Test 2

STAT218

2024-03-02

### Instructions

You have 48 hours from the release of this assignment to complete and submit your work. You may refer to all class materials, notes, and textbooks, but **must complete this assignment on your own**. By submitting your work, you are affirming that your work is your own and you have not consulted with anyone else in preparing your answers. Please use the word document provided (download from the class website) and write in your answers below each prompt. Please ensure that question numbering is preserved in your document. Submit your work via Gradescope.

The test comprises two parts: concepts, containing short-answer and multiple-choice questions; and applications, which requires some data analysis. Revisions will be allowed for the applications part, but not for the concepts part.

## Part I: Concepts

1. [L9] In an analysis of variance, the statistic measures:
   1. the variability between groups relative to the variability within groups
   2. the proportion of total variability due to differences in means
   3. the statistic isn’t directly interpretable
   4. the ratio of unexplained variability to total variability
   5. the sum of the squared differences in group means for all pairs of groups

* Answer: A.

1. [L4] A 95% confidence interval for a population mean is given as (18.985, 21.015). This confidence interval is based on a simple random sample of 36 observations. Calculate the sample mean and standard deviation. Assume that all conditions necessary for inference are satisfied. Use the t-distribution in any calculations.

* Answer: the sample mean is the midpoint, and the standard deviation is half the width, multiplied by , where is a -critical value.

# sample mean  
(21.015 + 18.985)/2

[1] 20

# sample SD  
6\*(21.015 - 18.985)/(2\*qt(0.975, df = 34))

[1] 2.996687

1. [L5] In each of the following scenarios, determine if the data are paired.
   1. Compare pre- (beginning of semester) and post-test (end of semester) scores of students.
   2. Assess gender-related salary gap by comparing salaries of randomly sampled men and women.
   3. Compare artery thicknesses at the beginning of a study and after 2 years of taking Vitamin E for the same group of patients.
   4. Assess effectiveness of a diet regimen by comparing the before and after weights of subjects.

* Answer:
  1. Paired
  2. Not paired
  3. Paired
  4. Paired

1. [L9] In post-hoc significance tests for pairwise comparisons between groups and , ordinarily the hypotheses tested are…
   1. vs.
   2. vs.
   3. vs.
   4. vs.

* Answer: C

1. [L9] Determine if the following statements are true or false in ANOVA, and explain your reasoning for statements you identify as false.
   1. As the number of groups increases, the modified significance level for pairwise tests increases as well.
   2. As the total sample size increases, the error degrees of freedom increases as well.
   3. As the number of groups increases, the error degrees of freedom increases as well, provided the sample size remains fixed.

* Answer:
  1. False. The tests become more stringent, so significance level decreases.
  2. True. Error df are , which increases in .
  3. False. Error df are , which decreases in .

1. [L9] For post-hoc comparisons relative to a control, the most appropriate multiple inference correction method is…
   1. Scheffe’s method
   2. Bonferroni correction
   3. Tukey’s method
   4. Sidak’s method
   5. Dunnett’s method

* Answer: E.

1. [L5] Suppose you conducted a hypothesis test to determine whether mean ice thickness differs significantly between two nearby creeks during a two-week period in winter, and obtained a -value of 0.37. Select the correct interpretation of this result.
   1. With 95% confidence, there is no difference in mean ice thickness between the two creeks.
   2. The data do not provide sufficient evidence to reject the null hypothesis that mean ice thickness differs between the two creeks.
   3. The data indicate that mean ice thickness is the same between the two creeks.
   4. The data provide sufficient evidence of a difference in mean ice thickness between the two creeks.
   5. The mean ice thickness is estimated to be at least 0.37.

* Answer: B.

## Part II: Applications

1. [L4] The tuition dataset contains in-state and out-of-state tuition at 25 public universities from 2011-2012. Calculate and interpret a 95% confidence interval for the mean difference between in-state and out-of-state tuition at U.S. public universities in 2011-2012.

With 95% confidence, the mean out-of-state 2011-2012 tuition at U.S. public universities is estimated to be between $7687.08 and $1.038972^{4} more expensive than mean in-state tuition.

