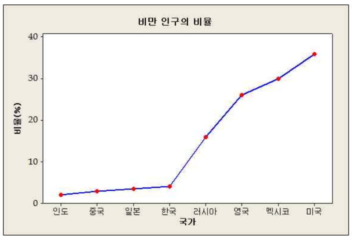
생생한 사례로 배우는 확률과 통계

연습문제 홀수번 풀이 이용 안내

- 본 문제 풀이의 저작권은 **이재원**과 **한빛아카데미(주)**에 있습니다.
- 이 자료를 무단으로 전제하거나 배포할 경우 저작권법 136조에 의거하여 최고 5년 이하의 징역 또는 5천만원 이하의 벌금에 처할 수 있고 이를 병과(倂科)할 수도 있습니다.

Chapter 01 연습문제 풀이

1.



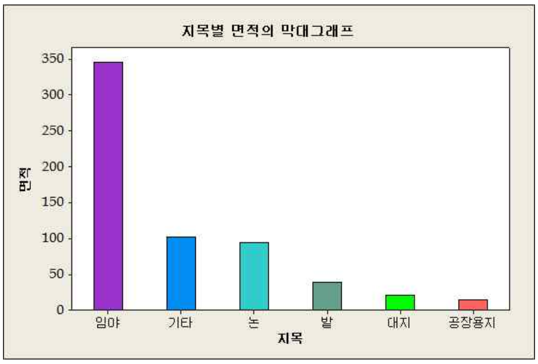
3. 전체 토지의 면적은 614.42이다. 따라서 다음 도수표를 얻는다.

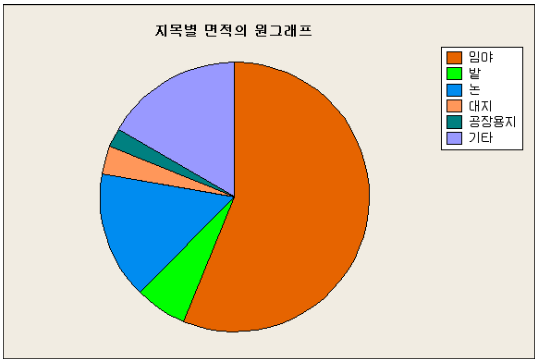
구분	면적	상대도수
임야	345.56	0.562
밭	38.56	0.063
논	94.11	0.153
대지	20.55	0.033
공장용지	13.94	0.023
기타	102.00	0.166

각 지목별 중심각은 다음과 같다.

임야: 202.3° 밭: 22.6° 논: 55.1° 1 대지: 12.0° 공장용지: 8.2° 기타: 59.8°

따라서 지목별 면적에 대한 막대그래프와 원그래프는 다음과 같다.





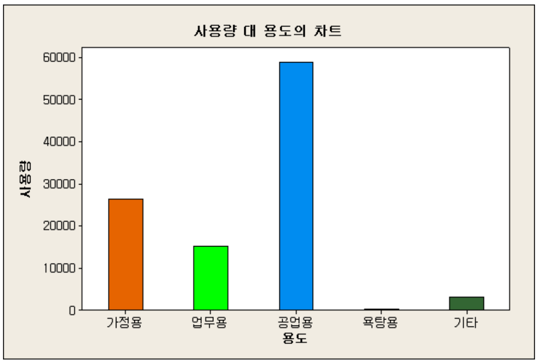
5. 전체 급수 사용량은103,622이다. 따라서 다음 도수표를 얻는다.

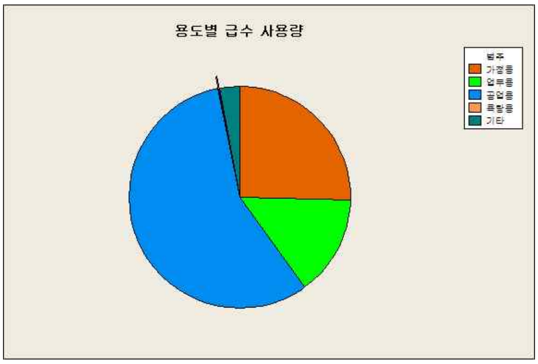
구분	급수 사용량	상대도수
가정용	26,384	0.255
업무용	15,145	0.145
공업용	58,821	0.568
욕탕용	207	0.002
기타	3,065	0.030

각 용도별 중심각은 다음과 같다.

가정용: 91.7° 업무용: 52.6° 공업용: 204.4° 욕탕용: 0.7° 기타: 10.6°

따라서 용도별 급수 사용량에 대한 막대그래프와 원그래프는 다음과 같다.





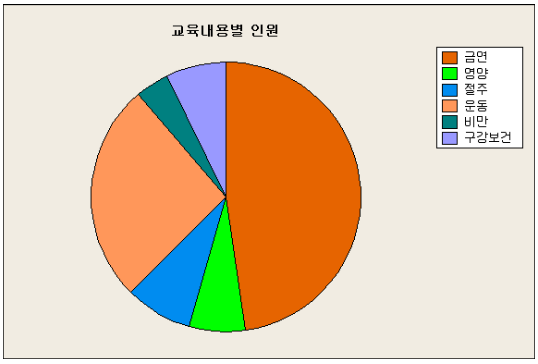
7. 교육을 받은 전체 인원은 104386이다. 따라서 다음 도수표를 얻는다.

교육 내용	인원	상대도수
금연	49,779	0.48
영양	7,013	0.07
절주	8,412	0.08
운동	27,420	0.26
비만	4,198	0.04
구강보건	7,564	0.07

각 교육 내용별 중심각은 다음과 같다.

금연: 171.7° 영양: 24.2° 절주: 29.0° 운동: 94.5° 비만: 14.5° 구강보건: 26.1°

따라서 교육 내용별 인원에 대한 막대그래프와 원그래프는 다음과 같다.



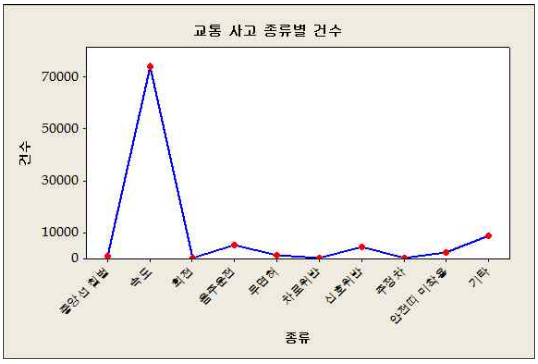
9. 전체 단속 및 처리 건수는 959344이다. 따라서 다음 도수표를 얻는다.

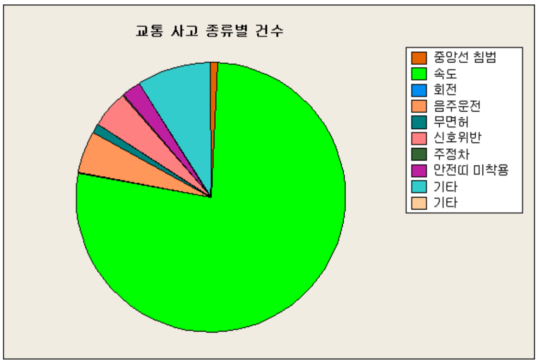
종류	건수	상대도수
중앙선침범	745	0.008
속도	73,919	0.771
회전	139	0.001
음주운전	4,837	0.050
무면허	1,013	0.011
차로위반	9	0.000
신호위반	4,390	0.046
주정차	111	0.001
안전띠미착용	2,208	0.023
기타	8,563	0.089

각 교통사고 종류별 중심각은 다음과 같다.

중앙선침범: 2.8° 속도: 277.4° 회전: 0.5° 음주운전: 18.1° 무면허: 3.8°
차로위반: 0.03° 신호위반: 16.5° 주정차: 0.4° 안전띠미착용: 8.3° 기타: 32.1°

따라서 교통사고 종류별 건수에 대한 막대그래프와 원그래프는 다음과 같다.





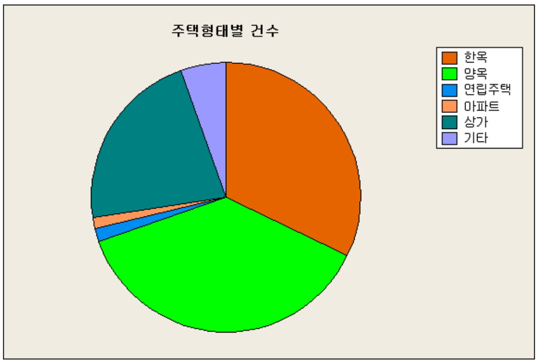
11. 전체 가구 수는 3267이다. 따라서 다음 도수표를 얻는다.

주택형태	건수	상대도수
한옥	1051	0.32
양옥	1225	0.38
연립주택	53	0.02
아파트	44	0.01
상가	717	0.22
기타	177	0.05

각 주택형태별 중심각은 다음과 같다.

한옥: 115.8° 양옥: 135.0° 연립주택: 5.8° 아파트: 4.9° 상가: 79.0° 기타: 19.5°

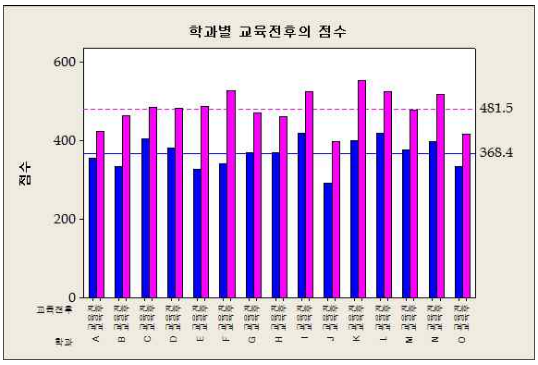
따라서 주택형태별 건수에 대한 막대그래프와 원그래프는 다음과 같다.

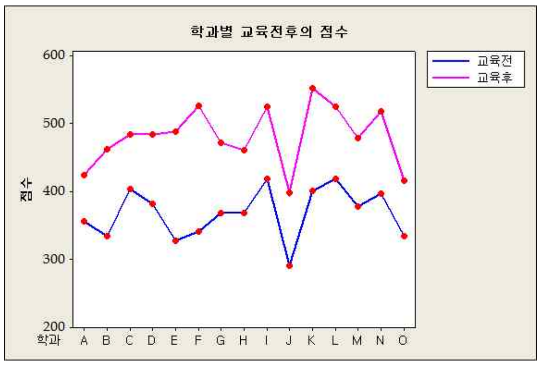


13. (a) 교육 전의 평균: $\bar{x} = \dfrac{1}{15}\sum_{i=1}^{15} x_i = \dfrac{5526}{15} = 368.4$

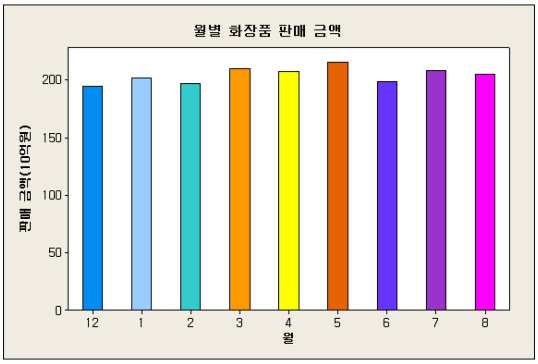
교육 후의 평균: $\bar{y} = \dfrac{1}{15}\sum_{i=1}^{15} y_i = \dfrac{7223}{15} = 481.5$

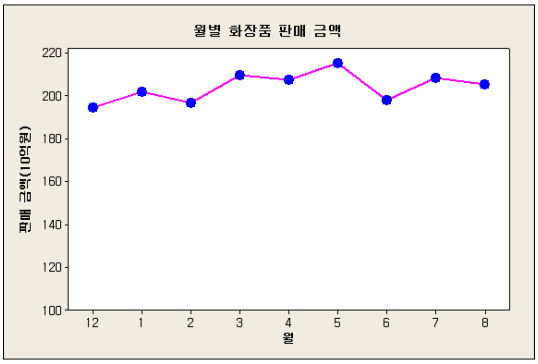
(b)





15. (a)





(b) 겨울의 평균 판매 금액: $\overline{x}_1 = \dfrac{1}{3}(194.7 + 201.9 + 196.7) = \dfrac{593.3}{3} = 197.77$

봄의 평균 판매 금액: $\overline{x}_2 = \dfrac{1}{3}(209.6 + 207.7 + 215.3) = \dfrac{632.6}{3} = 210.87$

여름의 평균 판매 금액: $\overline{x}_3 = \dfrac{1}{3}(198.1 + 208.4 + 205.3) = \dfrac{611.8}{3} = 203.93$

Chapter 02 연습문제 풀이

1. (a) 평균: $\overline{x} = \frac{1}{26}\sum_{i=1}^{26} x_i = \frac{167}{26} = 6.42$

분산: $s^2 = \frac{1}{25}\sum_{i=1}^{26}(x_i - 6.42)^2 = \frac{1428.35}{25} = 57.134$

표준편차: $s = \sqrt{57.134} = 7.5587$

(b) 전체 자료를 크기 순서로 재배열하면 다음과 같다.

0 0 0 1 1 1 1 1 2 2 2 2 2 4 6 6 6 8 8 10 11 12 15 16 17 33

전체 자료의 수가 26이므로 중위수는 13번째와 14번째의 평균 $M_e = \frac{x_{(13)} + x_{(14)}}{2} = \frac{2+4}{2} = 3$이다. 가장 많은 도수를 가진 자료는 1과 2이므로 최빈값은 1과 2이다.

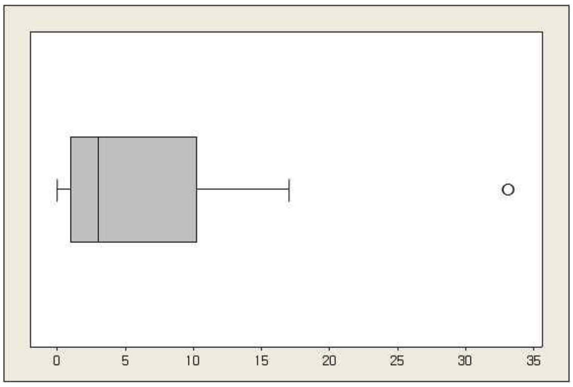
(c) 25-백분위수와 75-백분위수를 구한다. $26 \times 0.25 = 6.5$, $26 \times 0.75 = 19.5$이므로

$$Q_1 = P_{25} = x_{(7)} = 1,\ Q_3 = P_{75} = x_{(20)} = 10$$

(d) $26 \times 0.3 = 7.8$이므로 30-백분위수는 $P_{25} = x_{(8)} = 1$이다.

(e) 다음 순서에 따라서 상자그림을 그린다.

① 사분위수범위는 $\text{I.Q.R} = Q_3 - Q_1 = 10 - 1 = 9$이다.
② $Q_1 = 1$에서 $Q_3 = 10$까지 직사각형 모양의 상자로 연결하여 그리고, 중위수 $Q_2 = 3$ 위치인 상자 안에 수직선을 긋는다.
③ 안울타리와 인접값을 구한다. $f_l = Q_1 - (1.5)\text{I.Q.R} = -12.5$, $f_u = Q_3 + (1.5)\text{I.Q.R} = 23.5$이므로 인접값은 0과 17이다.
④ 위쪽 바깥울타리를 구한다. $f_U = Q_3 + 3\text{I.Q.R} = 37$이므로 보통 극단값 33의 위치에 “○”로 표시한다.



3. (a) 평균: $\overline{x} = \frac{1}{26}\sum_{i=1}^{26} x_i = \frac{5411}{26} = 208.1$

분산: $s^2 = \frac{1}{25}\sum_{i=1}^{26}(x_i - 208.1)^2 = \frac{925971}{25} = 37038.8$

표준편차: $s = \sqrt{37038.8} = 192.455$

(b) 전체 자료를 크기 순서로 재배열하면 다음과 같다.

11 16 21 25 29 41 46 52 80 80 89 133 148 153 202
209 238 257 267 281 397 415 422 529 572 698

전체 자료의 수가 26이므로 중위수는 13번째와 14번째 자료의 평균이다.

$$M_e = \frac{x_{(13)} + x_{(14)}}{2} = \frac{148+153}{2} = 150.5$$

가장 많은 도수를 가진 자료는 80이므로 최빈값은 80이다.

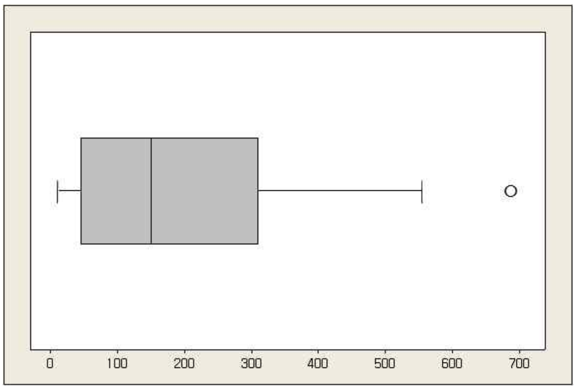
(c) 25-백분위수와 75-백분위수를 구한다. $26 \times 0.25 = 6.5$, $26 \times 0.75 = 19.5$이므로

$$Q_1 = P_{25} = x_{(7)} = 46,\ Q_3 = P_{75} = x_{(20)} = 281$$

(d) $26 \times 0.3 = 7.8$이므로 30-백분위수는 $P_{25} = x_{(8)} = 52$이다.

