Midterm

2021. 4. 19

1. (Computing, 120%)

(a) (25%) Let

$$G = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

Please write a program to compute:

(i) The eigenvalues and eigenvectors of G .

(ii) The determinant of G .

(iii) The solution of $Gx = b$, where $\mathrm{b} = (2 \quad 4 \quad 6)^t$.

(iv) $3G^6 - 2G^{-4} + 5G^t$

(v) The row variances and column means of G .

(b) (25%) Let

$$f(x) = 2x^2 + 4x^{-3} - 6x^{1/5}$$

and

$$g(x) = 20\pi e^{-x^2} - 5cos(x)log(x)$$
$$\mathrm{x} = 0.5, 1, \cdots, 9.5, 10.$$

Please write a program to plot the two functions with the following requirements:

- **X-label is "x" while Y-label is "Mathematical Functions".**
- **The title of this plot is "The Plot of Different Functions".**
- **Use two different kind of point types for the two functions.**
- **Use two different kind of lines types for the two functions.**
- **The legend associated with the two functions.**

(c) (25%) Let $X_1, X_2, \cdots, X_{50} \overset{i.i.d}{\longrightarrow} N(\mu, \sigma^2)$. Then, two estimators of ${\mu^2}/{\sigma^2}$ are

$$T_1 = \frac{50\bar{X}^2}{49S^2}$$

and

$$T_2 = \frac{\bar{X}^2 - {1}/{50}}{S^2},$$

where $\bar{X}$ and S^2 are sample mean and sample variance, respectively. Please sample 50 data from $N(3, 9)$. The sampling process is repeated 2000 times. Please write a program to do the following:

(i) Find the averages of the above two estimates.

(ii) Find the averages of the absolute differences between the above two estimates and the true value of the parameter.

(d) (25%) Please use Newton's method to find the at least *two roots* of

$$x^2 + y^2 - 2 = 0$$
$$x^2 - y = 0$$

with the stop criterion (error) equal to 0.00001.

(e) (20%) Please generate 100000 data from the standard normal random variable. Then, please write a program to find:

(i) the numbers of observations in the intervals,

$$(-\infty, 3), [-3, -2.99), \cdots, [-0.01, 0), [0, 0.01), \cdots, [2.99, 3), [3, \infty).$$

(ii) which intervals with the minimum and maximum number of observations.

2. (Statistics, 120%)

(a) (30%) Given that $X \sim N(1, 9)$, $Y_1 \sim Poisson(3)$ is a Poisson random variable with mean 3, $Y_2 \sim Binomial(5, 0.2)$ is a binomial random variable corresponding to 5 trials with the probability of success equal to 0.2. Please compute

(i) $P(-5 \le X \le 8.5)$.

(ii) $P(Y_1 < 7)$.

(iii) $P(Y_2 = 0 \ or \ Y_2 = 2 \ or \ Y_2 = 10)$

(iv) $t_{7,0.01}$.

(v) Generate a sample of 1000 data from the uniform random variable taking values on $[-5, 9]$.

(b) (20%) The data were selected from each of three normal populations with equal variances. The data obtained follow.

Observation	Sample 1	Sample 2	Sample 3
1	37	39	33
2	35	38	36
3	35	39	35
4	31	41	36
5	37	43	40

At the $\alpha = 0.05$ level of significance, find the p-value and test the null hypothesis that the three population means are equal?

(c) (30%)

(i) The following data have been collected for a sample from a normal population

5	9	6	4	6	8	9	11

find 90% confidence interval for population mean μ and the p-value for testing

$$H_0: \mu \geq 9 \ \ vs. \ \ H_1: \mu < 9.$$

(ii) For the following data taken from two normal populations with equal variances.

Sample 1	25	26	47	46	45	21	33
Sample 2	43	59	60	74	66		

Find the t-statistic and the p-value for

$$H_0: \mu_1 - \mu_2 \leq -15 \ \ vs. \ \ H_1: \mu_1 - \mu_2 > -15,$$

where μ_1 and μ_2 are the means of population 1 and population 2, respectively.

(iii) Consider the following data for two random samples taken from two normal populations with equal variances.

Sample 1	11	12	8	7	7	9
Sample 2	5	8	6	7	4	8

Consider the above data as the matched (paired) samples. Find the 85% confidence interval for $\mu_1 - \mu_2$ and the t-statistic for

$$H_0: \mu_1 - \mu_2 = 2 \ \ vs. \ \ H_1: \mu_1 - \mu_2 \neq 2,$$

where μ_1 and μ_2 are the means of population 1 and population 2, respectively.

(d) (20%)

(i) Please generate 500 data from a t distribution with the degree of freedom equal to 2. Please write a program to do the following:

- qq t plot for the generated data with 2 degrees of freedom.
- qq normal plot for the generated data with $: \mu = 2, \sigma^2 = 9.$

Put the above 2 plots in the same Figure.

(ii) Suppose we have the following data:

3	2	2	1	0
0	1	0	1	0
4	1	0	1	0
2	2	1	0	0

Test if the data is distributed as the Poisson distribution with mean equal to 0.7 ($\mathbf{Poisson(0.7)}$)) at $\alpha = 0.05$.
(Hint: For $X \sim Poisson(0.7)$, $P(X \geq 4) \approx 0$)

(e) (20%) The results of a recent poll on the preference of voters regarding two candidates are shown below:

Candidate	Voters Surveyed	Voters Favoring This Candidate
A	*400*	*192*
B	*450*	*225*

Please construct a 95% confidence interval for the difference between the preference for the two candidates $\mathbf{p_1 - p_2}$ and the p-value for

$$H_0: p_1 = p_2 \ \ vs. \ \ H_1: p_1 \neq p_2,$$

where p_1 and p_2 are the proportions of favoring candidate A and candidate B, respectively.