4장 확률변수의 기댓값

자료의 중심척도로서 **평균**과 산포의 척도로서 **분산**을 생각하듯이 확률분포의 특징을 측정하는 중심위치와 산포도는 평균(기댓값)과 분산이다. 평균이란 확률분포의 무게중심의 위치이다.

1. 확률변수의 기댓값

확률변수의 기댓값(expected value)은 각 사건이 일어날 때의 가치와 그 사건이 일어날 확률을 곱한 것을 표본공간 전체 사건에 대해 합한 값이다. 이것은 어떤 확률적 사건에 대한 평균(mean)의 의미로 생각할 수 있다.

[정의] (1) 이산확률변수 X의 확률질량함수가 $f(x)$일 때, X의 **기댓값** 또는 **평균**은

$$E(X) = \sum_x x\, f(x)$$

(2) 연속확률변수 X의 확률밀도함수가 $f(x)$일 때, X의 **기댓값**은

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

[예제1] (3장 예제1) 하나의 동전을 세 번 던지는 시행에서 앞면의 횟수를 확률변수 X라고 할 때, X의 기댓값을 구하여라.

X	0	1	2	3	합
$P(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

$$E(X) = 0\times\frac{1}{8} + 1\times\frac{3}{8} + 2\times\frac{3}{8} + 3\times\frac{1}{8} = \frac{3}{2}$$

```
library(prob)
# tosscoin()을 이용하여 표본공간을 생성
S <- tosscoin(3, makespace=T) ; S
# 앞면의 개수를 세는 함수를 정의
Hcount <- function(x) sum(x == "H")
```

```
# 확률변수 X의 정의 ⇒ apply( ) 함수를 행별로 적용
( X <- apply(S, 1, Hcount) )
( freq <- table(X) )
( prob <- freq/length(X) )
( prob <- freq/nrow(S) )
X <- c(0,1,2,3)
PX <- c(0.125, 0.375, 0.375, 0.125)
EX <- sum(X*PX)
```

참고 # example_1

```
library(prob)
# tosscoin()을 이용하여 표본공간을 생성
S <- tosscoin(3, makespace=T) ; S
# 앞면의 개수를 세는 함수를 정의
Hcount <- function(x) sum(x == "H")
# 확률변수 X의 정의 ⇒ apply( ) 함수를 행별로 적용
( X <- apply(S, 1, Hcount) )
# addrv( )을 이용하여 S에 X를 첨가함
( S1 <- addrv( S, X=X, name=c("X") ) )
# marginal( )을 이용하여 X의 주변확률분포를 구함
( S2 <- marginal( S1, var=c("X") ) )
# 기댓값을 구한다.
( S2$X*S2$probs )
( EX <- sum(S2$X*S2$probs) )
( EX <- sum( S2[,1]*S2[,2] ) )
# 분수로 표현하기
library(MASS)
as.fractions( EX )
```

[예제2] 연속확률변수 X의 확률밀도함수 f가

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

라고 할 때, X의 기댓값을 구하여라.

① X의 표본공간 $S = \{x \mid 0 < x < 1\}$에서 $f(x)$의 면적은 1인가?

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 2x\,dx = \left[2 \times \frac{1}{2}x^2\right]_0^1 = 1 - 0 = 1$$

② $X > 1/2$일 확률을 구하라.

$$P(X > 1/2) = \int_0^{1/2} 2x\,dx = \left[2 \times \frac{1}{2}x^2\right]_0^{1/2} = (1/2)^2 - 0 = 1/4$$

③ 누적분포함수(cdf)

$0 < x < 1$ 일 때,

$$F(x) = P(X \le x) = \int_0^x f(t)\,dt = \int_0^x 2t\,dt = \left[2 \times \frac{1}{2}t^2\right]_0^x = x^2$$

따라서

$$F(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \le x \le 1 \\ 1, & x \ge 1 \end{cases}$$

⑤ $\frac{d}{dx}F(x) = f(x)$ 인가?

$$\frac{d}{dx}F(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & otherwise \end{cases} = f(x)$$

④ 평균과 분산

$$\mu = E(X) = \int_0^1 x f(x)\,dx = \int_0^1 2x^2\,dx = 2\left[\frac{1}{3}x^3\right]_0^1 = \frac{2}{3}(1^3 - 0^3) = \frac{2}{3}$$

$$\sigma^2 = Var(X) = E(X^2) - \mu^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

```
# 확률밀도함수를 정의하고 x를 곱하여 객체에 할당함
xfx <- function(x)  x*2*x
# 기댓값을 구한다.
( EX <- integrate(xfx, 0, 1) )
# 분수로 표현하기
library(MASS)
```

as.fractions(0.6666667) # as.fractions(EX) ?