(e) 다음 순서에 따라서 상자그림을 그린다.

① 사분위수범위는 $\text{I.Q.R} = Q_3 - Q_1 = 281 - 46 = 235$이다.
② $Q_1 = 46$에서 $Q_3 = 281$까지 직사각형 모양의 상자로 연결하여 그리고, 중위수 $Q_2 = 150.5$ 위치인 상자 안에 수직선을 긋는다.
③ 안울타리와 인접값을 구한다. $f_l = Q_1 - (1.5)\text{I.Q.R} = -306.5$, $f_u = Q_3 + (1.5)\text{I.Q.R} = 633.5$이므로 인접값은 11과 572이다.
④ 위쪽 바깥울타리를 구한다. $f_U = Q_3 + 3\,\text{I.Q.R} = 986$이므로 보통 극단값 698의 위치에 "○"로 표시한다.



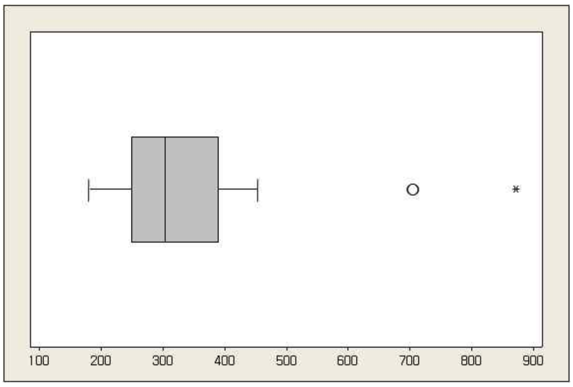
5. (a) 연도별 평균 강수량: 2002년 504, 2003년 384.7, 2005년 324.5, 2006년 249.8, 2011년 321.1, 2012년 264

(b) 25-백분위수, 50-백분위수와 75-백분위수는 각각 다음과 같다.

$$Q_1 = P_{25} = x_{(7)} = 255.5,\ \ Q_2 = P_{50} = x_{(13)} = 304,\ \ Q_3 = P_{75} = x_{(19)} = 382$$

그러므로 사분위수범위는 $\mathrm{I.Q.R} = Q_3 - Q_1 = 382 - 255.5 = 126.5$이다.

(c) $f_l = Q_1 - (1.5)\mathrm{I.Q.R} = 65.75$, $f_u = Q_3 + (1.5)\mathrm{I.Q.R} = 571.75$, $f_U = Q_3 + 3\,\mathrm{I.Q.R} = 761.5$이므로 인접값은 181과 453이다. 그러므로 보통 극단값은 712.5, 극단값은 870.5이다. 이제 보통 극단값 712.의 위치에 "○", 극단값 870.5의 위치에 "×"로 표시한다.

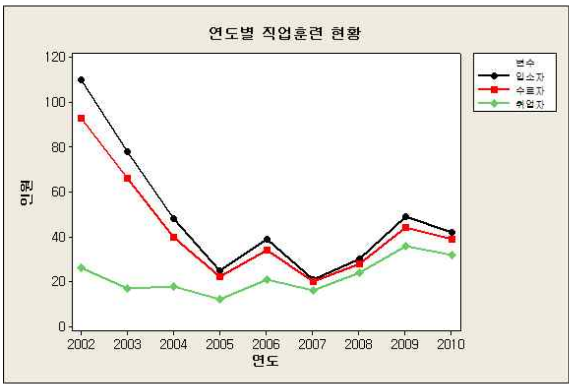


7. (a) 연평균 입소자 수: $\bar{x} = \frac{1}{9}\sum x_i = \frac{442}{9} = 49.11$

연평균 수료자 수: $\bar{y} = \frac{1}{9}\sum y_i = \frac{386}{9} = 42.89$

연평균 취업자 수: $\bar{z} = \frac{1}{9}\sum z_i = \frac{202}{9} = 22.44$

(b)



9. (a) 진도 3이상인 지진 발생 횟수에 대한 전체 자료의 수가 36이고
$36 \times 0.25 = 9$, $36 \times 0.5 = 18$, $36 \times 0.75 = 27$이므로 사분위수는 각각 다음과 같다.

$$Q_1 = \frac{x_{(9)} + x_{(10)}}{2} = \frac{7+7}{2} = 7,$$
$$Q_2 = \frac{x_{(18)} + x_{(19)}}{2} = \frac{9+10}{2} = 9.5,$$
$$Q_3 = \frac{x_{(27)} + x_{(28)}}{2} = \frac{11+12}{2} = 11.5$$

- 사분위수범위는 $\text{I.Q.R} = Q_3 - Q_1 = 11.5 - 7 = 4.5$이다.
- 안울타리: $f_l = Q_1 - (1.5)\text{I.Q.R} = 0.25$, $f_u = Q_3 + (1.5)\text{I.Q.R} = 18.25$
- 인접값: 2, 18
- 바깥울타리: $F_l = Q_1 - 3 \times \text{I.Q.R} = -6.5$, $F_u = Q_3 + 3 \times \text{I.Q.R} = 25$
- 극단값: 37

유감지진의 사분위수는 각각 다음과 같다.

$$Q_1 = \frac{x_{(9)} + x_{(10)}}{2} = \frac{4+5}{2} = 4.5,$$
$$Q_2 = \frac{x_{(18)} + x_{(19)}}{2} = \frac{6+7}{2} = 6.5,$$
$$Q_3 = \frac{x_{(27)} + x_{(28)}}{2} = \frac{8+9}{2} = 8.5$$

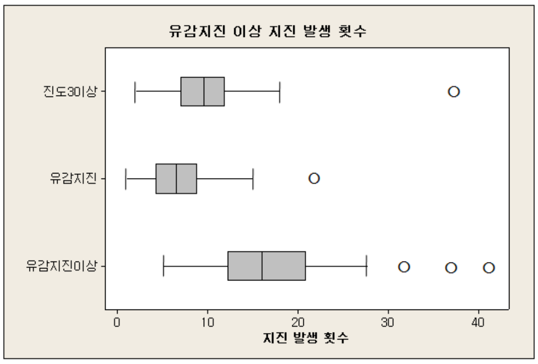
- 사분위수범위는 $\text{I.Q.R} = Q_3 - Q_1 = 8.5 - 4.5 = 4$이다.
- 안울타리: $f_l = Q_1 - 1.5 \times \text{I.Q.R} = -1.5$, $f_u = Q_3 + 1.5 \times \text{I.Q.R} = 14.5$
- 인접값: 1, 13
- 바깥울타리: $F_u = Q_3 + 3 \times \text{I.Q.R} = 20.5$
- 극단값: 22

유감지진 이상의 사분위수는 각각 다음과 같다.

$$Q_1 = \frac{x_{(9)} + x_{(10)}}{2} = \frac{12+13}{2} = 12.5,$$
$$Q_2 = \frac{x_{(18)} + x_{(19)}}{2} = \frac{16+16}{2} = 16,$$
$$Q_3 = \frac{x_{(27)} + x_{(28)}}{2} = \frac{20+21}{2} = 20.5$$

- 사분위수범위는 $\text{I.Q.R} = Q_3 - Q_1 = 20.5 - 12.5 = 8$이다.
- 안울타리: $f_l = Q_1 - 1.5 \times \text{I.Q.R} = 0.5$, $f_u = Q_3 + 1.5 \times \text{I.Q.R} = 32.5$

- 인접값: 5, 27
- 바깥울타리: $F_u = Q_3 + 3 \times \mathrm{I.Q.R} = 44.5$
- 보통 극단값: 33, 37, 41



(b) 진도 3이상인 연평균 발생 횟수: $\bar{x} = \dfrac{1}{36}\sum x_i = \dfrac{356}{36} = 10.14$

진도 3이상인 연평균 발생 횟수: $\bar{y} = \dfrac{1}{36}\sum y_i = \dfrac{253}{36} = 7.03$

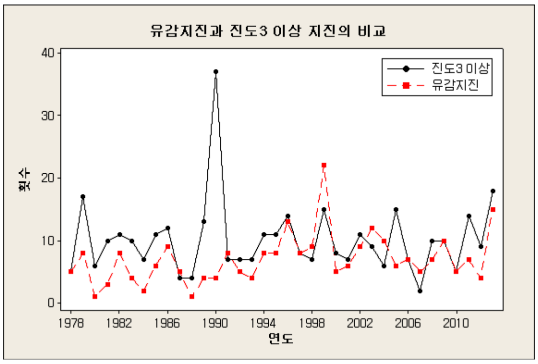
진도 3이상의 분산: $s_x^2 = \dfrac{1}{35}\sum (x_i - 10.14)^2 = \dfrac{1232.3}{35} = 35.209$,

표준편차: $s_x = \sqrt{35.209} = 5.9337$

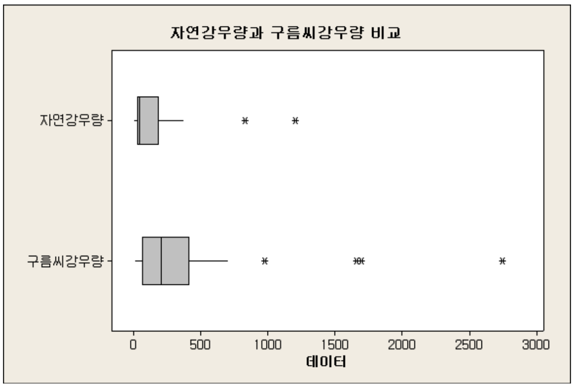
유감지진의 분산: $s_y^2 = \dfrac{1}{35}\sum (y_i - 7.03)^2 = \dfrac{570.97}{35} = 16.313$,

표준편차: $s_y = \sqrt{16.313} = 4.0389$

(c)



11.



13. (a) • 평균 물 소비량: $\bar{x}= \frac{1}{10}\sum x_i = \frac{450}{10}= 45.0$

• 평균 세탁량: $\bar{y}= \frac{1}{10}\sum y_i = \frac{37}{10}= 3.7$

• 부품 수의 분산: $s_x^2 = \frac{1}{9}\sum (x_i - 45)^2 = \frac{2750}{9}= 305.556$

• 표준편차: $s_x = \sqrt{305.556}= 17.48$

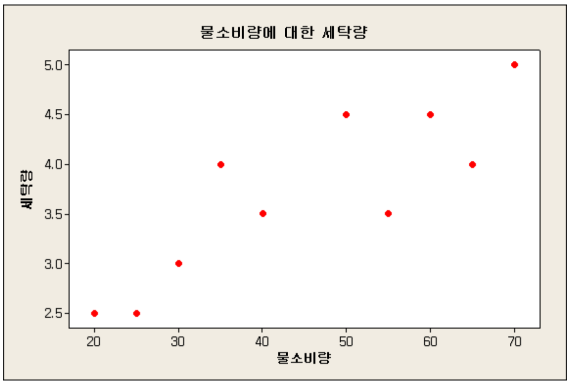
• 소요 시간의 분산: $s_y^2 = \frac{1}{9}\sum (y_i - 3.7)^2 = \frac{6.6}{9}= 0.733$

• 표준편차: $s_y = \sqrt{0.733}= 0.856$

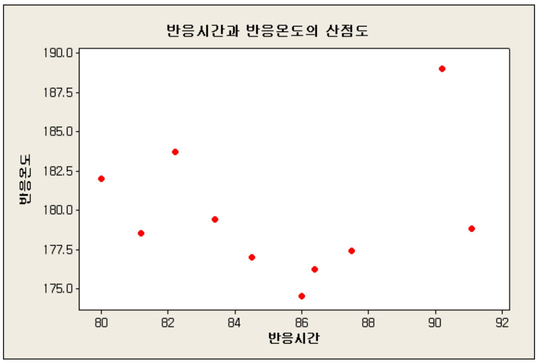
• 공분산: $s_{xy} = \frac{1}{9}\sum_{i=1}^{10} (x_i - \bar{x})(y_i - \bar{y}) = \frac{115}{9}= 12.7778$

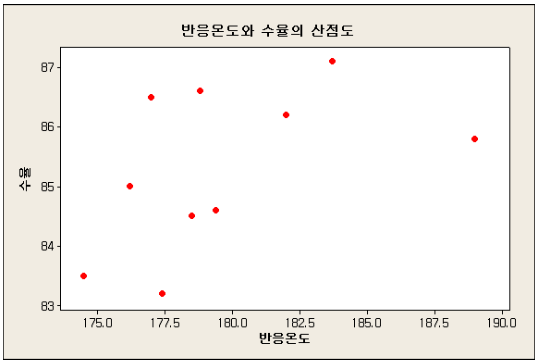
• 상관계수: $r_{xy} = \frac{s_{xy}}{s_x s_y}= \frac{12.7778}{17.48 \times 0.856}= 0.854$

(b)



15. (a) 산점도





(b) • 평균 반응시간: $\overline{x}=\dfrac{1}{10}\sum x_i=\dfrac{852.5}{10}=85.25$

• 평균 반응온도: $\overline{y}=\dfrac{1}{10}\sum y_i=\dfrac{1796.5}{10}=179.65$

• 반응시간의 분산: $s_x^2=\dfrac{1}{9}\sum(x_i-134.99)^2=\dfrac{122.92}{9}=13.6578$

• 표준편차: $s_x=\sqrt{13.6578}=3.6957$

• 반응온도의 분산: $s_y^2=\dfrac{1}{9}\sum(y_i-179.65)^2=\dfrac{161.96}{9}=17.9956$

• 표준편차: $s_y=\sqrt{17.9956}=4.2421$

• 공분산: $s_{xy}=\dfrac{1}{9}\sum_{i=1}^{10}(x_i-\overline{x})(y_i-\overline{y})=\dfrac{10.84}{9}=1.2044$

• 상관계수: $r_{xy}=\dfrac{s_{xy}}{s_x s_y}=\dfrac{1.2044}{3.6957\times 4.2421}=0.077$

(c) • 평균 수율: $\overline{x}=\dfrac{1}{10}\sum x_i=\dfrac{853}{10}=85.3$

• 수율의 분산: $s_x^2=\dfrac{1}{9}\sum(x_i-85.3)^2=\dfrac{16.381}{9}=1.8201$

• 표준편차: $s_x=\sqrt{1.8201}=1.3491$

• 공분산: $s_{xy}=\dfrac{1}{9}\sum_{i=1}^{10}(x_i-\overline{x})(y_i-\overline{y})=\dfrac{25.92}{9}=2.88$

• 상관계수: $r_{xy}=\dfrac{s_{xy}}{s_x s_y}=\dfrac{2.88}{1.3491\times 4.2421}=0.5032$

Chapter 03 연습문제 풀이

1. (a) A 도시에 물 공급이 원활하지 않는 사건은 두 파이프라인 1과 2에 이상이 있거나 파이프라인 3에 이상이 있는 경우이므로 $(A_1 \cap A_2) \cup A_3$이다.

   (b) (a)과 동일하게 B 도시에 물 공급이 원활하지 않는 사건은 $(A_1 \cap A_2) \cup A_3$이다. 그러므로 B 도시에 물 공급이 원활한 사건은 $[(A_1 \cap A_2) \cup A_4]^c$이다.

