2022-2 확률과통계 중간고사 모범답안

*** (문제의 풀이과정을 보여주시오. 풀이과정이 없는 경우, 0점 처리합니다.) ***

* 각 문제 끝에 표시된 숫자는 문제에 배정된 점수이다.

1. 한 TV 제조회사는 3개의 협력업체(A,B,C)로부터 동일한 부품을 40%,35%,25% 공급받는다. A회사의 불량률은 0.009, B회사의 불량률은 0.01, C회사의 불량률은 0.012 이다. 한 개의 부품을 임의로 골라서 검사를 했을 때 불량품으로 판정되었다. 이 불량품이 C회사에서 만들어졌을 확률을 구하라. (10)

$P(A) = 0.4$
$P(D \mid A) = 0.009$

$P(B) = 0.35$
$P(D \mid B) = 0.01$

$P(C) = 0.25$
$P(D \mid C) = 0.012$

$$\begin{aligned} P(D) &= P(D \mid A)P(A) + P(D \mid B)P(B) + P(D \mid C)P(C) \\ &= (0.009 \cdot 0.4) + (0.01 \cdot 0.35) + (0.012 \cdot 0.25) \\ &= 0.0101 \end{aligned}$$

따라서 불량품일 확률은 0.0101 이다.

이 불량품이 C회사에서 만들어졌을 확률은

$$P(C|D) = \frac{P(CD)}{P(D)} = \frac{P(D \mid C)P(C)}{P(D)}$$

$$= \frac{0.012 \cdot 0.25}{0.0101} = 0.2970$$

2. 서울에서 15가구(4인 가구 기준)를 랜덤 추출하여 7월 가구당 전력소비량(kWh)을 조사한 데이터이다.

데이터 : 745, 789, 685, 645, 749, 734, 756, 698, 765, 667, 812, 727, 701, 688, 699

(1) 평균을 구하라. (5) 724

평균을 구하면 $\frac{745+789+685+645+749+734+756+689+765+667+812+727+701+688+699}{15}$ = 724 이다.

(2) 중앙값을 구하라. (5) 727

주어진 값을 크기순으로 나열하면

645, 667, 685, 688, 698, 699, 701, 727, 734, 745, 749, 756, 765, 789, 812

이때 중앙값은 8번째 값인 727이 된다.

(3) 20% 절사평균을 구하라. (5) 721.8889

20% 절사평균을 구하기 위해 (2)에서 크기순으로 나열된 값 중 가장 작은 데이터 3개, 가장 큰 데이터 3개를 제외한 나머지의 평균을 구하면

$\frac{688+698+699+701+727+734+745+749+756}{9}$ = 721.8889 이다.

3. 복권 1,000장을 발행하여 1등 1명에게는 10,000,000원을 상금으로 지급하고, 2등 3명에게는 1,000,000원을 상금으로 지급하고, 3등 10명에게는 10,000원을 상금으로 지급한다. 이 복권의 기댓값을 구하라. (5)

x	f(x)	상금
1등	$\frac{1}{1000}$	10,000,000
2등	$\frac{3}{1000}$	1,000,000
3등	$\frac{10}{1000}$	10,000
없음	$\frac{986}{1000}$	0

$$E(X) = 10{,}000{,}000 \cdot \frac{1}{1000} + 1{,}000{,}000 \cdot \frac{3}{1000} + 10{,}000 \cdot \frac{10}{1000} = 13{,}100$$

따라서 이 복권의 기대값은 13,100원 이다.

4. 1개의 주사위를 두 번 굴렸을 때 나타난 숫자의 합이 4일 사상을 A, 두 번 다 나타난 숫자가 짝수일 사상을 B라고 할 때, 확률 P(A∪B)를 구하라. (5)

사상 A 일 경우는 (1,3), (2,2), (3,1)

사상 B 일 경우는 (2,2), (2,4), (4,6), (4,2), (4,4), (4,6), (6,2), (6,4), (6,6)

확률 P(A∪B) = P(A) + P(B) - P(A∩B) = $\frac{3}{36} + \frac{9}{36} - \frac{1}{36} = \frac{11}{36}$ 이다.

5. P(A|B) = 0.4, P(B) = 0.8, P(A) = 0.6 이다. 사상 A 와 사상 B 는 독립인가? 그 이유를 설명하라. (5)

사상 A 와 사상 B 가 독립이라면 P(A|B)=P(A)를 만족해야 한다.

하지만 주어진 확률로 계산하면 0.4 $\neq$ 0.6 이므로 독립이 아니다.

