

# Rであそぼ

## サンプリングをやってみる編

August 20, 2008

## 乱数を発生させる

- ❖ 一様分布
- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい
- ❖ 分布を比較する
- ❖ 分布を比較する
- ❖ 練習問題

サンプリング

## 乱数を発生させる

# 一様分布

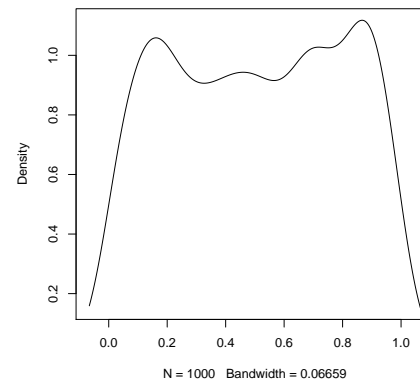
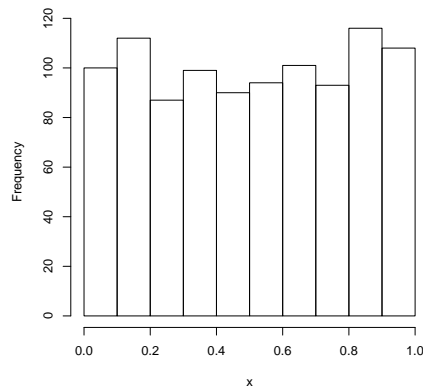
乱数を発生させる

## ❖ 一様分布

- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい
- ❖ 分布を比較する
- ❖ 分布を比較する
- ❖ 練習問題

サンプリング

```
> runif(8)
[1] 0.05894705 0.74916030 0.06164234 0.95508410 0.91973260
[6] 0.13457235 0.25409737 0.04707220
> runif(8,10,20)
[1] 10.33875 15.89624 14.44879 10.70855 13.06083 12.91227
[7] 15.49331 16.01266
> x<-runif(1000)
> hist(x)
> plot(density(x,cut=T))
```



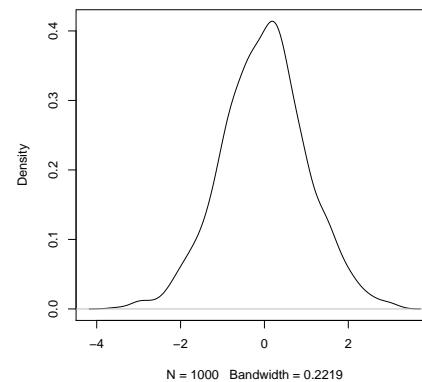
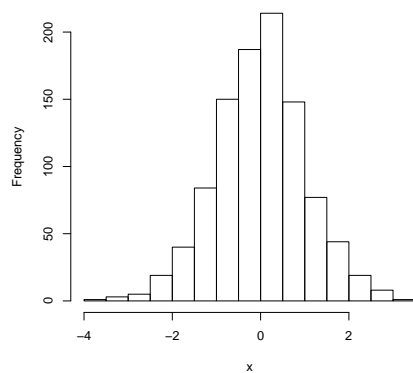
# 正規分布

## 乱数を発生させる

- ❖ 一様分布
- ❖ **正規分布**
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい
- ❖ 分布を比較する
- ❖ 分布を比較する
- ❖ 練習問題

## サンプリング

```
> rnorm(8)
[1]  1.387609674 -0.212740721  1.008203750  1.058998277
[5]  0.001050524  1.610348069 -0.247380784 -1.162312631
> rnorm(8, 30, 5)
[1] 30.96396 31.75022 30.82388 27.37124 30.11004 19.27867
[7] 32.54996 27.04288
> x<-runif(1000)
> plot(x)
> plot(density(x))
```



