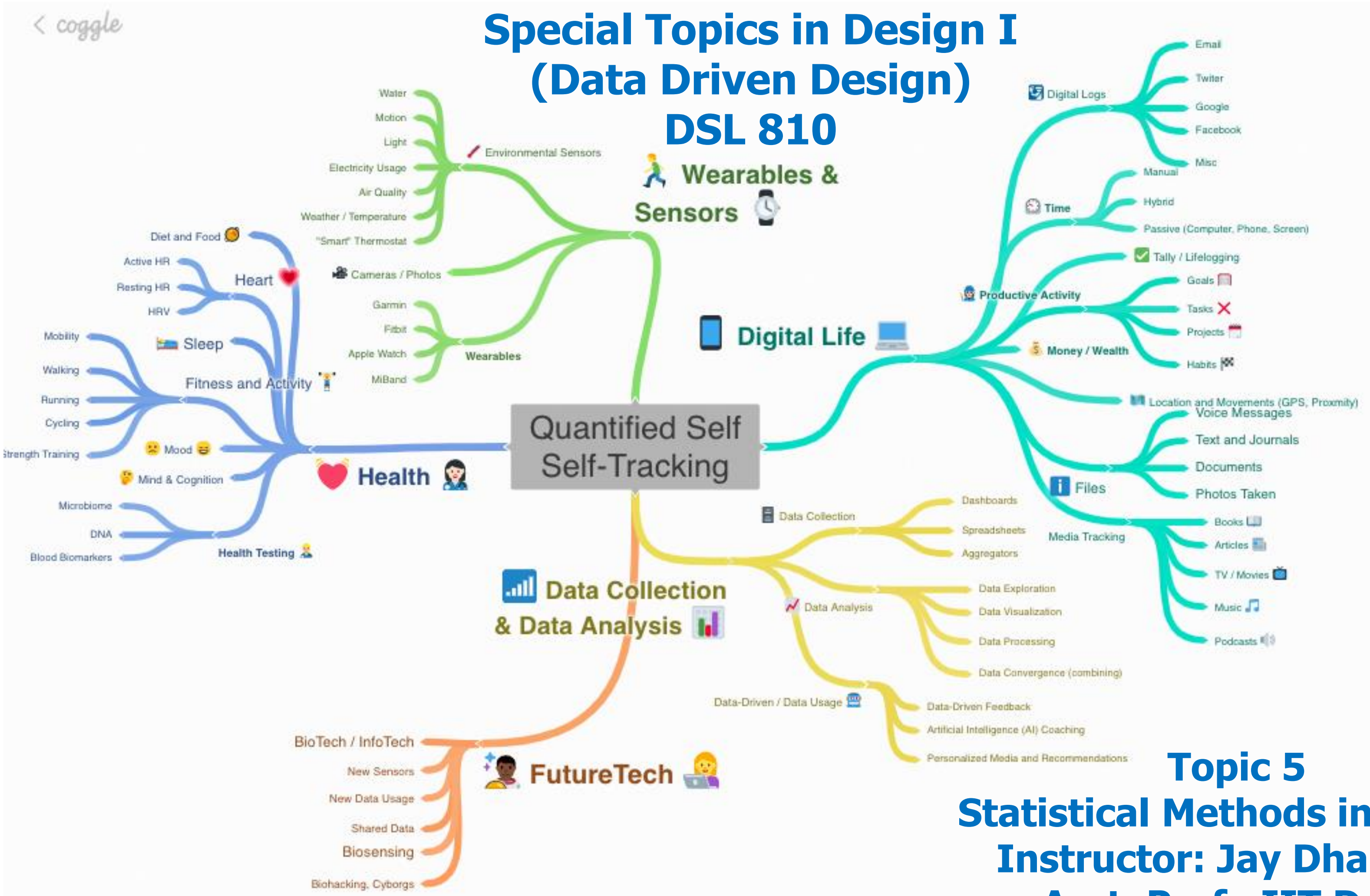
Special Topics in Design I (Data Driven Design) DSL 810



