Fall 2018 STSCI 5080 Discussion 3 (9/7)

Reviews of Lectures 4 and 5

PMF

For a discrete random variable X, the probability mass function (pmf) $p(x)$ is defined by

$$p(x) = P(X = x)$$

for any real number x.

PDF

A function f on $\mathbb{R}$ (the set of real numbers) is a probability density function (pdf) if it is non-negative, i.e., $f(x) \geq 0$ for any real x, and integrates to 1, i.e.,

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Contiuous random variable

A random variable X is continuous if there exists a pdf f such that

$$P(X \in B) = \int_B f(x)dx$$

for any $B \subset \mathbb{R}$. The function f is then called the density of X.

CDF

For a (discrete/continuous) random variable X, the cumulative distribution function (cdf) $F(x)$ is defined by

$$F(x) = P(X \leq x)$$

for any real x.

Problems

1. (**Rice 2.5.1**) Suppose that X is a discrete random variable with $P(X = 0) = 0.25, P(X = 1) = 0.125, P(X = 2) = 0.125$, and $P(X = 3) = 0.5$. Compute the cdf of X and draw its graph.

2. (**Rice 2.5.5**) For any (discrete/continuous) random variable, show that $P(x < X \leq y) = F(y) - F(x)$ for any $x < y$.

3. (**Rice 2.5.11**) Let $X \sim Bin(n, p)$, and suppose that $p(n + 1)$ is an integer. For what value of k is $P(X = k)$ maximized? (Hint). Let $p(k)$ denote the pmf and evaluate the ratio $p(k)/p(k - 1)$.

4. (**Rice 2.5.17**) Suppose that in a sequence of independent Bernoulli trials, each with probability of success probability p, the number of failures up to the first success is counted. What is the pmf for this random variable?

5. (**Rice 2.5.34**) Let $f(x) = (1 + \alpha x)/2$ for $-1 \leq x \leq 1$ and $f(x) = 0$ otherwise, where $-1 \leq \alpha \leq 1$. Show that f is a pdf and find the corresponding cdf.

6. (**Rice 2.5.38**) Let T be an exponential random variable with parameter λ. Let X be a discrete random variable defined by

   $$X = \text{integer part of } T.$$

   Find the pmf of X.

Solutions

1. (**Rice 2.5.1**) The cdf is

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.25 & 0 \le x < 1 \\ 0.375 & 1 \le x < 2 \\ 0.5 & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}.$$

2. (**Rice 2.5.5**) We have

$$\{X \le y\} = \{X \le x\} \cap \{x < X \le y\},$$

where the two events on RHS are disjoint. Hence,

$$P(X \le y) = P(X \le x) + P(x < X \le y).$$

3. (**Rice 2.5.11**) For $k = 1, \dots, n$, we have

$$\frac{p(k)}{p(k-1)} = \frac{\binom{n}{k}}{\binom{n}{k-1}} \cdot \frac{p^k(1-p)^{n-k}}{p^{k-1}(1-p)^{n-k+1}} = \cdots = \frac{n-k+1}{k} \cdot \frac{p}{1-p}.$$

The RHS is larger than 1 if and only if $k < (n+1)p$ and equal to 1 if and only if $k = (n+1)p$. This means that

$$p(1) < \cdots < p((n+1)p - 1) = p((n+1)p) < \cdots < p(n).$$

So $p(k)$ is maximized at $k = (n+1)p - 1, (n+1)p$.

4. (**Rice 2.5.17**) Let $X_1, X_2, \dots,$ be independent Bernoulli trials with success probability p, and let Y denote the number of failures up to the first success. The random variable Y takes values in $\{0, 1, 2, \dots\}$, and for $k = 0, 1, 2, \dots,$

$$\begin{aligned} Y = k &\Leftrightarrow \text{the first } k \text{ trials are failures and the } (k+1)\text{-th trial is a success} \\ &\Leftrightarrow X_1 = \cdots X_k = 0 \text{ and } X_{k+1} = 1. \end{aligned}$$

Hence, the pmf of Y is

$$\begin{aligned} p(k) = P(Y = k) &= P(X_1 = \cdots X_k = 0, X_{k+1} = 1) \\ &= P(X_1 = 0) \cdots P(X_k = 0) P(X_{k+1} = 1) = (1-p)^k p \end{aligned}$$

for $k = 0, 1, 2, \dots$.

5. (**Rice 2.5.34**) By its definition, f is non-negative, and

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{1} f(x)dx = \frac{1}{2}\int_{-1}^{1}(1 + \alpha x)dx = \frac{1}{2}\left[1 + \frac{\alpha x^2}{2}\right]_{-1}^{1} = 1.$$

So f is a pdf. The corresponding cdf is

$$F(x) = \int_{-1}^{x} f(y)dy = \frac{1}{2}\left[1 + \frac{\alpha x^2}{2}\right]_{-1}^{x} = \frac{1}{2}\left\{x + 1 + \frac{\alpha}{2}(x^2 - 1)\right\}$$

for $-1 \le x \le 1$. For $x < -1$, $F(x) = 0$, and for $x > 1$, $F(x) = 1$.

6. (**Rice 2.5.38**) The pdf of T is

$$f(t) = \lambda e^{-\lambda t}$$

for $t \geq 0$. Note that for given $k = 0, 1, 2, \ldots,$

$$X = k \Leftrightarrow k \leq T < k + 1,$$

so that the pmf of X is

$$p(k) = P(X = k) = P(k \leq T < k + 1) = \lambda \int_{k}^{k+1} e^{-\lambda t} dt = e^{-k\lambda} - e^{-(k+1)\lambda} = e^{-\lambda k}(1 - e^{-\lambda})$$

for $k = 0, 1, 2, \ldots.$