STA305/1004 Midterm Test Questions Preview

March 2, 2016

Name: ____________________________________

Student Number: __________________________

This document contains a preview of the scenarios for each of the four questions that will appear on the midterm test. The actual questions corresponding to the scenarios have been removed, but will appear on the actual test.

Instructions:

Time allowed: 90 minutes

Answer all four questions.

Complete all questions in pen. Any questions completed in pencil may not be eligible to be remarked even if there was a marking error.

Aids allowed: You are allowed to bring in one 8.5'x11' sheet with hand writing on both sides, and a calculator.

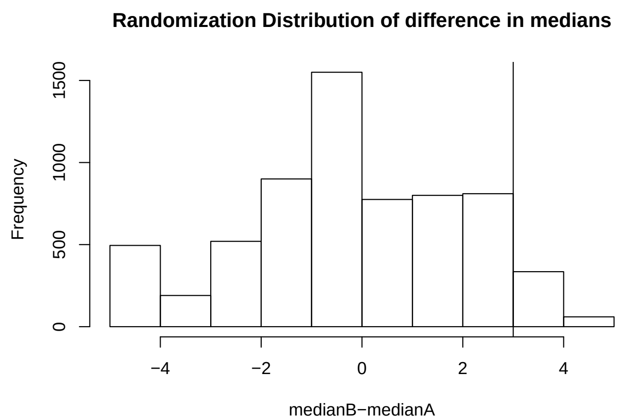
Question	Value	Mark
1	25	
2	20	
3	20	
4	20	
Total	85	

1. Fifteen judges were randomly allocated to judge one of two brands of beer, A or B, for taste. Eight judges will be assigned to Brand A and seven judges to brand B. The judges ranked the beer they tasted using a 10-point (Likert) scale with 1 representing "poor taste" and 10 representing "outstanding taste".
   The table below shows the rating from each judge. The number in brackets beside the rating indicates which judge gave the rating. For example, judge 1 gave a rating of 2 to brand A, and judge 9 gave a rating of 3 to brand B.

Brand A	2 (1)	4 (2)	2 (3)	1 (4)	9 (5)	9 (6)	2 (7)	2 (10)
Brand B	8 (8)	3 (9)	5 (11)	3 (12)	7 (13)	7 (14)	4 (15)	

   The randomization distribution and a one-sided p-value for testing $H_0 : \tilde{\mu}_A = \tilde{\mu}_b$ versus $H_1 : \tilde{\mu}_B > \tilde{\mu}_A$, where $\tilde{\mu}_A, \tilde{\mu}_B$ are the median taste rating for brands A and B respectively were calculated. Some of the R code is shown below.

```
yA <- c(2,4,2,1,9,9,2,2)
yB <- c(8,3,5,3,7,7,4)
beer <- c(yA,yB) #pool data
for (i in 1:N)
{
  res[i] <- median(beer[index[,i]])-median(beer[-index[,i]])
}
hist(res,xlab="medianB-medianA", main="Randomization Distribution of difference in medians")
observed <- median(yB)-median(yA) #store observed median difference
abline(v=observed) #add line at observed median diff
```




```
observed # Observed difference in medians
```

```
[1] 3
```

```
sum(res>=observed)/N
```

```
[1] 0.1002331
```

Answer the following questions based on this study. (5 marks each)

The questions have been removed.

2. A psychologist studying body language conducted an experiment on 20 subjects, and obtained a significant result from a two-sided z-test ($H_0 : \mu = 0$ vs. $H_1 : \mu \neq 0$). Let's call this experiment #1. The observed value of the z statistic from her experiment is $z = 2.5$ so the p-value=0.012. In order to confirm the results the psychologist is planning to run the same experiment on an additional 10 subjects (i.e., the same experiment will be done on 10 different subjects). Let's call this experiment #2.

   The following percentiles from the $N(0, 1)$ distribution might be required to carry out some of the calculations in the questions below.

