3.1: Random Variables

Taylor

University of Virginia

Motivation

Even though sometimes you might think of a random variable X as just a random number, technically they're actually functions. We'll define random variables in a more technical way so that a) you won't be totally confused when you get to measure-theoretic probability and b) we can classify rvs further (which helps us with modelling and doing stuff in statistics).

Definitions

A **random variable** (rv) is a function whose domain is the sample space $\mathcal{S}$ and whose range is a subset of the real numbers $\mathbb{R}$

A **bernoulli** rv is one that maps into $\{0, 1\}$

Example

Example 3.3: Let $\mathcal{S}$ be the set of outcomes of how many gas pumps are active at each of two gas stations. Let's say the the first gas station has five pumps, and the second has 6 pumps. Then $\mathcal{S} = \{1, 2, 3, 4, 5\} \times \{1, 2, 3, 4, 5, 6\}$. Let X be the sum of the number of active gas pumps, let Y be the absolute difference between the number of gas pumps active, and let U be the maximum number of gas pumps active. We can write X as $X(s_1, s_2) = s_1 + s_2$.

What if our sample space was $\mathcal{S}$ was just a set of configurations? Can we do this? Is there a unique $\mathcal{S}$?

Definitions

We can classify rvs by what their range is.

A **discrete** rv is one whose range is finite or countably infinite.

A **continuous** rv is one whose a) range is a union of real intervals, and b) has no mass on any single number (i.e. if we call our rv X then for any number c, $P(X = c) = 0$)

Note: If you've taken a real analysis course before, you might know about the different types of infinity. For this class, we won't be doing analysis; all this helps to recognize certain probability models/distributions. Also, when we do stuff with discrete random variables we're typically using sums and differences. If we're doing stuff with continous rvs, then we're probably going to be using integrals and derivatives.

Examples

Our last gas station example was a collection of discrete random variables.

Examples of continuous rvs: tomorrows return for a stock, or a randomly selected persons height.