

対応のある2群の分布の位置の比較 — ウィルコクソンの符号順位和検定

`wilcox.test`(中央値を比較したい変数1, 中央値を比較したい変数2, `paired=TRUE`)

`paired=TRUE` を指定するのを忘れないこと。忘れると対応のない検定になってしまう。

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
> d1 <- read.table("対応のある2群の分布位置データ.csv", header=TRUE, sep=",")
> head(d1)
  番号 患者 家族
1    1    5    4
2    2    3    2
3    3    5    3
4    4    3    3
5    5    2    3
6    6    2    2
>
> dtmp <- d1[,c(-1)]
> d2 <- stack(dtmp)
>
>
```

```
> # 標本サイズ, 平均, SD, 相関係数
> ntmp <- nrow(dtmp)
> mtmp <- colMeans(dtmp)
> stmp <- apply(dtmp, 2, sd)
> ctmp <- cor(dtmp)
> ktmp <- round(data.frame(ntmp, mtmp, stmp, ctmp), 2)
> colnames(ktmp) <- c("N", "Mean", "SD", colnames(ctmp))
> ktmp
      N Mean  SD 患者 家族
患者 50 2.74 1.12 1.00 0.22
家族 50 3.18 1.02 0.22 1.00
>
>
```

```
> # 度数分布表
> t1 <- table(d1$患者)
> p1 <- prop.table(t1)*100
> rbind(t1, p1)
   1  2  3  4  5
t1  7 16 12 13  2
p1 14 32 24 26  4
> t1p <- t1
>
> t1 <- table(d1$家族)
> p1 <- prop.table(t1)*100
> rbind(t1, p1)
   1  2  3  4  5
t1  1 13 18 12  6
p1  2 26 36 24 12
> t1f <- t1
>
```

```
> # ウィルコクソンの符号順位和検定
> wilcox.test(d1$患者, d1$家族, paired=TRUE, exact=TRUE, conf.int=T)
```

Wilcoxon signed rank test with continuity correction

```
data: d1$患者 and d1$家族
V = 220.5, p-value = 0.02527
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -1.000032e+00 -4.890761e-05
sample estimates:
(pseudo)median
      -0.5
```

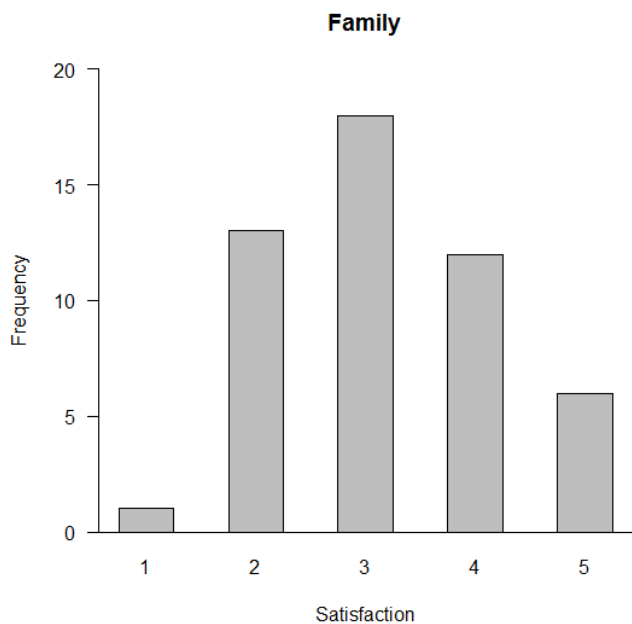
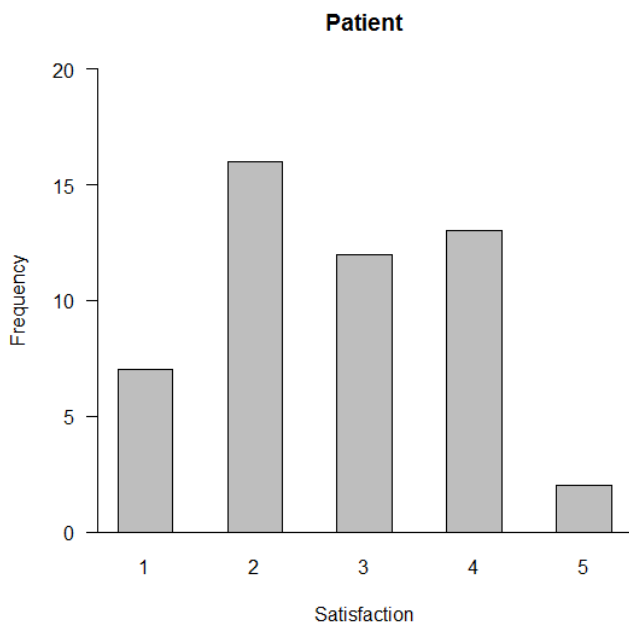
1	id	患者	家族
2	1	10	9
3	2	8	7
4	3	11	8
5	4	8	8
6	5	7	8
7	6	7	7
8	7	9	8
9	8	7	7
10	9	8	8
11	10	6	8
12	11	9	10
13	12	7	10
14	13	6	7
15	14	9	9
16	15	8	9
17	16	7	7
18	17	7	8
19	18	9	7
20	19	9	8
21	20	9	8

警告メッセージ:

