EXERCISES ON HYPOTHESIS TESTING

The excercises below are taken from:

Kottegoda, N.T. and Rosso, R., 1998, *Statistics, Probability, and Reliability for Civil and Environmental Engineers*, Mc.Graw-Hill, International Edition, New York.

SAMPLING DISTRIBUTION OF DIFFERENCES AND SUMS OF STATISTICS

Measurements of rainfall have been taken at a particular location, say, Station #1, over a period of $n_1 = 50$ years, and the annual mean and standard deviation are estimated as $\bar{x}_1 = 900$ mm and $s_1 = 80$ mm, respectively. At another location, say, Station #2, measurements cover a period of $n_2 = 40$ years, from which estimates of the annual mean and standard deviation are $\bar{x}_2 = 825$ mm and $s_2 = 90$ mm, respectively. Because annual rainfalls are the results of a large number of small causes, such an additive effect makes it plausible to assume that the distribution is normal (see Central Limit Theorem). There is sufficient empirical evidence to support this claim; and although in some cases non-normal distributions have been found, the approximation is a reasonable one. Determine the sampling distribution limits for the difference between the two means.

TESTS FOR PROPORTIONS USING THE BINOMIAL APPROXIMATION TO THE NORMAL DISTRIBUTION

Two types of plant are used to treat the sewage effluent from two similar areas of a city. Of 90 tests made on the output from plant X, 33 tests show that the pollution has been reduced significantly, whereas 44 tests out of 100 on the output from plant Y show that the pollution has been reduced to the same or lower levels. Are the effects of the plants in reducing pollution different?

CHANGE IN THE MEAN ANNUAL MAXIMUM FLOW WITH KNOWN STANDARD DEVIATION AND PRIOR MEAN

Annual maximum flows in the Pond Creek catchment area in the eastern United States are listed below in cubic meters per second for two periods of 12 years from 1945 to 1968. It is thought that improvements in the flow regime during the middle of this period have led to higher peak flows. This is equivalent to stating that the mean flow has increased from the first period to the second period. What would you say on that claim?

1st period	2000	1740	1460	2060	1530	1590	1690	1420	1330	607	1380	1660
2nd period	2290	2590	3260	2490	3080	2520	3360	8020	4310	4380	3220	4320

CHANGE IN THE MEAN ANNUAL MAXIMUM FLOW WITH KNOWN PRIOR MEAN AND UNKNOWN BUT CONSTANT STANDARD DEVIATION

In practice, we do not know the standard deviation in the above exercise. Construct your new hypothesis under this condition.

CHANGE IN THE MEAN ANNUAL MAXIMUM FLOW WITH UNKNOWN AND UNEQUAL VARIANCES

Construct your hypothesis considering that we do not know the variances but we know that the two variances are unequal.

A SIGNIFICANCE TEST ON THE CHANGE IN VARIANCE

From a long series of annual river flows, the variance is found to be 49 units. This can be treated as the population variance. However, new samples of 25 years give a value $\hat{s}^2 = 81$ units. Does it show a change in variance?

COMPARING VARIANCES OF TEST RESULTS ON AN EFFLUENT

A constituent in an effluent is analyzed seven and nine times through procedures *X* and *Y*, respectively. Test results have standard deviations of 1.9 and 0.8 mg/ℓ, respectively, by the two procedures. It is important to know whether the second method is more precise than the first (with less variance in the outcome).

MINIMUM SAMPLE SIZE FOR ESTIMATING MEAN DISSOLVED OXYGEN (DO) CONCENTRATION

Monitoring of pollution levels of similar streams in a region indicates that the standard deviation of DO is 1.95 mg/ℓ over a long period of time. (a) What is the minimum number of observations required to estimate the mean DO within ±0.5 mg/ℓ with 95 percent confidence? (b) If only 30 observations are taken, what should be the percentage level in the confidence limits for the same difference in mean?

CONFIDENCE LIMITS ON PROPORTIONS OF WET DAYS

A building contractor who works in a relatively dry area is planning to acquire additional work in a newly developing area but is somewhat doubtful of progress because of the adverse effects of rainfall in many months of the year. However, the contractor knows that March is a month of low rainfall with independently distributed daily rainfalls and no apparent relationship between the weather on successive days. Therefore, the thought is that this may be a suitable month to work on the foundations. The proportion of wet days in March is 0.10 from data of the past three years. Suppose it is possible to put off the decision for some time in order to make further observations of daily rainfalls in March. Determine the total number of years of data necessary before one can be 95 percent confident of estimating the true proportion of wet days to within 0.05.

COMPARING OUTPUTS OF WASTEWATER PLANTS

Two treatment plants are built in an area to treat waste-water from a city. Their relative performances are compared from the results of BOD tests made on the outputs. Eight preliminary results are listed here as differences in BOD between plant 1 and 2. Test the difference in the outputs at the 5 percent level of significance.

Test	1	2	3	4	5	6	7	8
Δ_{BOD}, mg/ℓ	+1.2	+0.2	–1.6	+0.7	+1.3	–0.9	–0.1	–1.9

---ist---