Math 583 Exam 3, Fall 2012 Name__________

```
$Ftable
            Fj        pvals
[1,] 82.147221 0.000000e+00
[2,] 58.448961 0.000000e+00
[3,] 15.700326 4.258563e-09
[4,]  9.072358 1.281220e-05
[5,] 45.364862 0.000000e+00

$MANOVA
      MANOVAF pval
[1,] 67.80145    0
```

1) The above output is for the R Seatbelts data set where $Y_1 = drivers$ = number of drivers killed or seriously injured, $Y_2 = front$ = number of front seat passengers killed or seriously injured, and $Y_3 = back$ = number of back seat passengers killed or seriously injured. The predictors were $x_2 = kms$ = distance driven, $x_3 = price$ = petrol price, $x_4 = van$ = number of van drivers killed, and $x_5 = law = 0$ if the law was in effect that month and 1 otherwise. The data consists of 192 monthly totals in Great Britain from January 1969 to December 1984, and the compulsory wearing of seat belts law was introduced in February 1983.

a) Do the MANOVA F test.

i) Ho the nontrivial predictors are not needed in the mreg model
H_1 at least one nontrivial predictor is needed

ii) $F_0 = 67.801$

iii) pval = 0

iv) reject Ho at least one of the nontrivial predictors is needed in the model (or there is an mreg relationship between Y_1, Y_2, Y_3 and the predictors $x_2 - x_5$.)

b) Do the F_4 test.

H_0 $\underline{b}_4^T = \underline{0}$ H_1 $\underline{b}_4^T \neq \underline{0}$

$F_4 = 9.072$

pval = 0

reject Ho, van is needed in the mreg model (for drivers, front and back given the other predictors are in the model)

20

Loadings:

	Factor1	Factor2
height	0.872	
arm.span	0.973	
forearm	0.938	
lower.leg	0.876	
weight		0.961
bitro.diameter		0.803
chest.girth		0.796
chest.width	0.125	0.611

	Factor1	Factor2
SS loadings	3.375	2.589
Proportion Var	0.422	0.324
Cumulative Var	0.422	0.745

2) The above output is for the factor analysis using a correlation matrix of eight physical measurements on 305 girls between ages seven and seventeen.

a) What is the cumulative variance explained by the 2 factors?

0.745

b) Which factor has a nonzero loading for weight?

Factor 2

c) Explain Factor 2.

a linear combination of weight, bitro.diameter, chest.girth and chest.width

21

$\underline{\bar{x}} = (185.72, 183.84)^T$

✓ 3) Data for first and second adult sons had $n = 25$ and variables X_1 = head length of first son and X_2 = head length of second son. ~~For this data set, $\overline{x} = (0.564, 0.603)^T$.~~ Suppose $\mu_0 = (182, 182)^T$ and $T_C^2 = 1.28$. Perform the one sample Hotelling's T^2 test.

i) $H_0\ \underline{\mu} = \underline{\mu}_0 \quad H_1\ \underline{\mu} \neq \underline{\mu}_0$

ii) $T_C^2 = 1.28$

iii) $\frac{n-p}{(n-1)p} T_C^2 = \frac{25-2}{24(2)} 1.28 = 0.6133$

$\text{pval} = P(0.613 < F_{2,23}) > 0.05$

	2
20	3.49
25	3.39

iv) fail to reject H_0, $\underline{\mu} = \begin{pmatrix} 182 \\ 182 \end{pmatrix}$

