

# 주사위 던지기 모의실험

📖 모의실험 개요

📖 주사위 던지기 모의실험과 나온 눈금의 기록

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

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 mediocram pro eu, no mei nostro sensibus platonem. Qui id sunno perpetus neglegentur. Vel ipsum novum copiosae ut. Quo et liber detracto probatus. Nam augue scriben- tur an. Sea oporteat percipitur incidereat ab. Qui viris nemore an.



# 모의실험 개요



## 실험방법

+ 공평한 주사위를  $n$ 번 던져서 각 눈금이 나오는 회수를 알아봄



## 실제실험 : 눈금 기록

+ 만일  $n=1,000$ 이면 주사위를 1,000번 던져 매번 눈금을 기록해야 함





# 주사위 던지기 모의실험과 나온 눈금의 기록



## 모의실험



### 주사위 던지기

공평한 주사위(각 눈금이 나올 확률이  $1/6$ 로 모두 같은 주사위)를 던져서 나오는 눈금을 기록한다면 각 눈금 1에서 6까지의 나올 확률이 모두 같은 모집단에서 복원 추출로 뽑으면 된다. 따라서 sample 함수나 sample.int 함수로 1,000개의 난수를 복원추출로 얻는 것과 같으므로 다음과 같은 명령을 사용하면 된다.



### sample 함수 사용 결과

```
> table(sample(seq(1,6), 1000, replace=T))
```

1	2	3	4	5	6
156	157	161	166	179	181

➤ 각 눈금이 나올 확률은  $1/6$ 이므로 1,000번 던지면 각 눈금이 약 166.67번 정도 나올 것임



# 주사위 던지기 모의실험과 나온 눈금의 기록



## 모의실험



### sample.int 함수 사용 결과

```
> table(sample.int(6, 1000, replace=T))
```

1	2	3	4	5	6
174	155	173	168	164	166



# 로또에 대한 모의실험

 로또 모의실험

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 dissonantia sea, facete tacinates definitiones et per. Nihil dicant mediocrem pro eu, no mei nostro sensibus platonem. Qui id sunno perpetua neglegentur. Vel ipsum novum copiosae ut. Quo et liber detracto probatus. Nam augue scribentur an. Sea oporteat percipitur incidere et. Qui viris nemore an.





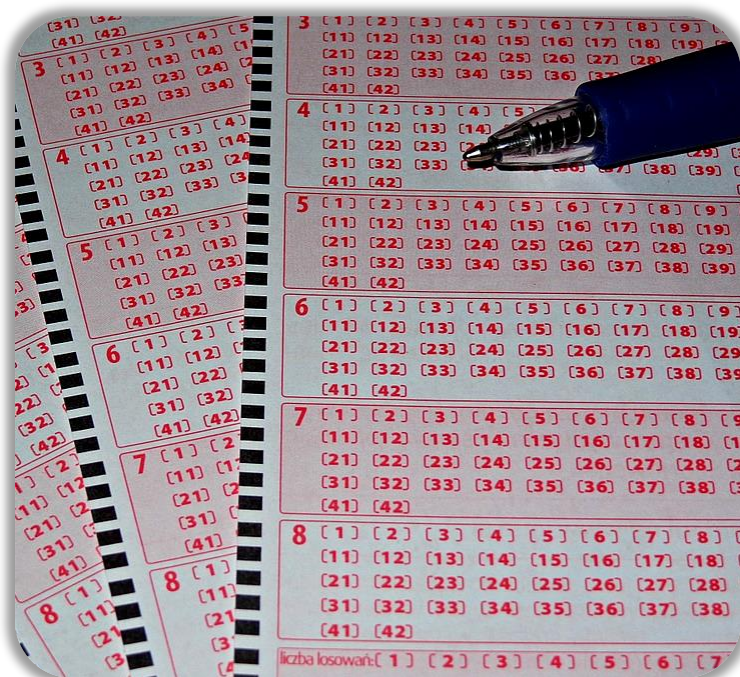
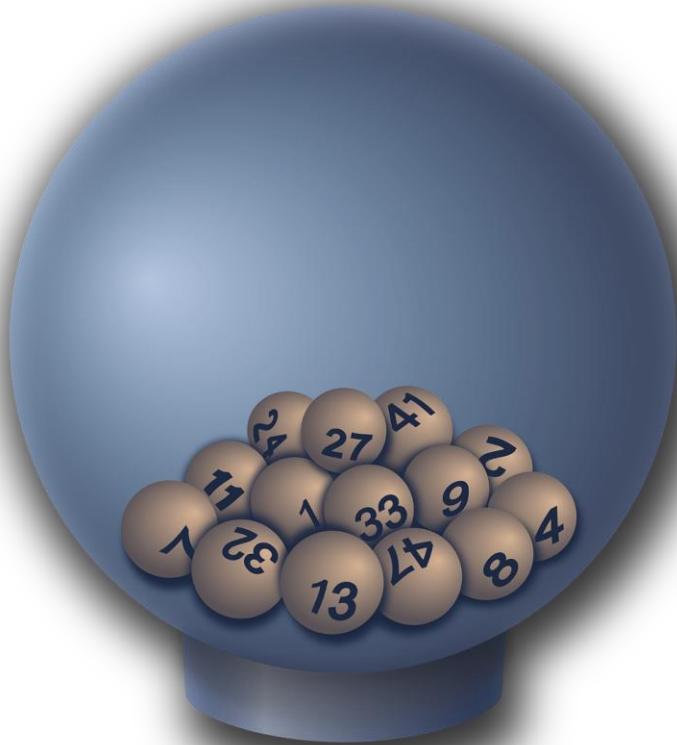