# $\chi^2$ 分布

乱数を発生させる

❖ 一様分布

❖ 正規分布

❖  $\chi^2$  分布

❖ t 分布

❖ F 分布

❖ ポアソン分布

❖ ヒストグラムと密度分布を一緒に描きたい

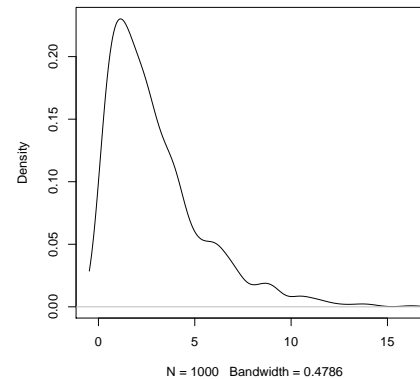
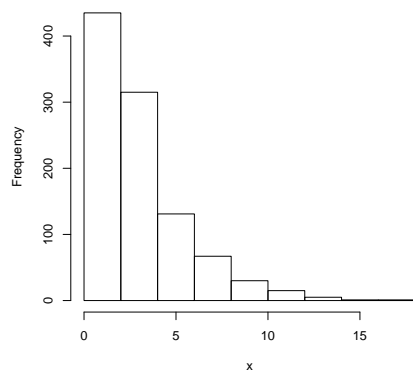
❖ 分布を比較する

❖ 分布を比較する

❖ 練習問題

サンプリング

```
> rchisq(8,1)
[1] 0.29669434 0.02767726 1.58979153 0.35215735 0.24488332
[6] 0.05255551 0.08142983 0.32275587
> x<-rchisq(1000,3)
> hist(x)
> plot(density(x,cut=T))
```



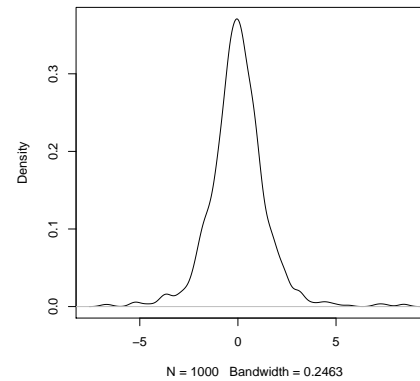
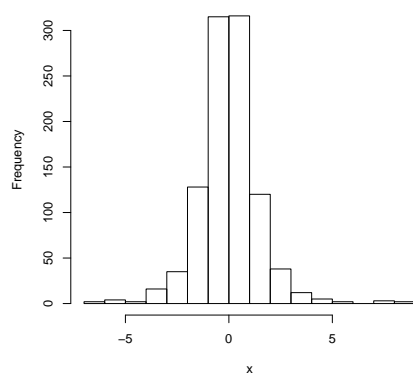
# t分布

## 乱数を発生させる

- ❖ 一様分布
- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布**
- ❖ F 分布
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい
- ❖ 分布を比較する
- ❖ 分布を比較する
- ❖ 練習問題

## サンプリング

```
> rt(8, 3)
[1] 0.86102824 -0.37013993 -0.79180135 0.88810515
[5] -0.35736905 -0.68623016 -0.91831565 2.06995619
> x<-rt(1000, 3)
> hist(x)
> plot(density(x))
```



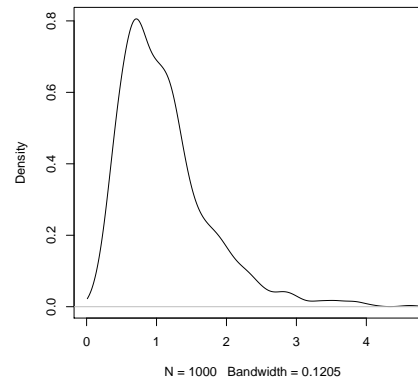
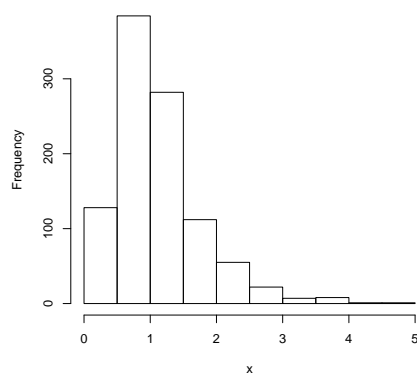
# F分布

## 乱数を発生させる

- ❖ 一様分布
- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布**
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい
- ❖ 分布を比較する
- ❖ 分布を比較する
- ❖ 練習問題

## サンプリング

```
> rf(10, 3, 5)
[1] 1.3804344 8.1968435 0.9039424 1.9950758 0.4637679
[6] 1.1678419 14.4302587 3.8488380 4.8697533 0.2572799
> x<-rf(1000,10,20)
> hist(x)
> plot(density(x, cut=T))
```



