MAT2410 (응용수학I)
중간시험 솔루션

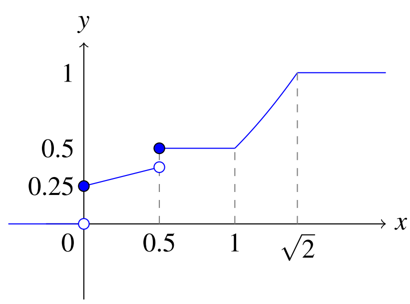
서강대학교 컴퓨터공학과 박수현 (20181634)

서강대학교 컴퓨터공학과

1. (12점) 확률변수(random variable) X의 분포함수(CDF)가 다음과 같을 때, $P\left(\frac{1}{2} \leq X < 1\right)$과 $f(x)$을 구하시오.

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.25x + 0.25 & 0 \leq x < 0.5 \\ 0.5 & 0.5 \leq x < 1 \\ 0.5x^2 & 1 \leq x < \sqrt{2} \\ 1 & x \geq \sqrt{2} \end{cases}$$

Solution. 위의 함수의 그래프는 다음과 같이 주어진다.



이 때

$$\begin{aligned} &P\left(\frac{1}{2} \leq X < 1\right) \\ =&P(X < 1) - P\left(X < \frac{1}{2}\right) \\ =&F\left(1^-\right) - F\left(\frac{1}{2}^-\right) \\ =&0.5 - 0.375 = 0.125 \end{aligned}$$

이다.

그래프에서, F는 $x = 0, x = 0.5$에서 불연속임을 확인할 수 있다. 그러므로

$$f(0) = F\left(0^+\right) - F\left(0^-\right) = 0.25$$
$$f(0.5) = F\left(0.5^+\right) - F\left(0.5^-\right) = 0.125$$

이고, 나머지 구간에서 f는 F를 x에 대해 미분해 얻을 수 있다. 따라서

$$f(x) = \begin{cases} 0 & x < 0 \\ 0.25 & 0 \le x < 0.5 \\ 0.125 & x = 0.5 \\ 0 & 0.5 < x < 1 \\ x & 1 < x < \sqrt{2} \\ 0 & x > \sqrt{2} \end{cases}$$

를 얻을 수 있다.

2. (6점) 명품을 취급하는 가게에는 60%가 진품인 것으로 알려져 있다. 이 가게의 진품 감별사는 진품을 모조품으로 감별할 확률이 0.05이고 모조품을 진품이라고 감별할 확률은 0.1이라고 한다. 이 감별사가 모조품이라고 감별한 제품이 실제로는 진품이었을 확률은 얼마인가?

Solution. 문제에서 주어진 사건들을 다음과 같이 정의하자.

$$\text{사건 } A = \text{명품이 진품인 사건}$$
$$\text{사건 } B = \text{명품을 진품으로 감별한 사건}$$

그러면 문제에서 등장한 명제들을 다음과 같이 표현할 수 있다.

$$P(A) = 0.6 \qquad P(B^c|A) = 0.05 \qquad P(B|A^c) = 0.1$$

우리가 구하고자 하는 값은 $P(A|B^c)$이다.

$$\begin{aligned}
&P(A|B^c)\\
=&\frac{P(A\cap B^c)}{P(B^c)}\\
=&\frac{P(A\cap B^c)}{P(A)}\times\frac{P(A)}{P(B^c)}\\
=&P(B^c|A)\times\frac{P(A)}{P(B^c)}\\
=&\frac{0.05\times 0.6}{P(B^c)}
\end{aligned}\tag{1}$$

또한 $P(B^c|A)=0.05$에서 $P(A\cap B^c)=0.05P(A)=0.03$을, $P(B|A^c)=0.1$에서 $P(A^c\cap B)=0.1P(A^c)=0.1[1-P(A)]=0.04$를 얻을 수 있다. 따라서

$$\begin{aligned}
&P(A)=P(A\cap B)+P(A\cap B^c)\\
\Rightarrow&0.6=P(A\cap B)+0.03\\
\Rightarrow&P(A\cap B)=0.57
\end{aligned}\tag{2}$$

이고, 또한

$$\begin{aligned}
&P(B)=P(A\cap B)+P(A^c\cap B)\\
\Rightarrow&P(B)=0.57+0.04 \qquad \because (2)\\
\Rightarrow&P(B)=0.61
\end{aligned}\tag{3}$$

이므로 (1)과 (3)에 의해

$$\begin{aligned}
P(A|B^c)=&\frac{0.05\times 0.6}{P(B^c)}\\
=&\frac{0.05\times 0.6}{1-P(B)}\\
=&\frac{0.05\times 0.6}{1-0.61}=\frac{0.03}{0.39}=\frac{1}{13}
\end{aligned}$$

이다.

