Applied Geostatistics
(응용지구통계학) (G17604)

- 2020 Midterm Examination -

Student ID: **Name:**

Notice

- Fill your name below:
  "I, ________________, swear I solve all problems by myself in this midterm examination. I will take any disadvantages if any dishonesty such as cheating is acted on my solution."
  5 points will be deducted from your total score if you do not fill in your name above.

- Submit your solution as *.pdf or *.word file on the cyber campus by May 9, 18:30.
- Please follow the format that gives a name to your report or report file:
  (Mid)-(Student ID)-(Last name)-(First name)
  For example, the file name must be Mid-XXXXXXX-Min-Baehyun.
- Due date: May 9, 2020, 18:30:00 PM GMT+9.
- No late submission is accepted.

Problem 1. [10 pts.]

1-1. Let N be a positive integer and let f be the function defined by:

$$f(x) = \begin{cases} \dfrac{2x}{N(N+1)}, & x = 1,2,\dots,N \\ 0, & \text{elsewhere} \end{cases}$$

Show that $f(x)$ is a discrete density and find its mean.

Hint:

$$\sum_{x=1}^{N} x = \frac{N(N+1)}{2} \text{ and } \sum_{x=1}^{N} x^2 = \frac{N(N+1)(2N+1)}{6}$$

1-2. The Cauchy cumulative distribution function $F(x)$ is

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x), \qquad -\infty < x < \infty.$$

i. Show that $F(x)$ is a cdf.
ii. Find the density function $f(x)$.
iii. Find x such that $P(X > x) = 0.1$.

Problem 2. [15 pts.]

2-1. Let

$$f(x,y) = xe^{-x(y+1)}, 0 \le x < \infty, 0 \le y < \infty.$$

a. Find the marginal densities of X and Y. Are X and Y independent?
b. Find the conditional densities of X and Y.

2-2. If X_1 is uniform on [0, 1] and conditional on X_1, X_2 is uniform on $[0, X_1]$, find the joint and marginal distributions of X_1 and X_2.

2-3. Let

$$f(x,y) = e^{-y}, 0 \le x \le y.$$

a. Find $\text{cov}(X, Y)$ and the correlation between X and Y.
b. Find the marginal densities
c. Find the conditional densities.
d. Find $E(X \mid Y = y)$ and $E(Y \mid X = x)$.
e. Find the density functions of $E(X \mid Y)$ and $E(Y \mid X)$.

Problem 3. [10 pts.]

Consider the following data pertaining to temperature in a given domain.

Observation number	Value
1	12.0
2	13.5
3	13.0
4	15.3
5	14.1
6	9.1
7	7.6
8	17.2
9	16.1
10	11.0
11	13.5
12	16.0
13	15.2
14	8.6
15	9.1
16	11.5
17	14.2
18	14.9
19	16.3
20	8.4
21	17.8
22	16.5
23	16.9
24	12.5
25	13.4
26	9.8
27	8.9
28	10.3
29	10.8
30	12.4

Discuss the graphical procedure to transform the data to Gaussian distributed values with unit mean and variance. Show your all procedures.

Problem 4. [15 pts.]

4-1. Prove the following:

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}, a > 0 .$$

4-2. Show that its cumulative distribution function value $F(\infty) = 1$.

Hint:

$$\left(\int_{-\infty}^{\infty} e^{-ax^2} dx\right)^2 = \int_{-\infty}^{\infty} e^{-ax^2} dx \int_{-\infty}^{\infty} e^{-ay^2} dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-a(x^2+y^2)} dxdy , a > 0$$

4-3. Show that the cumulative distribution function $F(z) = \int_{-\infty}^{z} f(z)dz$ can be expressed using the error function $erf(z)$ as follows:

$$F(z) = \begin{cases} \frac{1}{2}\left[1 + erf\left(\frac{z}{\sqrt{2}}\right)\right], if\ z \geq 0 \\ \frac{1}{2}\left[1 + erfc\left(-\frac{z}{\sqrt{2}}\right)\right], if\ z < 0 \end{cases} ,$$

where

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_{0}^{z} e^{-x^2} dx,$$

$$erfc(z) = 1 - erf(z).$$

Problem 5. [10 pts.]

Let the non-negative random variable w follow the log-normal distribution.

Using the log-normal probability density function $f(w)$, prove the following:

$$\mu_w = \exp(\mu_z + 0.5\sigma_z^2),$$
$$\sigma_w^2 = \mu_w^2[\exp(\sigma_z^2) - 1].$$

Note that

$$f(w) = g(z)\left|\frac{dz}{dw}\right|,$$

$$f(w) = \frac{1}{\sqrt{2\pi}\sigma_z w}\exp\left[-\frac{1}{2}\left(\frac{\ln(w) - \mu_z}{\sigma_z}\right)^2\right], w > 0\,,$$

$$z = \ln(w)\,, z \sim N(\mu_z, \sigma_z^2).$$

Problem 6. [10 pts.]

There is an infinite population composed of nonnegative values of which the mean and variance are μ and σ^2, respectively. Show that z^* tends to follow a log-normal distribution when z^* is the product of n data points sampled from the population.

Problem 7. [10 pts.]

Derive the relationship among variogram, variance, and covariance based on the 2^{nd}-order weak stationarity.

$$\gamma(h) = \sigma^2 - R(h),$$

where $\gamma(h)$ is the semi-variogram, σ^2 is the variance, and $R(h)$ is the covariance, and h is the separation distance (i.e., lag distance).

Problem 8. [10 pts.]

Explain nugget effects in variogram modeling.

Problem 9. [10 pts.]

Calculate experimental variogram values and experimental covariance values with the increment of separation distance of 2 ft (e.g., h = 0 ft, 2 ft, 4 ft, 6 ft, …). Draw a graph of experimental variogram values and that of experimental covariance values together in a plot.

cf.) Permeability in fluid mechanics and the earth sciences is a measure of the ability of a porous material to allow fluids to pass through it. [Wikipedia]

Depth (ft)	Permeability (md)	Depth (ft)	Permeability (md)
0.5	101.1	25.5	156.6
1.5	116.5	26.5	186.7
2.5	132.4	27.5	122.7
3.5	108.1	28.5	80.3
4.5	110.3	29.5	113.9
5.5	101.3	30.5	124.4
6.5	100.0	31.5	127.5
7.5	87.8	32.5	85.2
8.5	118.5	33.5	49.9
9.5	99.9	34.5	22.4
10.5	104.7	35.5	88.4
11.5	113.2	36.5	96.4
12.5	131.9	37.5	76.1
13.5	55.1	38.5	63.2
14.5	78.6	39.5	90.6
15.5	44.7	40.5	10.8
16.5	79.7	41.5	0.0
17.5	92.5	42.5	41.8
18.5	110.3	43.5	69.1
19.5	35.0	44.5	78.3
20.5	59.8	45.5	61.9
21.5	100.2	46.5	53.6
22.5	115.1	47.5	51.0
23.5	108.3	48.5	91.5
24.5	135.6	49.5	103.7

------------------- This is the End of the Midterm Examination --------------------