# 4강 통계계산(2)

한림대학교 금융정보통계학과 심송용교수

## 목 차



# 1 지수분포와 감마분포

# 지수분포와 감마분포

- 확률변수 $X$의 확률밀도함수가

$$f(x) = \lambda e^{-\lambda x}, \; x > 0$$

  인 분포를 **지수분포**. 기호로는 Exp($\lambda$) 로 표현.
- 지수분포의 기댓값과 분산은 각각 $1/\lambda$, $1/\lambda^2$.
- R-언어에서 지수분포 함수는 다음과 같음.

```
dexp(x, rate = 1)
pexp(q, rate = 1, lower.tail = TRUE)
qexp(p, rate = 1, lower.tail = TRUE)
rexp(n, rate = 1)
```

# 지수분포와 감마분포

- **감마분포**는 확률변수 $X$ 의 확률밀도함수가

$$f(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1}\, e^{-\beta x},\quad x > 0$$

  인 분포이며 기호로는 X ~ $\Gamma(\alpha,\beta)$ 로 표현.
- $\alpha = 1$ 이면 모수가 $\beta$인 지수분포임.
- 따라서 지수분포는 감마분포의 특수한 경우임.
- 감마분포의 기댓값과 분산은 각각 $\alpha/\beta$ 및 $\alpha/\beta^2$ 임.
- R–언어에서 감마분포에 대한 함수들은 다음과 같음.

```
dgamma(x, shape, rate = 1, scale = 1/rate)
pgamma(q, shape, rate = 1, scale = 1/rate, lower.tail = TRUE)
qgamma(p, shape, rate = 1, scale = 1/rate, lower.tail = TRUE)
rgamma(n, shape, rate = 1, scale = 1/rate)
```

# 지수분포와 감마분포

- x, q : 확률밀도함수값을 얻을 x 벡터, 누적확률을 얻을 q 벡터.
- p : 분위수를 얻을 확률값의 벡터.
- n : 발생할 난수의 개수.
- rate : 지수분포 및 감마분포의 모수 $\lambda$의 값.
- scale : 감마분포의 경우 rate $\frac{1}{\lambda}$ 의 값을 scale parameter라고 부르기도 하며 이 값을 설정. rate 또는 scale 중의 하나만 설정.
- shape : 감마분포의 모수 α의 값.
- lower.tail : 논리값을 설정하며 TRUE이면 확률은 $\Pr[X \leq x]$ 의 값으로, 그렇지 않으면, $\Pr[X > x]$을 계산함.

감마분포와 지수분포, 카이제곱 분포와의 관계는 다음과 같다.
① X ~ $\Gamma(1,\lambda)$ 이면 는 지수분포 $\mathrm{Exp}(\lambda)$와 같음.
② X ~ $\Gamma(\nu/2,1/2)$ 이면 X는 카이제곱 분포 $\chi^2(\nu)$와 같음.

# 지수분포와 감마분포

**예제 2.11.** x가 1/2, 1, 3/2에서 다음 확률밀도함수를 구해보자.
① $\alpha = 1,\ \beta = 2$ 일 때 지수분포(β)와 감마분포(α, β)의 확률밀도함수를 구해보자.
② $\lambda = 4$ 인 지수분포를 따르는 난수 1,000개를 생성하여 평균과 분산을 구해보자.
③ 자유도가 4인 카이제곱분포(4)와 감마분포(2, 1/2)의 확률밀도함수를 구해보자.

```
> x <- c(1/2, 1, 3/2)
> # 1
> dexp(x, 2)
[1] 0.73575888 0.27067057 0.09957414
> dgamma(x, 1, 2)
[1] 0.73575888 0.27067057 0.09957414
> #2
> x <- rexp(1000, 4)
```

# 지수분포와 감마분포

```
> mean(x)
[1] 0.2479532
> var(x)
[1] 0.06524859
> # 3
> dgamma(x, 2, rate = 1/2)
[1] 0.0973501 0.1516327 0.1771375
> dchisq(x, 4)
[1] 0.0973501 0.1516327 0.1771375
```

- 지수분포(2)는 감마분포(1,2)와 같음.
- 지수분포의 기댓값과 분산은 각각 0.2479532, 0.06524859 으로 $\lambda = 4$일 때 각각 $1/\lambda = 0.25$, $1/\lambda^2 = 0.0625$와 가까운 값.
- 또한 카이제곱(4)와 감마분포(2,1/2)는 같은 분포임을 확인.

