The University of British Columbia – Okanagan

DATA315

Assignment 1

Due on Feb 2 by uploading the file to Canvas that contains your assignment.

Question 1

Suppose $Y = 3X + Z$ where Z and X are independent random variables. Suppose X has mean 5 and variance 1 and Z has mean 0 and variance 9.

(a) (1 point) Find $E[Y]$.

(b) (1 point) Find $V(2X)$.

(c) (1 point) Find $E[X^2]$.

(d) (1 point) Find $E[XY]$.

(e) (1 point) Find $Cov(X, Y)$. Are X and Y positively linearly related?

(f) (1 point) Find $E[Y|Z = 2]$.

Question 2

(2 points) Suppose X is a normal random variable with mean 2.0 and standard deviation 0.8, and suppose $Y = 5X + 7$. Find the mean and standard deviation of Y.

Question 3

(3 points) Calculate the sample average and sample standard deviation of the level observations in `LakeHuron`. Use R for this question and submit the code together with your answers.

Question 4

(3 points) Suppose Z_1, Z_2 and Z_3 are independent standard normal random variables. Write down the joint pdf for Z_1, Z_2 and Z_3.

Question 5

(4 points) Suppose $Z_1, Z_2, \ldots, Z_n$ are time series with white noise model

$$z_t = \varepsilon_t$$

where ε_t is independent random variable with exponiential distribution $exp(\lambda)$.

Find the maximum likelihood estimator for λ.

Question 6

Suppose $Z_1, Z_2, \cdots, Z_n$ is a time series. Let

$$L(\rho) = \sum_{t=2}^{n} (Z_t - \rho Z_{t-1})^2$$

(a) (4 points) Using calculus, find a formula for $\hat{\rho}$, the value of ρ which minimizes $L(\rho)$. (This is the least -squares estimator for ρ, the so-called lag 1 autocorrelation.)

(b) (4 points) Suppose, in addition, that $Z_1, Z_2, \cdots, Z_n$ are independent normally distributed random variables with mean 0 and standard deviation is σ. Find $E[\sum_{t=2}^{n} Z_t Z_{t-1}]$ and $E[\sum_{t=1}^{n-1} Z_t^2]$.

(c) (3 points) Let

$$X_t = Z_{t-1}, \quad t = 2, \ldots, n$$

and

$$Y_t = Z_t, \quad t = 1, \ldots, n-1.$$

Write down the formula for the sample correlation between X and Y, expressed in terms of the Z's, and compare the formula you obtained with $\hat{\rho}$ obtained in part (a).

Question 7

(21 points) Z_t is the timeseries where t take $1, 2, \cdots, n$. Consider a time series where $n = 4$

- Find out the distribution of the statistic τ for Mann-Kendall test
- Verify in these cases that the variance of τ is $\frac{2(2n+5)/9}{n(n-1)}$.
- Suppose z_1, z_2, z_3, z_4 are 0.2, 0.25, 0,22, 0.3 . Do Mann-Kendall test using the exact distribution you have shown.

Question 8

(3 points) The data in `DAAG::vlt$prize` consist of a time series of observations of prize winnings (in dollars) taken on a video lottery terminal in a sequence of 345 plays. Obtain a trace plot of the series and conduct a trend test. Is there strong evidence of a decrease or increase in the prize winnings over time?

Question 9

Consider the time series: 13, 11, 14, 17, 16, 15. Conduct the Mann-Kendall trend test by hand, using the following steps.

(a) (3 points) Calculate the value of τ showing the steps that you are using.

(b) (1 point) Calculate the variance of τ under the null hypothesis .

(c) (3 points) Under the approximate normal assumption, calculate the test statistic using τ and the standard error. What would you conclude from this test?