1. 해저 케이블에 생긴 결함의 수는 1km당 발생 비율 $\lambda=0.15$인 푸아송과정에 따른다고 한다.

(1) 처음부터 3km지점까지 결함이 발견되지 않을 확률을 구하여라.

(2) 처음부터 3km지점까지 결함이 발견되지 않았다고 할 때, 3km지점부터 4km지점까지 결함이 발견되지 않을 확률을 구하여라.

(3) 처음부터 3km지점까지 결함이 1개, 3km지점부터 4km지점까지 결함이 1개 발견되지만, 4km지점부터 5km지점까지 결함이 발견되지 않을 확률을 구하여라.

(sol)

(1) 단위 시간당 발생비율이 $\lambda=0.15$이고, $(0,t)$에서 발견된 결함의 횟수를 $X_{(0,t)}$라 하면,

$X_{(0,t)} \sim Poisson\,P(0.15t)$이므로 $X_{(0,t)}$의 확률질량함수는

$$f_X(x)=\frac{(0.45t)^x}{x!}e^{-0.45t},\ x=0,1,2,\cdots$$

이다. 따라서 $P(X_{(0,3)}=0)=f_X(0)=\frac{(0.45)^0}{0!}e^{-0.45}=e^{-0.45}$이다.

(2) $P(X_{(3,4)}=0|X_{(0,t)}=0)=P(X_{(3,4)}=0)=P(X_{(0,1)}=0)=f_X(0)=\frac{(0.15)^0}{0!}e^{-0.15}=e^{-0.15}$이다.

(3) $$\begin{aligned}P(X_{(0,3)}=1,X_{(3,4)}=1,X_{(4,5)}=0)&=P(X_{(0,3)}=1)P(X_{(3,4)}=1)P(X_{(4,5)}=0)\\&=P(X_{(0,3)}=1)P(X_{(0,1)}=1)P(X_{(0,1)}=0)\\&=\frac{(0.45)^1}{1!}e^{-0.45}\frac{(0.15)^1}{1!}e^{-0.15}\frac{(0.15)^0}{0!}e^{-0.15}=0.0675e^{-0.75}\end{aligned}$$

2. 어느 집단의 구성원이 사망할 때까지 걸리는 시간은 평균 60년인 지수분포를 이룬다고 한다.

(1) 임의로 선정된 사람이 80세 이후까지 생존할 확률을 구하여라.

(2) 임의로 선정된 사람이 40세까지 생존했을 때, 이 사람이 50세 이전에 사망할 확률을 구하여라.

(sol)

(1) 구성원이 사망할 때까지 걸리는 시간을 T라 하면, 평균 $m=\frac{1}{\lambda}=60$인 지수분포를 따르므로,

T의 확률밀도함수와 누적분포함수는

$$f_T(t)=\frac{1}{60}e^{-\frac{t}{60}}\ \ (t>0)$$
$$F_T(t)=\int_0^t\frac{1}{60}e^{-\frac{x}{60}}dx=1-e^{-\frac{t}{60}}\ \ (t>0)$$

이고, 생존함수는 $S(t)=P(T\geq t)=e^{-\frac{t}{60}}\ \ (t>0)$ 이다.

따라서 $P(T\geq 80)=S(80)=e^{-\frac{80}{60}}=e^{-\frac{4}{3}}$이다.

(2) $P(T<50|T\geq 40)=\frac{P(40\leq T<50)}{P(T\geq 40)}=\frac{P(T<50)-P(T<40)}{P(T\geq 40)}=\frac{(1-e^{-5/6})-(1-e^{-2/3})}{e^{-2/3}}=1-e^{-1/6}$

3. 어떤 전자제품이 4개의 부품으로 구성되어 있고, 이 부품 중 적어도 2개만 기능을 유지하면 이 제품을 작동 될 수 있다고 한다. 각 부품이 독립적으로 기능을 유지할 확률을 0.7이라 한다. 이 제품이 작동할 확률을 구하시오.

(sol)

기능을 유지하고 있는 부품의 개수가 X이면, $X\sim B(4,\frac{7}{10})$이므로, X의 확률질량함수는

$$f_X(x)=\binom{4}{x}(\frac{7}{10})^x(\frac{3}{10})^{4-x}$$

이다. 따라서

$$P(X\geq 2)=1-P(X<2)=1-\{f_X(0)+f_X(0)\}=1-\{\binom{4}{0}(\frac{7}{10})^0(\frac{3}{10})^4+\binom{4}{1}(\frac{7}{10})^1(\frac{3}{10})^3\}$$
$$=1-0.0837=0.9163$$

4. 컴퓨터 시뮬레이션을 통하여 0에서 9까지의 숫자를 무작위로 선정하며, 각 숫자가 선정될 가능성은 동일하다고 한다.