6. 확률변수 X의 확률함수는 다음과 같다.

$$f(x) = c(x+1)^2 \qquad \text{for } x = 0,1,2,3$$

(1) 이때 c 값을 구하라. (5)

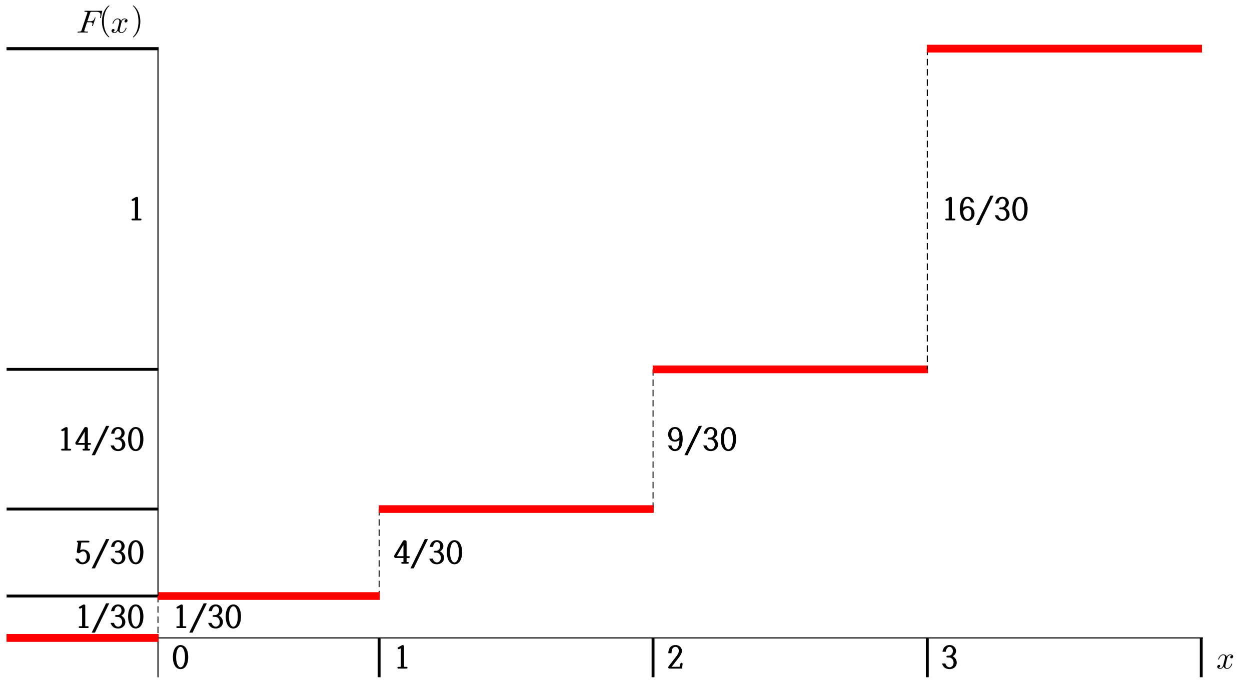
$f(x) = c(x+1)^2$,이산형확률변수이므로

$$c(0+1)^2 + c(1+1)^2 + c(2+1)^2 + c(3+1)^2$$

$$= c + 4c + 9c + 16c = 30c$$

$$30c = 1, c = \frac{1}{30}$$

(2) 분포함수 F(x)를 그림으로 표시하라. (5)



$0 \le F(x) \le 1,$
$-\infty < x < \infty$

7. 랜덤 샘플 $X_1, X_2, \ldots, X_7$ 은 평균이 μ 이고, 분산이 σ^2 인 정규분포를 따르고 있다.
모평균 μ 에 대한 두 개의 점추정량이 다음과 같다.

$$\hat{\theta}_1 = \frac{X_1 + X_2 + \ldots\ldots + X_7}{7}, \quad \hat{\theta}_2 = \frac{2X_1 - X_3 + 2X_7}{3}$$

점추정량 판정기준을 이용하여 어느 점추정량이 좋은지 평가하라. (10)

$E(\hat{\theta}_1) = E(\frac{X_1 + X_2 + \ldots\ldots + X_7}{7}) = \frac{1}{7} \cdot 7\mu = \mu$, $E(\hat{\theta}_2) = E(\frac{2X_1 - X_3 + 2X_7}{3}) = \frac{1}{3} \cdot 3\mu = \mu$ 이므로
둘 다 불편추정량이다. 분산을 구하면

$$Var(\hat{\theta}_1) = \frac{1}{49} \cdot 7\sigma^2 = \frac{1}{7}\sigma^2$$

$$Var(\hat{\theta}_2) = \frac{4}{9}\sigma^2 + \frac{1}{9}\sigma^2 + \frac{4}{9}\sigma^2 = \sigma^2$$

불편성에 의해 $E(\overline{X}) = \mu$가 되고, 효율성에 의해 분산이 작은$(\frac{\sigma^2}{7})$ $\hat{\theta}_1$이 좋다고 할 수 있다.

