

3. 확률변수와 확률분포함수 (2)

$$\frac{t(wt)}{[]}$$

이산형
w

Ber. Bin.

정규
C70

Normal

$\chi^2, t, F \dots$
Unif ..

주요 이산형 확률분포 함수

- 베르누이 분포(Bernoulli distribution)

- 베르누이 시행

매 시행마다,

- 1) '성공' 또는 '실패'의 오직 두 가지 가능한 결과만 가짐
- 2) '성공'의 확률이 p 로 일정함.

의 조건을 만족하는 실험.

- 베르누이 확률변수와 확률질량함수

- ♦ $S = \{success, failure\}$

- ♦ $X = \begin{cases} 1 & , \text{ if success} \\ 0 & , \text{ if failure} \end{cases}$

- ♦ $p(x) = P[X = x] = \begin{cases} p & , \text{ if } x = 1 \\ 1 - p & , \text{ if } x = 0 \end{cases}$

- ♦ 이 경우, $X \sim Ber(p)$ 라고 함.

주요 이산형 확률분포함수

- 베르누이 분포의 평균과 분산

$X \sim \text{Ber}(p)$ 인 경우,

- ♦ $E[X] = p$
- ♦ $V[X] = p(1 - p)$
- ♦ $S[X] = \sqrt{p(1 - p)}$

주요 이산형 확률분포함수

- 이항분포(binomial distribution)

- 이항확률변수와 확률질량함수

- ♦ 베르누이 시행을 독립적으로 n 번 반복 시행할 때,

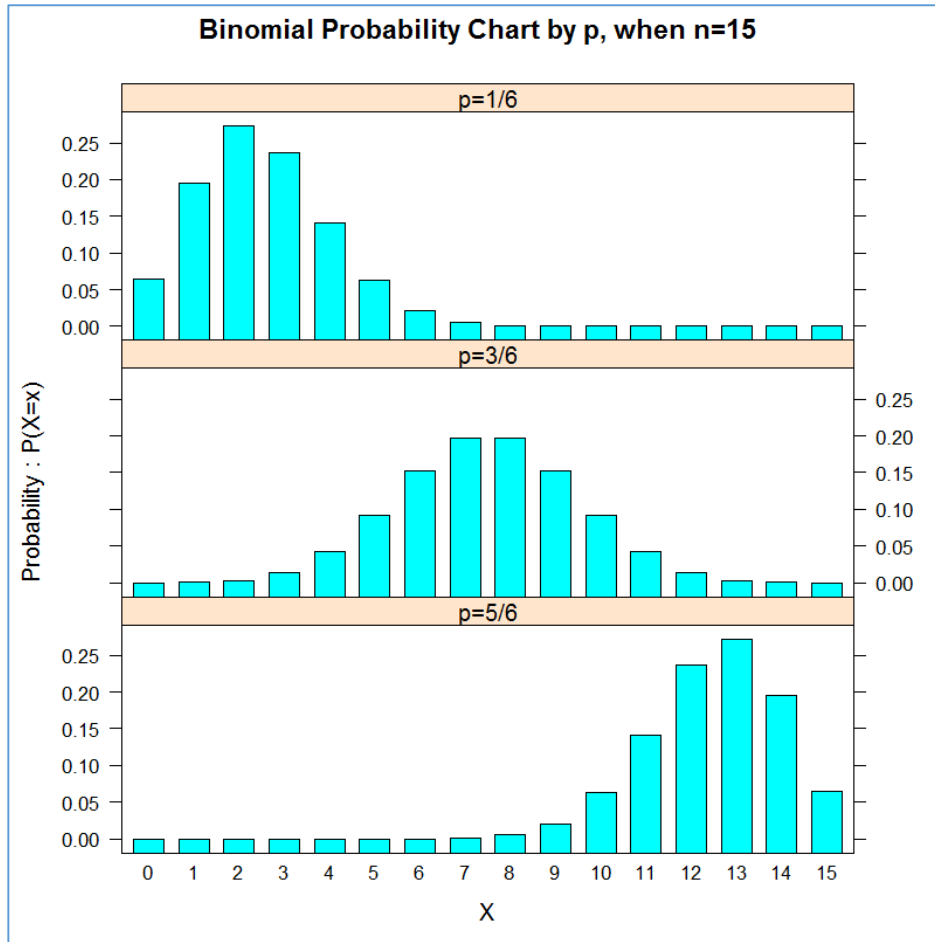
$X = n$ 번 베르누이 시행에서 '성공'의 횟수

로 정의하면,

- ♦ $p(x) = P[X = x] = \binom{n}{x} p^x (1 - p)^{n-x}, x = 0, 1, 2, \dots, n$
 - ♦ 이 경우, $X \sim \text{Bin}(n, p)$ 라고 함.

