Random variables - Class 8

September 19, 2013 Debdeep Pati

1 Random variables

1.1 Motivation

1. Sometimes events are not very convenient to use.
2. In an exam with ten problems of ten points each, the student may be only interested in how many problems he/she answered correctly.
3. We are interested in some functions of the outcome of an experiment.
4. Random variables: Real valued functions defined on sample space
5. Define functions and domain of functions

1.2 Example 1

1. Tossing 2 dice: $S = \{(i, j) : i, j = 1, 2, \cdots, 6\}$
2. What is the prob. that the sum of the two dice equals 7?
3. We may use events directly
4. Alternatively, let X denote the sum of the two dice. Then X is a random variable.
5. In fact, $X(i, j) = i + j$, $\forall\, (i, j) \in S$ (a function!!)
6. Possible values of X, $2, 3, 4, \cdots, 12$.
7. The prob. is given by $P(X = 7)$. $X = 7$ is actually $\{X = 7\}$, or $\{(i, j) \in S : X(i, j) = 7\}$, or $(i, j) \in S : i + j = 7$ - an event!

1.3 Example 2

1. A coin is tossed n times. What is P(# of heads is n tosses is 2)?
2. $S = \{(a_1, a_2, \cdots, a_n) : a_i = 0, 1\}$ (0: tails, 1: heads)
3. Let X denote the number of heads in the n experiments

4. X is a random variable, and can take any of the values $\{0, 1, \cdots, n\}$

5. Given any $(a_1, a_2, \cdots, a_n) \in S$, $X(a_1, a_2, \cdots, a_n) = \sum_{i=1}^{n} a_i$

1.4 Example 3

1. Flip a coin three times and record the flips.

2. Then $S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$.

3. Define a function X on S by X(s) = "the # of heads in the three flips." So $X(HHH) = 3, X(HTT) = 1, X(THT) = 1$, and $X(TTT) = 0$.

4. The range of the random variable X is $\{0, 1, 2, 3\}$. $P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 1/8 + 3/8 + 3/8 + 1/8 = 1$

1.5 Formal definition

1. What is a random variable?

2. In fact, a r.v. is just a function

3. Given S, if $X(\omega)$, $\forall\, \omega \in S$ is defined, then the function $X(\omega)$, or X is called a r.v.

4. A real valued function dened on the outcome of a probability experiment is called r.v.

5. X, Y, Z (capital letters) to denote r.v.s; x, y, z (lowercase letters) to represent outcomes that the r.v. may take. For example, we consider $P(X = x)$

6. In probability theory, we study the r.v.s (functions) using their prob of values but not their exact formulas.

1.6 Example

1. Three balls are randomly chosen from an urn containing 3 white, 3 red, and 5 black balls.

2. Suppose that we win 1 for each white ball selected and lose 1 for each red selected.

3. If we let X denote our total winning from the experiment, then X is a random variable taking on the possible values $0, \pm 1, \pm 2, \pm 3$ with respective probabilities, What is $P(X = k)$?

1.7 Benefits of using random variables

1. The concept of random variables helps us to formulate our problems better.
2. It allows us to apply all our knowledge of functions to the study of probability.
3. Also, it strips away real world differences to reveal situations that are probabilistically identical.
4. For instance, if X is the number of heads in three flips of a coin and Y is the number of girls in a family with three children, then X and Y are probabilistically identical (assuming the probability is 1/2 of each child being a girl) even though the physical situations they model are quite different.

1.8 Discrete r.v.

1. A random variable that can take at most a countable number of possible values is said to be discrete
2. $\{X(\omega) : \forall \omega \in S\}$ is finite or countably infinite
3. Examples $\mathbb{N}, \mathbb{Q}$ ($\mathbb{R}$: uncountably infinite!)
4. Countably Infinite: it is possible to make a list of the elements even though they are infinite. For instance the set of positive even numbers is countably infinite: $\{2, 4, 6, 8, \cdots\}$
5. The set of positive numbers of no more than two decimal digits is countably infinite $(0.01, 0.02, \cdots, 0.99, 1.00, 1.01, \cdots)$
6. An interval like $[0, 1]$ is uncountably iinfinite; it is impossible to make a list even an infinite list of the real numbers between zero and one.

1.8.1 Probability Mass Function (p.m.f.)

1. Given a discrete r.v. X, the p.m.f. of X is dened by $p(x) = P(X = x), \forall x$.
2. Example: the two dice example, $X(i, j) = i + j$
3. P(X = 2) = 1/36, P(X = 3) = 2/36, P(X = 3) = 2/36, P(X = 4) = 3/36, P(X = 5) = 4/36, P(X = 6) = 5/36, P(X = 7) = 6/36, P(X = 8) = 5/36, P(X = 9) = 4/36, P(X = 10) = 3/36, P(X = 11) = 2/36, P(X = 12) = 1/36

1.8.2 Properties of a Probability Mass Function (p.m.f.)

1. $p(x) \geq 0$
2. $\sum_{i:a_i \in S} a_i = 1$.
3. Example: The p.m.f. of a r.v. X is given by $p(i) = c\lambda^i/i!, i = 0, 1, \cdots$, where λ is some positive value.
4. Find $P(X = 0), P(X > 2)$

1.8.3 Cumulative Distribution Functions

1. Given a discrete r.v. X, the distribution (or c.d.f.) of X is dened by $F(x) = P(X \leq x)$, $x \in \mathbb{R}$.
2. Some properties: $F(-\infty) = 0, F(+\infty) = 1$, F is nondecreasing
3. F is right-continuous: $F(x+) = F(x)$, $\forall x \in \mathbb{R}$
4. If X is discrete, $F(\cdot)$ is a step function
5. Plot of cumulative distribution function: Let X be a random variable taking $-2.4, -1.4, -1.1, -1.0, -0.7, -0.5, 0.1, 0.2, 0.3, 0.8$ with equal probability $1/10$.
6. All prob questions can be answered in terms of c.d.f.! For example,
   (a) $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$.
   (b) $P(X < a) = \lim_{a_n \uparrow a} P(X \leq a_n) = \lim_{a_n \uparrow a} F(a_n) = F(a-)$ (left-limit)
   (c) $P(X = a) = P(X \leq a) - P(X < a) = F(a) - F(a-)$
   (d) If $F(\cdot)$ is continuous at a, $P(X = a) = F(a) - F(a-) = 0$, since $F(a-) = F(a) = F(a+)$.