## ❷ t-분포, F-분포

# t–분포와 F–분포

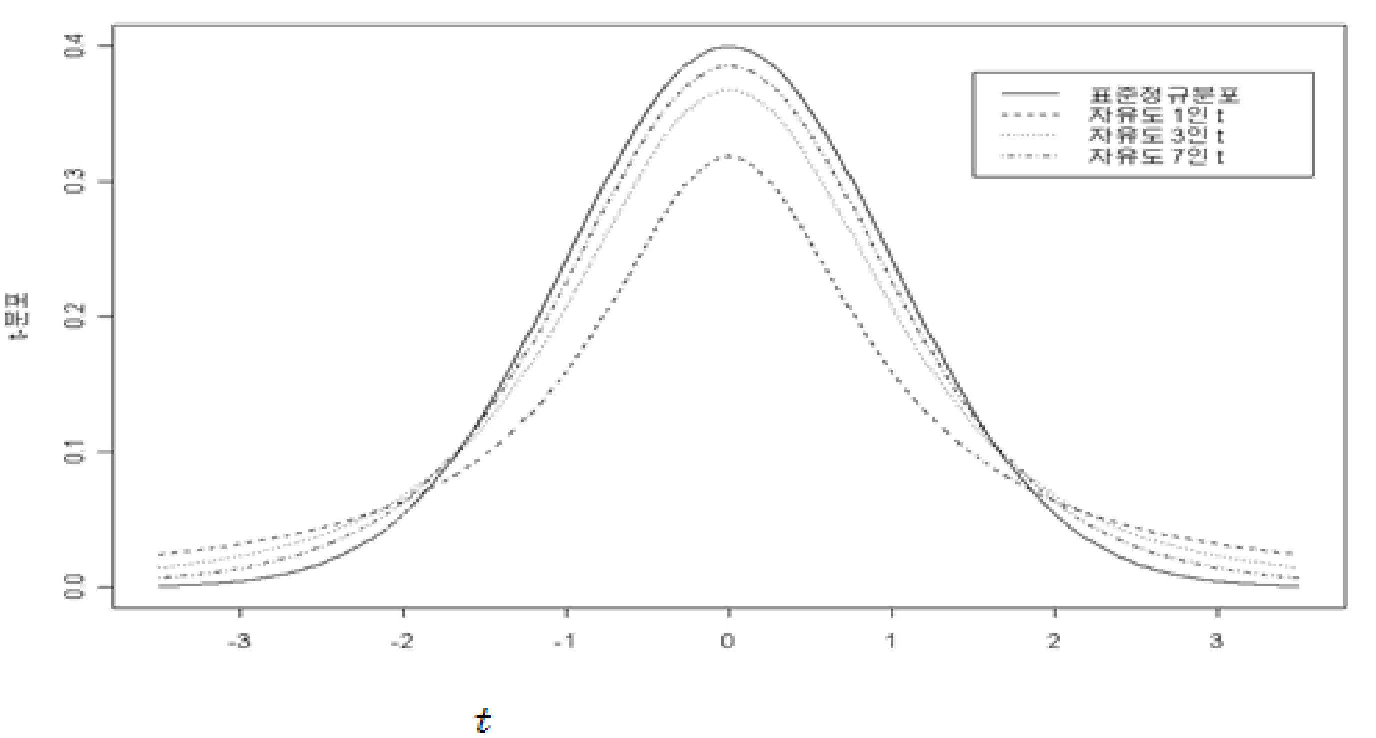
- $Z$가 표준정규분포를 따르는 확률변수, $\chi_\nu^2$를 자유도가 $\upsilon$인 카이제곱분포를 갖는 확률변수이고 이 둘이 독립이라면 확률변수

$$T = \frac{Z}{\sqrt{\chi_\nu^2 / \nu}}$$

  의 분포를 자유도가 $\upsilon$인 t–분포라고 하며, $t(\nu)$ 또는 $t_\nu$ 로 표시.
- 분자의 확률변수 $Z$가 정규분포이나 기댓값이 0이 아닌 경우의 분포를 비중심 t–분포
- t–분포는 비중심모수가 0인 경우이며, 이를 비중심 t–분포와 구분하기 위해 중심 t–분포라고 부르기도 함.
- t–분포는 0을 중심으로 좌우대칭.
- 표준정규분포보다 꼬리가 두터우며 자유도가 커질수록 표준정규분포로 수렴.