   (c) (b)와 동일하게 A 도시에 물 공급이 원활한 사건은 $[(A_1 \cap A_2) \cup A_3]^c$이다. 따라서 구하고자 하는 사건은 다음과 같다.

$$[(A_1 \cap A_2) \cup A_3]^c \cup [(A_1 \cap A_2) \cup A_4]^c = (A_1 \cap A_2)^c \cap (A_3^c \cup A_4^c)$$

3. (a) $P(A \cap B) = P(A) - P(A \cap B^c) = 0.4 - 0.3 = 0.1$
   (b) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.54 - 0.1 = 0.84$
   (c) $P(A^c \cap B^c) = P\left[(A \cup B)^c\right] = 1 - P(A \cup B) = 1 - 0.84 = 0.16$
   $P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{1}{4} + \frac{1}{3} - \frac{1}{2} = \frac{1}{12}$

5. (a) $S = \{(1,4,7,8,9), (1,4,5,8,9), (1,4,5,6,9), (1,2,5,8,9), (1,2,5,6,9), (1,2,3,6,9)\}$
   (b) $A = \{(1,4,5,8,9), (1,4,5,6,9), (1,2,5,8,9), (1,2,5,6,9)\}$
   (c) $P(A) = \frac{4}{6} = \frac{2}{3}$

7. (a) 각 부품이 작동하는 사건을 각각 A, B, C, D라 하면 $P(A) = 0.5$, $P(B) = 0.95$, $P(C) = 0.99$, $P(D) = 0.45$이다. 각 구성 요소들의 작동 여부가 독립이므로 구하고자 하는 확률은 다음과 같다.

$$P(A \cap B \cap C \cap D) = P(A)\,P(B)\,P(C)\,P(D) = 0.5 \times 0.95 \times 0.99 \times 0.45 = 0.2116$$

   (b) 네 부품이 독립적으로 작동하므로 다음 확률을 얻는다.

$$P(A \cap B) = P(A)\,P(B) = 0.5 \times 0.95 = 0.475$$

$$P(C \cap D) = P(C)\,P(D) = 0.99 \times 0.45 = 0.4455$$

$$P(A \cap B \cap C \cap D) = P(A)\,P(B)\,P(C)\,P(D) = 0.5 \times 0.95 \times 0.99 \times 0.45 = 0.2116$$

이 전기 장치가 작동하는 사건을 E 라 하면, $E=(A\cap B)\cup(C\cap D)$이므로 구하고자 하는 확률은 다음과 같다.

$$\begin{aligned}P(E)=P[(A\cap B)\cup(C\cap D)]&=P(A\cap B)+P(C\cap D)-P[(A\cap B)\cap(C\cap D)]\\&=P(A)P(B)+P(C)P(D)-P(A)P(B)P(C)P(D)\\&=0.475+0.4455-0.2116=0.7089\end{aligned}$$

(c) 네 부품이 독립적으로 작동하므로 구하고자 하는 확률은 다음과 같다.

$$\begin{aligned}P[A\cup(B\cap C\cap D)]&=P(A)+P(B\cap C\cap D)-P[A\cap(B\cap C\cap D)]\\&=P(A)+P(B)P(C)P(D)-P(A)P(B)P(C)P(D)\\&=0.5+0.95\times0.99\times0.45-0.2116\\&=0.7116\end{aligned}$$

(d) 네 부품이 독립적으로 작동하므로 구하고자 하는 확률은 다음과 같다.

$$\begin{aligned}P[A\cup B\cup(C\cap D)]&=P(A)+P(B)+P(C\cap D)-P(A\cap B)-P[A\cap(C\cap D)]\\&\quad-P[B\cap(C\cap D)]+P[A\cap B\cap(C\cap D)]\\&=P(A)+P(B)+P(C)P(D)-P(A)P(B)-P(A)P(C)P(D)\\&\quad-P(B)P(C)P(D)+P(A)P(B)P(C)P(D)\\&=0.5+0.95+0.99\times0.45-0.5\times0.95-0.5\times0.99\times0.45\\&\quad-0.95\times0.99\times0.45+0.2116\\&=0.9861\end{aligned}$$

9. (a) 공정한 동전을 네 번 던지는 게임에서 표본공간은 다음과 같다.

$$S=\left\{\begin{matrix}HHHH, & HHHT, & HHTH, & HTHH, & HHTT, & HTHT, & HTTH, & HTTT\\ THHH, & THHT, & THTH, & TTHH, & THTT, & TTHT, & TTTH, & TTTT\end{matrix}\right\}$$

$$A=\{HHHH,\ HHHT,\ HHTH,\ HTHH,\ HHTT,\ HTHT,\ HTTH,\ HTTT\}$$

$$B=\{HHTH,\ HHTT,\ HTTH,\ HTTT,\ THTH,\ THTT,\ TTTH,\ TTTT\}$$

$$C=\{HHTT,\ HTHT,\ HTTH,\ THHT,\ THTH,\ TTHH\}$$

따라서 $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{2}$, $P(C)=\frac{3}{8}$이다. 또한 각 사건들의 곱사건은 각각 다음과 같다.

$$A \cap B = \{HHTH,\ HHTT,\ HTTH,\ HTTT\}$$
$$B \cap C = \{HHTT,\ HTTH,\ THTH\}$$
$$C \cap A = \{HHTT,\ HTHT,\ HTTH\}$$
$$A \cap B \cap C = \{HHTT,\ HTTH\}$$

따라서 $P(A \cap B) = \frac{1}{4}$, $P(B \cap C) = \frac{3}{16}$, $P(C \cap A) = \frac{3}{16}$, $P(A \cap B \cap C) = \frac{1}{8}$이고 다음이 성립한다.

$$P(A \cap B) = P(A)P(B) = \frac{1}{4},$$
$$P(B \cap C) = P(B)P(C) = \frac{3}{16},$$
$$P(C \cap A) = P(A)P(C) = \frac{3}{16}$$

그러므로 세 사건은 쌍마다 독립이다.

(b) $P(A \cap B \cap C) = \frac{1}{8} \neq P(A)P(B)P(C) = \frac{1}{2} \times \frac{1}{2} \times \frac{3}{8} = \frac{3}{32}$이므로 세 사건은 독립이 아니다.

11. 선정된 사람이 미보균자일 사건을 A, 양성반응을 보일 사건을 B라 하면 다음 확률을 얻는다.

$$P(A) = \frac{95340}{100000} = 0.9534, \quad P(B) = \frac{9790}{100000} = 0.0979, \quad P(A \cap B) = \frac{5255}{100000} = 0.05255$$

(a) $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{0.05255}{0.9534} = 0.0551$

(b) $P(B^c|A^c) = \dfrac{P(A^c \cap B^c)}{P(A^c)} = \dfrac{0.00125}{0.0466} = 0.0268$

13. (a) 남자와 여자가 선정되는 사건을 각각 M, F라 하고 염색을 하고 싶은 성향을 보이는 사건을 A라 하면 구하고자 하는 확률은 다음과 같다.

$$P(A) = P(A \cap M) + P(A \cap F) = 0.08 + 0.15 = 0.23$$

(b) 선정한 사람이 남자일 확률은 다음과 같다.

$$P(M) = P(A \cap M) + P(A^c \cap M) = 0.08 + 0.40 = 0.48$$

따라서 구하고자 하는 확률은 다음과 같다.

$$P(A|M) = \frac{P(A \cap M)}{P(M)} = \frac{0.08}{0.48} = 0.167$$

(c) 선정한 사람이 남자일 확률은 다음과 같다.

$$P(F) = P(A \cap F) + P(A^c \cap F) = 0.15 + 0.37 = 0.52$$

따라서 구하고자 하는 확률은 다음과 같다.

$$P(A|F) = \frac{P(A \cap F)}{P(F)} = \frac{0.15}{0.52} = 0.288$$

15. 회사 A가 입찰서를 제출하는 사건을 A, 회사 B가 입찰서를 제출하고 선정되는 사건을 B라 하자. 그러면 $P(A) = \frac{1}{2}$, $P(B|A^c) = \frac{4}{5}$, $P(B|A) = \frac{1}{10}$이다. 따라서 전확률 공식에 의해 구하고자 하는 확률은 다음과 같다.

$$P(B) = P(A)P(B|A) + P(A^c)P(B|A^c) = \frac{1}{2} \times \frac{4}{5} + \frac{1}{2} \times \frac{1}{10} = \frac{9}{20}$$

17. (a) 임의로 선정한 사람이 전염병에 걸릴 사건을 A, 양성 반응을 보일 사건을 B라 하자. 그러면 $P(A) = 0.005$, $P(B|A) = 0.99$, $P(B|A^c) = 0.02$이다. 그러므로 임의로 선정한 사람이 전염병에 걸릴 확률은 다음과 같다.

$$\begin{aligned} P(B) &= P(A)P(B|A) + P(A^c)P(B|A^c) \\ &= 0.005 \times 0.99 + 0.995 \times 0.02 = 0.02485 \end{aligned}$$

(b) 어떤 사람이 양성 반응을 보였을 때, 이 사람이 건강한 사람일 확률은 베이즈 정리에 의해 다음과 같다.

$$P(A^c|B) = \frac{p(A^c)P(B|A^c)}{P(B)} = \frac{0.995 \times 0.02}{0.02485} = 0.8008$$

19. $P(A)=0.4$, $P(B)=0.3$, $P(C)=0.3$이고, 임의로 선정한 제품이 불량품일 사건을 D라 하면 $P(D|A)=0.02$, $P(D|B)=0.03$, $P(D|C)=0.05$이므로

(a) $P(D)=P(A)P(D|A)+P(B)P(D|B)+P(D|B)P(D|C)$

$=0.4\times0.02+0.3\times0.03+0.3\times0.05=0.032$

(b) 불량품이 A 공장에서 만들어졌을 확률 : $P(A|D)=\dfrac{0.4\times0.02}{0.032}=0.25$

불량품이 B공장에서 만들어졌을 확률 : $P(B|D)=\dfrac{0.3\times0.03}{0.032}=0.28125$

(c) 불량품이 A 또는 B 공장에서 만들어졌을 확률 : 두 사건 A와 B는 독립이므로

$$P(A\cup B|D)=P(A|D)+P(B|D)=0.25+0.28125=0.53125$$

Chapter 04 연습문제 풀이

1. (a) 0과 1의 발생이 거의 동등하므로 각각의 수가 발생할 확률은 동등하게 $\frac{1}{2}$이다. 따라서 구하고자 하는 확률질량함수는 다음과 같다.

$$f(x) = \begin{cases} \frac{1}{2}, & x = 0, 1 \\ 0, & \text{다른 곳에서} \end{cases}$$

(b) $x < 0$이면 $P(X \le x) = 0$이고 $0 \le x < 1$이면 $P(X \le x) = \frac{1}{2}$, $x \ge 1$이면 $P(X \le x) = 1$이다. 그러므로 구하고자 하는 분포함수는 다음과 같다.

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

3. 짝수들의 집합을 $A = \{2, 3, 6, \cdots\}$이라 하면, 구하고자 하는 확률은 다음과 같다.

$$P(X \in A) = f(2) + f(4) + f(5) + \cdots = \frac{2}{3^2} + \frac{2}{3^4} + \frac{2}{3^6} + \cdots$$
$$= \frac{2}{9}\left(1 + \frac{1}{3^2} + \frac{1}{3^4} + \cdots\right) = \frac{2}{9} \times \frac{1}{1-(1/9)} = \frac{1}{4}$$

5. 이산확률변수 X의 기댓값은 다음과 같다.

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x\,P(X=x) \\ &= 1 \cdot P(X=1) + 2 \cdot P(X=2) + 3 \cdot P(X=3) + 4 \cdot P(X=4) + \cdots \\ &= P(X=1) + P(X=2) + P(X=3) + 4 \cdot P(X=4) + \cdots \\ &\qquad + P(X=2) + P(X=3) + P(X=4) + \cdots \\ &\qquad\qquad + P(X=3) + P(X=4) + \cdots \\ &= P(X \ge 1) + P(X \ge 2) + \cdots = \sum_{x=1}^{\infty} P(X \ge x) \end{aligned}$$

7. 구하고자 하는 확률은 다음과 같다.

$$P(X>20)=\int_{20}^{\infty}\frac{1}{16}x\,e^{-x/4}\,dx=\left[-\frac{1}{4}(x+4)\,e^{-x/4}\right]_{20}^{\infty}=6e^{-5}$$

9. (a) 분포함수의 성질에 의해 $\lim_{x\to\infty}F(x)=\lim_{x\to\infty}(1-x^{-a})=1$이 성립해야 한다. 즉 $\lim_{x\to\infty}x^{-a}=0$이어야 하므로 $F(x)$가 분포함수이기 위한 상수 a의 조건은 $a>0$이다.

(b) $F(x)=1-x^{-a}$가 연속확률변수 X의 분포함수이면, X의 확률밀도함수는 다음과 같다.

$$f(x)=\frac{a}{x^{(a+1)}},\quad x>1$$

X의 기댓값이 존재한다면 다음과 같다.

$$E(X)=a\int_{1}^{\infty}\frac{x}{x^{a+1}}\,dx=a\int_{1}^{\infty}x^{-a}\,dx=\left[\frac{a}{1-a}x^{1-a}\right]_{1}^{\infty}$$

$$=\frac{a}{1-a}\lim_{x\to\infty}x^{1-a}-\frac{a}{1-a}$$

그러므로 X의 기댓값이 존재하기 위하여 $\lim_{x\to\infty}x^{1-a}$가 존재해야 하므로 $1-a<0$, 즉 $a>1$이어야 한다. 이때 X의 기댓값은 $E(X)=\dfrac{a}{a-1}$이다.

(c) X의 분산이 존재하기 위해 X의 2차 적률이 존재해야 한다.

$$E(X^2)=a\int_{1}^{\infty}\frac{x^2}{x^{a+1}}\,dx=a\int_{1}^{\infty}x^{-a+1}\,dx=\left[\frac{a}{2-a}x^{2-a}\right]_{1}^{\infty}$$

$$=\frac{a}{2-a}\lim_{x\to\infty}x^{2-a}-\frac{a}{2-a}$$

그러므로 X의 2차 적률이 존재하기 위하여 $\lim_{x\to\infty}x^{2-a}$가 존재해야 하므로 $2-a<0$, 즉 $a>2$이어야 한다. 이때 X의 2차 적률은 $E(X^2)=\dfrac{a}{a-2}$이다. 따라서 X의 분산은 다음과 같다.

$$\sigma^2=E(X^2)-E(X)^2=\frac{a}{a-2}-\left(\frac{a}{a-1}\right)^2=\frac{a}{(a-2)(a-1)^2}$$

11. (a) $u=\dfrac{x^2}{2\theta^2}$이라 하면 $\theta^2 du = x\,dx$이므로 다음을 얻는다.

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{\infty} \frac{x}{\theta^2} e^{-x^2/(2\theta^2)}\,dx = \int_0^{\infty} e^{-u}\,du = \left[-e^{-u}\right]_0^{\infty} = 1$$

따라서 $f(x)$는 확률밀도함수이다.

(b) X의 분포함수는 $x<0$이면 $F(x)=0$이고 $x \ge 0$이면 다음과 같다.

$$F(x) = \int_0^x \frac{u}{10^4} e^{-u^2/(2\cdot 10^4)}\,du = \left[-e^{-u^2/(2\cdot 10^4)}\right]_0^x = 1 - e^{-x^2/(2\cdot 10^4)}$$

(c) $P(X \ge 200) = 1 - F(200) = 1 - e^{-(4\cdot 10^4)/(2\cdot 10^4)} = 1 - e^{-2}$

(d) $P(100 \le X \le 200) = F(200) - F(100) = 1 - e^{-(4\cdot 10^4)/(2\cdot 10^4)}$

$$= (1-e^{-2}) - (1-e^{-1/2}) = e^{-1/2} - e^{-2}$$

13. $P(X \le a) = \displaystyle\int_0^a 2x\,dx = \left[x^2\right]_0^a = a^2 = \frac{1}{4}$이므로 $a = \dfrac{1}{2}$이다.

15. (a) $x<-1$이면 $F(x)=0$이고, $a \ge 2$이면 $F(x)=1$이다. $-1 \le x < 2$이면 $F(x)$는 다음과 같다.

$$F(x) = \int_{-1}^{x} \frac{u^2}{3}\,du = \left[\frac{u^3}{9}\right]_{-1}^{x} = \frac{1}{9}(x^3+1)$$

따라서 구하고자 하는 분포함수는 다음과 같다.

$$F(x) = \begin{cases} 0 & , x<-1 \\ \frac{1}{9}(x^3+1), & -1 \le x < 2 \\ 1 & , x \ge 2 \end{cases}$$

(b) X의 평균과 2차 적률은 각각 다음과 같다.

$$E(X) = \int_{-1}^{2} x \cdot \frac{x^2}{3}\,dx = \left[\frac{x^4}{12}\right]_{-1}^{2} = \frac{5}{4}$$

$$E(X^2) = \int_{-1}^{2} x^2 \cdot \frac{x^2}{3}\,dx = \left[\frac{x^5}{15}\right]_{-1}^{2} = \frac{11}{5}$$

따라서 평균은 $\mu = \frac{5}{4}$이고 분산은 $\sigma^2 = E(X^2) - E(X)^2 = \frac{11}{5} - \left(\frac{5}{4}\right)^2 = \frac{51}{80}$이다.

(c) 구하고자 하는 확률은 $P(-0.5 \le X \le 1) = F(1) - F(-0.5) = \frac{2}{9} - \frac{7}{72} = \frac{1}{8}$이다.

17. 분포함수를 구하면, $F(x) = \int_{-\infty}^{x} \frac{e^{-u}}{(1+e^{-u})^2}\, du = \frac{e^x}{1+e^x}$이므로 사분위수는 각각 다음과 같다.

$$F(q_1) = \frac{e^{q_1}}{1+e^{q_1}} = \frac{1}{4};\ e^{q_1} = \frac{1}{3};\ q_1 = \ln\frac{1}{3};\ q_1 = -1.09861$$

$$F(q_2) = \frac{e^{q_2}}{1+e^{q_2}} = \frac{1}{2};\ e^{q_2} = 1;\ q_2 = \ln 1;\ q_2 = 0$$

$$F(q_3) = \frac{e^{q_3}}{1+e^{q_3}} = \frac{3}{4};\ e^{q_3} = 3;\ q_3 = \ln 3;\ q_3 = 1.09861$$

19. (a) $x < -1$이면 $F(x) = 0$이고, $x \ge 1$이면 $F(x) = 1$이다. 그리고 $-1 \le x < 1$이면 $F(x)$는 다음과 같다.

$$F(x) = \int_{-1}^{x} \frac{3}{4}(1-u^2)\, du = \left[\frac{1}{4}u(3-u^2)\right]_{-1}^{x} = \frac{1}{4}(-x^3+3x+2)$$

(b) $P\left(|X| < \frac{1}{2}\right) = P\left(-\frac{1}{2} < X < \frac{1}{2}\right) = F(1/2) - F(-1/2) = \frac{27}{32} - \frac{5}{32} = \frac{11}{16} = 0.6875$

(c) 확률변수 X의 평균과 2차 적률은 다음과 같다.

$$\mu = E(x) = \int_{-1}^{1} \frac{3}{4}x(1-x^2)\, dx = \left[-\frac{3}{16}x^2(x^2-2)\right]_{-1}^{1} = 0$$

$$E(X^2) = \int_{-1}^{1} \frac{3}{4}x^2(1-x^2)\, dx = \left[-\frac{1}{20}x^3(3x^2-5)\right]_{-1}^{1} = \frac{1}{5}$$

따라서 분산은 $Var(X) = E(X^2) - E(X)^2 = \frac{1}{5}$이고 표준편차는 $\sigma = \frac{1}{\sqrt{5}}$이다.

