Read Carefully. Give an answer in the form of a number or numeric expression where possible. Use a value of 0.05 for α if not specified. t-tables, Chi square tables and F tables are provided separately. You may keep these tables. All multiple choice questions have only one answer unless otherwise indicated.

1) 2 points each – Answer the following questions as true (T) or false (F) by circling the appropriate letter.

T (F) a) A Chi square test of a variance against an hypothesized value is always a one tailed test.

T (F) b) The square root of values of the F distribution with 1, γ degrees of freedom will be equal to a Chi square value with γ degrees of freedom.

T (F) c) The t distribution is used to calculate confidence intervals for both means and variances.

(T) F d) The upper limits of the F distribution given in the F table can be used to calculate the lower limits as well as the upper limits.

(T) F e) It is not necessary to assume homogeneity of variance for two-sample t-tests since the test can also be done without equal variances.

(T) F f) The Satterthwaite approximation is used when a test of the equality of the variances in a two-sample t-test is rejected.

T (F) g) The expected value of the F distribution is zero (0.0).

(T) F h) The paired t-test is calculated as a one-sample t-test.

2) 4 points – Indicate **all** of the distributions below that are not symmetric about their expected value.

a) Z distribution

b) t distribution

(c) χ^2 distribution

(d) F distribution

3) 3 points – Which one of the following best describes a P value for a given value of a test statistic?

a) it is the probability of making a type I error

b) it is the probability of making a type II error

c) it is the probability of rejecting the null hypothesis

(d) it is the probability of getting a less likely value of the statistic

4) 3 points – The calculation of a confidence intervals for a variance uses which one of the following distributions.

a) Z distribution

b) t distribution

c) χ^2 distribution

d) F distribution

5) 3 points each — Find the t values or probabilities indicated by the expression below.
Place your answer in the blank space **keeping at least three decimal places**.

a) $P(|t| \le 2.921) = ?$ d.f. = 16 P = 0.990

b) $P(t \ge t_0) = 0.75$ d.f. = 15 t_0 = -0.691

c) $P(t \le t_0) = 0.20$ d.f. = 100 t_0 = -0.845

d) $P(-t_0 \le t \le t_0) = 0.90$ d.f. = 5 t_0 = 2.015

6) 3 points each — Find the Chi square values or probabilities indicated by the expression below.
Place your answer in the blank space **keeping at least three decimal places**.

a) $P(\chi^2 \ge 16.013) = ?$ d.f. = 7 P = 0.025

b) $P(29.707 \le \chi^2 \le 67.505) =$ d.f. = 50 P = 0.940

c) $P(\chi^2 \le 12.549) = ?$ d.f. = 10 P = 0.750

d) $P(\chi^2_1 \le \chi^2 \le \chi^2_2) = 0.500$ d.f. = 5 χ^2_1 = 2.6746 , χ^2_2 = 6.6257
assume symmetry for this problem

e) $P(\chi^2 \le 2.833) = ?$ d.f. = 7 P = 0.100

7) 3 points each — Find the F values or probabilities indicated by the expression below.
Place your answer in the blank space **keeping at least three decimal places**.

a) $P(F \le F_0) = 0.025$ $\gamma_{num}, \gamma_{den} = 8, 10$ d.f. $F_0 =$ 0.233

b) $P(F \ge 8.022) = ?$ $\gamma_{num}, \gamma_{den} = 5, 4$ d.f. $P =$ from 0.05 to 0.01
exact answers;
5, 4 d.f.=0.033 or 4, 5 d.f.=0.021

c) $P(F_0 \le F \le 5.523) = 0.95$ $\gamma_{num}, \gamma_{den} = 6, 9$ d.f. $F_0 =$ anything < 0.25
exact answers;
9, 6 d.f.=0.232 or 6, 9 d.f.=0.181

d) $P(F \le 2.111) = ?$ $\gamma_{num}, \gamma_{den} = 30, 120$ d.f. $P =$ anything > 0.99
exact answers;
30, 120 d.f.=0.9976 or 120, 30 d.f.=0.9900

e) $P(0.022 \le F \le 7.764) = ?$ $\gamma_{num}, \gamma_{den} = 3, 5$ d.f. $P =$ 0.970

For each of the questions below state the type of test that should be used to address the research question. You may assume that all relevant assumptions are met.

8) 3 points – If an experiment is to be conducted on two samples to test the hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$, which of the statistical tests below will be used?

Type of test (**circle one**) 2-sample t-test paired t-test Chi square test (F test)

9) 3 points – If an experiment is to be conducted on a sample to test the hypothesis $H_0 : \sigma^2 = \sigma_0^2$, which of the statistical tests below will be used?

Type of test (**circle one**) 2-sample t-test paired t-test (Chi square test) F test

10) 3 points – In many manufacturing processes it is necessary to insure that the product is very uniform with very little variation between units. A manufacturer of ball bearings wants to test his product to unsure that the variance of the 12 mm bearings he produces is no more than 0.003 mm^2. After measuring a random sample, what type of test would he use to test this hypothesis?

Type of test (**circle one**) 2-sample t-test paired t-test (Chi square test) F test

11) 3 points – A student of advertizing wants to compare the mean amount of advertizing in two popular magazines, Time and Newsweek. He gets data for the dollar value of advertizing in each of the magazines for the last 20 months? What type of test would he use to compare the advertizing?

Type of test (**circle one**) (2-sample t-test) paired t-test Chi square test F test

The computer output below will provide information to answer the last questions on the exam.