# t-분포와 F-분포

## t-분포의 모양



# t-분포와 F-분포

```
dt(x, df, ncp)
pt(q, df, ncp, lower.tail = TRUE)
qt(p, df, ncp, lower.tail = TRUE)
rt(n, df, ncp)
```

- x, q : 확률밀도함수값을 얻을 x 벡터, 누적확률을 얻을 q 벡터.
- p : 분위수를 얻을 확률값의 벡터.
- n : 발생할 난수의 개수.
- df : 사용할 t 분포의 자유도. 무한대(Inf)값을 사용할 수 있음(표준정규분포).
- ncp : 비중심모수(noncentrality parameter)의 값.
- lower.tail : 논리값을 설정하며 TRUE이면 확률은 Pr[X ≤ x] 의 값으로, 그렇지 않으면, Pr[X 〉 x]을 계산.

# t-분포와 F-분포

**예제 2.12.** t-분포에서 x = 0.025에서 자유도 10인 경우와 자유도가 무한대인 값을 구해보고, 자유도가 무한대로 수렴하는 경우 표준정규분포로 수렴함을 살펴보자.

```
> qt(.025, 10) #  자유도가 10인 경우의 값
[1] -2.228139
> qt(.025, Inf)
[1] -1.959964
> qnorm(.025)
[1] -1.959964
```

- qt(.025, 10) 는 x = 0.025, 자유도가 10인 t-분포 값.
- 자유도가 무한대인 t-분포 qt(.025, Inf) 의 값 -1.959964는 표준정규분포 qnorm(.025)와 같은 값임.

# t–분포와 F–분포

- $X_1$, $X_2$가 독립이고 각각 자유도가 $\nu_1$, $\nu_2$인 카이제곱분포를 따르는 확률변수라고 할 때

$$F_{\nu_1,\nu_2} = \frac{X_1/\nu_1}{X_2/\nu_2}$$

의 분포를 **자유도가 $\nu_1$, $\nu_2$인 F–분포**라고 하고, F($\nu_1$,$\nu_2$) 또는 $F_{\nu_1,\nu_2}$로 표시.

- F–분포는 정의에 의해서

$$F_{\nu_1,\nu_2} = \frac{1}{F_{\nu_2,\nu_1}}$$

임을 알 수 있음.

- 분자 또는 분모의 카이제곱분포가 비중심 카이제곱분포일 때의 F–분포를 비중심 F–분포라 함. 대부분 분자가 비중심, 분모가 중심카이제곱분포일 때 비중심 F–분포라 함.

# t-분포와 F-분포

- $F_{\nu_1,\nu_2;\,\alpha} = \dfrac{1}{F_{\nu_2,\nu_1;1-\alpha}}$ 임을 아래와 같이 확인.

```
> qf(0.05, 10, 5)
[1] 0.3006764
> 1/qf(0.95, 5, 10)    # qf(0.05, 10, 5) 와 같은 값
[1] 0.3006764
```

- t-분포의 정의에 의해

$$T_v^2 = F_{1,v} = \frac{\chi_1^2/1}{\chi_v^2/v}$$

임을 확인.

# t-분포와 F-분포

```
> qf(.95, 1, 10)
[1] 4.964603
> qt(.975, 10)^2
[1] 4.964603
```

t-분포는 0.95가, F-분포에서는 0.975가 사용된 것은 t-분포의 제곱이 F-분포이기 때문.