3. (12점) 두 확률변수 X와 Y의 결합확률밀도함수(joint PDF)가 다음과 같을 때,

$$f(x,y) = 2x \qquad 0 < x < 1, 0 < y < 1$$

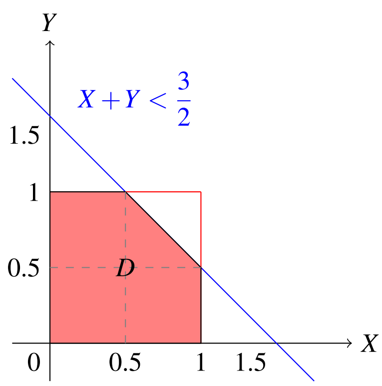
$P\left(X^2+Y^2 < \frac{1}{2}\right)$와 $P\left(X+Y < \frac{3}{2}\right)$을 각각 구하시오.

Solution.

(1)

$$\begin{aligned}
&P\left(X^2+Y^2 < \frac{1}{2}\right)\\
=&\int_0^{\frac{1}{\sqrt{2}}}\int_0^{\sqrt{\frac{1}{2}-x^2}} f(x,y)\,dy\,dx\\
=&\int_0^{\frac{\pi}{2}}\int_0^{\frac{1}{\sqrt{2}}} r f(x,y)\,dr\,d\theta \qquad r^2 = x^2+y^2, \tan\theta = \frac{y}{x}\\
=&\int_0^{\frac{\pi}{2}}\int_0^{\frac{1}{\sqrt{2}}} 2rx\,dr\,d\theta\\
=&2\int_0^{\frac{\pi}{2}}\int_0^{\frac{1}{\sqrt{2}}} r^2\cos\theta\,dr\,d\theta\\
=&2\int_0^{\frac{\pi}{2}}\cos\theta\,d\theta\int_0^{\frac{1}{\sqrt{2}}} r^2\,dr\\
=&2\left[\sin\theta\right]_0^{\frac{\pi}{2}}\left[\frac{1}{3}r^3\right]_0^{\frac{1}{\sqrt{2}}}\\
=&\frac{2}{3}\times\frac{1}{2\sqrt{2}} = \frac{1}{3\sqrt{2}}
\end{aligned}$$

(2) $P\left(X+Y < \frac{3}{2}\right)$에 해당하는 영역을 그래프로 나타내면 다음과 같다.



따라서

$$
\begin{aligned}
&P\left(X+Y<\frac{3}{2}\right)\\
&=\iint_D f(x,y)\,dy\,dx\\
&=\int_0^1\int_0^{\min\left(1,\,\frac{3}{2}-x\right)} f(x,y)\,dy\,dx\\
&=\int_0^{\frac{1}{2}}\int_0^1 2x\,dy\,dx+\int_{\frac{1}{2}}^1\int_0^{\frac{3}{2}-x} 2x\,dy\,dx\\
&=\int_0^{\frac{1}{2}} 2x\,dx+\int_{\frac{1}{2}}^1 2x\left(\frac{3}{2}-x\right)dy\,dx\\
&=x^2\Big|_0^{\frac{1}{2}}+\left[\frac{3}{2}x^2-\frac{2}{3}x^3\right]_{\frac{1}{2}}^1\\
&=\frac{1}{4}+\left(\frac{3}{2}-\frac{2}{3}\right)-\left(\frac{3}{8}-\frac{1}{12}\right)=\frac{19}{24}
\end{aligned}
$$

4. (15점) 도심의 순찰차는 호출이 있을 때마다 순찰을 하는데 그 순찰 횟수는 시간당 평균 5회인 포아송(Poisson) 분포를 따른다고 할 때

(1) 2시간 동안 한 번의 호출도 없을 확률을 구하시오.

(2) 순찰차가 호출을 3번 받을 때까지의 시간을 나타내는 확률변수를 T라고 할 때, T의 분포를 구하시오.

(3) 호출이 없었던 상태로 3시간이 지난 상태에서 다음 호출까지의 총 시간이 8시간을 넘길 확률을 구하시오.

Solution.

(1) 2시간 동안 일어나는 호출의 기댓값은 10회이다. 따라서 $\lambda = 10$인 포아송 분포 Poi(10)에 대해

$$f(0)=\frac{10^0e^{-10}}{0!}=e^{-10}$$

이다.

