MAST90105 Methods of Mathematical Statistics

Assignment 4, Semester 1 2019

Due: Sunday 9 June, end of day. Please submit a scanned or other electronic copy of your work via the Learning Management System - see this link for instructions

Note: You may use R and/or Mathematica for any questions but must include your commands and reasoning.

Problems:

1. Carbon monoxide concentrations (in $\mu g/m^3$) are measured on 6 different days in city A and on 4 different days in city B. The following measurements have been obtained in cities A and B, respectively: 3.9, 4,7, 7.1, 6.9, 4.3, 6.3 and 7.3, 6.9, 7.6, 9.1. We assume that these two samples are independent.

   (a) Assume that both samples are from normally distributed populations. We also assume that the same type of measurement device was used in both cities so that the measurement error (variance) is the same for both samples. Construct the 95% confidence interval for $\delta = \mu_A - \mu_B$ where μ_A and μ_B are mean concentrations of carbon monoxide in cities A and B, respectively.

   (b) Do we reject $H_0 : \delta = 0$ in favor of $H_1 : \delta \neq 0$ at the 5% significance level? Why or why not? Find the p-value of this test.

   (c) Now assume that both samples are from normally distributed populations, as before, but two different measurement devices were used and their measurements errors are known. We assume that the variances are 2 and 1, respectively, for the two samples. Construct the 95% confidence interval for δ.

   (d) Compare confidence intervals you obtained in (a) and (c). Which one is narrower? Briefly explain why.

   (e) Show that $(3.9, 7.1)$ is an approximate 97% confidence interval for the median concentration of carbon monoxide in city A.

2. Let $X_1, X_2, \ldots, X_{10}$ be a sample of size 10 from an exponential distribution with the density function

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

   We reject $H_0 : \lambda = 1$ in favor of $H_1 : \lambda = 2$ if the observed value of $Y = \sum_{i=1}^{10} X_i$ is smaller than 6.

   (a) Find the probability of type 1 error for this test.

   (b) Find the probability of type 2 error for this test.

(c) Let $y = 5$ be the observed value of Y. Find the p-value for this test.

(d) Let $y = 5$ be the observed value of Y. Construct the exact 95% confidence interval for λ. *Hint:* $\lambda Y \sim Gamma(10, 1)$.

3. The number of faults over a period of time was collected for a sample of 100 data-transmission lines. We want to test if the data come from a Poisson distribution.

number of faults	0	1	2	3	4	5	> 5
number of lines	38	30	16	9	5	2	0

(a) Assuming the number of faults for a data-transmission line, Y_i, $i = 1, \ldots, 100$, follows a Poisson distribution with parameter λ, find the maximum likelihood estimate of λ for these data.

(b) Test the hypothesis that the number of faults for a data-transmission line follows a Poisson distribution using the chi-squared test. Do we reject the null hypothesis at the 5% significance level? *Hint:* You should combine the observed data into several groups such that expected frequencies are greater or equal to 5 for each group.