# ③ 포아송, 베르누이, 이항분포

# 포아송(Poisson)분포

- X가 0과 자연수의 값을 가질 수 있는 확률변수이고

$$\Pr[X = x] = f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \; x = 0, 1, 2, \ldots$$

  이면 X의 분포를 포아송분포라고 하며 일정한 시간 동안 발생한 사건, 결함 등의 개수의 확률모형으로 많이 사용함.
- X의 기댓값과 분산은 모두 $\lambda$로 같음.
- 분포함수, 분위수, 난수발생 함수 등은 아래와 같음.

```
dpois(x, lambda)
ppois(q, lambda, lower.tail = TRUE)
qpois(p, lambda, lower.tail = TRUE)
rpois(n, lambda)
```

# 포아송(Poisson)분포

- x, q : 각각 확률밀도함수의 값을 얻을 x 벡터 또는 누적확률을 얻을 q 벡터.
- p : 분위수를 얻을 확률값의 벡터.
- n : 발생할 난수의 개수.
- lower.tail : 논리값을 설정하며 TRUE이면 확률은 $Pr[X \leq x]$ 의 값으로,
  그렇지 않으면, $Pr[X > x]$을 계산.
- lambda : 사용할 포아송분포의 기댓값.

**예제 2.14.** 어느 도시에서 빈 택시는 시간당 평균 10대꼴로 지나며 빈 택시의 댓수는 포아송 분포를 따른다고 한다. 한 시간 동안 한 대의 빈 택시도 만나지 못할 확률은 얼마인가?

풀이 : X를 빈 택시의 댓수라고 하면 $Pr[X = 0]$을 원함.

# 포아송(Poisson)분포

```
> dpois(0, 10)
[1] 4.539993e-05
```

이 경우 Pr[X ≤ 0]을 계산하여도 되므로 ppois(0,10)을 사용하여도 같은 결과를 얻음.

# 베르누이(Bernoulli) 시행과 이항분포

- **베르누이 시행**은 관찰 가능한 값이 두 가지 뿐인 이산형 분포 중의 하나로 이 두 관찰 가능한 값은 보통 '성공'과 '실패'로 표현.
- 이를 수치로 표현할 때는 성공은 1로 실패는 0으로 표현.
- 성공 확률이 $p$이고 실패 확률이 $(1-p)$인 베르누이 시행의 $(0 < p < 1)$ 확률분포는

$$\Pr[X = x] = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \end{cases}$$

로 표현.

- 이는

$$\Pr[X = x] = p^x(1-p)^{1-x}, \quad x = 0, 1.$$

로도 표현.

# 베르누이(Bernoulli) 시행과 이항분포

- **이항분포**는 독립인 $n$개의 베르누이 시행에서의 성공 확률이 $p$일 때 이 베르누이 시행들의 합의 분포.
- $B(n,p)$로 표시
- $X$가 이항분포일 때 확률질량함수는

$$f(x) = \Pr[X = x] = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0,1,\ldots,n$$

로 표현.
- 이항분포의 기댓값과 분산은 각각 $np$, $np(1-p)$.
- 분포함수, 난수, 분위수 함수 등은 다음과 같음.

# 베르누이(Bernoulli) 시행과 이항분포

```
dbinom(x, size, prob)
pbinom(q, size, prob, lower.tail = TRUE)
qbinom(p, size, prob, lower.tail = TRUE)
rbinom(n, size, prob)
```

- x, q : 확률밀도함수값을 얻을 x 벡터, 누적확률을 얻을 분위수 q 벡터.
- p : 분위수를 얻을 확률값의 벡터.
- n : 발생할 난수의 개수.
- lower.tail : 논리값을 설정하며 TRUE이면 확률은 $Pr[X \le x]$ 의 값으로,
  그렇지 않으면, $Pr[X > x]$을 계산.
- size : 이항분포 B(n, p)에서 n값을 설정함.
  size = 1이면 베르누이 시행.
- prob : 이항분포 B(n, p)에서 p값을 설정함.

# 베르누이(Bernoulli) 시행과 이항분포

**예제 2.15.** $X$가 이항분포 $B(10, 0.2)$ 이면 Pr[X = 2]는

```
> dbinom(2, 10, 0.2)
[1] 0.3019899
```

로 얻으며 Pr[X ≤ 2]는

```
> pbinom(2, 10, 0.2)
[1] 0.6777995
```

이며 이 값은 dbinom(0, 10, 0.2) + dbinom(1, 10, 0.2) + dbinom(2, 10, 0.2)와 같음.

