7주2강 복습 및 문제풀이(6장)

6장 복습

균등분포(uniform distribution)

1. **균등분포(uniform distribution)**는 연속형 분포에서 가장 단순한 분포형태로 특정구간 내의 값들이 가능성이 균등한 분포를 말한다.

- 연속형 확률변수 X가 실수구간 $[a, b]$에서 나타날 가능성이 균등할 때, X는 균등분포를 따른다고 하며 $\boldsymbol{X \sim U(a, b)}$로 표현한다. X의 확률밀도함수는 다음과 같다.

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le X \le b \\ 0, & \text{다른 곳에서} \end{cases}$$

2. 균등분포의 평균과 분산

- 확률변수 X가 $X \sim U(a, b)$라고 할 때, X의 평균과 분산은 다음과 같다.

① 평균 : $E(X) = \frac{b+a}{2}$

② 분산 : $Var(X) = \frac{1}{12}(b-a)^2$

3. 균등분포의 확률계산

- 구간 $\alpha \leq X \leq \beta$에서의 확률은 다음과 같다.

$$P_r(\alpha \leq X \leq \beta) = \frac{\beta - \alpha}{b - a}, \qquad a \leq \alpha, \qquad \beta \leq b$$

정규분포

- **정규분포**란 가능한 값이 $(-\infty, \infty)$사이의 모든 실수값이고 분포의 형태가 종모양(bell shape)인 분포를 말한다. 정규분포(normal distribution)는 통계이론에 있어서 가장 중요한 확률분포로 **대부분의 통계분석은 수집된 자료가 정규분포를 따른다고 전제**한다.
- 확률변수 X가 평균 μ와 분산 σ^2을 갖는 정규분포를 따른다고 할 때, X 확률밀도함수는 다음과 같다.

$$X \sim N(\mu, \sigma^2) \qquad f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \qquad -\infty < x < \infty$$

- 정규분포는 분포의 중심이 가장 높으며 이 **중심을 축으로 좌우대칭인 형태**를 갖고 있다. 정규분포의 특성은 중심축의 위치와 분포가 중심축을 중심으로 하여 흩어진 정도이며, 각각 정규분포의 **모수인 평균(μ)과 분산(σ^2)**으로 인해 나타나고 분포의 모양이 결정되어진다.

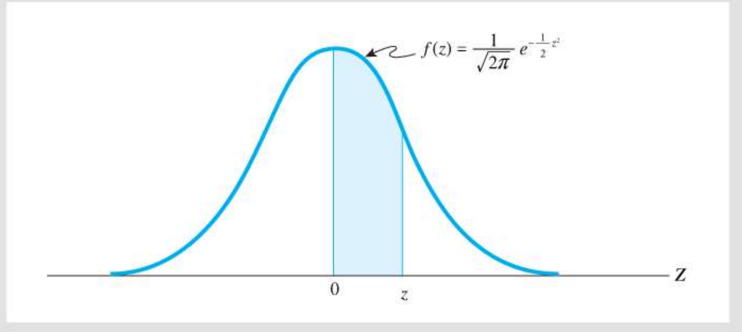
표준정규분포

- 확률변수 Z가 평균이 0이고 분산이 1인 정규분포를 따를 때 Z는 **표준정규분포**를 따른다고 하며, $Z \sim N(0,1)$으로 표현한다. Z의 확률밀도함수는 다음과 같다.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \qquad -\infty < z < \infty$$

- 표준정규분포의 형태는 다음과 같으며, 중심 0에서부터 양의 값 Z까지의 확률은 색칠한 부분의 넓이와 같다. 수식을 통해 확률값을 구하는 것이 쉽지 않으므로 확률 $P(0 \le Z \le z)$은 주로 표준정규분포표를 이용하여 확률값을 구한다.

그림 8-5 표준정규분포



정규분포의 확률계산

표준화

- 정규분포를 따르는 확률변수 $X \sim N(\mu, \sigma^2)$에서 $Z = \frac{X-\mu}{\sigma}$로 정의하면, $Z \sim N(0, 1)$이다.

- 위의 표준화 공식을 이용하여 정규분포를 따르는 확률변수를 표준정규분포로 표준화시켜 특정구간의 확률을 구할 수 있다.