# ポアソン分布

## 乱数を発生させる

- ❖ 一様分布
- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布
- ❖ **ポアソン分布**

❖ ヒストグラムと密度分布を一緒に描きたい

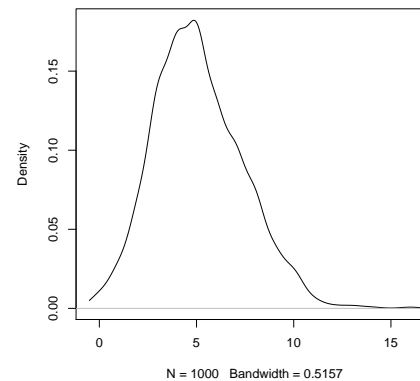
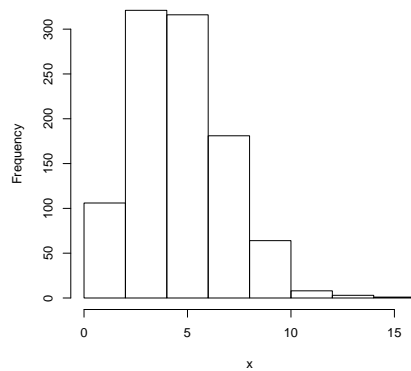
❖ 分布を比較する

❖ 分布を比較する

❖ 練習問題

## サンプリング

```
> rpois(8, 3)
[1] 2 2 3 1 4 3 5 2
> x<-rpois(1000, 5)
> hist(x)
> plot(density(x, cut=T))
```





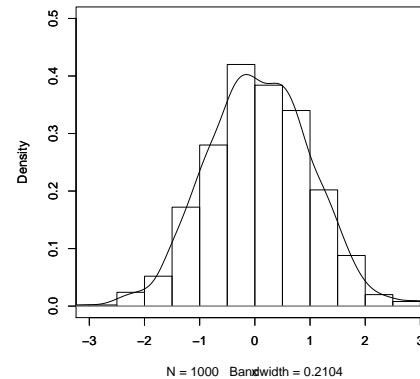
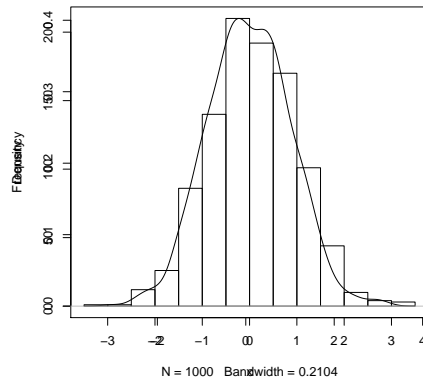
# ヒストグラムと密度分布を一緒に描きたい

## 乱数を発生させる

- ❖ 一様分布
- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい
- ❖ 分布を比較する
- ❖ 分布を比較する
- ❖ 練習問題

## サンプリング

```
> x<-rnorm(1000)
> hist(x)
> par(new=T)
> plot(density(x))
> hist(x,prob=T,xlim=c(-3,3),ylim=c(0,0.5))
> par(new=T)
> plot(density(x),xlim=c(-3,3),ylim=c(0,0.5))
```



# 分布を比較する

## 乱数を発生させる

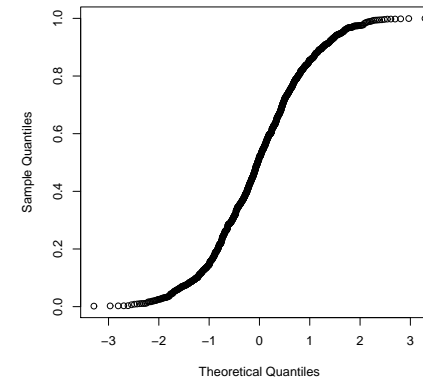
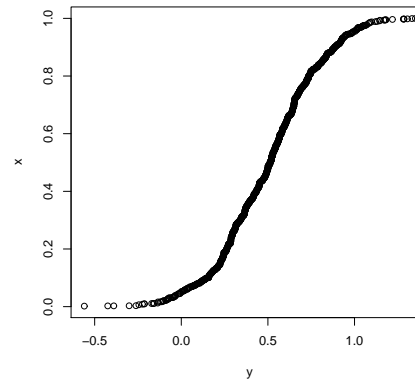
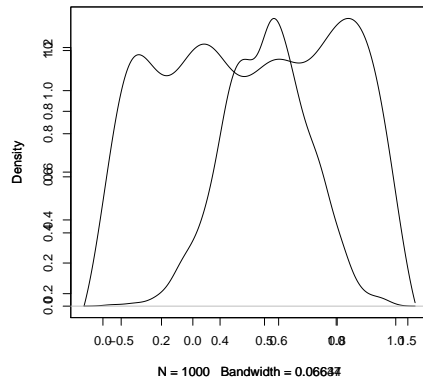
- ❖ 一様分布
- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい

