

# Homework – Week 2

Dominic Pazzula

## Question 1

Given the dataset in problem1.csv:

- calculate the first four moments values by using normalized formula in the "Week1 - Univariate Stats".
- calculate the first four moments values again by using your chosen statistical package
- Is your statistical package functions biased? Prove or disprove your hypothesis. Explain your conclusion.

The table below shows my self-implemented functions which are the biased estimators for Variance, Skewness, and Kurtosis. It then shows the values calculated by the Julia functions. Finally it shows the difference between the two:

	Mean	Variance	Skewness	Kurtosis
My Biased Functions	1.04897039	5.421793461	0.880608643	23.12220079
Julia Function	1.04897039	5.427220682	0.880608643	23.12220079
Diff	0	-0.00542722	-9.992E-16	0

The only difference between the functions is in the Variance. This means that the Skewness and Kurtosis functions as implemented by Julia are unbiased.

The difference in the variance leads me to believe that the Julia function is unbiased. I can check this by multiplying my value by  $n/(n-1)$ .  $N=1000$  so that ratio is 1.001001001001.

$$5.421793461 * 1.001001001001 = 5.427220681881728$$

That is the Julia value, therefore the Julia value is unbiased.

## Question 2

- Fit using OLS and MLE using a normal assumption. Compare Beta values and std of the errors vs fitted std from MLE.

	B0	B1	std (error)
OLS	-0.0873845	0.775274099	1.00627516
MLE Normal	-0.0873845	0.775274099	1.00375632

The fitted betas are identical, as expected. The standard deviation of the OLS error is not the same as the fitted MLE error. This is because the MLE estimator is biased.

- b. Fit with a T distribution. Show the fitted parameters. Compare the fits and determine which fits best.

	B0	B1	ll	AICC
MLE Normal	-0.0873845	0.775274099	-284.5375631	575.1975751
MLE T	-0.0972694	0.675009126	-281.2934032	570.7919346

The MLE with the assumption of t distributed errors fits best because the AICC is lower.

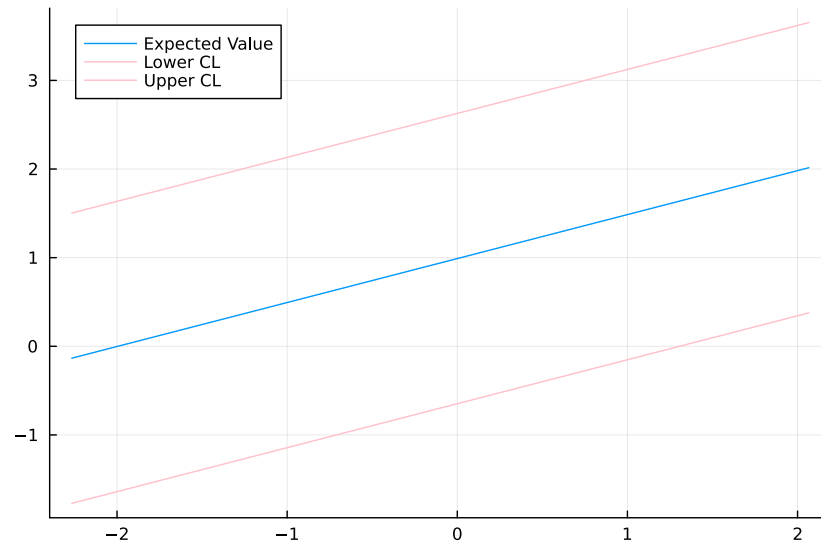
- c. Fit the data in problem2\_x.csv assuming multivariate normality. Given the values in problem2\_x1.csv, what are the expected values and 95% confidence interval.

MLE fitting should give us the same values as fitting with a calculated covariance matrix and mean.

Variable	Statistic	X1	X2
	MEAN	0.001022695	0.990243819
X1	COV	1.06977	0.530685
X2	COV	0.530685	0.961473

The top row is the mean and the 2<sup>nd</sup> and 3<sup>rd</sup> row represents the covariance.