```
1: In wilcox.test.default(d1$患者, d1$家族, paired = TRUE, exact = TRUE, :
   タイがあるため、正確な p 値を計算することができません
2: In wilcox.test.default(d1$患者, d1$家族, paired = TRUE, exact = TRUE, :
   タイがあるので、正確な信頼区間を計算することができません
3: In wilcox.test.default(d1$患者, d1$家族, paired = TRUE, exact = TRUE, :
   ゼロ値のため、正確な p 値を計算することができません
4: In wilcox.test.default(d1$患者, d1$家族, paired = TRUE, exact = TRUE, :
   ゼロ値のため、正確な信頼区間を計算することができません
>
>
```

> # 各群の分布

```
> barplot(t1p, ylim=c(0,20), width=0.7, space=1, xlab="Satisfaction",
+   ylab="Frequency", main="Patient", las=1)
> segments(0,0, 10,0)
>
> barplot(t1f, ylim=c(0,20), width=0.7, space=1, xlab="Satisfaction",
+   ylab="Frequency", main="Family", las=1)
> segments(0,0, 10,0)
>
```



>

対応のある多群の分布の位置の比較 — フリードマンの検定

```
friedman.test(as.matrix(データフレーム名))
```

as.matrix でデータ形式をmatrix形式にする必要がある。

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
> d1 <- read.table("対応のある多群の分布位置データ.csv", header=TRUE, sep=",")
> head(d1)
  番号 お使い下さい 迷惑 ありがとう
1    1             5    5          4
2    2             4    4          3
3    3             4    5          4
4    4             4    4          4
5    5             4    3          3
6    6             5    4          3
>
> dtmp <- d1[,c(-1)]
> d2 <- stack(dtmp)
>
>
```

```
> # 標本サイズ, 平均, SD, 相関係数
> ntmp <- nrow(dtmp)
> mtmp <- colMeans(dtmp)
> stmp <- apply(dtmp, 2, sd)
> ctmp <- cor(dtmp)
> ktmp <- round(data.frame(ntmp, mtmp, stmp, ctmp), 2)
> colnames(ktmp) <- c("N", "Mean", "SD", colnames(ctmp))
> ktmp
      N Mean   SD お使い下さい 迷惑 ありがとう
お使い下さい 129 5.03 1.14      1.00 0.24      0.31
迷惑          129 4.61 1.31      0.24 1.00      0.22
ありがとう    129 4.05 1.41      0.31 0.22      1.00
>
```

