

1. 참, 거짓 판별하는 문제(이유 없이 진리값만 적으시면 됩니다)

2. 전확률의 법칙, 베이즈의 법칙을 이용한 확률 구하기

3. 확률변수 $Z=h(X)$ 또는 $Z=h(X, Y)$ 의 확률밀도함수 및 평균 구하기

3-1

Example 2.33 흰 공이 5개 검은 공이 4개가 있는 주머니에서 동시에 임의로 3개의 공을 꺼내서 그 속에 포함된 흰 공의 개수를 X 라 하자.

(1) 확률변수 X 의 확률질량함수를 구하시오.

(2) 확률변수 X 의 평균을 구하시오.

solution. X 의 치공간은 $R_X = \{0, 1, 2, 3\}$ 으로 X 는 이산확률변수이다.

(1) 확률변수 X 의 확률질량함수:

$$f_X(x) = P[X = x] = \frac{\binom{5}{x} \binom{4}{3-x}}{\binom{9}{3}} \cdot I_{\{0, 1, 2, 3\}}(x)$$

$$= \begin{cases} \frac{\binom{5}{x} \binom{4}{3-x}}{\binom{9}{3}}, & x \in R_X = \{0, 1, 2, 3\} \\ 0, & \text{otherwise} \end{cases}$$

x	0	1	2	3
$P[X = x]$	$\frac{4}{84}$	$\frac{30}{84}$	$\frac{40}{84}$	$\frac{10}{84}$

(2) 확률변수 X 의 평균:

$$E[X] = \sum_{x \in R} x \cdot f_X(x) = \sum_{x=0}^3 x \cdot \frac{\binom{5}{x} \binom{4}{3-x}}{\binom{9}{3}}$$

$$= 0 \times \frac{1 \cdot 4}{84} + 1 \times \frac{5 \cdot 6}{84} + 2 \times \frac{10 \cdot 4}{84} + 3 \times \frac{10 \cdot 1}{84}$$

$$= \frac{140}{84} = \frac{5}{3}$$

3-2

Example 2.45 연속 확률변수 X 의 확률밀도함수가 $f_X(x) = \begin{cases} \frac{1}{3}, & -2 < x < 1 \\ 0, & \text{otherwise} \end{cases}$ 이라 하자. Theorem 2.44를 이용하여 $Y = X^2$ 의 평균 $E[Y] = E[X^2]$ 을 구하시오. (Example 2.43)

solution.

$$E(Y) = E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_{-2}^1 x^2 \cdot \frac{1}{3} dx = \left[\frac{1}{9} x^3 \right]_{-2}^1 = \frac{1 - (-8)}{9} = 1$$

Example 3.5 한 주머니에 흰 공이 3개, 검은 공이 2개, 붉은 공이 3개가 들어 있다.
이 주머니에서 2개의 공을 뽑을 때 X 를 흰 공의 개수, Y 를 검은 공의 개수라 한다.

- (1) 확률벡터 (X, Y) 의 치공간을 구하시오.
- (2) 확률벡터 (X, Y) 의 결합확률질량함수를 구하시오.

solution. 흰 공(W_1, W_2, W_3), 검은 공(B_1, B_2), 붉은 공(R_1, R_2)
표본공간 $\Omega = \{W_1W_2, W_1W_3, W_2W_3, W_1B_1, \dots, R_1R_2\}$

- (0) X, Y 의 치공간과 확률질량함수

$$R_X = \{0, 1, 2\}, \quad R_Y = \{0, 1, 2\}$$

$$f_X(x) = \begin{cases} \frac{\binom{3}{x}\binom{2}{2-x}}{\binom{5}{2}}, & x \in R_X \\ 0, & \text{otherwise} \end{cases} \quad f_Y(y) = \begin{cases} \frac{\binom{2}{y}\binom{6}{2-y}}{\binom{8}{2}}, & y \in R_Y \\ 0, & \text{otherwise} \end{cases}$$

- (1) (X, Y) 의 치공간

$$R_{X \times Y} = \{(0,0), (0,1), (0,2), (1,0), (1,1), (2,0)\}$$

여기서 $R_{X \times Y} \subseteq R_X \times R_Y$ 임을 알 수 있다.

