Topic 7

Random Variables and Distribution Functions

Mass Functions and Density Functions

Outline

Mass Functions

Density Functions

Mass Functions

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 $x \in S$, the state space.
  - Probabilities are non-negative.
- $\sum_x f_X(x) = 1$.
  - The collection

    $$C_x = \{\omega; X(\omega) = x\}$$

    for all $x \in S$, forms a partition of the probability space, Ω.

Mass Functions

Let's make tosses of a biased coin whose outcomes are independent. Let X denote the random variable that gives the number of tails before the first head and p denote the probability of heads in any given toss. Then

$$
\begin{aligned}
f_X(0) &= P\{X = 0\} &&= P\{H\} = p \\
f_X(1) &= P\{X = 1\} &&= P\{TH\} = (1-p)p \\
f_X(2) &= P\{X = 2\} &&= P\{TTH\} = (1-p)^2 p \\
\vdots & \qquad \vdots && \qquad \vdots \\
f_X(x) &= P\{X = x\} &&= P\{T \cdots TH\} = (1-p)^x p
\end{aligned}
$$

Because the terms in this mass function form a geometric sequence, X is called a geometric random variable

Mass Functions

Recall that a geometric sequence $c, cr, cr^2, \ldots, cr^n$ has sum

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

for $r \neq 1$. If $|r| < 1$, then $\lim_{n\to\infty} r^n = 0$ and thus s_n has a limit as $n \to \infty$. In this case, the infinite sum is the limit

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

Thus, $\sum_{x=0}^{\infty} f_X(x) = \sum_{x=0}^{\infty}(1-p)^x p = \frac{p}{1-(1-p)} = 1$.

$$\bar{F}(b) = P\{X > b\} = \sum_{x=b+1}^{\infty} f_X(x) = \sum_{x=b+1}^{\infty}(1-p)^x p = \frac{(1-p)^{b+1}p}{1-(1-p)} = (1-p)^{b+1}$$

and $F_X(b) = 1-(1-p)^{b+1}$ for $b = 0, 1, 2, \ldots$.

Mass Functions

Exercise. We use R to investigate a geometric random variable with $p = 1/4$. Enter the commands

```
> x<-c(0:10)      #creates a sequence from 0 to 10
> f<-dgeom(x,1/4) #gives the mass function for these values
> F<-pgeom(x,1/4) #gives the distribution function for these values
> data.frame(x,f,F)
```

- Check that the jumps in the cumulative distribution function $F_X(x) - F_X(x-1)$ is equal to the values of the mass function.
- Find
  1. $P\{X \leq 4\}$,
  2. $P\{2 < X \leq 5\}$, and
  3. $P\{X \geq 5\}$.

Density Functions

For X 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, the density function is the derivative of the distribution function.

$$f_X(x) = \lim_{\Delta x \to 0} \frac{F_X(x + \Delta x) - F_X(x)}{\Delta x} = F_X'(x).$$

In other words, if Δx is small,

$$F_X(x + \Delta x) - F_X(x) \approx f_X(x)\Delta x.$$

Density Functions

We can compute probabilities by evaluating definite integrals

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

The density function has two basic properties that mirror the properties of the mass function:

- $f_X(x) \ge 0$ for all x in the state space.
- $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$.

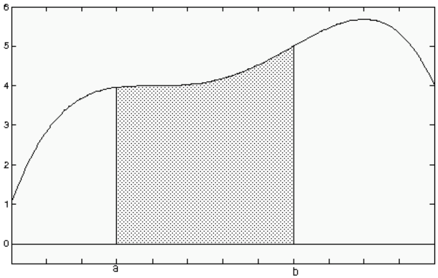

Figure: Integral of density function $f_X(x)$. Shaded region has area $P\{a < X \le b\}$.

Density Functions

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

Exercise. Let X be a continuous random variable with density

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

1. Sketch a graph of f_X. Notice that f_X is not bounded.
2. Show that f_X is a density function.
3. Found the distribution function F_X for X.
4. Use distribution function to find $P\{X \leq 1/4\}$ and $P\{1/4 \leq X \leq 1/2\}$.