

```

1   ###第5回クロス集計とグラフ化###
2   ##グラフ化はモザイクプロットとJitter(点プロットグラフ)##
3   install.packages("vcd")
4   install.packages("zoo")
5   install.packages("descr")
6   install.packages("gmodels")
7   dat <- read.csv("data/cars93.csv",
8                 row.names = 1 ,
9                 header = T )
10  head(dat , 3)
11  ###Cross Table###
12  names(dat)                #項目名確認
13  table.s <- table(dat$Origin, #table()関数,引数(表側,表頭)
14                dat$AirBags)
15  table.s
16  addmargins(table.s)      #合計を追加
17  library(vcd)
18  assocstats(table.s)     #assocstats()関数:連関係数
19  #library(gmodels)
20  library(descr)
21  library(vcd)
22  table.c <- CrossTable(dat$Origin , #CrossTable()関数,表頭
23                    dat$AirBags , #表側
24                    prop.r = F, #列%
25                    prop.c = T, #行%
26                    prop.t = T, #表%
27                    prop.chisq = F, #χ2統計量への寄与
28                    #各セルのχ2統計量への寄与:(観測度数-期待度数)^2/期待度数
29                    chisq = T ) #χ2検定
30  table.c
31  mosaicplot(~ AirBags + Origin, #mosaicplot()関数:~表頭+表側
32            data = dat, #data名
33            cex.axis = 1.1, #Category文字サイズ
34            xlab="Air Bag",
35            ylab="Origin",
36            main = "Mosaic Plot", #タイトル
37            shad=T) #shad=T:χ2値の標準化残差
38                    #残差2以下白,2~4:負の値赤,正の値ブルー
39  mosaic(~ DriveTrain + Origin + #mosaicplot()関数:3元クロス表の場合
40         Man.trans.avail,
41         data = dat,
42         main = "Mosaic Plot",
43         shade = T)
44  table.3 <- xtabs(~ DriveTrain + Origin+ #多元クロス表の作成
45                Man.trans.avail, #次元数をあげる場合+で追加
46                data = dat)
47  pairs(table.3, #pairst()関数:モザイクPマトリックス
48        main = "Mosaic Plot",
49        shad=T, labels) #labels:下三角に
50
51  assoc(table.3,shade=T) #assoc()関数:associationグラフ,矩形の高さは残差,
52  pairs(table.3, #幅は期待度数の平方根,面積は観測度数と期待度数の差
53        upper_panel = pairs_assoc, #上側三角:associationグラフ
54        shade=T)
55
56  ##オッズ比: =(6*22)/(39*26)
57  library(vcd)
58  table.3 <- table(dat$Origin,
59                dat$Man.trans.avail)
60  table.3

```

```
61  od <- oddsratio(table.3,          #oddsratio()関数:loddsratio()も同じ
62                        log=F)      #log=T:対数オッズ
63  od                                #オッズ比を返す
64  confint(od)                       #オッズ比の95%信頼区間
65  summary(od)                       #Z検定
66
67  #χ2検定とオッズ比(上記table.3を解析)
68  x <- matrix(c(6,26,39,22),      #matrix作成
69                nrow=2, ncol=2)
70  x
71  chisq.test(x)                     #χ2検定
72
73  x.lod <- loddsratio(x, log=F)     #loddsratio()関数
74  x.lod                             #オッズ比を返す
75  confint(x.lod)                   #オッズ比の95%信頼区間
76  summary(x.lod)                  #Z検定
77
78  ###Jitter(点プロットグラフ)を使うと視覚的にクロス表を表現できる:次回説明###
79  ##library(psych)に格納されているbfiデータを使う
80  #因子に対応した質問項目A1-A5, C1-C5, E1-E5, N1-N5, O1-O5, 性, 教育歴, 年齢
81  library(psych)
82  dat <- bfi
83  head(dat)
84  #dataは6段階の評定尺度, 散布図は使えない(クロス集計表でみる)
85  table.s <- table(dat$A1,         #table()関数, 引数(表側, 表頭)
86                  dat$A2)
87  table.s                            #このクロス表をJitterで:dataの方向要注意
88  plot(jitter(bfi$A1,2),          #table.sをjitterで表現
89        jitter(bfi$A2,2))
90  mosaicplot(~dat$A1+dat$A2)      #mosaicplot関数で:dataの方向要注意
91  plot(table.s)                   #plot関数で:dataの方向要注意
92
```