## ❖ 分布を比較する

- ❖ 分布を比較する
- ❖ 練習問題

## サンプリング

```
> x<-runif(1000)
> y<-rnorm(1000,mean(x),sd(x))
> plot(density(x,cut=T))
> par(new=T)
> plot(density(y))
> qqplot(y,x)
> qqnorm(x)
```



# 分布を比較する

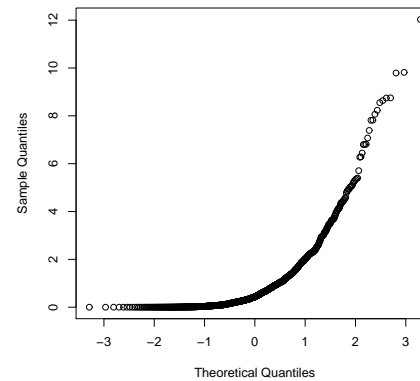
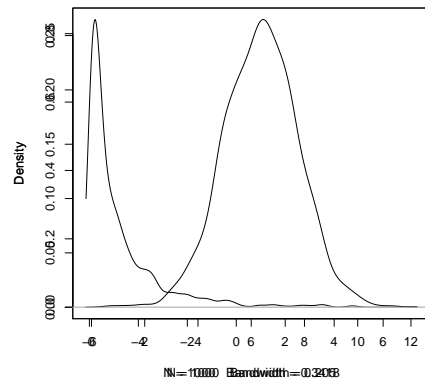
## 乱数を発生させる

- ❖ 一様分布
- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい
- ❖ 分布を比較する
- ❖ 分布を比較する**
- ❖ 練習問題

## サンプリング

## 分布と比較する

```
> x<-rchisq(1000,1)
> y<-rnorm(1000,mean(x),sd(x))
> plot(density(x,cut=T))
> par(new=T)
> plot(density(y))
> qqnorm(x)
```



# 練習問題

## 乱数を発生させる

- ❖ 一様分布
- ❖ 正規分布
- ❖  $\chi^2$  分布
- ❖ t 分布
- ❖ F 分布
- ❖ ポアソン分布
- ❖ ヒストグラムと密度分布を一緒に描きたい
- ❖ 分布を比較する
- ❖ 分布を比較する

## ❖ 練習問題

## サンプリング

### 練習問題

1. 平均-3, 標準偏差 4 の正規乱数 1000 個を発生させ, 変数  $x$  に代入しなさい。
2. 平均 4, 標準偏差 0.5 の正規乱数 1000 個を発生させ, 変数  $y$  に代入しなさい。
3. 変数  $x$  と  $y$  の和 ( $x+z$ ) を  $z$  に代入しなさい。
4. 変数  $x$  と  $y$  と  $z$  の密度分布を同時に描きなさい。
5. 変数  $z$  の分布が正規分布に近似できるか `qqnorm` で確認しなさい。

乱数を発生させる

## サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

# サンプリング

# sample 関数

乱数を発生させる

サンプリング

❖ sample 関数

❖ とったサンプルを蓄積する

❖ for 関数を使う

❖ for 関数の中で for 関数を使う

❖ 標本平均をとる

❖ 標本平均の平均をとる

❖ データフレームからサンプルをとる

❖ とったサンプルを蓄積する

❖ 2 標本平均の分布

❖ 5 標本平均の分布

❖ OLS 推定量の分布

❖ 関数をつくる

❖ OLS 推定量を発生させる関数

❖ 関数を使ってみる

❖ 練習問題

❖ 練習問題

```
> x<-1:10
> x
[1] 1 2 3 4 5 6 7 8 9 10
> sample(x,5)
[1] 5 9 6 7 1
> sample(x,5)
[1] 1 3 2 10 6
> sample(x,5,replace=T)
[1] 2 2 2 2 6
> sample(x,5,replace=T)
[1] 1 9 6 6 8
```

