260762008 Tianze Lin **Math 204 Assignment 1**

8.122

a.

Hypotheses

H_0: $\sigma=0.1$, equivalent to $\sigma^2=0.01$

H_a: $\sigma\neq0.1$, equivalent to $\sigma^2\neq0.01$

Test Statistic

$W=(n-1)*s^2/\sigma_0^2=(6-1)*0.011^2/0.01=0.0605$

Rejection Region:

```
> qchisq(c(0.025,0.975),5)
[1]  0.8312116 12.8325020
```

0.0605<0.83

Conclusion: We reject the null hypothesis at the significance level of $\alpha=0.05$.

b.

In 7.111, the confidence interval is (0.0069,0.0270).

The σ_0 is 0.1, and it is not in the confidence interval, so we should reject the null hypothesis at the significance level of $\alpha=0.05$. This result agrees with the result in a.

8.124

Hypotheses:

H_0:$\sigma^2=9$

H_a:$\sigma^2<9$

Test Statistic:

$W=(n-1)*s^2/\sigma_0^2=55.18$

```
> n <- length(data$Content)
> s <- sd(data$Content)
> (n-1)*s^2/9
[1] 55.17574
```

Rejection Region:

```
> qchisq(c(0.01),49)
[1] 28.94065
```

55.18>28.94

Conclusion: We cannot reject the null hypothesis at the significance level of $\alpha=0.05$.

9.114

From Math 203, if we want to use t-test, we need to assume $\sigma_1=\sigma_2$.

Hypotheses

H_0: $\sigma_1^2/\sigma_2^2=1$

H_a: $\sigma_1^2/\sigma_2^2\neq1$

Test Statistic: $F=S_{large}^2/S_{small}^2=0.011^2/0.002^2=30.25$

Assume the significance level is 0.05.

Rejection Region:

```
> qf(0.975,5,5)
[1] 7.146382
```

30.25>7.15

Conclusion: We reject the null hypothesis at the significance level of α=0.05. Therefore, we cannot prove $\sigma_1^2/\sigma_2^2=1$, so I don't recommend the researchers carry out the analysis.

9.116

If we want to use t-test, we need to assume $\sigma_1=\sigma_2$.

$F_1= S_{large}^2/S_{small}^2=4^2/2^2=4$

$F_2= S_{large}^2/S_{small}^2=15^2/10^2=2.25$

The rejection regions in both scenarios are $F_1>F_{\sigma/2}$ and $F_2>F_{\sigma/2}$, where $F_{\sigma/2}$ is a constant. Since $F_1>F_2$, it has a larger chance to be rejected. Therefore, the assumption required for a t-test to compare means is more likely to be violated in the first scenario.

10.16

a. The response variable is the tablet dissolution time.

b. Factors are binding agent, binding concentration and relative density. Each factor has two levels. The first factor (binding agent) has two levels, khaya gum and PVP. The second factor (binding concentration) has two levels, 0.5% and 4%. The third factor (relative density) has two levels, low and high.

c. There are 8 treatments. (khaya, 0.5%, low), (gum, 0.5%, low), (khaya, 4%, low), (gum, 4%, low), (khaya, 0.5%, high), (gum, 0.5%, high), (khaya, 4%, high), (gum, 4%, high).

10.40

a. A completely randomized design is employed in this study since students were randomly assigned.

b. H_0 : $\mu_1 =\mu_2=\mu_3=\mu_4$ H_a: At least one mean is different.

```
> data2 <- read_csv("D:/学习/McGill/U0/Winter/Math204/Assignment/A1/DRINKS.csv")
Parsed with column specification:
cols(
  GROUP = col_character(),
  SCORE = col_double()
)

> data_aov <- aov(data2$SCORE~data2$GROUP,data=data2)
> summary(data_aov)
             Df Sum Sq Mean Sq F value   Pr(>F)    
data2$GROUP   3 0.9506  0.3169   10.29 3.76e-05 ***
Residuals    40 1.2317  0.0308                     
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> qf(0.95,3,40)
[1] 2.838745
```

Since F=10.29>2.84, we reject H_0 when α=0.05. It is possible that there are differences among the mean task scores for the four groups when α=0.05.

c. 1. Students are randomly assigned to different groups and randomly sampled from the population. 2. Variances within each group is equal. 3. Outcome within each group is normally distributed.

10.42

$H_0 : \mu_1 = \mu_2 = \mu_3$

H_a: At least one mean is different.

Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
Column 1	50	1532	30.64	401.4188		
Column 2	42	1101	26.21429	561.7822		
Column 3	47	711	15.12766	246.592		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	6109.714	2	3054.857	7.687191	0.000687	3.0627
Within Groups	54045.83	136	397.3958			
Total	60155.54	138				

Since our test statistics F=7.69 is greater than F(K-1,n-K)=3.0627, we reject H_0 at the significance level α=0.05. Therefore, the mean percentages of names recalled differ for the three name-retrieval methods when α=0.05.

10.64

From 10.43, we get this table.

	Sample Size	Mean Drop	Std. Dev.
Group T: Volunteer + Trained Dog	26	10.5	7.6
Group V: Volunteer only	25	3.9	7.5
Group C: Control group (no visit)	25	1.4	7.5

$SSE=(n_1-1)s_1^2+(n_2-1)s_2^2+(n_3-1)s_3^2$

MSE=SSE/n=56.77

```
> (7.6^2*26+7.5^2*25+7.5^2*25)/76
[1] 56.76658
```

s=sqrt(MSE)=7.534

$\alpha^*=2\alpha/(k*(k-1))=2*0.03/6=0.01$

$t(\alpha^*/2)=t(0.005,26+25+25-3)=2.64$

```
> qt(1-0.005,26+25+25-3)
[1] 2.644869
```

The confidence interval for $\mu_T-\mu_V$ is $((x_T-x_V)-t*s*sqrt(1/n_T+1/n_V), (x_T-x_V)+t*s*sqrt(1/n_T+1/n_V))$=(10.5-3.9-2.64*7.534*sqrt(1/26+1/25), 10.5-3.9+2.64*7.534*sqrt(1/26+1/25))=(1.02,12.18)

The confidence interval for $\mu_T-\mu_C$ is $((x_T-x_C)-t*s*sqrt(1/n_T+1/n_C), (x_T-x_C)+t*s*sqrt(1/n_T+1/n_C))$=(10.5-1.4-2.64*7.534*sqrt(1/26+1/25), 10.5-1.4+2.64*7.534*sqrt(1/26+1/25))=(3.52,14.68)
The confidence interval for $\mu_V-\mu_C$ is $((x_V-x_C)-t*s*sqrt(1/n_V+1/n_C), (x_V-x_C)+t*s*sqrt(1/n_V+1/n_C))$=(3.9-1.4-2.64*7.534*sqrt(1/25+1/25), 3.9-1.4+2.64*7.534*sqrt(1/25+1/25))=(-3.14,8.14)
0 is an element of (-3.14,8.14). 0 is not an element of (1.02,12.18). 0 is not an element of (3.52,14.68).
Therefore, at the significance level $\alpha=0.05$, we conclude that there is a difference between T's mean and V's mean and between T's mean and C's mean, and we cannot conclude there is a difference between V's mean and C's mean.