Test 2 Form A - Statistics 212-500 (Fall 2003)

- This test consists of 8 numbered pages. Make sure you have all 8 pages. It is your responsibility to inform me if a page is missing!!!
- You have 50 minutes to complete this test.
- Make sure you fill out and "bubble in" the following items on the scantron sheet:
  - Last Name, First Name, MI.
  - Dept (STAT), Course No (212), Section (500).
  - Social Security Number (enter your Student ID).
  - Test Form (A,B,C, or D).
  - Enter the exam number, sign and date the form.

  You **will** lose credit if you fail to fill out your scantron form correctly.
- You may use all the formula sheets from the web site and a calculator. This test is **not** open book.
- If I provide partial results—assume they are correct and use them even if they are not.
- If there is no correct answer or if multiple answers are correct, select the *best* answer.
- There is no penalty to wrong answers... so guess if you do not know an answer.
- It is your responsibility to look at the overhead or blackboard about every 15 minutes and to incorporate any relevant information into your test.
- You will not receive extra time to bubble in your answers on the scantron should you fail to do so as you were taking the exam.

An operations analyst collected data on the number of acceptable units produced from equal amounts of raw material by 24 entry-level piece work employees who had received special training. Four training levels were used (6, 8, 10, and 12 hours) with 6 employees randomly assigned to each level. The resulting means and ANOVA table were obtained:

Treatment	6	8	10	12
Mean ($\bar{x}_{i\cdot}$)	62	72	85	86

Source	SS	df	MS
Treatment	2356.5	3	785.5
Error	2014.9	20	100.74
Total	4371.4	23	

Answer the next 5 questions based on these data.

[1] The experimental design used for this experiment and the reason for using it are

a) the completely randomized design since the experimental units are homogeneous.
b) the randomized complete block design since the experimental units are homogeneous.
c) the completely randomized design to control for the variation in employees.
d) the randomized complete block design to control for the variation of employees.
e) a 4×6 factorial experiment since we wish to study the effects of both training levels and employees.

[2] What is the f-value for the above ANOVA?

a) 1.17
b) 23.39
c) 0.13
d) 6.67
e) 7.80

[3] What is the rejection region for the above ANOVA? That is, if $f \geq F_{.05}, v_1, v_2$ we reject H_0. What is $F_{.05}, v_1, v_2$?

a) 2.78
b) 2.87
c) 3.03
d) 3.10
e) 3.49

[4] A comparison that would compare the mean production of workers with shorter training (6 or 8 hours) to those with longer training (10 or 12 hours) is

a) $\ell = \mu_6 - \mu_8$
b) $\ell = \mu_{10} - \mu_{12}$
c) $\ell = 3\mu_6 - \mu_8 - \mu_{10} - \mu_{12}$
d) $\ell = \mu_6 - \mu_8 + \mu_{10} - \mu_{12}$
e) $\ell = \mu_6 + \mu_8 - \mu_{10} - \mu_{12}$

[5] The calculated value of Tukey's W at level 0.05 is 16.23. The significant differences in mean production levels found by using Tukey's method are

a) $\mu_6 \neq \mu_{12}$
b) $\mu_6 \neq \mu_{10},\ \mu_6 \neq \mu_{12}$
c) $\mu_6 \neq \mu_{10},\ \mu_6 \neq \mu_{12},\ \mu_8 \neq \mu_{12}$
d) $\mu_6 \neq \mu_{10},\ \mu_6 \neq \mu_{12},\ \mu_8 \neq \mu_{10},\ \mu_8 \neq \mu_{12}$
e) $\mu_6 \neq \mu_8,\ \mu_6 \neq \mu_{10},\ \mu_6 \neq \mu_{12},\ \mu_8 \neq \mu_{10},\ \mu_8 \neq \mu_{12}$

[6] The purpose of one-way analysis of variance (ANOVA) with fixed effects is to

a) compare the variances for several treatments.
b) compare the means for several treatments.
c) compare proportions for several treatments.
d) randomly assign experimental units to treatments.
e) to form pairs of similar experimental units.

[7] In a one-way analysis of variance, the number of observations for each treatment is increased while the number of treatments and the value of σ remain constant. Then the power of the F-test for a given set of $\alpha_1, \ldots, \alpha_I$ for which H_0 is not true will

a) become smaller.
b) become larger.
c) not be changed.
d) have an unknown change. The effect can not be determined using the above information.
e) become either larger or smaller depending on how much the number of observations is increased.

[8] In a one-way ANOVA, the expected value of the mean square for error is

a) always less than $MSTr$.
b) less than the expected value of the mean square for treatments when H_0 is true.
c) σ^2, whether or not H_0 is true.
d) $\sigma^2 + \frac{J}{I-1}\sum \alpha_i^2$, when H_a is true.
e) none of the above.

