

# Problem Set3

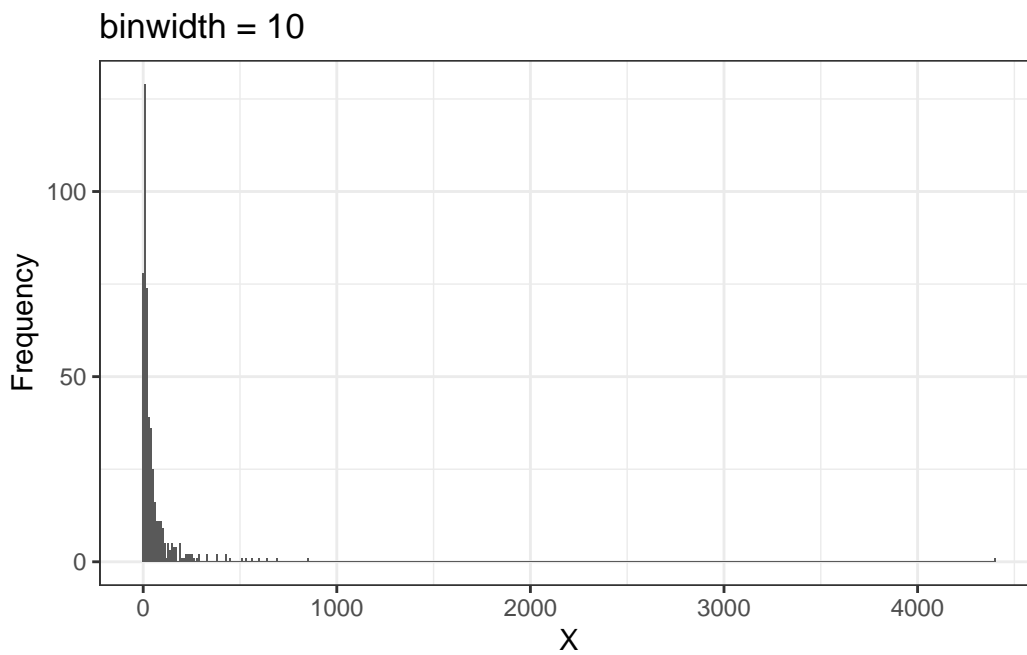
Yamato Igarashi (2125701)

