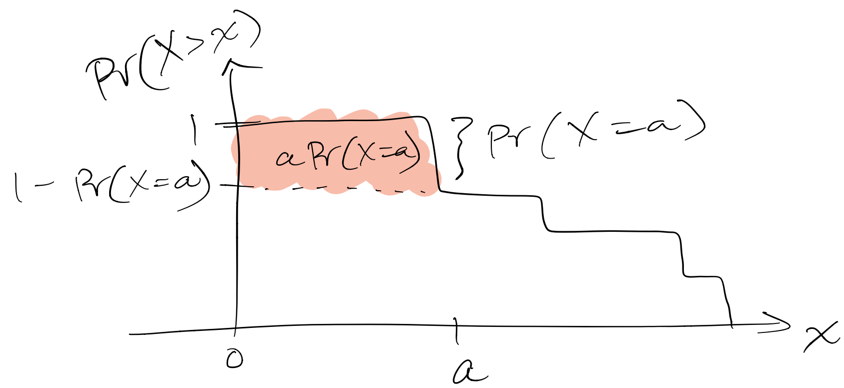
First we show that the expectation of a r.v. is equal to the area under its complementary cdf, i.e $\Pr(X > x) = 1 - F_X(x)$.

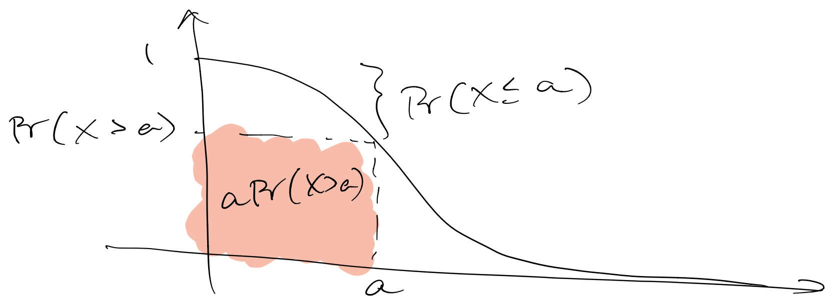
Suppose X is a non-negative discrete r.v. with complementary cdf:



The area of the shaded region is given by $a \cdot \Pr(X=a)$. In this way, we can calculate the total area under the curve:

$$\sum_x x \Pr(X=x) = E[X].$$

In the case of a continuous random variable, the complementary cdf looks like:



It's clear that the shaded area is less than the total area under the curve, i.e.

$$aPr(X > a) \leq E[X]$$

$$= \quad Pr(X > a) \leq \frac{E[X]}{a}$$

which gives us Markov's inequality.

10-27-2020