(1) 처음으로 숫자 0이 나올 때까지 시뮬레이션을 반복한 횟수에 관한 확률분포를 구하여라.

(2) (1)의 확률분포에 대한 평균과 분산을 구하여라.

(3) 시뮬레이션을 10번 실시해서야 비로소 4번째 0이 나올 확률을 구하여라.

(4) 4번째 0을 얻기 위하여 시뮬레이션을 반복한 평균 횟수를 구하여라.

(sol)

(1) 숫자 0이 나올 확률이 $p=\frac{1}{10}$이고, 숫자 0이 나올 때까지 시뮬레이션을 반복한 횟수를 X라 하면,

$X \sim G(\frac{1}{10})$이므로 X의 확률질량함수는

$$f_X(x)=(\frac{1}{10})^x(\frac{9}{10})^{x-1},\ x=1,2,\cdots$$

이다.

(2) $E[X]=\frac{1}{p}=10,\ Var[X]=\frac{q}{p^2}=90$

(3) 4번째 0이 나올 때까지 반복 시행한 횟수를 X라 하면, $X \sim NB(4,\frac{1}{10})$이므로 X의 확률밀도함수는

$$f_X(x)=\binom{x-1}{3}(\frac{1}{10})^4(\frac{9}{10})^{x-4},\ x=4,5,\cdots$$

이다. 따라서 $f_X(10)=\binom{9}{3}(\frac{1}{10})^4(\frac{9}{10})^6=0.0044641$

(4) $E[X]=\frac{r}{p}=\frac{4}{1/10}=10$

5. (1) 지하수 오염 실태를 조사하기 위하여 30곳의 구멍을 뚫어 수질을 조사하였다. 그 결과 19곳은 오염이 매우 심각하였고, 6곳을 약간 오염되었다고 보고하였다. 그러나 채취한 지하수 병들이 섞여 있어 어느 지역이 깨끗한 지하수를 갖고 있는지 모르는 상황에서 5곳을 선정하였을 때, 오염정도에 따른 확률질량함수를 구하여라.

(2) 선정된 5곳 중에서 매우 심각하게 오염된 곳이 3곳, 약간 오염된 지역이 1곳일 확률을 구하여라.

(3) 5곳 중에서 적어도 4곳에서 심각하게 오염되었을 확률을 구하여라.

(sol)

(1) 선정된 5개 병 중에서 오염이 매우 심각한 곳, 약간 오염된 곳, 그리고 청정한 곳의 수를 각각 X_1, X_2, X_3 라 하면, 결합확률질량함수는

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3)=\frac{\binom{19}{x_1}\binom{6}{x_2}\binom{5}{x_3}}{\binom{30}{5}},\ x_1+x_2+x_3=5,\ x_1,x_2,x_3=0,1,2,3,4,5$$

이다.

(2) 오염이 매우 심각한 지역이 3, 약간 오염된 지역이 1이면, 청정한 지역은 1이므로,

$$f_{X_1,X_2,X_3}(3,1,1)=\frac{\binom{19}{3}\binom{6}{1}\binom{5}{1}}{\binom{30}{5}}=\frac{1615}{7917}=0.20399$$

이다.

(3) 선정된 5개 병 중에서 오염이 매우 심각한 곳의 수를 X라 하면, $X \sim H(30;19,5)$이므로 X의 확률질량함수는

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

이다. 따라서 $P(X \ge 4)=f_X(4)+f_X(5)=\frac{\binom{19}{4}\binom{11}{1}}{\binom{30}{5}}+\frac{\binom{19}{5}\binom{11}{0}}{\binom{30}{5}}=\frac{7016}{23751}+\frac{1938}{23751}=\frac{1292}{3393}=0.3808$

6. $X \sim B(200, 0.55)$일 때, 확률 $P[X \leq 80]$을 다음 방법들을 통하여 구하여라.