(d) $\mu = 0$이고 $\sigma = \frac{1}{\sqrt{5}}$이므로 체비쇼프 부등식에 의해 다음을 얻는다.

$$P\left(|X| < \frac{1}{2}\right) = P\left(|X| < \frac{1}{\sqrt{5}} \cdot \frac{\sqrt{5}}{2}\right) \ge 1 - \frac{1}{(\sqrt{5}/2)^2} = \frac{1}{5}$$

Chapter 05 연습문제 풀이

1. (a) $\{(x, y) | x + y \le 5\} = \{(1,1), (1,2), (1,3), (1,4), (2,2), (2,3)\}$이므로 구하고자 하는 확률은 다음과 같다.

$$P(X+Y \le 5) = f(1,1) + \cdots + f(2,3) = \frac{1}{16} + \frac{1}{12} + \frac{1}{16} + \frac{1}{8} + \frac{1}{4} + \frac{1}{8} = \frac{17}{24}$$

(b) $X+Y \le 5$인 사건을 A라 하면, $P(A) = \frac{17}{24}$이고 $f(y|A) = \frac{P(X+Y \le 5,\ Y=y)}{P(A)}$이므로 Y의 조건부 확률질량함수는 다음 표와 같다.

Y	1	2	3	4
$f(y\|A)$	$\frac{3}{34}$	$\frac{16}{34}$	$\frac{9}{34}$	$\frac{6}{34}$

(c) Y의 조건부 평균과 조건부 2차 적률은 각각 다음과 같다.

$$\mu_{Y|X} = 1 \times \frac{3}{34} + 2 \times \frac{3}{34} + 3 \times \frac{3}{34} + 4 \times \frac{3}{34} = \frac{43}{17}$$

$$E(Y^2 | X) = 1^2 \times \frac{3}{34} + 2^2 \times \frac{3}{34} + 3^2 \times \frac{3}{34} + 4^2 \times \frac{3}{34} = \frac{122}{17}$$

그러므로 조건부 분산은 $\sigma^2_{Y|X} = E(Y^2 | X) - \mu^2_{Y|X} = \frac{122}{17} - \left(\frac{43}{17}\right)^2 = \frac{225}{289}$이다.

3. 구하고자 하는 확률은 $P(X+Y \le 1) = \int_0^1 \int_0^{1-x} 4xy\, dy\, dx = \frac{1}{6}$이다.

5. (a) 구하고자 하는 확률은 다음과 같다.

$$P(X \ge 4) = \int_4^\infty \int_0^\infty y e^{-y(1+x)} dy\, dx = \int_4^\infty \left[-\frac{xy+y+1}{(1+x)^2} e^{-y(1+x)} \right]_{y=0}^{y=\infty} dx$$

$$= \int_4^\infty \frac{1}{(1+x)^2} dx = \left[-\frac{1}{1+x} \right]_4^\infty = \frac{1}{5}$$

(b) X와 Y의 주변확률밀도함수는 각각 다음과 같다.

$$f_X(x) = \int_0^{\infty} y e^{-y(1+x)} dy = \left[-\frac{xy+y+1}{(1+x)^2} e^{-y(1+x)} \right]_{y=0}^{y=\infty} = \frac{1}{(1+x)^2}, \ x > 0$$

$$f_Y(y) = \int_0^{\infty} y e^{-y(1+x)} dx = \left[-e^{-y(1+x)} \right]_{x=0}^{x=\infty} = e^{-y}, \ y > 0$$

(c) Y의 평균과 2차 적률은 각각 다음과 같다.

$$\mu_Y = \int_0^{\infty} y e^{-y} dy = \left[-(1+y) e^{-y} \right]_0^{\infty} = 1$$

$$E(Y^2) = \int_0^{\infty} y^2 e^{-y} dy = \left[-(2+2y+2y^2) e^{-y} \right]_0^{\infty} = 2$$

따라서 분산은 $\sigma_Y^2 = E(Y^2) - \mu_Y^2 = 2 - 1 = 1$이다.

7. (a) X와 Y의 주변확률질량함수는 각각 다음과 같다.

$$f_X(x) = \begin{cases} \frac{1}{3}, & x = 0, 1, 2 \\ 0, & \text{다른 곳에서} \end{cases} \quad ; \quad f_Y(y) = \begin{cases} \frac{1}{3}, & y = 0 \\ \frac{2}{3}, & y = 1 \\ 0, & \text{다른 곳에서} \end{cases}$$

(b) $f(0,0) = \frac{1}{3} \neq f_X(0) f_Y(0) = \frac{1}{9}$이므로 X와 Y는 독립이 아니다.

(c) X와 Y의 기대값은 각각 다음과 같다.

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

또한 XY의 기대값은 $E(XY) = 0 \times 1 \times \frac{1}{3} + 1 \times 0 \times \frac{1}{3} + 2 \times 1 \times \frac{1}{3} = \frac{2}{3}$이다. 그러므로 공분산은 $Cov(X, Y) = E(XY) - E(X)E(Y) = \frac{2}{3} - 1 \cdot \frac{2}{3} = 0$이다.

9. (a) X와 Y의 주변질량함수는 각각 다음과 같다.

$$f_X(x) = \int_0^3 \frac{x+y}{27}\,dy = \left[\frac{1}{54}y(2x+y)\right]_0^3 = \frac{1}{18}(2x+3),\ 0<x<3$$

$$f_Y(y) = \int_0^3 \frac{x+y}{27}\,dx = \left[\frac{1}{54}x(x+2y)\right]_0^3 = \frac{1}{18}(2y+3),\ 0<y<3$$

(b) $0<x<3$에서 $f_X(x) = f_Y(x)$이므로 X와 Y는 항등적으로 분포한다. 그러나 $0<x<3$, $0<y<3$에서 다음 이유로 독립이 아니다.

$$f(x,y) = \frac{x+y}{27} \neq f_X(x)f_Y(y) = \frac{1}{324}(2x+3)(2y+3)$$

그러므로 X와 Y가 i.i.d 확률변수가 아니다.

(c) X와 Y는 항등적으로 분포하므로 평균과 2차 적률은 다음과 같다.

$$\mu_X = \mu_Y = \int_0^3 \frac{1}{18}x(2x+3)\,dx = \left[\frac{1}{108}x^2(4x+9)\right]_0^3 = \frac{7}{4}$$

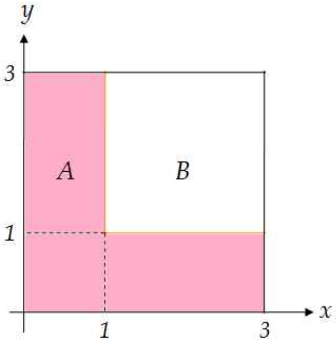
$$E(X^2) = E(Y^2) = \int_0^3 \frac{1}{18}x^2(2x+3)\,dx = \left[\frac{1}{36}x^3(x+2)\right]_0^3 = \frac{15}{4}$$

따라서 X와 Y의 분산은 다음과 같다.

$$\sigma_X^2 = \sigma_Y^2 = E(X^2) - \mu_X^2 = \frac{15}{4} - \left(\frac{7}{4}\right)^2 = \frac{11}{16}$$

(d) 이 장치가 1시간 안에 멈출 사건은 아래 그림과 같다. 따라서 이 장치가 1시간 안에 작동이 멈출 확률은 다음과 같다.

$$P(A) = 1 - P(B) = 1 - \int_1^3\int_1^3 \frac{x+y}{27}\,dx\,dy = 1 - \frac{32}{54} = \frac{11}{27}$$



11. (a) $0 < x < 2$, $0 < y < \infty$에서 X와 Y의 결합밀도함수는 다음과 같다.

$$f(x,y) = \frac{\partial^2}{\partial y \partial x} F(x,y) = \frac{\partial^2}{\partial y \partial x}\left[\frac{1}{4}x^2(e^{2y} - 1 - 2y)e^{-2y}\right]$$

$$= \frac{\partial}{\partial y}\left[\frac{1}{2}x(e^{2y} - 1 - 2y)e^{-2y}\right] = \frac{x}{2}\frac{d}{dx}(1 - e^{-2y} - 2ye^{-2y})$$

$$= \frac{x}{2}(4ye^{-2y}) = 2xye^{-2y}$$

(b) X의 주변분포함수는 다음과 같다.

$$F_X(x) = \lim_{y \to \infty} F(x,y) = \lim_{y \to \infty} \frac{1}{4}x^2(e^{2y} - 1 - 2y)e^{-2y} = \frac{1}{4}x^2, \quad 0 < x < 2$$

(c) X의 주변밀도함수는 다음과 같다.

$$f_X(x) = \frac{d}{dx}F_X(x) = \frac{d}{dx}\left(\frac{1}{4}x^2\right) = \frac{x}{2}, \quad 0 < x < 2$$

(d) 구하고자 하는 확률은 다음과 같다.

$$P(0 < X < 1, 0 < Y < 1) = F(1,1) - F(0,1) - F(1,0) + F(0,0)$$

$$= \frac{1}{4}(1 - 3e^{-2}) - 0 - 0 + 0 = \frac{1}{4}(1 - 3e^{-2})$$

13. (a) 상태공간은 $S_{XY} = \{(x,y) | 3y < x < 6, 0 < y < 2\}$이므로 다음을 만족해야 한다.

$$\int_{S_{XY}} f(x,y)\,dx\,dy = \int_0^2 \int_{3y}^6 k\,dx\,dy = \int_0^2 [kx]_{x=3y}^{x=6}\,dy$$

$$= \int_0^2 3k(2-y)\,dy = \left[\frac{3}{2}k(y^2 - 4y)\right]_0^2 = 6k = 1$$

따라서 $k = \frac{1}{6}$이다.

(b) X와 Y의 주변확률질량함수는 각각 다음과 같다.

$$f_X(x) = \int_0^{x/3} \frac{1}{6}\,dy = \left[\frac{1}{6}y\right]_{y=0}^{y=x/3} = \frac{1}{18}x, \quad 0 < x < 6$$

$$f_Y(y) = \int_{3y}^6 \frac{1}{6}\,dx = \left[\frac{1}{6}x\right]_{x=0}^{x=3y} = \frac{1}{2}y, \quad 0 < y < 2$$

(c) $X=3$일 때, Y의 조건부 확률밀도함수는 다음과 같다.

$$f_{Y|X}(y|3)=\frac{f(3,y)}{f_X(3)}=\frac{1/6}{1/6}=1, \quad 0<y<1$$

(d) Y의 조건부 평균과 조건부 2차 적률은 각각 다음과 같다.

$$\mu_{Y|X}=\int_0^1 y f_{Y|X}(y|3)\,dy=\int_0^1 y\,dy=\left[\frac{1}{2}y^2\right]_0^1=\frac{1}{2}$$

$$E(Y^2|X=3)=\int_0^1 y^2 f_{Y|X}(y|3)\,dy=\int_0^1 y^2\,dy=\left[\frac{1}{3}y^3\right]_0^1=\frac{1}{3}$$

따라서 조건부 분산은 $\sigma^2_{Y|X}=E(Y^2|X=3)-\mu^2_{Y|X}=\frac{1}{3}-\left(\frac{1}{2}\right)^2=\frac{1}{12}$이다.

15. (a) X와 Y는 각각 비복원추출에 의해 공 두 개를 뽑을 때, 빨간 공과 파란 공이 나온 횟수이므로 X와 Y가 취할 수 있는 값은 0과 1뿐이다.

(i) $[X=0,\ Y=0]$인 사건은 처음에 빨간 공이 나오지 않고 두 번째 파란 공이 나오지 않는 사건이므로 다음과 같이 구분하여 생각한다.

처음에 파란 공이 나오고 두 번째 파란 공이 아닌 빨간 공 또는 노란 공이 나오는 경우와 처음에 노란 공이 나오고 두 번째 파란 공이 아닌 빨간 공 또는 노란 공이 나오는 경우이다. 처음에 파란 공이 나올 확률은 $\frac{4}{12}$이고 이 조건 아래서 두 번째 빨간 공이 나올 확률은 $\frac{4}{11}$, 두 번째 노란 공이 나올 확률은 $\frac{4}{11}$이다. 그리고 처음에 노란 공이 나올 확률은 $\frac{4}{12}$이고 이 조건 아래서 두 번째 빨간 공이 나올 확률은 $\frac{4}{11}$, 두 번째 노란 공이 나올 확률은 $\frac{3}{11}$이다. 따라서 곱의 법칙에 의해 $[X=0,\ Y=0]$일 확률은 다음과 같다.

$$P(X=0,\ Y=0)=\frac{4}{12}\cdot\frac{4}{11}+\frac{4}{12}\cdot\frac{4}{11}+\frac{4}{12}\cdot\frac{4}{11}+\frac{4}{12}\cdot\frac{3}{11}=\frac{60}{132}$$

(ii) $[X=0,\ Y=1]$인 사건은 처음에 빨간 공이 나오지 않고 두 번째 파란 공이 나오는 사건이므로 다음과 같이 구분하여 생각한다.

처음에 파란 공이 나오고 두 번째 파란 공이 나오는 경우와 처음에 노란 공이 나오고 두 번째 파란 공이 나오는 경우이다. 처음에 파란 공이 나올 확률은 $\frac{4}{12}$이고 이 조건 아래서 두 번째 파란 공이 나올 확률은 $\frac{3}{11}$이다.

그리고 처음에 노란 공이 나올 확률은 $\frac{4}{12}$이고 이 조건 아래서 두 번째 파란 공이 나올 확률은 $\frac{4}{11}$이다. 따라서 $[X=0,\ Y=1]$일 확률은 다음과 같다.

$$P(X=0,\ Y=1)=\frac{4}{12}\cdot\frac{3}{11}+\frac{4}{12}\cdot\frac{4}{11}=\frac{28}{132}$$

(iii) $[X=1,\ Y=0]$인 사건은 처음에 빨간 공이 나오고 두 번째 파란 공이 나오지 않는 사건이므로 다음과 같이 구분하여 생각한다.

처음에 빨간 공이 나오고 두 번째 빨간 공이나 노란 공이 나오는 경우이다. 처음에 빨간 공이 나올 확률은 $\frac{4}{12}$이고 이 조건 아래서 두 번째 빨간 공이 나올 확률은 $\frac{3}{11}$이다. 그리고 처음에 빨간 공이 나올 확률은 $\frac{4}{12}$이고 이 조건 아래서 두 번째 노란 공이 나올 확률은 $\frac{4}{11}$이다. 따라서 $[X=1,\ Y=0]$일 확률은 다음과 같다.

$$P(X=1,\ Y=0)=\frac{4}{12}\cdot\frac{3}{11}+\frac{4}{12}\cdot\frac{4}{11}=\frac{28}{132}$$

(iv) $[X=1,\ Y=1]$인 사건은 처음에 빨간 공이 나오고 두 번째 파란 공이 나오는 사건이므로 구하고자 하는 확률은 다음과 같다.

$$P(X=1,\ Y=1)=\frac{4}{12}\cdot\frac{4}{11}=\frac{16}{132}$$

따라서 X와 Y의 결합확률을 나타내는 다음 표를 얻는다.

$X \backslash Y$	0	1	$f_X(x)$
0	$\frac{15}{33}$	$\frac{7}{33}$	$\frac{2}{3}$
1	$\frac{7}{33}$	$\frac{4}{33}$	$\frac{1}{3}$
$f_Y(y)$	$\frac{2}{3}$	$\frac{1}{3}$	1

(b) 표로부터 X와 Y의 주변확률질량함수는 다음과 같다.

$$f_X(x)=\begin{cases}\frac{2}{3}, & x=0\\ \frac{1}{3}, & x=1\\ 0, & \text{다른 곳에서}\end{cases}\quad,\quad f_Y(y)=\begin{cases}\frac{2}{3}, & y=0\\ \frac{1}{3}, & y=1\\ 0, & \text{다른 곳에서}\end{cases}$$

(c) $x=0, 1$에 대해 $f_X(x)=f_Y(x)$이므로 X와 Y는 항등분포이지만 다음 이유로 독립이 아니다. 그러므로 X와 Y가 i.i.d 확률변수가 아니다.

$$f(0,0)=\frac{15}{33}\neq f_X(0)f_Y(0)=\left(\frac{2}{3}\right)^2$$

(d) X와 Y가 항등분포이므로 평균과 분산은 각각 다음과 같다.

$$\mu_X=\mu_Y=0\times\frac{2}{3}+1\times\frac{1}{3}=\frac{1}{3}$$

$$\sigma_X^2=\sigma_Y^2=E(X^2)-\mu_X^2=1^2\times\frac{1}{3}-\left(\frac{1}{3}\right)^2=\frac{2}{9}$$

(e) $E(XY)=0\times0\times\frac{15}{33}+0\times1\times\frac{7}{33}+1\times0\times\frac{7}{33}+1\times1\times\frac{4}{33}=\frac{4}{33}$이므로 공분산은 다음과 같다.

$$Cov(X,Y)=E(XY)-\mu_X\mu_Y=\frac{4}{33}-\left(\frac{1}{3}\right)^2=\frac{1}{99}$$

(f) X와 Y의 표준편차가 $\sigma_X=\sigma_Y=\frac{\sqrt{2}}{3}$이므로 상관계수는 다음과 같다.

$$Corr(X,Y)=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}=\frac{1/99}{(\sqrt{2}/3)(\sqrt{2}/3)}=\frac{1}{66}$$

(g) $f_X(0)=\frac{2}{3}$이므로 구하고자 하는 조건부 확률질량함수는 다음과 같다.

$$f_{Y|X}(y|0)=\frac{f(0,y)}{f_X(0)}=\begin{cases}\frac{15/33}{2/3}, & y=0\\ \frac{7/33}{2/3}, & y=1\\ 0, & \text{다른 곳에서}\end{cases}=\begin{cases}\frac{15}{22}, & y=0\\ \frac{7}{22}, & y=1\\ 0, & \text{다른 곳에서}\end{cases}$$

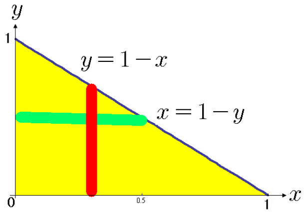
17. (a) X와 Y의 상태공간이 그림과 같다. 따라서 주변확률밀도함수는 각각 다음과 같다.

