quasi-ind

#Quasi-indenpendence

```
datagss <- data.frame(
  Now = c("NortheastN", "MidwestN", "SouthN", "WestN",
          "NortheastN", "MidwestN", "SouthN", "WestN",
          "NortheastN", "MidwestN", "SouthN", "WestN",
          "NortheastN", "MidwestN", "SouthN", "WestN"),
  Age16 =rep( c("Northeast16", "Midwest16", "South16", "West16"),each=4),
  Count = c(394,17,81,38,8,596,74,59,29,32,769,35,10,24,35,417)
)
fit1<-glm(Count ~ Now + Age16,family=poisson,data=datagss)
summary(fit1)
```

```
Call:
glm(formula = Count ~ Now + Age16, family = poisson, data = datagss)

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)       5.23821    0.04970 105.407  < 2e-16 ***
NowNortheastN    -0.41674    0.06133  -6.795 1.09e-11 ***
NowSouthN         0.36011    0.05037   7.149 8.76e-13 ***
NowWestN         -0.19769    0.05759  -3.433 0.000597 ***
Age16Northeast16 -0.32971    0.05695  -5.789 7.07e-09 ***
Age16South16      0.16014    0.05013   3.195 0.001400 **
Age16West16      -0.41638    0.05843  -7.126 1.03e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)
```

```
    Null deviance: 4236.9  on 15  degrees of freedom
Residual deviance: 3871.7  on  9  degrees of freedom
AIC: 3981

Number of Fisher Scoring iterations: 6
```

```r
datagss_quasi <- data.frame(
  Now = c("NortheastN", "MidwestN", "SouthN", "WestN",
          "NortheastN", "MidwestN", "SouthN", "WestN",
          "NortheastN", "MidwestN", "SouthN", "WestN",
          "NortheastN", "MidwestN", "SouthN", "WestN"),
  Age16 =rep( c("Northeast16", "Midwest16", "South16", "West16"),each=4),
  Count = c( 394,17,81,38,8,596,74,59,29,32,769,35,10,24,35,417),
  diag_index = c(1,0,0,0,0,2,0,0,0,0,3,0,0,0,0,4) )
fit2 <- glm(Count ~ factor(Now) + factor(Age16) + factor(diag_index),family=poisson, data=dat
summary(fit2)
```

```
Call:
glm(formula = Count ~ factor(Now) + factor(Age16) + factor(diag_index),
    family = poisson, data = datagss_quasi)

Coefficients:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                3.38251    0.16019  21.115  < 2e-16 ***
factor(Now)NortheastN     -0.48251    0.19270  -2.504  0.01228 *
factor(Now)SouthN          0.97109    0.15062   6.447 1.14e-10 ***
factor(Now)WestN           0.42567    0.15172   2.806  0.00502 **
factor(Age16)Northeast16 -0.11300    0.12406  -0.911  0.36237
factor(Age16)South16       0.03512    0.14716   0.239  0.81137
factor(Age16)West16       -0.59721    0.15254  -3.915 9.04e-05 ***
factor(diag_index)1        3.18935    0.18796  16.968  < 2e-16 ***
factor(diag_index)2        3.00773    0.16535  18.191  < 2e-16 ***
factor(diag_index)3        2.25637    0.15333  14.716  < 2e-16 ***
factor(diag_index)4        2.82212    0.17038  16.563  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 4236.927  on 15  degrees of freedom
Residual deviance:   29.803  on  5  degrees of freedom
```

```
AIC: 147.09

Number of Fisher Scoring iterations: 4
```

```r
datagss2 <- matrix(c(
  394, 8, 29, 10,
  17, 596, 32, 24,
  81, 74, 769, 35,
  38, 59, 35, 417
), nrow = 4, byrow = F)
row_sums <- rowSums(datagss2)
col_sums <- colSums(datagss2)
total_sum <- sum(datagss2)
datagss_with_totals <- cbind(datagss2, row_sums)
datagss_with_totals <- rbind(datagss_with_totals, c(col_sums, total_sum))

rownames(datagss_with_totals) <- c("Northeast16", "Midwest16", "South16", "West16", "Column T
colnames(datagss_with_totals) <- c("NortheastN", "MidwestN", "SouthN", "WestN", "Row Total")
datagss_with_totals
```

```
             NortheastN MidwestN SouthN WestN Row Total
Northeast16         394       17     81    38       530
Midwest16             8      596     74    59       737
South16              29       32    769    35       865
West16               10       24     35   417       486
Column Total        441      669    959   549      2618
```

#bradley-Terry

```r
datagss2 <- matrix(c(
  394, 8, 29, 10,
  17, 596, 32, 24,
  81, 74, 769, 35,
  38, 59, 35, 417
), nrow = 4, byrow = F)
row_sums <- rowSums(datagss2)
col_sums <- colSums(datagss2)
total_sum <- sum(datagss2)
datagss_with_totals <- cbind(datagss2, row_sums)
datagss_with_totals <- rbind(datagss_with_totals, c(col_sums, total_sum))
```

```r
rownames(datagss_with_totals) <- c("Northeast16", "Midwest16", "South16", "West16", "Column T
colnames(datagss_with_totals) <- c("NortheastN", "MidwestN", "SouthN", "WestN", "Row Total")
datagss_with_totals
```

```
             NortheastN MidwestN SouthN WestN Row Total
Northeast16         394       17     81    38       530
Midwest16             8      596     74    59       737
South16              29       32    769    35       865
West16               10       24     35   417       486
Column Total        441      669    959   549      2618
```

```r
library(psych)
dat<-matrix(datagss_quasi$Count,ncol=4,byrow = T)
cohen.kappa(dat)
```

```
Call: cohen.kappa1(x = x, w = w, n.obs = n.obs, alpha = alpha, levels = levels,
    w.exp = w.exp)

Cohen Kappa and Weighted Kappa correlation coefficients and confidence boundaries
                 lower estimate upper
unweighted kappa  0.75     0.77  0.79
weighted kappa    0.71     0.74  0.77

 Number of subjects = 2618
```