(1) 연속성 수정을 사용하지 않은 정규 근삿값

(2) 연속성 수정을 사용한 정규 근삿값

(sol)

$np = 200 \times 0.55 = 110 \geq 5$, $nq = 200 \times 0.45 = 99 \geq 5$이므로 $X_1 \sim N(110, 49.5)$ 이다.

(1) $P(X \leq 80) \approx P(X_1 \leq 80) = P(Z \leq \frac{80-110}{\sqrt{49.5}}) = P(Z \leq 4.264) \approx 0.0000100$

(2) $P(X \leq 80) \approx P(X_1 \leq 80.5) = P(Z \leq \frac{80.5-110}{\sqrt{49.5}}) = P(Z \leq 4.193) \approx 0.0000138$

※ $P(X \leq 80) = 0.0000144$

7. 확률밀도함수가 $f_X(x) = \frac{1}{2}\lambda e^{-\lambda|x-\theta|}$, $-\infty < x < \infty$인 확률분포를 모수 λ와 θ인 라플라스 분포(Laplace distribution)라 한다.

(1) X의 누적분포함수 $F_X(x)$를 구하고, $f_X(x)$와 $F_X(x)$의 그래프를 그려라.

(2) $\lambda = 3$, $\theta = 1$일 때, $P(X \leq 0)$와 $P(0 \leq X \leq 2)$을 구하여라.

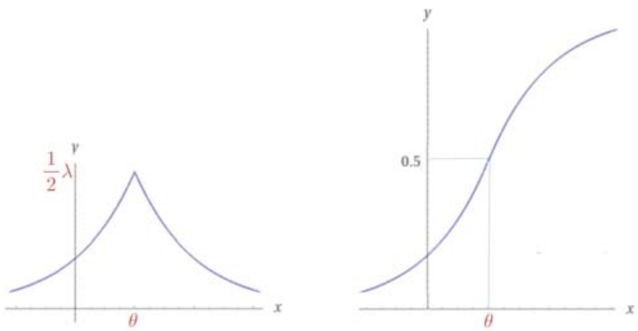
(sol)

(1) $-\infty < x < \theta$이면, $F_X(x) = \int_0^x \frac{1}{2}\lambda e^{-\lambda|t-\theta|}dt = \int_0^x \frac{1}{2}\lambda e^{-\lambda(\theta-t)}dt = \left[\frac{1}{2}e^{-\lambda(\theta-t)}\right]_{-\infty}^{x} = \frac{1}{2}e^{-\lambda(\theta-x)}$

$\theta < x < -\infty$이면, $F_X(x) = \int_0^x \frac{1}{2}\lambda e^{-\lambda|t-\theta|}dt = \int_0^\theta \frac{1}{2}\lambda e^{-\lambda(\theta-t)}dt + \int_0^x \frac{1}{2}\lambda e^{-\lambda(t-\theta)}dt$

$$= \left[\frac{1}{2}e^{-\lambda(\theta-t)}\right]_{-\infty}^{\theta} + \left[-\frac{1}{2}e^{-\lambda(t-\theta)}\right]_{\theta}^{x}$$

$$= \frac{1}{2} - \left(\frac{1}{2}e^{-\lambda(t-\theta)} - \frac{1}{2}\right) = 1 - \frac{1}{2}e^{-\lambda(t-\theta)}$$



(2) $0 < \theta = 1$이므로, $P(X \leq 0) = \frac{1}{2}e^{-3(1-0)} = \frac{1}{2}e^{-3}$

$0 < \theta = 1$이므로, $P(0 \leq X \leq 2) = P(X \leq 2) - P(X \leq 0) = 1 - \frac{1}{2}e^{-3(2-1)} - \frac{1}{2}e^{-3(1-0)} = 1 - e^{-3}$

8. 버스가 정류장에 9시와 9시 30분 사이에 도착하는데 그 시점은 균일분포를 따른다고 한다.

(1) 9시 정각에 이 정류장에 도착한 사람이 10분 이내에 버스를 탈 수 있을 확률을 구하여라.

(2) 만약 9시 15분까지 아직 버스가 도착하지 않았을 때 앞으로 적어도 10분은 더 기다려야 할 확률을 구하여라.

(sol)

버스를 기다리는 시간을 X라 하면, X의 확률밀도함수는 $f_X(x) = \frac{1}{30}$ $(0 < x < 30)$ 이다.