The rainfall in acre-feet from 52 clouds, 26 of which were chosen at random and seeded with silver nitrate, is show. The results of a comparison of the two groups using SAS® proc t-test is given below.

```
21    DATA Students; INFILE CARDS MISSOVER;
22       INPUT Unseeded_Clouds Seeded_Clouds;
23         label Unseeded_Clouds = 'rainfall from unseeded clouds with silver
nitrate'
24                Seeded_Clouds = 'rainfall from seeded clouds with silver nitrate'
25                Seeded = 'seeded clouds with silver nitrate; Yes or No'
26                Rainfall = 'Amount of rainfall in acre-feet';
27       seeded = 'Yes';  Rainfall = Seeded_Clouds; output;
28       seeded = 'No';  Rainfall = Unseeded_Clouds; output;
29       Keep seeded rainfall;
30    CARDS;
NOTE: The data set WORK.STUDENTS has 52 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds
57     ;
58    PROC ttest DATA=Students;
59       class Seeded;
60       VAR Rainfall;
61    RUN;
NOTE: The PROCEDURE TTEST printed page 1.
NOTE: PROCEDURE TTEST used (Total process time):
      real time           0.12 seconds
      cpu time            0.06 seconds

Rainfall from Cloud-Seeding

The TTEST Procedure

Variable:  Rainfall  (Amount of rainfall in acre-feet)
Seeded          N        Mean     Std Dev     Std Err     Minimum     Maximum
No             26       164.6       278.4     54.6039      1.0000      1202.6
Yes            26       442.0       650.8       127.6      4.1000      2745.6
Diff (1-2)             -277.4       500.5       138.8

Seeded      Method            Mean        95% CL Mean    Std Dev    95% CL Std Dev
No                           164.6    52.1296    277.0     278.4    218.4    384.3
Yes                          442.0      179.1    704.8     650.8    510.4    898.4
Diff (1-2)  Pooled          -277.4     -556.2   1.4319     500.5    418.8    622.2
Diff (1-2)  Satterthwaite   -277.4     -559.6   4.7643

Method           Variances        DF    t Value    Pr > |t|
Pooled           Equal            50      -2.00      0.0511
Satterthwaite    Unequal      33.855      -2.00      0.0538

             Equality of Variances
Method      Num DF    Den DF    F Value    Pr > F
Folded F        25        25       5.46    <.0001
```

12) The computer output provided has a program segment and output from a SAS® proc ttest. Answer the questions below using this computer information. The null hypothesis is that there is "no difference in raifall between seeded and unseeded clouds" and all tests are to be done at $\alpha = 0.05$.

a) 4 points – Using appropriate considerations for the choice of a test statistic, does there appear to be a statistically significant difference in the means between the two groups?

Circle one: **YES** **NO** P value (4 decimal places) 0.0538

b) 3 points – Using appropriate considerations for the choice of a test statistic, does there appear to be a statistically significant difference in the variances for the two groups?

Circle one: **YES** **NO** P value (4 decimal places) <0.0001

c) 3 points – Is the question in part "a" above a one-tailed test or a two-tailed test?

Circle one: **one-tailed** **two-tailed**

The computer output below will provide information to answer the last questions on the exam.

Groups of dolphins were observed off the coast of Iceland near Keflavik in 1998. The data give the time of the day and the main activity of the group, whether **travel**ling quickly, **feed**ing or **social**izing. The dolphin groups varied in size - usually feeding or socializing groups were larger than travelling groups..

```
22    options nocenter ps=256 ls=99 nodate nonumber nolabel;
23    data one; length activity period $ 10;
24       input Activity $ Period $ Groups;
25       label Activity = 'Main activity of group: travelling, feeding or socializing'
26             Period = 'Time of the day: Morning, Noon, Afternoon or Evening'
27             Groups = 'Number of groups observed';
28    cards;

NOTE: The data set WORK.ONE has 12 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds

45    proc freq;
46      weight groups;
47      tables period*activity / chisq expected cellchi2 norow nocol nopercent;
48    run;



Activities of Dolphin Groups

Data listing

The FREQ Procedure
```

Table of period by activity

period Frequency Expected Cell Chi-Square	activity Feed	Social	Travel	Total
Afternoon	0 10.709 10.709	9 7.545 0.2806	14 4.746 18.044	23
Evening	56 36.783 10.04	10 25.915 9.7741	13 16.302 0.6687	79
Morning	28 33.524 0.9102	38 23.619 8.7561	6 14.857 5.2802	72
Noon	4 6.9841 1.275	5 4.9206 0.0013	6 3.0952 2.726	15
Total	88	62	39	189

For office use only

Do not write in this space

Statistics for Table of period by activity

Statistic	DF	Value	Prob
Chi-Square	6	68.4646	<.0001
Likelihood Ratio Chi-Square	6	74.4301	<.0001
Mantel-Haenszel Chi-Square	1	1.7891	0.1810
Phi Coefficient		0.6019	
Contingency Coefficient		0.5157	
Cramer's V		0.4256	

WARNING: 25% of the cells have expected counts less
than 5. Chi-Square may not be a valid test.
Sample Size = 189

13) The computer output provided has a program part and output from a SAS® proc freq. Answer the questions below using this computer information. All tests are to be done at $\alpha = 0.05$.

a) 4 points – Using appropriate considerations for the test, does there appear to be a statistically significant difference in dolphin activities during different times of day?

Circle one: **YES** **NO** P value (4 decimal places) <0.0001

b) 3 points – Whether significant or not, what <u>one</u> combination of activity and time of day appears to be the most inconsistent with the expectation and contributes the most to the chi square value?

Indicate your choice by circling one box with the combination that answers the question.

Morning Travel	Morning Feed	Morning Social	Noon Travel	Noon Feed	Noon Social	Afternoon Travel	Afternoon Feed	Afternoon Social	Evening Travel	Evening Feed	Evening Social

c) 3 points – Is the question in part "a" above a one-tailed test or a two-tailed test?

Circle one: **one-tailed** **two-tailed**