- (2) (X, Y) 의 결합확률질량함수 $f_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \in \mathbb{R}^2$ 에 대하여

$$f_{X,Y}(x,y) = P[X=x, Y=y]$$

$$= \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}}, \quad (x, y) \in R_{X \times Y}$$

$$= \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}} \cdot I_{R_{X \times Y}}(x, y)$$

$x \setminus y$	0	1	2
0	$\frac{3}{28}$	$\frac{6}{28}$	$\frac{1}{28}$
1	$\frac{9}{28}$	$\frac{6}{28}$	0
2	$\frac{1}{28}$	0	0

4. 확률벡터에서 조건부 확률질량(밀도)함수 및 조건부 기댓값 구하기

Example 3.47 연속확률벡터 (X, Y) 의 결합확률밀도함수가

$$f_{X,Y}(x,y) = \begin{cases} \frac{2}{5}(x+4y), & 0 < x < 1, 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

일 때, 다음 물음에 답하시오.(Example 3.17)

- (1) $Y = y$ 에 대한 X 의 조건부 확률밀도함수를 구하시오.
- (2) (1)을 이용하여 확률 $P[X \leq \frac{1}{2} \mid Y = \frac{1}{2}]$ 를 구하시오.

solution. Example 3.17에서 다음과 같이 X 와 Y 의 주변확률밀도함수를 얻었다.

$$f_X(x) = \begin{cases} \frac{2}{3}(x+2), & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{3}(8y+1), & 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

- (1) $Y = y$ 에 대한 X 의 조건부 확률밀도함수: $0 < y < 1$ 일 때,

$$\frac{f_{X,Y}(x|y)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{\frac{2}{5}(x+4y)}{\frac{1}{3}(8y+1)}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{2x+8y}{1+8y}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

(2) (1)을 이용하여 확률 $P[X \leq \frac{1}{2} | Y = \frac{1}{2}]$ 를 구하시오.

$$\begin{aligned} P\left[X \leq \frac{1}{2} \mid Y = \frac{1}{2}\right] &= \int_{-\infty}^{1/2} f_{XY}(x|y=\frac{1}{2}) dx \\ &= \int_0^{1/2} \frac{2x+4}{5} dx \\ &= \left[\frac{x^2 + 4x}{5} \right]_0^{1/2} = \frac{9}{20} \end{aligned}$$

5. 감마분포와 푸아송분포 사이의 관계를 이용한 확률 구하기

5장 연속확률분포
└ 감마분포

Example $X \sim \text{Pois}(15)$ $T \sim G_a(1, 4)$
 $X_{\text{평}} \sim \text{Pois}(\frac{15}{4})$ $G = T \sim \text{Exp}(4)$

1. 어떤 가게에서 만 월 이상 구매하는 고객들의 수는 시간당 15 명의 비율로 도착하는 푸아송분포를 따른다고 하자. 만 월 이상 구매하는 첫 고객이 생길 때 까지 가게 종업원이 기다려야 하는 시간이 4분 이상일 확률을 구하시오.

$$P[T \geq 4] = \int_4^\infty \frac{1}{4} e^{-t/4} dt$$

2. 어떤 햄버거 가게에는 시간당 평균 30명의 손님이 오는 비율로 푸아송분포를 따른다고 한다. $X \sim \text{Pois}(30) = X_{\text{평}} \sim \text{Pois}(\frac{30}{4})$

- (1) 처음 두 명의 손님이 방문하기까지 가게 주인이 기다리는 시간이 7분 이상일 확률을 구하시오. $T \sim G_a(2, 2)$
- (2) 처음 두 명의 손님이 방문하기까지 가게 주인이 기다리는 시간의 평균과 분산을 구하시오. $E(X) = \alpha\beta = 4$

$$P[T \geq 7] = \int_7^\infty \frac{1}{7} e^{-t/7} dt$$

$$V(X) = \alpha\beta^2 = 8$$

20/20

5.2 # 14. 우리나라 동남부 지역은 매년 2건의 비율로 개간이 일어나며, 각각 발생 횟수는 폭발성 과정에 따른다고 한다.

지점이 일어나는 횟수를 X 라하자

$$\begin{matrix} X \sim P_{0.1}(2) \\ (0,1) \\ 10\% \end{matrix}$$

(1) $t=0$ 이후 3번 째 사건이 발생할 때까지 걸리는 시간이 예상 확률 분포를 구하라.

