## Supplementary Exercises 8.S.1–8.S.23

**8.S.1** A volunteer working at an animal shelter conducted a study of the effect of catnip on cats at the shelter. She recorded the number of "negative interactions" each of 15 cats made in 15-minute periods before and after being given a teaspoon of catnip. The paired measurements were collected on the same day within 30 minutes of one another; the data are given in the accompanying table.[29]

(a) Construct a 95% confidence interval for the difference in mean number of negative interactions.

(b) Construct a 95% confidence interval the wrong way, using the independent-samples method. How does this interval differ from the one obtained in part (a)?

| Cat | Before ($Y_1$) | After ($Y_2$) | Difference |
|---|---|---|---|
| Amelia | 0 | 0 | 0 |
| Bathsheba | 3 | 6 | −3 |
| Boris | 3 | 4 | −1 |
| Frank | 0 | 1 | −1 |
| Jupiter | 0 | 0 | 0 |
| Lupine | 4 | 5 | −1 |
| Madonna | 1 | 3 | −2 |
| Michelangelo | 2 | 1 | 1 |
| Oregano | 3 | 5 | −2 |
| Phantom | 5 | 7 | −2 |
| Posh | 1 | 0 | 1 |
| Sawyer | 0 | 1 | −1 |
| Scary | 3 | 5 | −2 |
| Slater | 0 | 2 | −2 |
| Tucker | 2 | 2 | 0 |
| Mean | 1.8 | 2.8 | −1 |
| SD | 1.66 | 2.37 | 1.20 |

**8.S.2** Refer to Exercise 8.S.1. Compare the before and after populations using a $t$ test at $\alpha = 0.05$. Use a nondirectional alternative.

**8.S.3** Refer to Exercise 8.S.1. Compare the before and after populations using a sign test at $\alpha = 0.05$. Use a nondirectional alternative.

**8.S.4** Refer to Exercise 8.S.1. Construct a scatterplot of the data. Does the appearance of the scatterplot indicate that the pairing was effective? Explain.

**8.S.5** As part of a study of the physiology of wheat maturation, an agronomist selected six wheat plants at random from a field plot. For each plant, she measured the moisture content in two batches of seeds: one batch from the "central" portion of the wheat head, and one batch from the "top" portion, with the results shown in the following table.[30] Construct a 90% confidence interval for the mean difference in moisture content of the two regions of the wheat head.

| | Percentage moisture | |
|---|---|---|
| Plant | Central | Top |
| 1 | 62.7 | 59.7 |
| 2 | 63.6 | 61.6 |
| 3 | 60.9 | 58.2 |
| 4 | 63.0 | 60.5 |
| 5 | 62.7 | 60.6 |
| 6 | 63.7 | 60.8 |

**8.S.6** Biologists noticed that some stream fishes are most often found in pools, which are deep, slow-moving parts of the stream, while others prefer riffles, which are shallow, fast-moving regions. To investigate whether these two habitats support equal levels of diversity (i.e., equal numbers of species), they captured fish at 15 locations along a river. At each location, they recorded the number of species captured in a riffle and the number captured in an adjacent pool. The following table contains the data.[31] Construct a 90% confidence interval for the difference in mean diversity between the types of habitats.

| Location | Pool | Riffle | Difference |
|---|---|---|---|
| 1 | 6 | 3 | 3 |
| 2 | 6 | 3 | 3 |
| 3 | 3 | 3 | 0 |
| 4 | 8 | 4 | 4 |
| 5 | 5 | 2 | 3 |
| 6 | 2 | 2 | 0 |
| 7 | 6 | 2 | 4 |
| 8 | 7 | 2 | 5 |
| 9 | 1 | 2 | −1 |
| 10 | 3 | 2 | 1 |
| 11 | 4 | 3 | 1 |
| 12 | 5 | 1 | 4 |
| 13 | 4 | 3 | 1 |
| 14 | 6 | 2 | 4 |
| 15 | 4 | 3 | 1 |
| Mean | 4.7 | 2.5 | 2.2 |
| SD | 1.91 | 0.74 | 1.86 |

**8.S.7** Refer to Exercise 8.S.6. What conditions are necessary for the confidence interval to be valid? Are those conditions satisfied? How do you know?

**8.S.8** Refer to Exercise 8.S.6. Compare the habitats using a $t$ test at $\alpha = 0.10$. Use a nondirectional alternative.

**8.S.9** Refer to Exercise 8.S.6.

(a) Compare the habitats using a sign test at $\alpha = 0.10$. Use a nondirectional alternative.

(b) Use the binomial formula to calculate the exact $P$-value for part (a).

**8.S.10** Refer to Exercise 8.S.6. Analyze these data using a Wilcoxon signed-rank test.

**8.S.11** Refer to the Wilcoxon signed-rank test from Exercise 8.S.10. On what grounds could it be argued that the $P$-value found in this test might not be accurate? This is, why might it be argued that the Wilcoxon test $P$-value is not a completely accurate measure of the strength of the evidence against $H_0$ in this case?