지수분포

- 지수분포란 어떤 사건이 포아송분포에 의하여 발생될 때, 지정된 시점으로부터 이 사건이 처음 발생될 때까지 걸린 시간을 측정하는 확률분포이다.
- 확률변수 T가 다음과 같은 확률밀도함수를 가질 때 T는 모수 λ를 갖는 지수분포를 따른다고 하며 다음과 같이 표현한다.

$$f(t) = \lambda e^{-\lambda t}, \qquad t > 0$$

지수분포의 평균과 분산

$$E(T) = \frac{1}{\lambda}, \qquad Var(X) = \frac{1}{\lambda^2}$$

이항분포의 근사확률

① 이항분포의 정규근사

- 확률변수 X가 $X \sim B(n,p)$인 이항분포를 따를 때, X의 평균과 분산은 각각 다음과 같음을 알고 있다.

$$E(X) = np, \qquad Var(X) = npq$$

- 이러한 이항분포에서 n이 크고 p가 $\frac{1}{2}$에 가까운 경우에 X 의 분포는 평균 np 와 분산 npq를 갖는 정규분포와 근사적으로 동일하다고 할 수 있으며, X에 대한 확률은 $X \sim N(np, npq)$를 이용하여 근사적으로 구할 수 있다.
- 이항분포는 이산형 확률변수이고 정규분포는 연속형 확률변수 이므로 확률계산에서 약간의 조정을 하는 것이 필요하다. 정규분포를 이용하여 이항분포의 근사확률을 구하는 경우에 있어서 0.5를 조정하는 것을 이산형 확률분포의 확률을 연속형 확률분포를 이용하여 구하는 데 따르는 수정이라고 하여 **연속수정(continuity correction)**이라고 한다.

이항분포의 근사확률

② 이항분포의 포아송근사

- 이항분포 $X \sim B(n,p)$에서 n이 크고 p가 작은 경우에 있어서의 확률은 포아송분포를 이용하여 근사적으로 구할 수 있다.
- 포아송분포에서는 평균 μ가 모수이고, 이항분포 $X \sim B(n,p)$ 에서는 평균이 np이므로 $X \sim B(n,p)$ 는 $X \sim \text{Poisson}\ (np)$를 이용하여 근사확률을 구한다.

6장 연습문제 풀이

연습문제1. 확률변수 X가 -1과 1 사이에서 균등분포 $U(-1, 1)$을 따를 때 다음 질문에 답하여라.

(a) X의 확률밀도함수를 구하고, 이를 그림으로 표현하여라.

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

(b) $\{-0.5 \le X \le 0.5\}$의 확률은 얼마인가?

풀이) $P_r(-0.5 \le X \le 0.5) = \int_{-0.5}^{0.5} \frac{1}{2}\, dx = \frac{1}{2}$

(c) $\{0.5 \le X \le 1.0\}$의 확률은 얼마인가?

풀이) $P_r(0.5 \le X \le 1.0) = \int_{0.5}^{1.0} \frac{1}{2}\, dx = \frac{1}{4}$

(d) $\{X \le 0\}$의 확률은 얼마인가?

풀이) $P_r(X \le 0) = \int_{-1}^{0} \frac{1}{2}\, dx = \frac{1}{2}$

연습문제9(b). 확률변수 X가 $\mu = 40, \sigma = 9$인 정규분포 $N(40, 81)$을 따를 때 다음 확률조건을 만족하는 a를 구하여라.

(b) $P_r(X \le a) = 0.025$.

풀이)

$$P_r(X \le a) = 0.025 \quad \Longrightarrow \quad P_r\left(Z \le \frac{a-40}{9}\right) = 0.025$$

$$\frac{a-40}{9} = -1.96$$

$$a = 22.36$$

연습문제11. 한 음식점의 일별 매상액은 평균이 53만원이고 표준편차가 12만원인 정규분포에 근사한다고 한다.

힌트) $X \sim N(53, 12^2)$

(a) 어느 날 이 음식점의 매상액이 70만원 이상일 확률은 얼마인가?