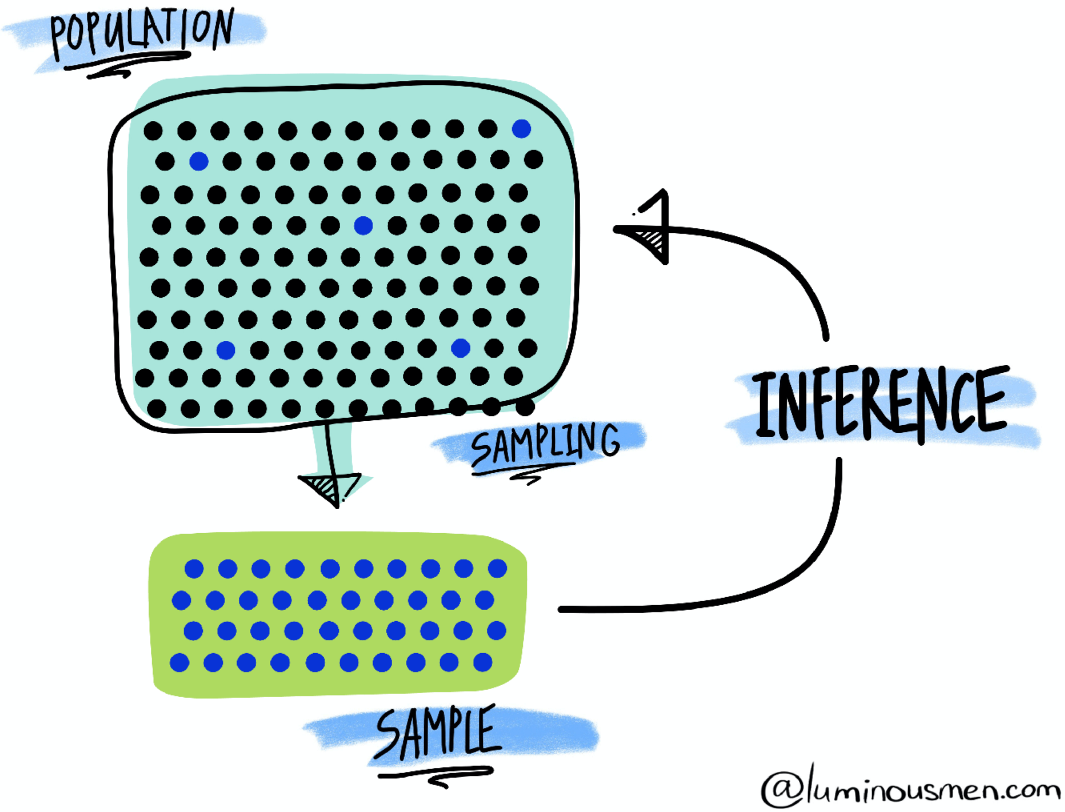
Topic 5
Statistical Methods in Design

Instructor: Jay Dhariwal, Asst. Prof., IIT Delhi

Dated: 2nd November, 2020

Source: http://www.markwk.com/quantified-self-mind-map.html

Application of Statistics

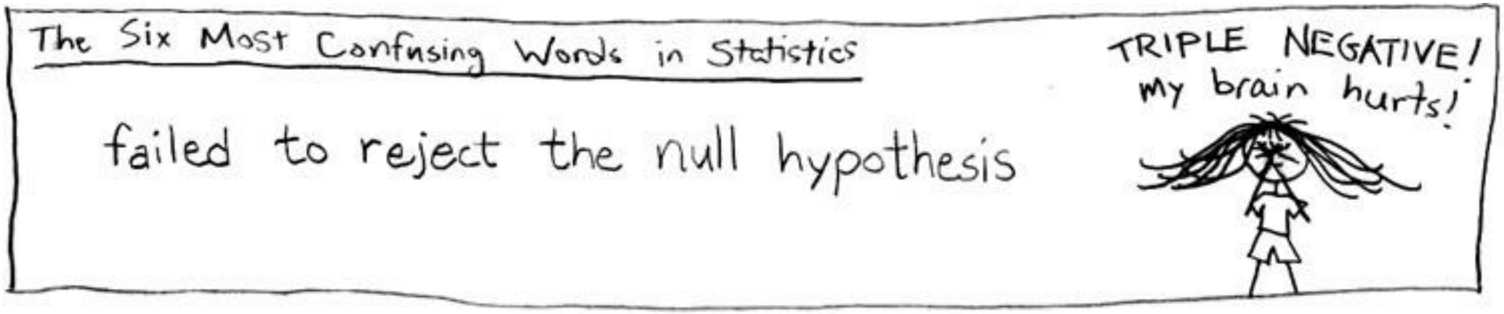


- Fields of application of Statistics
- Opinion Polls/Exit Polls for Elections Article
- User research – sample size Example
- Hypothesis Testing (Effect of certain exercises on health)
- Design of Experiments (Optimizing agriculture yield)
- Machine Learning (Regression, y=f(x))