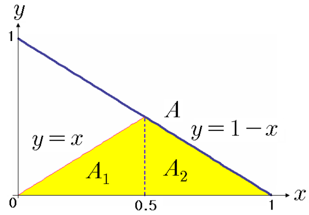
$$f_X(x)=\int_0^{1-x}24xy\,dy=12xy^2\Big|_{y=0}^{1-x}=12x(1-x)^2,\quad 0<x<1$$

$$f_Y(y)=\int_0^{1-y}24xy\,dx=12x^2y\Big|_{x=0}^{1-y}=12y(1-y)^2,\quad 0<y<1$$

(b) 상태공간 안에서 $X > Y$인 영역은 그림의 A와 같이 $x = \frac{1}{2}$에 의하여 분할된다. 그러므로 구하고자 하는 확률은 다음과 같다.

$$P(X > Y) = \int_0^{1/2} \int_0^{x} 24xy\,dy\,dx + \int_{1/2}^{1} \int_0^{1-x} 24xy\,dy\,dx = \frac{3}{16} + \frac{5}{16} = \frac{1}{2}$$





(c) X와 Y가 항등분포를 이루므로 평균과 분산은 각각 다음과 같다.

$$\mu_X = \mu_Y = \int_0^1 12x^2(1-x)^2\,dx = \left[\frac{12}{5}x^5 - 6x^4 + 4x^3\right]_0^1 = \frac{2}{5}$$

$$E(X^2) = \int_0^1 12x^3(1-x)^2\,dx = \left[2x^6 - \frac{24}{5}x^5 + 3x^4\right]_0^1 = \frac{1}{5}$$

$$\sigma_X^2 = \sigma_Y^2 = E(X^2) - \mu_X^2 = \frac{1}{5} - \left(\frac{2}{5}\right)^2 = \frac{1}{25}$$

(d) 먼저 $E(XY)$를 구한다.

$$E(XY) = \int_0^1 \int_0^{1-x} 24x^2y^2\,dy\,dx = \int_0^1 \left[8x^2y^3\right]_{y=0}^{y=1-x} dx = \int_0^1 8(1-x)^3x^2\,dx$$

$$= \left[8\left(\frac{x^3}{3} - \frac{3}{4}x^4 + \frac{3}{5}x^5 - \frac{x^6}{6}\right)\right]_0^1 = \frac{2}{15}$$

따라서 공분산은 $Cov(X, Y) = E(XY) - \mu_X\mu_Y = \frac{2}{15} - \left(\frac{2}{5}\right)^2 = -\frac{2}{75}$이다.

(e) $\sigma_X = \sigma_Y = \frac{1}{5}$이므로 상관계수는 $Corr(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = -\frac{2/75}{(1/5)^2} = -\frac{2}{3}$이다.

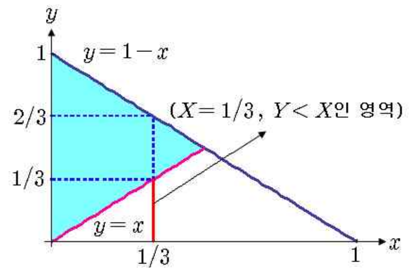
19. X의 주변확률밀도함수는 $f_X(x) = \int_0^{1-x} 24xy\,dy = 12x(1-x)^2,\ 0 < x < 1$이다.

따라서 $f_X(1/3) = \frac{16}{9}$이다. 그러므로 $X = \frac{1}{3}$일 때, Y의 조건부 확률밀도함수는 다음과 같다.

$$f\left(y \mid x = \frac{1}{3}\right) = \frac{f(1/3,\ y)}{f_X(1/3)} = \frac{8y}{16/9} = \frac{9}{2}y, \quad 0 < y < \frac{2}{3}$$

따라서 구하고자 하는 확률은 다음과 같다.

$$P\left(Y < X \mid X = \frac{1}{3}\right) = \int_0^{1/3} \frac{9}{2}y\,dy = \left[\frac{9}{4}y^2\right]_0^{1/3} = \frac{1}{4}$$



Chapter 06 연습문제 풀이

1. X의 1차 적률과 2차 적률을 구하면 각각 다음과 같다.

$$E(X)=\sum_{x=-3}^{3} x f(x)=\frac{1}{7}\sum_{x=-3}^{3} x=0$$

$$E(X^2)=\sum_{x=-3}^{3} x f(x)=\frac{1}{7}\sum_{x=-3}^{3} x^2=\frac{28}{4}=7$$

따라서 평균은 $\mu=E(X)=0$이고 표준편차는 $\sigma^2=E(X^2)-\mu^2=7$이다.

3. (a) 노이즈 지수가 3dB을 초과하는 증폭기 수를 X라 하면, $X\sim B(20,\,0.05)$이므로 구하고자 하는 평균과 분산은 각각 다음과 같다.

$$\text{평균 : } \mu=20\times 0.05=1$$
$$\text{분산 : } \sigma^2=20\times 0.05\times 0.95=0.95$$

(b) 이항누적확률표를 이용하면 구하고자 하는 확률은 다음과 같다.

$$P(X=3)=P(X\le 3)-P(X\le 2)=0.9841-0.9245=0.0596$$

(c) 이항누적확률표를 이용하면 구하고자 하는 확률은 다음과 같다.

$$P(X\ge 2)=1-P(X\le 1)=1-0.7358=0.2642$$

5. (a) 매우 오염된 표본과 약간 오염된 표본 그리고 청정한 표본의 수를 각각 $X,\ Y,\ Z$라 하면, $r_1=19,\ r_2=6,\ r_3=5$이고 $n=5$이므로 확률함수는 다음과 같다.

$$P(X=x,\ Y=y,\ Z=z)=\frac{\binom{19}{x}\binom{6}{y}\binom{5}{z}}{\binom{30}{5}},\quad \begin{array}{l} x+y+z=5 \\ x,\ y,\ z=0,\ 1,\ 2,\ 3,\ 4,\ 5 \end{array}$$

(b) 매우 심각하게 오염된 표본이 3, 약간 오염된 표본이 1이면, 청정한 표본은 1이므로 구하고자 하는 확률은 다음과 같다.

$$P(X=3,\ Y=1,\ Z=1)=\frac{\binom{19}{3}\binom{6}{1}\binom{5}{1}}{\binom{30}{5}}=\frac{1615}{7917}\approx 0.204$$

(c) 매우 심각하게 오염된 표본의 수에 관점을 둔다면, $X \sim H(30, 19, 5)$이고, X의 확률함수는 다음과 같다.

$$P(X=x) = \frac{\binom{19}{x}\binom{11}{5-x}}{\binom{30}{5}}, \quad x = 0,\ 1,\ 2,\ 3,\ 4,\ 5$$

그러므로 5개 표본 중에서 적어도 4개에서 심각하게 오염되었을 확률은 다음과 같다.

$$P(X=4 \text{ 또는 } X=5) = \frac{\binom{19}{4}\binom{11}{1}}{\binom{30}{5}} + \frac{\binom{19}{5}\binom{11}{0}}{\binom{30}{5}} = \frac{7106}{23751} + \frac{1938}{23751} = \frac{1292}{3393} = 0.3808$$

7. (a) $X \sim G(0.2)$이므로 $\mu = \frac{1}{0.2} = 5$이다.

(b) 구하고자 하는 확률은 $P(X=4) = f(4) = (0.2)(0.8)^3 = 0.1024$이다.

9. (a) 0.001㎣의 혈액을 채취하여 백혈구의 수를 X라 하면, 1㎣ 당 평균 6,000개의 백혈구가 있으므로 0.001㎣의 혈액 안에 평균 6개의 백혈구가 있다.

(b) $X \sim P(6)$이므로 혈액 안에서 기껏해야 2개가 관찰될 확률은 $P(X \le 2) = 0.062$이다.

11. (a) $X \sim B(100,\ 0.02)$이므로 확률질량함수는

$$f(x) = \binom{100}{x}(0.02)^x(0.98)^{100-x}, \quad x = 0,\ 1,\ \cdots,\ 100$$

이다. 그러므로 구하고자 하는 확률은 $P(X=3) = f(3) = \binom{100}{3}(0.02)^3(0.98)^{97} = 0.182276$이다.

(b) $\mu = 100 \times 0.02 = 2$이므로 $X \approx P(2)$이다. 따라서 구하고자 하는 근사확률은 다음과 같다.

$$P(X=3) = P(X \le 3) - P(X \le 2) = 0.857 - 0.677 = 0.180$$

13. (a) 실험결과가 나타날 때까지 걸리는 시간을 X라 하면, $X \sim U(12,\ 14)$이므로 확률밀도함수는 다음과 같다.

$$f(x) = \frac{1}{14-12} = \frac{1}{2}, \quad 12 \le x \le 14$$

(b) 평균과 분산은 각각 다음과 같다.

$$\text{평균 : } \mu = \frac{12+14}{2} = 13$$
$$\text{분산 : } \sigma^2 = \frac{(14-12)^2}{12} = \frac{1}{3}$$

(c) 12시간 30분은 12.5시간이므로 구하고자 하는 확률은 다음과 같다.

$$P(X < 12.5) = \frac{1}{2} \times (12.5 - 12) = \frac{1}{4}$$

15. (a) $X \sim \mathrm{Exp}(1/\lambda)$이므로 평균거리는 $\mu = \lambda$이다.

(b) $X \sim \mathrm{Exp}(1/10)$이므로 생존함수는 $S(x) = e^{-x/10}$, $x > 0$이다. 따라서 비기억성 성질에 의해 구하고자 하는 확률은 다음과 같다.

$$P(X > 5 | X > 3) = P(X > 2) = S(2) = e^{-2/10} = 0.8187$$

17. (a) 1분에 평균 3개의 입자가 푸아송분포에 따라 들어오므로 X는 모수 $\alpha = 4$, $\beta = \frac{1}{3}$인 감마분포에 따른다. 그러므로 X의 확률밀도함수는 다음과 같다.

$$f(x) = \frac{1}{\Gamma(4)(1/3)^4} x^{4-1} e^{-3x} = \frac{27}{2} x^3 e^{-3x}, \quad 0 < x < \infty$$

(b) 구하고자 하는 평균과 분산은 각각 $\mu = \alpha\beta = \frac{4}{3}$, $\sigma^2 = \alpha\beta^2 = \frac{4}{9}$이다.

(c) 구하고자 하는 확률은 다음과 같다.

$$P(X \geq 5) = \int_5^\infty \frac{27}{2} x^3 e^{-3x}\, dx = -\frac{1}{2}\left[(9x^3 + 9x^2 + 6x + 2)e^{-3x}\right]_5^\infty = 691 e^{-15} = 0.0002$$

19. 구하고자 하는 평균과 분산은 각각 다음과 같다.

$$\text{평균 : } E(X) = \frac{3.5}{3.5 + 4.4} = 0.4430$$
$$\text{분산 : } Var(X) = \frac{3.5 \times 4.4}{(3.5 + 4.47)^2 \times (3.5 + 4.4 + 18)} = 0.0277$$

Chapter 07 연습문제 풀이

1. (a) $P(Z \geq 2.65) = 1 - \Phi(2.65) = 1 - 0.9960 = 0.0040$

   (b) $P(Z < -1.11) = 1 - \Phi(1.11) = 1 - 0.8665 = 0.1335$

   (c) $P(Z > -1.27) = P(Z < 1.27) = \Phi(1.27) = 0.8980$

   (d) $P(-1.02 \leq Z \leq 1.02) = 2P(0 \leq Z \leq 1.02) = 2\Phi(1.02) - 1 = 2(0.8461 - 0.5) = 0.6922$

3. (a) <부록 A.5>로부터 $t_{0.1}(7) = 1.415$이다.

   (b) $t_{0.99}(7) = -t_{0.01}(7) = -2.998$이다.

5. $X \sim N(4, 4)$이므로 표준화하면 $Z = \dfrac{X-4}{2} \sim N(0, 1)$이다.

   (a) $z = \dfrac{4.5-4}{2} = 0.25$이므로 $P(X \geq 4.5) = P(Z \geq 0.25) = 1 - \Phi(0.25) = 1 - 0.5987 = 0.4013$이다.

   (b) $z = \dfrac{6.5-4}{2} = 1.25$이므로 $P(X < 6.5) = \Phi(1.25) = 0.8944$이다.

   (c) $z = \dfrac{2.5-4}{2} = -0.75$이므로 구하고자 하는 확률은 다음과 같다.

$$P(X \leq 2.5) = P(Z \leq -0.75) = 1 - \Phi(0.75) = 1 - 0.7734 = 0.2266$$

   (d) $z_l = \dfrac{3-4}{2} = -0.5$, $z_r = \dfrac{7-4}{2} = 1.5$이므로 구하고자 하는 확률은 다음과 같다.

$$\begin{aligned} P(3 \leq X \leq 7) &= P(-0.5 \leq Z \leq 1.5) = \Phi(1.5) - \Phi(-0.5) \\ &= \Phi(1.5) - (1 - \Phi(0.5)) = 0.9332 - (1 - 0.6915) = 0.6247 \end{aligned}$$

7. (a) 표준정규분포표로부터 $z_0 = 2.98$이다.

(b) $P(Z \le z_0) = P(Z \ge -z_0) = 1 - P(Z \le -z_0) = 0.0154$이므로 $P(Z \le -z_0) = 0.9846$이다.
따라서 $-z_0 = 2.16$, 즉 $z_0 = -2.16$이다.

(c) $P(0 \le Z \le z_0) = 0.3554$이므로 $P(Z \le z_0) = 0.8554$이고 $z_0 = 1.06$이다.

(d) $P(-z_0 \le Z \le z_0) = 2[P(Z \le z_0) - 0.5] = 0.9030$이므로 $P(Z \le z_0) = 0.9515$이고 $z_0 = 1.66$이다.

(e) $P(-z_0 \le Z \le z_0) = 2[P(Z \le z_0) - 0.5] = 0.2052$이므로 $P(Z \le z_0) = 0.6026$이고 $z_0 = 0.26$이다.

(f) $P(Z \ge z_0) = P(Z \le -z_0) = 0.6915$이므로 $-z_0 = 0.50$, 즉 $z_0 = -0.50$이다.

9. $i = 1, 2, \cdots, 100$에 대하여 $X_i \sim \mathrm{Exp}(1/4)$이므로 $\mu = E(X_i) = 4$, $\sigma^2 = Var(X_i) = 16$이다. 그러므로 중심극한정리에 의하여 $\overline{X} \approx N(4, 0.16)$이다. 그러므로 구하고자 하는 확률은 다음과 같다.

$$\begin{aligned} P(|\overline{X} - 4| \ge 0.5) &= P(\overline{X} \le 3.5) + P(\overline{X} \ge 4.5) \\ &\approx P\left(Z \le \frac{3.5 - 4}{\sqrt{0.16}}\right) + P\left(Z \ge \frac{4.5 - 4}{\sqrt{0.16}}\right) \\ &= P(Z \le -1.25) + P(Z \ge 1.25) = 2[1 - P(Z \le 1.25)] \\ &= 2(1 - 0.8944) = 0.2112 \end{aligned}$$

11. $T \sim t(r)$이므로 $Var(T) = \dfrac{r}{r-2} = 1.25$; $r = (1.25) \cdot (r-2) = 1.25r - 2.5$; $r = 10$이다.
또한 $t_\alpha(10) = 2.228$를 만족하는 $\alpha = 0.025$이다. 즉, $P(T > 2.228) = 0.025$이고 $P(T < -2.228) = 0.025$이다. 그러므로 구하고자 하는 확률은 다음과 같다.

$$P(|T| \le 2.228) = 1 - 2 \times 0.025 = 0.95$$

13. (a) $X \sim \chi^2(4)$이면, $X \sim \Gamma(2, 2)$이므로 구하고자 하는 확률밀도함수는 다음과 같다.

$$f(x) = \frac{1}{\Gamma(2) 2^2} x e^{-x/2} = \frac{1}{4} x e^{-x/2}, \quad x > 0$$

(b) $r = 4$이므로 구하고자 하는 평균과 분산은 각각 $\mu = 4$, $\sigma^2 = 8$이다.

(c) 구하고자 하는 확률은 다음과 같다.

$$P(2 < X < 4) = \int_2^4 \frac{1}{4} x e^{-x/2} = \left[-\frac{1}{2}(x+2)e^{-x/2}\right]_2^4 = -3e^{-2} + 2e^{-1} = (2e-3)e^{-2}$$

(d) 카이제곱분포표로부터 95% 백분위수는 $\chi^2_{0.05}(4) = 9.49$이다.