주요 이산형 확률분포함수

■ 이항분포의 그래프 개형



주요 이산형 확률분포함수

- 이항분포의 평균과 분산

$X \sim \text{Bin}(n, p)$ 인 경우,

- ♦ $E[X] = np$
- ♦ $V[X] = np(1 - p)$
- ♦ $S[X] = \sqrt{np(1 - p)}$

주요 이산형 확률분포함수

■ 예제

주사위 6개를 던졌을 때, 각 주사위의 눈이 1이 나오는 경우가 어느정도 인지에 관심이 있다고 하자.

- ♦ X 를 6개 중 1이 나온 경우의 수로 정의했을 때, X 가 가질 수 있는 값들은 무엇인가?

- ♦ X 의 확률분포함수는 무엇인가?

- ♦ 6개의 주사위 중 1이 나오는 경우는 몇 개로 기대되는가?

- ♦ 6개의 주사위의 눈이 모두 1이 나올 확률은 무엇인가?

```
> dbinom(6, size=6, prob=1/6)
[1] 2.143347e-05
```

주요 이산형 확률분포함수

■ 예제

- ◆ 어느 기업의 연간 수익이 증가할 가능성은 63%로 알려져 있으며, 이 회사는 연간 수익이 증가하거나 떨어지거나 두 가지 상황만 존재한다고 가정하자. 또한 매년 연간 수익이 증가하는지 떨어지는지 여부는 독립적이라고 하자. 이 경우에,
 - 5년 동안 이 회사의 성장 횟수에 대한 확률분포는 무엇인가?
 - 5년 동안 이 회사가 적어도 두 번 성장할 확률을 구하여라.
 - 5년 동안 이 회사의 성장 횟수에 대한 기대값과 표준편차를 구하여라.

주요 이산형 확률분포함수

```
> 1-pbinom(1, size=5, prob=0.63)
[1] 0.9340295
>
> Ex <- 5*0.63
> Ex
[1] 3.15
>
> Sx <- sqrt( 5*0.63*(1-0.63) )
> Sx
[1] 1.079583
```

주요 이산형 확률분포함수

- 예제

- ♦ 어느 금융서비스를 제공하는 회사의 고객 중 30%는 서비스에 만족하지 못한다고 한다. 이 주장이 사실인 경우, 이 회사 고객 중 15명을 무작위로 선택하였을 때, 그 중 서비스에 만족하지 못한다고 고객의 수가 3명 이하일 확률은 얼마인가?

```
> pbinom(3, size=15, prob=0.3)
[1] 0.2968679
```

주요 이산형 확률분포함수

- 예제

- ♦ 어느 회사에서 생산 한 디스크는 서로 독립적으로 0.1의 확률로 결함이 있는 것으로 알려져 있다. 이 회사는 디스크를 4개씩 묶어서 패키지로 판매하고 있으며 한 패키지에 디스크가 하나라도 결함이 있을 경우 환불을 보증한다고 한다. 반품되는 패키지의 비율은 얼마인가?

```
> 1 - dbinom(0, size=4, prob=0.1)
[1] 0.3439
```

주요 연속형 확률분포함수

- 정규분포 (normal distribution)

- 정규확률변수와 확률밀도함수

- ♦ 확률변수 X 가 평균이 μ , 분산이 σ^2 이고, 다음의 확률밀도함수를 가질 때, X 는 정규분포를 따른다고 하며, 이 경우, $X \sim N(\mu, \sigma^2)$ 라고 함.