[성질] X가 확률변수일 때, X의 함수로 표현되는 모든 $Y=g(X)$도 확률변수가 된다.

$Y=g(X)$의 값이 X의 값과는 달라지지만, $Y=g(X)$의 각각 값이 발생할 확률은 본래 확률변수 X가 발생할 확률과 동일하다는 것이다.

[정리] (1) 이산확률변수 X의 확률질량함수가 $f(x)$일 때, $Y=g(X)$의 **기댓값** 또는 **평균**은

$$E(Y)=E[g(X)]=\sum_x g(x)f(x)$$

(2) 연속확률변수 X의 확률밀도함수가 $f(x)$일 때, $Y=g(x)$의 기댓값은

$$E(Y)=E[g(X)]=\int_{-\infty}^{\infty} g(x)f(x)dx$$

[예제4] 확률변수 X는 두 주사위 눈의 합이라 하고 $Y=g(X)=100X$라 하면, Y는 상금액수를 의미한다고 할 수 있다. 이때 평균적으로 기대할 수 있는 상금액수를 구하여라.

```
# library( ) 함수를 이용하여 패키지 ‘prob’를 불러오기
library(prob)
# rolldie( )를 이용하여 표본공간 S를 구하기
( S <- rolldie(2, makespace=TRUE) )
# 두 주사위 눈의 합에 100을 곱한 확률변수 Y를 정의하자.
Y <- 100*(S$X1 + S$X2)
# addrv( )을 이용하여, 표본공간 S에 확률변수 Y를 추가하자.
( S1 <- addrv(S, Y=Y) )
( S1 <- addrv(S, Y=100*(X1 + X2) ) )
# 확률변수 Y의 기댓값을 구하기
( EY <- sum(S1$Y*S1$probs) )
( EY <- sum(S1[ ,3]*S1[ ,4]) )
```

[정리] 기댓값의 성질

확률변수 X, Y와 상수 a, b에 대하여

(1) $E(a) = a$

(2) $E(aX+b) = aE(X)+b$

(3) $E(aX+bY) = aE(X)+bE(Y)$

두 확률변수의 합이나 차의 기댓값은 각 변수의 기댓값들의 합이나 차와 같다.

4.3 확률변수의 분산

확률변수 X의 기댓값으로는 단지 분포의 중심 위치에 대한 정보만을 얻을 수 있으며 분포의 흩어진 정도를 판별할 수는 없다. 실제로 평균, 즉 분포의 중심 위치는 같으면서도 분포의 모습이 서로 다른 경우들이 많이 있기 때문에 확률분포가 중심위치 주변에 어느 정도로 밀집되어 있는가를 파악할 수 있는 적절한 측도가 필요하다.
분산은 평균으로부터 확률분포의 흩어진 정도를 측정하는 값이다. 따라서 확률분포에서 분산이 크다는 것은 그 확률분포가 넓게 흩어져 있음을 의미하고, 분산이 작다는 것은 확률분포가 평균(기댓값)에 가깝게 분포하고 있다는 것을 의미한다.

[정의] 확률변수의 분산

(1) 이산확률변수 X의 기댓값을 $\mu = E(X)$라고 할 때, X의 분산과 표준편차는

$$V(X) = E[(X-\mu)^2] = \sum_x (x-\mu)^2 P(X=x),$$

$$\sigma = \sigma(X) = \sqrt{V(X)}$$

(2) 연속확률변수 X의 확률밀도함수가 $f(x)$일 때, X의 분산은 다음과 같다.

$$V(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

Note. $(X-\mu)^2 \geq 0$ 이므로 $V(X) = E[(X-\mu)^2] \geq 0$ 이다.