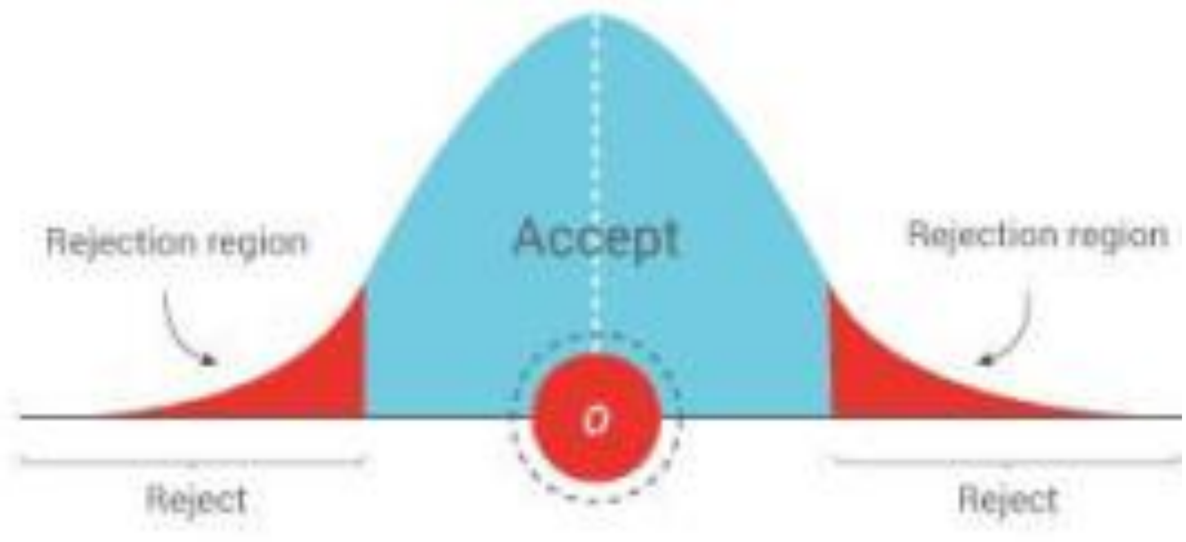
Image source: https://luminousmen.com/post/descriptive-and-inferential-statistics

Hypothesis Testing

- Virtual try-on of clothes, other projects – feedback from users
- Effect of interventions on air pollution = odd-even scheme or red signal on, gaadi off
- Effect of meditation on stress (J Clin Psychiatry. 2013 Aug; 74(8): 786–792)



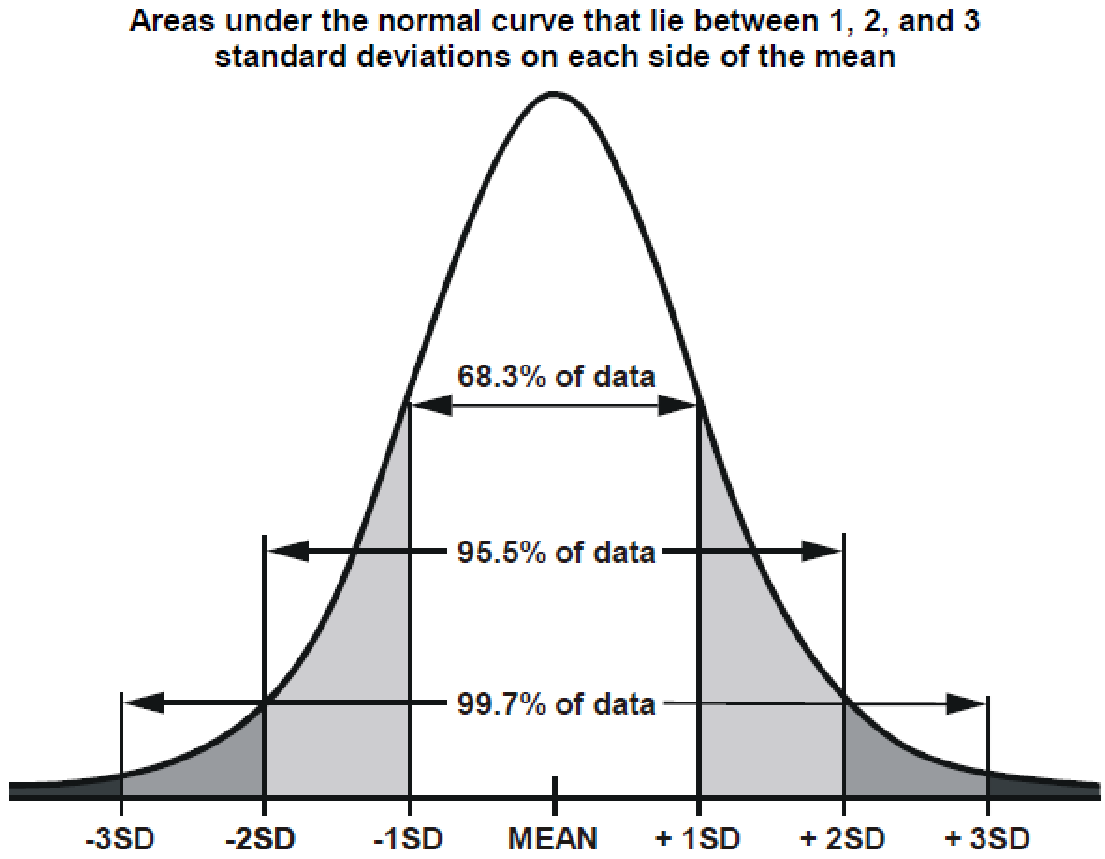
Source of image: http://web.stanford.edu/~rosenman/CME195/slides/Lecture7_Hypothesis_testing_and_classification.html