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

- 정규분포의 평균과 분산

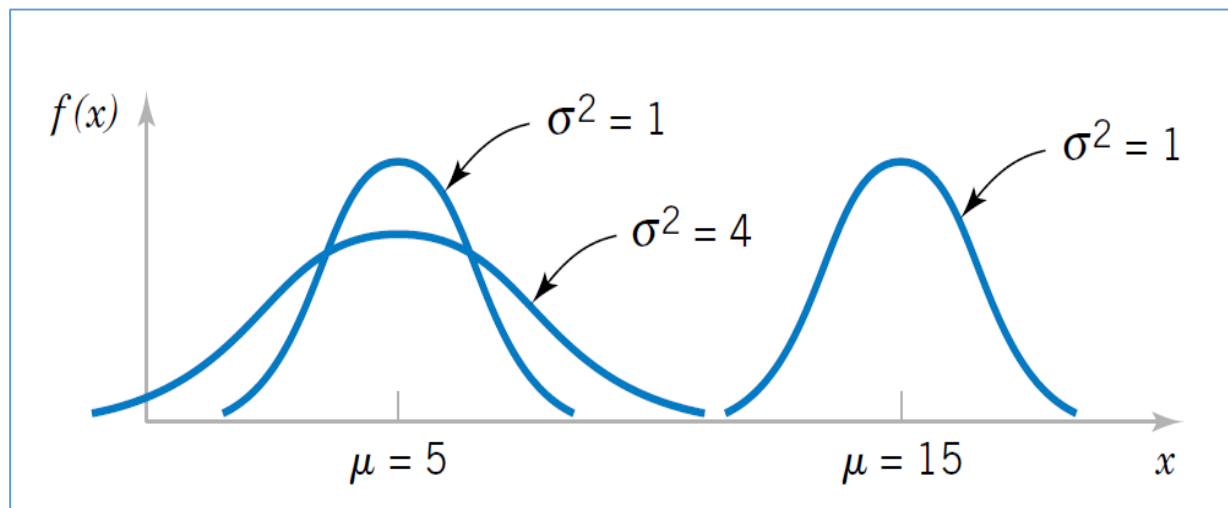
$X \sim N(\mu, \sigma^2)$ 인 경우,

- ♦ $E[X] = \mu$
 - ♦ $V[X] = \sigma^2$

주요 연속형 확률분포함수

■ 정규분포 확률밀도함수의 개형

- ♦ μ 를 중심으로 대칭이며, μ 에서 가장 큰 값을 가지는 종 모양의 분포임.
→ μ 보다 클 확률과 작을 확률 모두 1/2임.
- ♦ σ^2 이 크면 분포의 산포가 커지고, σ^2 이 작으면 분포의 산포가 작아짐.



주요 연속형 확률분포함수

- 정규분포의 성질(선형성1)

- ♦ $X \sim N(\mu, \sigma^2)$ 일 때, $Y(= aX + b) \sim N(a\mu + b, a^2\sigma^2)$ 임. (단, a, b 는 상수)

- 예. $X \sim N(0, 1^2)$ 일 때, $Y(= 2X + 1)$ 의 분포는?

주요 연속형 확률분포함수

▪ 정규분포의 성질(선형성2)

- ♦ X_1, \dots, X_n 이 서로 독립인 정규 확률변수로 $X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, \dots, n$ 이라고 할 때,
 $\sum_{i=1}^n X_i \sim N(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$.

- 예. $X_1 \sim N(20, 3^2), X_2 \sim N(40, 4^2)$ 이고, X_1 과 X_2 는 서로 독립일 때, $(X_1 + X_2)$ 의 분포는?

주요 연속형 확률분포함수

- 표준정규분포 (standard normal distribution)

- 표준정규 확률변수와 확률밀도함수

- ♦ $X \sim N(\mu, \sigma^2)$ 일 때, 정규분포의 선형성 성질에 의해

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

이 되므로, Z 를 평균(μ)이 0 분산(σ^2)이 1인 정규분포를 따르는 확률변수로 볼 수 있음.

- ♦ 이 확률변수의 분포를 표준정규분포라 정의.
 - ♦ 표준정규분포의 확률밀도함수는 다음과 같음.

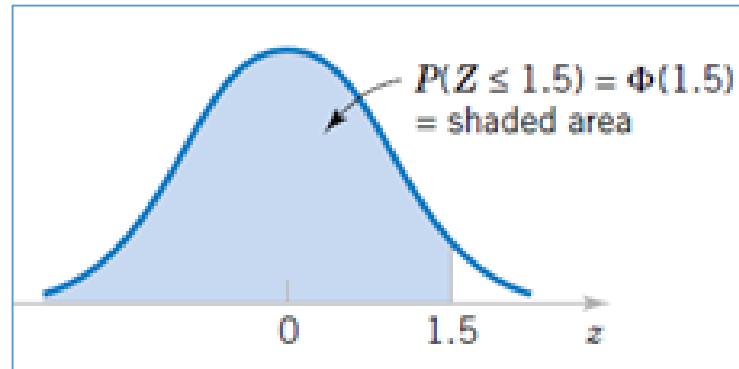
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, -\infty < z < \infty$$

주요 연속형 확률분포함수

- 표준정규 누적분포함수

- ♦ 일반적으로 표준정규분포의 누적분포함수를 $\Phi(\cdot)$ 로 표기함.

