THE UNIVERSITY OF MINNESOTA
Statistics 5021 May 5, 2003
Sample Final Examination Solutions

1. During the 1996 presidential campaign, one of the Republican candidates based his campaign on a proposal for a "flat tax" (income tax with just one rate).

A polling organization interviewed random samples of Republicans in Delaware and Arizona, both on the same day. In both states those interviewed were asked the identical question “Do you favor the flat tax proposal?”. Out of 260 Delaware Republicans interviewed, 90 favored the flat tax and 170 did not, while out of 140 Arizona Republicans interviewed, 62 favored the flat tax and 78 did not.

(a) Is there evidence the actual proportion of Delaware Republicans favoring the flat tax differs from the actual proportion of Arizona Republicans that favor it? Use $\alpha = .05$

H_0: $p_1 = p_2$ H_a: $p_1 \neq p_2$

```
Cmd> n1 <- 260; n2 <- 140; x1 <- 90; x2 <- 62 # enter data

Cmd> phat1 <- x1/n1;phat2 <- x2/n2;vector(phat1,phat2)
(1)      0.34615      0.44286  Sample  proportions

Cmd> phatc <- (x1 + x2)/(n1 + n2) # combined estimate

Cmd> sec <- sqrt(phatc*(1 - phatc)*(1/n1 + 1/n2)); sec
(1)     0.050882        Standard  error  for  test

Cmd> z <- (phat1 - phat2)/sec; z # Z-statistic
(1)      -1.9005

Cmd> 2*(1 - cumnor(abs(z)))#two tail P-value;2*(1-.9713) (Table A)
(1)     0.057364
```

Not significant; not sufficient evidence to reject H_0

(b) Find a 95% confidence interval for the difference between Delaware and Arizona in the actual proportions of Republicans favoring the flat tax.

```
Cmd> se1 <- sqrt(phat1*(1-phat1)/n1 + phat2*(1 - phat2)/n2); se1
(1)     0.051312        Standard  error  for  C.I.

Cmd> phat1 - phat2 # difference in sample proportions
(1)    -0.096703

Cmd> margin <- 1.96*se1; margin # margin of error
(1)      0.10057

Cmd> phat1 - phat2 + vector(-1,1)*margin #C.I.
(1)     -0.19727    0.0038677
```

(c) Suppose now the actual proportion of Arizona Republicans who favor the flat tax is .45. How many Arizona Republicans must be sampled for the 95% margin of error in estimating this proportion to be 4% (.04)?

```
Cmd> m <- .04 # aimed for margin of error

Cmd> p0 <- .45 # presumed value for p

Cmd> 1.96^2*p0*(1 - p0)/m^2
(1)        594.25
```

2. University researchers put 12 SafeHouse brand radon detectors in a chamber that exposed them to 105 picocuries per liter of radon. The readings were as follows

91.9	97.8	111.4	122.3	105.4	95.0
103.8	99.6	96.6	119.3	104.8	101.7

Exhibit 2 (Originally on separate exhibit sheet)

```
Cmd> radon <- vector(91.9,97.8,111.4,122.3,105.4,95,103.8,\
       99.6,96.6,119.3,104.8,101.7)

Cmd> sum(radon)
(1)       1249.6

Cmd> sum((radon - sum(radon)/12)^2)
(1)       971.43
```

(a) Is there convincing evidence that SafeHouse brand radon detectors are biased, that is their mean measurement differs from 105? Use $\alpha = .05$. Be sure to state the null and alternative hypotheses.

H_0: $\mu = 105$, H_a: $\mu \neq 105$ (two sided alternative)

```
Cmd> n <- 12; ybar <- 1249.6/n; ybar # sum(radon)/n
(1)        104.13

Cmd> s <- sqrt(971.43/(n-1)); s # sqrt(sum((y-ybar)^2)/(n-1))
(1)        9.3974

Cmd> mu0 <- 105; se <- s/sqrt(n); se # std error of ybar
(1)        2.7128

Cmd> t_stat <- (ybar - mu0)/se; t_stat # t-statistics
(1)     -0.31947

Cmd> twotailt(t_stat,n-1) #2-tail P-value
(1)       0.75535

Cmd> invstu(1 - .05/2,n-1)#2-tail critical value (see also Table D)
(1)         2.201
```

Since $|t| = 0.31947 < 2.201$, there is not convincing evidence of bias. You cannot reject H_0.

The use of a Z-test and normal P-value is wrong. This is truly a small sample with unknown σ.

(b) Find a 95% confidence interval for the mean measurement of SafeHouse brand radon detectors when exposed to 105 picocuries per liter of radon.

```
Cmd> margin <- 2.201 *se;margin # margin of error
(1)       5.9708

Cmd> ybar + vector(-1,1)*margin # 95% confidence interval
(1)       98.162        110.1
```

The use of the normal critical value 1.96 is wrong since this is a small sample with unknown σ.