Source of image: https://www.iedunote.com/hypothesis-testing

Hypothesis Testing

- Weight Loss for Diet vs Exercise
- Did dieters lose more fat than the exercisers?
- Source:https://www2.stat.duke.edu/courses/Fall11/sta10/STA10lecture21.pdf
- How do we approach this problem?



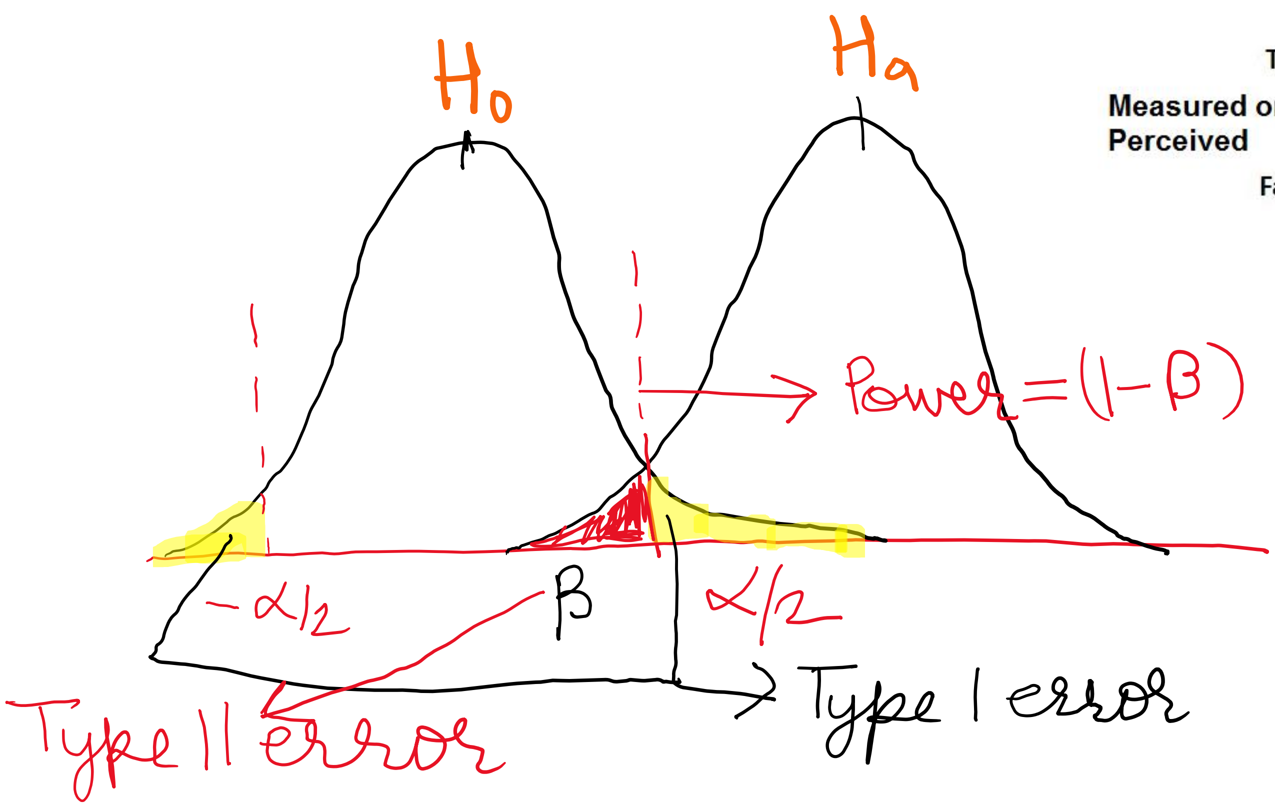
Normal distribution

- User testing example
- Example: height of students at school.
- Notation: N(μ, σ^2)
- PDF, CDF
- Standard normal distribution
- Z test statistic = (x - μ)/σ
- Z table
- Central Limit Theorem
- Assumptions of random sampling, i.i.d.

Weight loss for Diet vs Exercise?

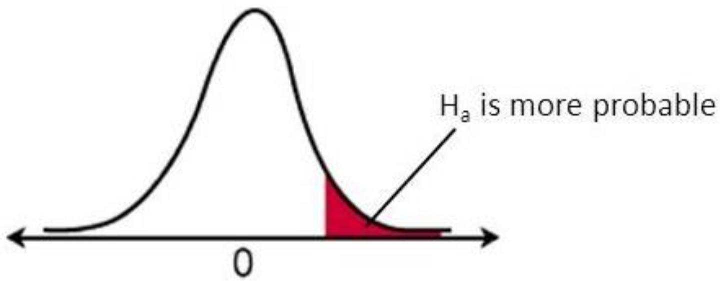