# 베르누이(Bernoulli) 시행과 이항분포

Pr[X ≤ a]=0.5 가 되는 a는

```
> qbinom(0.5, 10, 0.2)
[1] 2
```

이다.

**예제 2.16.** 공평한 주사위를 던지면 1의 눈금이 나올 확률이 1/6이다. 주사위를 다섯 번 던질 때 1의 눈금이 나올 회수를 확률변수 X라고 하면 X의 분포는 $B(5,1/6)$이며 , 기댓값과 분산은 각각 $np = 5/6 = 0.8333$, $np(1-p) = 25/36 = 0.6944$이다.

주사위를 다섯 번 던지는 실험을 100회 반복할 때 1의 눈금이 나온 회수의 평균과 분산이 이 값들과 얼마나 가까운지  모의실험을 해보자.

# 베르누이(Bernoulli) 시행과 이항분포

```
binom.par <- function(nrep=100, n=5, p=1/6) { # binom.par.r
    x <- rbinom(nrep, n, p)      # B(n,p)에서 nrep 개의 난수 생성
    meanx <- mean(x)             # 이들 난수의 평균
    varx <- var(x)               # 이들 난수의 분산
    list(meanx = meanx, varx = varx)
} # end function

binom.par()
$meanx
[1] 0.75
$varx
[1] 0.7348485
```

# 베르누이(Bernoulli) 시행과 이항분포

X가 $B(n,p)$이면 $n$이 충분히 클 때 X의 분포는 정규분포 $N(np,npq)$에 근사하다는 것이 알려져 있으며 이를 이항분포의 정규근사라고 한다. 확률변수 X가 B(50,1/2)일 때 Pr[X ≤ x] 의 확률값을 이항분포에서와 정규근사에서 각각 구해보자. (normAbinom.r)

```
normAbinom <- function() { # normAbinom.r
  x <- seq(5, 50, by=5)                   # 5, 10, ... 50인 x에 대해
  p.binom <-  pbinom(x, 50, 0.5)      # 누적 이항분포 확률
  p.norm <- pnorm(x, 25, sqrt(12.5)) # 누적 정규분포확률
  # 평균 np=25, 분산 np(1-p)=12.5인 정규분포
  diff <- p.binom-p.norm                  # 누적 분포의 차이
  list(p.binom = round(p.binom,4), p.norm = round(p.norm,4), diff = round(diff,4))
 } # end function
```

# 베르누이(Bernoulli) 시행과 이항분포

```
normAbinom()
$p.binom
 [1] 0.0000 0.0000 0.0033 0.1013 0.5561 0.9405 0.9987 1.0000 1.0000 1.0000
$p.norm
 [1] 0.0000 0.0000 0.0023 0.0786 0.5000 0.9214 0.9977 1.0000 1.0000 1.0000
$diff
 [1] 0.0000 0.0000 0.0010 0.0227 0.0561 0.0192 0.0010 0.0000 0.0000 0.0000
```

위의 경우 이항분포와 정규분포의 차이가 없거나 미미하다는 것을 알 수 있음.
(소수점 넷째 자리까지)

# 4 기하분포와 음이항분포

# 기하분포와 음이항분포

- 성공확률이 p일 때 첫 번째 성공을 관찰할 때까지의 시행회수 X의 분포를 기하분포라고 함.
- 확률변수 X가 기하분포를 따른다면 Pr[X = x]는 처음의 (x − 1)개의 실패와 x번째의 성공을 관찰할 확률이므로

$$p(x) = \Pr[X = x] = (1 - p)^{x-1}p, \quad x = 1,2,3,\ldots$$

- 첫 번째 성공을 관찰할 때까지의 실패의 횟수의 분포를 기하분포라고 정의하기도 하는데 R–언어에서는 이 실패의 회수를 기하분포로 정의.
- 따라서 첫 번째 성공을 관찰할 때까지 실패의 횟수가 x번일 확률 Pr[X = x]는 x번의 실패와 마지막 한 번의 성공을 관찰할 확률이므로

$$p(x) = \Pr[X = x] = (1 - p)^{x}p, \quad x = 0,1,2,\ldots$$

임.

