MATH 461 LECTURE NOTE
WEEK 4

DAESUNG KIM

1. INDEPENDENT EVENTS (SEC 3.4)

Consider tossing two coins. The sample space is $S = \{(H,H),(H,T),(T,H),(T,T)\}$ and assume that the probability of each outcome is uniform, $\frac{1}{4}$. Let E (and F) be the events that the first coin (respectively the second coin) is heads. It is natural to expect that the two events are independent, that is, one event does not affect on the other. In fact, the conditional probabilities $\mathbb{P}(E|F)$ and $\mathbb{P}(F|E)$ are both $\frac{1}{2}$, which are the same as $\mathbb{P}(E)$ and $\mathbb{P}(F)$ "unconditional" probabilities.

Definition: Independent Events
Events E and F are independent if $\mathbb{P}(EF) = \mathbb{P}(E)\mathbb{P}(F)$.

Note also that this is effectively a consequence of the model. If we believe or data show that there is independence, the model has to incorporate this. If you had to construct a mathematical model for events E and F, as described below, would you assume that they were independent events? Explain your reasoning.

(i) E is the event that a businesswoman has blue eyes, and F is the event that her secretary has blue eyes.
(ii) E is the event that a man is under 6 feet tall, and F is the event that he weighs over 200 pounds.

Example 1. Suppose two cards are drawn at random from a 52-card deck. Let $E = \{\text{first card is black}\}$, $F = \{\text{second card is black}\}$. Are E and F independent?

Example 2. Suppose we roll 2 dice. Let E_1 denote the event that the sum of the dice is 6 and F denote the event that the first die equals 4. Are E_1 and F independent? Suppose that we let E_2 be the event that the sum of the dice equals 7. Is E_2 independent of F?

Example 3. (i) Suppose E and F are independent. Are E and F^c independent?
(ii) Suppose E and F are disjoint. Are E and F independent?

Independent Events: More than two events
Events E, F and G are independent if

$$\mathbb{P}(EF) = \mathbb{P}(E)\mathbb{P}(F), \qquad \mathbb{P}(FG) = \mathbb{P}(F)\mathbb{P}(G),$$
$$\mathbb{P}(GE) = \mathbb{P}(G)\mathbb{P}(E), \qquad \mathbb{P}(EFG) = \mathbb{P}(E)\mathbb{P}(F)\mathbb{P}(G).$$

In general, a sequence of events $E_1, E_2, \cdots, E_n$ are independent if, for every subsequence $E_{r_1}, E_{r_2}, \cdots, E_{r_k}$,

$$\mathbb{P}(E_{r_1}E_{r_2}\cdots E_{r_k}) = \mathbb{P}(E_{r_1})\mathbb{P}(E_{r_2})\cdots\mathbb{P}(E_{r_k})$$

Example 4. An infinite sequence of independent trials is to be performed. Each trial results in a success with probability p and a failure with probability $1-p$. What is the probability that

(i) at least 1 success occurs in the first n trials;
(ii) exactly k successes occur in the first n trials;
(iii) all trials result in successes?

2. RANDOM VARIABLES (SEC 4.1)

Suppose that our experiment consists of tossing 3 fair coins. Let X denote the number of heads that appear. For instance, if the outcome is (H,H,T), then the corresponding X is 2. That means, X is a function of outcomes in the sample space.

Definition: Random Variables
A random variable is a real-valued function defined on the sample space.

If X is 2, then possible outcomes are $(H,H,T),(H,T,H),(T,H,H)$. We use the notation $\{X=2\}=\{(H,H,T),(H,T,H),(T,H,H)\}$. The probability that $X=2$ is then defined by

$$\mathbb{P}(X=2)=\mathbb{P}(\{(H,H,T),(H,T,H),(T,H,H)\})=\frac{3}{8}.$$

Example 5. Three balls are to be randomly selected without replacement from an urn containing 20 balls numbered 1 through 20. Let X be the largest ball selected. Indicate what values it takes and with what probabilities.

Example 6. Independent trials consisting of the flipping of a coin having probability p of coming up heads are continually performed until either a head occurs or a total of n flips is made. Let X be the number of times the coin is flipped. Indicate what values it takes and with what probabilities.

Definition: Discrete Random Variables
A discrete random variable is a random variable that takes at most a countable number of possible values. For a discrete random variable, we define the probability mass function $p(a)$ by $p(a)=\mathbb{P}(X=a)$. If X takes the values $x_1,x_2,\cdots$, then

$$\begin{aligned}&p(x_i)\geq 0,\ \text{for } i=1,2,\cdots,\\&p(x)=0,\ \text{otherwise.}\end{aligned}$$

Example 7. Independent trials consisting of the flipping of a coin having probability p of coming up heads are continually performed until either a head occurs or a total of n flips is made. Let X be the number of times the coin is flipped. Indicate what values it takes and with what probabilities.

