

# Stats 531 Homework 1

Yirui Gao

Jan 17

## Contents

### Question 1.1.

First, we should start by expanding the variance equation for the estimator of random variable  $Y_{1:N}$ . We can use the sample mean provided:

$$\begin{aligned}\text{Var}(\hat{\mu}(Y_{1:N})) &= \text{Var}\left(\frac{1}{N} \sum_{n=1}^N Y_n\right) \\ &= \text{Cov}\left(\frac{1}{N} \sum_{m=1}^N Y_m, \frac{1}{N} \sum_{n=1}^N Y_n\right) \text{ (P1.)} \\ &= \frac{1}{N^2} \text{Cov}\left(\sum_{m=1}^N Y_m, \sum_{n=1}^N Y_n\right) \text{ (P3.)} \\ &= \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N \text{Cov}(Y_m, Y_n) \text{ (P4.)} \\ &= \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N E[(Y_m - \mathbb{E}[Y_m])(Y_n - \mathbb{E}[Y_n])].\end{aligned}$$

The mean stationary property holds since the mean function is constant. And according to the definition of autocovariance, we can see

$$E[(Y_m - \mathbb{E}[Y_m])(Y_n - \mathbb{E}[Y_n])] = E[(Y_m - \mu_m)(Y_n - \mu_n)] = \gamma_{m,n}.$$

The covariance stationarity moves the derivations forward. Moreover, we can see

$$\gamma_0 = \gamma_{0,0} = \frac{1}{N} \sum_{n=1}^N (Y_n - \mu_n)^2.$$

Therefore, we can further derive:

$$\begin{aligned}\text{Var}(\hat{\mu}(Y_{1:N})) &= \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N \gamma_{m,n} \\ &= \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N (\gamma_0 + \gamma_{m,n,m \neq n}) \\ &= \frac{1}{N^2} \sum_{m,n=1}^N \gamma_0 + \frac{1}{N^2} \sum_{m=1}^N \sum_{n \neq m}^N \gamma_{m,n} \\ &= \frac{1}{N} \gamma_0 + \frac{1}{N^2} \sum_{m=1}^N \sum_{n \neq m}^N \gamma_{m,n}.\end{aligned}$$

It's not hard to get that inside  $\sum_{m=1}^N \sum_{n \neq m}^N \gamma_{m,n}$ , if we introduce lag  $h$  as the difference between  $m$  and  $n$  (assume that  $m$  is larger than  $n$ ), then when  $h = 1$ , there are  $N - 1$  pairs of  $(m, n)$ . When  $h = 2$ , there are

$N - 1$  pairs of  $(m, n) \dots$  When  $h = N - 1$ , there are only 1 pairs of  $(m, n)$ . Considering that  $n$  can also be larger than  $m$ , we need to double the number of pairs. Thus,

$$\sum_{m=1}^N \sum_{n \neq m} \gamma_{m,n} = 2 \sum_{h=1}^{N-1} (N-h) \gamma_h.$$

So in the end, the equation holds:

$$\begin{aligned} \text{Var}(\hat{\mu}(Y_{1:N})) &= \frac{1}{N} \gamma_0 + \frac{1}{N^2} \sum_{m=1}^N \sum_{n \neq m} \gamma_{m,n} \\ &= \frac{1}{N} \gamma_0 + \frac{2}{N^2} \sum_{h=1}^{N-1} (N-h) \gamma_h. \end{aligned}$$

### Question 1.2.

A.

In this case, in the Taylor series approximation scenario, we replace  $Y$  with our targeted random variable, the sample autocorrelation function,  $\hat{\rho}_h(Y_{1:N})$ . Then the function  $g()$  becomes the definition of the sample ACF.

According to the expression of Taylor series approximation and consider the formula  $g(\mu_X)$ , here  $X$  should be replaced by the i.i.d. random variables  $Y_1, \dots, Y_N$  with 0 mean and finite variance  $\sigma^2$ .

Then, since we cannot directly derive the derivative for the expression for  $Y$ , we can use the knowledge of linear approximation in multivariable Taylor series approximation, which states that "take the constant and linear terms from the Taylor series. In a neighborhood of  $(x, y) = (a, b)$ ,

$$f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)."$$

with two random variables  $x, y$  approximately centered at  $a, b$ . In our case here, we can consider here random variable  $A = \frac{1}{N} \sum_{n=1}^{N-h} Y_n Y_{n+h}$  as the numerator for sample ACF, and random variable  $B = \frac{1}{N} \sum_{n=1}^N Y_n^2$  as the denominator. For the two centered point, we can pick out the mean function for these random variables,  $a = \mu_A$  and  $b = \mu_B$ . We can calculate the mean functions based on the truth of i.i.d.  $Y$ s:

$$a = \mu_A = \mathbb{E}[A] = \frac{1}{N} \sum_{n=1}^{N-h} \mathbb{E}[Y_n Y_{n+h}] = 0,$$

$$\begin{aligned} b = \mu_B &= \mathbb{E}[B] = \frac{1}{N} \sum_{n=1}^N \mathbb{E}[Y_n^2] \\ &= \frac{1}{N} \sum_{n=1}^N \mathbb{E}[Y_n^2] \\ &= \frac{1}{N} \sum_{n=1}^N \text{Var}[Y_n] \\ &= \sigma^2. \end{aligned}$$

Therefore by applying the formula of multiple variable Taylor series approximation, we can have:

$$\begin{aligned} \hat{\rho}_h(Y_{1:N}) &\approx \frac{\mu_A}{\mu_B} + (A - \mu_A) \frac{1}{\mu_B} - (B - \mu_B) \frac{\mu_A}{\mu_B^2} \\ &= 0 + \frac{A}{\sigma^2} \\ &= \frac{\sum_{n=1}^{N-h} Y_n Y_{n+h}}{N \sigma^2}. \end{aligned}$$

From the above equation, we can see that  $\mu_{\hat{\rho}_h(Y_{1:N})} = 0$ , which means that the sample ACF has the zero mean. Now for the variance of sample ACF:

$$\begin{aligned}\text{Var} \hat{\rho}_h(Y_{1:N}) &= \text{Var} \left[ \frac{\sum_{n=1}^{N-h} Y_n Y_{n+h}}{N\sigma^2} \right] \\ &= \frac{1}{N^2\sigma^4} \sum_{n=1}^{N-h} \text{Var}[Y_n Y_{n+h}] \\ &= \frac{1}{N^2\sigma^4} \sum_{n=1}^{N-h} \{ \mathbb{E}[Y_n^2 Y_{n+h}^2] - \mathbb{E}[Y_n]^2 \mathbb{E}[Y_{n+h}]^2 \} \\ &= \frac{1}{N^2\sigma^4} \sum_{n=1}^{N-h} \text{Var}[Y_n] \text{Var}[Y_{n+h}] \\ &= \frac{1}{N^2\sigma^4} \sum_{n=1}^{N-h} \sigma^4 \\ &= \frac{N-h}{N^2}.\end{aligned}$$

Therefore, when the sample size  $N$  is large, the variance of the sample ACF can be reduced to  $\frac{1}{N}$  and thus the standard deviation to be  $\frac{1}{\sqrt{N}}$ .

**B.**

A confidence interval “gives an estimated range of values which is likely to include an unknown population parameter, the estimated range being calculated from a given set of sample data.”

Seen in the  $R$  command, we notice that it uses definition “coverage probability”. According to the definition of coverage probability, it “refers to the probability that a procedure for constructing random regions will produce an interval containing, or covering, the true value. It is a property of the interval producing procedure, and is independent of the particular sample to which such a procedure is applied.”

Therefore we can see the difference of these two terms: the confidence interval makes heavy use of the known sample data, it is constructed by covering repeated sampling. But coverage probability we derive, is made use of the property of the model itself, like in the above example, we approximated and found the standard deviation of sample ACF to be  $\frac{1}{\sqrt{N}}$  when the sample size is large enough. In this case, the coverage probability is totally based on the mean and variance we derive, but has nothing to do with a true data sample.

Therefore in conclusion, these lines do not construct a confidence interval actually.

**Question 1.3.**

**Question 1.1.**

For this assignment, I did Question 1.1. totally by myself without any reference since it is quite simple.

**Question 1.2 Part A**

I use the resource for Taylor series approximation for multivariables from the url <http://www.math.ucdenver.edu/~esulliva/Calculus3/Taylor.pdf>;

to better understand the sample ACF concepts, I read materials from [https://www.stat.berkeley.edu/~arturof/Teaching/STAT248/lab05\\_part2.html](https://www.stat.berkeley.edu/~arturof/Teaching/STAT248/lab05_part2.html);

and to refresh the variance of two producted random variables, I looked into a homework assignment from NYU at <http://people.stern.nyu.edu/wgreene/MathStat/Problem3-Solutions.pdf>.

**Question 1.2 Part B**

I get the definition of confidence interval from <http://www.stat.yale.edu/Courses/1997-98/101/confint.htm>,

and the definition of coverage probability together with the differences between them at <https://www.stats.ox.ac.uk/pub/bdr/IAUL/Course1Notes5.pdf>.