

# Statistics/Probability

Statistics is the study of the collection, organization, analysis, interpretation and presentation of data. It deals with all aspects of data, including the planning of data collection in terms of the design of surveys and experiments. Probability theory is the branch of mathematics concerned with probability, the analysis of random phenomena. The central objects of probability theory are random variables, stochastic processes, and events.

Knowledge item description	References
Beginner	
Visualizing relationships in data	[1.1]
Probability; Bayes Rule; Correlation vs. Causation	[1.2]
Maximum Likelihood Estimation; Mean, Median, Mode; Standard Deviation, Variance	[1.3]
Outliers, Quartiles; Binomial Distribution; Central Limit Theorem; Manipulating Normal Distribution	[1.4]
Confidence intervals; Hypothesis Testing	[1.5]
Linear regression; correlation	[1.6]
General concepts of probability	[3.1]
Competent	
Experimental research	[2.1]
Correlational research	[2.2]
Variables and distributions	[2.3]
Summary statistics	[2.4]
Correlation	[2.5], [2.21]
Measurement	[2.6]
Basics of regression	[2.7]
Null Hypothesis Significance Tests (NHST)	[2.8]
Central limit theorem	[2.9]
Confidence intervals	[2.10], [2.11]
Multiple regression	[2.12]
Moderation	[2.13]
Mediation	[2.14]
Expert	
Group comparisons (t-tests)	[2.15], [2.23]
Group comparisons (ANOVA)	[2.16], [2.23]
Factorial ANOVA	[2.17], [2.23]
Repeated measures ANOVA	[2.18], [2.23]
Chi-square	[2.19]
Binary logistic regression	[2.20]
Generalized Linear Model	[2.22]
Non-parametrics (Mann-Whitney U, Kruskal-Wallis)	[2.24]
Markov Models, Hidden Markov Models	[3.2]
Multivaried Gaussian Distributions	[3.3], [3.4]

## References

#	Reference	Link
1.	Udacity Class: Introduction to Statistics	<a href="#">Link</a>
1.1	Lesson 1: Visualizing relationships in data	
1.2	Lesson 2: Probability	
1.3	Lesson 3: Estimation	
1.4	Lesson 4: Outliers and Normal Distribution	
1.5	Lesson 5: Inference	
1.6	Lesson 6: Regression	
2.	Coursera Class: Statistics One	<a href="#">Link</a>
2.1	Lecture 1: Experimental research	
2.2	Lecture 2: Correlational research	
2.3	Lecture 3: Variables and distributions	
2.4	Lecture 4: Summary statistics	
2.5	Lecture 5: Correlation	
2.6	Lecture 6: Measurement	
2.7	Lecture 7: Introduction to regression	
2.8	Lecture 8: Null Hypothesis Significance Tests (NHST)	
2.9	Lecture 9: Central limit theorem	
2.10	Lecture 10: Confidence intervals	
2.11	Lecture 11: Multiple regression	
2.12	Lecture 12: Multiple regression continued	
2.13	Lecture 13: Moderation	
2.14	Lecture 14: Mediation	
2.15	Lecture 15: Group comparisons (t-tests)	
2.16	Lecture 16: Group comparisons (ANOVA)	
2.17	Lecture 17: Factorial ANOVA	
2.18	Lecture 18: Repeated measures ANOVA	
2.19	Lecture 19: Chi-square	
2.20	Lecture 20 Binary logistic regression	
2.21	Lecture 21: Assumptions revisited (correlation and regression)	
2.22	Lecture 22: Generalized Linear Model	
2.23	Lecture 23: Assumptions revisited (t-tests and ANOVA)	
2.24	Lecture 24: Non-parametrics (Mann-Whitney U, Kruskal-Wallis)	

3.	Stanford CS229 Lecture Notes and Handout Materials	<a href="#">Link</a>
3.1	Probability Theory Review	<a href="#">Section Note</a>
3.2	Hidden Markov Models	<a href="#">Section Note</a>
3.3	The Multivariate Gaussian Distribution	<a href="#">Section Note</a>
3.4	More on Gaussian Distribution	<a href="#">Section Note</a>
3.5	Gaussian Processes	<a href="#">Section Note</a>