Chap04. Central Limit Theorem

node 4 and 5 are most central nodes.

```
library(sand)
```

```
## Loading required package: igraph
```

```
##
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
```

```
## The following object is masked from 'package:base':
##
##     union
```

```
## Loading required package: igraphdata
```

```
##
## Statistical Analysis of Network Data with R
## Type in C2 (+ENTER) to start with Chapter 2.
```

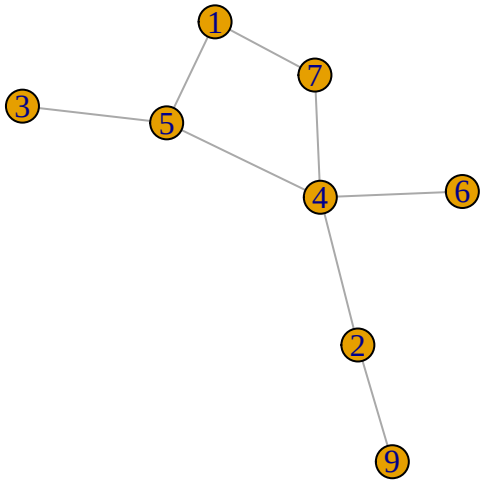
```
library(igraph)
g = graph.formula(1-5, 1-7, 2-9, 2-4, 3-5, 4-5, 4-6, 4-7)
V(g)
```

```
## + 8/8 vertices, named, from c05e06b:
## [1] 1 5 7 2 9 4 3 6
```

```
E(g)
```

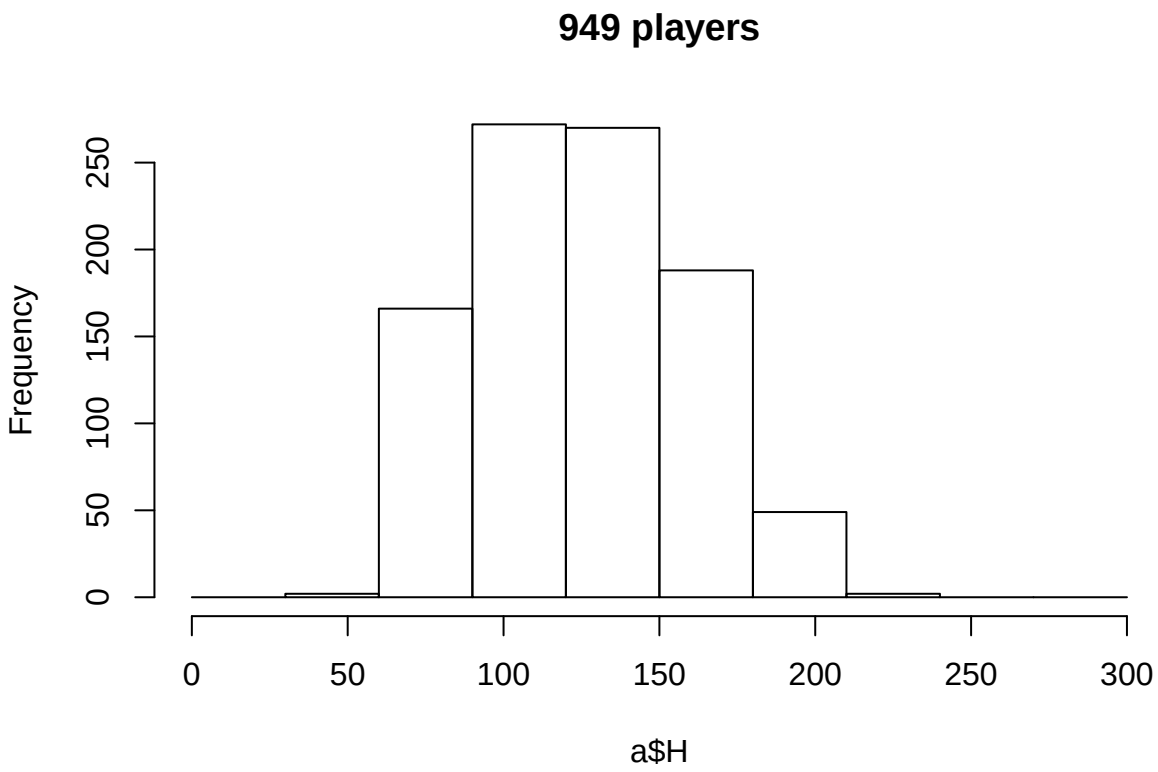
```
## + 8/8 edges from c05e06b (vertex names):
## [1] 1--5 1--7 5--4 5--3 7--4 2--9 2--4 4--6
```

```
plot(g)
```