Note. 기댓값의 성질을 이용하여, 다음과 같은 분산의 간편한 계산법을 얻을 수 있다.

$$V(X)=E[(X-\mu)^2]=E(X^2-2\mu X+\mu^2)$$
$$=E(X^2)-2\mu E(X)+\mu^2$$
$$=E(X^2)-\mu^2$$

[예제5] 확률변수 X가 두 주사위의 눈의 합을 나타낸다고 할 때, X의 분산을 구하여라.

1단계: 표본공간과 확률변수 X의 확률분포
2단계: 기댓값을 $\mu=E(X)$ 을 구하자.
3단계: $V(X)=E(X^2)-\mu^2$ 을 계산하자.

```
# example_5
# library( ) 함수를 이용하여 패키지 'prob'를 불러오기
library(prob)
# rolldie( )를 이용하여, 표본공간 S를 구하자.
( S <- rolldie(2, makespace=T) )
# 두 주사위 눈의 합인 확률변수 X를 정의하자.
X <- S$X1 + S$X2
# addrv( )을 이용하여, 표본공간 S에 확률변수 X를 추가하자.
( S1 <- addrv(S, X=X) )
( S1 <- addrv(S, X=(X1 + X2) ) )
# 확률변수 X의 기댓값을 구하기
( EX <- sum(S1$X*S1$probs) )
( EX <- sum(S1[ ,3]*S1[ ,4]) )
# 확률변수 X의 분산을 구하자.
( VX <- sum( ( (S1$X)-EX )^2*S1$probs ) )   # 분산의 정의
( VX <- sum(S1$X^2*S1$probs) - EX^2)
# as.fractions(  )을 이용하여 분수로 나타내기
library(MASS)
as.fractions(5.833333)
```

[예제6] 예제2에서 연속확률변수 X의 확률밀도함수 f가

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

일 때, X의 분산을 구하여라.

```
# example_6
# 확률밀도함수에 x를 곱하여 객체에 할당함
xfx <- function(x)  x*2*x
# X의 기댓값을 구한다.
( EX <- integrate(xfx, 0, 1) )
# X의 분산을 구한다.
x2fx <- function(x)  x^2*2*x
( integrate(x2fx, 0, 1) )
( VX <- 0.5 -(0.6666667)^2 )
( VX <- integrate(x2fx, 0, 1) -(0.6666667)^2 )
# 분수로 표현하기
library(MASS)
as.fractions(0.05555551)
```

[정리] 분산의 성질

확률변수 X, Y와 상수 a, b에 대하여

(1) $V(a) = 0$

(2) $V(aX) = a^2 V(X)$

(3) $V(aX + b) = a^2 V(X)$

(4) $V(aX + bY) = a^2 V(X) + b^2 V(Y)$ 단, X와 Y는 서로 독립이다.

[예제7] 키와 몸무게처럼 다른 척도를 가진 자료들이거나 범위가 매우 다른 자료들을 비교하는데 있어서 유용하게 사용되는 개념이 자료의 표준화이다. 확률변수 X에서 그 확률변수의 기댓값을 뺀 것을 표준편차로 나눈 변수를 표준화 변수 Z라고 하고, -4.0에서 $+4.0$ 사이의 값을 갖으며, 원점수(X)를 Z점수로 변환하면 두 개의 다른 분포에서 나온 점수들을 비교할 때 매우 용이하다. 확률변수 X의 평균이 μ, 분산이 σ^2일 때, $Z = \dfrac{X - m}{\sigma}$의 평균은 0 이고, 분산은 1 이다.

풀이: $E(X)=\mu$ 이므로

$$E(Z)=E\left(\frac{X-m}{\sigma}\right)=\frac{1}{\sigma}E(X)-\frac{m}{\sigma}=0$$

$V(X)=\sigma^2$, $V(X-m)=V(X)$ 이므로

$$V(Z)=V\left(\frac{X-m}{\sigma}\right)=\frac{1}{\sigma^2}V(X)=1$$