# とったサンプルを蓄積する

乱数を発生させる

サンプリング

❖ sample 関数

❖ とったサンプルを蓄積する

❖ for 関数を使う

❖ for 関数の中で for 関数を使う

❖ 標本平均をとる

❖ 標本平均の平均をとる

❖ データフレームからサンプルをとる

❖ とったサンプルを蓄積する

❖ 2 標本平均の分布

❖ 5 標本平均の分布

❖ OLS 推定量の分布

❖ 関数をつくる

❖ OLS 推定量を発生させる関数

❖ 関数を使ってみる

❖ 練習問題

❖ 練習問題

```
> y<-matrix(nrow=3,ncol=5)
> y
      [,1] [,2] [,3] [,4] [,5]
[1,]  NA  NA  NA  NA  NA
[2,]  NA  NA  NA  NA  NA
[3,]  NA  NA  NA  NA  NA
> y[1,]<-sample(x,5)
> y[2,]<-sample(x,5)
> y[3,]<-sample(x,5)
> y
      [,1] [,2] [,3] [,4] [,5]
[1,]    7   10    5    9    1
[2,]    2    7    1    8    3
[3,]    9   10    1    5    3
> for(i in 1:3) y[i,]<-sample(x,5)
> y
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    5    3    9    4
[2,]    7    4    1    2    6
[3,]   10    8    5    9    3
```

# for関数を使う

乱数を発生させる

サンプリング

❖ sample 関数

❖ とったサンプルを蓄積する

❖ for 関数を使う

❖ for 関数の中で for 関数を使う

❖ 標本平均をとる

❖ 標本平均の平均をとる

❖ データフレームからサンプルをとる

❖ とったサンプルを蓄積する

❖ 2 標本平均の分布

❖ 5 標本平均の分布

❖ OLS 推定量の分布

❖ 関数をつくる

❖ OLS 推定量を発生させる関数

❖ 関数を使ってみる

❖ 練習問題

❖ 練習問題

```
> x<-rep(NA,10)
> x
[1] NA NA NA NA NA NA NA NA NA NA
> for(i in 1:10) x[i]<-i*10
> x
[1] 10 20 30 40 50 60 70 80 90 100
> for(i in 1:10){
+ x[i]<-i*10
+ }
> x
[1] 10 20 30 40 50 60 70 80 90 100
```



## for関数の中でfor関数を使う

乱数を発生させる

サンプリング

❖ sample 関数

❖ とったサンプルを蓄積する

❖ for 関数を使う

❖ for 関数の中で for 関数を使う

❖ 標本平均をとる

❖ 標本平均の平均をとる

❖ データフレームからサンプルをとる

❖ とったサンプルを蓄積する

❖ 2 標本平均の分布

❖ 5 標本平均の分布

❖ OLS 推定量の分布

❖ 関数をつくる

❖ OLS 推定量を発生させる関数

❖ 関数を使ってみる

❖ 練習問題

❖ 練習問題

```
> x<-matrix(nrow=3,ncol=7)
> for(i in 1:3){
+   for(j in 1:7){
+     x[i,j]<-i*10+j
+   }
+ }
> x
      [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,]   11   12   13   14   15   16   17
[2,]   21   22   23   24   25   26   27
[3,]   31   32   33   34   35   36   37
```

# 標本平均をとる

乱数を発生させる

サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる**
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

## 標本平均をとる

```
> x<-1:10
> y<-matrix(nrow=3,ncol=5)
> for(i in 1:3) y[i,]<-sample(x,5)
> y
      [,1] [,2] [,3] [,4] [,5]
[1,]   10    5    9    4    8
[2,]    7    9    6    1    3
[3,]    5    8    7   10    2
> apply(y,1,mean)
[1] 7.2 5.2 6.4
> z<-apply(y,1,mean)
> z
[1] 7.2 5.2 6.4
> mean(z)
[1] 6.266667
```