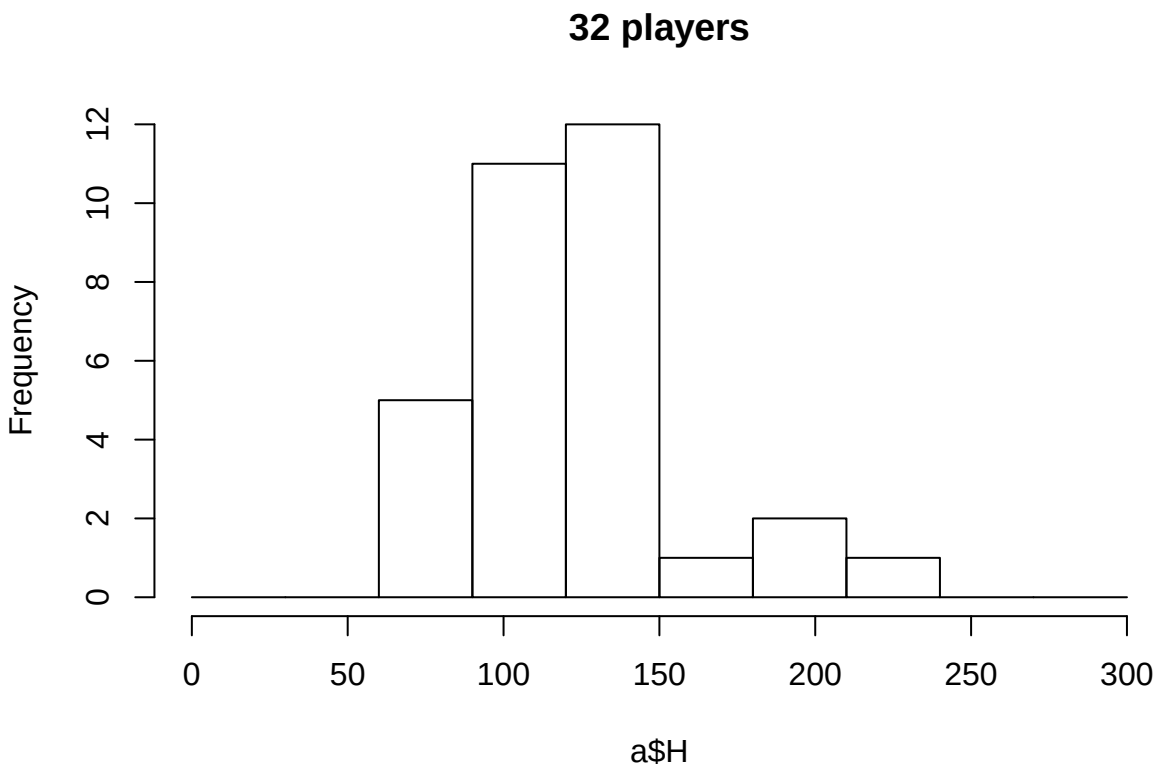
If there are many samples

```
library(Lahman)
a = subset(Batting, yearID > 2011 & yearID < 2016 & AB >=300)
hist(a$H, main='949 players', breaks=seq(from=0, to=300, by=30))
```



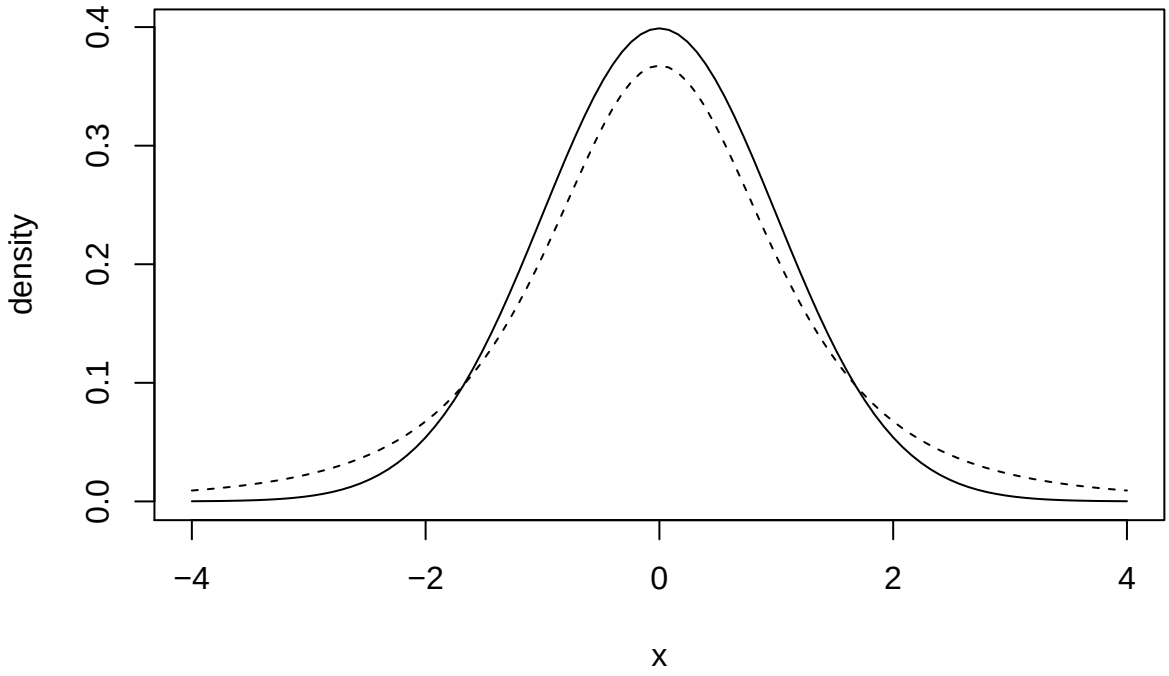
If there are not many samples

```
a = subset(Batting, yearID>2011&yearID<2016&AB>=300&teamID=='NYA')
hist(a$H, main='32 players', breaks=seq(from=0, to=300, by=30))
```



student-t is used when the number of samples is less than 30 and populations' variance is unknown.

```
curve(dnorm(x), -4, 4, ylab='density')
curve(dt(x, df=3), add=TRUE, lty=2)
```


density
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
x