1	番号	お使い下さい	迷惑	ありがとう
2	1	5	5	4
3	2	4	4	3
4	3	4	5	4
5	4	4	4	4
6	5	4	3	3
7	6	5	4	3
8	7	6	5	5
9	8	7	7	5
10	9	5	5	2
11	10	5	6	4
12	11	5	6	4
13	12	3	2	6
14	13	6	6	4
15	14	6	5	3
16	15	5	5	2
17	16	5	4	5
18	17	7	4	6
19	18	5	6	5
20	19	6	5	4
21	20	5	5	4

```
> # 度数分布表
> t1 <- table(d1$お使い下さい)
> p1 <- prop.table(t1)*100
> round(rbind(t1, p1), 2)
      2    3    4    5    6    7
t1 1.00 12.0 25.00 49.00 28.00 14.00
p1 0.78  9.3 19.38 37.98 21.71 10.85
> t1c <- rbind(c(1:7), c(0, t1))
> colnames(t1c) <- c(1:7)
>
> t1 <- table(d1$迷惑)
> p1 <- prop.table(t1)*100
> round(rbind(t1, p1), 2)
      1    2    3    4    5    6    7
t1 1.00 5.00 21.00 32.00 36.00 25.00 9.00
p1 0.78 3.88 16.28 24.81 27.91 19.38 6.98
> t1i <- t1
>
> t1 <- table(d1$ありがとう)
> p1 <- prop.table(t1)*100
> round(rbind(t1, p1), 2)
      1    2    3    4    5    6    7
t1 1.00 15.00 35.00 34.00 20.0 17.00 7.00
p1 0.78 11.63 27.13 26.36 15.5 13.18 5.43
> t1t <- t1
>
```

カテゴリ 1 の人数が 0 人だが、グラフにこのカテゴリも
含めるための応急処置

> #フリードマンの検定

```
> friedman.test(as.matrix(dtmp), conf.int=T)
```

Friedman rank sum test

```
data: as.matrix(dtmp)
```

```
Friedman chi-squared = 49.14, df = 2, p-value = 2.135e-11
```

>

> #各群の分布

```
> barplot(t1c[,], ylim=c(0,50), xlim=c(0,10), width=0.7, space=1, xlab="",
+   ylab="Frequency", main="Use cleanly", las=1)
> segments(0,0, 10,0)
```

>

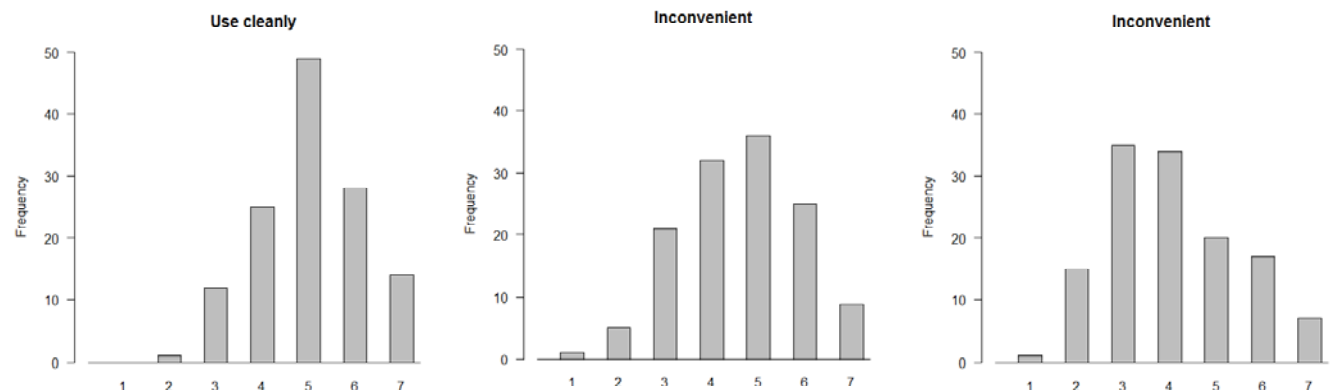
```
> barplot(t1i, ylim=c(0,50), xlim=c(0,10), width=0.7, space=1, xlab="",
+   ylab="Frequency", main="Inconvenient", las=1)
> segments(0,0, 10,0)
```

>

```
> barplot(t1t, ylim=c(0,50), xlim=c(0,10), width=0.7, space=1, xlab="",
+   ylab="Frequency", main="Inconvenient", las=1)
> segments(0,0, 10,0)
```

>

>



対応のない2群の分布の位置の比較 — ウィルコクソンの順位和検定(マン・ホイットニーの検定)

`wilcox.test`(中央値を比較したい変数 ~ 群分け変数, データフレーム名, `paired=FALSE`) または
`wilcox.test`(中央値を比較したい変数1, 中央値を比較したい変数2, `paired=FALSE`)

`paired=FALSE` は省略可.

