Section 4.1-4.2 Cumulative distribution functions

Tuesday, October 11, 2016 12:44 PM

Section 4.1-4.2 Cu...

4 Continuous Distributions

Complementary reading from WMS: Chapter 4.

4.1 Introduction

RECALL: In Chapter 3, we focused on discrete random variables. A discrete random variable Y can assume a finite or (at most) a countable number of values. We also learned about probability mass functions (pmfs). These functions tell us what probabilities to assign to each of the support points in R (a countable set).

PREVIEW: **Continuous random variables** have support sets that are not countable. In fact, most often, the support set for a continuous random variable Y is an interval of real numbers; e.g., $R = \{y : 0 \leq y \leq 1\}$, $R = \{y : 0 < y < \infty\}$, $R = \{y : -\infty < y < \infty\}$, etc. Thus, probabilities of events involving continuous random variables must be assigned in a different way.

4.2 Cumulative distribution functions (cdf)

binomcdf

poisson cdf

TERMINOLOGY: The **(cumulative) distribution function (cdf)** of a random variable Y, denoted by $F_Y(y)$, is given by the probability

$$F_Y(y) = P(Y \leq y),$$

for all $-\infty < y < \infty$. Note that the cdf is defined for all $y \in \mathcal{R}$ (the set of all real numbers), not just for those values of $y \in R$ (the support of Y). Every random variable, discrete or continuous, has a cdf.

$y \to +\infty$ $\quad F_Y(y) = P(Y \leq y) \to 1$

$y_1 < y_2$ $\quad F_Y(y_1) \leq F_Y(y_2)$

$\quad P(Y \leq y_1) \quad ? \quad P(Y \leq y_2)$

$y \to -\infty$ $\quad F_Y(y) \overset{?}{\to} 0$

Example 4.1. Suppose that the random variable Y has pmf

$$p_Y(y) = \begin{cases} \frac{1}{6}(3-y), & y = 0, 1, 2 \\ 0, & \text{otherwise.} \end{cases}$$

$P_Y(0) = \frac{3-0}{6} = \frac{1}{2}$

$P_Y(1) = \frac{3-1}{6} = \frac{1}{3}$

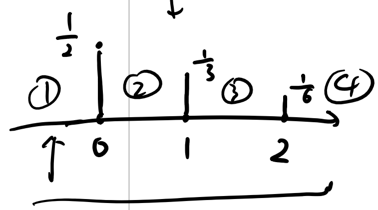
$P_Y(2) = \frac{3-2}{6} = \frac{1}{6}$

pmf

$P_Y(1) =$

$P_Y(2) = \frac{3-2}{6} = \frac{1}{6}$

pmf



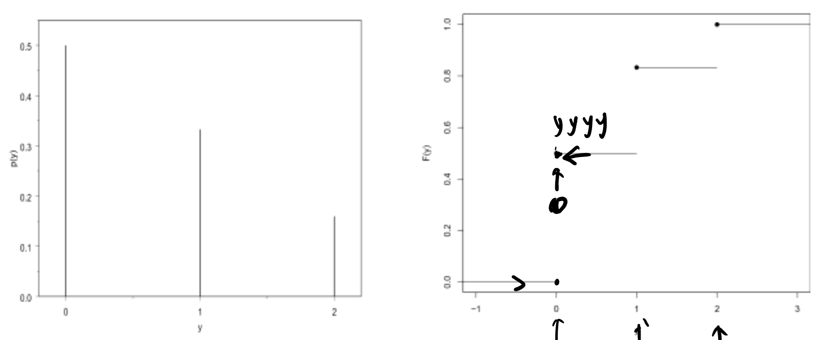
We now compute probabilities of the form $P(Y \le y)$: $-\infty < y < +\infty$

- for $y < 0$, $F_Y(y) = P(Y \le y) = 0$
- for $0 \le y < 1$, $F_Y(y) = P(Y \le y) = P(Y = 0) = \frac{3}{6}$
- for $1 \le y < 2$, $F_Y(y) = P(Y \le y) = P(Y = 0) + P(Y = 1) = \frac{3}{6} + \frac{2}{6} = \frac{5}{6}$
- for $y \ge 2$, $F_Y(y) = P(Y \le y) = P(Y = 0) + P(Y = 1) + P(Y = 2) = \frac{3}{6} + \frac{2}{6} + \frac{1}{6} = 1$.

Putting this all together, we have the cdf for Y:

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \frac{3}{6}, & 0 \le y < 1 \\ \frac{5}{6}, & 1 \le y < 2 \\ 1, & y \ge 2. \end{cases}$$ ← how you display a cdf.

It is instructive to plot the pmf of Y and the cdf of Y side by side.



pmf, $p_Y(y)$ cdf, $F_Y(y)$

Figure 4.5: *Probability mass function* $p_Y(y)$ *and cumulative distribution function* $F_Y(y)$ *in Example* 4.1.

- **PMF**
  - The height of the bar above y is the probability that Y assumes that value.
  - For any y not equal to $0, 1,$ or 2, $p_Y(y) = 0$.

- **CDF**
  - $F_Y(y)$ is a nondecreasing function.
  - $0 \leq F_Y(y) \leq 1$; this makes sense since $F_Y(y) = P(Y \leq y)$ is a probability!
  - The cdf $F_Y(y)$ in this example takes a "step" at the support points and stays constant otherwise. The height of the step at a particular point is equal to the probability associated with that point. □

CDF PROPERTIES: Let Y be a random variable (discrete or continuous) and suppose that $F_Y(y)$ is the cdf for Y. Then

(i) $F_Y(y)$ satisfies the following:

$$\lim_{y \to -\infty} F_Y(y) = 0 \quad \text{and} \quad \lim_{y \to +\infty} F_Y(y) = 1.$$

(ii) $F_Y(y)$ is a right continuous function; that is, for any real a,

$$\lim_{y \to a^+} F_Y(y) = F_Y(a).$$

(iii) $F_Y(y)$ is a non-decreasing function; that is,

$$y_1 \leq y_2 \Longrightarrow F_Y(y_1) \leq F_Y(y_2).$$

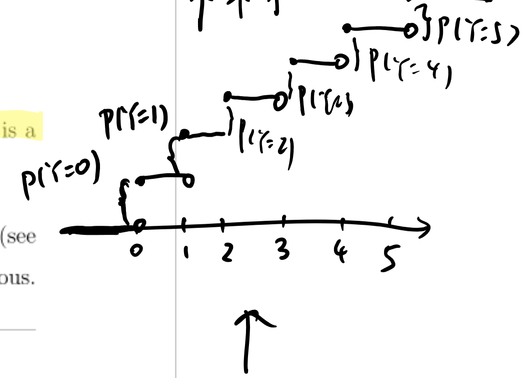
non-decreasing.

$= (.8)^5$

$Y \sim b(5, 0.2)$

$R = \{0, 1, 2, 3, 5\}$ 1

EXERCISE: Graph the cdf for (a) $Y \sim b(5, 0.2)$ and (b) $Y \sim \text{Poisson}(2)$.



4.3 Continuous random variables

TERMINOLOGY: A random variable Y is said to be **continuous** if its cdf $F_Y(y)$ is a continuous function of y.

REMARK: The cdfs associated with discrete random variables are step functions (see Example 4.1). Such functions are not continuous; however, they are still right continuous.