(2) $T\sim\Gamma\left(3,\frac{1}{5}\right)$이다.

(3) 포아송 분포는 무기억성을 가지므로 호출이 없었던 상태로 3시간이 지난 상태에서 다음 호출까지의 총 시간이 8시간을 넘길 확률은 5시간 동안 한 번의 호출도 없었을 확률과 같다. 따라서 이 때의 확률을 (1)과 같은 방법으로 계산하면 e^{-25}이다.

5. (15점) 이산인 두 확률변수 X와 Y의 결합확률질량함수(joint PMF)가 다음과 같을 때,

$$f_{X,Y}(x,y) = \frac{2}{n(n+1)} \qquad x = 1, 2, \cdots, n \qquad y = 1, 2, \cdots, x$$

(1) 두 확률변수 X와 Y의 표본공간(sample space)를 구하고 결합확률질량함수(joint PMF)임을 확인하시오.

(2) 각각의 확률변수에 대한 주변확률질량함수(marginal PMF)를 구하고 이를 이용하여 X와 Y의 독립성 여부를 결정하시오.

Solution.

(1) $x = 1, 2, \cdots, n$이고 $y = 1, 2, \cdots, x$이므로 표본공간 Ω는 다음과 같이 나타낼 수 있다.

$$\Omega = \left\{(x,y) \middle| 1 \le x \le n,\ 1 \le y \le x,\ (x,y) \in \mathbb{Z}^2\right\}$$

f가 확률질량함수라면 Ω의 모든 원소 (x,y)에 대해 $\sum_{(x,y)\in\Omega} f(x,y) = 1$이어야 한다.

$$\begin{aligned}
&\sum_{(x,y)\in\Omega} f(x,y) \\
=&\sum_{x=1}^{n}\sum_{y=1}^{x} f(x,y) \\
=&\sum_{x=1}^{n}\sum_{y=1}^{x} \frac{2}{n(n+1)} \\
=&\sum_{x=1}^{n} \frac{2x}{n(n+1)} \\
=&\frac{2}{n(n+1)} \times \frac{n(n+1)}{2} = 1
\end{aligned}$$

따라서 f는 확률질량함수이다.

(2)

$$\begin{aligned}
f_X(y) &= \sum_{y=1}^{x} f(x,y) \\
&= \sum_{y=1}^{x} \frac{2}{n(n+1)} \\
&= \frac{2x}{n(n+1)} \qquad 1 \le x \le n, x \in \mathbb{Z}
\end{aligned}$$

$$
\begin{aligned}
f_Y(y) &= \sum_{x=y}^{n} f(x,y) \\
&= \sum_{x=y}^{n} \frac{2}{n(n+1)} \\
&= \frac{2(n-y+1)}{n(n+1)} \qquad 1 \le y \le n,\, y \in \mathbb{Z}
\end{aligned}
$$

X와 Y가 독립일 경우 $f_{X,Y} = f_X f_Y$이어야 한다. 그러나 $f_{X,Y} \neq f_X f_Y$이므로 X와 Y는 독립이 아니다.

6. (10점) 두 확률변수 X와 Y의 결합확률밀도함수(joint PDF)가

$$
f_{X,Y}(x,y) = \frac{e^{-\frac{x}{y}-y}}{y} \qquad x>0,\, y>0
$$

일 때, 조건부 확률밀도함수(conditional PDF) $f(x|y)$을 구하고 이를 이용하여 $P(X>1|Y=y)$를 구하시오.

Solution.

(1)

$$
\begin{aligned}
& f(x|y) \\
=& \frac{f(x,y)}{f_Y(y)} \\
=& \frac{f(x,y)}{\int_0^\infty f(x,y)\,dx} \\
=& \frac{f(x,y)}{\int_0^\infty \frac{e^{-\frac{x}{y}-y}}{y}\,dx} = \frac{f(x,y)}{\frac{e^{-y}}{y}\int_0^\infty e^{-\frac{x}{y}}\,dx} \\
=& \frac{f(x,y)}{\frac{e^{-y}}{y}\left[-ye^{-\frac{x}{y}}\right]_{x=0}^{x=\infty}} \\
=& \frac{f(x,y)}{\frac{e^{-y}}{y}\left(0+ye^{0}\right)} = \frac{\frac{e^{-\frac{x}{y}-y}}{y}}{e^{-y}} = \frac{e^{-\frac{x}{y}}}{y}
\end{aligned}
$$