- Null hypothesis (Ho: $\mu_d = \mu_e$ or $\mu_d - \mu_e = 0$) = There is no difference in average fat lost in population for two methods.
- Alternative hypothesis (Ha: $\mu_d \neq \mu_e$) = There's a difference in average fat lost in population for two methods.

		Reality	
		True	False
Measured or Perceived	True	Correct ☺	Type 1 error False Positive
	False	Type 2 error False Negative	Correct ☺



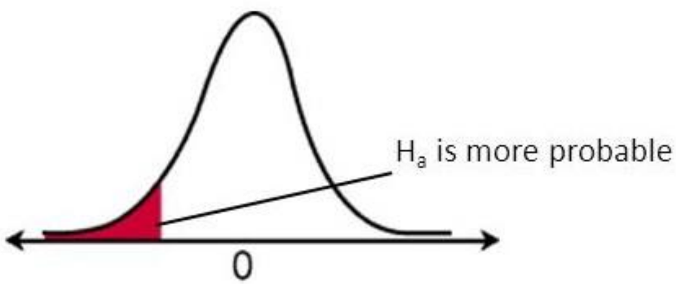
α ↓ β ↓ desired

One and two tailed tests



Right-tail test

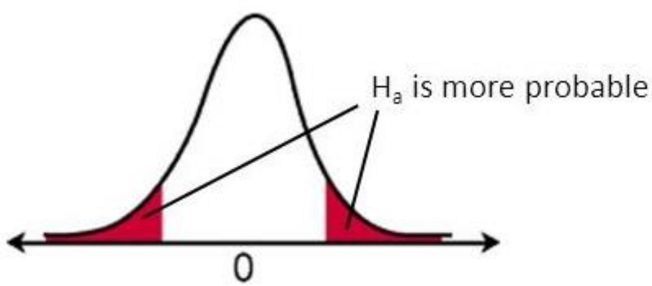
H_a: $\mu >$ value



Left-tail test

H_a: $\mu <$ value

[Power calculation for one tailed test]