1. [L9] The anorexia dataset contains observations of weight change after therapeutic treatment for 72 female anorexia patients. Each patient was prescribed one of two therapies, cognitive behavioral therapy (CBT) and family therapy (FT), or assigned to a control group. Treatment allocations were random. Quantify the effect of each therapy relative to the control by estimating the difference in mean weight change between each of the two treatment groups and the control group. Provide simultaneous 95% confidence intervals for each difference.

Numerical answers are sufficient here; a plot is also acceptable.

| contrast | estimate | SE | df | lower.CL | upper.CL |
| --- | --- | --- | --- | --- | --- |
| CBT - Cont | 3.457 | 2.033 | 69 | -1.157 | 8.071 |
| FT - Cont | 7.715 | 2.348 | 69 | 2.386 | 13.04 |

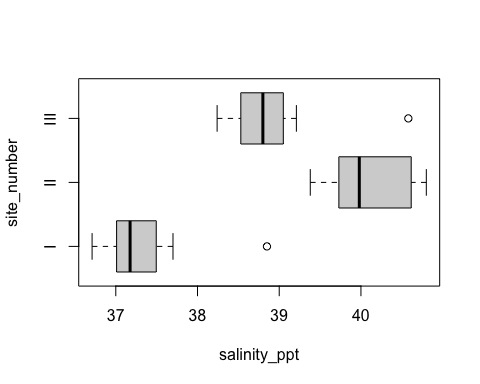
1. [L4, L5] The dataset lizards contains running speeds measured in a laboratory race track for two species of lizards, Western Fence (WF) and Sagebrush (S) lizards. Speeds are recorded in meters per second. Compute point estimates and standard errors for the mean top speed for each species, and test for a difference in mean top speed between species at significance level 0.01. Write a narrative summary of the test result and, if a difference is identified, construct and interpret a 99% confidence interval for the difference in mean top speed.

| type | mean | mean.se |
| --- | --- | --- |
| WF | 1.613 | 0.06356 |
| S | 2.315 | 0.1184 |

Welch Two Sample t-test  
  
data: top.speed by type  
t = -5.2217, df = 32.57, p-value = 9.939e-06  
alternative hypothesis: true difference in means between group WF and group S is not equal to 0  
99 percent confidence interval:  
 -1.0695322 -0.3341741  
sample estimates:  
mean in group WF mean in group S   
 1.612692 2.314545

Answer: Results suggest that sagebrush lizards run faster than western fence lizards. We tested the hypothesis that mean top speed differs between the two species; the data provide sufficiently strong evidence to reject the null hypothesis of no difference in mean speed in favor of the alternative of a difference (\*T = -5.2217\* on 32.57 degrees of freedom, \*p = 0.00000994\*). With 99% confidence, the mean top speed for sagebrush lizards is estimated to be between 0.33 and 1.07 meters per second faster than the mean top speed for western fence lizards.

1. [L9] The salinity dataset contains replicate measurements of salinity in parts per thousand for three water masses in Bahamas. Is the mean salinity different among the sites, and if so, how? Carry out a full analysis, and provide graphical and tabular data summaries, inference and estimates for site means, and comparisons between sites. Write a short summary of the results, interpreting inferences in context using technically appropriate language.



| site\_number | mean.salinity | mean.salinity.se |
| --- | --- | --- |
| I | 37.31 | 0.1654 |
| II | 40.1 | 0.1879 |
| III | 38.94 | 0.2044 |

Analysis of Variance Model

|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| --- | --- | --- | --- | --- | --- |
| **site\_number** | 2 | 39.13 | 19.56 | 56.51 | 2.238e-10 |
| **Residuals** | 27 | 9.347 | 0.3462 | NA | NA |

$emmeans  
 site\_number emmean SE df lower.CL upper.CL  
 I 37.3 0.170 27 37.0 37.7  
 II 40.1 0.208 27 39.7 40.5  
 III 38.9 0.186 27 38.6 39.3  
  
Confidence level used: 0.95   
  
$contrasts  
 contrast estimate SE df lower.CL upper.CL  
 I - II -2.79 0.269 27 -3.46 -2.12  
 I - III -1.63 0.252 27 -2.25 -1.00  
 II - III 1.16 0.279 27 0.47 1.85  
  
Confidence level used: 0.95   
Conf-level adjustment: tukey method for comparing a family of 3 estimates

Results suggest a difference in salinity between sites. The data provide sufficient evidence to reject the null hypothesis of no difference in mean salninity among the sites (\*F = 56.51\* on 2 and 27 degrees of freedom, \*p < 0.0001\*). Site II had the highest estimated mean salinity, followed by site III, followed by site I. Post-hoc comparisons indicated that all three sites differ significantly from one another.