T 를 지점 간 측후 ~~증가~~ 증가하기 전에 걸리는 시간이라니라.
세 번째

$$T \sim Ga(3, \frac{1}{2})$$

$$\therefore f_T(x) = \begin{cases} \frac{x^2 e^{-\frac{x}{2}}}{\Gamma(3) \cdot \frac{1}{2}} = 4x^2 e^{-2x}, & x > 0 \\ 0, & \text{o/w} \end{cases}$$

(2) 3번 째 사건이 $t=0.5$ 과 $t=1.5$ 사이에 발생할 확률을 구하라.

$$P[0.5 < T < 1.5] = \int_{0.5}^{1.5} 4x^2 e^{-2x} dx = [-(2x^2 + 2x)]_{0.5}^{1.5}$$

$$\approx 0.4965$$

6. 모평균, 모비율, 모분산, 모평균 차의 점추정 및 신뢰구간 모평균(분산이 알려진 경우)

Example 7.34 어떤 전자 제품의 수명을 X 라 하자. X 가 균사적으로 정규분포를 따르므로 부품 100개를 검사하여 $\bar{x} = 301.2$ (시간)를 얻었다고 하자. $\sigma^2 = 16$ 으로 알려졌다고 할 때 μ 의 95% 신뢰구간을 구하시오.

solution. 어떤 전자 제품의 수명을 X 라 하자. X 가 균사적으로 정규분포를 따르므로 표본평균 \bar{X} 는 균사적으로 정규분포를 따른다.

주어진 통계값은 $n = 100$, $\bar{x} = 301.2$, $\sigma^2 = 16$ 이므로

$$\boxed{\bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(\mu, \frac{16}{100}\right) = N\left(\mu, \left(\frac{4}{10}\right)^2\right)}$$

이제 Theorem 7.33에 의해 μ 의 95% 신뢰구간은 다음과 같다.

$$\left(301.2 - 1.96 \times \frac{4}{10}, 301.2 + 1.96 \times \frac{4}{10}\right) = (300.4, 302.0)$$

모평균(분산이 미지인 경우)

Example 7.37 어떤 공장에서 생산하는 비누 무게의 평균을 알아보기 위하여 10개를 추출하여 비누 무게를 알아본 결과 다음과 같았다. 실제 평균 비누 무게 μ 의 95% 신뢰구간을 구하시오. 단, 비누 무게는 정규분포를 따른다고 한다.

94 96 97 85 86 79 93 89 92 78

solution. 어떤 공장에서 생산하는 비누 무게를 X 라 하자. $n = 10$ 인 표본평균을 \bar{X} , 표본분산을 S^2 이라 하면,

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(10-1) \quad (t(9))$$

이다. 표본평균(\bar{X})과 표본분산(S^2)을 구하면 다음과 같다.

$$\begin{aligned} \bar{X} &= \frac{1}{n} \sum_{i=1}^n x_i = \frac{94 + 96 + 97 + 85 + 86 + 79 + 93 + 89 + 92 + 78}{10} = 88.9 \\ S^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2 = \frac{1}{9} \left(\sum_{i=1}^n x_i^2 - n(\bar{X})^2 \right) \\ &= \frac{1}{9} \left\{ (94 - 88.9)^2 + (96 - 88.9)^2 + (97 - 88.9)^2 + (85 - 88.9)^2 + (86 - 88.9)^2 \right. \\ &\quad \left. + (79 - 88.9)^2 + (93 - 88.9)^2 + (89 - 88.9)^2 + (92 - 88.9)^2 + (78 - 88.9)^2 \right\}^2 \\ &= 45.4333 \\ s &= \sqrt{45.4333} \approx 6.74042 \end{aligned}$$

$n = 10, t_{0.025}(9) = 2.262$ 이므로 Theorem 7.36에 의해 95% 신뢰구간은 다음과 같다.

$$\begin{aligned} &\left(\bar{X} - t_{0.025}(n-1) \cdot \frac{s}{\sqrt{n}}, \bar{X} + t_{0.025}(n-1) \cdot \frac{s}{\sqrt{n}} \right) \\ &= \left(88.9 - 2.262 \times \frac{\sqrt{45.4333}}{\sqrt{10}}, 88.9 + 2.262 \times \frac{\sqrt{45.4333}}{\sqrt{10}} \right) \\ &= (88.9 - 4.821, 88.9 + 4.821) \\ &= (84.079, 93.721) \end{aligned}$$