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
> d1 <- read.table("対応のない2群の分布位置データ.csv", header=TRUE, sep=",")
> head(d1)
```

```
  番号 居住形態 バイト量
1     1     自宅         1
2     2     自宅         2
3     3     自宅         3
4     4     自宅         4
5     5     自宅         3
6     6     自宅         2
```

```
>
>
> # 度数分布表
> (t1 <- table(d1$居住形態, d1$バイト量))
```

```
      1  2  3  4  5
自宅   5 11 20 23 20
自宅外  4  3 14 25 20
```

```
>
> (p1 <- round(prop.table(t1, 1), 2))
```

```
      1    2    3    4    5
自宅  0.06 0.14 0.25 0.29 0.25
自宅外 0.06 0.05 0.21 0.38 0.30
```

```
>
>
>
```

```
> # 各群の記述統計量
> library(psych)
> describeBy(d1$バイト量, d1$居住形態)
group: 自宅
```

```
vars  n mean  sd median trimmed  mad min max range  skew kurtosis  se
1     1 79 3.53 1.2      4    3.62 1.48   1  5    4 -0.43   -0.77 0.13
```

```
-----
group: 自宅外
```

```
vars  n mean  sd median trimmed  mad min max range  skew kurtosis  se
1     1 66 3.82 1.11      4    3.96 1.48   1  5    4 -0.91    0.32 0.14
```

```
>
>
```

```
> #ウィルコクソンの順位和検定の実行
> wilcox.test(バイト量~居住形態, d1, paired=FALSE)
```

Wilcoxon rank sum test with continuity correction

```
data: バイト量 by 居住形態
W = 2241, p-value = 0.1321
alternative hypothesis: true location shift is not equal to 0
```

```
>
>
```

```
> #群ごとのデータフレームを作成
> d11 <- d1[d1$居住形態=="自宅",]
> d12 <- d1[d1$居住形態=="自宅外",]
>
```

1	番号	居住形態	バイト量
2	1	自宅	5
3	2	自宅外	2
4	3	自宅外	3
5	4	自宅	3
6	5	自宅外	4
7	6	自宅外	5
8	7	自宅外	1
9	8	自宅	4
10	9	自宅	3
11	10	自宅	4
12	11	自宅	3
13	12	自宅	5
14	13	自宅	3
15	14	自宅外	4
16	15	自宅外	4
17	16	自宅外	3
18	17	自宅	5
19	18	自宅外	4
20	19	自宅外	5
21	20	自宅外	4

```
> #ウィルコクソンの順位和検定の実行
> wilcox.test(d11$バイト量, d12$バイト量, paired=FALSE)
```

Wilcoxon rank sum test with continuity correction

data: d11\$バイト量 and d12\$バイト量

W = 2241, p-value = 0.1321

alternative hypothesis: true location shift is not equal to 0

```
> # 各群の分布
```

```
> d1.v <- d1[d1$居住形態=="自宅",]
```

```
> d1.n <- d1[d1$居住形態=="自宅外",]
```

```
>
```

```
> t1h <- table(d1.v$バイト量)
```

```
> t1o <- table(d1.n$バイト量)
```

```
>
```

```
> barplot(t1h, ylim=c(0, 25), xlim=c(0, 7), width=0.7, space=1, xlab="Amount",
+ ylab="Frequency", main="Home", las=1)
```