3. Two astronomers, H and M measured the corrected red shift (km/sec) of 10 type SO galaxies. Here is a table of their results:

	Galaxy Number									
Astronomer	1	2	3	4	5	6	7	8	9	10
H x_1	1507	858	1205	832	2206	924	4607	2592	1930	5378
M x_2	1471	778	1155	915	2194	1033	4430	2670	2050	5278
Difference d	36	80	50	-83	12	-109	177	-78	-120	100

Exhibit 3 (Originally on separate exhibit sheet)

```
Cmd> x1 <- vector(1507,858,1205,832,2206,924,4607,2592,1930,5378)

Cmd> x2 <- vector(1471,778,1155,915,2194,1033,4430,2670,2050,5278)

Cmd> d <- x1 - x2

Cmd> vector(sum(x1),sum(x2),sum(d))
(1)       22039       21974          65

Cmd> xbar1 <- describe(x1,mean:T); xbar2 <- describe(x2,mean:T)

Cmd> write(vector(sum((x1 - xbar1)^2),sum((x2 - xbar2)^2)))
VECTOR:
(1)       22890618.9       21349016.4

Cmd> write(sum((x1 - xbar1)*(x2 - xbar2)))
NUMBER:
(1)       22074567.4

Cmd> dbar <- describe(d,mean:T)

Cmd> write(sum((d - dbar)^2))
NUMBER:
(1)          90500.5
```

(a) Is this a paired or a two sample comparison? Explain your answer.

It is a paired comparison, since there were 10 galaxies with two measurements each. You cannot ignore the pairing and treat it as a two-sample problem.

For the remainder of this question, your answers must be consistent with your answer to (a).

(b) H was a fairly inexperienced graduate student and M was an experienced professor. Is there evidence at the 10% level of significance that H's measurements tended to be systematically greater than those of M.

H_0: $\mu_H = \mu_M$, H_a: $\mu_H > \mu_M$

```
Cmd> n <- 10 # sample size
(1)          10

Cmd> dbar <- 65/n; dbar # sample mean of differences
(1)         6.5

Cmd> sd <- sqrt(90500.5/(n-1)); sd # sample standard deviation of d
(1)      100.28

Cmd> se <- sd/sqrt(n);se # standard error of dbar
(1)      31.711

Cmd> t_stat <- dbar/se; t_stat # t-statistic
(1)     0.20498

Cmd> 1 - cumstu(t_stat,n-1) # 1-tail P-value
(1)     0.42107

Cmd> invstu(1 - .10,n-1) # 1-tail critical value (see Table D)
(1)       1.383
```

Since $t = 0.20498 < t_{.10} = 1.383$, you cannot reject H_0 and the 10% significance level.

Using a Z-test instead of a t-test would be wrong since the sample is small and σ is not known

(c) Find a 95% confidence interval for the difference between the average of measurements by H and the average of measurements by M?

```
Cmd> t_025 <- invstu(1 - .05/2,n-1); t_025
(1)      2.2622

Cmd> margin <- t_025*se; margin # margin of error
(1)      71.734

Cmd> dbar + vector(-1,1)*margin # 95% C.I.
(1)     -65.234       78.234
```

As before, using the normal critical value 1.96 would an error.

4. The Cheese data set in the Appendix to *IPS* contains data relating the taste of cheese to the concentrations of three chemicals, acetic acid, hydrogen sulfide and lactic acid. The variables are

`taste`	Taste score, an average of scores from several tasters
`acetic`	natural log of the concentration of acetic acid
`h2s`	natural log of the concentration of hydrogen sulfide
`lactic`	concentration of lactic acid

There are 30 samples of cheese, all manufactured by the same process.

