

# 정규분포에서의 평균의 분포

📖 정규분포에서의 평균

📖 모의실험을 이용한 평균의 분포 확인

Loren ipsum dolor sit amet, ius an molestie  
facilisi erroribus, nislai nalerun delectus ei  
vis. Has ornatus conclusionemque id, an videri  
maiestatis sit. In etqui praesent sit. An vel  
agan porro comprehensan, ad ludus constituto  
nea, et ius utroque scaevola assuaverit.

Vis cu nodus nulla feugait, oratio facilisi ex  
usu, eili vitae sea te. Ea fabulas accusamus  
dissentias sea, facete tacinates definitiones  
et per. Nihil dicant mediocrem pro eu, no mei  
nostro sensibus platonem. Qui id sunno perpatas  
neglegantur. Vel ipsum novum copiosae ut. Quo  
et liber detracto probatus. Nam augue scriben-  
tur an. Sea oporteat percipitur incidere at.  
Qui viris nemore an.



# 정규분포에서의 평균



## 가정

+  $n$ 개의 자료  $x_1, x_2, \dots, x_n$ 가 모두 독립이고 기댓값이  $\mu$ 이고 분산이  $\sigma^2$ 인 정규분포



## 표본 평균 $\bar{X}$ 의분포

+ 기댓값은  $\mu$ , 분산은  $\sigma^2/n$ 인 정규분포

+ 기댓값은 변함이 없음

+ 분산은 작아짐

# 모의실험을 이용한 평균의 분포 확인



## 모의실험 1

### 평균에 대한 모의실험

1. 10개의 표준정규분포에서의 난수를 평균을 계산하는 것을
2. 100번 반복하여 100개의 평균을 얻어
3. 이들 평균의 평균과 분산을 계산해보자. 이 값이 0과 1/10에 근접하는지 확인

```
meannorm <- function(n=10, nrep=100) {
```

```
  mx <- rep(0, nrep)  * mx를 nrep개의 표본평균을 저장할 벡터로 사용하며 초기값으로 0을 설정함
```

```
  for (i in 1:nrep) { * for 반복문은 n개의 표준정규분포에서의 난수 생성, 이의 평균을 mx에 저장함
```

```
    x <- rnorm(n)
```

```
    mx[i] <- mean(x)
```

```
  } # end for
```

```
  list(mmx = mean(mx), sdmx = var(mx)) * list 명령으로 평균과 표준편차의 값을 출력으로 얻음
```

```
} # end function
```

# 모의실험을 이용한 평균의 분포 확인



## 모의실험 1

### 함수 호출 결과

```
> meannorm()
```

```
$mmx
```

```
[1] 0.03829618
```

```
$sdmx
```

```
[1] 0.3595127
```

\* 평균은  $\mu = 0$ 에 가까운 -0.038을 표준편차는  $1/\sqrt{n} = 1/\sqrt{100} = 0.316$ 에 근접한 값을 확인할 수 있음



# 모의실험을 이용한 평균의 분포 확인



## 모의실험 1

 평균을 1000번 얻어 이들의 평균과 표준편차를 반복한 결과

```
> meannorm(n=10, nrep=1000)
```

```
$mmx
```

```
[1] -0.001268805
```

```
$sdmx
```

```
[1] 0.3213833
```

**\* 비슷한 결과를 얻음**



# 모의실험을 이용한 평균의 분포 확인

## 모의실험 2



### 정규분포에서 표본평균의 분포

정규분포의 표본평균이 정규분포인지 확인하기 위해

1. 정규분포 난수를 사용하여 표본평균의 값을 모의실험으로 생성하고
2. 이들 표본평균에 대한 히스토그램을 그려보자.

```
meanhist <- function(n=10, nrep=100) {
```

```
  mx <- rep(0, nrep)
```

\* mx를 nrep개의 표본평균을 저장할 벡터로 사용하며 초기값으로 0을 설정함

```
  for (i in 1:nrep) {
```

\* for 반복문은 n개의 표준정규분포에서의 난수 생성, 이의 평균을 mx에 저장함

```
    x <- rnorm(n)
```

```
    mx[i] <- mean(x)
```

```
  } # end for
```

```
  hist(mx, prob = T)
```

\* nrep개의 평균을 얻은 다음 mx로 히스토그램을 그림

```
  lines(xx<- seq(-3,3, by=0.02), dnorm(xx, 0, sqrt(1/n)), col="red")
```

