Question 1

Introduction:

Dialysis is a treatment method which acts as a blood filtration alternative when the kidneys do not function correctly. There are two types of dialysis for patients: Hemodialysis (blood filtration via the vessels in your arm) and Peritoneal Dialysis (blood filtration via the vessels in your kidneys). Diabetes is a widespread disease that affects blood glucose intake by the cells. It was interest to determine whether or not there was a significant difference in patients who are affected by diabetes and who receive peritoneal dialysis versus patients with diabetes and receive hemodialysis.

Methods:

The frequency of patients who reported they are currently on dialysis treatment is reported for both the hemodialysis (n=529) and peritoneal dialysis (n=326) groups, and the proportion of those subjects reporting diabetes in both groups are summarized with sample proportions and a *95%* confidence intervals. The difference in sample proportions is also presented, as a 95% confidence intervals on the difference between the two groups. We test the null hypothesis of difference between these two groups (Ho = p1 - p2 = 0)against a two sided alternative hypothesis that the difference in diabetes in proportion of patients who receive dialysis is not equal to zero (Ho = p1-p2 ≠ 0) by using a two sided z-test with a significance level a =0.05. We will reject the null hypothesis in favor of the alternative hypothesis if the *p-value* is less than a; otherwise we will not reject the null hypothesis. The R statistical software (version 2.11.1) was used for all statistical analyses.

Results:

The data are summarized in Table 1 below. Assuming that the two samples are representative, subjects are independent, the two samples are large enough to conduct the statistical analysis. The observed proportion of patients with diabetes and using hemodialysis (0.471, 95% CI: 0.428, 0.513) is greater than observed proportion of patients with diabetes and using peritoneal dialysis (0.411, 95% CI: 0.358, 0.464) with an observed difference in proportions of 0.0597. The two-sided z-test yielded p-value =0.0884. The p-value is larger than alpha,thus we fail to reject the null hypothesis.

Table 1. Contingency Table for Diabetes observed in patients receiving Dialysis treatment

Treatment group	Patients with Diabetes	Sample Size	Observed	95% confidence Interval
Hemodialysis	249	529	0.4701	0.428,0.513
Peritoneal Dialysis	134	326	0.4110	0.358,0.464
Total	383	855	0.060	-0.009, 0.12

Discussion:

The sample data suggests that the proportion of patients with diabetes and who use hemodialysis is greater than those patients who use peritoneal dialysis. Thus clinicians, and practitioners interested in treating patients with diabetes should take into account which dialysis treatment the patient is currently on.

Question 2

Introduction:

Obesity is an epidemic known to inflict individuals who live in westernized nations. Obesity is diagnosed when the individual's Body Mass Index is above 30. Due to the differences in fat deposits and the differences in heights among men, and women we wanted to take a closer look at obesity and its prevalence among the genders between ages of 20 and 75. It was interest to determine whether or not there was a significant difference in men who are affected by obesity and women who are affected by obesity.

Methods:

The frequency of patients who reported they are overweight is reported for both the genders, male (*n*=150) and female (*n*=200) groups, and the proportion of those subjects reporting diabetes in both groups are summarized with sample proportions and a *95%* confidence intervals. The difference in sample proportions is also presented, as a 95% confidence intervals on the difference between the two groups. We test the null hypothesis of difference between these two groups (Ho = p1 - p2 = 0)against a two sided alternative hypothesis that the difference in males who are overweight and females who are overweight is not equal to zero (Ho = p1-p2 ≠ 0) by using a two sided z-test with a significance level *a* =0.05. We will reject the null hypothesis in favor of the alternative hypothesis if the *p-value* is less than a; otherwise we will not reject the null hypothesis. The R statistical software (version 2.11.1) was used for all statistical analyses.

Results:

The data are summarized in Table 1 below. Assuming that the two samples are representative, subjects are independent, the two samples are large enough to conduct the statistical analysis. The observed proportion of patients who are male and overweight (0.140, 95% CI: 0.084, 0.196) is less than observed proportion of patients who are female and overweight (0.240, 95% CI: 0.181, 0.299). with an observed difference in proportions of -0.10. The two-sided z-test yielded p-value = 0.01996, the p-value is lower than alpha, thus we fail to reject the null hypothesis.