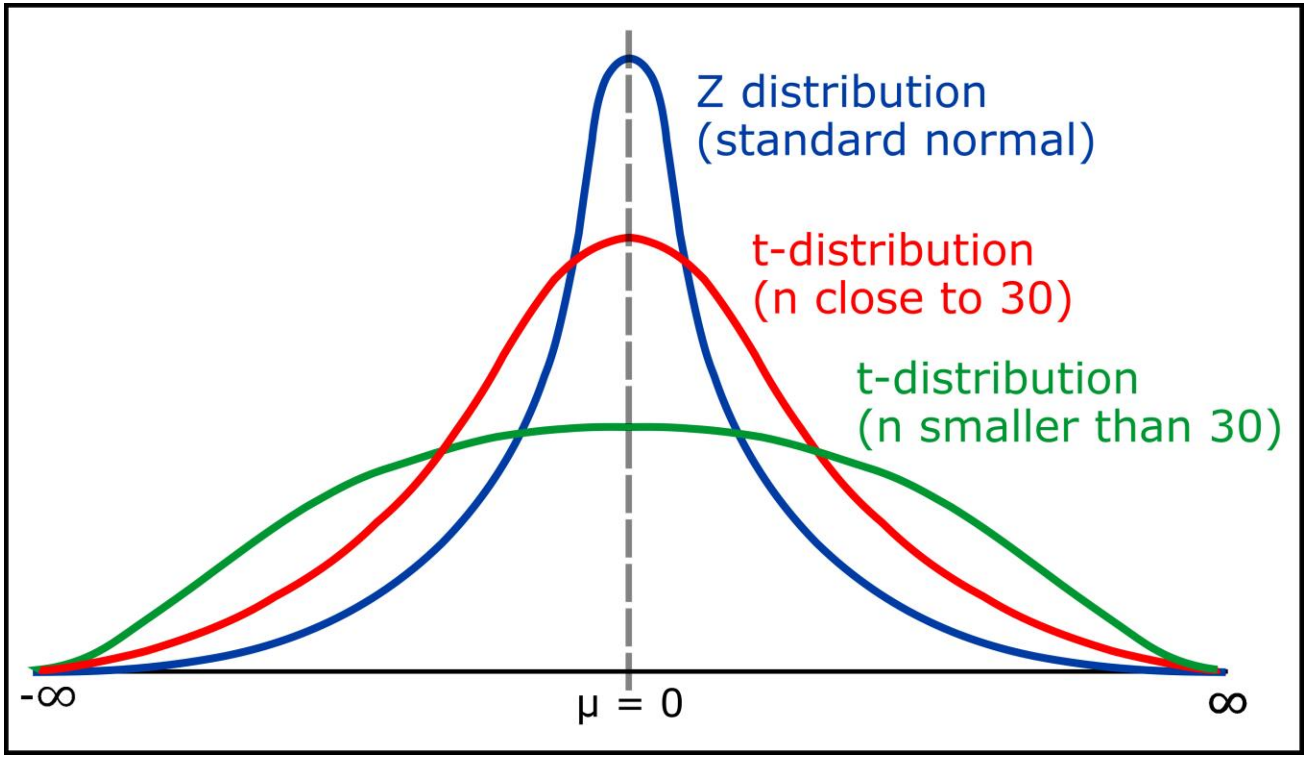
Two-tail test

H_a: $\mu \neq$ value

[Power calculation for two tailed test]

Source: https://www.fromthegenesis.com/difference-between-one-tail-test-and-two-tail-test/

Student's t-distribution



If the variance is unknown or the sample size is low, t distribution should be used.
t table read t table

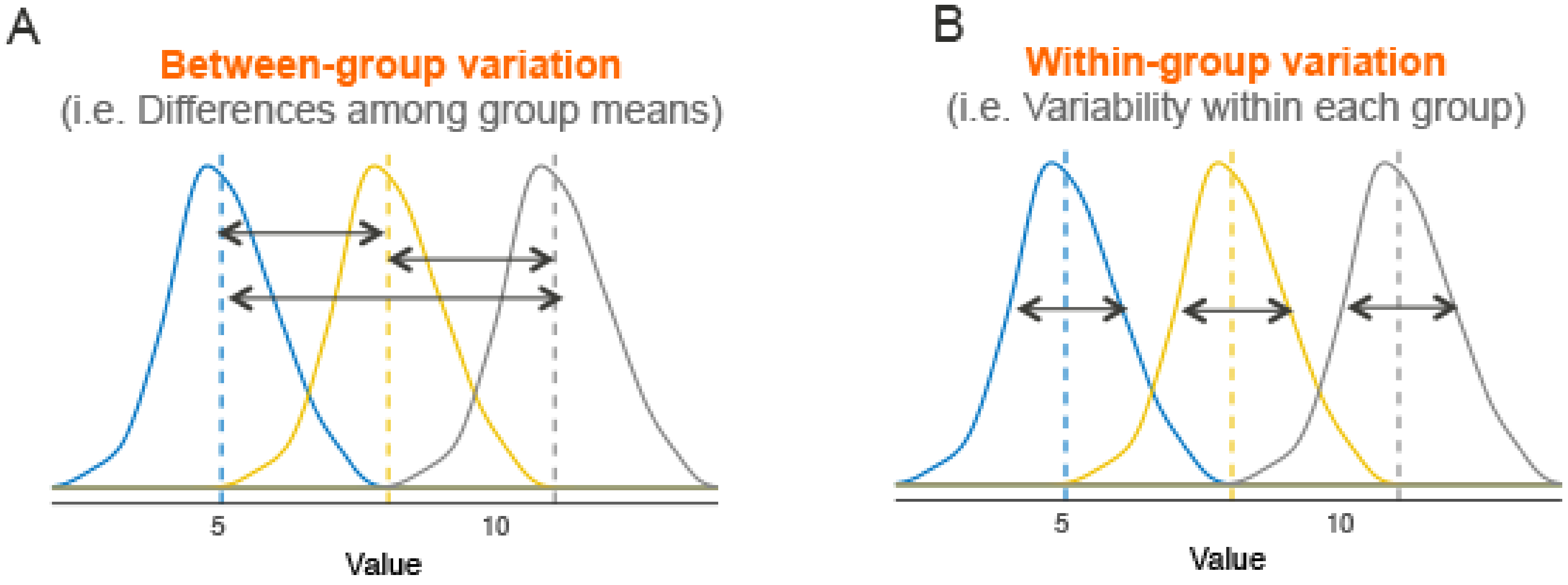
Source: https://www.geeksforgeeks.org/students-t-distribution-in-statistics/