Example 8. Let $\lambda>0$. The probability mass function of a random variable X is given by

$$p(k)=\begin{cases}c\frac{\lambda^k}{k!} & \text{if } k=0,1,2,\cdots,\\ 0, & \text{otherwise}\end{cases}$$

for some c.

(i) Find c in terms of λ.
(ii) Find $\mathbb{P}(X=0)$.
(iii) Find $\mathbb{P}(X>2)$.

Distribution function
The (cumulative) distributtion function (d.f.) is

$$F(x)=\mathbb{P}(X\leq x)$$

for $-\infty<x<\infty$. We have the following properties:

(i) F is nondecreasing.
(ii) $\lim_{b\to\infty}F(b)=1$.
(iii) $\lim_{b\to-\infty}F(b)=0$.
(iv) F is right continuous. (That is, for any $x\in\mathbb{R}$ and a sequence $\{x_n\}$ with $x_n\geq x$ and $\lim_n x_{n\to\infty}=x$, we have $F(x)=\lim_{n\to\infty}F(x_n)$.)

Remark 9. If we define the cdf by $F(x)=\mathbb{P}(X<x)$, then $F(x)$ is left continuous. The right continuity of the cdf is useful when we determine if a given function is the cdf of some random variable.

Example 10. If the distribution function of X is given by

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{2}, & 0 \le x < 1, \\ \frac{3}{5}, & 1 \le x < 2, \\ \frac{4}{5}, & 2 \le x < 3, \\ \frac{9}{10}, & 3 \le x < 3.5, \\ 1, & x \ge 3.5, \end{cases}$$

calculate the probability mass function of X.

3. Expectation (Sec 4.3, 4, 6)

Definition
If X is a discrete random variable taking values x_i with probability $p(x_i)$, its expected value (or mean or expectation) is defined as

$$\mathbb{E}[X] = \sum_i x_i p(x_i).$$

One can think that the expected value of X is a weighted average of the possible values that X taks on.

Example 11. Find $\mathbb{E}[X]$, where X is the outcome when we roll a fair die.

The expectation can be understood as a long run average of values of X in n repeated experiments. That is,

$$\begin{aligned} \mathbb{E}[X] &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X^{(i)} \\ &= \lim_{n\to\infty} x_k \frac{\text{Number of } x_k \text{ in } n \text{ repeated experiments}}{n} \\ &= \sum_k x_k p(x_k). \end{aligned}$$

Example 12. We say that I_A is an indicator variable for the event A if

$$I_A = \begin{cases} 1, & \text{if } A \text{ occurs}, \\ 0, & \text{if } A^c \text{ occurs}. \end{cases}$$

Find $\mathbb{E}[I_A]$.

Example 13. If X is a discrete random variable and g is a function, then $g(X)$ is also a discrete random variable. Suppose X takes values $-1, 0$ and 1 with probabilities 0.2, 0.5 and 0.3. Let $Y = X^2$. Then, Y takes values either 0 or 1 with probabilities 0.5, 0.5. Thus, the expected value is $\mathbb{E}[Y] = 0.5$. Indeed, we have

$$\mathbb{E}[Y] = 1 \cdot 0.5 + 0 \cdot 0.5 = 1^2 \cdot 0.3 + 0^2 \cdot 0.5 + (-1)^2 \cdot 0.2.$$

Expectation of a function of RV
If X is a discrete random variable taking values x_i with probability $p(x_i)$, and g is a function, then

$$\mathbb{E}[g(X)] = \sum_i g(x_i) p(x_i).$$

In particular, for $g(x) = x^n$ and a positive integer n, we call $\mathbb{E}[g(X)] = \mathbb{E}[X^n]$ the n-th moment of X.

Proof. Let $Y = g(X)$ and $p_X(a), p_Y(a)$ be the probability mass functions of X and Y respectively. Then,

$$p_Y(y_j) = \sum_{i: g(x_i) = y_j} p_X(x_i).$$

$$\begin{aligned}\mathbb{E}[g(X)] = \mathbb{E}[Y] = \sum_j y_j p_Y(y_j) &= \sum_j y_j \sum_{i:g(x_i)=y_j} p_X(x_i) \\ &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p_X(x_i) \\ &= \sum_i g(x_i)p_X(x_i).\end{aligned}$$

□

Linearity of Expectation
Let X be a discrete random variable with probability mass function $p(a)$. If a and b are real numbers, then

$$\mathbb{E}[aX+b] = a\mathbb{E}[X] + b.$$

Proof. Using $g(x) = ax + b$ and the fact that $\sum_i p(x_i) = 1$, we have

$$\mathbb{E}[aX+b] = \sum_i (ax_i + b)p(x_i) = a\sum_i x_i p(x_i) + b\sum_i p(x_i) = a\mathbb{E}[X] + b.$$

□

References

[SR] Sheldon Ross, *A First Course in Probability*, 9th Edition, Pearson

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu