

**Aryan Sharma**  
**Problem Statement**

Download the data set [LungCancer.txt](#). It contains data from a Veteran's Administration Lung Cancer Trial, where 137 patients with advanced, inoperable lung cancer were treated with chemotherapy (standard treatment) vs chemotherapy combined with a new drug (test treatment). Your objective is to do survival analysis on this data.

1. We would like to see Kaplan-Meier survival graphs for patients with the test vs standard treatment.

Use this data to assess:

- What is the probability that the patient will survive for 6 months (183 days) and 1 year (365 days) on the standard treatment vs the test treatment?
- What is the median number of days where a patient can be expected to survive if they are on the standard vs the test treatment?

2. Create three semi-parametric and parametric models to estimate the marginal effects of relevant predictors on survival outcomes. Interpret the precise effects of standard vs test treatment and other model predictors.

Page limit for this assignment is 3 pages. Points will be assigned as follows:

- Kaplan-Meier curves and their interpretation
- Parametric and semi-parametric models and their interpretation
- Completeness of analysis