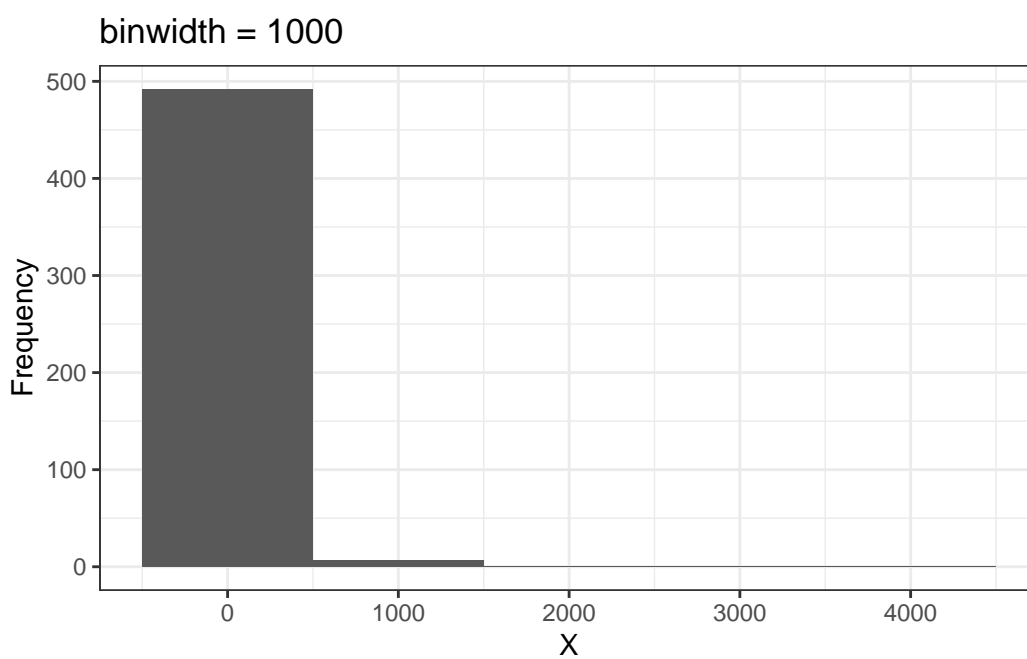
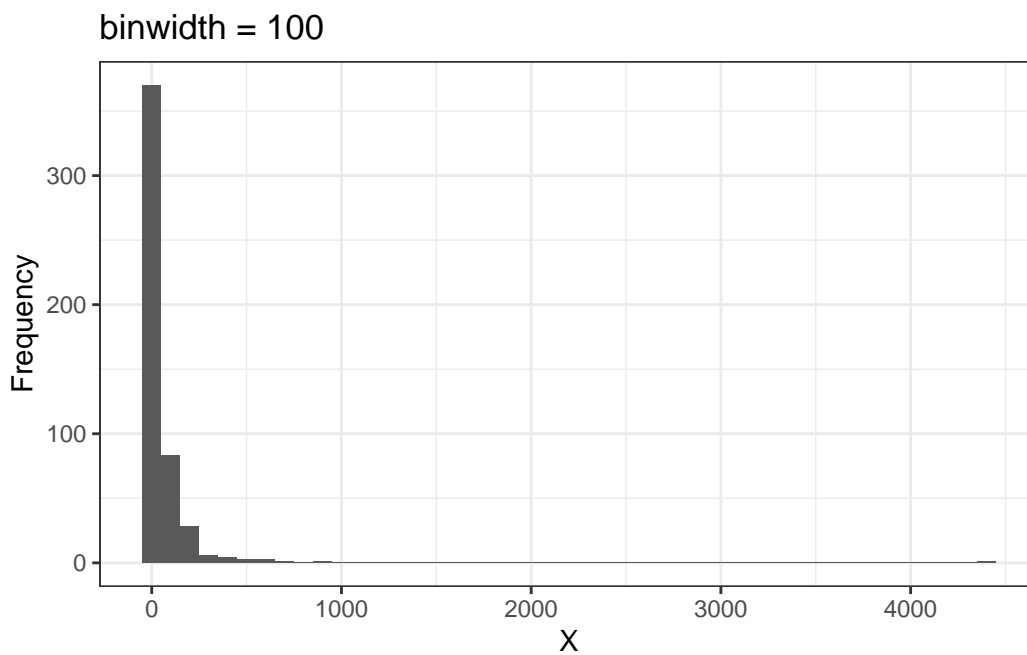
GitHub URL: [https://github.com/yamato5810/MicroDataScience\\_Intermediate\\_ProblemSet3](https://github.com/yamato5810/MicroDataScience_Intermediate_ProblemSet3)