The strength of concrete used in commercial construction varies from one batch to another. The concrete samples in small test cylinders are cured for a 28-day period in a controlled environment before strength measurements are made. An experiment was carried out to compare three different curing methods with respect to compressive strength (MPa). Three samples were taken from each of the 10 batches of concrete, and the three curing methods were randomly assigned to different concrete samples from each batch. The resulting means and ANOVA table were obtained:

Method	A	B	C
Mean ($\bar{x}_{i\cdot}$)	29.49	31.31	31.40

Source	SS	df	MS
Method	23.23	2	11.615
Batch	86.79	9	9.643
Error	24.04	18	1.336
Total	134.07	29	

Answer the next 4 questions based on these data.

[9] The experimental design used for this experiment and the reason for its use are

a) the completely randomized design since the experimental units are homogeneous.
b) the randomized complete block design since the experimental units are homogeneous.
c) the completely randomized design to control for the batch-to-batch variation in concrete.
d) the randomized complete block design to control for the batch-to-batch variation in concrete.
e) a 3×3 factorial experiment.

[10] The investigator wished to test the equality of mean strengths for the three methods of curing. Which of the following conclusions is correct?

a) Since $F = 8.69 > 3.55 = F_{.05,2,18}$, reject H_0 and conclude that the mean strengths differ for the three methods of curing.
b) Since $F = 8.69 > 3.55 = F_{.05,2,18}$, do not reject H_0 and conclude that the mean strengths do not differ significantly for the three methods of curing.
c) Since $F = 7.22 > 2.46 = F_{.05,9,18}$, reject H_0 and conclude that the mean strengths differ for the three methods of curing.
d) Since $F = 7.22 > 2.46 = F_{.05,9,18}$, do not reject H_0 and conclude that the mean strengths do not differ significantly for the three methods of curing.
e) Since $F = 1.20 < 4.26 = F_{.05,2,9}$, do not reject H_0 and conclude that the mean strengths do not differ significantly for the three methods of curing.

[11] The value of W for making multiple comparisons of the mean strength levels at the $\alpha = 0.05$ level using Tukey's procedure is

a) 1.060
b) 1.086
c) 1.320
d) 1.718
e) 3.891

[12] Suppose that A is the curing method that is commonly used and that B and C are new curing methods developed by the engineer. A 95% confidence interval for $\ell = \mu_A - \frac{1}{2}(\mu_B + \mu_c)$ is

a) -1.865 ± 0.941
b) -1.865 ± 0.919
c) -1.865 ± 0.877
d) -1.865 ± 0.421
e) -1.865 ± 0.411

[13] In a two-factor experiment with fixed effects, one should

a) test the main effects first, then test the interaction.
b) main effects and interaction simultaneously.
c) test the main effects using $MSAB$ as the denominator in the F statistic.
d) test the main effects only if the interaction is significant.
e) test the main effects only if the interaction is not significant.

[14] The model for one-way ANOVA with a fixed treatment effect is $X_{ij} = \mu_i + \epsilon_{ij},\ i = 1, \ldots, I,\ j = 1, \ldots, J_i$. In this model the random error terms are

a) X_{ij}
b) μ_i
c) ϵ_{ij}
d) i
e) J_i

[15] The experimentwise Type I error rate in a multiple comparison procedure is

a) the error due to improper randomization.
b) the probability of incorrectly declaring at least one pair of means significantly different.
c) the probability of incorrectly failing to declare a pair of means significantly different
d) always $\alpha = .05$.
e) none of the above

[16] The model for one-way ANOVA with random effects is $X_{ij} = \mu + A_i + \epsilon_{ij},\ i = 1, \ldots, I,\ j = 1, \ldots, J$. The estimator of the variance due to treatments is

a) $MSTr$
b) MSE
c) $\bar{X}_{..}$
d) $\frac{MSTr - MSE}{J}$
e) $MSTr/MSE$

An occupational health study was carried out to evaluate the health status of workers. The response is forced expiratory volume (FEV), a measure of respiratory function. Observations are made on n=12 individuals in each of three plants where workers are exposed primarily to one of three toxic substances, depending on the section of the plant where they work. Suppose that the toxic substances are the only ones of interest and the plants were randomly selected from a large population of similar plants. The data resulted in the following ANOVA table: ß

Source	SS	df	MS
Substance	66.89	2	33.445
Plant	13.30	2	6.65
Interaction	24.50	4	6.125
Error	26.58	99	.2685
Total	131.27	107	

Answer the next 5 questions based on these data.

[17] In this experiment, the number of treatments is

a) 3
b) 6
c) 8
d) 9
e) 108

[18] Using the above data, one can estimate variance components for

a) error and plant
b) error and substance
c) error, plant, and interaction
d) error, substance, and interaction
e) error, substance, plant, and interaction

[19] In the test for an effect due to interaction, one would

a) reject H_0 : All $\gamma_{ij} = 0$ at level 0.05 since $F = 22.81 > 2.46 = F_{.05,4,99}$.
b) reject $H_0 : \sigma_G^2 = 0$ at level 0.05 since $F = 22.81 > 2.46 = F_{.05,4,99}$.
c) not reject H_0 : All $\gamma_{ij} = 0$ at level 0.05 since $F = 22.81 > 2.46 = F_{.05,4,99}$.
d) not reject $H_0 : \sigma_G^2 = 0$ at level 0.05 since $F = 22.81 > 2.46 = F_{.05,4,99}$.
e) not have enough information to come to a conclusion.