15. (a) $\mu_X = \exp\left(\mu + \frac{\sigma^2}{2}\right) = \exp\left(1 + \frac{1.5}{2}\right) = e^{1.75} = 5.765$

(b) $\sigma_X^2 = \left(e^{\sigma^2} - 1\right)\exp(2\mu + \sigma^2) = \left(e^{1.5} - 1\right)e^{2+1.5} = 115.298$

(c) $x_{0.25} = e^{1+(1.5)\cdot(-z_{0.75})} = e^{1+(1.5)\cdot(-0.675)} = 0.9876$;

$x_{0.5} = e^{1+(1.5)\cdot(-z_{0.5})} = e^{1+(1.5)\cdot(0)} = 2.7183$;

$x_{0.75} = e^{1+(1.5)\cdot(z_{0.75})} = e^{1+(1.5)\cdot(0.675)} = 7.482$

(d) $P(4 \le X \le 6) = F(6) - F(4) = \Phi\left(\frac{(\ln 6) - 4}{\sqrt{1.5}}\right) - \Phi\left(\frac{(\ln 4) - 4}{\sqrt{1.5}}\right)$

$$= \Phi(-1.80) - \Phi(-2.13) = \Phi(2.13) - \Phi(1.80)$$

$$= 0.9834 - 0.9641 = 0.0193$$

17. 저항의 크기를 X라 하면, $X \sim N(4, 0.25^2)$이고, 4.32를 표준화하면 다음과 같다.

$$z_r = \frac{4.32 - 4}{0.25} = 1.28$$

따라서 이 공장에서 생산된 저항을 전기회로에 사용할 확률은 다음과 같다.

$$P(X \le 4.32) = P(Z \le 1.28) = \Phi(1.28) = 0.8997$$

그러므로 이 공장에서 생산된 10,000개의 저항 중에 전기회로에 사용할 수 있는 저항의 수는 8,997개이다.

19. $\mu = np = 8$, $\sigma^2 = npq = 4.8$이므로 $X \approx N(8, 2.19^2)$에 근사한다.

(a) $z = (10.5 - 8)/2.19 = 1.14$이므로 다음을 얻는다.

$$P(X \le 10) = P(X \le 10.5) = P(Z \le 1.14) \approx 0.8729$$

(b) $z_l = (6.5-8)/2.19 = -0.68$, $z_r = (11.5-8)/2.19 = 1.60$이므로 다음을 얻는다.

$$\begin{aligned} P(7 \le X \le 11) &= P(6.5 \le X \le 11.5) = P(-0.68 \le Z \le 1.60) \\ &= P(Z \le 1.60) + P(Z \le 0.68) - 1 \\ &= 0.9452 + 0.7517 - 1 = 0.6969 \end{aligned}$$

(c) $z = (14.5-8)/2.19 = 2.97$이므로 다음을 얻는다.

$$\begin{aligned} P(X \ge 15) &= P(X \ge 14.5) = P(Z \ge 2.97) \\ &= 1 - \Phi(2.97) = 1 - 0.9985 = 0.0015 \end{aligned}$$

Chapter 08 연습문제 풀이

1. (a) 복원추출에 의하여 두 개를 꺼내므로 표본으로 나올 수 있는 모든 경우는 다음과 같다.

$$\{1,1\}\ \ \{1,2\}\ \ \{2,1\}\ \ \{2,2\}$$

(b) (a)에서 구한 4개의 표본평균을 구하면, 1, 1.5, 2이다.

(c) $\overline{x}=1$인 경우, $\{1,1\}$이 나올 확률은 $0.6\times0.6=0.36$
$\overline{x}=1.5$인 경우, $\{1,2\}$, $\{2,1\}$이 나올 확률은 $2\times0.6\times0.4=0.48$
$\overline{x}=2$인 경우, $\{2,2\}$가 나올 확률은 $0.4\times0.4=0.16$

따라서 표본평균의 확률분포는 다음과 같다.

$\overline{x}$	1	1.5	2
$p_{\overline{X}}$	0.36	0.48	0.16

(d) 표본평균 $\overline{X}$의 평균은 $\mu_{\overline{X}}=\sum\overline{x}\,p_{\overline{x}}=1.4$이고 분산은 $\sigma^2_{\overline{X}}=E(\overline{X}^2)-E(\overline{X})^2=0.12$이다.

3. $\sigma^2=36$, $n=25$이므로 $\overline{X}\sim N(\mu,1.44)$, 즉 $\dfrac{\overline{X}-\mu}{1.2}\sim N(0,1)$이다. 그러므로 구하고자 하는 확률은 다음과 같다.

$$P(|\overline{X}-\mu|\geq 3)=P\left(|Z|\geq\frac{3}{1.2}\right)=P(|Z|\geq 2.5)=2P(Z\geq 2.5)$$

$$=2[1-\Phi(2.5)]=0.0124$$

5. (a) 배기가스에 포함된 질소산화물의 양을 X라 하면, $X\sim N(0.47,0.05^2)$이므로 구하고자 하는 확률은 다음과 같다.

$$P(X\leq 0.5)=P\left(Z\leq\frac{0.5-0.47}{0.05}\right)=P(Z\leq 0.6)=0.7257$$

(b) $n=16$이므로 $\overline{X}$의 분산은 $\sigma_{\overline{X}}^2 = \dfrac{0.05^2}{16} = 0.0125^2$이고 $\overline{X} \sim N(0.47, 0.0125^2)$이다.

따라서 구하고자 하는 확률은 다음과 같다.

$$P(\overline{X} \le 0.5) = P\left(Z \le \frac{0.5-0.47}{0.0125}\right) = P(Z \le 2.4) = 0.9918$$

7. 표본분산 $s^2=9$, $n=16$이고 모분산을 모르므로 $T = \dfrac{\overline{X}-\mu}{3/4} = \dfrac{\overline{X}-\mu}{0.75} \sim t(15)$이다. 따라서 다음을 얻는다.

$$P(|\overline{X}-\mu| < k) = P\left(\frac{|\overline{X}-\mu|}{0.75} < \frac{k}{0.75}\right) = 0.95;$$

$$P\left(T \ge \frac{k}{0.75}\right) = P(T \ge t_{0.05}(15)) = 0.05$$

그러므로 구하고자 하는 상수는 다음과 같다.

$$\frac{k}{0.75} = t_{0.05}(15) = 1.753;\ k = 1.753 \times 0.75 = 1.31475$$

9. 표본평균 $\overline{X}$의 확률분포는 중심극한정리에 의하여 평균 $\mu_{\overline{X}} = 40$, 분산 $\sigma_{\overline{X}}^2 = \dfrac{4}{100} = 0.2^2$인 정규분포에 근사한다. 그러므로 구하고자 하는 근사확률은 다음과 같다.

$$P(39.45 < \overline{X} < 40.25) = P\left(\frac{39.45-40}{0.2} < Z < \frac{40.25-40}{0.2}\right) = P(-2.75 < Z < 1.25)$$

$$\approx \Phi(1.25) - [1-\Phi(2.75)] = 0.8944 - (1-0.9970) = 0.8914$$

11. (a) $p=0.11$, $q=0.89$, $n=200$이므로 $\sqrt{pq/n} \fallingdotseq \sqrt{0.00049} = 0.0221^2$이다. 그러므로 표본비율 $\hat{p}$의 근사확률분포는 $\hat{p} \approx N(0.11, 0.0221^2)$이다.

(b) $\dfrac{\hat{p}-0.11}{0.0221} \approx N(0,1)$이므로 구하고자 하는 확률은 다음과 같다.

$$P(\hat{p} < 0.07) = P\left(Z < \frac{0.07-0.11}{0.0221}\right) = P(Z < -1.81) = P(Z > 1.81)$$

$$= 1 - \Phi(1.81) = 1 - 0.9649 = 0.0351$$

(c) 95% 백분위수를 $p_{0.05}$라 하면, $\dfrac{p_{0.05}-0.11}{0.0221} = z_{0.05} = 1.645$이므로 구하고자 하는 백분위수는 $p_{0.05} = 0.11 + 1.645 \times 0.0221 = 0.14635$이다.

13. (a) $\mu_1 - \mu_2 = 2$이고 $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{9}{16} + \frac{9}{25} = 0.96^2$이므로 $\overline{X} - \overline{Y} \sim N(2,\ 0.96^2)$이다.

(b) 구하고자 하는 확률은 다음과 같다.

$$P(1 < \overline{X} - \overline{Y} < 3) = P\left(\frac{1-2}{0.96} < Z < \frac{3-2}{0.96}\right) = P(-1.04 < Z < 1.04)$$
$$= 2\Phi(1.04) - 1 = 2 \times 0.8508 - 1 = 0.7016$$

(c) 95% 백분위수를 x_0이라 하면 다음을 얻는다.

$$\frac{x_0 - 2}{0.96} = z_{0.05} = 1.645;\ \ x_0 = 2 + 0.96 \times 1.645 = 3.5792$$

15. (a) $\sigma^2 = 0.04$이고 두 표본의 크기가 각각 10이므로 합동표본분산은 $\frac{18}{0.04} S_p^2 \sim \chi^2(18)$이다.
그러므로 구하고자 하는 근사확률은 다음과 같다.

$$P(S_p^2 \le 0.0139) = P\left(\frac{18}{0.04} S_p^2 \le \frac{18 \times 0.0139}{0.04}\right) = P(V \le 6.255) \approx 0.995$$

(b) $\sigma_1^2 = \sigma_2^2 = 0.04$이고 두 표본의 크기가 각각 10이므로 $\frac{S_2^2/0.04}{S_1^2/0.04} = \frac{S_2^2}{S_1^2} \sim F(9,\ 9)$이다.
그러므로 구하고자 하는 확률은 $P\left(S_2^2 \ge (10.11) S_1^2\right) = P\left(\frac{S_2^2}{S_1^2} \ge 10.11\right) = 0.001$이다.

(c) 두 지역의 풍속에 대한 합동표본분산은 $s_p^2 = \frac{9}{10 + 10 - 2}(0.21^2 + 0.23^2) = 0.0485$이다.

(d) A와 B 지역의 평균 풍속을 각각 μ_A, μ_B라 하면 $\mu_B - \mu_A = 0$이고 $s_p = \sqrt{0.0485} = 0.22$이다.
두 지역의 표본 평균 풍속을 $\overline{X}$, $\overline{Y}$라 하면 $\frac{\overline{Y} - \overline{X}}{(0.22) \cdot \sqrt{\frac{1}{10} + \frac{1}{10}}} = \frac{\overline{Y} - \overline{X}}{0.0984} \sim t(18)$이다.
따라서 구하고자 하는 확률은 다음과 같다.

$$P(\overline{Y} - \overline{X} \ge 0.251) = P\left(\frac{\overline{Y} - \overline{X}}{0.0984} \ge \frac{0.251}{0.0984}\right) = P(T \ge 2.551) \approx 0.01$$

17. (a) r을 먼저 구하면 다음과 같다.

$$r = \frac{[(s_1^2/n)+(s_2^2/m)]^2}{\dfrac{(s_1^2/n)^2}{n-1}+\dfrac{(s_2^2/m)^2}{m-1}}$$

$$= \frac{(4/12+16/15)^2}{\dfrac{(4/12)^2}{11}+\dfrac{(16/15)^2}{14}} = \frac{33957}{1583} = 21.45$$

따라서 소숫점 이하를 잘라내면, 자유도는 $r \approx 21$이다. 또한 $\mu_1 - \mu_2 = 2$이고 다음을 얻는다.

$$\sqrt{\frac{S_1^2}{n}+\frac{S_2^2}{m}} = \sqrt{\frac{4}{12}+\frac{16}{15}} = \sqrt{1.4} = 1.1832$$

그러므로 $U = \dfrac{(\overline{X}-\overline{Y})-2}{1.1832} \sim t(21)$이다.

(b) $U = \dfrac{(\overline{X}-\overline{Y})-2}{1.1832} \sim t(21)$이므로 다음이 성립한다.

$$P(\overline{X}-\overline{Y} > t_0) = P\left(\frac{(\overline{X}-\overline{Y})-2}{1.1832} > \frac{t_0-2}{1.1832}\right)$$

$$= P(T > t_{0.05}(21)) = P(T > 1.721) = 0.05$$

따라서 구하고자 하는 t_0은 다음과 같다.

$$\frac{t_0-2}{1.1832} = 1.721 \;;\; t_0 = 2 + 1.1832 \times 1.721 = 4.036$$

18. (a) 기존 방법에 의한 불량률을 $p_1 = 0.05$, 새로운 방법에 의한 불량률을 $p_2 = 0.02$라 하면, $p_1 - p_2 = 0.02$이고 $\sqrt{\dfrac{0.05 \times 0.95}{1500}+\dfrac{0.03 \times 0.97}{2000}} = 0.0068$이므로 $\hat{p}_1 - \hat{p}_2$의 확률분포는 $\hat{p}_1 - \hat{p}_2 \approx N(0.02,\, 0.0068^2)$이다.

(b) 구하고자 하는 확률은 다음과 같다.

$$P(\hat{p}_1 - \hat{p}_2 \ge 0.03) = P\left(\frac{(\hat{p}_1-\hat{p}_2)-0.02}{0.0068} \ge \frac{0.03-0.02}{0.0068}\right) = P(Z \ge 1.47)$$

$$= 1 - \Phi(1.47) = 1 - 0.9292 = 0.0708$$

(c) 관찰된 표본비율 $\hat{p}_1 = \dfrac{72}{1500} = 0.048$, $\hat{p}_2 = \dfrac{68}{2000} = 0.034$이므로 $\hat{p}_1 - \hat{p}_2 = 0.014$이다. 따라서 $\hat{p}_1 - \hat{p}_2$이 0.014보다 클 확률은 다음과 같다.

$$\begin{aligned} P\left(\hat{p}_1 - \hat{p}_2 > 0.014\right) &= P\left(\frac{\hat{p}_1 - \hat{p}_2 - 0.02}{0.0068} > \frac{0.014 - 0.02}{0.0068}\right) \\ &= P(Z > -0.88) \\ &= \Phi(0.88) = 0.8106 \end{aligned}$$

(d) $z_{0.025} = 1.96$이므로 다음이 성립한다.

$$\begin{aligned} P\left(\hat{p}_1 - \hat{p}_2 > p_0\right) &= P\left(Z > \frac{p_0 - 0.02}{0.0068}\right) \\ &= P(Z > z_{0.025}) = P(Z > 1.96) = 0.025 \end{aligned}$$

따라서 구하고자 하는 p_0은 다음과 같다.

$$\frac{p_0 - 0.02}{0.0068} = 1.96; \quad p_0 = 0.02 + 1.96 \times 0.0068 = 0.0333$$

Chapter 09 연습문제 풀이

1. (a) 각 추정량의 평균을 구하면 다음과 같다.

$$E(\hat{\mu}_1) = \frac{1}{3}E(X_1 + X_2 + X_3) = \frac{1}{3}(\mu + \mu + \mu) = \mu$$

$$E(\hat{\mu}_2) = \frac{1}{4}E(X_1 + 2X_2 + X_3) = \frac{1}{4}(\mu + 2\mu + \mu) = \mu$$

$$E(\hat{\mu}_3) = \frac{1}{5}E(X_1 + 2X_2 + X_3) = \frac{1}{5}(\mu + 2\mu + \mu) = \frac{4}{5}\mu$$

그러므로 각 추정량의 편의는 다음과 같다.

$$b_1 = E(\hat{\mu}_1) - \mu = 0,\ b_2 = E(\hat{\mu}_2) - \mu = 0,\ b_3 = E(\hat{\mu}_3) - \mu = -\frac{1}{5}\mu$$

(b) 불편추정량 : $\hat{\mu}_1$, $\hat{\mu}_2$ 편의추정량 : $\hat{\mu}_3$

(c) $Var(\hat{\mu}_1) = \frac{1}{9}Var(X_1 + X_2 + X_3) = \frac{1}{9}(1+1+1) = \frac{1}{3}$

$Var(\hat{\mu}_2) = \frac{1}{16}Var(X_1 + X_2 + 2X_3) = \frac{1}{16}(1 + 4 \cdot 1 + 1) = \frac{3}{8}$

따라서 $Var(\hat{\mu}_1) < Var(\hat{\mu}_2)$이므로 최소분산불편추정량은 $\hat{\mu}_1$ 이다.