(1) $P(X < 10) = \frac{10-0}{30} = \frac{1}{3}$

(2) $P(X \geq 15 + 10 | X \geq 15) = \frac{P(X \geq 25)}{P(X \geq 15)} = \frac{1 - P(X \leq 25)}{1 - P(X \leq 15)} = \frac{1 - \frac{25}{30}}{1 - \frac{15}{30}} = \frac{1}{3}$

9. 전화교환대에 1분당 평균 2번의 비율로 신호가 들어오고 있으며, 교환대에 도착한 신호의 횟수는 푸아송과정에 따른다고 한다.

(1) 교환대에 들어오는 두 신호 사이의 평균 시간을 구하여라.

(2) 2분과 3분 사이에 신호가 없을 확률을 구하여라.

(3) 교환원이 교환대에 앉아서 3분 이상 기다려야 첫 번째 신호가 들어올 확률을 구하여라.

(4) 처음 2분 동안 신호가 없으나 2분과 4분 사이에 4건의 신호가 있을 확률을 구하여라.

(5) 처음 신호가 15초 이내에 들어오고, 그 이후 두 번째 신호가 들어오기까지 3분 이상 걸릴 확률을 구하여라.

(sol)

(1) 1분당 평균 2번 비율로 신호가 들어오는 교환대에서 들어온 두 신호사이의 대기시간을 T라 하면,

$T \sim \mathrm{Exp}(\lambda=2)$이므로, $E[T]=\frac{1}{\lambda}=\frac{1}{2}$(분) 이다.

(2) 단위 시간당 발생비율이 $\lambda=2$이고, $(0,t)$에서 들어온 신호 횟수를 $X_{(0,t)}$라 하면,

$X_{(0,t)} \sim Poisson\,P(2t)$이므로 $X_{(0,t)}$의 확률질량함수는

$$f_X(x)=\frac{(2t)^x}{x!}e^{-2t},\ x=0,1,2,\cdots$$

이다. 따라서 $P(X_{(2,3)}=0)=P(X_{(0,1)}=0)=\frac{2^0}{0!}e^{-2}=e^{-2}$ 이다.

(3) $T \sim \mathrm{Exp}(\lambda=2)$이므로, T의 확률밀도함수와 누적분포함수는

$$f_T(t)=2e^{-2x}\ \ (t>0)$$

$$F_T(t)=\int_0^t 2e^{-2x}\,dx=1-e^{-2t}\ \ (t>0)$$

이고, 생존함수는 $S(t)=P(T\geq t)=e^{-2t}(t>0)$이다. 따라서 $P(T\geq 3)=e^{-6}$이다.

(4) $P(X_{(0,2)}=0, X_{(2,4)}=4)=P(X_{(0,2)}=0)P(X_{(2,4)}=4)$

$$=P(X_{(0,2)}=0)P(X_{(0,2)}=4)=\frac{4^0}{0!}e^{-4}\frac{4^4}{4!}e^{-4}=\frac{32}{3}e^{-8}$$

(5) 처음 신호가 들어올 때까지 걸린 시간을 T_1, 처음 신호이후 다음 신호가 들어올 때까지 걸린 시간을 T_2라 하면,

T_1과 T_2는 독립인 지수분포 $\mathrm{Exp}(2)$를 이룬다.

$$P(T_1<\frac{1}{4}, T_2>3)=P(T_1<\frac{1}{4})P(T_2>3)=F_T(\frac{1}{4})S(3)=(1-e^{-\frac{1}{2}})(e^{-6})$$

10. $X \sim NB(r,p)$일 때, $m_X(t)=\left(\frac{pe^t}{1-(1-p)e^t}\right)^r$임을 보여라.

(sol)

$$\begin{aligned} m_X(t) &= \sum_{x=r}^{\infty} e^{tx}\binom{x-1}{r-1}p^r q^{x-r} \\ &= \sum_{x=r}^{\infty}\binom{x-1}{r-1}(qe^t)^x\left(\frac{p}{q}\right)^r \\ &= \sum_{x=r}^{\infty}\binom{x-1}{r-1}(qe^t)^{x-r}(qe^t)^r\left(\frac{p}{q}\right)^r \\ &= \sum_{x=r}^{\infty}\binom{x-1}{r-1}(1-qe^t)^r(qe^t)^{x-r}\frac{(qe^t)^r}{(1-qe^t)^r}\left(\frac{p}{q}\right)^r \\ &= \frac{(pe^t)^r}{(1-qe^t)^r} \end{aligned}$$