```
> segments(0, 0, 10, 0)
```

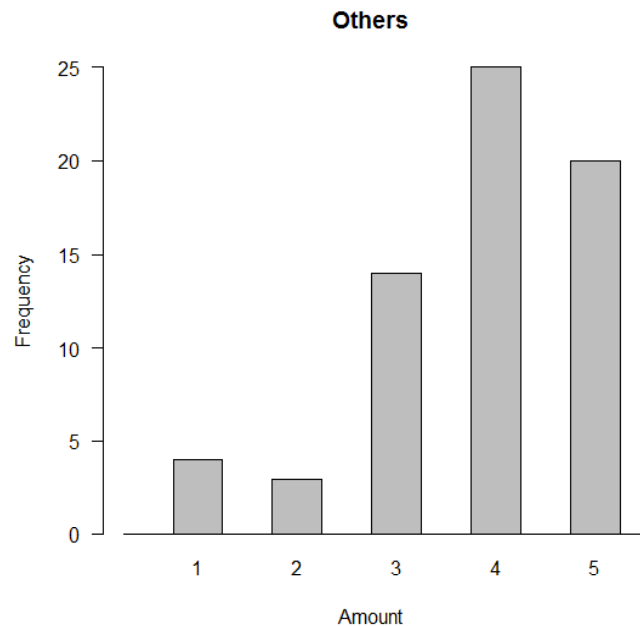
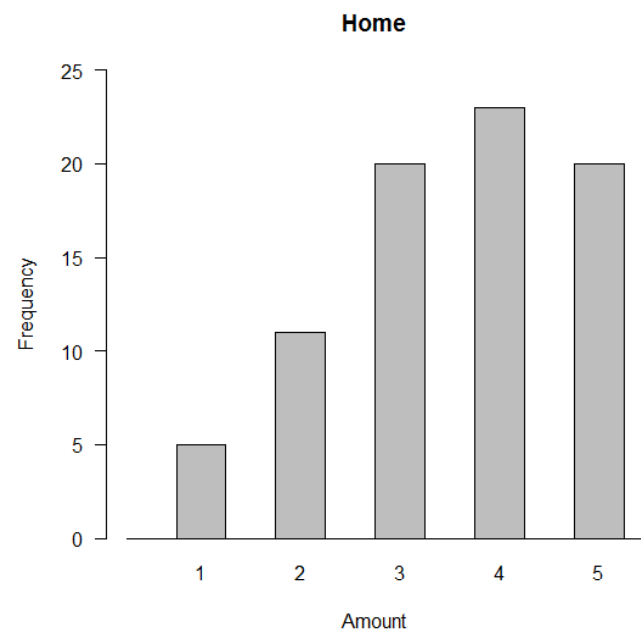
```
>
```

```
> barplot(t1o, ylim=c(0, 25), xlim=c(0, 7), width=0.7, space=1, xlab="Amount",
+ ylab="Frequency", main="Others", las=1)
```

```
> segments(0, 0, 10, 0)
```

```
>
```

```
>
```



```
>
```

```
>
```

```
>
```

対応のない多群の分布の位置の比較 — クラスカル・ウォリスの検定

`kruskal.test`(中央値を比較したい変数 ~ 群分け変数, データフレーム名) または
`kruskal.test`(データフレーム名\$中央値を比較したい変数, データフレーム名\$群分け変数)

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
> d1 <- read.table("対応のない多群の分布位置データ.csv", header=TRUE, sep=",")
> head(d1)
  番号 区分 嫌悪感
1    1 喫煙      1
2    2 禁煙      2
3    3 卒煙      5
4    4 喫煙      3
5    5 喫煙      1
6    6 喫煙      5
```

1	番号	区分	嫌悪感
2	1	喫煙	1
3	2	禁煙	2
4	3	卒煙	5
5	4	喫煙	3
6	5	喫煙	1
7	6	喫煙	5
8	7	禁煙	4
9	8	喫煙	2
10	9	禁煙	1
11	10	卒煙	1
12	11	卒煙	3
13	12	喫煙	2
14	13	卒煙	2
15	14	卒煙	2
16	15	喫煙	3
17	16	禁煙	2
18	17	禁煙	4
19	18	卒煙	4
20	19	卒煙	4
21	20	卒煙	4

```
> # 度数分布表
> (t1 <- table(d1$区分, d1$嫌悪感))
```

```
      1  2  3  4  5
喫煙  7 10  7  3  2
禁煙  6  6  4  5  7
卒煙  4  5  8  9  7
```

```
>
> (p1 <- round(prop.table(t1, 1), 2))
```

```
      1    2    3    4    5
喫煙 0.24 0.34 0.24 0.10 0.07
禁煙 0.21 0.21 0.14 0.18 0.25
卒煙 0.12 0.15 0.24 0.27 0.21
```

```
> # 各群の記述統計量
```

```
> library(psych)
```

```
> describeBy(d1$嫌悪感, d1$区分)
```

```
group: 喫煙
  vars  n mean   sd median trimmed  mad min max range skew kurtosis   se
1    1 29 2.41 1.18      2    2.32 1.48   1  5    4 0.58   -0.56 0.22
```

```
group: 禁煙
  vars  n mean   sd median trimmed  mad min max range skew kurtosis   se
1    1 28 3.04 1.53      3    3.04 1.48   1  5    4  0   -1.57 0.29
```

```
group: 卒煙
  vars  n mean   sd median trimmed  mad min max range skew kurtosis   se
1    1 33  3.3 1.31      3    3.37 1.48   1  5    4 -0.31  -1.08 0.23
```

```
> # クラスカル・ウォリスの検定
```

```
> kruskal.test(嫌悪感 ~ 区分, d1)
```

```
Kruskal-Wallis rank sum test
```

```
data: 嫌悪感 by 区分
```

```
Kruskal-Wallis chi-squared = 6.6299, df = 2, p-value = 0.03634
```

```
> # 各群の分布
```

```
> d1k <- d1[d1$区分=="喫煙",]
```

```
> d1s <- d1[d1$区分=="卒煙",]
```

