Mass Functions and Density Functions

September 11, 2009

Even though the cumulative distribution function is defined for every random variable, we will use the mass function for discrete random variable and the density function for continuous random variables more often. Indeed, we typically will introduce a random variable via one of these two functions.

1 Mass Functions

Definition 1. *A* **discrete random variable** *is a random variable whose range is finite or countably infinite.*

Definition 2. *The* **(probability) mass function** *of a discrete random variable X is*

$$f_X(x) = P\{X = x\}.$$

The mass function has two basic properties:

- $f_X(x) \geq 0$ for all x in the state space.
- $\sum_x f_X(x) = 1$.

Example 3. *Let's make independent tosses of a biased coin until we obtain a toss of heads. This number X is a random variable. Let p denote the probability of heads in any given toss. Then*

$$\begin{aligned}
f_X(1) &= P\{X = 1\} = P\{H\} = p \\
f_X(2) &= P\{X = 2\} = P\{TH\} = (1-p)p \\
f_X(3) &= P\{X = 3\} = P\{TTH\} = (1-p)^2 p \\
\vdots &= \vdots \\
f_X(x) &= P\{X = x\} = P\{T \cdots TH\} = (1-p)^{x-1} p
\end{aligned}$$

So, the probability mass function $f_X(x) = (1-p)^{x-1}p$. Because the terms in this mass function form a geometric sequence, X is called a **geometric random variable**. *Recall that a geometric sequence $c, cr, cr^2, \ldots, cr^n$ has sum*

$$c + cr + cr^2 + \cdots + cr^n = \frac{c(1 - r^{n+1})}{1 - r}.$$

If $|r| < 1$, then the infinite sum

$$c + cr + cr^2 + \cdots + cr^n + \cdots = \frac{c}{1 - r}.$$

Consequently, for positive integers a and b,

$$\begin{aligned} P\{a < X \le b\} &= \sum_{x=a+1}^{b} (1-p)^{x-1}p = (1-p)^a p + \cdots + (1-p)^{b-1}p \\ &= \frac{(1-p)^a p - (1-p)^b p}{1-(1-p)} = (1-p)^a - (1-p)^b \end{aligned}$$

Example 4. *An urn contains balls numbers 1 through n. From these r are removed at random. Let X be the maximum value of the removed balls.*

If they are removed with replacement, then the distribution function

$$F_X(x) = P\{X \le x\} = \left(\frac{x}{n}\right)^r$$

for x in the state space and constant in between these values. The mass function

$$f_x(x) = P\{x-1 < X \le x\} = F_X(x) - F_X(x-1) = \left(\frac{x}{n}\right)^r - \left(\frac{x-1}{n}\right)^r.$$

If they are removed without replacement, then the distribution function

$$F_X(x) = P\{X \le x\} = \frac{\binom{x}{r}}{\binom{n}{r}} = \frac{(x)_r}{(n)_r}$$

for x in the state space and constant in between these values. The mass function

$$\begin{aligned} f_x(x) &= P\{x-1 < X \le x\} = F_X(x) - F_X(x-1) \\ &= \frac{\binom{x}{r}}{\binom{n}{r}} - \frac{\binom{x-1}{r}}{\binom{n}{r}} = \frac{\binom{x-1}{r-1}}{\binom{n}{r}} \end{aligned}$$

by the Pascal triangle identity.

Exercise 5. *Give a mass function that is a reasonable definition of uniform on the integers $\{a, a+1, \ldots, b\}$ and describe the graph of the distribution function.*

Example 6 (mixtures)**.** *Let X_1 and X_2 be discrete random variables with respective mass functions f_1 and f_2. Flip a biased coin that lands heads with probability p. With heads, observe X_1 and with tails observe X_2. By the law of total probability, the mass function of the observation X is*

$$\begin{aligned} f_X(x) &= P\{X = x\} \\ &= P\{X = x|\text{coin lands heads}\}P\{\text{coin lands heads}\} + P\{X = x|\text{coin lands tails}\}P\{\text{coin lands tails}\} \\ &= pf_1(x) + (1-p)f_2(x). \end{aligned}$$

2 Density Functions

Definition 7. *Let X be a random variable whose distribution function F_X has a derivative. The function f_X satisfying*

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\ dt$$

is called the **probability density function** *and X is called a* **continuous random variable***.*

By the fundamental theorem of calculus, $F'_X(x) = f_X(x)$ We can compute compute probabilities using

$$P\{a < X \leq b\} = F_X(b) - F_X(a) = \int_a^b f_X(t)\ dt.$$

The density function has two basic properties:

- $f_X(x) \geq 0$ for all x in the state space.
- $\int_{-\infty}^{\infty} f_X(t)\ dt = 1$.

Return to the dart board example, letting X be the distance from the center of a dartboard having unit radius. Then,

$$P\{x < X \leq x + \Delta x\} = F_X(x + \Delta x) - F_X(x) \approx F'_X(x)\Delta x = 2x\Delta x$$

and X has density

$$f_X(s) = \begin{cases} 0 & \text{if } x < 0, \\ 2x & \text{if } 0 \leq x \leq 1, \\ 0 & \text{if } x > 1. \end{cases}$$

Exercise 8. *Let f_X be the density for a random variable X and pick a number x_0. Explain why $P\{X = x_0\} = 0$.*

Example 9 (Pareto random variable)**.** *X is said to be a Pareto random variable if for some $\alpha > 0$ and $\beta > 0$, its density*

$$f_X(x) = \begin{cases} 0 & \text{if } x < \alpha, \\ \frac{\beta\alpha^\beta}{x^{\beta+1}} & \text{if } x \geq \alpha. \end{cases}$$

Note that

$$\int_{-\infty}^{\infty} f_X(t)\ dt = \int_{\alpha}^{\infty} \frac{\beta\alpha^\beta}{t^{\beta+1}}\ dt = -\alpha^\beta t^{-\beta}\Big|_{\alpha}^{\infty} = 1 - 0 = 1.$$

Density functions do not deed to be bounded, for example, if we take

$$f_X(x) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{c}{\sqrt{x}} & \text{if } 0 \leq x \leq 1, \\ 0 & \text{if } x > 1. \end{cases}$$

Then

$$1 = \int_0^1 \frac{c}{\sqrt{t}}\ dt = 2c\sqrt{t}\Big|_0^1 = 2c.$$

So $c = 1/2$.

For $0 \le a < b \le 1$,

$$P\{a < X \le b\} = \int_0^1 \frac{1}{2\sqrt{t}}\, dt = \sqrt{t}\Big|_a^b = \sqrt{b} - \sqrt{a}.$$

Exercise 10. *Let X have density function*

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$

Verify that this is a bona fide density function. Compute the probability $P\{-1 \le X \le 1\}$. Indicate this on a graph of both the density and the distribution function.

Exercise 11. *Derive an analagous formula for mixtures with X_1 and X_2 continuous random variables.*