

## Homework 1: due Monday, 30 January

Homework is due at the start of class, and we will discuss the homework problems as soon as they have been collected. Please photocopy your homework paper before class so that you'll be able to check your answers.

One objective of this class is that you learn to extract what you need from the output for tables and written results.

**Please do not attach output at the back of your homework as an appendix.**

Convert SAS output in your HW paper to Courier or Consolas font.

1. **Dietary study.** Clinical investigators recruited 134 healthy adults with moderately high cholesterol for a clinical trial to compare three diets: A, B, and Control, for 6 weeks. Primary endpoint was change in cholesterol and C-reactive protein from baseline (= pre-diet measurement) to study end.

Please download from the course website the data `hw1_CRP.csv` and the SAS program, `hw1_CRP.sas`, and save both on the desktop. Start PC-SAS, open `hw1_CRP.sas`, and run this program. Use the output to make these tables.

- a. Complete a table of baseline demographic and clinical means and standard deviations, with overall  $p$ -values from the  $F$ -test in the ANOVA table. For gender, get the  $p$ -value from the chi-square test.

	Diet A	Diet B	Control	$P$ value
Participant numbers				
Females $n$ (percent)				
Age (years)				
Weight (kg)				
Body Mass Index (BMI)				
Cholesterol (mg/dl)				
C-Reactive Protein (mg/l)				

- b. Complete a table of results for the two endpoints, cholesterol and C-reactive protein. Give the means and standard errors *based on the group standard deviations*. Use letters to show significant differences *between groups* within each row. Use an asterisk ( \* ) to show significant *within group* change from baseline.

Change from Baseline	Diet A	Diet B	Control
Cholesterol (mg/dl)			
C-Reactive Protein (mg/l)			

Values are mean  $\pm$  standard error. Pairwise comparisons are within rows: means that do not share letters are significantly different ( $p < 0.05$ ).

\*Significant within-group change from baseline ( $p < 0.05$ ).

2. **Weight Loss.** Two sentences, [a] and [b] below, appeared in a paper reporting on a weight loss trial:

[a] After 6 months on the study diets, the 100 participants in the placebo group had lost  $1.2 \pm 2$  kg (mean  $\pm$  standard error) while the 100 participants in the test-diet group had lost  $10.3 \pm 2$  kg.

Can you tell from [a] whether the two groups had significantly different mean weight loss at six months (use  $p < .05$ )?

[b] After 12 months on the study diets, the 100 participants in the test-diet group had lost 2.1 kg more than the 100 participants in the placebo group (95% confidence interval:  $-0.7$  to  $4.9$  kg).

Can you tell from [b] whether the test-diet group experienced significantly greater mean weight loss at 12 months than the placebo group ( $p < .05$ )?

3. **Newborns and cocaine.** Physicians in New York measured healthy newborn infants whose mothers may have used cocaine during pregnancy (source: *Pediatrics*, 2000; 3:106–111.) Infant’s exposure to cocaine was measured from cocaine metabolite in the mother’s hair, and infants were divided into three categories: High, Low, and No cocaine exposure. Below is part of the table comparing infants in these three categories.

TABLE 2. Maternal and Infant Characteristics According to Level of Cocaine Metabolite in Maternal Hair (Continuous Variables)

Variable	Level 0 No Cocaine <i>n</i> = 136 Mean $\pm$ SE	Level 1 Low Cocaine <i>n</i> = 52 Mean $\pm$ SE	Level 2 High Cocaine <i>n</i> = 52 Mean $\pm$ SE	95% CIs for Pairwise Differences Between Specified Means*
Infant characteristics				
Gestational age (wk)	39.0 $\pm$ .21	38.6 $\pm$ .32	38.6 $\pm$ .29	0–1 [–.50, 1.30] 0–2 [–.45, 1.35] 1–2 [–1.03, 1.14]
Birth weight (g)	3369 $\pm$ 40	3271 $\pm$ 76	2925 $\pm$ 58	0–1 [–88, 284] 0–2 [257, 629]**** 1–2 [121, 569]****
Length (cm)	51.4 $\pm$ .22	50.4 $\pm$ .42	49.0 $\pm$ .36	0–1 [–.07, 2.02] 0–2 [1.35, 3.44]**** 1–2 [.16, 2.67]****
Head circumference (cm)	34.7 $\pm$ .12	34.3 $\pm$ .21	33.4 $\pm$ .19	0–1 [–.12, .96] 0–2 [74, 1.81]**** 1–2 [.22, 1.50]****
Ponderal Index ( $100 \times$ birth weight/length <sup>3</sup> )	2.47 $\pm$ .02	2.56 $\pm$ .06	2.51 $\pm$ .4	0–1 [–.21, .04] 0–2 [–.17, .09] 1–2 [–.11, .20]

CI indicates confidence interval.

