Lecture Notes for
Introduction to Probability and Statistics
© 2013, Sungkuk Lim
Hanyang University, Ansan, 425-791, Korea
calltimy@hanyang.ac.kr

Probability and Statistics / 확률과 통계

강의노트 08

통계 - 이산확률분포 2

60. 교재 중요 부분

Definition 3.1.1 (Discrete random variable). A random variable is discrete if it can assume at most a finite or a countably infinite number of possible values.

▶ 이산 - 셀 수 있고 (countable), 측정 가능(finite)

Definition 3.2.1 (Discrete density). Let X be a discrete random variable. The function f given by

$$f(x) = P[X = x]$$

for x real is called the density function for X.

▶ 확률밀도함수

Necessary and Sufficient Conditions for a Function to be a Discrete Density

1. $f(x) \geq 0$

2. $\sum_{\text{all } x} f(x) = 1$

Definition 3.2.2 (Cumulative distribution—discrete). Let X be a discrete random variable with density f. The cumulative distribution function for X, denoted by F, is defined by

$$F(x) = P[X \leq x] \quad \text{for } x \text{ real}$$

▶ CDF : 누적분포함수

Definition 3.3.1 (Expected value). Let X be a discrete random variable with density f. Let $H(X)$ be a random variable. The expected value of $H(X)$, denoted by $E[H(X)]$, is given by

$$E[H(X)] = \sum_{\text{all } x} H(x)f(x)$$

provided $\sum_{\text{all } x} |H(x)| f(x)$ is finite. Summation is over all values of X that occur with nonzero probability.

▶ 기대값 : (평균)

61. 평균, 기대값

Expected Value of X

$$E[X] = \sum_{\text{all } x} x f(x)$$

62. 주사위를 던질 때

1) 눈의 수(X)의 기댓값은?

E[X] = 1*1/6+2*1/6+...+6*1/6
= 3.5

2) 만약 던져서 나오는 눈의 100배만큼 돈을 받는다면 그때 한번 주사위를 던질 때 받을 수 있는 돈의 기댓값은?

H(X) = X * 100
H(1)=100, H(2)=200, H(3)=300,...
E[H(X)]=H(1)*1/6+H(2)*1/6+...
= (100+200+300+...+600)/6
= 350

63. 기댓값에 대한 기본 공식

Theorem 3.3.1 (Rules for expectation). Let X and Y be random variables and let c be any real number.

1. $E[c] = c$ (The expected value of any constant is that constant.)
2. $E[cX] = cE[X]$ (Constants can be factored from expectations.)
3. $E[X + Y] = E[X] + E[Y]$ (The expected value of a sum is equal to the sum of the expected values.)

64. 분산의 정의

Definition 3.3.2 (Variance). Let X be a random variable with mean μ. The variance of X, denoted by Var X, or σ^2, is given by

$$\text{Var } X = \sigma^2 = E[(X - \mu)^2]$$

65. 프로그램에서의 분산값 계산방법

Theorem 3.3.2 (Computational formula for σ^2)

$$\sigma^2 = \text{Var } X = E[X^2] - (E[X])^2$$

66. 예제 (p.56)

Example 3.3.4. To find σ_X^2 and σ_Y^2 for the variables of Example 3.3.3, we first use Table 3.6 to find $E[X^2]$ and $E[Y^2]$. We know that $E[X] = E[Y] = 70$.

$$\begin{aligned} E[X^2] &= \sum_{\text{all } x} x^2 f(x) \\ &= (40^2)(.01) + (60^2)(.04) + \cdots + (100^2)(.01) \\ &= 4926.4 \\ E[Y^2] &= \sum_{\text{all } y} y^2 f(y) \\ &= (40^2)(.40) + (60^2)(.05) + \cdots + (100^2)(.40) \\ &= 5630.32 \end{aligned}$$

$$\begin{aligned} \text{Var } X &= E[X^2] - (E[X])^2 \\ &= 4926.4 - 70^2 = 26.4 \\ \text{Var } Y &= E[Y^2] - (E[Y])^2 \\ &= 5630.32 - 70^2 = 730.32 \end{aligned}$$

67. 표준편차

Definition 3.3.3 (Standard deviation). Let X be a random variable with variance σ^2. The standard deviation of X, denoted by σ, is given by

$$\sigma = \sqrt{\text{Var } X} = \sqrt{\sigma^2}$$

68. 분산의 기본 성질

Theorem 3.3.3 (Rules for variance). Let X and Y be random variables and c any real number. Then

1. Var $c = 0$
2. Var $cX = c^2$ Var X
3. If X and Y are independent, then Var$(X + Y)$ = Var X + Var Y

(Two variables are independent if knowledge of the value assumed by one gives no clue to the value assumed by the other.)

Example 3.3.6. Let X and Y be independent with $\sigma_X^2 = 9$ and $\sigma_Y^2 = 3$. Then

$$\begin{aligned} \text{Var}[4X - 2Y + 6] &= \text{Var}[4X] + \text{Var}[-2Y] + \text{Var } 6 && \text{Rule 3} \\ &= 16 \text{ Var } X + 4 \text{ Var } Y + \text{Var } 6 && \text{Rule 2} \\ &= 16 \text{ Var } X + 4 \text{ Var } Y + 0 && \text{Rule 1} \\ &= 16(9) + 4(3) = 156 \end{aligned}$$

E[(X-m)^2] 계산

1. m 계산

```
xm = 0
   // read x1 to xn (반복 n회)
   // loop start
   i == 1 to n
   xm += xi
   // loop end
m = xm / n
```

계산량 : 덧셈n회, 나눗셈1회

2. E[(X-m)^2] 계산

```
em = 0
   // loop x1 to xn (반복n회)
   em += (xi-m) * (xi-m)
   // loop end
em = em/n
```

계산량 : 뺄셈n회, 곱셈n회, 나눗셈1 or n회

E[X^2] - (E[X])^2 계산

1. m 계산 (동일)

계산량 : 덧셈n회, 나눗셈1회

2. E[X^2] 계산

```
em = 0
// loop x1 to xn
em += xi * xi
// loop end
em = em/n
```

계산량 : 곱셈n회, 나눗셈1 or n 회

3. (E[X])^2 계산

```
m*m
```

계산량 : 곱셈1회

추가 생각할 것

▶ 저장공간
▶ 정수형 계산 및 실수형 계산
▶ 각 시행의 확률값이 다르면?