

Statistics 514
Review for Second Midterm Exam

1. Answer the following questions. Please use complete sentences.

- (a) The inspector general is coming, and an officer wishes to arrange some soldiers for inspection. In the officer's command are men and women of three different ranks, who come from six different states. The officer is trying to arrange 36 soldiers for inspection in a six-by-six square with one soldier from each state-rank-gender combination. Furthermore, the idea is to arrange soldiers so that no matter which rank or file (row or column) is inspected by the general, the general will see someone from each of the six states, one woman from each rank, and one man from each rank. Why is this officer so frustrated?
- (b) If a log transformation is desirable, what trend would be expected in the untransformed data?
- (c) A Latin square design is run with $a = 4$ treatments. Suppose $MS_E = 20$ and there is interest in comparing the contrast $\mu_1 - 2\mu_2 + 2\mu_3 - \mu_4$ to zero. Which of the following are true?
 - i. The degrees of freedom for the overall F -test are 1 and 7.
 - ii. The standard error of the grand mean is $\sqrt{12.5}$.
 - iii. The degrees of freedom for a t -test on the contrast is 6.
 - iv. The standard error of the contrast is $\sqrt{50}$.

2. Short answer.

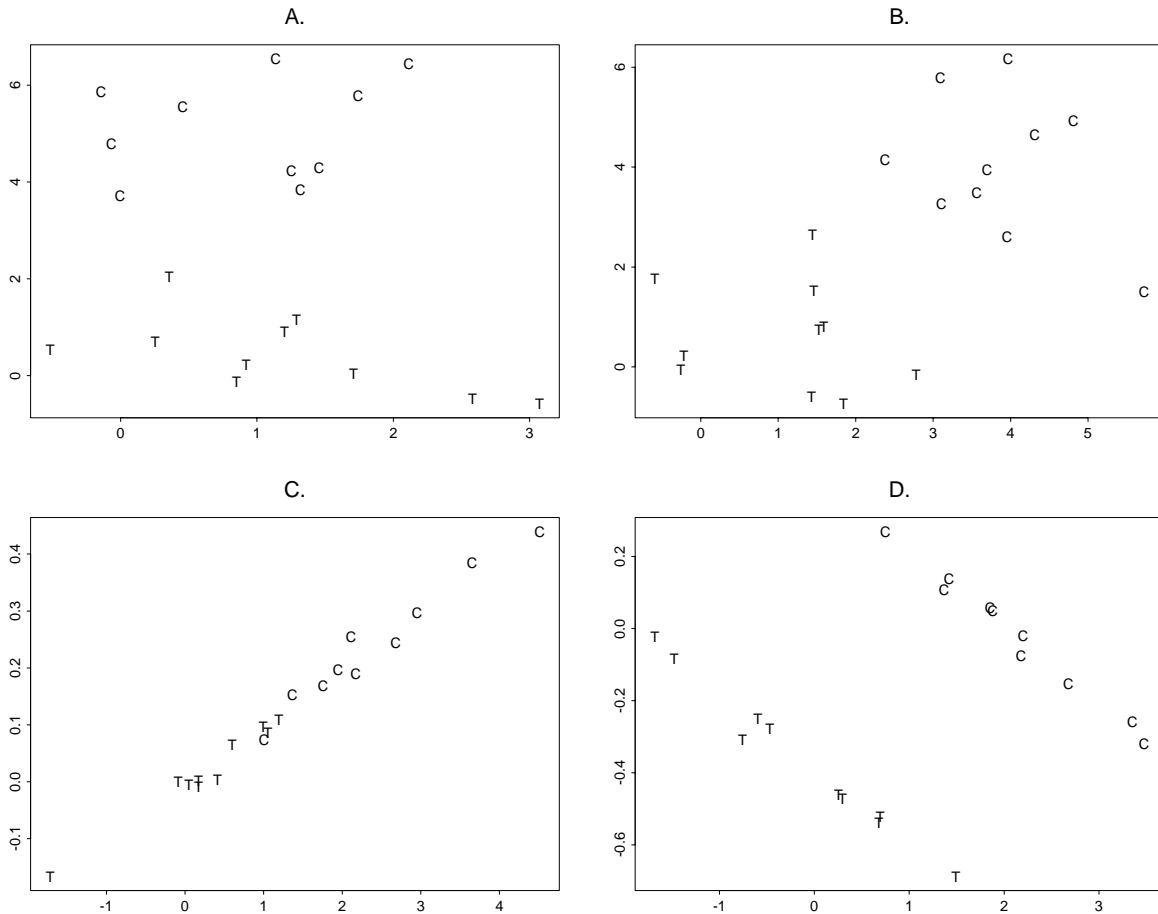
- (a) When using a pretest score, one can include the score as a covariate in an ANCOVA analysis, or one can do a "gains analysis" – using the difference in the test scores as the response. It can be shown that if the response variances are equal, then the error variances for ANCOVA and gain score analysis are