[20] In the test for an effect due to substance, one would

a) reject $H_0 : \sigma_A^2 = 0$ at level 0.05 since $F = 124.56 > 3.088 = F_{.05,2,99}$.
b) reject H_0 : All $\alpha_i = 0$ at level 0.05 since $F = 124.56 > 3.088 = F_{.05,2,99}$.
c) do not reject $H_0 : \sigma_A^2 = 0$ at level 0.05 since $F = 5.46 < 6.94 = F_{.05,2,4}$.
d) do not reject H_0 : All $\alpha_i = 0$ at level 0.05 since $F = 5.46 < 6.94 = F_{.05,2,4}$.
e) not have enough information to come to a conclusion.

[21] The estimated component of variance due to plants is

a) 0.7589
b) 0.4880
c) 0.2685
d) 0.0438
e) 0.0146

Stability data were generated for 2-mL vials manufactured with 30 mg/mL of active ingredient of a drug product. The factors in the study were time of storage (1, 3, 6 and 9 months) and two temperatures (30° C, 40°) at which the vials were stored. Three vials were analyzed in a laboratory at each time and temperature combination. The response was the pH measure in the vial at the time of the analysis. Use the SAS output on page 8 to answer the following questions.

Answer the next 3 questions based on these data.

[22] This is a two-factor experiment with

a) both time and temperature fixed effects.
b) time a random factor and temperature a fixed effect.
c) time a fixed effect and temperature a random effect.
d) both time and temperature random effects.
e) both vial and time fixed effects.

[23] Which is the appropriate value of the test statistic and conclusion for the test of interaction at level 0.05?

a) $F = 3.93$ $\Rightarrow$ Time and temperature interact significantly.
b) $F = 4.62$ $\Rightarrow$ Time and temperature interact significantly.
c) $F = 5.34$ $\Rightarrow$ Time and temperature interact significantly.
d) $F = 4.15$ $\Rightarrow$ Time and temperature interact significantly.
e) $F = 4.05$ $\Rightarrow$ Time and temperature interact significantly.

[24] Which is the appropriate interpretation of the means plot?

a) The pH level remained fairly constant for the vials stored at 30° C. but tended to decrease for the vials stored at 40° C.
b) The pH level remained fairly constant for the vials stored at 40° C. but tended to decrease for the vials stored at 30° C.
c) The pH level tended to increase for the vials stored at 30° C. but tended to decrease for the vials stored at 40° C.
d) The pH level tended to increase for the vials stored at 40° C. but tended to decrease for the vials stored at 30° C.
e) There is no clear-cut pattern in the means.

[25] Let $X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$ where $\sum \alpha_i = \sum \beta_j = 0$ be the model used to represent a randomized complete block design. The estimator of α_i is

a) MSE.
b) $\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot}$
c) $\bar{X}_{i\cdot}$
d) $\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}$
e) $\bar{X}_{\cdot\cdot}$

SAS Output for Drug Stability Data

```
Class Level Information
Class    Levels    Values
TIME          4    1 3 6 9
TEMP          2    30 40

Number of observations in data set = 24

General Linear Models Procedure

Dependent Variable: PH

Source                DF            Sum of Squares               Mean Square   F Value     Pr > F
Model                  7                0.07460000                0.01065714      4.62     0.0054
Error                 16                0.03693333                0.00230833
Corrected Total       23                0.11153333


Source                DF               Type III SS               Mean Square   F Value     Pr > F
TIME                   3                0.03696667                0.01232222      5.34     0.0097
TEMP                   1                0.01041667                0.01041667      4.51     0.0496
TIME*TEMP              3                0.02721667                0.00907222      3.93     0.0281

Tukey's Studentized Range (HSD) Test for variable: PH

Alpha= 0.05  df= 16  MSE= 0.002308
Critical Value of Studentized Range= 4.046
Minimum Significant Difference= 0.0794

Means with the same letter are not significantly different.

Tukey Grouping             Mean      N  TIME

                 A           3.80333      6  1
                 A
         B       A           3.77000      6  3
         B       A
         B       A           3.75833      6  6
         B
         B                   3.69500      6  9

Alpha= 0.05  df= 16  MSE= 0.002308
Critical Value of Studentized Range= 2.998
Minimum Significant Difference= 0.0416

Means with the same letter are not significantly different.

Tukey Grouping             Mean      N  TEMP

                 A           3.77750     12  30

                 B           3.73583     12  40


Level of   Level of       --------------PH-------------
TIME       TEMP       N     Mean              SD
1          30         3   3.77000000      0.06082763
1          40         3   3.83666667      0.03511885
3          30         3   3.79000000      0.03605551
3          40         3   3.75000000      0.05000000
6          30         3   3.80000000      0.04000000
6          40         3   3.71666667      0.07637626
9          30         3   3.75000000      0.02645751
9          40         3   3.64000000      0.04000000
```