Exhibit 4 (Originally on separate exhibit sheet)

```
Cmd> readdata("",caseno,taste,acetic,h2s,lactic) # file cheese.txt
# File cheese.txt from publisher's web site
# Data for Appendix data set CHEESE, p. D-1 of IPS4
# Col. 1: id = case number (1-30)
# Col. 2: taste = taste score combined from several tasters
# Col. 3: acetic = log acetic acid concentration
# Col. 4: h2s = log hydrogen sulfide concentration
# Col. 5: lactic = lactic acid concentration
Read from file "TP1:Stat5021:Data:Appendix:cheese.txt"
Column 1 saved as REAL vector id
Column 2 saved as REAL vector taste
Column 3 saved as REAL vector acetic
Column 4 saved as REAL vector h2s
Column 5 saved as REAL vector lactic

Cmd> regress("taste=acetic+h2s+lactic",pval:T)
Model used is taste=acetic+h2s+lactic
                Coef      StdErr           t     P-Value
CONSTANT     -28.877      19.735     -1.4632      0.1554
acetic       0.32774      4.4598    0.073489     0.94198
h2s           3.9118      1.2484      3.1334   0.0042471
lactic        19.671      8.6291      2.2796    0.031079

N: 30, MSE: 102.63, DF: 26, R^2: 0.65177
Regression F(3,26): 16.221, P-value: 3.8102e-06, Durbin-Watson:
1.5751
To see the ANOVA table type 'anova()'

Cmd> SS
    CONSTANT      acetic         h2s      lactic      ERROR1
       18057      2314.1        2147      533.32      2668.4

Cmd> DF
    CONSTANT      acetic         h2s      lactic      ERROR1
           1           1           1           1          26

Cmd> regpred(vector(6,8,1.6),estimate:F)# acetic=6,h2s=8,lactic=1.6
component: SEest
(1)     2.7302
component: SEpred
(1)     10.492
```

(a) What is the null hypothesis H_0 tested by F = 16.221 (underlined in `regress()` output)? What is the alternative hypothesis H_a? What critical value should be used for a test when $\alpha = .01$?

H_0: $\beta_1 = \beta_2 = \beta_3 \neq 0$: H_a: At least two coefficients non-zero

```
Cmd> df <- 26

Cmd> invF(1 - .01,3,df) # or interpolate in Table E
(1)       4.6366
```

(b) Find a 95% confidence interval for the coefficient of `lactic`.

```
Cmd> b3 <- 19.671; se_b3 <- 8.6291 # from regress() output

Cmd> t_025 <- invstu(1 - .05/2,df); t_025 # or from Table D
(1)       2.0555   2-tail  critical  value

Cmd> margin3 <- t_025*se_b3;margin3 # margin of error
(1)       17.737

Cmd> b3 + vector(-1,1)*margin3 # 95% C.I.
(1)       1.9336       37.408
```

As before, it would be wrong to use the normal critical value 1.96 instead of Student's t.

(c) Estimate with a 95% confidence interval the mean taste score when log acetic acid concentration is 6, log hydrogen sulfide concentration is 8, and lactic acid concentration is 1.6.

> This asks for an interval for a *mean* tast score, so the standard error is component `SEest` from `regpred()` output. Since the `estimate` component was suppressed, you had to calculate it directly using the estimated coefficients from the `regress()` output.

```
Cmd> regpred(vector(6,8,1.6),estimate:T)# acetic=6,h2s=8,lactic=1.6
component: estimate
(1)       35.857
component: SEest
(1)       2.7302
component: SEpred
(1)       10.492

Cmd> mu_y <- -28.877 + 0.32774*6 + 3.9118*8 + 19.671*1.6;mu_y
(1)       35.857   Estimate

Cmd> seest <- 2.7302 # SEest component

Cmd> margin_mu <- t_025*seest; margin_mu # margin of error
(1)        5.612

Cmd> mu_y + vector(-1,1)*margin_mu # 95% confidence interval
(1)       30.245       41.469
```

Using 1.96 would be an error.

5. An experiment was conducted to compare birch plywood made with 6 different resin glues, numbered here 1 through 6. 10 pieces of plywood made with each glue were tested and the shear strength (PSI = pounds per square inch) measured. In the following table are the values of PSI for each of the 60 pieces of plywood.

Glue	1	2	3	4	5	6
Mean shear strength	502	470	500	520	551	620
	458	483	502	510	556	643
	445	478	480	582	592	589
	479	493	519	530	562	576
	468	498	459	495	566	576
	463	466	499	543	558	581
	517	492	500	513	573	606
	494	479	509	540	557	633
	499	534	528	523	633	562
	463	531	538	532	638	579

Exhibit 5 (Originally on separate exhibit sheet) MacAnova output

```
Cmd> psi <- vector(502,458,445,479,468,463,517,494,499,463,\
       470,483,478,493,498,466,492,479,534,531,\
       500,502,480,519,459,499,500,509,528,538,\
       520,510,582,530,495,543,513,540,523,532,\
       551,556,592,562,566,558,573,557,633,638,620,\
       643,589,576,576,581,606,633,562,579)

Cmd> glue <- factor(vector(rep(1,10),rep(2,10),rep(3,10),\
       rep(4,10),rep(5,10),rep(6,10)))

Cmd> print(format:"1.0f",glue)
glue:
 (1) 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3
(31) 4 4 4 4 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 5 6 6 6 6 6 6 6 6 6 6

Cmd> tabs(psi,glue,mean:T,count:T,stddev:T)
component: mean
(1)       478.8       492.4       503.4       528.8       578.6
(6)       596.5
component: count
(1)          10          10          10          10          10
(6)          10
component: stddev
(1)       23.16      23.396      22.707       23.63      32.139
(6)      27.428

Cmd> anova("psi=glue")
Model used is psi=glue
                DF          SS          MS
CONSTANT         1    16838104    16838104
glue             5      115280     23056.1
ERROR1          54     35486.9     657.165
```

(a) On the basis of these data, can you statistically conclude that the six glues differ with respect to the PSI of plywood made using them? Use $\alpha = .05$. Be sure to give the critical value used.