$$\begin{aligned}\sigma_{ANCOVA}^2 &= \sigma^2(1 - \rho^2) \\ \sigma_{gains}^2 &= \sigma^2 2(1 - \rho),\end{aligned}$$

where ρ is the population correlation within groups between the pretest and the posttest.

- i. For what value(s) of ρ is ANCOVA most powerful? Least powerful?
 - ii. For which value(s) of ρ is ANCOVA more powerful than gains analysis?
 - iii. For which value(s) of ρ is gains analysis more powerful than simple one-way ANOVA analysis (without the pretest as a covariate)?
- (b) Each scatterplot below is for a data set with the ANCOVA factor structure. The points in the treatment group are plotted as T 's, those in the control group are plotted as C 's. Match each scatterplot with one of the following verbal descriptions.
 - i. Response and covariate show strong linear relationships within groups, and the two groups have very different sets of covariate values. Covariance adjustment will drastically reduce the residual sum of squares and substantially increase the estimated treatment differences.

- ii. The response and covariate show a strong positive relationship *between* groups, but no linear relationship within groups. Fitted parallel lines would be pretty much horizontal: ANCOVA and ordinary ANOVA would be essentially equivalent.
- iii. The response and covariate show a strong positive relationship. The unadjusted SD will be a lot larger than the adjusted SD. Although ordinary ANOVA would find fairly substantial treatment effects, the group differences “go away” if you adjust for the covariate.
- iv. The response and covariate show almost no linear relationship. There are big treatment differences, whether you adjust for the covariate or not. The SD is pretty much unaffected by the covariance adjustment.



- (c) Birth weights in pounds of five pigs were recorded from each of six randomly chosen litters. If the between litters sums of squares is 6.0 and the within litters sum of squares is 9.6, what is the percentage of total variance that can be attributed to litter-to-litter differences? What is the ANOVA estimate of the variance component for litters?
3. For each of the following, describe the experimental design used and give a skeleton ANOVA (sources and degrees of freedom only).
- (a) A grocery store chain is experimenting with its weekly advertising, trying to decide among cents-off coupons, regular merchandise sales, and special-purchase merchandise sales. There are two cities about 100 km apart in which the chain operates, and the chain will always run one advertisement in each city on Wednesday, with the offer good

for 1 week. The response of interest is total sales in each city, and large city-to-city differences in total sales are expected due to population differences. Furthermore, week-to-week differences are expected. The chain runs the experiment on 12 consecutive weeks, randomizing the assignment of advertising method to each city, subject to the restrictions that each of the three methods is used eight times, four times in each city, and each of the three pairs of methods is used an equal number of times.

- (b) Metribuzin is an agricultural chemical that may accumulate in soils. To see if the amount retained in the soil depends on the amount applied to the soil, the following experiment was performed. Three different levels of metribuzin were equally and randomly assigned to a total of 24 plots. After one growing season, the amount of metribuzin in the first three cm of soil was measured. The pH of the soil in each plot was also measured as pH may affect the ability of the soil to retain metribuzin.