(1, 4.5pts) **(a)** The experiment includes 2 treatment factors, *brand* of fuel and *amount* of ethanol.

(b) *Brand* is a qualitative factor with 4 levels, and *amount* is a quantitative factor with 3 equally-spaced levels.

(c) a fixed portion of standard fuel

(d) An acceptable conceptual model is

$$response \sim \underbrace{brand}_{3 \text{ d.f.}} + \underbrace{amount}_{2 \text{ d.f.}} + \underbrace{brand : amount}_{6 \text{ d.f.}} + \underbrace{\epsilon}_{12 \text{ d.f.}}$$

where *brand* should use dummy variables to define its effects, and *amount* should use linear and quadratic codings to define its effects. The interactions $brand : amount$ can be coded as the component-wise product of *brand* and *amount* main-effect codings.

(e) This is a completely randomized design with two-way layout.

(2, 4.5pts) **(a)** The experiment includes one treatment factor: *drug* (or *formulation*), and two block factors: *subject* and *time* (i.e., the number of times taking a drug).

(b) *Drug* and *subject* are qualitative, each with 3 levels. *Time* is a quantitative factor with 3 levels. However, because *time* is treated as a block factor, its effects are merely used to remove blocks' variation from the error. Therefore, *time* can be considered as qualitative in the modeling and analysis.

(c) a subject taking a single dose of drug

(d) An acceptable conceptual model is

$$response \sim \underbrace{subject}_{2 \text{ d.f.}} + \underbrace{time}_{2 \text{ d.f.}} + \underbrace{drug}_{2 \text{ d.f.}} + \underbrace{\epsilon}_{2 \text{ d.f.}},$$

where *subject*, *time*, and *drug* represent the main-effect codings of the three factors.

(e) Latin-square design

(3, 1pt) It is because of the ties in the data in diets C.

(4, 1.5pts) Outliers, skewness showing a lack of normality, and equality of variance in these groups.

(5, 1pt) Diet D has a different number of replicates compared to diets B and C.

(6, 2pts) Let $\hat{\mu}_1, \ldots, \hat{\mu}_4$ denote the estimated mean response for diets A, ..., D, respectively. Under treatment coding, we have $\hat{\mu}_1 = 61$, $\hat{\mu}_2 - \hat{\mu}_1 = 5$, $\hat{\mu}_3 - \hat{\mu}_1 = 7$, and $\hat{\mu}_4 - \hat{\mu}_1 = 0$. Additionally, $var(\hat{\mu}_i - \hat{\mu}_1) = var(\hat{\mu}_i) + var(\hat{\mu}_1)$, $i = 2, 3, 4$, as observations are independent of each other. Because the standard error of dietD is smaller than that of dietB (or dietC), diet D has more replicates than diet B (or diet C). Also,

$$var(\hat{\mu}_2) = var(\hat{\mu}_2 - \hat{\mu}_1) - var(\hat{\mu}_1) = 1.5275^2 - 1.1832^2 = 0.9324 < 1.1832^2 = var(\hat{\mu}_1).$$

Similarly $var(\hat{\mu}_3) < var(\hat{\mu}_1)$. As replicates decrease, variance increases; therefore, diet A has the smallest number of replicates.

(7, 1pt) The test statistic is 13.6, and the sum of squares for residual is $2.3664^2 \times 20 = 112$.

(8, 1pt) Following the answer to question (6), we find that $\hat{\mu}_1 = 61$, $\hat{\mu}_2 = 66$, $\hat{\mu}_3 = 68$, and $\hat{\mu}_4 = 61$. Therefore, the answer is

$$68 - \frac{61 + 66 + 68 + 61}{4} = 4.$$

(9, 1pt) Only the A-D and B-C differences are not significant, as the corresponding intervals contain zero.

(10, 1pt) We can use B-A to illustrate. The adjusted p-value of the Bonferroni method is $0.00380 \times 6 = 0.0228$, which is greater than that of the Tukey method, 0.01833. Because both methods ensure that the overall Type-I error probability is controlled below 0.05, the test with the smaller p-value (i.e., the Tukey method) will have higher power.

(11, 1pt) Two

(12, 1pt) All the treatment contrasts (i.e., the parwise comparisons of treatment means) are estimated with the same variance, so that their confidence intervals are all the same length. FYI, BIBD also tend to give the shortest confidence intervals on the average for any large number of contrasts.

(13, 1pt) When all main effects of treatments and blocks are included in the model, there are not enough degrees of freedom to explore all interactions. The interactions require $9 \times 5 = 45$ degrees of freedom, where there are only 15 degrees of freedom left in residuals.

(14, 1pt) In BIBD, due to incomplete blocking, the spaces spanned by treatment effects and block effects are not orthogonal with each other. In sequential ANOVA, when effects are not orthogonal to each other, the difference in the orders in which effects enter the model will result in different ANOVA tables.

(15, 1pt) The null model is: gain ~ 1, and the alternative model is: gain $\sim 1 +$ treat.

(16, 3pts) The F-value is

$$\frac{293.38/5}{(595.74 + 150.77)/(9 + 15)} = \frac{293.38/5}{746.51/24} = \frac{58.676}{31.105} = 1.8864,$$

which is much smaller (and actually insignificant) than the F-values in the two ANOVA tables containing block. This is because block is highly significant, meaning that the between-block variation is much larger than the within-block variation, indicating successful blocking. When block is not included in the model, its effect enters the error term, causing an overestimation of $\hat{\sigma}^2 = 31.105$ and resulting in a much smaller F-statistic.

(17, 1.5pts) In the one-way ANOVA, $\hat{\sigma}^2 = (595.74 + 150.77)/(9 + 15) = 31.105$, while in the ANOVA tables for BIBD, $\hat{\sigma}^2 = 10.052$. The relative efficiency is $31.105/10.052 = 3.0944$, much larger than 1.

(18, 2pts) A BIBD must satisfy two conditions: (i) $rt = bk$, and (ii) $\lambda(t-1) = r(k-1)$. Here, we have $t = 6$, $b = 12$, and $k = 3$. We can set $r = 6$ to make (i) hold. But, by (ii), we get $\lambda = r(k-1)/(t-1) = (6 \times 2)/5 = 2.4$, which is not an integer. So, BIBD does not exist in this case.