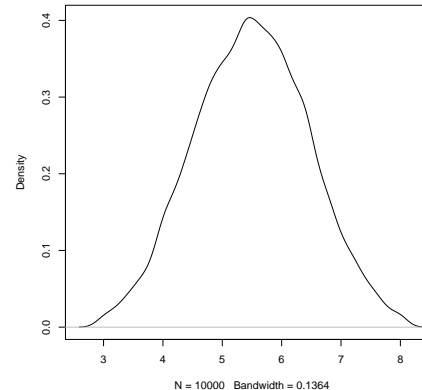
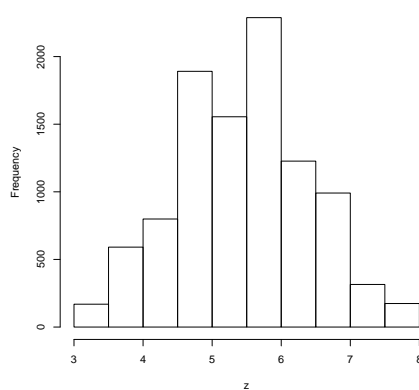
# 標本平均の平均をとる

乱数を発生させる

サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる**
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

```
> x<-1:10
> y<-matrix(nrow=10000,ncol=5)
> for(i in 1:10000) y[i,]<-sample(x,5) for(i in 1:10000) y[i,]<-sam
> z<-apply(y,1,mean)
> hist(z)
> plot(density(z))
> mean(z)
[1] 5.4987
> mean(x)
[1] 5.5
```



# データフレームからサンプルをとる

乱数を発生させる

サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

```
> x<-data.frame(aaa=c(1:5),bbb=c(11:15),ccc=c(101:105))
> x
  aaa bbb ccc
1   1  11 101
2   2  12 102
3   3  13 103
4   4  14 104
5   5  15 105
> x[sample(1:5,2),]
  aaa bbb ccc
1   1  11 101
3   3  13 103
> x[sample(1:5,2),]
  aaa bbb ccc
2   2  12 102
4   4  14 104
```

# とったサンプルを蓄積する

乱数を発生させる

サンプリング

❖ sample 関数

❖ とったサンプルを蓄積する

❖ for 関数を使う

❖ for 関数の中で for 関数を使う

❖ 標本平均をとる

❖ 標本平均の平均をとる

❖ データフレームからサンプルをとる

❖ とったサンプルを蓄積する

❖ 2 標本平均の分布

❖ 5 標本平均の分布

❖ OLS 推定量の分布

❖ 関数をつくる

❖ OLS 推定量を発生させる関数

❖ 関数を使ってみる

❖ 練習問題

❖ 練習問題

```
> x<-data.frame(aaa=c(1:5),bbb=c(11:15),ccc=c(101:105))
```

```
> y<-array(dim=c(2,3,4))
```

```
> for(i in 1:4) y[,,i]<-as.matrix(x[sample(1:5,2),])
```

```
> y
```

```
, , 1
```

```
      [,1] [,2] [,3]
[1,]     1    11   101
[2,]     5    15   105
```

```
, , 2
```

```
      [,1] [,2] [,3]
[1,]     3    13   103
[2,]     4    14   104
```

```
, , 3
```

```
      [,1] [,2] [,3]
[1,]     3    13   103
[2,]     2    12   102
```

```
, , 4
```

```
      [,1] [,2] [,3]
[1,]     2    12   102
[2,]     3    13   103
```

## 2 標本平均の分布

乱数を発生させる

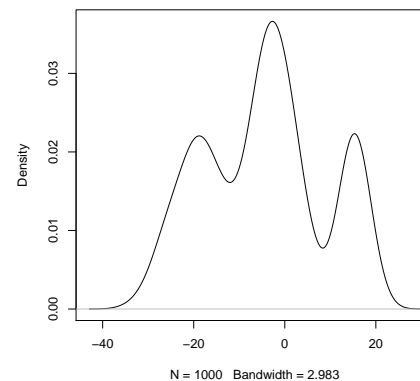
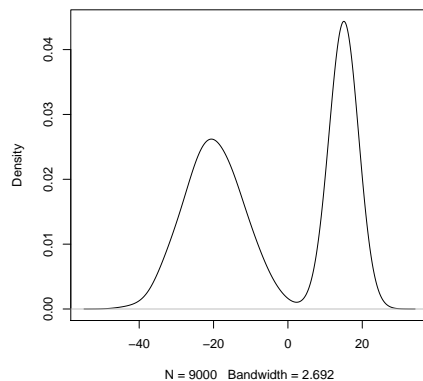
サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する

