In this quarter's DRP, I explored survival analysis by reading the Chapter 7 of the online textbook "Introduction to Regression Methods for Public Health Using R". This chapter primarily focused on the survival function (Kaplan-Meier estimate) and Cox regression (hazard ratio), with a stronger emphasis on Cox regression. It introduced relevant definitions and demonstrated how to analyze time-to-event datasets, using preterm birth as the primary example. Also, it provided R code to implement these analyses.

The first two sections 7.1 and 7.2 briefly introduced what is survival analysis and provides several examples: time to preterm birth, time to heart attack, transition to heroin. And then explained the concept of time-to-event data and censoring, where censoring occurs when the exact event time is not observed for an individual.

Then 7.3 introduced the survival function, which represents the probability of surviving beyond a given time and complements the cumulative distribution function (CDF). Using preterm birth data as an example, the text explains how survival probabilities decrease over time with events and demonstrates the Kaplan-Meier survival curve.

Section 7.4 moves on to the hazard function, defining it as the instantaneous risk of an event and highlighting its relationship to the survival function. It illustrates how hazard rates change over time and differ between groups, again using preterm birth as a case study.

7.5 details the structure of datasets required for survival analysis in R, emphasizing the need for an event time variable and a binary event indicator and using preterm birth dataset to show proper format in R.

7.6 further explored The Kaplan-Meier method, with step-by-step explanations of how to estimate survival probabilities nonparametrically and show the plot of survival functions look like.

7.7 introduced Cox regression, which is a semi-parametric method for modeling the relationship between predictors and the hazard function, with an emphasis on the proportional hazards assumption and the interpretation of hazard ratios.

Section 7.8 - 7.11 all focused on coding in R, they showed how to fit cox regression in R, the techniques for visualizing hazard ratios through forest plots, predicting survival probabilities and hazard ratios for specific individuals, and plotting estimated survival functions to compare survival curves across groups.

7.12 it discusses incorporating interaction terms into Cox regression models to evaluate effect modification, with examples of testing and visualizing these interactions.