3. $E(X) = 8p$, $Var(X) = 8p(1-p)$이므로 $E(\hat{p}) = E\left(\frac{X}{10}\right) = \frac{8p}{10} = \frac{4p}{5}$이다. 그러므로 $\hat{p}$의 편의와 분산은 각각 다음과 같다.

$$\text{bias} = E(\hat{p}) - p = \frac{4p}{5} - p = -\frac{p}{5},\quad Var(\hat{p}) = Var\left(\frac{X}{10}\right) = \frac{Var(X)}{100} = \frac{2p(1-p)}{25}$$

5. (a) $e = 1.96 \times \sqrt{\frac{5}{50}} = 0.6198$ (b) $e = 1.96 \times \sqrt{\frac{15}{50}} = 1.0735$

(c) $e = 1.96 \times \frac{5}{\sqrt{50}} = 1.3859$ (d) $e = 1.96 \times \sqrt{\frac{35}{50}} = 1.6399$

7. $\overline{x}=99$, $z_{0.025}=1.96$, $\sigma=5$ $n=25$이므로 μ에 대한 95% 신뢰구간은 다음과 같다.

$$\left(\overline{x}-z_{0.025}\frac{\sigma}{\sqrt{n}}\ ,\ \overline{x}+z_{0.025}\frac{\sigma}{\sqrt{n}}\right)=\left(150-(1.96)\cdot\frac{5}{\sqrt{30}}\ ,\ 150+(1.96)\cdot\frac{5}{\sqrt{30}}\right)$$
$$=(148.211,\ 151.789)$$

9. 표본으로부터 얻은 표본평균은 $\overline{x}=334.13$이다. 한편 $\sigma=8.16$, $n=30$, $z_{0.025}=1.96$이므로 95% 신뢰구간은 다음과 같다.

$$\left(\overline{x}-z_{0.025}\frac{\sigma}{\sqrt{n}},\ \overline{x}+z_{0.025}\frac{\sigma}{\sqrt{n}}\right)=\left(334.13-(1.96)\cdot\frac{8.16^2}{\sqrt{30}},\ 334.13+(1.96)\cdot\frac{8.16^2}{\sqrt{30}}\right)$$
$$=(310.303,\ 357.957)$$

11. $\overline{x}=100$, $\overline{y}=94$이므로 $\mu_1-\mu_2$의 점추정값은 $\overline{x}-\overline{y}=6$이다. $\sigma_1^2=25$, $\sigma_2^2=16$이고 $n=16$, $m=25$이므로 표본오차는 $\mathrm{S.E.}(\overline{X}-\overline{Y})=\sqrt{\frac{25}{16}+\frac{16}{25}}=\sqrt{2.2025}=1.484$이다. 그리고 95% 오차한계는 $e=1.96\times\mathrm{S.E.}(\overline{X}-\overline{Y})=1.96\times1.484=2.91$이다. 따라서 구하고자 하는 신뢰구간은 다음과 같다.

$$(6-2.91,\ 6+2.91)=(3.09,\ 8.91)$$

13. 약품 A와 B의 표본평균을 각각 $\overline{x}$, $\overline{y}$라 하면 $\mu_1-\mu_2$에 대한 점추정값은 $\overline{x}-\overline{y}=2.2$이다. $\sigma_1^2=1.44$, $\sigma_2^2=2.25$이고 $n=15$, $m=12$이므로 표준오차는 다음과 같다.

$$\mathrm{S.E.}(\overline{X}-\overline{Y})=\sqrt{\frac{1.44}{15}+\frac{2.25}{12}}=0.532$$

그리고 95% 오차한계는 $e=1.96\times\mathrm{S.E.}(\overline{X}-\overline{Y})=1.96\times0.532=1.0427$이므로 구하고자 하는 신뢰구간은 $(2.2-1.0427,\ 2.2+1.0427)=(1.1573,\ 2.2427)$이다.

15. (a) 45번의 발사 중에서 성공한 횟수가 38번이므로 표본비율은 $\hat{p}=\frac{40}{45}=0.889$이므로 모비율 p에 대한 점추정값은 $\hat{p}=0.889$이다. 또한 90% 오차한계는 $e=1.645\times\sqrt{\frac{0.889\times0.111}{45}}=0.077$이므로 90% 신뢰구간은 $(0.889-0.077,\ 0.889+0.077)=(0.812,\ 0.966)$이다.

(b) 새로운 추진체에 의한 90% 신뢰구간의 하한이 기존 방법에 의한 성공률 $p=0.08$보다 크므로 새로운 추진체에 의한 로켓 발사가 기존의 방법보다 좋다고 할 수 있다.

17. $\hat{p}_1 = \frac{12}{40} = 0.3$, $\hat{p}_2 = \frac{6}{20} = 0.3$이므로 $\hat{p}_1 - \hat{p}_2 = 0$이고, 90% 오차한계는 다음과 같다.

$$e = 1.645\sqrt{\frac{0.3 \times 0.7}{40} + \frac{0.3 \times 0.7}{20}} = 0.2447$$

그러므로 90% 신뢰구간은 $(-0.2447, 0.2447)$이다.

19. $\sigma_1^2 = 2.4336$, $\sigma_2^2 = 2.1025$이고 오차범위가 $d = 0.85$이므로 95% 신뢰구간을 얻기 위한 표본의 크기는 다음과 같다.

$$n = m \geq \left(\frac{1.96}{0.85}\right)^2 (2.4336 + 2.1025) = 18.8 ; \; n = m = 19$$

Chapter 10 연습문제 풀이

1. ① 귀무가설 $H_0 : \mu = 902$와 대립가설 $H_0 : \mu \neq 902$를 설정한다.
   ② 유의수준 $\alpha = 0.05$에 대한 양측검정의 기각역은 $Z \leq -1.96$ 또는 $Z \geq 1.96$이다.
   ③ 표본표준편차가 $s = 25.1$, $n = 50$이므로 검정통계량은 $Z = \dfrac{\overline{X} - 902}{25.1/\sqrt{50}}$이고 $\overline{x} = 895$이므로 검정통계량의 관찰값은 $z_0 = \dfrac{895 - 902}{25.1/\sqrt{50}} = -1.972$이다.
   ④ 검정통계량의 관찰값이 기각역 안에 들어가므로 이 지역의 대기배출 농도가 평균 902ppm이라는 결론은 불충분하다.

3. (a) $\alpha = 0.05$에 대한 양측검정의 기각역은 $Z \leq -1.96$ 또는 $Z \geq 1.96$이다.

   (b) 모표준편차가 $\sigma = 5$이므로 검정통계량은 $Z = \dfrac{\overline{X} - 50}{5/\sqrt{36}}$이고 $\overline{x} = 48.5$이므로 검정통계량의 관찰값은 $z_0 = \dfrac{48.5 - 50}{5/\sqrt{36}} = -1.8$이다.

   (c) p-값$= 2[1 - P(Z < 1.8)] = 2(1 - 0.9641) = 0.0718$

   (d) 검정통계량의 관찰값 $z_0 = -1.8$은 기각역 안에 놓이지 않으므로 유의수준 5%에서 귀무가설 $H_0 : \mu = 50$을 기각할 수 없다. 즉, 모평균이 $\mu = 50$이라는 주장은 타당성이 있다.

   (e) p-값$= 0.0718 > \alpha = 0.05$이므로 귀무가설 $H_0 : \mu = 50$을 유의수준 5%에서 기각할 수 없다.

5. ① 평균 절단강도가 180을 초과한다는 회사의 주장이므로 대립가설로 설정한다.
   즉, 귀무가설 $H_0 : \mu \leq 180$과 대립가설 $H_1 : \mu > 180$을 설정한다.
   ② 유의수준 $\alpha = 0.1$에 대한 상단측검정이므로 기각역은 $Z \geq 2.326$이다.
   ③ $\sigma = 6.2$이므로 검정통계량은 $Z = \dfrac{\overline{X} - 180}{6.2/\sqrt{25}} = \dfrac{\overline{X} - 180}{1.24}$이다.
   ④ $\overline{x} = 182.7$이므로 검정통계량의 관찰값은 $z_0 = \dfrac{182.7 - 180}{1.24} = 2.177$이다.
   ⑤ 검정통계량의 관찰값 $z_0 = 2.177$은 기각역 안에 놓이지 않으므로 이 회사의 주장은 타당성이 없다.

7. ① 비타민 처리 그룹과 그렇지 않은 그룹의 평균을 각각 μ_1, μ_2라 하면 귀무가설 $H_0 : \mu_1 - \mu_2 = 0$과 대립가설 $H_1 : \mu_1 - \mu_2 \neq 0$을 설정한다.
② 유의수준 $\alpha = 0.05$에 대한 양측검정의 기각역은 $Z \leq -1.96$ 또는 $Z \geq 1.96$이다.
③ 표본표준편차가 각각 $\sigma_1^2 = 1.4$, $\sigma_1^2 = 2.2$이고 $n = 100$, $m = 65$므로 검정통계량은 다음과 같다.

$$Z = \frac{(\overline{X} - \overline{Y}) - 0}{\sqrt{(1.4^2/100) + (2.2^2/65)}} = \frac{\overline{X} - \overline{Y}}{0.3067}$$

④ $\overline{x} = 5.2$, $\overline{y} = 5.8$이므로 검정통계량의 관찰값은 $z_0 = \dfrac{5.2 - 5.8}{0.3067} = -1.956$이다.
⑤ $z_0 = -1.956$이므로 검정통계량의 관찰값은 기각역 안에 놓이지 않는다. 따라서 회복기간이 같다는 근거는 충분하다.

9. ① 여성과 남성의 평균 수명을 각각 μ_1, μ_2라 하면, 귀무가설 $H_0 : \mu_1 - \mu_2 \geq 8$과 대립가설 $H_1 : \mu_1 - \mu_2 < 8$을 설정한다.
② 유의수준 $\alpha = 0.05$에 대한 하단측검정의 기각역은 $Z \leq -1.645$이다.
③ 모표준편차가 각각 $\sigma_1 = 5.2$, $\sigma_2 = 6.8$이고 $n = m = 50$이므로 검정통계량은 다음과 같다.

$$Z = \frac{(\overline{X} - \overline{Y}) - 8}{\sqrt{(5.2^2 + 6.8^2)/50}} = \frac{(\overline{X} - \overline{Y}) - 8}{1.21}$$

④ $\overline{x} - \overline{y} = 5.9$이므로 검정통계량의 관찰값은 $z_0 = \dfrac{5.9 - 8}{1.21} = -1.73$이다.
⑤ 관찰값 $z_0 = -1.73$은 기각역 안에 놓이므로 $H_0 : \mu_1 - \mu_2 \geq 8$를 기각한다. 즉, 여성이 남성보다 8년 이상 오래 산다는 근거는 신빙성이 없다.

11. ① 귀무가설은 $H_0 : p = 0.85$이고 대립가설은 $H_1 : p \neq 0.85$이다.
② 유의수준 $\alpha = 0.1$에 대한 양측검정 기각역은 $Z \leq -1.645$ 또는 $Z \geq 1.645$이다.
③ 검정통계량은 $Z = \dfrac{\hat{p} - 0.85}{\sqrt{0.85 \times 0.15/40}} = \dfrac{\hat{p} - 0.85}{0.0565}$이다.
④ 표본비율이 $\hat{p} = \dfrac{37}{40} = 0.925$이므로 검정통계량의 관찰값은 $z_0 = \dfrac{0.925 - 0.85}{0.0565} = 1.33$이다.
⑤ 검정통계량의 관찰값 $z_0 = 1.33$은 기각역 안에 놓이지 않으므로 건축물의 85%가 열펌프를 설치했다는 주장은 타당성이 있다.

13. ① 귀무가설은 $H_0 : p \le 0.9$이고 대립가설은 $H_1 : p > 0.9$이다.

② 유의수준 $\alpha = 0.05$에 대한 상단측검정 기각역은 $Z \ge z_{0.05} = 1.645$이다.

③ 검정통계량은 $Z = \dfrac{\hat{p} - 0.9}{\sqrt{0.9 \times 0.1/45}} = \dfrac{\hat{p} - 0.9}{0.0447}$이다.

④ 표본비율이 $\hat{p} = \dfrac{43}{45} = 0.956$이므로 검정통계량의 관찰값은 $z_0 = \dfrac{0.956 - 0.9}{0.0447} = 1.253$이다.

⑤ 검정통계량의 관찰값 $z_0 = 1.253$은 기각역 안에 놓이지 않으므로 로켓 발사 성공률이 90%를 초과한다는 주장은 타당성이 없다.

15. ① 두 의학 단체 A와 B가 주장하는 흡연가의 비율을 각각 p_1, p_2라 하면, A의 주장이 B의 주장보다 큰가를 검정하므로 귀무가설은 $H_0 : p_1 - p_2 \le 0$이고 대립가설은 $H_1 : p_1 - p_2 > 0$이다.

② 유의수준 5%에서 상단측검정이므로 기각역은 $Z \ge 1.645$이다.

③ 합동표본비율은 $\hat{p} = \dfrac{648 + 551}{2000 + 1800} = 0.3155$이므로 검정통계량은 다음과 같다.

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{0.3155 \times 0.6845 \left(\frac{1}{2000} + \frac{1}{1800} \right)}} = \frac{\hat{p}_1 - \hat{p}_2}{0.0151}$$

④ 두 생산라인의 표본비율은 각각 $\hat{p}_1 = \dfrac{648}{2000} = 0.324$, $\hat{p}_2 = \dfrac{551}{1800} = 0.306$이므로 검정통계량의 관찰값은 다음과 같다.

$$z_0 = \frac{0.324 - 0.306}{0.0151} = 1.192$$

⑤ 검정통계량의 관찰값 $z_0 = 1.192$은 기각역 안에 놓이지 않으므로 귀무가설을 기각하지 않는다. 즉, A가 주장하는 흡연가의 비율이 더 크다는 주장이 설득력이 없다.

17. ① 각 처방약의 신뢰 비율을 각각 p_1, p_2, p_3이라 하면, 귀무가설 $H_0 : p_1 = p_2 = p_3$과 대립가설 "H_1 : H_0가 아니다."를 설정한다.

② 범주의 수가 3개이므로 자유도 2인 카이제곱분포에서 유의수준 $\alpha = 0.05$에 대한 기각역은 $\chi^2 \ge 5.99$이다.

③ 합동표본비율은 $\hat{p} = \dfrac{74}{390} = 0.19$이다. 따라서 각 처방전에 대한 기댓값은 각각 다음과 같다.

$$132 \times 0.19 = 25.05,\ 139 \times 0.19 = 26.41,\ 114 \times 0.19 = 21.66$$

이제 검정통계량 $\chi^2 = \displaystyle\sum_{i=1}^{3} \frac{(o_i - e_i)^2}{e_i}$의 값을 구한다.

범주	관찰도수(o_i)	비율($\hat{p}$)	기대도수 ($e_i = n_i\hat{p}$)	$o_i - e_i$	$(o_i - e_i)^2$	$\frac{(o_i - e_i)^2}{e_i}$
1	24	0.19	25.05	-1.05	1.10	0.044
2	38	0.19	26.41	11.59	134.33	5.086
3	12	0.19	21.66	-9.66	93.32	4.308
합계						9.438

④ 검정통계량의 관찰값 $\chi_0^2 = 9.438$은 기각역 안에 놓이므로 귀무가설 H_0을 기각한다. 즉, 각 처방전에 대한 신뢰 정도가 동일하다는 근거는 불충분하다.

Chapter 11 연습문제 풀이

1. $\overline{x}=8.2$, $s=0.32$이고 $t_{0.025}(12)=2.179$이므로 오차한계는 $e=2.179\times\dfrac{0.32}{\sqrt{13}}=0.193$이다.

따라서 평균 ph에 대한 95% 신뢰구간은 $(8.2-0.193,\ 8.2+0.193)=(8.007,\ 8.393)$이다.

3. 두 표본에 대한 합동표본분산과 합동표본표준편차는 다음과 같다.

$$s_p^2=\frac{5\times1.9^2+7\times1.6^2}{6+8-2}=2.9975,\ \ s_p=\sqrt{2.9975}=1.731$$

자유도 12인 t-분포에서 $t_{0.025}(12)=2.179$이므로 95% 신뢰구간에 대한 오차한계는 다음과 같다.

$$e=2.179\times1.731\times\sqrt{\frac{1}{6}+\frac{1}{8}}=2.037$$

이때 $\overline{x}-\overline{y}=2.7$이므로 $\mu_1-\mu_2$에 대한 95% 신뢰구간은 다음과 같다.

$$(2.7-2.037,\ 2.7+2.037)=(0.663,\ 4.737)$$

5. (a) 표본평균과 표본분산은 각각 다음과 같다.

$$\overline{x}=\frac{1}{12}(32.5+32.9+\cdots+32.3)=32.85$$
$$s^2=\frac{1}{11}\sum_{i=1}^{10}(x_i-32.85)^2=0.717$$

그리고 자유도 11인 χ^2-분포에서 $\chi^2_{0.025}(11)=21.92$, $\chi^2_{0.975}(11)=3.82$이므로 모분산 σ^2에 대한 95% 신뢰구간의 하한과 상한은 각각 다음과 같다.

$$l=\frac{11\times0.717}{21.92}=0.3598,\ \ u=\frac{11\times0.717}{3.82}=2.0647$$

따라서 구하고자 하는 신뢰구간은 $(0.3598,\ 2.0647)$이다.

(b) $\sqrt{0.3598}=0.5998$, $\sqrt{2.0647}=1.4369$이므로 모표준편차 σ에 대한 95% 신뢰구간은 $(0.5998,\ 1.4369)$이다.

7. (a) 평균과 분산을 구하면 각각 다음과 같다.

$$\overline{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 14.61$$

$$s^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i - 14.61)^2 = \frac{7.069}{9} = 0.7854$$

$s = 0.886$이고 $t_{0.025}(9) = 2.262$이므로 오차한계는 $e = 2.262 \times \frac{0.886}{\sqrt{10}} = 0.634$이다. 그러므로 평균 당분함량에 대한 90% 신뢰구간은 다음과 같다.

$$(14.61 - 0.634,\ 14.61 + 0.634) = (13.976,\ 15.244)$$

(b) 자유도 9인 χ^2-분포로부터 $\chi^2_{0.975}(9) = 2.70$, $\chi^2_{0.025}(9) = 19.02$이므로 σ^2에 대한 95% 신뢰구간의 상한과 하한은 각각 $l = \frac{9 \times 0.7854}{19.02} = 0.570$, $u = \frac{9 \times 0.7854}{2.70} = 2.618$이다. 따라서 σ에 대한 95% 신뢰구간의 상한과 하한은 각각 $l = \sqrt{0.570} = 0.755$, $u = \sqrt{135.14} = 1.618$이고, 95% 신뢰구간은 $(0.755,\ 1.618)$이다.