1. [L9] In the FAMuSS study, test for a difference in mean percent change in nondominant arm strength *and* in dominant arm strength according to genotype and estimate means of each variable for each group. If differences are identified for either response variable, determine which groups differ. Conduct significance tests at the 0.05 significance level and interpret the results of the test. Briefly, discuss the strength of evidence associated with your findings. If differences were identified, how convincing are the results?

ANOVA table for nondominant arm strengths

|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| --- | --- | --- | --- | --- | --- |
| **actn3.r577x** | 2 | 7043 | 3522 | 3.231 | 0.04022 |
| **Residuals** | 592 | 645293 | 1090 | NA | NA |

ANOVA table for dominant arm strengths

|  | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
| --- | --- | --- | --- | --- | --- |
| **actn3.r577x** | 2 | 2040 | 1020 | 3.157 | 0.04328 |
| **Residuals** | 592 | 191265 | 323.1 | NA | NA |

Pairwise comparisons for nondominant arm strength

| contrast | estimate | SE | df | t.ratio | p.value |
| --- | --- | --- | --- | --- | --- |
| CC - CT | -4.355 | 3.237 | 592 | -1.345 | 0.3706 |
| CC - TT | -9.19 | 3.615 | 592 | -2.542 | 0.03029 |
| CT - TT | -4.835 | 3.309 | 592 | -1.461 | 0.3104 |

Pairwise comparisons for dominant arm strength

| contrast | estimate | SE | df | t.ratio | p.value |
| --- | --- | --- | --- | --- | --- |
| CC - CT | 2.226 | 1.762 | 592 | 1.263 | 0.4167 |
| CC - TT | -2.264 | 1.968 | 592 | -1.15 | 0.4835 |
| CT - TT | -4.49 | 1.801 | 592 | -2.493 | 0.03454 |

Differences were identified for both variables; for nondominant arm strength, the data suggest mean change in strength differs between genotypes CC and TT; for dominant arm strength, the data suggest mean change in strength differs between genotypes CT and TT. Significance tests for each of these findings produced only moderate evidence; in particular, the ANOVA tests both yielded $p$-values around 0.04. Considering these were not adjusted for multiplicity, the evidence is rather weak.

1. [L5] The Sleuth3::case0101 dataset contains data from an experiment on the effect of intrinsic vs. extrinsic motivation on creativity. 47 creative writing students were randomly assigned to one of two groups, extrinsic and intrinsic; each subject was instructed to write two short poems, but those in the extrinsic motivation group were primed on the task in a way that oriented them to external motivations for writing, and those in the intrinsic group were primed on the task in a way that oriented them to internal motivations for writing. Poems were scored by judges for creativity on a 40-point scale, and each subject received an average score. Test whether the framing led to differences in creativity; if so, which type of motivation was more effective with respect to eliciting creativity?

Welch Two Sample t-test  
  
data: Score by Treatment  
t = -2.9153, df = 43.108, p-value = 0.005618  
alternative hypothesis: true difference in means between group Extrinsic and group Intrinsic is not equal to 0  
95 percent confidence interval:  
 -7.010803 -1.277603  
sample estimates:  
mean in group Extrinsic mean in group Intrinsic   
 15.73913 19.88333

The data provide sufficiently strong evidence to reject the null hypothesis of no difference in mean creativity scores between motivation-type groups (\*T = -2.9153\* on 43.11 degrees of freedom, \*p = 0.005618\*). With 95% confidence, the mean creativity score for the intrinsic group is estimated to be between 1.28 and 7.01 higher than the mean creativity score for the extrinsic group. This suggests intrinsic motivation improves creativity relative to extrinsic motivation.