풀이) $P_r(X \geq 70) = P_r\left(Z \geq \frac{70-53}{12}\right) = P_r(Z \geq 1.42) = 0.5 - 0.4222 = 0.0778$

(b) 이 음식점은 매일매일 최소한 30만원 이상의 매상이 있어야 이윤이 남는다고 할 때, 어느 특정한 날에 이윤이 남지 않을 확률은 얼마인가?

풀이) $P_r(X \leq 30) = P_r\left(Z \leq \frac{30-53}{12}\right) = P_r(Z \leq -1.92) = 0.5 - 0.4726 = 0.0274$

연습문제17. 한 회사원이 집에서 회사까지 출근하는데 걸리는 시간은 평균이 한 시간이고 표준편차가 0.3시간인 정규분포에 근사한 분포를 갖는다고 할 때,

힌트) $X \sim N(1, 0.3^2)$

(a) 이 사람이 회사에 출근하는데 걸린 시간이 한 시간 반 이상일 확률은 얼마인가?

풀이) $P_r(X \geq 1.5) = P_r\left(Z \geq \frac{1.5-1}{0.3}\right) = P_r(Z \geq 1.67) = 0.5 - 0.4525 = 0.0475$

(b) 회사의 업무가 오전 9시에 시작된다고 할 때, 이 사람이 집에서 7시 40분에 출발했다면 이 사람이 회사에 지각할 확률은 얼마인가?

풀이) $P_r\left(X \geq 1 + \frac{1}{3}\right) = P_r\left(Z \geq \frac{4/3-1}{0.3}\right) = P_r(Z \geq 1.11) = 0.5 - 0.3665 = 0.1335$

연습문제19. 확률변수 X가 $\lambda = 0.5$인 지수분포 Exponential(0.5)를 따를 때,

(a) X의 평균과 분산을 구하여라.

풀이) $E(X) = \frac{1}{0.5} = 2, \quad Var(X) = \frac{1}{0.5^2} = 4$

(b) X가 1 이상일 확률은 얼마인가?

풀이) $P_r(X \geq 1) = e^{-0.5}$

(b) X가 2 이하일 확률은 얼마인가?

풀이) $P_r(X \leq 2) = 1 - e^{-1}$

연습문제20. 한 낚시꾼이 낚시를 하는데 고기가 평균 한 시간에 두 마리가 잡힌다고 한다.

힌트) 이 문제는 $\lambda = 2$인 지수분포 문제이다.

(a) 한 번 고기를 잡은 후에 다음 고기를 잡을 때까지 걸리는 평균시간은 얼마인가?

풀이) 0.5시간

(b) 한 번 고기를 잡은 후에 두 시간 동안 전혀 고기를 잡지 못할 확률은 얼마인가?

풀이) $P_r(X \geq 2) = e^{-4}$

(b) 한 번 고기를 잡은 후에 30분 만에 다시 고기를 잡을 확률은 얼마인가?

풀이) $P_r(X \leq 0.5) = 1 - e^{-1}$

연습문제26. 한 자동차 외판원은 판매상담한 사람 중에서 30%의 사람들이 실제로 자동차를 구입한다고 한다. 어느 한달 동안에 100명과 판매상담을 했다고 할 때 최소한 20명 이상의 고객이 자동차를 구매할 확률은 얼마인가?

풀이) $X \sim B(100, 0.3)$

$E(X) = 100 \times 0.3 = 30, \;\; Var(X) = 100 \times 0.3 \times 0.7 = 21$

$\text{표준편차} = \sqrt{21} \approx 4.58$

$P_r(X \geq 20) = P_r\left(Z \geq \frac{19.5-30}{4.58}\right) = P_r(Z \geq -2.29) = 0.5 + 0.489 = 0.989$

연습문제31. 어느 사격선수가 표적을 맞추지 못할 확률이 0.48%라고 한다. 어느 사격대회에서 이 선수가 500회의 사격을 실시했다고 할 때 표적을 맞추지 못한 경우가 2회 이내일 확률은 얼마인가? 이항분포에 대한 포아송근사를 이용해 구하여라.

풀이) $X \sim B(500, 0.0048)$

$E(X) = 500 \times 0.0048 = 2.4$

$P_r(X \leq 2)$는 <부록 V>의 [표 2]에서 $\mu = 2.4$에 의해 $P_r(X \leq 2) = 0.570$이다.

끝~~ ♥♥