❖ 2 標本平均の分布

- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

```
> x<-c(rnorm(5000,-20,8),rnorm(4000,15,3))
> plot(density(x))
> y<-rep(NA,1000)
> for(i in 1:1000) y[i]<-mean(sample(x,2))
> plot(density(y))
> mean(x)
[1] -4.432267
> mean(y)
[1] -4.427477
```



## 5 標本平均の分布

乱数を発生させる

サンプリング

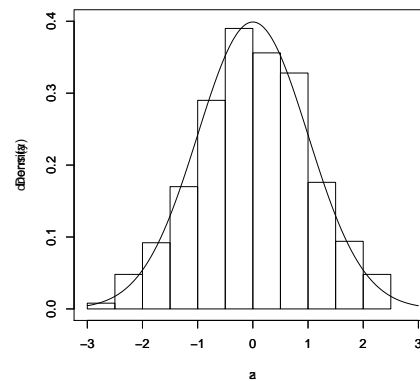
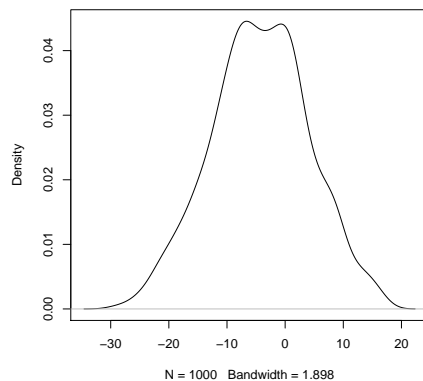
- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する

❖ 2 標本平均の分布

**❖ 5 標本平均の分布**

- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

```
> for(i in 1:1000) y[i]<-mean(sample(x,5))
> plot(density(y))
> mean(y)
[1] -4.281371
> z<-(y-mean(y))/sd(y)
> hist(z,xlim=c(-3,3),ylim=c(0,.4),prob=T)
> par(new=T)
> a<-seq(-3,3,0.01)
> plot(a,dnorm(a),xlim=c(-3,3),ylim=c(0,.4),type="l")
```



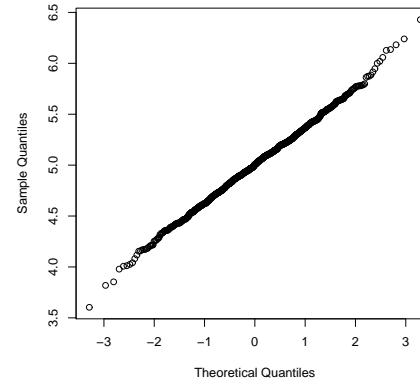
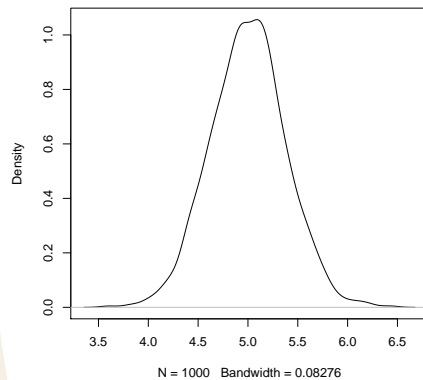
# OLS 推定量の分布

乱数を発生させる

サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布**
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

```
> x<-matrix(1:20,20,1)
> y<-matrix(nrow=20,ncol=1000)
> for(i in 1:1000) y[,i]<-5*x+rnorm(20,0,20)
> for(i in 1:1000) z[i]<-coef(lm(y[,i]~x-1))
> plot(density(z))
> qqnorm(z)
> mean(z)
[1] 5.001683
> sd(z)
[1] 0.3775910
> summary(lm(y[,1]~x-1))
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
x      5.0469      0.3319   15.21 4.33e-12 ***
```





# 関数をつくる

## 乱数を発生させる

## サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる**
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

```
> x<-3
> y<-5
> z<-x/y
> z
[1] 0.6
> y<-2
> x<-4
> z<-x/y
> z
[1] 2
> f01<-function(x,y) {
+ z<-x/y
+ z
+ }
> f01(3,2)
[1] 1.5
> f01
function(x,y) {
z<-x/y
z
}
```

