

Ma 322: Biostatistics

Homework Assignment 8

Prof. Wickerhauser

Due Friday, March 30th, 2012

Read Chapter 14, “Hypothesis Testing,” pages 240–262 of our text.

- Following are 14 samples from a normal population with mean μ and unknown standard deviation σ variance:

2.59 2.67 2.16 1.95 2.61 1.11 2.62 2.06 2.06 1.66 2.16 3.35 2.46 2.55

- Test the hypothesis $H_0 : \mu = 2.0$, using the significance level $\alpha = 0.05$.
 - Test the hypothesis $H_0 : \mu \leq 2.0$, using the significance level $\alpha = 0.05$.
- Using the sample standard deviation from Exercise 1 and a significance level of $\alpha = 0.05$, determine:
 - The power $1 - \beta$ of the t -test to reject the two-sided null hypothesis on the mean in Exercise 1a when there is a true difference $\delta = 0.4$.
 - The power $1 - \beta$ of the t -test to reject the one-sided null hypothesis on the mean in Exercise 1b when there is a true difference $\delta = 0.4$.
 - The number of samples needed to get a power $1 - \beta = 99\%$ in the t -test of the two-sided null hypothesis on the mean in Exercise 1a when there is a true difference $\delta = 0.4$.
 - The number of samples needed to get a power $1 - \beta = 99\%$ in the t -test of the one-sided null hypothesis on the mean in Exercise 1b when there is a true difference $\delta = 0.4$.
 - (a) Using the following data, and assuming that both populations are normal with equal variance, test the null hypothesis that male and female turtles have the same mean serum cholesterol concentrations.

Serum cholesterol (mg/100 ml) of turtles.

Male	248,329,223,313,271,324,255,255,423,332,311,264
Female	341,311,362,371,419,366,246,273,312,331

- The following data were found in Table 1 of C. M. Holcomb, C. G. Jackson, Jr., and M. M. Jackson, “Serum Cholesterol Values in Three Species of Turtles,” *J. Wildlife Diseases* 8(1972), pp.181–182. < www.jwildlifedis.org/cgi/reprint/8/2/181.pdf>

Serum cholesterol (mg/100 ml) in turtles.

Species	n	Mean	S.E.	Range	Coef. of Var.
<i>C. scripta</i>	8	290.0	± 42.3	174–512	41.2%
<i>T. carolina</i>	31	339.7	± 15.6	178–511	25.6%

Assuming that both populations are normal with equal variance, test the alternative hypothesis that *T. carolina* has higher mean serum cholesterol concentrations than *C. scripta*.

4. State the null and alternative hypotheses, and use a paired-sample t -test at the 5% significance level to compare the marksmanship of fathers and their daughters, using the data in the following table:

Scores after four 5-shot targets (out of 200).

Team Number	Father	Daughter
1	120	160
2	155	160
3	184	158
4	190	172
5	125	185
6	88	134
7	155	155

5. For a fair coin, expect a binomial distribution with “heads” probability $p = 1/2$. A certain guilder coin is tossed 1000 times and comes up heads just 461 times.
- (a) Rosencrantz does not believe that this guilder is a fair coin. Use the experimental data and a significance threshold of $\alpha = 0.01$ to test Rosencrantz’s hypothesis that heads are less likely than tails in a toss of that coin.
- (b) Guildenstern does not share Rosencrantz’s suspicions about the coin. Use the experimental data and a significance threshold of $\alpha = 0.01$ to test Guildenstern’s hypothesis that heads and tails are equally likely in a toss of that coin.
6. (a) Using the data for Problem 3, part a, test the null hypothesis that male and female turtles have the same serum cholesterol variance.
- (b) Using the data for Problem 3, part b, test the alternative hypothesis that *C. scripta* has a higher serum cholesterol variance than *T. carolina*.
7. (a) Use the Mann-Whitney (or Wilcoxon) rank test to compare male and female turtle cholesterol, using the data for Problem 3, part a, without the normality and equivariance assumptions.
- (b) Use the Mann-Whitney (or Wilcoxon) rank test to compare the marksmanship of fathers and their daughters, using the data for Problem 4, without the normality and equivariance assumptions.
8. (a) Test the hypothesis that nucleotides a,c,g,t are equally likely in the GenBank sequence NM_005369, using the χ^2 goodness-of-fit method. Use significance level $\alpha = 0.01$.
- (b) Test the hypothesis that nucleotides a,c,g,t are equally likely in the GenBank sequence NM_005367, using the χ^2 goodness-of-fit method. Use significance level $\alpha = 0.01$.
9. (a) How many 2×2 contingency tables are there with row sums (2, 5) and column sums (3, 4)? (Hint: Write down all the solutions.)
- (b) Assuming that the rows and columns are independent, compute the exact hypergeometric probability of each 2×2 contingency table in part a.