\* line함수는 정규분포의 곡선을 추가하는 명령

```
} # end function
```

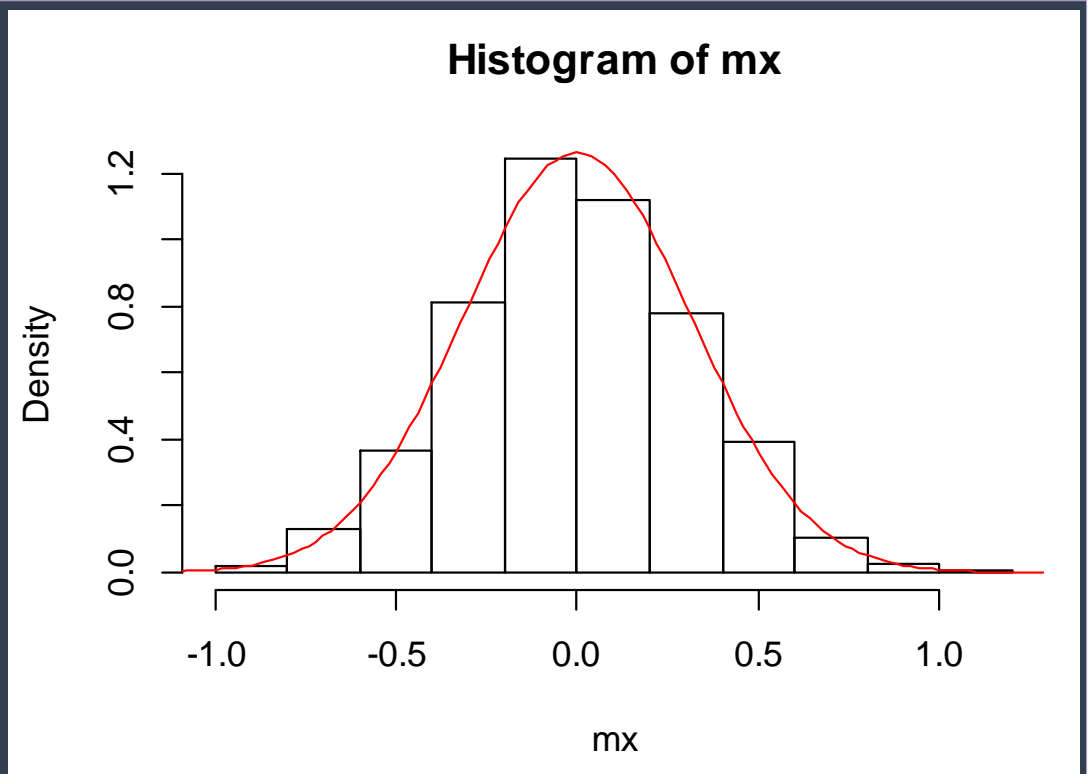
# 모의실험을 이용한 평균의 분포 확인

## 모의실험 2

### + 함수 호출 결과

> meanhist()

\* 정규분포와 가까운 것을 확인할 수 있음





# 이항분포의 정규근사

- 📖 이항분포와 정규분포의 근사
- 📖 모의실험을 이용한 이항분포의 정규근사 확인

Loren ipsum dolor sit amet, ius an molestie facilisi erroribus, mutat nalerum delectus ei vis. Has ornatus conclusionemque id, an videri molestatis sit. In etqui praesent sit. An vel agan porro comprehensan, ad ludus constituto nea, et ius utroque scaevola assuaverit.

Vis cu nodus nulla feugait, oratio facilisi in usu, eili vitae sea te. Ea fabulas accusamus dissentias sea, facete tacinates definitiones et per. Nihil dicant mediocrem pro eu, no mei nostro sensibus platonem. Qui id sunno perpetus neglegantur. Vel ipsum novum copiosae ut. Quo et liber detracto probatus. Nam augue scriben- tur an. Sea oporteat percipitur incidereat ab. Qui viris nemore an.





# 이항분포와 정규분포의 근사



## 자료 $X$

- +  $n$  개 중에서 성공의 개수인  $X$ 의 분포가 성공확률  $p$ 인 이항분포인 경우



## $X$ 의 분포

- + 기댓값  $np$ , 분산  $np(1-p)$ 인 정규분포로 근사할 수 있음( $n$ 이 충분히 클 때)
- + 이를 **이항분포의 정규근사** 라고 함



# 모의실험을 이용한 이항분포의 정규근사 확인



## 모의실험



### 이항분포의 정규근사

이항분포의 정규근사를 모의실험으로 확인해 보기로 한다.

```
norm.binom <- function(ndata=1000, nn=100, p=0.5) {  
# -----#  
# B(nn,p)에서 ndata개의 난수를 생성하여 이를 N(np, npq)와 비교  
# B(nn,p)의 확률 히스토그램과 정규분포의 pdf 그림  
# -----#  
  x <- rbinom(ndata, nn, p)  
  substr <- paste("n=", as.character(nn), ", p=", as.character(p), sep="")  
  hist(x, prob=T, main="B(n,p)와 N(np, npq)의 비교", sub=substr)
```

- x 에 이항분포에서의 난수를 ndata 개수만큼 생성하여 저장함
- hist 함수로 이에 대한 히스토그램을 그림 : 이때 주제목과 보조제목(substr)을 설정함



# 모의실험을 이용한 이항분포의 정규근사 확인



## 모의실험

```
mx <- nn*p  
sdx <- sqrt(nn*p*(1-p))  
xmin <- min(x)  
xmax <- max(x)  
xx <- seq(from=xmin, to=xmax, length.out=100)  
lines(xx, dnorm(xx, mx, sdx))  
} # end function
```

- 히스토그램을 정규분포에 비교하기 위해 평균을 mx, 표준편차를 sdx라고 하였음
- 정규분포의 곡선을 겹쳐 그리기 위해 자료 중 최소와 최대를 얻어 이 사이에 100개의 값을 얻어 xx라고 하였음
- xx에서의 dnorm 값을 얻어 정규분포 xx에서 정규분포의 확률밀도함수를 계산하였음
- 마지막으로 lines 함수를 사용하여 정규분포 곡선을 히스토그램에 겹쳐 그렸음



# 모의실험을 이용한 이항분포의 정규근사 확인



## 모의실험



## 함수 호출 결과

> norm.binom()

\* 이항분포가 정규분포에 근사함을 그림으로 확인할 수 있음

