

Knowledge Base in Regression and Inference

Below is a list of the minimum knowledge requirements for enrolling in graduate statistical courses in MSc Data Science. In brackets is the initial estimated time that a student would need to grasp the concept (since you already have appropriate math skills such as knowing how to take derivatives, integrals, etc..).

Probability concepts

- [30 min] Comprehend the notions of random variables, probability distributions, expected value and variance (discrete and continuous)
[30 min] Compute simple and joint probabilities from distributions (discrete and continuous) and tables (discrete)
[60 min] Be comfortable with well-known discrete probability distributions, such as binomial and Poisson distributions.
[30 min] Master the Normal and t-distributions, be aware of their shapes and properties.
- [30 min] Know the definition of moment generating function (MGF) and characteristic function (CGF).

Inference concepts

- [15 min] Know how to estimate the mean and variance for a normal distribution
- [60 min] Know how to construct a confidence interval and carry out a hypothesis test on a mean, proportion, variance.
- [30 min] Interpret confidence intervals and hypothesis tests and use them to make statistical decisions.
- [15 min] Know the main idea of the Central Limit Theorem (CLT).
- [15 min] Comprehend the notion of a sampling distribution and know the sampling distribution of the sample mean under various circumstances.
- [30 min] Know the likelihood definition and how to find Maximum likelihood estimate for a given likelihood function.
- [60 min] Know the likelihood ratio test.

The link below points to an online course which covers the most fundamental topics for the inference course (STAT 5310):

<https://online.stat.psu.edu/stat200/home>

For other advance topics, such as MGF, CF, likelihood, MLE and likelihood ratio test, please watch the following videos or online lecture notes:

CF and example:

1. Intro (6 minutes)

<https://www.youtube.com/watch?v=mYhca1p26n4>

2. CF for Normal distribution -part 1 (5 minutes)

<https://www.youtube.com/watch?v=-glT8cCczfw>

3. CF for Normal distribution - part 2 (4.5 minutes)

<https://www.youtube.com/watch?v=105aWG54AL8>

MGF and examples

<https://www.youtube.com/watch?v=cbmFYoepHPk> (~17 minutes)

<https://www.youtube.com/watch?v=rIdqnwEj7w0> (~5 minutes)

Point Estimation, likelihood and MLE

http://www.stats.ox.ac.uk/~dlunn/b8_02/b8pdf_5.pdf

Likelihood ratio test: theory and examples

http://www.stats.ox.ac.uk/~dlunn/b8_02/b8pdf_8.pdf