**MATH 560 MidTerm** **Department of Mathematics and Statistics KFUPM**

Name: ----------ID: ----------

**Q1. (15 points).**

## Circle the correct option.

**1**.

In the ANOVA, treatment refers to
a. experimental units
b. different levels of a factor
c. a factor
d. blocks

**2**.

In designing an experiment, blocking is used
a. to reduce bias.
b. to reduce variation.
c. as a substitute for a control group.
d. as a first step in randomization.

**3.**

An experimental design that permits statistical conclusions about two or more factors is termed as

a. randomized block design
b. factorial design
c. completely randomized design
d. randomized design

**4.**

The number of times each experimental condition is observed in a factorial setup is known as
a. factor
b. replications
c. Treatment.
d. experimental units

**5.**

The required condition for using ANOVA procedure on data from several populations is that the
a. the selected samples are dependent on each other
b. sampled populations have equal variances
c. samples populations are uniform
d. sampled populations have equal means

**6.**

The non-parametric tests, as compared to parametric test, are often
a. more efficient, more practical
b. less efficient, more practical
c. equally efficient, equally practical
d. none

**7.**

The Kruskal-Wallis test uses the information on which of the following scales of measurement
a. Ratio
b. Ordinal
c. interval
d. nominal

**8**.
In experimental studies, the covariate effect is
a. controllable but not measureable
b. measureable and controllable
c. measureable but not controllable
d. neither measureable nor controllable

**9**.
In general, an **F-statistic** is a ratio of two quantities that are expected to be roughly --------- under the null hypothesis
a. one
b. unequal
c. equal
d. zero

**10**.
If the value of the Friedman test statistics is very large, what will be its approximate p-value?
a. -1
b. 0.5
c. zero
d. one

**11**.
The non-parametric alternative to two independent sample t-test is:
a. Sign test
b. Friedman test
c. Mann-Whitney test
d. Wilcoxon test

**12**.
Which of the following multiple comparison tests control individual error rate:
a. Tukey
b. Duncan
c. Fisher
d. None

**13**.
Which of the following is an unbiased estimator of $\sigma^2$:
a. MST
b. MSB
c. MSE
d. SSE

**14**.
For a 3x4 factorial experiment conducted under RCBD using 4 blocks, the error degrees of freedom will be:
a. 48
b. 44
c. 33
d. 36

**15**.
In order to secure a more accurate estimate of experimental error, which of the following basic principle plays the most vital role:
a. randomization
b. blocking
c. replication
d. none

Name: ----------ID: ----------

**Q2. (10 points).**

Four different formulations of an industrial glue are being tested. The tensile strength of the glue when it is applied to join parts is also related to the application thickness. Five observations on strength ($y$) in pounds and thickness ($x$) in 0.01 inches are obtained for each formulation. The data are shown in the following table. Analyze these data and draw appropriate conclusions.

| | | | | Glue Formulation | | | |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 2 | 3 | 3 | 4 | 4 |
| *y* | *x* | *y* | *x* | *y* | *x* | *y* | *x* |
| 46.5 | 13 | 48.7 | 12 | 46.3 | 15 | 44.7 | 16 |
| 45.9 | 14 | 49.0 | 10 | 47.1 | 14 | 43.0 | 15 |
| 49.8 | 12 | 50.1 | 11 | 48.9 | 11 | 51.0 | 10 |
| 46.1 | 12 | 48.5 | 12 | 48.2 | 11 | 48.1 | 12 |
| 44.3 | 14 | 45.2 | 14 | 50.3 | 10 | 48.6 | 11 |

Using this information, we carried out MINITAB analysis and the output is given below. Some part of the output is missing that you are required to fill in based on your DOE knowledge. (an estimated p-value is needed).

```
Analysis of Variance for Strength, using Adjusted SS for Tests


Source     DF     Seq SS     Adj SS     Adj MS       F          P

Thick    -------  68.852     59.566     ---------   ------      -----

Glue     -------  1.771      1.771      ---------   -------     0.740

Error    -------  -------    20.962     ---------

Total    -------  91.585
```

**Q3. (10 points).**

Complete the following ANOVA tables, showing the partitioning of degrees of freedom for different DOE:

i). Latin Square Design (single factor experiment with t treatments),

| **Source of Variation** | ***df*** |
|---|---|
| Rows | |
| Cols | |
| Treatments | |
| Error | |
| Total | |

ii). Graeco Latin Square Design (single factor experiment with t treatments),

| **Source of Variation** | ***df*** |
|---|---|
| Rows | |
| Columns | |
| Greeks | |
| Treatments | |
| Error | |
| Total | |

iii). Randomized complete block design (Two factor factorial experiment with “a” levels of Factor A, “b” levels of Facor B, and “m” blocks),

| **Source of Variation** | ***df*** |
|---|---|
| A | |
| B | |
| AB | |
| Blocks | |
| Error | |
| Total | |

Name: ----------ID: ----------

**Q4. (5+5=10 points).**

a). Write the statistical model of a single factor experiment for completely randomized design. Provide a brief outline of the derivation of the least square estimates for the parameters by clearly mentioning the constraints.

b). Write the statistical model of a single factor experiment for Latin square design. Provide a brief outline of the derivation of the expected value of MSE.

Name: ----------ID: ----------

**Q5. (7+5+2+6=20 points).**

Three different washing solutions are being compared to study their effectiveness in retarding bacteria growth in 5-gallon milk containers. The analysis is done in a laboratory, and only three trials can be run on any day. Because days could represent a potential source of variability, the experimenter decides to use a randomized block design. Observations are taken for four days, and the data are shown here.

| | Days | | | |
|---|---|---|---|---|
| **Solution** | **1** | **2** | **3** | **4** |
| 1 | 13 | 22 | 18 | 39 |
| 2 | 16 | 24 | 17 | 44 |
| 3 | 5 | 4 | 1 | 22 |

i). Using an appropriate parametric technique, test the hypothesis that the average effectiveness of the three washing solutions is same (at $\alpha = 0.05$). (Provide only the required information as outlined below).

**Hypotheses: H0:** ----------

**H1:** ----------

**Name of the test:** ----------

**Sampling Distribution of test:** ----------

**Value of the test statistic:** ----------

**p-value:** ----------

**Decision:** ----------

**Conclusion:** ----------

---

---

---

ii). Based on the decision of part (i), carry out pair wise comparisons (if needed) such that the individual error rates are 0.05. Provide the name of the pair wise comparison test and indicate the pairs that are significantly different.

**Name of the pair wise comparison test:** ----------

**Significance of the pair wise comparisons:**

iii). Based on the results of part (ii), what is the overall error rate?

iii). Perform residual analysis and comment on the assumptions of random error term.

iv). Assuming non-normality, analyze the data from this experiment (at $\alpha = 0.05$) using an appropriate non-parametric technique. (Provide only the required information as outlined below).

**Hypotheses: H0:** ----------

**H1:** ----------

**Name of the test:** ----------

**Value of the test statistic:** ----------

**p-value:** ----------

**Decision:** ----------