**8.S.12** In a study of the effect of caffeine on muscle metabolism, nine male volunteers underwent arm exercise tests on two separate occasions. On one occasion, the volunteer took a placebo capsule an hour before the test; on the other occasion he received a capsule containing pure caffeine. (The time order of the two occasions was randomly determined.) During each exercise test, the subject's respiratory exchange ratio (RER) was measured. The RER is the ratio of carbon dioxide produced to oxygen consumed and is an indicator of whether energy is being obtained from carbohydrates or from fats. The results are presented in the accompanying table.[32] Use a $t$ test to assess the effect of caffeine. Use a nondirectional alternative and let $\alpha = 0.05$.

| | RER(%) | |
|---|---|---|
| Subject | Placebo | Caffeine |
| 1 | 105 | 96 |
| 2 | 119 | 99 |
| 3 | 92 | 89 |
| 4 | 97 | 95 |
| 5 | 96 | 88 |
| 6 | 101 | 95 |
| 7 | 94 | 88 |
| 8 | 95 | 93 |
| 9 | 98 | 88 |

**8.S.13** For the data of Exercise 8.S.12, construct a display like that of Figure 8.1.1.

**8.S.14** Refer to Exercise 8.S.12. Analyze these data using a sign test.

**8.S.15** Certain types of nerve cells have the ability to regenerate a part of the cell that has been amputated. In an early study of this process, measurements were made on the nerves in the spinal cord in rhesus monkeys. Nerves emanating from the left side of the cord were cut, while nerves from the right side were kept intact. During the regeneration process, the content of creatine phosphate (CP) was measured in the left and the right portion of the spinal cord. The following table shows the data for the right (control) side ($Y_1$), and for the left (regenerating) side ($Y_2$). The units of measurement are mg CP per 100 gm tissue.[33] Consider a $t$ test to compare the two sides at $\alpha = 0.05$ using a nondirectional alternative.

| Animal | Right side (Control) | Left side (Regenerating) | Difference |
|---|---|---|---|
| 1 | 16.3 | 11.5 | 4.8 |
| 2 | 4.8 | 3.6 | 1.2 |
| 3 | 10.9 | 12.5 | −1.6 |
| 4 | 14.2 | 6.3 | 7.9 |
| 5 | 16.3 | 15.2 | 1.1 |
| 6 | 9.9 | 8.1 | 1.8 |
| 7 | 29.2 | 16.6 | 12.6 |
| 8 | 22.4 | 13.1 | 9.3 |
| Mean | 15.50 | 10.86 | 4.64 |
| SD | 7.61 | 4.49 | 4.89 |

(a) What is the value of the $t$ test statistic?

(b) In the context of this study, state the null and alternative hypotheses.

(c) The $P$-value for the $t$ test is 0.031. If $\alpha = 0.05$, what is your conclusion regarding the hypotheses in (b)?

(d) Compute a 95% confidence interval for the mean difference in creatine phosphate between the cut and control sides.

(e) A 90% confidence interval for the mean difference in creatine phosphate between the cut and control sides is (1.36, 7.91) mg/100gm. Interpret this interval in the context of the research.

**8.S.16** Aldosterone is a hormone involved in maintaining fluid balance in the body. In a veterinary study, six dogs with heart failure were treated with the drug Captopril, and plasma concentrations of aldosterone were measured before and after the treatment. The results are given in the table below.[34] Consider a sign test, using a nondirectional alternative, to investigate the claim that Captopril affects aldosterone level.

| Animal | Before | After | Difference |
|---|---|---|---|
| 1 | 749 | 374 | 375 |
| 2 | 469 | 300 | 169 |
| 3 | 343 | 146 | 197 |
| 4 | 314 | 134 | 180 |
| 5 | 286 | 69 | 217 |
| 6 | 223 | 20 | 203 |
| Mean | 397.3 | 173.8 | 223.5 |
| SD | 190.5 | 136.4 | 76.1 |

(a) What is the value of the test statistic?

(b) The $P$-value for the sign test is 0.031. If $\alpha = 0.10$, what is your conclusion?

**8.S.17** Refer to Exercise 8.S.16. Analyze these data using a Wilcoxon signed-rank test.

**8.S.18** Refer to Exercise 8.S.16. Note that the dogs in this study are not compared to a control group. How does this weaken any inference that might be made about the effectiveness of Captopril?

**8.S.19 (Computer exercise)** For an investigation of the mechanism of wound healing, a biologist chose a paired design, using the left and right hindlimbs of the salamander *Notophthalmus viridescens*. After amputating each limb, she made a small wound in the skin and then kept the limb for 4 hours in either a solution containing benzamil or a control solution. She theorized that the benzamil would impair the healing. The accompanying table shows the amount of healing, expressed as the area ($mm^2$) covered with new skin after 4 hours.[35]

(a) Assess the effect of benzamil using a $t$ test at $\alpha = 0.05$. Let the alternative hypothesis be that the researcher's expectation is correct.

(b) Proceed as in part (a) but use a sign test.

(c) Construct a 95% confidence interval for the mean effect of benzamil.

(d) Construct a scatterplot of the data. Does the appearance of the scatterplot indicate that the pairing was effective? Explain.

