

# Statistical Inference and Multivariate Analysis (MA324)

## LECTURE SLIDES Lecture 01

Dr. Palash Ghosh



Indian Institute of Technology Guwahati

Jan-May 2023

# Syllabus

- **Review** of different transformation techniques, modes of convergence, law of large numbers, and central limit theorem; Sampling distributions based on normal distributions, multivariate normal distribution;
- **Point estimation:** sufficiency, Neymann-Fisher factorization theorem, unbiased estimation, method of moments, maximum likelihood estimation, consistency and asymptotic normality of maximum likelihood estimator;
- **Interval estimation:** confidence coefficient and confident level, pivotal method, asymptotic confidence interval, Bootstrap confidence interval;
- **Hypothesis testing:** type-I and type-II errors, power function, size and level, test function and randomized test, most powerful test and Neyman-Pearson lemma, likelihood ratio test, p-value; Multiple linear regression:
- **Multiple linear regression:** least squares estimation, estimation of variance, tests of significance, interval estimation, multicollinearity, residual analysis, PRESS statistic, detection and treatment of outliers, lack of fit; Multivariate analysis:
- **Multivariate analysis:** principle component analysis, factor analysis, canonical correlations, cluster analysis

# Books

- Text Books

- R. V. Hogg, J. W. McKean and A. T. Craig, Introduction to Mathematical Statistics, 7th Ed., Pearson, 2013.
- D. C. Montgomery, E. A. Peck and G. G. Vining, Introduction to Linear Regression Analysis, 5th Ed., Wiley, 2012.
- R. A. Johnson and D. W. Wichern, Applied Multivariate Statistical Analysis, 6th Ed., Prentice Hall of India, 2012.
- R. V. Hogg, E. A. Tanis, D. L. Zimmerman, **Probability and Statistical Inference**, Pearson, 2015.

- Reference Books

- V. K. Rohatgi and A. K. Saleh, An Introduction to Probability and Statistics, 3rd Ed., Wiley, 2015.
- G. Casella and R. L. Berger, Statistical Inference, 2nd Ed., Cengage Learning, 2006.
- N. R. Draper and H. Smith, Applied Regression Analysis, 3rd Ed., Wiley, 2000.
- S. Weisberg, Applied Linear Regression, 1st Ed., Wiley, 2005.
- T. W. Anderson, An Introduction to Multivariate Statistical Analysis, 3rd Ed., Wiley, 2012.

# Grading Policy

- Weights in different examination are as follows:
  - Quiz I: 15%
  - Mid-semester Examination: 30%
  - Quiz II: 15%
  - End-semester Examination: 40%
- Penalty for illogical arguments during copy showing from obtained marks : 5%
- An **F** grade will be awarded if you obtain less than 20% of total marks after the end semester examination.

# Why should we study Probability and Statistics?

# Why should we study Probability and Statistics?

- Lets a play a game!

# Why should we study Probability and Statistics?

- Lets a play a game!
- You are familiar with Ludo (board game). Similar to that, consider a game having a die with four outcomes (instead of six in Ludo-die): 1, 2, 3 and 4.

# Why should we study Probability and Statistics?

- Lets a play a game!
- You are familiar with Ludo (board game). Similar to that, consider a game having a die with four outcomes (instead of six in Ludo-die): 1, 2, 3 and 4.
- Here, the board is  $\mathbb{R}^2$ . At the beginning, you are at  $(0, 0)$ . If you are at  $(x, y)$ , then your (the token) next move is determined by the following Table<sup>1</sup> using the outcome by rolling the die:

---

<sup>1</sup>Probability and Statistics by Dr. Arnab Chakraborty



# Why should we study Probability and Statistics?

- Lets a play a game!
- You are familiar with Ludo (board game). Similar to that, consider a game having a die with four outcomes (instead of six in Ludo-die): 1, 2, 3 and 4.
- Here, the board is  $\mathbb{R}^2$ . At the beginning, you are at  $(0, 0)$ . If you are at  $(x, y)$ , then your (the token) next move is determined by the following Table<sup>1</sup> using the outcome by rolling the die:

Outcome of the die	New position in $\mathbb{R}^2$
1	$(0.8x + 0.1, 0.8y + 0.04)$
2	$(0.5x + 0.25, 0.5y + 0.4)$
3	$(0.355(x-y) + 0.266, 0.355(x+y) + 0.078)$
4	$(0.355(x+y) + 0.378, 0.355(y-x) + 0.434)$

<sup>1</sup>Probability and Statistics by Dr. Arnab Chakraborty

# Statistical Regularity: Random to Deterministic...

- From totally meaningless movement to something meaningful in long run!
- See the final outcome-graph (repeating many times) of the above play in R-software.
- The beauty of Probability and Statistics is to get valuable information from Random Experiment which may seems meaningless at the beginning!
- The notion of getting something meaningful (regularity) from a random phenomena (experiment) is called **Statistical Regularity**.