(2)

$$
\begin{aligned}
&P(X>1|Y=y)\\
&=\int_1^\infty f(x|y)\,dx\\
&=\int_1^\infty \frac{e^{-\frac{x}{y}}}{y}\,dx=\frac{1}{y}\int_1^\infty e^{-\frac{x}{y}}\,dx\\
&=-e^{-\frac{x}{y}}\Big|_{x=1}^{x=\infty}\\
&=0+e^{-\frac{1}{y}}=e^{-\frac{1}{y}}
\end{aligned}
$$

7-1. (10점) 연속확률변수 $X_1, X_2, \cdots, X_n$에 대하여

(1) $\mathrm{E}(X_1+X_2)=\mathrm{E}(X_1)+\mathrm{E}(X_2)$임을 증명하시오.
(2) $\mathrm{Var}(X_1+X_2)=\mathrm{Var}(X_1)+\mathrm{Var}(X_2)+2\mathrm{Cov}(X_1,X_2)$임을 증명하시오.

Solution. X_1의 PDF를 f_1, X_2의 PDF를 f_2라고 하자.

(1)

$$
\begin{aligned}
&\mathrm{E}(X_1)+\mathrm{E}(X_2)\\
&=\int_{-\infty}^\infty xf_1(x)\,dx+\int_{-\infty}^\infty xf_2(x)\,dx\\
&=\int_{-\infty}^\infty x\left[f_1(x)+f_2(x)\right]dx\\
&=\mathrm{E}(X_1+X_2) \qquad \square
\end{aligned}
$$

(2)

$$
\begin{aligned}
&\mathrm{Var}(X_1)+\mathrm{Var}(X_2)+2\mathrm{Cov}(X_1,X_2)\\
&=\left[\mathrm{E}\left(X_1^2\right)-[\mathrm{E}(X_1)]^2\right]+\left[\mathrm{E}\left(X_2^2\right)-[\mathrm{E}(X_2)]^2\right]+2\mathrm{E}\left[(X_1-\mathrm{E}(X_1))(X_2-\mathrm{E}(X_2))\right]\\
&=\left[\mathrm{E}\left(X_1^2\right)+\mathrm{E}\left(X_2^2\right)\right]-\left[[\mathrm{E}(X_1)]^2+[\mathrm{E}(X_2)]^2\right]+2\mathrm{E}\left[X_1X_2-X_1\mathrm{E}(X_2)-X_2\mathrm{E}(X_1)+\mathrm{E}(X_1)\mathrm{E}(X_2)\right]\\
&=\left[\mathrm{E}\left(X_1^2\right)+\mathrm{E}\left(X_2^2\right)\right]-\left[[\mathrm{E}(X_1)]^2+[\mathrm{E}(X_2)]^2\right]+2\left[\mathrm{E}(X_1X_2)-\mathrm{E}(X_1)\mathrm{E}(X_2)\right]\\
&=\left[\mathrm{E}\left(X_1^2\right)+2\mathrm{E}(X_1X_2)+\mathrm{E}\left(X_2^2\right)\right]-\left[[\mathrm{E}(X_1)]^2+2\mathrm{E}(X_1)\mathrm{E}(X_2)+[\mathrm{E}(X_2)]^2\right]\\
&=\mathrm{E}\left(X_1^2+2X_1X_2+X_2^2\right)-\left[[\mathrm{E}(X_1)]^2+2\mathrm{E}(X_1)\mathrm{E}(X_2)+[\mathrm{E}(X_2)]^2\right]\\
&=\mathrm{E}\left[(X_1+X_2)^2\right]-[\mathrm{E}(X_1)+\mathrm{E}(X_2)]^2\\
&=\mathrm{Var}(X_1+X_2) \qquad \square
\end{aligned}
$$

7-2. (25점) N명의 사람들이 자신의 소지품을 항아리에 하나씩 담고 임의로 한 개를 뽑을 때 자신의 것을 선택하게 되는 사람의 수를 확률변수 X라 하자.

(1) i번째 사람이 자신의 소지품을 선택하면 $X_i = 1$, 그렇지 않으면 $X_i = 0$이라 하면 확률변수 X_i가 베르누이(Bernoulli) 분포를 갖는다고 할 수 있는지 설명하고, 확률변수 X를 X_i의 함수로 표시하시오.
(2) $\mathrm{E}(X_i) = \dfrac{1}{N}$와 $\mathrm{Var}(X_i) = \dfrac{N-1}{N^2}$임을 보이시오.
(3) 자신의 소지품을 선택하게 될 기댓값 $\mathrm{E}(X)$와 분산 $\mathrm{Var}(X)$를 구하시오.

