-Sample t-Test for Unpaired Data

Welch's t-test using Welch-Sattherthwaite equation for df

$H_0\colon\ \mu_1 = \mu_2$; $H_1\colon\ \mu_1 \neq \mu_2$; Test statistic, $t = \frac{(\overline{x_1} - \overline{x_2})}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}}$

for **unequal standard deviations** for the two populations.

The degrees of freedom in this case are calculated as:

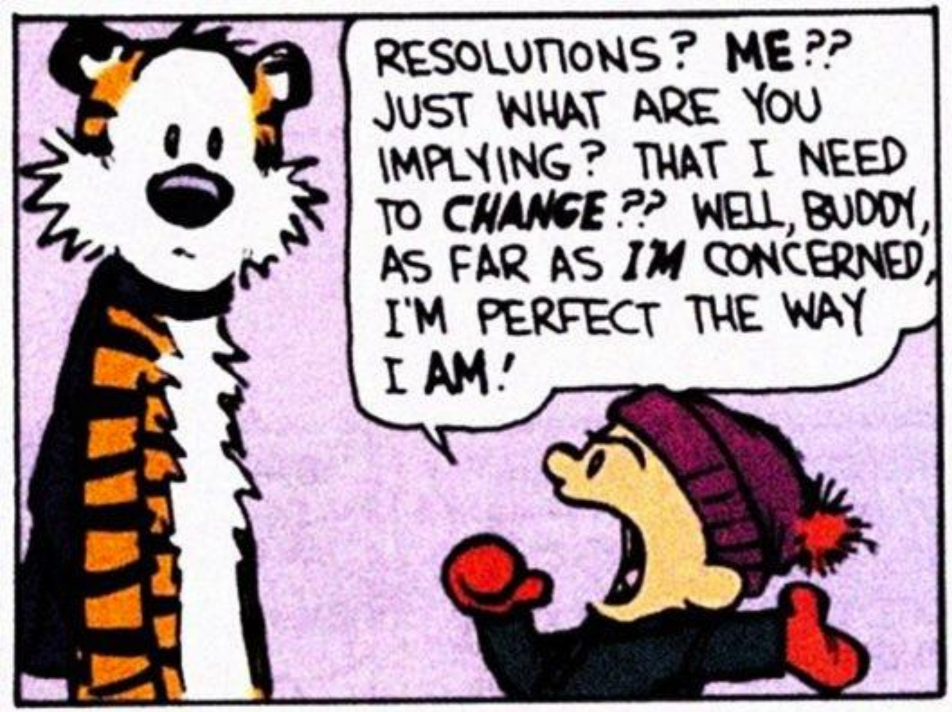
$\nu = \frac{\left(\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}\right)^2}{\left[\frac{\left(\frac{{s_1}^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{{s_2}^2}{n_2}\right)^2}{n_2 - 1}\right]}$, rounded off to the nearest integer.

R code: t.test(data1, data2, alternative="two.sided", var.equal=FALSE)

Summary

- Confidence Intervals
- t-Distribution
- Hypothesis testing
- Two-sample t-Tests


Hi-res JPG file download - Resolution 5000x3750px - www.psdgraphics.com
Happy New Year


RESOLUTIONS? ME?? JUST WHAT ARE YOU IMPLYING? THAT I NEED TO CHANGE?? WELL, BUDDY, AS FAR AS I'M CONCERNED, I'M PERFECT THE WAY I AM!

CSE 7315c

Resources

- Standard deviation – why use (n-1) instead of n?
  - https://www.khanacademy.org/video/review-and-intuition-why-we-divide-by-n-1-for-the-unbiased-sample-variance
  - http://nebula.deanza.edu/~bloom/math10/m10divideby_nminus1.pdf

- Z-statistic vs t-statistic: https://www.khanacademy.org/video/z-statistics-vs-t-statistics
- Hypothesis testing: https://www.khanacademy.org/video/hypothesis-testing-and-p-values

International School of Engineering

Plot 63/A, Floors 1&2, Road # 13, Film Nagar, Jubilee Hills, Hyderabad - 500 033

For Individuals:	+91-9502334561/63 or 040-65743991
For Corporates:	+91-9618483483
Web:	http://www.insofe.edu.in
Facebook:	https://www.facebook.com/insofe
Twitter:	https://twitter.com/Insofeedu
YouTube:	http://www.youtube.com/InsofeVideos
SlideShare:	http://www.slideshare.net/INSOFE
LinkedIn:	http://www.linkedin.com/company/international-school-of-engineering