MATH 461 LECTURE NOTE
WEEK 5

DAESUNG KIM

1. VARIANCE (SEC 4.5)

Let X be a random variable with its distribution function F. If X is a discrete random variable, then the expectation is defined by $\mathbb{E}[X] = \sum_i x_i p(x_i)$. The expectation tells us what is the "typical" value of the random variable. It is natural to ask how the random variable is concentrated around the typical value. Consider the following three random variables

$$W = 0 \text{ with probability } 1,$$
$$Y = \begin{cases} -1 & \text{with probability } \frac{1}{2}, \\ 1 & \text{with probability } \frac{1}{2}, \end{cases}$$
$$Z = \begin{cases} -100 & \text{with probability } \frac{1}{2}, \\ 100 & \text{with probability } \frac{1}{2}. \end{cases}$$

One can see that $\mathbb{E}[W] = \mathbb{E}[Y] = \mathbb{E}[Z] = 0$.

Let X be a random variable. In order to measure how X is concentrated around the expectation, we consider the distance between the random variable and its expectation, $|X - \mathbb{E}[X]|$. Since the distance is again a random variable, we introduce the following quantity

$$\mathrm{Var}(X) = \mathbb{E}[|X - \mathbb{E}[X]|^2],$$

which is called the variance. The standard deviation of X is defined by

$$(X) = \sqrt{\mathrm{Var}(X)}.$$

Example 1. Suppose W, Y, Z are the random variables as in the previous example. One can directly compute the variances of W, Y, Z, that is,

$$\mathrm{Var}(W) = 0,$$
$$\mathrm{Var}(Y) = 1,$$
$$\mathrm{Var}(Z) = 10000.$$

Proposition 2. *Let X be a random variable with $\mathrm{Var}(X) < \infty$.*

(i) $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.
(ii) $(\mathbb{E}[X])^2 \leq \mathbb{E}[X^2]$.
(iii) *For real numbers a and b,* $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$.

Proof. (i) Note that if $X_1, X_2, \cdots, X_n$ are random variables, then

$$\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]$$

(without proof at this moment). Then,

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}(X - \mathbb{E}X)^2 \\ &= \mathbb{E}[X^2 - 2X\mathbb{E}X + (\mathbb{E}X)^2] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}X\mathbb{E}X + (\mathbb{E}X)^2 \\ &= \mathbb{E}[X^2] - (\mathbb{E}X)^2. \end{aligned}$$

(ii) It follows from $\mathrm{Var}(X) \geq 0$.

(iii) It is straightforward to see that

$$\begin{aligned}\mathrm{Var}(aX+b) &= \mathbb{E}[(aX+b)^2] - (\mathbb{E}[aX+b])^2 \\ &= a^2\mathbb{E}[X^2] + 2ab\mathbb{E}[X] + b^2 - (a^2(\mathbb{E}X)^2 + 2ab\mathbb{E}X + b^2) \\ &= a^2(\mathbb{E}[X^2] - (\mathbb{E}X)^2).\end{aligned}$$

□

Example 3. Calculate $\mathrm{Var}(X)$ if X represents the outcome when a fair die is rolled.

2. Bernoulli and Binomial Random Variables (Sec 4.6)

We consider an experiment with two outcomes, success with probability p and failure with probability $1-p$, for some $p \in [0,1]$. This is called *Bernoulli experiment*. Let X be a random variable such that it takes 1 if the outcome is success and 0 otherwise. We say X is a *Bernoulli random variable* with success probability p.

Example 4. Tossing a fair coin is a Bernoulli experiment with success probability $\frac{1}{2}$. If we define a randome variable X by $X = 1$ if the outcome is heads, and $X = 0$ otherwise, then X is a Bernoulli random variable.

Suppose we perform n independent Bernoulli experiment with probability p. Let X be the number of successes that occur in the n experiment. This random variable is called *Binomial random variable* with parameter (n,p) and denoted by $X \sim \mathrm{Bin}(n,p)$. In particular, if a Bernoulli random variable X is a special case of Binomial random varialbe $X \sim \mathrm{Bin}(1,p)$.