$$\Phi(a) = P(Z \leq a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-z^2/2} dz, -\infty < a < \infty$$

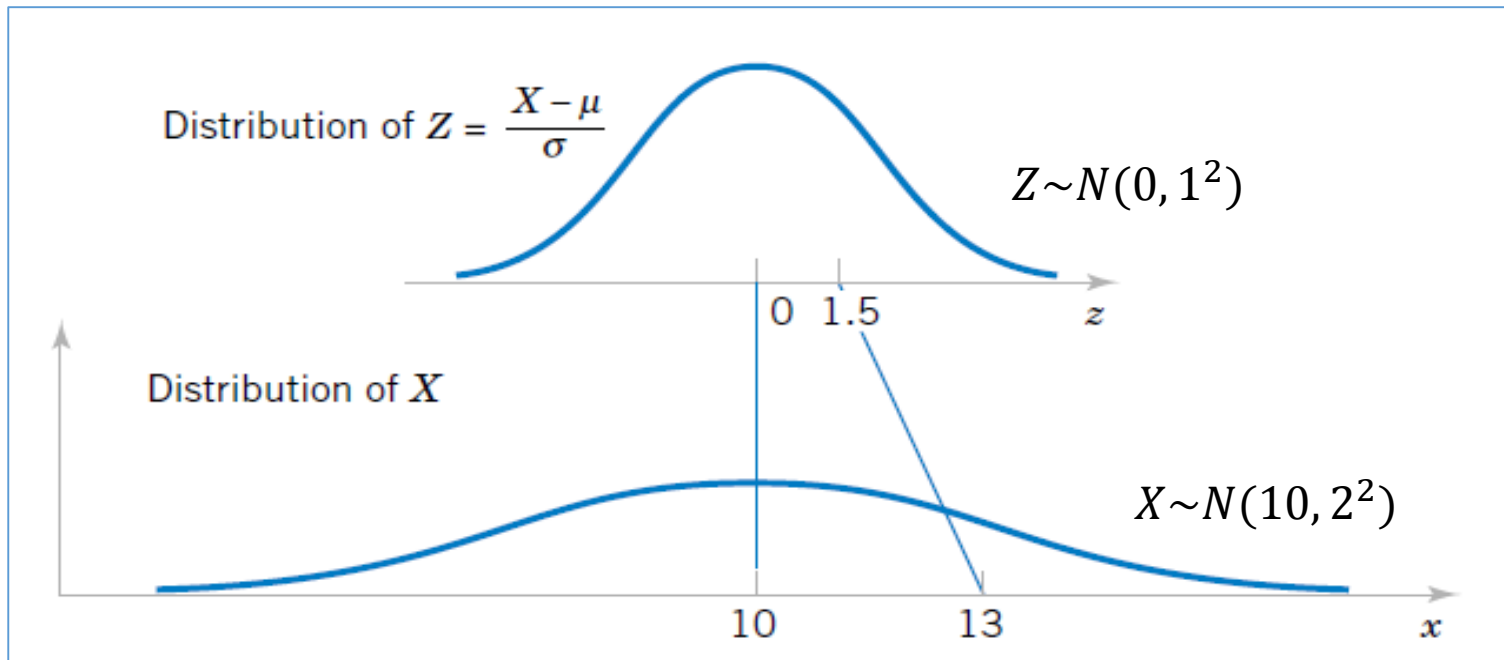


- $P(Z < 0) = P(Z > 0) = \Phi(0) = 0.5$
- $P(Z > a) = 1 - \Phi(a)$
- $P(a < Z < b) = \Phi(b) - \Phi(a)$

주요 연속형 확률분포함수

- 임의의 정규 확률변수 $X \sim N(\mu, \sigma^2)$ 가 어느 구간에 속할 확률은 모두 표준정규분포의 누적 확률분포함수 $\Phi(\cdot)$ 를 이용하여 표현할 수 있음.

