ECS20 Handout — An introduction to Discrete Probability – Part II

1. A **random variable** X is a rule that assigns a numerical value to each outcome in the sample space S.

   The set of numerical values assigned by X is called the *range space*, denoted as $R_X = \{x_1, \ldots, x_t\}$.

   Example 1. A pair of *fair* dice is tossed. The sample space S consists of the 36 outcomes (i, j), where $1 \leq i, j \leq 6$. Let X assign to each outcome in S the sum of two numbers, then X is a random variable with range space $R_X = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. □

2. Let X be a random variable on the sample space S with range space $R_X = \{x_1, \ldots, x_t\}$. Then X induces "a function f" which assigns probabilities p_k to the value $x_k \in R_X$ as follows:

   $$p_k = \underline{p(X = x_k)} = \text{sum of probabilities of outcomes in } S \text{ whose value is } x_k.$$

   Notation: $\underline{p(X = x_k)} \equiv \underline{p(\{s \,|\, X(s) = x_k, s \in S\})}$.

   Properties of p_k:

   (1) $p_k \geq 0$

   (2) $\displaystyle\sum_k p_k = 1$

3. The set of ordered pairs (x_k, p_k) for $k = 1, \ldots, t$ is called the **distribution** of the random variable X.

4. If S is a finite sample space of equally likely outcomes, also called *equiprobable space*, and let (x_k, p_k) be the distribution of a random variable X on S with the range space $R_X = \{x_1, \ldots, x_t\}$. Then

   $$p_k = \frac{\text{number of outcomes in } S \text{ whose value is } x_k}{\text{number of outcomes in } S}. \quad (1)$$

5. Example 2. Continue Example 1, the pair of "fair" dice is meant that we have the equiprobable (sample) space S. By expression (1), then the distribution of X is as follows:

   $p(X = 2) = 1/36$, since there is one outcome (1,1) whose sum is 2.
   $p(X = 3) = 2/36$, since there are two outcomes (1,2) and (2,1) whose sum is 3.
   $p(X = 4) = 3/36$, since there are three outcomes (1,3), (2,2) and (3,1) whose sum is 4.
   $p(X = 5) = 4/36$, since there are four outcomes (1,4), (2,3), (3,2) and (4,1) whose sum is 5.

   and $p(X = 6) = 5/36$, $p(X = 7) = 6/36$, $p(X = 8) = 5/36$, $p(X = 9) = 4/36$,
   $p(X = 10) = 3/36$, $p(X = 11) = 2/36$, $p(X = 12) = 1/36$.

   The distribution is often written in a table as follows:

x_k	2	3	4	5	6	7	8	9	10	11	12
p_k	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/26	3/36	2/36	1/36

□

6. **Expectation (mean)** of X:

$$E(X) \equiv \sum_{k=1}^{t} x_k p_k$$

7. Example 3. Suppose a fair coin is tossed six times. Then the number of heads which can occur with their respective probabilities is as follows:

x_k	0	1	2	3	4	5	6
p_k	1/64	6/64	15/64	20/64	15/64	6/64	1/64

The expected number of heads is

$$E(X) = 0 \cdot (1/64) + 1 \cdot (6/64) + 2 \cdot (15/64) + 3 \cdot (20/64) + 4 \cdot (15/64) + 5 \cdot (6/64) + 6 \cdot (1/64) = 3.$$

This agrees with our intuition that we expect that half of the tosses to be heads. □

Example 4. Three horses a, b and c are in a race, and suppose that their respective probabilities of winning are $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{6}$. Let X denote the payoff function for the winning horse, and suppose X pays \$2, \$6 and \$9 according as a, b or c wins the race. The expected payoff for the race is

$$E(X) = X(a)p(a) + X(b)p(b) + X(c)p(c) = 2 \cdot \frac{1}{2} + 6 \cdot \frac{1}{3} + 9 \cdot \frac{1}{6} = 4.5.$$

□

8. The **variance** of X:

$$\text{Var}(X) \equiv \sum_{k=1}^{t} (x_k - E(X))^2 \, p_k$$

The **standard drivation** of X:

$$\sigma = \sqrt{\text{Var}(X)}$$

9. Example 5, continue Example 3, the variance of X is

$$\begin{aligned}\text{Var}(X) =&(0-3)^2 \cdot (1/64) + (1-3)^2 \cdot (6/64) + (2-3)^2 \cdot (15/64) + \\ &(3-3)^2 \cdot (20/64) + (4-3)^2 \cdot (15/64) + (5-3)^2 \cdot (6/64) + (6-3)^2 \cdot (1/64) \\ =&1.5.\end{aligned}$$

The standard derivation of heads is $\sigma = \sqrt{\text{Var}(X)} = \sqrt{1.5} \approx 1.225$. □

10. Chebyshev's inequality.

Let X be a random variable with expectation μ and standard derivation σ. Then for any positive number k, the probability that a value of X lies in the interval $[\mu - k\sigma, \mu + k\sigma]$ is at least $1 - \dfrac{1}{k^2}$. That is

$$p\left(|X - \mu| \leq k\sigma\right) \geq 1 - \frac{1}{k^2}.$$

Chebyshev's inequality justifies our intuition that for smaller σ, we would expect that X will be closer to μ.