Proposition 5. *Let* $X \sim \mathrm{Bin}(n,p)$.

(i) $p(k) = \mathbb{P}(X=k) = \binom{n}{k}p^k(1-p)^{n-k}$.
(ii) $\mathbb{E}[X] = np$.
(iii) $\mathrm{Var}(X) = np(1-p)$.

Proof. (i) In the n experiment, we have $\binom{n}{k}$ different sequences of lenth n consisting of successes and failures. For each sequence, we have a proability $p^k(1-p)^{n-k}$.

(ii) Using the Binomial theorem and the fact that $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$, we have

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k} \\ &= n\sum_{k=1}^{n}\binom{n-1}{k-1}p^k(1-p)^{n-k} \\ &= np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{(n-1)-k} \\ &= np(p+(1-p))^{n-1} \\ &= np.\end{aligned}$$

(iii) Similarly, we have

$$\begin{aligned}\mathbb{E}[X(X-1)] &= \sum_{k=2}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k} \\ &= n(n-1)\sum_{k=2}^{n}\binom{n-2}{k-2}p^k(1-p)^{n-k} \\ &= n(n-1)p^2.\end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - (\mathbb{E}[X])^2 \\ &= n(n-1)p^2 + np - n^2p^2 \\ &= np(1-p).\end{aligned}$$

□

Example 6. It is known that screws produced by a certain company will be defective with probability .01, independently of one another. The company sells the screws in packages of 10 and offers a money-back guarantee that at least 2 of the 10 screws is defective. What proportion of packages sold must the company replace?

Proposition 7. *Let $n_1, n_2 \in \mathbb{N}$ and $p \in (0,1)$. Suppose $X_1 \sim \mathrm{Bin}(n_1, p)$ and $X_2 \sim \mathrm{Bin}(n_2, p)$ are independent. Then, $X = X_1 + X_2 \sim \mathrm{Bin}(n_1 + n_2, p)$.*

Proof. For $k = 0, 1, \cdots, n_1 + n_2$, we have

$$\begin{aligned}\mathbb{P}(X = k) &= \sum_{m=0}^{k} \mathbb{P}(X_1 = m, X_2 = k - m) \\ &= \sum_{m=0}^{k} \mathbb{P}(X_1 = m)\mathbb{P}(X_2 = k - m) \\ &= \sum_{m=0}^{k} \binom{n_1}{m} p^m (1-p)^{n_1 - m} \binom{n_2}{k-m} p^{k-m} (1-p)^{n_2 - k + m} \\ &= \left(\sum_{m=0}^{k} \binom{n_1}{m} \binom{n_2}{k-m} \right) p^k (1-p)^{n_1 + n_2 - k} \\ &= \binom{n_1 + n_2}{k} p^k (1-p)^{n_1 + n_2 - k}.\end{aligned}$$

□

Remark 8. In particular, if $X_i \sim \mathrm{Bin}(1, p)$ are independent Bernoulli random variables for $i = 1, 2, \cdots, n$, then $X = \sum_{i=1}^{n} X_i \sim \mathrm{Bin}(n, p)$. As a consequence,

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np.$$

Example 9. A communication system consists of n components, each of which will, independently, function with probability p. The total system will be able to operate effectively if at least one-half of its components function.

(i) For what values of p is a 5-component system more likely to operate effectively than a 3-component system?

(ii) Let $k \in \mathbb{N}$. Show that a $(2k+1)$-component system is better than a $(2k-1)$-component system if and only if $p > \frac{1}{2}$.