## 2. 分布の推定

### a. ヒストグラム

下記のグラフのように、ヒストグラムは、階級の幅を狭くすると滑らかな密度関数が描けず、幅を広くすると分布の形状を特定しづらくなる。

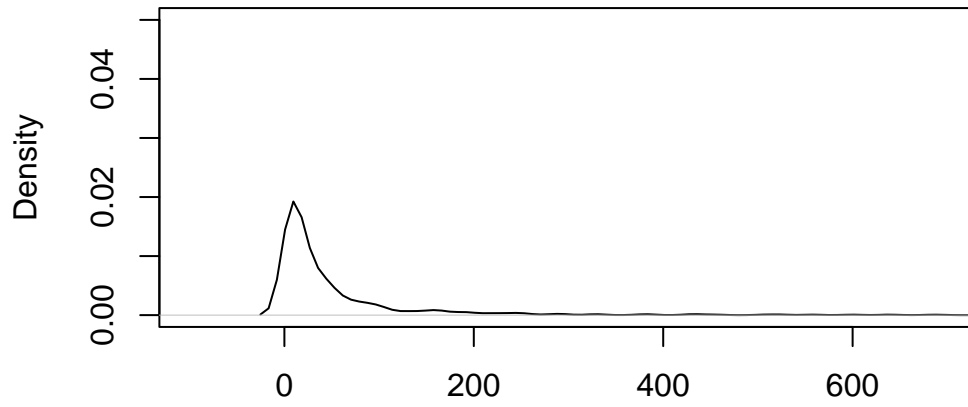




## b. カーネル密度

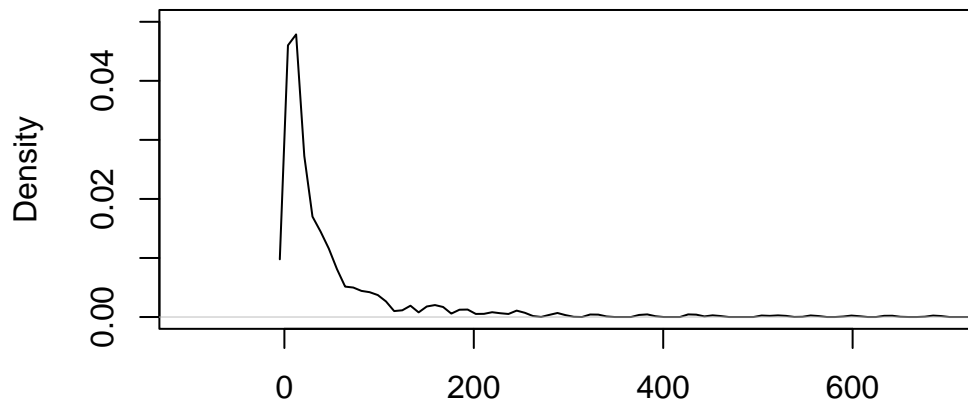
下記のグラフのように、bandwidth を狭くすると形状が少しいびつになり、bandwidth を広くすると滑らかになりすぎる。

**bandwidth = 1\*default**



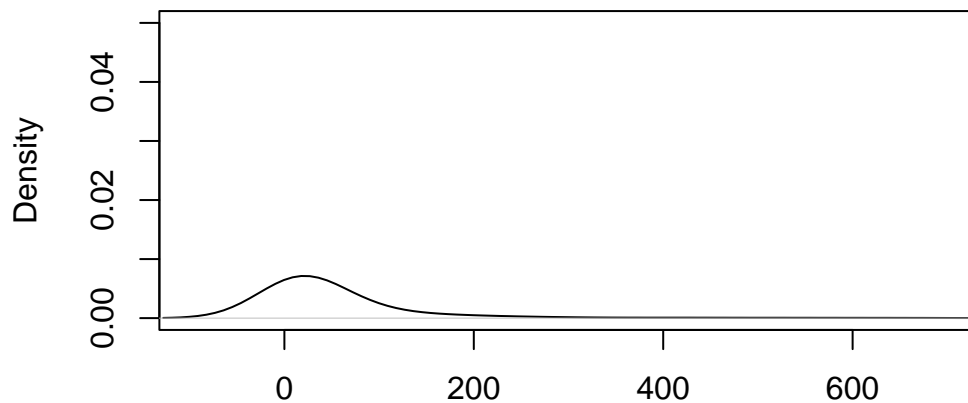
N = 500 Bandwidth = 8.545

**bandwidth = 0.2\*default**



N = 500 Bandwidth = 1.709

**bandwidth = 5\*default**



N = 500 Bandwidth = 42.73

### c. 分位回帰

以下は、線形回帰の結果である。

	Estimate	Std. Error	t value	Pr(> t )	CI Lower
(Intercept)	-0.01034966	0.0789366154	-0.1311136	0.8957384	-0.1654395
X	2.00044584	0.0005703533	3507.3800609	0.0000000	1.9993252

	CI Upper	DF
(Intercept)	0.1447402	498
X	2.0015664	498

以下は、中央値の分位回帰の結果である。

Call:

```
quantreg::rq(formula = Z ~ X, tau = Tau, data = Simulation_Data)
```

Coefficients:

	X
(Intercept)	-0.01625955
	2.00074829

Degrees of freedom: 500 total; 498 residual

以下は、下位 25% に関する分位回帰の結果である。

```
Call:
quantreg::rq(formula = Z ~ X, tau = Tau, data = Simulation_Data)
```

```
Coefficients:
(Intercept)          X
-0.9770873    1.9992207
```

```
Degrees of freedom: 500 total; 498 residual
```

以下は、上位 25% に関する分位回帰の結果である。

```
Call:
quantreg::rq(formula = Z ~ X, tau = Tau, data = Simulation_Data)
```

```
Coefficients:
(Intercept)          X
  1.024205    2.000512
```

```
Degrees of freedom: 500 total; 498 residual
```

これらからわかるように、切片は上位 25%, 中央値≒平均値, 下位 25% の順に、小さくなっている。一方で、傾きはそれぞれの結果で大きな差は見受けられない。

### 3. 非線形回帰 [ボーナス]

#### a. プロビット回帰分析

以下のように Probit で推定すると、係数は 0.245036 となっており、真の値である 1/4 にほぼ等しい。

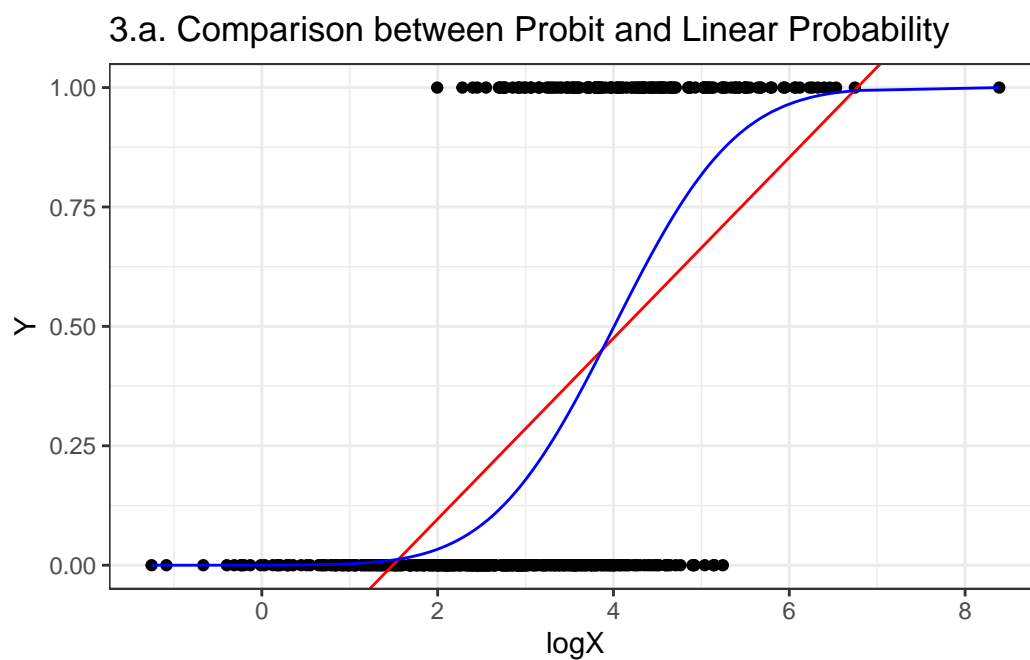
```
Call:
mfx::probitmfx(formula = Y ~ logX, data = Simulation_Data)
```

```
Marginal Effects:
      dF/dx Std. Err.      z    P>|z|
logX 0.245036  0.020031 12.233 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

一方、線形回帰では、0.1892428 と係数が小さく推定されている。

	Estimate	Std. Error	t value	Pr(> t )	CI Lower	CI Upper
(Intercept)	-0.2817887	0.026828992	-10.50314	1.914006e-23	-0.3345007	-0.2290767
logX	0.1892428	0.009697096	19.51541	2.035779e-63	0.1701906	0.2082951
DF						
(Intercept)	498					
logX	498					

これは、以下のグラフからもわかる通り、線形回帰の場合は外れ値に引っ張られ、傾きがより緩やかになっているからである。



## b. LASSO 回帰分析

下記のような分析結果が得られた。罰則項が大きい ( $\lambda = 0.1$ ) と、線形回帰での係数の値より小さく推定された。罰則項が小さいと、3.a. で求めた線形回帰の結果とほぼ等しい結果になった (Probit や真の値よりは依然として小さい)。

一方で、どのモデルにおいても、足の速さ、シュートの精度、背の高さ、年齢、チームワークのよさなどの他の項の影響は、一部残るものの、ほぼ 0、または、排除されていた。

以下参考 (LASSO 回帰)

■(1)  $\lambda = 0.1$

[[1]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -0.07331189

logX 0.11888789

speed .

shoot .

height .

age .

teamwork .

[[2]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -0.07335325

logX 0.12106066

speed .

shoot .

height .

age .

teamwork .

[[3]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -0.06694582

logX 0.11695236

speed .

shoot .

height .

age .

teamwork .

[[4]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -0.06603136

logX 0.11853582

speed .

shoot .

height .