모분산

Example 7.44 어떤 공장에서 생산된 철판의 두께는 정규분포를 따른다고 한다. 16개의 임의 표본을 추출된 결과 두께의 표준편차가 2.2 있다면 철판 두께의 분산에 대한 95% 신뢰구간을 구하시오.

solution. 공장에서 생산된 철판 두께의 분산을 σ^2 이라 하고, $n = 16$ 인 표본분산을 S^2 이라 하자. 그러면

$$V = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) = \chi^2(15) \quad \frac{15 \cdot S^2}{\sigma^2} \sim \chi^2(15)$$

이다.

$$S^2 = 2.2^2, \quad \chi^2_{0.025}(15) = 27.4884, \quad \chi^2_{0.975}(15) = 6.2621$$

이므로, 분산 σ^2 의 95% 신뢰구간은 다음과 같다.

$$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \right) = \left(\frac{15 \times 2.2^2}{27.4884}, \frac{15 \times 2.2^2}{6.2621} \right) = (2.6411, 11.5936)$$

모비율

모평균의 차

3. 어떤 지역의 응답자 중에서 암이로 150명을 선정하여 이 지역
주민대체상 투쟁 참여에 대한 자세여부를 조사한다고 한다.
90명이 자리를 나타냈을 때 투쟁 참여자의 비율이 대략 95%
신뢰구간을 보아라.

$$X \sim \text{Bin}(p) \xrightarrow{n=150} \bar{X} \approx N\left(p, \frac{p(1-p)}{n}\right) \quad (\bar{p} = \hat{p} = p)$$

$$\begin{aligned} p &= \frac{90}{150}, z_{0.025} = 1.96 \\ \therefore \left(\bar{p} - z_{0.025} \times \sqrt{\frac{p(1-p)}{n}}, \bar{p} + z_{0.025} \times \sqrt{\frac{p(1-p)}{n}} \right) \\ &= \left(\frac{90}{150} - 1.96 \times \sqrt{\frac{\frac{90}{150} \times \frac{60}{150}}{150}}, \frac{90}{150} + 1.96 \times \sqrt{\frac{\frac{90}{150} \times \frac{60}{150}}{150}} \right) \\ &= \left(\frac{3}{5} - 1.96 \times \frac{1}{25}, \frac{3}{5} + 1.96 \times \frac{1}{25} \right) \\ &= (0.5216, 0.6784) \end{aligned}$$

7. 다음은 남자와 여자의 생존 연령을 조사한 자료이다.
남자와 여자의 생존 연령은 각각 동일한 분포를 갖는 정규분포를
다른다고 한다. 남자와 여자의 평균 생존 연령과 $\mu_1 - \mu_2$ 에 대한
95% 신뢰구간을 구하시오.

$$\begin{array}{ll} X & \text{남자 } 52, 60, 55, 46, 33, 75, 58, 45, 51, 88 \\ Y & \text{여자 } 62, 58, 65, 56, 53, 45, 56, 65, 77, 47 \end{array}$$

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(18)$$

$$\begin{aligned} \bar{x} &= \frac{1}{m} \sum_{i=1}^m x_i = 56.9 \\ \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i = 58.4 \\ S_x^2 &= \frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})^2 = 240.544 \\ S_y^2 &= \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = 88.489 \\ S_p^2 &= \frac{(9)S_x^2 + (10)S_y^2}{18} = 56.5165 \end{aligned}$$

$$\begin{aligned} &\therefore \left((\bar{X} - \bar{Y}) - t_{0.025}(18) \sqrt{\frac{1}{m} + \frac{1}{n}}, (\bar{X} - \bar{Y}) + t_{0.025}(18) \sqrt{\frac{1}{m} + \frac{1}{n}} \right) \\ &= (-1.5 - 2.101 \times \sqrt{\frac{56.5165}{9}}, -1.5 + 2.101 \times \sqrt{\frac{56.5165}{9}}) \\ &= (-8.564, 5.564) \end{aligned}$$

7. 모평균, 모비율, 모분산, 모평균 차(대응비교 포함), 모분산 비의 검정