# 기하분포와 음이항분포

```
dgeom(x, prob)
pgeom(q, prob, lower.tail = TRUE)
qgeom(p, prob, lower.tail = TRUE)
rgeom(n, prob)
```

- x, q : 분포함수를 계산하거나 누적확률을 계산할 벡터.
- p : 분위수를 계산할 확률값의 벡터.
- n : 난수를 발생할 개수.
- prob : 각 시행에서의 성공확률.
- lower.tail : 논리값을 설정하며 TRUE이면 확률은 Pr[X ≤ x] 의 값으로,
  그렇지 않으면, Pr[X 〉 x]을 계산.

# 기하분포와 음이항분포

**예제 2.18.** 전체 자동차 중 빨간색을 가진 자동차는 5%라고 한다. 도로에서 자동차의 색을 관찰할 때 20번째에 처음으로 빨간색 자동차를 관찰할 확률은 얼마인가?

20번째 시행에서 처음으로 성공을 관찰할 확률이므로 실패의 회수는 19인 경우이다. 따라서

```
> dgeom(19,0.05)
[1] 0.01886768
```

이 값은 앞의 19대가 빨간색 이외의 자동차, 20번째가 빨간색일 확률 $0.95^{19} \times 0.05$ 와 같음을 알 수 있다.

# 기하분포와 음이항분포

- 음이항분포는 r번째 성공을 관찰할 때까지의 시행회수의 분포.
- r번째 성공이 x번째 시행에서 일어나기 위해서는 x번째는 성공이고 그 이전의 (x − 1)번의 시행에서 (r − 1)번의 성공이 있어야 함.
- 따라서 X를 성공확률이 p인 음이항분포를 따르는 확률변수라고 하면 x번째 시행에서 r번째 성공이 관찰될 확률 Pr[X = x]는

$$\begin{aligned}\Pr[X=x] &= \binom{x-1}{r-1} p^{r-1}(1-p)^{x-r} \cdot p \\ &= \binom{x-1}{r-1} p^{r}(1-p)^{x-r}, \quad x = r, r+1, r+2 \ldots \end{aligned}$$ (교재 오류)

이다.

# 기하분포와 음이항분포

- R–언어에서는 실패의 횟수를 기준으로 함.
- r개의 성공을 관찰할 때까지 실패의 횟수를 x라고 하면 전체 (x + r)개의 베르누이 시행 중 마지막은 r번째 성공이라야 함.
- 이전의 (x + r − 1) 개의 시행에서 (r − 1)번의 성공과 x번의 실패.
- 이 확률은

$$\Pr[X = x] = \binom{x + r - 1}{r - 1}(1 - p)^x p^{r-1} \times p$$
$$= \binom{x + r - 1}{r - 1}(1 - p)^x p^r, \quad x = 0,1,\ldots,n$$

으로 얻음.

# 기하분포와 음이항분포

```
dnbinom(x, size, prob, mu)
pnbinom(q, size, prob, mu, lower.tail = TRUE)
qnbinom(p, size, prob, mu, lower.tail = TRUE)
rnbinom(n, size, prob, mu)
```

- x, q : 분포함수를 계산하거나 누적확률을 계산할 벡터.
- p : 분위수를 계산할 확률값의 벡터.
- n : 난수를 발생할 개수.
- size : 양수값을 설정하며 분포함수에서 r에 해당하는 값이다.
  즉 r 번째 실패를 할 때까지의 분포에서 사용한 값 r을 설정.
- mu : 음이항분포의 기댓값은 mu= size*(1-prob)/prob 인데 이 mu를 설정.
  이 때는 size나 prob 중의 한 값만 설정함.
- prob : 각 시행에서의 성공확률.
- lower.tail : 논리값을 설정하며 TRUE이면 확률은 $Pr[X \le x]$ 의 값으로,
  그렇지 않으면, $Pr[X > x]$을 계산.