$\frac{10}{}$

✓ 4) Suppose the 15 units are 1 Adatorwovor, 2 Adhikari, 3 Alanzi, 4 Alsibiani, 5 Al-Talib, 6 Fan, 7 Kuo, 8 Lamsal, 9 Liu, 10 Meyer, 11 Peiris, 12 Rathnayake, 13 Rupasinghe, 14 Schroeppel and 15 Watagoda. Use the following output to allocate the 15 units to three groups of 5. Show the three groups.

```
> sample(15)
13  9  12  6  5 / 11  2  3  15  7 / 14  4  10  8  1
```

group 1 : Rupasinghe, Liu, Rathnayake, Fan, Al Talib

group 2 : Peiris, Adhikari, Alanzi, Watagoda, Kuo

group 3 : Shroeppel, Alsibiani, Meyer, Lamsal, Adatorwovor

$\frac{}{7}$

5) The cork data records weights of cork deposits for 28 trees in the directions $x_1 =$ *north*, $x_2 =$ *east*, $x_3 =$ *south* and $x_4 =$ *west*. Do the four step repeated measurements test if $T_R^2 = 20.74$. Show how an F table is used.

i) H_0 $\mu_y = 0$ H_A $\mu_y \neq 0$

ii) $T_R^2 = 20.74$

iii) $\frac{n-p+1}{(n-1)(p-1)} T_R^2 = \frac{28-4+1}{(28-1)(4-1)} 20.74 = 6.401$

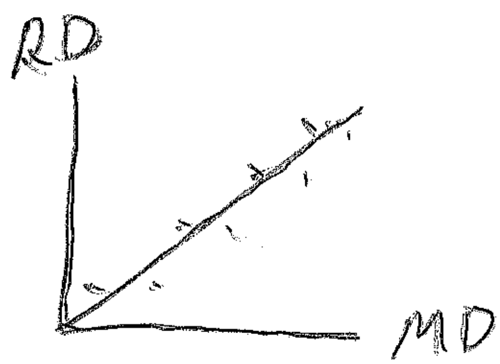
pval $= P(6.401 < F_{3,25}) < .05$

	3
25	2.99

iv) reject H_0 the means of the cork deposits are different (so depend on the direction)

10

6) a) Sketch a DD plot of the residual vectors $\hat{\epsilon}_i$ for the multivariate linear regression model if the error vectors ϵ_i are iid from a multivariate normal distribution. b) Does the DD plot change if the one way MANOVA model is used instead of the multivariate linear regression model?

a) 

b) no

10

```
Response       = pass   Terms          = (hscalc survey)
Coefficient Estimates
Label       Estimate         Std. Error     Est/SE     p-value
Constant    0.875469         0.532291        1.645      0.1000
hscalc      10.3274          54.7562         0.189      0.8504
survey     -2.26176          1.23828        -1.827      0.0678
```

Exam 2 problem in 2024

7) The group = response variable *pass* was a 1 if the Math 150 (intro calc) student got a C or higher on the combined final and a 0 (withdrew, D or F) otherwise. Data was collected at the beginning of the semester on 31 students who took a section of Math 150 in Fall, 2002. Here $x_1 = hscalc$ was coded as a 1 if the student said that their last math class was high school calculus and as a 0 otherwise. Here $x_2 = survey$ was coded as a 1 if the student failed to turn in the survey, 0 otherwise. Assume the case to be classified has $\boldsymbol{x}$ with $hscalc = x_1 = 1.0$ and $survey = x_2 = 0.0$.

a) Find ESP for $\boldsymbol{x}$. $= \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 =$

$.875469 + 10.3274 + 0 = \boxed{11.2029}$

b) Is $\boldsymbol{x}$ classified in group 0 or group 1?

$\boxed{\text{group 1}}$ since ESP > 0

12

8) Krzanowski (1998, p. 381) describes a one way MANOVA model for producing plastic film. Samples were tested in 2 groups using a low rate or a high rate. using $X_1 = tear$, $X_2 = gloss$, and $X_3 = opacity$. There were $n_i = 10$ samples for $i = 1, 2$. Suppose $t_0 = 7.561$ and pvalue $= 0.0022$. Do a 4 step one way MANOVA test.

$H_0\ \underline{\mu}_1 = \underline{\mu}_2$ $\qquad H_1$ not H_0

$\underline{\mu}_1 \neq \underline{\mu}_2$ (since only 2 groups)

$t_0 = 7.561$

pval $= 0.0022 < \alpha = 0.05$

reject H_0, the mean film measurements are not the same (so depend on high or low rate).

10