Name: Akond Rahman
Lab: 001B
HW#1

Question-1

Researchers ran a quality control study to evaluate the quality of plastic irrigation pipes under different conditions of use. The goal of the study was to compare compressive strength of the pipes under a given water temperature (L and H) and hardener type (H1, H2, H3). The study design involved a total of 24 pipes, with 12 pipes randomly selected from each of two manufacturing plants (M1, M2). For the 12 pipes from each plant, the combinations of water and hardness were each assigned to two of the pipes.

a. Identify the response and explanatory variables that were measured in the study.

Answer

The response variable was compressive strengths of the pipes. The explanatory variables were water temperature and hardener type.

b. Identify the factors

Answer

Water temperature and hardener type

c. Identify the levels of factors.

Answer

For water temperature the levels were L and H. For hardener type the levels were H1, H2, and H3.

d. Identify the treatments and state the total number of treatments.

Answer

The treatments were water temperature H, water temperature H, hardener type H1, hardener type H2, and hardener type H3. Total numbers of treatments were: six.

e. Identify the confounding variables.

Answer

None.

f. Identify the objects that are experimental units.

Answer

The pipes used in the study.

Question-2

Researchers record the ohms passing through 6 different types of resistors. For each resistor type, 4 replications were done. The overall goal was to determine if the mean number of ohms was the same for the different types of resistors.

a. A model for this experiment is $Y_{ij} = \mu_i + \varepsilon_{ij}$, $\varepsilon_{ij} \sim^{iid} N(0,\sigma^2)$ Describe the following terms: Y_{ij}, μ_i, ε_{ij}, σ^2

Answer

Y_{ij} : the j^{th} observation selected from treatment group i. For example, y_{34} denotes the fourth observation drawn from treatment group three.

μ_i : the mean of all observations in group i. For example, μ_5 presents the mean of all observations in treatment group five

ε_{ij} : the deviation of y_{ij} about the mean of group i that is u_i. For example, ε_{12} presents the deviation of y_{12} about the mean of group i, μ_i

σ^2 : The variances of each treatment group in the study. In ANOVA, the variances of each group i must be equal.

b. Output from SAS is given below. Some portions of the ANOVA table are missing. Recreate the ANOVA table with these values filled in

Answer

Source	DF	Sum of Squares	Mean Square	F-Value	Pr > F
Model	**5**	**15.75052**	**3.15010**	**0.35139**	0.8746
Error	18	161.35994	8.96444		
Corrected Total	23	177.11046			

c. Sum of Squares for Treatment. What does this quantity represent?

Answer

This quantity measures the variability between or among the mean number of ohms for each resistor type

d. What does the value of Root MSE estimate?

Answer

This quantity measures the square root of the Mean Square Error and defines the standard deviation of an observation about the predicted value. Root MSE gives an idea of the overall variance in the data.

e. Conduct a hypothesis test at the 0.01 significance level to determine if the mean yields are plausibly different. Be sure to state the null and alternative hypotheses, the assumptions made on the model, a test statistic, either a rejection region or p-value with conclusion, and an interpretation of the conclusion in the context of the problem.

Answer

Hypotheses for parameter of interest
NULL Hypothesis, H_0 = mu_1 = mu_2 = mu_3 = mu_4 = mu_5 = mu_6
Alt. Hypothesis, H_A = at least one of mu_1 , mu_4, mu_3, mu_4, mu_5, mu_6 is different

Assumptions made on the model:
i. single, continuous response variable obtained
from a simple random sample, OR, a completely
randomized design is used
ii. Within each treatment group the residuals are
homoscedastic, independent, and normally distributed

The test statistic is: 0.35139

The p-value: 0.8746

Conclusion: As p-value > 0.01 (the significance level), we fail to reject the null hypothesis. We cannot state that there is any statistical evidence to suggest that
at least one of the mean ohms of different six types of resistor types is different.

f. Comment on the assumption of homogeneity of variances.

Answer

From the 'Residual vs. Predicted Value' plot we observe that except for five observations, all other residuals fit within the range (-2.5, 2.5). Overall, the study maintains the assumptions of homogeneity of variances, but while drawing conclusions the five residuals that are out of the range should be kept in account.

g. Comment on the assumption of normality on the errors.

Answer

We observe that more than 10% of the total observations not fitting the straight line in the 'Quantile vs. Residual' plot. We cannot claim that all data points in the dataset satisfy the condition of Normality. While drawing conclusions we should keep this fact in mind. From the 'Percent vs. Residual' plot we observe that the overall data distribution is normal.

h. Your friend reads your notes and says ANOVA stands for Analysis of Variance, but your hypothesis test is about means. That doesn't make sense! Explain why we call it Analysis of Variance.

Answer

In analysis of variance, all differences in sample means are judged statistically significant (or not) by comparing them to the variation within samples. Hypotheses test might be about means but ANOVA still measures the variances within the sample groups, as well as, between the sample groups. The ANOVA test is the first step to identify factors that can influence a certain variability of interest. After performing the ANOVA analysis we can perform further analysis on factors that are statistically contributing to the variability.