# 로또 모의실험



## 방법

+ 로또를 nn번 구매할 때 6개 당첨번호 중 0개, 1개, ..., 6개 모두가 맞는 경우를 알아봄





# 로또 모의실험



## 모의실험



### 실험 방법

로또를 nn번 구매할 때 하나도 맞지 않을 경우, 1만 일치하는 경우, ... 6개 모두가 일치하는 경우가 몇 번이나 나오는지 모의실험을 통해 알아보자.

```
lotto.sim <- function(hitnos, nn=100000) {  
#-----#  
# 45개의 번호 중 0개, 1개, ..., 6개의 번호를 모두 맞춘 회수와 실제  
#   맞출 확률을 계산  
# INPUT: hitnos = 6개의 당첨번호  
#         nn = 반복회수  
# OUTPUT: nhit = 0, 1, ..., 6개의 당첨번호를 맞춘 회수  
#          probhits = 0, 1, ..., 6개의 당첨번호를 맞춘 비율  
#          true.prob = 0, 1, ..., 6개의 당첨번호를 맞출 확률  
#-----#
```

- 매개변수로 당첨 번호 hitnos와 반복할 회수 nrep를 받음
- 기본값 100,000번으로 이 프로그램은 반복회수가 많아서 약간의 시간이 소요됨  
→ hitnos는 1-45 사이의 6개의 숫자를 가진 벡터임(당첨번호이므로 중복불가)



# 로또 모의실험



## 모의실험

```
nhits <- rep(0, 7)
for (i in 1:nn) {
  ncorrect <- sum(sample.int(45,6) %in% hitnos)
  nhits[ncorrect+1] <- nhits[ncorrect+1] + 1
}
```

- nhits에는 각각 당첨번호 hitnos와 임의로 뽑은 6개의 로또번호와 0개, 1개, ... 6개 일치된 회수를 기록함(이를 nn 만큼 반복함)
- sample.int(45,6) %in% hitnos 명령으로 hitnos 와 1-45 사이의 6개의 번호가 일치되는지 확인하고 이를 합함(일치=1, 불일치=0)
- 위의 결과를 nhits에 저장함







# 로또 모의실험



## 모의실험

```
prob.true <- rep(0,7)
for (i in 0:6) {
  prob.true[i+1] <- choose(6,i)*choose(39, 6-i)/ choose(45,6)
}

names(nhits) <- as.character(seq(0,6))
names(prob.true) <- as.character(seq(0,6))
prob.true <- round(prob.true, 8)
list(nhits=nhits, probhits=nhits/nn, prob.true=prob.true)
} # end function
```

- x개가 일치할 이론적 확률은  $\frac{\binom{6}{x} \binom{39}{6-x}}{\binom{45}{6}}$  이므로 이를 계산하여 prob.true에 저장
- nhits에 개수로 이름을 설정함



# 로또 모의실험



## 모의실험



### 함수 호출 결과

```
> source("c:/hwp/STI/lotto.sim.r", encoding="UTF-8")
```

```
> lotto.sim(c(1, 11, 21, 32, 33, 40))
```

```
$nhits
```