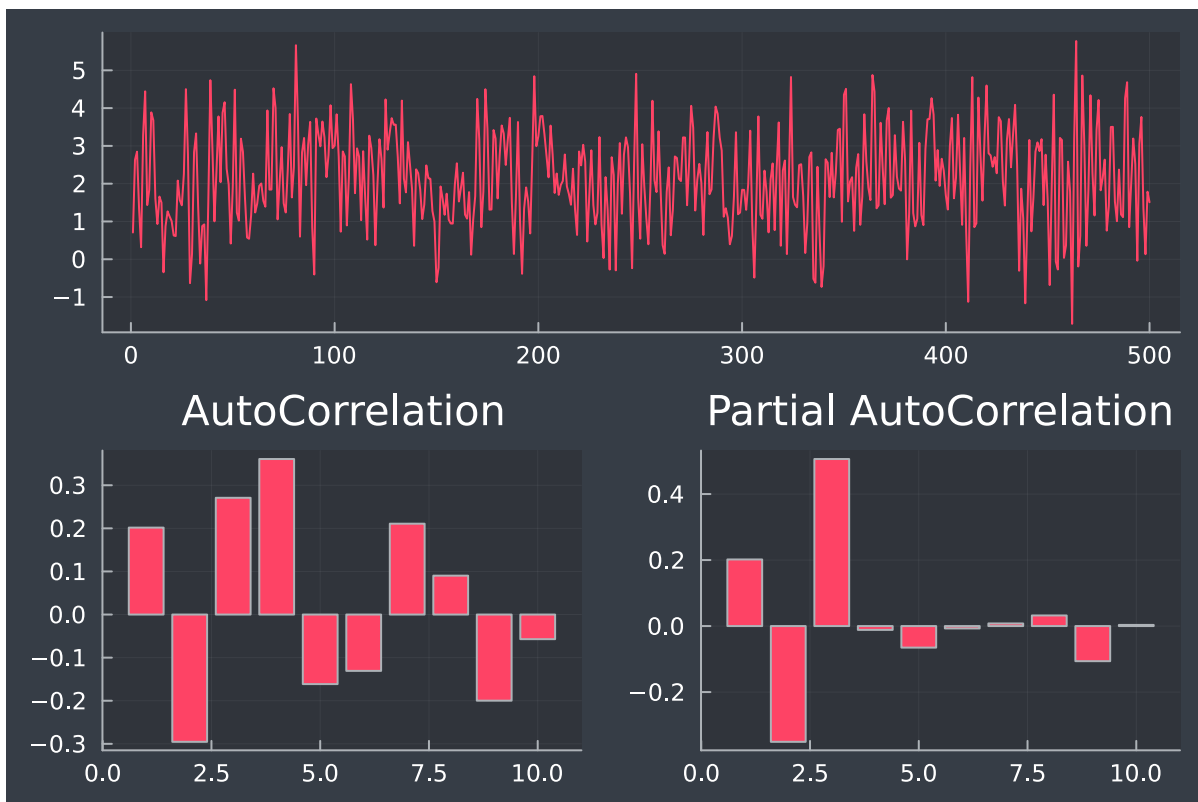
Using the conditional expectation and variance given in the notes, I can calculate the expected values and 95% confidence intervals around that using the quantile function (critical values [.975,.025]). Plotted:



### Problem 3

Fit AR(1) – AR(3) and MA(1) – MA(3) and decide what fits best:

First lets look at the ACF/PACF graphs:



The absolute value of the ACF is slowly decreasing. There are 3 non-zero values in the PACF. I expect this to be an AR(3) model.

I fit the models and output the AICC values:

Model	AICC
AR(1)	1644.7039
AR(2)	1581.1601
<b>AR(3)</b>	<b>1436.7813</b>
MA(1)	1567.452
MA(2)	1538.0219
MA(3)	1536.9892

The AR(3) model has the lowest AICC and therefore fits the best. This is expected from the ACF/PACF graphs.