# OLS 推定量を発生させる関数

乱数を発生させる

サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

```
> x<-runif(50,0,20)
> y<-5*x+rnorm(50,0,2)
> f02<-function(d01,p01){
+   y<-5*d01+rnorm(length(d01),0,p01)
+   z<-lm(y~d01-1)
+   a<-coef(z)
+   a
+ }
> f02(x,10)
      d01
5.018042
> f03<-function(d01,p01){
+   z<-rep(NA,1000)
+   for(i in 1:1000) z[i]<-f02(d01,p01)
+   z
+ }
```

# 関数を使ってみる

## 乱数を発生させる

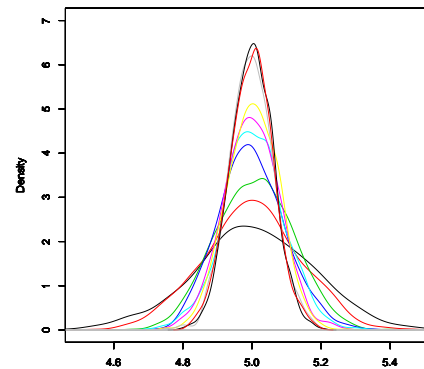
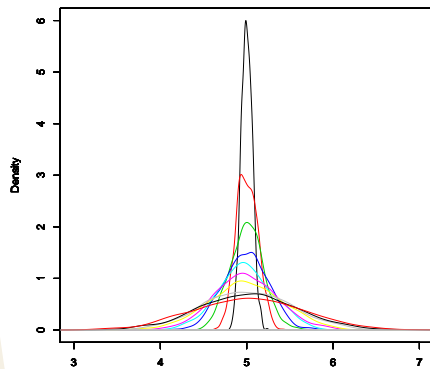
## サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数

## ❖ 関数を使ってみる

- ❖ 練習問題
- ❖ 練習問題

```
> oldpar<-par()
> x<-runif(50,0,20)
> for(i in 1:10){
+   z<-f03(x,i*5)
+   plot(density(z),col=i,xlim=c(3,7),ylim=c(0,5))
+   par(new=T)
+ }
> par(oldpar)
> for(i in 1:10){
+   x<-runif(20*i,0,20)
+   z<-f03(x,10)
+   plot(density(z),col=i,xlim=c(4.5,5.5),ylim=c(0,7))
+   par(new=T)
+ }
```



# 練習問題

乱数を発生させる

サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

## 練習問題

1. 平均 0, 標準偏差 5 の正規乱数 5000 個と [10,20] の一様乱数 5000 個を結合した 10000 個の乱数を母集団として, 変数  $x$  に代入しなさい。
2.  $x$  の密度分布を描きなさい。
3.  $x$  から 2 個だけ無作為抽出した値の平均を 1000 通り求め, 変数  $y$  に代入しなさい。
4.  $y$  の密度分布を描きなさい。
5.  $y$  の平均と分散を求めて,  $x$  の平均と分散と比較しなさい。
6.  $x$  から 10 個無作為抽出した値の平均を 1000 通り求め, 変数  $y$  に代入しなさい。
7.  $y$  の密度分布を描きなさい。
8.  $y$  の平均と分散を求めて,  $x$  の平均と分散と比較しなさい。
9. 上の 3. または 6. の作業を任意のサンプル数で行い,  $y$  の平均値と分散を返す関数を定義し, この関数を用いて, サンプル数が 2~10 の各数値の場合に, 標本平均の平均と分散がどのように代わるかを調べなさい。

# 練習問題

乱数を発生させる

サンプリング

- ❖ sample 関数
- ❖ とったサンプルを蓄積する
- ❖ for 関数を使う
- ❖ for 関数の中で for 関数を使う
- ❖ 標本平均をとる
- ❖ 標本平均の平均をとる
- ❖ データフレームからサンプルをとる
- ❖ とったサンプルを蓄積する
- ❖ 2 標本平均の分布
- ❖ 5 標本平均の分布
- ❖ OLS 推定量の分布
- ❖ 関数をつくる
- ❖ OLS 推定量を発生させる関数
- ❖ 関数を使ってみる
- ❖ 練習問題
- ❖ 練習問題

練習問題

OLS でいう  $E(\varepsilon|x) = 0$  を再現しなさい。