0	1	2	3	4	5	6
40084	42568	15000	2194	151	3	0

```
$probhits
```

0	1	2	3	4	5	6
0.40084	0.42568	0.15000	0.02194	0.00151	0.00003	0.00000

```
$prob.true
```

0	1	2	3	4	5	6
0.40056464	0.42412726	0.15147402	0.02244060	0.00136463	0.00002873	0.00000012

**\* 이론적인 당첨번호와 실제 확률이 가까운 값으로 얻어짐**

➤ 당첨번호는 임의의 서로 다른 6개의 번호를 사용하여도 결과는 같음(난수이므로 수치는 비슷함)

# 정규분포의 평균에 대한 신뢰구간 모의실험

📖 정규분포의 평균에 대한 신뢰구간

📖 신뢰구간에 대한 모의실험

Loren ipsum dolor sit amet, ius an molestie  
facilisi erroribus, mutat nalerum delectus ei  
vis. Has ornatus conclusionemque id, an vide  
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 mediocram 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 incidereat ab  
Qui viris nemore an.


# 정규분포의 평균에 대한 신뢰구간





 표준정규분포로부터 난수를 n개 만든 경우 모평균에 대한 95% 신뢰구간

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = \bar{x} \pm z_{\alpha/2} \frac{1}{\sqrt{n}}$$

$*\alpha = 0.05$

 자료를 얻어서 한 개의 신뢰구간을 얻으면 모집단의 평균은 해당 신뢰구간에 포함되거나 포함되지 않거나 둘 중의 하나이며 모평균이 자료에서 구한 신뢰구간에 포함될 확률이 95%인 것은 아님

 신뢰도(신뢰수준)이 95%라는 말의 뜻

 여러 번 신뢰구간을 얻으면 이들 신뢰구간의 95% 정도는 모평균을 포함하고 나머지 5% 정도는 신뢰구간을 포함하지 않는다는 의미임





# 신뢰구간에 대한 모의실험



## 모의실험



### 자료

$n = 10$ 인 경우의 95% 신뢰구간을 1000번 계산하여 1000개의 신뢰구간 중 모평균 0을 포함하는 경우가 몇 번이나 되는지 알아보자.

```
z.ci <- function(alpha = 0.05, ndata=10, nrep = 1000) {  
#-----#  
# 정규분포에서 ndata개의 자료를 사용하여 100(1-alpha)% 신뢰구간을  
# 계산하는 것을 nrep 만큼 반복할 때 모평균을 포함하는 신뢰구간의 수  
# source("d:/hwp/STI/z.ci.r", encoding="UTF-8")  
#-----#
```



# 신뢰구간에 대한 모의실험



## 모의실험

```
qz <- qnorm(1-alpha/2)
se <- 1/sqrt(ndata)
ncover <- 0
for (i in 1:nrep) {
  x <- rnorm(ndata)
  meanx <- mean(x)
  ubound <- meanx + qz*se
  lbound <- meanx - qz*se
  if (ubound > 0 & lbound < 0) ncover = ncover + 1
  # 신뢰구간에 포함되는 회수를 계산
} # end for
list(ncover=ncover) # 출력
} # end function
```

# 신뢰구간이 0을 포함하는 회수  
# nrep 번 (기본값 1000번) 반복  
# ndata 개(기본값 10개)의 난수 생성  
# ndata 개의 평균  
# 신뢰상한  
# 신뢰하한

- qz와 se로 각각  $1 - \alpha/2$  분위수와 표준오차  $\sigma/\sqrt{n}$ 를 구함
- 표준정규분포에서 ndata 만큼의 난수를 얻어 신뢰구간( $lbound <- meanx - qz*se$ ,  $ubound <- meanx + qz*se$ )을 구함
- 이 신뢰구간이 모평균 0을 포함하는지 확인
- 신뢰구간을 구하고 모평균 0을 포함하는지 확인하는 과정을 nrep 만큼 반복하여 0을 포함한 회수를 기록





# 신뢰구간에 대한 모의실험



## 모의실험



유의수준과 반복회수를 기본값으로 설정하여 호출한 결과

```
z.ci()
```

```
$ncover
```

```
[1] 943
```

- \* 1,000개 신뢰구간 중 943번은 실제 모평균  $\mu$ 를 포함하고 57개의 신뢰구간은 모평균을 포함하지 못하였음
- \* 난수를 이용한 모의실험이므로 프로그램을 실행할 때마다 다른 결과를 얻으나 약 95%인 950 내외의 값을 얻음