α	z_α
0.450	-0.13
0.100	1.28
0.050	1.64
0.025	1.96
0.010	2.33

   z_α is the $100(1 - \alpha)^{th}$ percentile of the $N(0, 1)$. For example, the 90^{th} percentile is $z_{0.10} = 1.28$.

The questions have been removed.

3. What is the effect of smoking on weight gain? Data was used from a voluntary survey to assess this question. Smoking status (cessation/ no cessation) was recorded in 1971, and weight (kg), the outcome of interest, was recorded in 1982. The survey has baseline (1971) information on weight, sex, race, height, education, alcohol use, and intensity of smoking.

   The table below shows the distribution of baseline covariates in the two groups. 403 people are in the smoking group (No cessation T=0) and 1163 people are in the smoking cessation group (T=1).

	Cessation (T=1)	No cessation (T=0)
age, years (mean)	46.2	42.8
men, %	54.6	46.6
white, %	91.1	85.4
university, %	15.4	9.9
weight, kg (mean)	72.4	70.3
Cigarettes/day (mean)	18.6	21.2
year smoking (mean)	26.0	24.1
little/no exercise, %	40.7	37.9
inactive daily life, %	11.2	8.9

   The propensity score was estimated using a logistic regression model based on all 9 covariates; no interactions were included in the propensity score model. Three propensity score methods were used to estimate the average treatment effect (the average difference in 1982 weight between the smoking cessation group and the smoking group): propensity score matching; stratifying on the propensity score; and regression adjustment using the propensity score. Unadjusted means were compared using the two-sample t-test. The propensity score methods successfully balanced all the covariates (i.e., the absolute standardized differences are less than 10%).

   The table below shows the average treatment effect for each method with 95% confidence interval. The propensity score matching method was able to match the 403 subjects in the smoking group to 403 subjects in the smoking cessation group.

Method	Average Treatment Effect	95% Confidence Interval
Matched	2.93	1.8 - 4.0
Stratified	3.26	1.7 - 3.4
Regression	3.40	2.5 - 4.3
Unadjusted	2.54	1.7 - 3.4

Answer the following questions. (5 marks each)

The questions have been removed.

4. In order to determine the effect of diet on blood coagulation 18 animals were randomized to three diets: A, B, C. The table below gives coagulation times for blood samples drawn from 18 animals receiving three different diets A, B, and C.

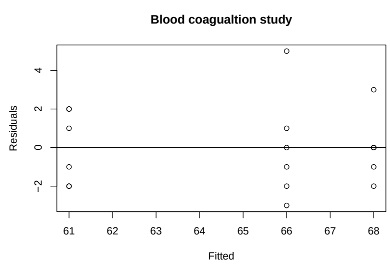
	A	B	C
	60	65	71
	63	66	66
	59	67	68
	63	63	68
	62	64	67
	59	71	68
Treatment Average	61	66	68
Grand Average	64	64	64
Difference	-3	2	4

The data was analyzed using R. Some of the output is shown below.

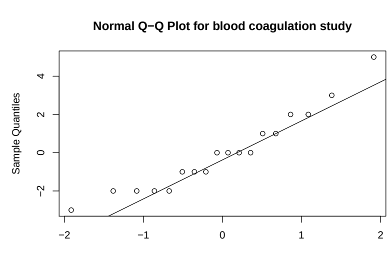
```
aov.diets <- aov(y~diets,data=dietdat)
summary(aov.diets)
```

```
            Df Sum Sq Mean Sq F value   Pr(>F)
diets          156
Residuals       72
---
```

```
plot(aov.diets$fitted.values,aov.diets$residuals,ylab="Residuals",
     xlab="Fitted",main="Blood coagulation study")
abline(h=0)
```




```
qqnorm(aov.diets$residuals, main="Normal Q-Q Plot for blood coagulation study")
qqline(aov.diets$residuals)
```



Answer the following questions based on the blood coagulation study.

The questions have been removed.