H_0: $\mu_1 = \mu_2 = ... = \mu_6$; H_a: at least two means differ

```
Cmd> f_stat <- 23056/657.16; f_stat # ratio of MS from output
(1)       35.084

Cmd> I <- 6; N <- length(psi); N
(1)           60

Cmd> df_glue <- 5;df_error <- N - I; df_error
(1)           54

Cmd> 1 - cumF(f_stat,df_glue,df_error) # P-value
(1)            0

Cmd> invF(1 - .05,df_glue,df_error) #(from Table E)
(1)       2.3861
```

You can conclude the glues differ; F is highly significant.

(b) Glues 1 and 3 were made by different manufacturers and were intended to have the same properties. Is there statistical evidence at the 5% level of significance that their mean strengths were not the same?

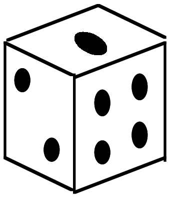
H_0: $\mu_1 = \mu_3$; H_a: $\mu_1 \neq \mu_3$

This is like a two-sample t except you use s_p^2 = MSE from ANOVA, pooling the information about the variance from all 6 samples.

```
Cmd> ybar1 <- 478.8;ybar3 <- 503.4 # from tabs() output)

Cmd> s_pooled <- sqrt(657.16) # from anova() output

Cmd> n1 <- n2 <- n3 <- n4 <- n5 <- n6 <- 10 # sample sizes

Cmd> se <- s_pooled*sqrt(1/n1 + 1/n3); se #Std error ybar1-ybar3
(1)       11.464

Cmd> t_stat <- (ybar1 - ybar3)/se; t_stat 3 t-statistic
(1)      -2.1458

Cmd> twotailt(t_stat,N-I) # two-tail P-value
(1)     0.036403

Cmd> invstu(1 - .05/2,N-I) # two-tail critical value (see Table C)
(1)       2.0049
```

There is significant evidence at the 5% level that the means differed.

6.



A die is a cube whose faces have 1, 2, 3, 4, 5, or 6 spots with 6 opposite 1, 5 opposite 2, and 4 opposite 3. When a fair die is rolled, each face should have equal probability 1/6 of being on top and hence each number should have probability 1/6. A particular die has its spots deeply cut into the surface. It is conjectured that it is biased in favor of higher numbers, because lower numbers have less material cut away and are thus more likely to be on the bottom. If the die is not biased, a 5 or a 6 should occur with probability p = 2/6 = 1/3. If the die is biased as conjectured then p > 1/3.

The following experiment procedure is proposed to examine this conjecture.

> The die is to be rolled independently 7 times and the number of times (X) a 5 or a 6 is on top will be counted. If $X \geq 5$ (that is, at least five 5's or 6's appeared out of 7 trials), the die will be considered to be biased.

Here is an table of the probability distributions of X when p = 1/3 and when p = 3/4.

X	0	1	2	3	4	5	6	7
p = 1/3	0.0585	0.2048	0.3073	0.2561	0.1280	0.0384	0.0064	0.0005
p = 3/4	0.0001	0.0013	0.0115	0.0577	0.1730	0.3115	0.3115	0.1335

(a) The procedure described can be considered to be a test of a particular null hypothesis H_0 against a particular alternative hypothesis H_a. State H_0 and H_a.

> H_0: X is binomial with p = 1/3
> H_a: X is binomial with p > 1/3

(b) Find the *exact* value of the significance level α for this test. *Do not use a normal approximation.*

$$\alpha = P(\text{Type I error}) = P(\text{reject } H_0 \mid H_0 \text{ true}) = P(X = 5 \text{ or } 6 \text{ or } 7 \mid p = 1/3) = .0384 + .0064 + .0005 = .0453.$$

Finding the standard error of the estimate is part of using the normal approximation which was against the rules. Besides, n = 7 is too small to use the normal approximation

(c) Suppose in fact that p = 3/4. Find the *exact* value of β for this test. *Do not use a normal approximation.*

$$\alpha = P(\text{Type II error}) = P(\text{not reject } H_0 \mid H_a \text{ true}) = 1 - P(X = 5 \text{ or } 6 \text{ or } 7 \mid p = 3/4) = 1 - .3115 - .3115 - .1335 = .2435$$

Any attempt to find a probability using Table A would be wrong here.