8-1 이론적 확률분포(연속형) 1

제주권역
대학 e-러닝 지원센터

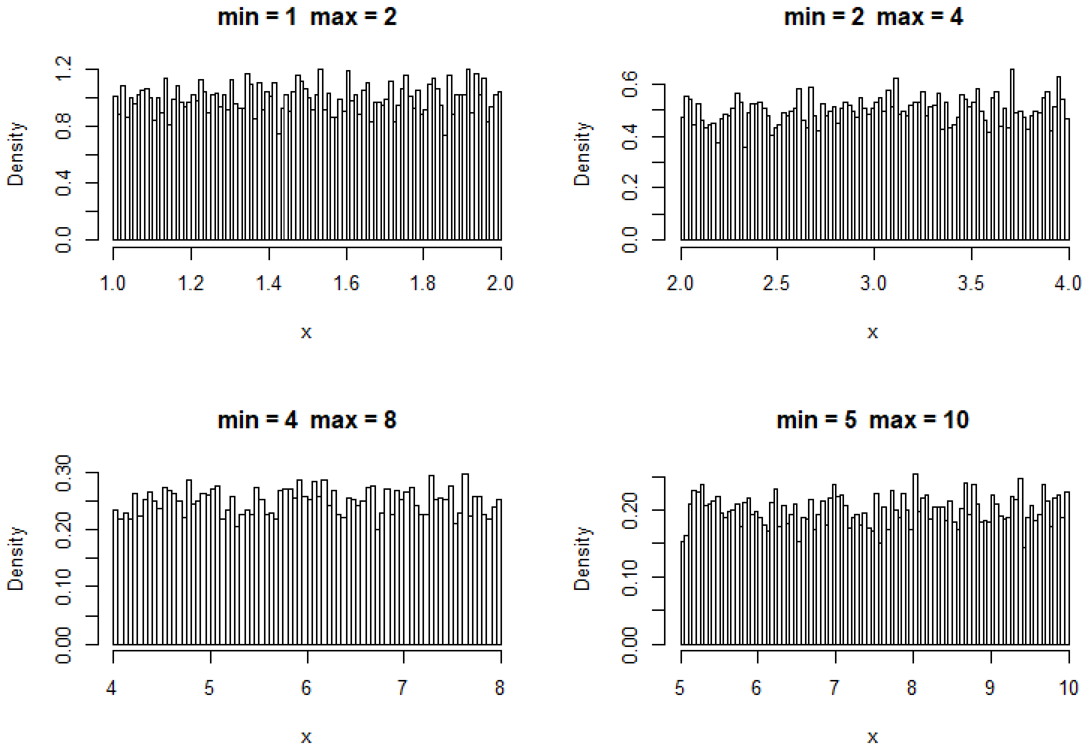
1. 균등분포(uniform distribution)

- 연속형 분포에서 가장 단순한 분포형태로 특정구간 내의 값들이 나타날 가능성이 균등한 분포를 말한다.

- 연속형 확률변수 X가 실수구간 [a, b] 에서 나타날 가능성이 균등할 때, X는 균등분포를 따른다고 하며 X~U(a, b)로 표현한다.

X 의 확률밀도함수는 다음과 같다.

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


min = 1 max = 2
Density
0.0
0.4
0.8
1.2
1.0
1.2
1.4
1.6
1.8
2.0
x
min = 2 max = 4
Density
0.0
0.2
0.4
0.6
2.0
2.5
3.0
3.5
4.0
x
min = 4 max = 8
Density
0.00
0.10
0.20
0.30
4
5
6
7
8
x
min = 5 max = 10
Density
0.00
0.10
0.20
5
6
7
8
9
10
x

2. 균등분포의 평균과 분산

· 확률변수 X가 X~U(a, b) 라고 할 때, X의 평균과 분산은 다음과 같다.

① 평균 : $E(X) = \dfrac{b+a}{2}$

② 분산 : $Var(X) = \dfrac{1}{12}(b-a)^2$

```
> (munif1<-mean(unif1))
[1] 1.500681

> (munif2<-mean(unif2))
[1] 2.993355

> (munif3<-mean(unif3))
[1] 5.999272

> (munif4<-mean(unif4))
[1] 7.474975

> (vunif1<-var(unif1))
[1] 0.08240007

> (vunif2<-var(unif2))
[1] 0.3336517

> (vunif3<-var(unif3))
[1] 1.320054

> (vunif4<-var(unif4))
[1] 2.077339
```

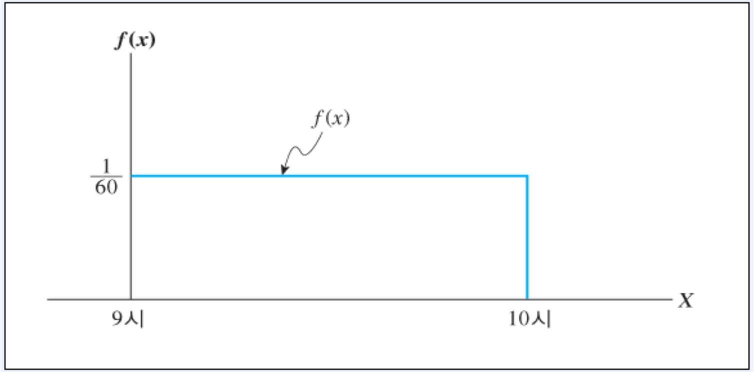
3. 균등분포의 확률계산

구간 $\alpha \le X \le \beta$ 에서의 확률은 다음과 같다.