9. ① 귀무가설 $H_0 : \mu = 8$과 대립가설 $H_1 : \mu \neq 8$을 설정한다.

② $t_{0.025}(12) = 2.179$이므로 $\alpha = 0.05$에 대한 양측검정의 기각역은 $T \leq -2.179$ 또는 $T \geq 2.179$이다.

③ $s = 0.32$이므로 검정통계량은 $T = \frac{\overline{X} - 8}{0.32/\sqrt{13}}$이다.

④ $\overline{x} = 8.2$이므로 검정통계량의 관찰값은 $t_0 = \frac{8.2 - 8}{0.32/\sqrt{13}} = 2.25$이다.

⑤ 관찰값이 기각역 안에 놓이므로 귀무가설을 기각한다. 즉, 하천 물의 평균 ph 농도가 8이라는 주장은 근거가 불충분하다.

11. ① 귀무가설 $H_0 : \mu \leq 48$과 대립가설 $H_1 : \mu > 48$(주장)을 설정한다.

② $t_{0.05}(14) = 1.761$이므로 $\alpha = 0.05$에 대한 상단측검정의 기각역은 $T > 1.761$이다.

③ $s = 2.2$이므로 검정통계량은 $T = \frac{\overline{X} - 48}{2.2/\sqrt{15}}$이다.

④ $\overline{x} = 49.1$이므로 검정통계량의 관찰값은 $t_0 = \frac{49.1 - 48}{2.2/\sqrt{15}} = 1.936$이다.

⑤ 관찰값이 기각역 안에 놓이므로 귀무가설을 기각한다. 즉, 배터리의 평균수명이 48개월을 초과한다는 주장은 타당성이 있다.

13. 대도시와 소도시의 가솔린 평균 가격을 각각 μ_1, μ_2라 하자.

① 귀무가설 $H_0 : \mu_1 - \mu_2 \leq 0$와 대립가설 $H_1 : \mu_1 - \mu_2 > 0$(주장)를 설정한다.
② $n = 15$, $m = 15$이므로 $t_{0.01}(28) = 2.467$이고 $\alpha = 0.01$에 대한 상단측검정이므로 기각역은 $T \geq 2.467$이다.
③ 합동표본분산과 합동표준편차는 각각 다음과 같다.

$$s_p^2 = \frac{14 \times 18^2 + 14 \times 14^2}{15 + 15 - 2} = 260, \ s_p = \sqrt{260} = 16.12$$

따라서 검정통계량은 다음과 같다.

$$T = \frac{(\overline{X} - \overline{Y}) - 0}{16.12 \times \sqrt{\frac{1}{15} + \frac{1}{15}}} = \frac{\overline{X} - \overline{Y}}{5.886}$$

④ $\overline{x} - \overline{y} = 16$이므로 검정통계량의 관찰값은 $t_0 = \frac{16}{5.886} = 2.718$이다.
⑤ 관찰값이 기각역 안에 놓이므로 귀무가설을 기각한다. 즉, 대도시의 가솔린 평균 가격이 소도시보다 높다고 할 수 있다.

15. 우선 실험 전후의 측정값에 대한 차를 구한다.

	1	2	3	4	5	6	7	8	9	10
d_i	0.3	0.5	0.3	1.2	0.4	−0.1	−0.2	0.6	0.4	0.3

① 실험 전후의 평균을 각각 μ_1, μ_2라 하면, 귀무가설은 $H_0 : \mu_1 = \mu_2$, 대립가설은 $H_1 : \mu_1 \neq \mu_2$이다.
② 유의수준 $\alpha = 0.05$에 대한 양측검정이고 이때 자유도 9인 t-분포를 사용하므로 기각역은 $R : |T| > t_{0.025}(9) = 2.262$이다.
③ d_i에 대한 평균과 표준편차를 구한다.

$$\overline{d} = \frac{1}{10}(0.3 + 0.5 + 0.3 + \cdots + 0.3) = 0.37$$

그리고 $\sum d_i^2 = 2.69$, $(\sum d_i)^2 = 13.69$이므로 타수의 차에 대한 표준편차는 다음과 같다.

$$s_d = \sqrt{\frac{10 \times 2.69 - 13.69}{9 \times 10}} = 0.3831$$

④ 검정통계량 $T = \frac{\overline{d}}{0.3831/\sqrt{10}} = \frac{\overline{d}}{0.1211}$의 관찰값은 $t_0 = \frac{0.37}{0.1211} = 3.0553$이므로 기각역 안에 놓인다.
⑤ 따라서 귀무가설을 기각한다. 즉, 실험 전후의 평균에 차이가 있다고 할 수 있다.

17. ① 귀무가설 $H_0 : \sigma \le 2$과 대립가설 $H_1 : \sigma > 2$(주장)를 귀무가설 $H_0 : \sigma^2 \le 4$와 대립가설 $H_1 : \sigma^2 > 4$(주장)로 변경하여 설정한다.

② 크기 10이므로 $\chi^2_{0.05}(9) = 16.92$이고 $\alpha = 0.05$에 대한 상단측검정의 기각역은 $V > 16.92$이다.

③ 검정통계량은 $V = \dfrac{9S^2}{4}$이고 $s = 2.5$이므로 관찰값은 $v_0 = \dfrac{9 \times 2.5^2}{4} = 14.0625$이다.

④ 관찰값이 기각역 안에 놓이지 않으므로 귀무가설을 기각할 수 없다. 즉, 모표준편차가 2보다 크다는 주장은 근거가 불충분하다.

19. ① 귀무가설 $H_0 : \sigma_1^2 = \sigma_2^2$과 대립가설 $H_0 : \sigma_1^2 \ne \sigma_2^2$을 설정한다.

② $n = 10,\ m = 7,\ \alpha = 0.05$이므로 양측검정의 기각역은 다음과 같다.

$$F < f_{0.975}(9,\ 6) = \frac{1}{f_{0.025}(6,\ 9)} = \frac{1}{4.32} = 0.2315 \text{ 또는 } F > f_{0.025}(9,\ 6) = 5.52$$

③ 검정통계량 $F = S_1^2 / S_2^2$을 선택한다.

④ $s_1 = 3.1,\ s_2 = 2.7$이므로 검정통계량의 관찰값은 $f_0 = 3.1^2/1.7^2 = 3.325$이다.

⑤ 검정통계량의 관찰값 $f_0 = 3.325$는 기각역 안에 놓이지 않으므로 유의수준 5%에서 두 모분산이 동일하다는 주장은 타당성이 있다.

Chapter 12 연습문제 풀이

1. $[0, t]$에서 상점에 찾아온 손님의 수를 $N(t)$라 하면, 발생비율이 $\lambda = 4$이므로 $N(t) \sim P(4t)$이다. 한편 30분은 1/2시간이므로 8시부터 8시 30분 사이에 상점에 찾아오는 손님의 수에 대한 확률분포는 $N(1/2) \sim P(2)$이다. 그러므로 이 시간에 손님 1명이 찾아올 확률은 다음과 같다.

$$P(N(1/2) = 1) = \frac{2^1}{1!}e^{-2} = 2e^{-2}$$

한편 8시 30분까지 꼭 한 사람이 찾아오고 11시까지 찾아온 손님이 모두 5 사람이 되려면, 8시 30분부터 11시까지 4명이 상점을 찾아와야 하므로

$$\begin{aligned} P(N(1/2) = 1,\, N(3) - N(1/2) = 4) &= P(N(1/2) = 1)\, P(N(3) - N(1/2) = 4) \\ &= P(N(1/2) = 1)\, P(N(5/2) = 4) \end{aligned}$$

이다. 이때 $N(5/2) \sim P(10)$이고 따라서 구하고자 하는 확률은 다음과 같다.

$$\begin{aligned} P(N(1/2) = 1,\; N(3) - N(1/2) = 4) &= P(N(1/2) = 1)\, P(N(5/2) = 4) \\ &= (2e^{-2}) \times \left(\frac{10^4}{4!}e^{-10}\right) = 0.2707 \times 0.0189 = 0.0051 \end{aligned}$$

3. (a) 판넬이 시간당 평균 1.8의 비율인 푸아송과정에 따라 도착하므로 푸아송과정의 비율은 $\lambda = 1.625$이다.

(b) 판넬생산공정에 도착하는 두 판넬 사이의 대기시간을 T라 하면, $T \sim \mathrm{Exp}(\lambda)$이므로 T는 모수 1.625인 지수분포를 이룬다. 따라서 T의 평균시간은 $E(T) = 1/1.625 = 0.6154$이다.

(c) $T \sim \mathrm{Exp}(1.625)$이므로 생존함수는 $S(x) = e^{-1.625x},\ x > 0$이고 따라서 구하고자 하는 확률은 $P(X \geq 1) = S(1) = e^{-1.625} = 0.1969$이다.

(d) t시간 동안 도착한 판넬의 수 $N(t)$는 모수 $\lambda t = 1.625t$인 푸아송분포에 따르므로 $N(4) \sim P(6.5)$이다.

(e) $N(4) \sim P(6.5)$이므로 <부록 A.2>로부터 다음을 얻는다.

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.043 = 0.957$$

5. 1㎠ 당 평균 5개의 흔적을 가지고 있으므로 2㎠에 평균 10개의 흔적을 가지고 있으며, 분열 흔적의 수는 $X \sim P(10)$이다. 그러므로 2㎠에 많아야 3개의 흔적을 가질 확률은 다음과 같다.

$$P(X \le 3) = \frac{10^0}{0!}e^{-10} + \frac{10^1}{1!}e^{-10} + \frac{10^2}{2!}e^{-10} + \frac{10^3}{3!}e^{-10}$$

$$= 0.00005 + 0.00045 + 0.00227 + 0.00757 = 0.01034$$

7. (a) $[0, t]$에서 계수기를 통과한 방사능 물질의 수를 $N(t)$라 하면, 발생비율이 $\lambda = 3$이므로 $N(t) \sim P(3t)$이다. 그러므로 $N(1) \sim P(3)$이다. 따라서 구하고자 하는 확률은 <부록 A.2>로부터 $P(N(1) = 2) = P(N(1) \le 2) - P(N(1) \le 1) = 0.423 - 0.199 = 0.224$이다.

(b) $P(N(1) \ge 5) = 1 - P(N(1) \le 4) = 1 - 0.815 = 0.185$

9. (a) $N(t) \sim P(3t)$이므로

$$P[N(1) \ge 4] = 1 - e^{-3} - 3e^{-3} - \frac{9}{2}e^{-3} - \frac{9}{2}e^{-3} = 1 - 13e^{-3} = 0.3528$$

(b) $P[N(1) \ge 4 | N(1) \ge 1] = \dfrac{P[N(1) \ge 1,\ N(1) \ge 4]}{P[N(1) \ge 1]} = \dfrac{P[N(1) \ge 4]}{P[N(1) \ge 1]}$

$$= \frac{1 - 13e^{-3}}{1 - e^{-3}} = \frac{0.3528}{0.9502} = 0.0345$$

(c) $P[N(5) - N(3) \ge 1] = P[N(2) \ge 1] = 1 - e^{-6} = 0.9975$

(d) $E[N(5) - N(3) \,|\, N(2) = 2] = E[N(5) - N(3)] = E[N(2)] = 6$

11. (a) $n = 0, 1, 2, \cdots$에 대하여 $X_n \sim B(1, p)$이므로 $N_n \sim B(n, p)$이다. 따라서 $E(N_n) = np$, $Var(N_n) = npq$이다. 그러므로 N_n^2의 기댓값은 다음과 같다.

$$E(N_n^2) = Var(N_n) + E(N_n)^2 = npq + n^2p^2$$

(b) 전확률 공식에 의해 다음을 얻는다.

$$P(N_{n+1} = k) = \sum_i P(N_n = i)P(N_{n+1} = k | N_n = i)$$

그리고 X_{n+1}은 X_1, $\cdots$, X_n과 독립이므로 X_{n+1}은 $N_n = X_1 + \cdots + X_n$과 독립이다. 따라서 다음을 얻는다.

$$P(N_{n+1} = k \mid N_n = i) = P(X_{n+1} + N_n = k \mid N_n = i) = P(X_{n+1} = k - i \mid N_n = i)$$

$$= P(X_{n+1} = k - i) = \begin{cases} p, & i = k-1 \\ q, & i = k \\ 0, & \text{다른 곳에서} \end{cases}$$

따라서 N_n을 이용하여 $P(N_{n+1} = k)$를 나타내면 다음과 같다.

$$P(N_{n+1} = k) = p\,P(N_n = k-1) + q\,P(N_n = k)$$

(c) $N_n \sim B(n, p)$이므로 다음을 얻는다.

$$P(N_n = k-1) = \binom{n}{k-1} p^{k-1} q^{n-k+1}, \quad P(N_n = k) = \binom{n}{k} p^k q^{n-k}$$

따라서 N_{n+1}의 확률분포는 다음과 같다.

$$\begin{aligned} P(N_{n+1} = k) &= p\,P(N_n = k-1) + q\,P(N_n = k) \\ &= p\binom{n}{k-1} p^{k-1} q^{n-k+1} + q\binom{n}{k} p^k q^{n-k} \\ &= n!\left(\frac{1}{(k-1)!\,(n-k+1)!} + \frac{1}{k!\,(n-k)!}\right) p^k q^{n+1-k} \\ &= \frac{(n+1)!}{k!\,(n+1-k)!} p^k q^{n+1-k} = \binom{n+1}{k} p^k q^{n+1-k} \end{aligned}$$

즉, $N_{n+1} \sim B(n+1, p)$이다.

13. 점심시간 1시간 동안 이 식당을 찾아온 사람에 따른 자동차의 수를 N^1, $\cdots$, N^7이라 하자. 그러면 N^1, $\cdots$, N^7는 각각 모수 4, 6, 12, 14, 2, 1.2, 0.8인 푸아송과정에 따른다. 따라서 다음 기댓값을 얻는다.

$$E(N^1) = 4,\ E(N^2) = 6,\ E(N^3) = 12,\ E(N^4) = 14,\ E(N^5) = 2,\ E(N^6) = 1.2,\ E(N^7) = 0.8$$

따라서 구하고자 하는 평균은 다음과 같다.

$$\begin{aligned} &E(N^1 + 2N^2 + 3N^3 + 4N^4 + 5N^5 + 6N^6 + 7N^7) \\ &\qquad = 4 + 2\times 6 + 3\times 12 + 4\times 14 + 5\times 2 + 6\times 1.2 + 7\times 0.8 = 131.8 \end{aligned}$$

15. (a) $0 \leftrightarrow 1 \leftrightarrow 2$이므로 기약 동치류이다.
(b) $0 \leftrightarrow 2$이므로 동치류 $\{0, 2\}$와 $\{1\}$로 구분된다.
(c) $0 \to 1 \to 2 \to 2$이므로 동치류가 없다.

17. (a) $0 \leftrightarrow 1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 4 \leftrightarrow 5$이므로 상태공간은 기약동치류이다.

(b) 구하고자 하는 추이행렬은 다음과 같다.

$$\mathrm{P} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{3}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 \\ \frac{3}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 \\ \frac{3}{4} & 0 & 0 & 0 & \frac{1}{4} & 0 \\ \frac{3}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} \\ \frac{3}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} \end{pmatrix}$$

(c) 두 번 이동한 추이행렬은 다음과 같다.

$$\mathrm{P}^2 = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{3}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 \\ \frac{3}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 \\ \frac{3}{4} & 0 & 0 & 0 & \frac{1}{4} & 0 \\ \frac{3}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} \\ \frac{3}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} \end{pmatrix} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{3}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 \\ \frac{3}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 \\ \frac{3}{4} & 0 & 0 & 0 & \frac{1}{4} & 0 \\ \frac{3}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} \\ \frac{3}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} & \frac{3}{16} & \frac{1}{16} & 0 & 0 & 0 \\ \frac{3}{4} & \frac{3}{16} & 0 & \frac{1}{16} & 0 & 0 \\ \frac{3}{4} & \frac{3}{16} & 0 & 0 & \frac{1}{16} & 0 \\ \frac{3}{4} & \frac{3}{16} & 0 & 0 & 0 & \frac{1}{16} \\ \frac{3}{4} & \frac{3}{16} & 0 & 0 & 0 & \frac{1}{16} \\ \frac{3}{4} & \frac{3}{16} & 0 & 0 & 0 & \frac{1}{16} \end{pmatrix}$$

따라서 구하고자 하는 확률은 다음과 같다.

$$P(X_2 = 2 | X_0 = 3) = \mathrm{P}_{32}^2 = \frac{3}{16}$$

(d) $0 \to 0$, $0 \to 1 \to 0$, $\cdots$, $0 \to 1 \to 2 \to 3 \to 4 \to 5 \to 5 \to 0$이므로 $d(0) = 1$이고 기약 동치류이므로 $d(0) = d(1) = d(2) = d(3) = d(4) = d(5) = 1$이다.