Proof. Let $X \sim \mathrm{Bin}(2k-1, p)$ and $Y \sim \mathrm{Bin}(2, p)$ be independent binomial random variables. Let A and B be the events that $(2k+1)$-component system ($(2k-1)$-component system) operates effectively. Then,

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(X + Y \geq k + 1) \\ &= \mathbb{P}(X \geq k + 1) + \mathbb{P}(X = k)\mathbb{P}(Y \geq 1) + \mathbb{P}(X = k - 1)\mathbb{P}(Y = 2)\end{aligned}$$

and $\mathbb{P}(B) = \mathbb{P}(X \geq k) = \mathbb{P}(X \geq k+1) + \mathbb{P}(X = k)$. Thus,

$$\begin{aligned}
\mathbb{P}(A) - \mathbb{P}(B) &= \mathbb{P}(X = k-1)\mathbb{P}(Y = 2) - \mathbb{P}(X = k)\mathbb{P}(Y = 0) \\
&= \binom{2k-1}{k-1} p^{k+1}(1-p)^k - \binom{2k-1}{k} p^k (1-p)^{k+1} \\
&= \binom{2k-1}{k-1} p^k (1-p)^k \left(p - (1-p)\right) > 0
\end{aligned}$$

if and only if $p > \frac{1}{2}$. □

3. Poisson Random Variables (Sec 4.7)

A random variable X is called a *Poisson random variable* with parameter $\lambda > 0$ if the probability mass function is given by

$$p(i) = \mathbb{P}(X = i) = e^{-\lambda} \frac{\lambda^i}{i!},$$

and denoted by $X \sim \text{Poisson}(\lambda)$. Note that

$$\sum_{i=0}^{\infty} p(i) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}. = e^{-\lambda} e^{\lambda} = 1.$$

Example 10. Suppose that the number of typographical errors on a single page of this book has a Poisson distribution with $\lambda = \frac{1}{2}$. Calculate the probability that there is at least one error on this page.

Other examples of random variables that usually obey Poisson law are

(i) the number of misprints on a page of a book;
(ii) the number of people in a community living to 100 years;
(iii) the number of customers entering a post office on a given day; etc.

Relation to Binomial Random Variables. Consider $Y \sim \text{Bin}(n, p)$ with $np = \lambda$. If $n \to \infty$ and $p \to 0$ maintaining $np = \lambda$, one can show that

$$\mathbb{P}(Y = k) \to \mathbb{P}(X = k)$$

for each $k = 0, 1, 2, \cdots$. Thus, a Poisson random variable can be thought of as a limit of Binomial random variables. Since $\mathbb{E}[Y] = np = \lambda$ and $\text{Var}(Y) = np(1-p) = \lambda(1-p) \to \lambda$, it is expected that $\mathbb{E}[X] = \text{Var}(X) = \lambda$.

Proposition 11. *If* $X \sim \text{Poisson}(\lambda)$*, then* $\mathbb{E}[X] = \text{Var}(X) = \lambda$.

Proof. The results follow from

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[X(X-1)] &= \sum_{k=2}^{\infty} k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} \\
&= \lambda^2.
\end{aligned}$$

□

Example 12. If you buy a lottery ticket in 50 lotteries, in each of which your chance of winning a prize is 1/100, what is the (approximate) probability that you will win a prize (a) at least once? (b) exactly once? (c) at least twice?

Example 13. The monthly worldwide average number of airplane crashes of commercial airlines is 3.5. What is the probability that there will be at least 2 such accidents in the next month? Explain your reasoning!

Poisson Paradigm. Recall that $\mathrm{Poisson}(\lambda) \approx \mathrm{Bin}(n, p)$ with $\lambda = np$, n large, p small. Suppose there are n event and the probability of each event is p_i. If p_i's are small and the trials are independent (or weakly dependent), then the number of these events that occur approximately has a Poisson distribution with mean $\lambda = \sum_{i=1}^{n} p_i$.

Example 14. A total of $2n$ people, consisting of n married couples, are randomly seated (all possible orderings being equally likely) at a round table. Let C_i denote the event that the members of couple i are seated next to each other, $i = 1, 2, \cdots, n$.

(i) Find $\mathbb{P}(C_i)$.
(ii) For $j \neq i$, find $\mathbb{P}(C_j | C_i)$.
(iii) Approximate the probability, for n large, that there are no married couples who are seated next to each other.

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