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

예 A교수가 학교에 출근하는 시간이 오전 9시에서 10시 사이이며 그 시간 안에서 특정시간에 출근할 가능성이 동일하다고 할 때, 확률변수 X를 A교수의 출근시간이라고 정의하면 X는 9시에서 10시 사이에 균등한 확률분포를 가진다.

균등분포인 경우, 확률밀도함수 f(x) 가 1이어야 하므로 확률분포형태가 직사각형으로 나타난다. 가로 길이의 단위를 분으로 하면 가로 길이가 60이 되므로 세로 길이인 f(x)도 1/60 이 되며, 이 경우의 확률밀도함수와 분포의 그림은 다음과 같다.

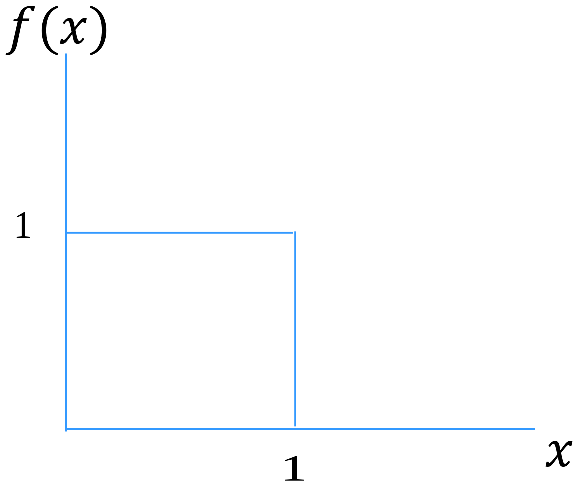


· 한편, 가로 길이의 단위를 시간으로 하면 가로 길이가 1 이 되므로 세로 길이인 f(x)도 1 이 되며, 이 경우의 확률밀도함수와 분포의 그림은 다음과 같다.