Solution. $A_1 \cdots A_N$의 소지품이 있을 때, i번째 사람이 각각 소지품 A_j를 고르는 것을 i번째 원소가 A_j인 집합 P_k라고 정의하면 P_k는 $\{A_1, A_2, \cdots, A_N\}$의 원소들의 순서를 섞은 순열이다. 그러면 각 순열이 나오는 경우가 균등분포를 따른다고 생각할 수 있을 것이다.

(1) 문제에서 X_i의 정의는 베르누이 분포의 정의와 같다. 따라서 X_i는 베르누이 분포를 갖는다고 할 수 있다. 이 경우

$$X = \sum_{i=1}^{N} X_i$$

이다.

(2) 각 순열이 나오는 경우는 균등분포를 따르므로, i번째 사람이 A_i번 소지품을 고를 확률은 $\dfrac{1}{N}$으로 일정하다. 따라서 기댓값은

$$\begin{aligned}\mathrm{E}(X) =& 0 \times \left(1 - \frac{1}{N}\right) + 1 \times \frac{1}{N} \\ =& \frac{1}{N}\end{aligned}$$

이다. 또한 분산은

$$\begin{aligned}\mathrm{Var}(X) =& \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2 \\ =& \left[0^2 \times \left(1 - \frac{1}{N}\right) + 1^2 \times \frac{1}{N}\right] - \frac{1}{N^2} \\ =& \frac{1}{N} - \frac{1}{N^2} = \frac{N-1}{N^2}\end{aligned}$$

이다.

(3) (1)에서 $X = \sum_{i=1}^{N} X_i$임을 보였다. 이를 이용해 기댓값을 구하면

$$\begin{aligned}\mathrm{E}(X) =& \mathrm{E}\left(\sum_{i=1}^{N} X_i\right) \\ =& \sum_{i=1}^{N} \mathrm{E}(X_i) \\ =& \sum_{i=1}^{N} \frac{1}{N} = 1\end{aligned}$$

이다.

문제 **7-1**에서 증명한 내용을 귀납적으로 확장하면, 여러 확률변수의 합에 대해서 분산은 일반적으로

$$\mathrm{Var}(X) = \mathrm{Var}\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} \mathrm{Var}(X_i) + 2\sum_{1\le i<j\le N} \mathrm{Cov}(X_i, X_j)$$

이다.

X의 분산을 구하기 위해 임의의 $1 \le i < j \le N$에 대한 공분산 $\mathrm{Cov}(X_i, X_j) = \mathrm{E}(X_iX_j) - \mathrm{E}(X_i)\mathrm{E}(X_j)$을 먼저 구하자. X_i와 X_j는 모두 베르누이 분포를 따르므로 $\mathrm{E}(X_iX_j)$는 i번째 사람과 j번째 사람이 모두 자신의 소지품을 선택하는 확률을 뜻한다. 각 순열이 나오는 경우가 균등분포를 따름을 이용한다면 i번째 사람과 j번째 사람이 모두 자신의 소지품을 선택하는 확률은 먼저 i번째 사람이 N개의 소지품 중 자신의 소지품 1개를 고르고, j번째 사람이 남은 $N-1$개의 소지품 중 자신의 소지품 1개를 고르는 확률로 생각할 수 있으므로

$$\mathrm{E}(X_iX_j) = \frac{1}{N} \times \frac{1}{N-1}$$

이고, 이를 통해 공분산을 구하면 공분산은

$$\begin{aligned}\mathrm{Cov}(X_i, X_j) =& \mathrm{E}(X_iX_j) - \mathrm{E}(X_i)\mathrm{E}(X_j)\\ =& \frac{1}{N} \times \frac{1}{N-1} - \frac{1}{N} \times \frac{1}{N}\\ =& \frac{1}{N^2(N-1)}\end{aligned}$$

으로 일정하다. 따라서 X의 분산은

$$\begin{aligned}\mathrm{Var}(X) =& \mathrm{Var}\left(\sum_{i=1}^{N} X_i\right)\\ =& \sum_{i=1}^{N} \mathrm{Var}(X_i) + 2\sum_{1\le i<j\le N} \mathrm{Cov}(X_i, X_j)\\ =& \sum_{i=1}^{N} \frac{N-1}{N^2} + 2\sum_{1\le i<j\le N} \frac{1}{N^2(N-1)}\\ =& \frac{N-1}{N} + 2\binom{N}{2}\frac{1}{N^2(N-1)}\\ =& \frac{N-1}{N} + \frac{N(N-1)}{N^2(N-1)}\\ =& \frac{N-1}{N} + \frac{1}{N} = 1\end{aligned}$$

이다.