

*BIOS 522: Survival Analysis Methods*

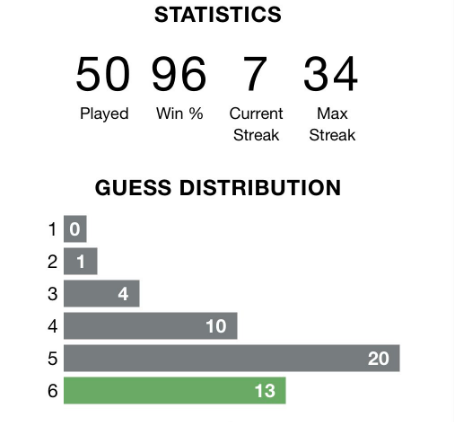
**Activity 4:**

**The hazard and cumulative hazard functions**

*This week, we defined the hazard and cumulative hazard functions and their relationship to the survival function. We studied three key parametric distributions used for time-to-event data. We reviewed maximum likelihood estimation for survival data. We estimated the cumulative hazard function using the Nelson-Aalen estimator.*

Problem 1. Wordle and discrete time survival

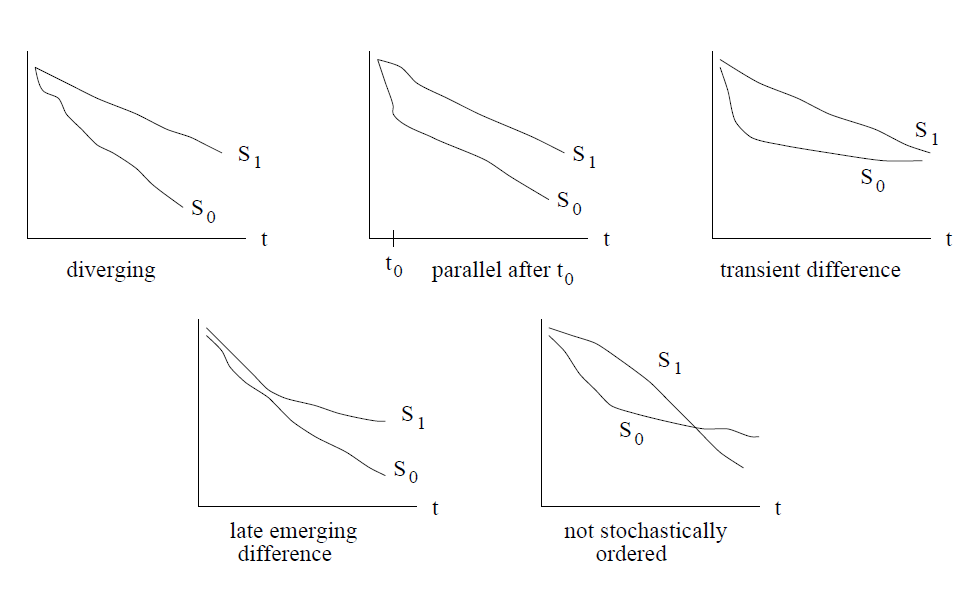
Wordle is a popular word game. Each day, players are given six guesses to guess a secret five-letter word. If the player cannot guess the word in six guesses, they lose. The game tracks statistics for how often a player wins, and, among the wins, how often the player gets the answer in 1, 2, 3, 4, 5 or 6 guesses. An example of a player’s win statistics are provided below.



1. To the extent allowable by the data, estimate the density function for this player’s guess distribution. What happens after 6 guesses?
2. In discrete time, the hazard is the conditional probability of the event occurring at time given that the event has not yet occurred immediately before time . Estimate the hazard function for times using the player’s guess distribution.
3. Do the density function and hazard function maximize at the same time? Why or why not?

Problem 2. Sketching the hazard functions

For each of the five pairs of survival curves shown below, sketch the corresponding pairs of hazard functions.



Problem 3. Log-logistic likelihood function

Recall that the log-logistic distribution has rate parameter and shape parameter . It has hazard function:

And survival function:

Define the likelihood function for right-censored data when the underlying failure time random variable follows the log-logistic distribution.