Hypothesis Testing

- **Hypothesis testing** determines whether there is enough **statistical evidence** in favour of a hypothesis about a parameter.
- Please work on the user testing data to find out whether there is difference between concept A and concept B or not? Which statistical distribution would you use? What is the power of your test? What sample size to use to increase power > 90%?

Analysis of Variance (ANOVA)

- What if we wanted to find out if there were statistically significant differences between groups of three or more means?



ANOVA components

- Do all children from school A, B and C have equal mean IQ scores? Link
- $SS_T = SS_{treatments} + SS_E$
- If F-statistic > $F_{\alpha,\, k-1,\, N-k}$ then we reject H_o and conclude that there is a difference in the means

Source of variation	Sum of squares (SS)	Degrees of freedom(DF)	Mean Square (MS)	F-statistic
Treatments	$SS_{between}$ (SS_b)	k-1	$MS_b = SS_b/(k-1)$	$F = MS_b/MS_w$
Error (or Residual)	SS_{Within} (SS_W)	N-k	$MS_W = SS_W/(N-k)$	
Total	SS_{Total} (SS_T)	N-1		

F table

When to use one way ANOVA

- one categorical independent variable and one quantitative dependent variable.
- For e.g. - your independent variable is social media platform, and you assign levels to different kinds to rate the UX on them (dependent variable).
- independent variable >= three levels
- ANOVA tells you if the dependent variable changes with level of the independent variable
- Assumptions: $N(\mu, \sigma^2)$, i.i.d.

Source: https://www.scribbr.com/statistics/one-way-anova/

Two way ANOVA

- A two-way ANOVA estimates how the mean of a quantitative variable changes according levels of two categorical variables.
- Two-way ANOVA used when we want to know how two independent variables, in combination, affect a dependent variable. Example
- Example 2: If we are researching which type of fertilizer and planting density produces the greatest crop yield in a field experiment, we assign different plots in a field to a combination of fertilizer type (1, 2, or 3) and planting density (1=low density, 2=high density), and measure the final crop yield at harvest time.
- We can use a two-way ANOVA to find out if fertilizer type and planting density have an effect on average crop yield.