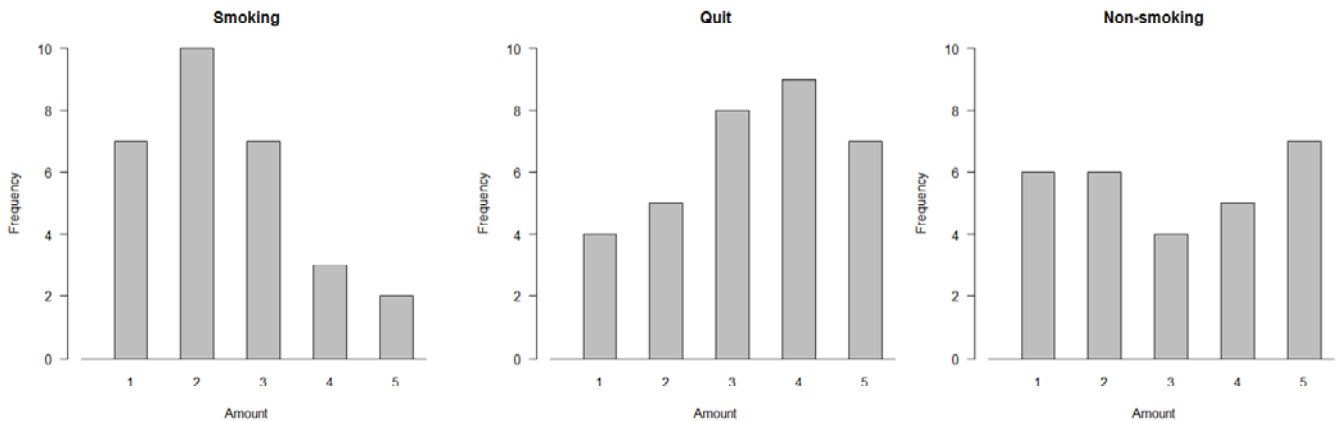
```
> d1n <- d1[d1$区分=="禁煙",]
```

```
>
```

```

> t1s <- table(d1k$嫌悪感)
> t1q <- table(d1s$嫌悪感)
> t1n <- table(d1n$嫌悪感)
>
> barplot(t1s, ylim=c(0,10), xlim=c(0,7), width=0.7, space=1, xlab="Amount",
+ ylab="Frequency", main="Smoking", las=1)
> segments(0,0, 10,0)
>
> barplot(t1q, ylim=c(0,10), xlim=c(0,7), width=0.7, space=1, xlab="Amount",
+ ylab="Frequency", main="Quit", las=1)
> segments(0,0, 10,0)
>
> barplot(t1n, ylim=c(0,10), xlim=c(0,7), width=0.7, space=1, xlab="Amount",
+ ylab="Frequency", main="Non-smoking", las=1)
> segments(0,0, 10,0)
>

```



```

>
> #3つのグラフを縦横に並べて描く設定
> par(mar=c(5,4,2,1))
> par(mfrow=c(3,1))
> hist(d1k$嫌悪感, breaks=seq(0.5,5.5,1), ylim=c(0,10), ylab="人数", main="喫煙群", xlab="")
> hist(d1s$嫌悪感, breaks=seq(0.5,5.5,1), ylim=c(0,10), ylab="人数", main="卒煙群", xlab="")
> hist(d1n$嫌悪感, breaks=seq(0.5,5.5,1), ylim=c(0,10), ylab="人数", main="禁煙群", xlab="")
> par(mfrow=c(1,1))

```