\* Means of each variable are compared by 1-way analysis of variance. Ninety-five percent simultaneous CIs for specified pairwise differences in group means are calculated by the Tukey method. CIs excluding 0 are flagged by \*\*\*\*.

a. The head circumference means are shown below. Use letters to indicate significant differences at the 5% level. Assume that the “Tukey method” confidence intervals are the same as ordinary confidence intervals.

	No Cocaine	Low Cocaine	High Cocaine
<i>Head Circumference (cm)</i>	34.7	34.3	33.4

b. Compare the no-cocaine group (Level 0) to the high-cocaine group (Level 2): list all the endpoints where there were significant differences. Explain how you distinguish these.

c. *True or False:* For all endpoints in the table, infants with no cocaine exposure (Level 0) were indistinguishable from those with low cocaine exposure (Level 1).

4. **Tooth Loss.** This problem is based on: RJ Hunt, CW Drake, JD Beck (1995) Eighteen-month incidence of tooth loss among older adults in North Carolina, *American Journal of Public Health*, 85:561–563. Please use the Library E-Journals website ( [tc.liblink.umn.edu/sfx\\_local/azlist/default](http://tc.liblink.umn.edu/sfx_local/azlist/default) ) to download this paper.

This paper reports a dental substudy of elderly adults who were participating in the large Piedmont Health Study of the Elderly in North Carolina. More than a thousand people volunteered to be in this dental substudy, and in the *Results* this was called the “older adult cohort.”

The plan of the study was to count how many teeth each person lost over an 18 month period. From the “older adult cohort,” the investigators excluded people who had no teeth left, and therefore could not lose any teeth. The investigators were able to enroll several hundred black and white adults over age 65 who still had at least one tooth, and results from these are reported in this paper. We’ll call the participants in Tables 1 and 2 the “study sample.” They counted how many teeth each participant had at baseline and after 18 months, to see how many had been lost during the interval.

Please read the *Abstract*, *Introduction*, *Subjects* (under *Methods*), and the first two paragraphs of *Results*.

- a. Using Table 1, describe the shapes of the histograms of numbers of teeth lost by blacks and whites. Check whether these two samples satisfy the assumptions of the two-sample *t*-test: Do the samples appear to follow the Normal curve? Do they have about the same SD? What populations were the adults selected from? How were they selected—simple random samples?

- b. The flow of participants in this study from the original samples to the study samples is obscure. According to the first sentence of the *Results*, 46% of the blacks and 29% of whites were edentulous and dropped from the “older adult cohort” to form the “study sample.” There were 335 blacks remaining in the study sample, so:

$$(1.00 - 0.46) * (\text{number of blacks in “older adult cohort”}) = 335.$$

Solve this equation to estimate the number of blacks in the older adult cohort. Use the same method to figure out the number of whites in the older adult cohort.

- c. Find the percent of edentulous blacks and whites in the older adult cohort. Use the *Z*-test to compare these proportions. Write a sentence summarizing this result.
- d. From Table 2, find the *total* number of teeth at baseline among the 335 blacks together, and among the 284 whites together. Now use your estimates of the older adult cohort and the total numbers of teeth to find the mean number of teeth at baseline among the blacks and whites for the older adult cohort.
- e. Organize your results in a table, and add the *p*-values for each column where you can:

<i>Baseline Values</i>					
<i>Group</i>	<i>Older Adult Cohort, n</i>	<i>Edentulous Adults in Cohort, n (%)</i>	<i>Mean Teeth in Cohort</i>	<i>Mean Teeth in Study Sample</i>	<i>Study Endpoint: Mean Teeth Lost</i>
<b>Blacks</b>					
<b>Whites</b>					
<i>P</i> value					

- f. The abstract summarizes the results from the study, mainly from Table 2. Do you think the abstract characterizes “tooth loss in North Carolina adults 65 and older”? Say yes or no, and explain briefly using the other results you have found.