① 확률밀도함수

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

② 균등분포의 그림

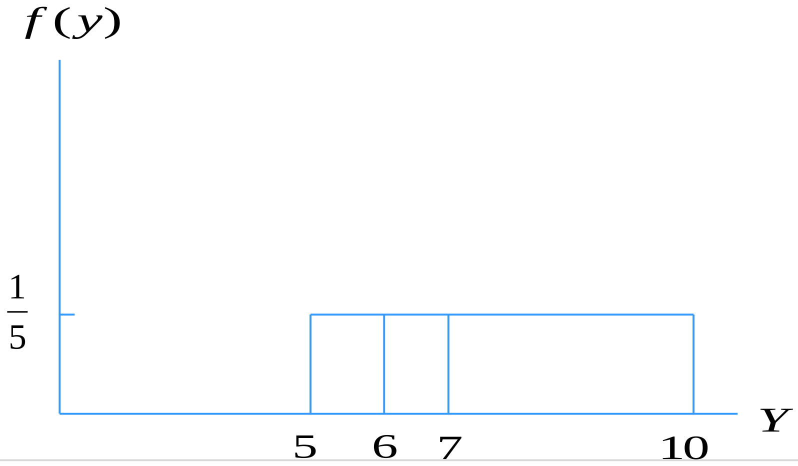


예

확률변수 Y가 구간 [5,10]에서 균등한 분포를 가질 때 $6 \le Y \le 7$의 확률을 계산하라

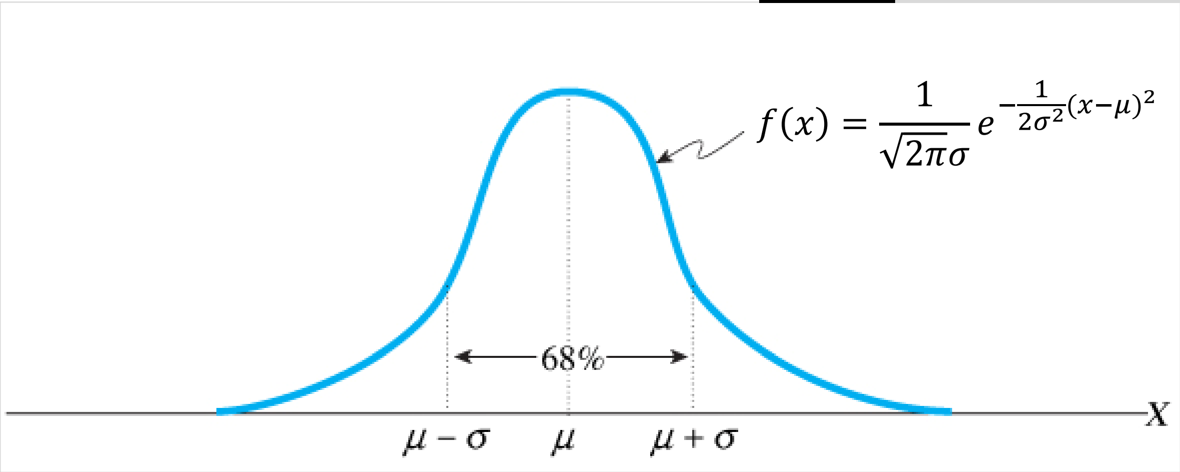
풀이

Y의 확률밀도함수가 $f(y) = \dfrac{1}{10-5} = \dfrac{1}{5}, 5 \le Y \le 10$이므로 $P_r(6 \le Y \le 7) = \dfrac{1}{5} \times 1 = \dfrac{1}{5}$ 이다.



8-1-2. 정규분포

그림 정규분포



정규분포

정규분포는 가능한 값이 $-\infty < x < \infty$ 사이의 모든 실수 값이며, 분포의 형태가 그림과 같이 종모양인 분포를 말한다.

정규분포

· 정규분포란 분포의 형태가 종을 엎어 놓은 모양인 분포를 말하며, 분포의 형태는 평균 μ 와 분산 σ^2 에 의해 결정된다. 확률변수 X가 평균 μ 와 분산 σ^2을 갖는 정규분포를 따른다면

$X \sim N(\mu, \sigma^2)$ 이라고 표현한다.

· X의 확률밀도함수는 다음과 같다.

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

여기에서 e는 자연대수의 밑수로 $e = 2.71828\cdots$ 이다.

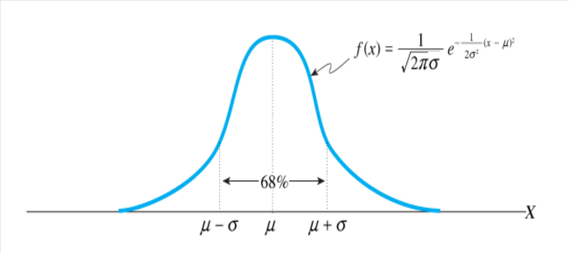
- 확률변수 X가 정규분포를 따를 때 X의 일차함수로 표시되는 새로운 변수 Y=aX+b역시 정규분포를 한다.

$$E(Y) = E(aX + b) = a\mu + b$$

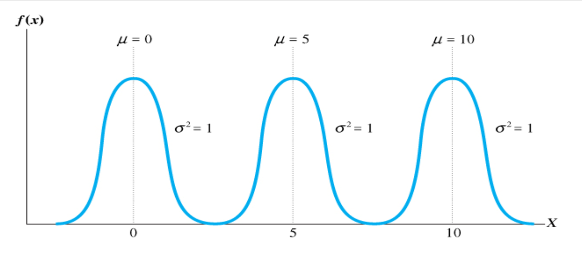
$$Var(Y) = Var(aX + b) = a^2\sigma^2$$

즉, $X \sim N(\mu, \sigma^2)$일 때, $Y \sim N(a\mu + b, a^2\sigma^2)$

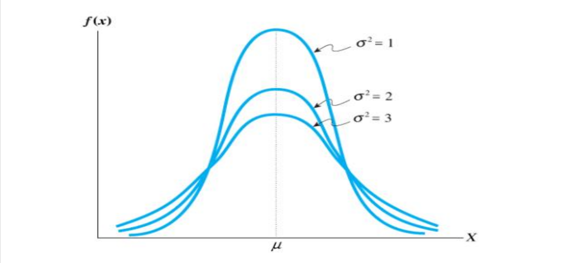
① 정규분포의 일반적 형태



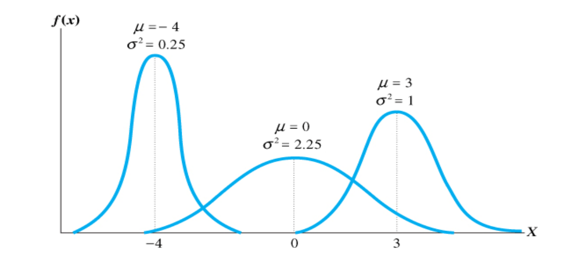
② 평균이 다르고 분산이 같은 정규분포



③ 평균은 같으나 분산이 다른 정규분포



④ 평균과 분산이 각각 다른 경우의 정규분포



분산의 값이 클수록 분포가 평균을 중심으로 넓게 흩어지며, 분산의 값이 작을수록 분포의 형태가 평균에 집중되어 있음을 알 수 있다.