What is a hazard rate?

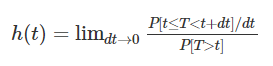
Many of the statistical analyses of mortality mention the hazard rate or the hazard function. What is this, exactly?

First, you have to remember the difference between a probability and a rate. Both involve counts, but a probability involves one count divided by another count. This ratio is always between zero and one because the count in the numerator represents the count of a subset.

In contrast, a rate is a count divided by a measure of time. A rate can never be negative, but it can be larger than one, depending on the units of time that you use.

The hazard rate is the short term rate of death (or the short term failure rate, if you are talking about a mechanical device).

The mathematical definition is



and if this reminds you of your college Calculus class, you remember more about college than I do. But you can (and should) think about the hazard function without the rigor of mathematics.

In some settings, the hazard rate increases with time. This means that the death or failure rate gets larger as the person/device gets older. This is the most common pattern for a hazard rate, and it implies that new is better than used.

In some situations, the reverse is true and the hazard rate decreases with time. This implies that a person or device is likely to see early mortality or failure, but that the person or device toughens over time and becomes less likely to fail. This is a pattern you see with some electronic equipment, which if it doesn’t fail early, is likely to last you through the long haul. This represents the case where used is better than new.

If the hazard rate is constant over time, you have a setting where the short term failure rate is independent of time. A new device is equivalent to a used device and it doesn’t matter how long the device has been around. This is sometimes referred to as a “memoryless” process. THe device doesn’t need to remember how old it is because its short term failure rate remains unchanged with age.

If you are a normal human being, you have a “bathtub” shaped hazard function. Your hazard starts out high. In fact, your first day of your life was the most dangerous on, and even the first week and month were quite perilous. The older you got, the tougher you got and all was good until you hit your teenage years. Then, things started to slowly get worse again. Your hazard rate slowly increased because you started doing dangerous adult things like driving a car. Then as the decades moved along, the aging process in your body contributed even more to the hazard.

Why is the hazard function important? In an industrial setting, a hazard function tells you when to use a system of preventive replacement of aging parts. If you make some reasonable assumptions about the form of the hazard function, you can (carefully) extrapolate survival probabilities to longer time frames. If you are planning to study survival probabilities, knowledge of the hazard function can help you plan the sample size and duration of follow-up for your research. Finally, assumptions about the hazard function are critical for regression models of survival.