Table 1. Contingency Table for Overweight patients reported in both genders

Treatment group	Overweight	Sample Size	Observed	95% confidence interval
Male	21	150	0.14	0.0844, 0.1955
Female	48	200	0.24	0.1803, 0.2991
Total	69	350	-0.1	-0.0188, -0.1811

Discussion:

The sample data suggests that the proportion of patients who are male and reported as being overweight is less than the proportion of patients who are female and reported being overweight. Thus clinicians, and practitioners interested in treating obesity and overweight patients should individualize their approaches based on the gender of the patient.

Appendix: R code

Question 1.

```
Table1<-matrix(c(249, 134, 529, 326), nrow=2,ncol=2)

Table1

expval1<-chisq.test(Table1)

expval1$expected
```

Output

```
> Table1<-matrix(c(249, 134, 529, 326), nrow=2,ncol=2)

> Table1

     [,1] [,2]

[1,]  249  529

[2,]  134  326

> expval1<-chisq.test(Table1)

> expval1$expected

         [,1]     [,2]

[1,] 240.6898 537.3102

[2,] 142.3102 317.6898
```

//

```
prop.test(c(249,134),c(529,326),alternative="two.sided",correct=FALSE)
```

2-sample test for equality of proportions without continuity

correction

Output:

```
data:  c(249, 134) out of c(529, 326)
```

X-squared = 2.9028, df = 1, p-value = 0.08843

alternative hypothesis: two.sided

95 percent confidence interval:

 -0.008621297 0.127934273 <-CI of the difference

sample estimates:

   prop 1   prop 2

0.4706994 0.4110429

//

Confidence Interval for p-hat 1

```
#Finding the CI for Sample1
phat<-0.4701
CV<-1.96
n<-529
se<-sqrt(phat*(1-phat)/n)
ci_1<-phat + (CV * se)
ci_2<-phat - (CV * se)
ci_1
ci_2
```

Output:

```
>ci_2
 [1] 0.5126324
> ci_2
[1] 0.4275676
```

//

Confidence interval for p-hat2

```
phat<-0.4110

CV<-1.96

n<-326

se<-sqrt(phat*(1-phat)/n)

ci_1<-phat + (CV * se)

ci_2<-phat - (CV * se)

ci_1

ci_2
```

Output:

```
> ci_1

[1] 0.4644104

> ci_2

[1] 0.3575896
```

Question 2.

```
Table1<-matrix(c(21, 48, 150, 200), nrow=2,ncol=2)

Table1

expval1<-chisq.test(Table1)

expval1$expected
```

Output:

```
> Table1<-matrix(c(21, 48, 150, 200), nrow=2,ncol=2)
```

```
> Table1
     [,1] [,2]
[1,]   21  150
[2,]   48  200
> expval1<-chisq.test(Table1)
> expval1$expected
        [,1]     [,2]
[1,] 28.1599 142.8401
[2,] 40.8401 207.1599
```

//

```
prop.test(c(21,48),c(150,200),alternative="two.sided",correct=FALSE)
2-sample test for equality of proportions without continuity
        correction

data:  c(21, 48) out of c(150, 200)
X-squared = 5.4154, df = 1, p-value = 0.01996
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.18115923 -0.01884077
sample estimates:
```

prop 1 prop 2

  0.14   0.24

//

CI for phat 1

```
phat<-0.14

CV<-1.96

n<-150

se<-sqrt(phat*(1-phat)/n)

ci_1<-phat + (CV * se)

ci_2<-phat - (CV * se)

ci_1

ci_2
```

Output:

```
> ci_1

[1] 0.1955295

> ci_2

[1] 0.08447051
```

//

CI for phat 2

```
phat<-0.24

CV<-1.96

n<-200

se<-sqrt(phat*(1-phat)/n)

ci_1<-phat + (CV * se)
```

ci_2<-phat - (CV * se)

ci_1

Ci_2

Output:

> ci_1

[1] 0.2991907

> ci_2

[1] 0.1808093