Source: https://www.scribbr.com/statistics/two-way-anova/

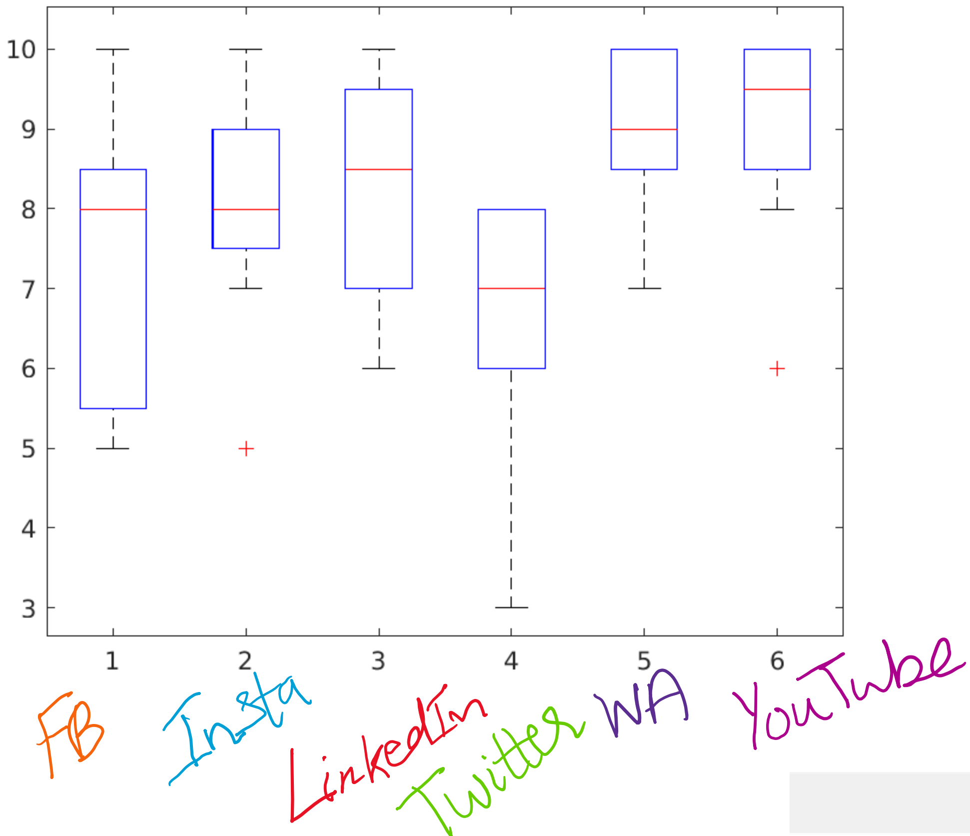
Two way ANOVA table

- Independent variables = Car models, Factories
- Dependent variable = Mileage

Source	SS	df	MS	F	*p*-value
Columns	SS_A	$k - 1$	MS_A	MS_A/MSE	$P(F_{k-1,mk(R-1)}) > F$
Rows	SS_B	$m - 1$	MS_B	MS_B/MSE	$P(F_{m-1,mk(R-1)}) > F$
Interaction	SS_{AB}	$(m-1)(k-1)$	MS_{AB}	MS_{AB}/MSE	$P(F_{(m-1)(k-1),mk(R-1)}) > F$
Error	SSE	$mk(R - 1)$	MSE		
Total	SST	$mkR - 1$			

Source: https://in.mathworks.com/help/stats/two-way-anova.html

UX for Social Media Platform



- The null hypothesis of the means of all Social Media Platforms being the same is rejected.
- At 5% α level, Twitter UX is felt to be different from WhatsApp and YouTube.
- At 10% α level, FB differs from WhatsApp and YouTube differ; LinkedIn and Twitter UX differ. At 20% α level, Instagram and Twitter UX differ.
- We can conclude that Twitter (except FB) UX is least liked in this class.
- More data is needed for further clarity.
- Kruskal-Wallis Test for Normality assumption deviation

Social Media Platform Type → (Columns)

Engg. vs. Non-Engg. → (Rows)

ANOVA Table

Source	SS	df	MS	F	Prob>F
Columns	53.736	5	10.7472	5.32	0.0004
Rows	5.014	1	5.0139	2.48	0.1204
Interaction	10.403	5	2.0806	1.03	0.4082
Error	121.167	60	2.0194		
Total	190.319	71			

Assumptions not valid

Two way ANOVA example

- We have marks from a project review given to students (non-engg. and engg. orientation) by different faculty members.
- Is there any difference between the marks given by different faculty to the students?
- Is there any difference between the marks given by faculty to an engg. vs. a non-engg. student?
- What do you infer from this analysis?
- Dataset

References

1. William W. Hines, Douglas C. Montgomery, David M. Goldsman, Connie M. Borror, Probability and Statistics in Engineering, 4th edition, Wiley, 2009.
2. Douglas C. Montgomery and George C. Runger, Applied Statistics and Probability for Engineers, 6th edition, Wiley, 2016
3. Douglas C Montgomery, Elizabeth A Peck, et al. Introduction to Linear Regression Analysis, 3rd edition, Wiley, 2006
4. Douglas C. Montgomery, Design and Analysis of Experiments, 8th edition, Wiley, 2013

Announcements

- Interim Project Presentation guidelines
- Evaluation criteria