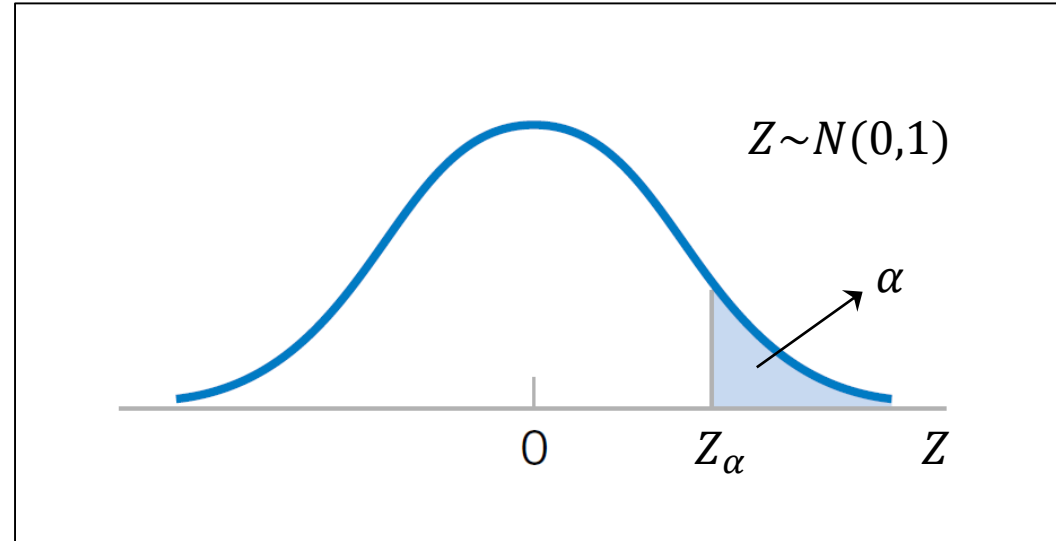
- $P(a < X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$



주요 연속형 확률분포함수

- 표준정규 확률변수의 $(1 - \alpha)$ 분위수 : Z_α

$Z \sim N(0,1)$ 일 때, $P[Z > c] = \alpha$ 를 만족하는 Z 의 $(1 - \alpha)$ 분위수 c 를 Z_α 으로 표기함.



주요 연속형 확률분포함수

■ 예제

◆ $Z \sim \text{Normal}(0, 1)$ 일 때, 다음을 구하여라.

- $P[Z < -1.645] + P[Z > 1.645]$
- $P[-1.645 < Z < 1.96]$
- $P[|Z| < a] = 0.9$ 를 만족하는 a .

◆ 다음의 Z_α 값을 구하여라.

- $Z_{0.025}$
- $Z_{0.95}$

주요 연속형 확률분포함수

```
> pnorm(-1.645)*2  
[1] 0.09996981  
>  
> pnorm(1.96) - pnorm(-1.645)  
[1] 0.9250172  
>  
> qnorm(0.95)  
[1] 1.644854  
>  
> qnorm(0.975)  
[1] 1.959964  
>  
> qnorm(0.05)  
[1] -1.644854
```

주요 연속형 확률분포함수

- 예제

- ♦ 어느 포트폴리오의 연 수익률에 대한 분포는 평균이 45이고, 표준편차가 39인 정규분포를 따른다. 1년 후에 포트폴리오의 가치가 39와 43 사이에 있을 확률을 구하여라. (단위: 천만 달러)

```
> pnorm( 43, mean=45, sd=39)-pnorm( 39, mean=45, sd=39)  
[1] 0.04068486
```

주요 연속형 확률분포함수

■ 예제

- ◆ 어느 투자의 수익률은 평균이 10%이고 표준 편차가 4% 인 정규 분포를 따른다고 할 때, 이 투자로 인해 돈을 잃을 확률은 얼마인가?
- ◆ 만일 이 투자 수익률의 표준편차가 7%였다면, 돈을 잃을 확률은 얼마인가?

```
> pnorm( 0, mean=10, sd=4)
```

```
[1] 0.006209665
```

```
> pnorm( 0, mean=10, sd=7)
```

```
[1] 0.07656373
```

주요 연속형 확률분포함수

- 예제

- ◆ 한 대학에서 새로운 Executive MBA 프로그램을 승인하였다. 이 프로그램의 협의회는 입학 요건 중 하나로 지원자가 GMAT 점수의 상위 1% 이내여야 한다고 결정하였다. GMAT 점수가 평균 490, 표준 편차 61인 정규분포를 가진다고 할 때, 이 프로그램 입학을 위한 최소 GMAT 점수는 얼마인가?

```
> qnorm(0.99, mean=490, sd=61)  
[1] 631.9072
```

주요 연속형 확률분포함수

■ 예제

- ◆ 어느 은행의 카드 소지자가 매월 지불하는 이자 금액을 분석한 결과 해당 금액은 평균 \$27, 표준 편차 \$7로 정규 분포를 따르는 것으로 나타났다. 이 은행 카드 소지자 가운데 지불하는 이자 금액이 하위 20%에 해당하는 사람이 있다면, 그 사람이 지불하는 이자 금액은 얼마인가?

```
> qnorm(0.20, mean=27, sd=7)  
[1] 21.10865
```