8. 확률변수 X_1 , X_2, X_3는 서로 독립이고 다음과 같은 분포를 따른다.
$X_1 \sim N(5,\ 1)$, $X_2 \sim N(0,\ 9)$, $X_3 \sim N(3,\ 2^2)$.
새로운 확률변수 $Y = X_1 + 2X_2 + 3X_3$ 이다.

확률변수 Y의 분포를 구하라. (5)

(1) 세 변수가 정규분포를 따르고, Y는 세 변수의 선형결합이므로 정규분포를 따르며, 평균과 분산은 다음과 같다.

(2) $E(Y) = EX_1 + 2EX_2 + 3EX_3$

$$= 5 + 2 \cdot 0 + 3 \cdot 3 = 14$$

$$Var(Y) = VarX_1 + 4VarX_2 + 9VarX_3$$

$$= 1 + 4 \cdot 9 + 9 \cdot 4 = 73$$

$$Y \sim N(14, 73)$$

9. 환자가 특정 질병에 대한 치료를 받고 회복할 확률은 동일하게 0.8 이라고 알려져 있다. 이 질병 환자 10명이 현재 치료를 받고 있다.

(1) 8 명이 회복될 확률을 구하라. (5)

$p=0.8, n=8$

$$P(X=8) = \binom{10}{8} \cdot (0.8)^8 \cdot (0.2)^2 = 0.30199$$

(2) 2 명 이상이 회복될 확률을 구하라. (5)

$$P(X \geq 2) = 1 - P(X<2)$$
$$= 1 - P(X=0) - P(X=1)$$
$$= 1 - \binom{10}{0} \cdot (0.8)^0 \cdot (0.2)^{10} - \binom{10}{1} \cdot (0.8)^1 \cdot (0.2)^9$$
$$= 0.9999$$

10. 자동차 A/S 고객센터에 찾아오는 고객의 수는 포아송 분포를 따른다고 가정을 한다. 시간당 평균 방문 고객의 수는 3명이다. 다음의 문제에 답하라.

(1) 두 시간 동안 5명의 고객이 방문하는 확률을 구하라. (5)

$t=2, \lambda=3, \lambda t=6, X \sim Possion(6)$

$$P(X=5) = \frac{e^{-5}(6)^5}{6!} = 0.1606$$

(2) 업무 개시(오전 9시) 후 30분 동안 고객이 한 명도 오지 않는 확률을 구하라. (5)

$t=0.5, \lambda=3, \lambda t = 1.5$ 따라서 $X \sim Possion(1.5)$

$$P(X=0) = \frac{e^{-1.5}(1.5)^0}{0!} = 0.22313$$

11. 확률변수 X와 확률변수 Y는 무상관 즉 상관계수 Corr(X,Y) = 0 이라고 한다. 그렇다면 확률변수 X와 Y는 서로 독립인가? (1) YES (2) NO (5)

아니다,
두 변수가 독립이라면 상관계수는 0이지만 상관계수가 0이라고 해서 두 변수가 독립이 되지 않는다.

12. 랜덤샘플 $X_1, X_2, \ldots\ldots, X_n \sim$ iid $N(\mu, \sigma^2)$ 이다. $\overline{X}, S^2$ 은 표본평균, 표본분산을 나타낸다.

(1)확률변수 $T = \dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ 임을 증명하라. (5)

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{이고, } \frac{(n-1)S^2}{\sigma^2} \sim X^2_{(n-1)} \text{이다.}$$

$$T = \frac{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}/(n-1)}} \sim t_{(n-1)} \text{인데 이를 정리하면 } = \frac{\overline{X}-\mu}{S/\sqrt{n}} \text{ 이므로,}$$

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{(n-1)} \text{이다.}$$

(2) S^2 은 σ^2 에 대한 불편추정량임을 증명하라. (5)

$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$ 이고, $\frac{(n-1)S^2}{\sigma^2} \sim X^2_{(n-1)}$ 이다.

$E\left(\frac{(n-1)S^2}{\sigma^2}\right) = n-1$ 그러므로 $E(S^2) = \sigma^2$ 이다.