# 기하분포와 음이항분포

**예제 2.19.** 예제 2.18.과 같이 빨간색 자동차의 비율이 5%일 때 세 번째 빨간색 자동차를 관찰할 때까지 관찰한 자동차의 수가 20대 이하일 확률은 얼마인가?

**풀이 :** 세 번째 빨간색 자동차를 관찰할 때까지의 자동차수가 20대라는 것은 세 번째 빨간색 자동차를 관찰할 때까지 빨간색 이외의 자동차를 관찰한 자동차가 17대임을 의미하므로 빨간색 이외의 자동차가 17대 이하로 관찰될 확률을 뜻한다. 따라서

```
> pnbinom(17, 3, .05)
[1] 0.07548367
```

음이항분포는 기하분포의 확장이므로 예제 2.18.의 20번째 처음 성공은 첫 번 성공까지 19번의 실패로 바꿀 수 있고 이는

```
> dnbinom(19, 1, .05)
[1] 0.01886768
```

로 예제 2.18.의 결과와 같음을 알 수 있다.

# 5 초기하분포

# 초기하분포

- 주머니에 $m$개의 흰색 공과 $n$개의 검은 공이 있을 때 임의로 $k$개를 비복원 추출로 꺼내는 경우 흰색 공의 개수가 $x$개가 될 확률은

$$f(x) = \Pr[X = x] = \frac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}, \quad \max(0, \mathrm{n} - \mathrm{k}) \le x \le \min(m, k)$$

으로 얻음. 이 때 확률변수 $X$의 분포를 초기하분포라고 함.

- 분포함수, 난수발생, 분위수 계산 등의 함수는 다음과 같음.

```
dhyper(x, m, n, k)
phyper(q, m, n, k, lower.tail = TRUE)
qhyper(p, m, n, k, lower.tail = TRUE)
rhyper(nn, m, n, k)
```

# 초기하분포

- x, q : 분포함수를 계산하거나 누적확률을 계산할 벡터.
- p : 100p% 백분위를 계산할 확률값 p를 저장한 벡터의 이름.
- m : 주머니 속의 흰색공의 개수.
- n : 주머니 속의 검은색 공의 개수.
- k : 주머니에서 비복원 추출하는 공의 수.
- p : 분위수를 계산할 확률값의 벡터
- nn : 난수를 발생할 개수.
- lower.tail : 논리값을 설정하며 TRUE이면 확률은 $Pr[X \leq x]$ 의 값으로,
  그렇지 않으면, $Pr[X > x]$을 계산.

# 초기하분포

**예제 2.20.** 주머니 속에 빨간 공 5개 파란 공 6개가 있다. 이 주머니에서 3개의 공을 비복원으로 추출할 때 빨간 공의 수가 2일 확률은

```
> dhyper(2, 5, 6, 3)
[1] 0.3636364
```

이며 이 값은 choose(5,2)*choose(6,1)/choose(11,3) 로 얻어도 같은 결과.

빨간 공의 개수가 2 이상일 확률은 $1-\Pr[X \le 1]$ 로도 얻을 수 있으므로

```
> 1-phyper(1,5,6,3)
[1] 0.4242424
```

이며

# 초기하분포

이 값은 직접 빨간 공의 수가 2 또는 3일 확률을

```
choose(5,2)*choose(6,1)/choose(11,3) + choose(5,3)*choose(6,0)/choose(11,3)
```

또는

```
dhyper(2,5,6,3) + dhyper(3,5,6,3)
```

로도 얻을 수 있다.

**예제 2.21.** 초기하분포의 기댓값과 분산은 각각 $k\frac{m}{m+n}$ 및 $k\frac{m}{m+n}\frac{n}{m+n}\frac{m+n-k}{m+n-1}$이다. 따라서 예제 2.17.에서 $m = 5,\ n = 6,\ k = 3$이므로 기댓값과 분산은 각각 $15/11 = 1.364$ 및 $3\frac{5}{11}\frac{6}{11}\frac{8}{10} = 0.595$ 이다. 이 분포에서 1,000개의 난수를 생성하여 평균과 분산을 계산하여 이 값들과 얼마나 가까운지 확인해보자.

# 초기하분포

```
x <- rhyper(1000, 5,6,3)
> mean(x)
[1] 1.37
> var(x)
[1] 0.5816817
```

**로 이론적인 값과 비슷한 값을 얻음.**

## 다음시간 안내

# 통계추론(1)