**8.S.20 (Computer exercise)** In a study of hypnotic suggestion, 16 male volunteers were randomly allocated to an experimental group and a control group. Each subject participated in a two-phase experimental session. In the first phase, respiration was measured while the subject was awake and at rest. (These measurements were also described in Exercises 7.5.6 and 7.10.4.) In the second phase, the subject was told to imagine that he was performing muscular work, and respiration was measured again.

For subjects in the experimental group, hypnosis was induced between the first and second phases; thus, the suggestion to imagine muscular work was "hypnotic suggestion" for experimental subjects and "waking suggestion" for control subjects. The accompanying table shows the measurements of total ventilation (liters of air per minute per square meter of body area) for all 16 subjects.[36]

(a) Use a $t$ test to compare the mean resting values in the two groups. Use a nondirectional alternative and let $\alpha = 0.05$. This is the same as Exercise 7.5.6(a).

(b) Use suitable paired and unpaired $t$ tests to investigate (i) the response of the experimental group to suggestion; (ii) the response of the control group to

| Animal | Control limb | Benzamil limb | Animal | Control limb | Benzamil limb |
|---|---|---|---|---|---|
| 1 | 0.55 | 0.14 | 10 | 0.42 | 0.21 |
| 2 | 0.15 | 0.08 | 11 | 0.49 | 0.11 |
| 3 | 0.00 | 0.00 | 12 | 0.08 | 0.03 |
| 4 | 0.13 | 0.13 | 13 | 0.32 | 0.14 |
| 5 | 0.26 | 0.10 | 14 | 0.18 | 0.37 |
| 6 | 0.07 | 0.08 | 15 | 0.35 | 0.25 |
| 7 | 0.20 | 0.11 | 16 | 0.03 | 0.05 |
| 8 | 0.16 | 0.00 | 17 | 0.24 | 0.16 |
| 9 | 0.03 | 0.05 | | | |

suggestion; (iii) the difference between the responses of the experimental and control groups. Use directional alternatives (suggestion increases ventilation, and hypnotic suggestion increases it more than waking suggestion) and let $\alpha = 0.05$ for each test.

| Experimental group | | | Control group | | |
|---|---|---|---|---|---|
| Subject | Rest | Work | Subject | Rest | Work |
| 1 | 5.74 | 6.24 | 9 | 6.21 | 5.50 |
| 2 | 6.79 | 9.07 | 10 | 4.50 | 4.64 |
| 3 | 5.32 | 7.77 | 11 | 4.86 | 4.61 |
| 4 | 7.18 | 16.46 | 12 | 4.78 | 3.78 |
| 5 | 5.60 | 6.95 | 13 | 4.79 | 5.41 |
| 6 | 6.06 | 8.14 | 14 | 5.70 | 5.32 |
| 7 | 6.32 | 11.72 | 15 | 5.41 | 4.54 |
| 8 | 6.34 | 8.06 | 16 | 6.08 | 5.98 |

(c) Repeat the investigations of part (b) using suitable nonparametric tests (sign and Wilcoxon-Mann-Whitney tests).

(d) Use suitable graphs to investigate the reasonableness of the normality condition underlying the $t$ tests of part (b). How does this investigation shed light on the discrepancies between the results of parts (b) and (c)?

**8.S.21** Suppose we want to test whether an experimental drug reduces blood pressure more than does a placebo. We are planning to administer the drug or the placebo to some subjects and record how much their blood pressures are reduced. We have 20 subjects available.

(a) We could form 10 matched pairs, where we form a pair by matching subjects, as best we can, on the basis of age and sex, and then randomly assign one subject in each pair to the drug and the other subject in the pair to the placebo. Explain why using a matched pairs design might be a good idea.

(b) Briefly explain why a matched pairs design might *not* be a good idea. That is, how might such a design be inferior to a completely randomized design?

**8.S.22** A group of 20 postmenopausal women were given transdermal estradiol for one month. Plasma levels of plasminogen-activator inhibitor type 1 (PAI-1) went down for 10 of the women and went up for the other 10 women.[37] Use a sign test to test the null hypothesis that transdermal estradiol has no effect on PAI-1 level. Use $\alpha = 0.05$ and use a nondirectional alternative.

**8.S.23** Six patients with renal disease underwent plasmapheresis. Urinary protein excretion (grams of protein per gram of creatinine) was measured for each patient before and after plasmapheresis. The data are given in the following table.[38] Use these data to investigate whether or not plasmapheresis affects urinary protein excretion in patients with renal disease. (*Hint*: Graph the data and consider whether a $t$ test is appropriate in the original scale.)

| Patient | Before | After | Difference |
|---|---|---|---|
| 1 | 20.3 | 0.8 | 19.5 |
| 2 | 9.3 | 0.1 | 9.2 |
| 3 | 7.6 | 3.0 | 4.6 |
| 4 | 6.1 | 0.6 | 5.5 |
| 5 | 5.8 | 0.9 | 4.9 |
| 6 | 4.0 | 0.2 | 3.8 |
| Mean | 8.9 | 0.9 | 7.9 |
| SD | 5.9 | 1.1 | 6.0 |