

# Homework 3

STAT 5333 (Spring 2021)

Md Abul Hayat

## Problem 2.22 (a)

Diagnosis	Drugs	No Drugs	Row Sum
Schizophrenia	105	8	113
Affective Disorder	12	2	14
Neurosis	18	19	37
Personality Disorder	47	52	99
Special Symptoms	0	13	13
Column Sum	182	94	276

```
# General Code
n_table= matrix(c(105,8,12,2,18,19,47,52,0,13), nrow=5, byrow=TRUE)
n_rows = dim(n_table)[1]
n_cols = dim(n_table)[2]
```

Finding  $\{\hat{\mu}_{i,j}\}$

```
mu_table = matrix(rep(0,n_rows*n_cols), nrow=n_rows, byrow=TRUE)
for (i in 1:n_rows){
  for (j in 1:n_cols){
    mu_table[i,j]=(sum(n_table[i,])*sum(n_table[,j]))/sum(n_table)
  }
}
(mu_table)
```

```
##           [,1]      [,2]
## [1,] 74.514493 38.485507
## [2,]  9.231884  4.768116
## [3,] 24.398551 12.601449
## [4,] 65.282609 33.717391
## [5,]  8.572464  4.427536
```

Pearsons  $\chi^2_{I-1,J-1}$  test statistic for  $(I \times J)$  table

$$\chi^2_{I-1,J-1} = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}}$$

```
chi2_stat = sum((n_table-mu_table)^2/mu_table)
(chi2_stat)
```

```
## [1] 84.18847
```

p-value of test statistic

```
df = (n_rows-1)*(n_cols-1)
p_value = 1-pchisq(chi2_stat,df)
(p_value)
```

```
## [1] 0
```

There is significant evidence against  $H_0$  hypothesis. Hence, drugs are *not* being prescribed independent of diseases. The exact same results can be obtained using built-in function of R

```
(builtin_func = chisq.test(n_table))
```

```
## Warning in chisq.test(n_table): Chi-squared approximation may be incorrect
```

```
##
## Pearson's Chi-squared test
##
## data:  n_table
## X-squared = 84.188, df = 4, p-value < 2.2e-16
```

```
builtin_func$expected
```

```
##           [,1]      [,2]
## [1,] 74.514493 38.485507
## [2,]  9.231884  4.768116
## [3,] 24.398551 12.601449
## [4,] 65.282609 33.717391
## [5,]  8.572464  4.427536
```

## Problem 2.30

Treatment	Cancer Controlled	Cancer Not Controlled	Row Sum
Surgery	21	2	23
Radiation Therapy	15	3	18
Column Sum	36	5	41

Here,  $H_0: \theta = 1$  and  $H_\alpha: \theta > 1$

```
(p_value = 1-phyper(20,23,18,36))
```

```
## [1] 0.3808337
```

As  $p\text{-value} = 0.3808337 > 0.05$ , we do not have significant evidence against  $H_0$ . Hence, controlling cancer is independent of treatment. The same results are obtained using built-in function

```
data = matrix(c(21,2,15,3),2,2,byrow=TRUE)
fisher.test(data,alternative = "greater")
```

```
##
## Fisher's Exact Test for Count Data
##
## data:  data
## p-value = 0.3808
## alternative hypothesis: true odds ratio is greater than 1
## 95 percent confidence interval:
##  0.2864828      Inf
## sample estimates:
## odds ratio
##  2.061731
```