```
age      .
teamwork .
```

```
[[5]]
7 x 1 sparse Matrix of class "dgCMatrix"

              s0
(Intercept) -0.07391989
logX        0.12061220
speed       .
shoot       .
height      .
age         .
teamwork    .
```

■(2)  $\lambda = 0.01$

```
[[1]]
7 x 1 sparse Matrix of class "dgCMatrix"

              s0
(Intercept) -2.602049e-01
logX        1.798756e-01
speed       .
shoot       4.140434e-05
height      .
age         .
teamwork    .
```

```
[[2]]
7 x 1 sparse Matrix of class "dgCMatrix"

              s0
(Intercept) -2.603685e-01
logX        1.820568e-01
speed       .
shoot       2.962928e-05
height      .
age         .
teamwork    .
```

```
[[3]]
```



7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -2.575351e-01  
logX 1.799267e-01  
speed .  
shoot 9.347162e-07  
height .  
age .  
teamwork .

[[4]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -0.2596696  
logX 0.1821298  
speed .  
shoot .  
height .  
age .  
teamwork .

[[5]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -0.2663547  
logX 0.1837996  
speed .  
shoot .  
height .  
age .  
teamwork .

■(3) l<sub>lambda</sub> = 0.001

[[1]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -2.750232e-01  
logX 1.841276e-01

```

speed      .
shoot      7.491747e-05
height     .
age        .
teamwork   .

```

```

[[2]]
7 x 1 sparse Matrix of class "dgCMatrix"
              s0
(Intercept) -2.751212e-01
logX         1.862730e-01
speed        .
shoot        6.217188e-05
height       .
age          .
teamwork     .

```

```

[[3]]
7 x 1 sparse Matrix of class "dgCMatrix"
              s0
(Intercept) -0.2723614257
logX         0.1842503317
speed        .
shoot        0.0000295428
height       .
age          .
teamwork     .

```

```

[[4]]
7 x 1 sparse Matrix of class "dgCMatrix"
              s0
(Intercept) -2.750628e-01
logX         1.866445e-01
speed        .
shoot        2.714117e-05
height       .
age          .
teamwork     .

```

```

[[5]]

```

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -2.838309e-01  
logX 1.893091e-01  
speed .  
shoot 1.102245e-05  
height .  
age .  
teamwork .

■ optimal lambda

[[1]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -2.740218e-01  
logX 1.838402e-01  
speed .  
shoot 7.265282e-05  
height .  
age .  
teamwork .

[[2]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -2.740752e-01  
logX 1.859740e-01  
speed .  
shoot 5.986424e-05  
height .  
age .  
teamwork .

[[3]]

7 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -2.716060e-01  
logX 1.840300e-01

```

speed      .
shoot      2.808644e-05
height     .
age        .
teamwork   .

```

```

[[4]]
7 x 1 sparse Matrix of class "dgCMatrix"
      s0
(Intercept) -2.742908e-01
logX        1.864213e-01
speed       .
shoot       2.561525e-05
height      .
age         .
teamwork    .

```

```

[[5]]
7 x 1 sparse Matrix of class "dgCMatrix"
      s0
(Intercept) -2.830300e-01
logX        1.890766e-01
speed       .
shoot       9.554914e-06
height      .
age         .
teamwork    .

```