Question 1

Introduction:

Nonsteroidal anti-inflammatory (NSAID) is a category of drugs taken by children and adults in order to relieve muscle pain and reduce fever. This medicine is ubiquitous, and has been regarded safe to take for a short period of time. However, the side effects of NSAIDs include stomach ulcers, kidney disease, and liver disease. In 1996, the American College of Rheumatology issued a series of guidelines recommending baseline tests of a Complete Blood Cell Count, Liver Function series, and Renal series. These tests can be performed with a simple blood draw, and results can be reported within a week. It was of interest to test whether or not the physicians who order these baseline tests when prescribing NSAIDS is the same across all disciplines.

Methods:

A total of 13,009 physicians who took part in the national database participated in this study. We report the frequency of physicians who stated whether or not they ordered the baseline tests (Yes or No) along with the total number of physicians and the proportion stating "Yes" for each discipline of practice (internal medicine, family practice, academic family practice, and multispecialty groups) for those physicians who ordered the recommended tests. A test of homogeneity on the "Yes" proportions for each discipline is conducted using a chi-square test with 3 degrees of freedom. The null hypothesis ($Ho : Py_{IM}=Py_{FM}=Py_{AFM}=Py_{MS}$) is that the ordered baseline tests proportions for each discipline are equal, while the alternative hypothesis ($Ha : P_{yi} \neq P_{yj}$) is that at least two of those proportions differ. We will reject the null hypothesis if the resulting p-value is less than the significance level $\alpha = 0.05$, and we will fail to reject the null

hypothesis otherwise. The R statistical software (version 2.11.1) was used for all statistical analyses.

Results:

The data are summarized in Table 1.1 below for the 13,009 physicians included in this study. Assuming that the data are representative and subjects are independent, the sample size is large enough to conduct statistical analysis since all expected cell frequencies are greater than 5. The test produced the following results (χ_3 2 = 816.41, p-value < 2.2e-16), since our p-value is less than alpha we reject the null hypothesis in favor of the alternative and claim that there is a significant difference between at least two of the "Yes" proportions. By comparing the observed with expected cell frequencies, we report that the observed "Yes" are greater for Internal Medicine then all other disciplines, therefore the proportion of "Yes" is not equal among the disciplines.

Practice Type	Yes	No	Total	Proportion "Yes"
Internal Medicine	294	921	1,215	0.24
Family Practice	98	2,862	5,825	0.016
Academic Family Practice	50	3,064	3,114	0.016
Multi Speciality Practice	203	2,652	2,855	0.071
Total	645	9,499	13,009	

Discussion:

The proportion of physicians who ordered baseline tests increased in Internal Medicine and Multispecialty. Patients in these two disciplines tend to have co-mobilities and it appears that physicians are taking more precautions when ordering NSAIDs for this group of patients. Clinic directors can use this information in order to ensure that physicians across all specialties are adhering the recommended guidelines before prescribing medicine.

Table 1.1 Contingency Table for Physicians in Four Disciplines and Whether or Not They Ordered the Recommended Baseline Tests When Prescribing NSAIDS.

Question 2

Introduction:

Cigarette smoking has long been regarded as an epidemic among both teens and adults. There was a time in which smoking was designated for all areas, including operating rooms. The amount of restrictions on smoking is likely tied to the highest education level. We test the hypothesis that the education level and smoking policy favored are related to each other.

Methods:

A total of 300 adults participated in this study. The frequency and proportion of cigarette policy favored with their highest education level (College Graduate, High School Graduate and Grade School Graduate) as well as the total number of adults are reported for each level of education. A test of association between highest education level cigarette policies favored is conducted using a chi-square test with 6 degrees of freedom. The null hypothesis is that there is highest education level is unrelated to cigarette policy favored, while the alternative hypothesis is that the cigarette policy favored depends on highest education level. We will reject the null hypothesis if the resulting p-value is less than the significance level 0.05, and we will fail to reject the null hypothesis otherwise. The R statistical software (Version 2.11.1) was used for all statistical analyses.

Results:

The data are summarized in Table 2.1 below for the 300 patients included in this study. Assuming that the data are representative and subjects are independent, the sample size is large enough to conduct statistical analysis as fewer than 20 % of all cells have expected frequencies less than 5 (1／12 = 0.083). The test produced the following results (χ_6 2 = 22.502, p-value = 0.00098), since the p-value is less than alpha we reject the null hypothesis in favor of the alternative and claim that there is a relationship highest education level and cigarette

policies favored. Comparing the observed and expected cell frequencies, we see that participants with grade school as their highest education more often have no opinion on the matter versus those whose are college educated. Interestly, those who have college as their highest education level are less in favor of no restrictions in smoking than those who are a grade school graduate.

Table 2.1 Contingency Table of Highest Education Level and Policy Favored

Highest Education Level	No Restrictions on Smoking	Smoking Allowed in Designated Areas Only	No Smoking at All	No Opinion	Total
College	5 6.6%	44 58.6%	23 30.6%	3 4%	75
High School Graduate	15 10.0%	100 66.6%	30 20.0%	5 3.3%	150
Grade School Graduate	15 20.0%	40 53.3%	10 13.3%	10 13.3%	75
Total	35	184	63	18	300

Discussion:

In general, those with a college degree and a high school diploma favored highest in smoking allowed in designated areas only when compared to grade school graduates. In addition, participants with a grade school graduate favored no smoking restrictions when compared to college graduates.

R-Code

```
#homework 4
#question 1
Table1<-matrix(c(294,98,50,203,921,2862,3064,2652),nrow = 4,ncol=2)
Table1
#[,1] [,2]
#[1,]  294  921
#[2,]   98 2862
#[3,]   50 3064
#[4,]  203 2652
prop.table(Table1,1)
#[,1]      [,2]
#[1,] 0.24197531 0.7580247
#[2,] 0.03310811 0.9668919
#[3,] 0.01605652 0.9839435
#[4,] 0.07110333 0.9288967
expval1<-chisq.test(Table1)
expval1
#Pearson's Chi-squared test
#
#data:  Table1
#X-squared = 816.41, df = 3, p-value < 2.2e-16
expval1$expected
```

```
expval1$expected

#[,1]     [,2]

#[1,]  77.25503 1137.745

#[2,] 188.20978 2771.790

#[3,] 198.00177 2915.998

#[4,] 181.53342 2673.467

################################################################################

##############

#Question 2

Table2<-matrix(c(5,15,15,44,100,40,23,30,10,3,5,10),nrow = 3,ncol=4)

Table2

#[,1] [,2] [,3] [,4]

#[1,]    5   44   23    3

#[2,]   15  100   30    5

#[3,]   15   40   10   10

expval2<-chisq.test(Table2)

expval2

#Warning message:

#In chisq.test(Table2) : Chi-squared approximation may be incorrect

#waring because expected values is less than 5.

#Pearson's Chi-squared test

#

#data:  Table2

#X-squared = 22.502, df = 6, p-value = 0.0009817
```

```
expval2$expected

#[,1] [,2]  [,3] [,4]

#[1,]  8.75   46 15.75  4.5

#[2,] 17.50   92 31.50  9.0

#[3,]  8.75   46 15.75  4.5

prop.table(Table2,1)

#> > prop.table(Table2,1)

#[,1]      [,2]      [,3]       [,4]

#[1,] 0.06666667 0.5866667 0.3066667 0.04000000

#[2,] 0.10000000 0.6666667 0.2000000 0.03333333

#[3,] 0.20000000 0.5333333 0.1333333 0.13333333
```