Stat 436/536 (An Introduction to) Time Series Analysis Fall 2016

Instructor: Mark Greenwood **Email**: greenwood@montana.edu (please use this email)
Office: Wilson 2-228 **Office Phone:** 994-1962 (rarely answered – use email!)

Prerequisites: Stat 436: Stat 411/511 or consent of instructor
Stat 536: (Stat 411/511 **and** consent of instructor) or STAT 506

Recommended: STAT 412/512

Course Description: An introduction to time series analysis using time domain methods (time series regression, autoregressive, moving average, and ARIMA models) and explores estimation, model building and forecasting. Additionally, frequency domain methods (spectral analysis) will be briefly introduced. The goal of the course is to acquire the skills required to investigate data collected as a univariate time series.

Learning outcomes for STAT 436/536:
Upon successful completion of the course, students will be able to:

- graphically explore and describe common features of interest in time series.
- estimate and describe autocorrelation and partial autocorrelation.
- perform model selection, estimation, and interpretation of ARIMA models.
- diagnose potential issues with using regression models for correlated data without accounting for the correlation.
- build regression models for correlated data that account for correlation.

Required Texts (They are available bound and printed in black and white for $24.99 if you go through the "MyCopy" options. It takes about a week to arrive. The pdf version of the book is are available for free. You may need to be on a campus computer or VPN'd in to the campus network to get the books ordered. Let me know if you have any problems!)

Cowpertwait and Metcalfe (2009): Introductory time series with R [electronic resource]
http://link.springer.com/book/10.1007%2F978-0-387-88698-5

Cryer and Chan (2008): Time series analysis [electronic resource]: with applications in R
http://link.springer.com/book/10.1007%2F978-0-387-75959-3

Schedule: T,Th 9:25-10:40 in Wilson 1-144 except holidays (9/27 for MT ASA meeting, 11/8 for voting, and possibly 11/22: I might have activities for the MT ASA and day before Thanksgiving but these likely will not be in-class days).

Office Hours: Monday, Wednesday at 2, and by appointment

Cut-offs for grade assignment:

A 93%	C+ 77%	F <60%
A- 90%	C 73%	
B+ 87%	C- 70%	
B 83%	D+ 67%	
B- 80%	D 60%	

Course Evaluation:
There will be approximately weekly or bi-weekly homeworks, possibly one project, two exams, and a self-graded component

	STAT 436	STAT 536
Midterm Exam	25%	20%
Final Project	---	20%
Final Exam	30%	25%
Homework	40%	30%
Self-Graded Component	5%	5%

HW: We will have approximately 9 homework assignments that will be equally weighted to find your overall score on Homework. Late HW will be 10% off the total possible per day late (including weekends) and will only be accepted late until the assignment is graded. Once grading is completed, no further homework will be accepted unless prior arrangements are made. Email submissions will include a small penalty unless accompanied by a reasonable excuse (needed to be in color to grade, printer not working, etc.). For late HW, this penalty will not be applied. You may work in groups as instructed on particular (portions of) homework assignments.

Exams: Both exams will be take-home exams and provide you with a chance to demonstrate what you have learned and what you can do. They will not be group work. You will have a few days to complete the exam to allow you to work on it around other responsibilities.

Project: Students taking STAT 536 will need to complete a project along with all other components of STAT 436. This project can involve a sophisticated analysis of a data set *of interest to you* using the methods in the course or be a review of methods related to the course but not covered in lectures with an application to a data set. Students in 436 may complete a short project involving a write-up of an analysis of a real data set using the methods discussed in the course to replace their worst homework assignment score.

Self-Graded Component: You will also evaluate your own performance in the course on the following criteria.

- Preparation for class (reading the book and engaging the material).
- Participation in discussions, answering questions when posed, or asking questions outside of class to get clarifications.
- Physical *and mental* attendance in class.
- Amount of presented material that you learned and will retain beyond this semester.

I reserve the right to modify this component of your grade but have yet to need to do that in any courses that I have used this grading component.

Attendance: I expect you to be in class or let me know why you are not there. Whether you are an undergraduate or graduate student, I expect you to treat the class periods with respect – show up prepared and on time or let me know in advance of any issues with attendance.

Academic Dishonesty: It is expected that students will abide by university regulations regarding academic dishonesty. Instructions will be given on each type of activity regarding the use of outside resources and working together.

Group Work: You may work in small groups on some or all of the homework unless otherwise instructed. ***Please note the individual(s) that you have worked with if you are working together and make sure you understand everything that is turned in. If you turn in separate assignments but discussed with other students, you need to that you got help from others and how that help impacted your answers.*** Exams will not be group work.

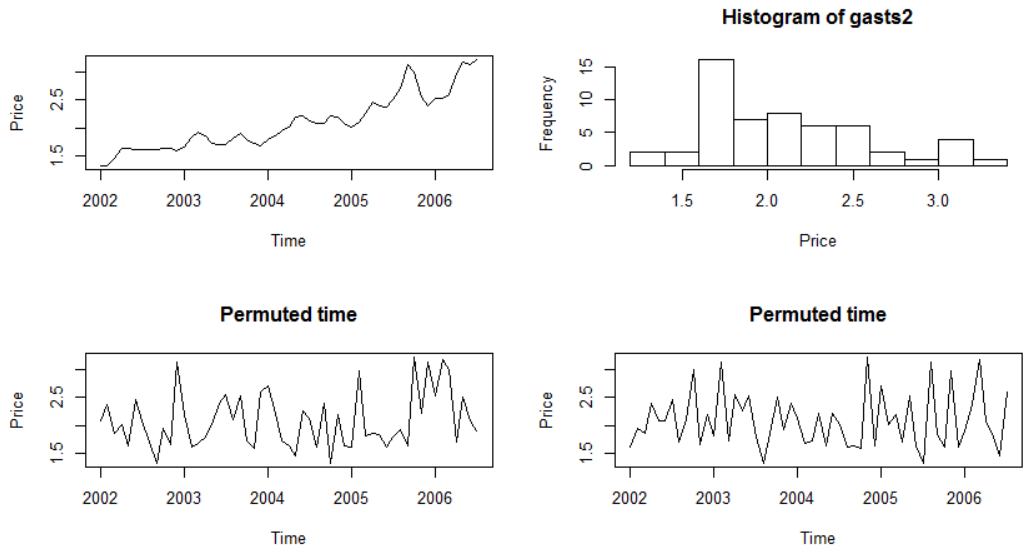
Computer package: We will exclusively use R for the homework and projects. You will need to remember your basic R skills to successfully complete this class. The STAT 217 book (https://scholarworks.montana.edu/xmlui/handle/1/2999) or the end of Cryer and Chan are nice starting points. Make sure you are working from a relatively up to date version of R. I highly recommend R-studio as an interface to R if you are not already using it and using R-markdown for all of your homework and exams.

Outline of Topics:

Cryer/Chan Chapter	Topic	Cowpertwaite/Metcalf Chapter
CC Ch 1	What is Time Series Analysis? (and some basic R tips and plotting information)	CM Ch 1
CC Ch 2	Time Series and expectation, variance and covariance	CM Ch 2
CC Ch 3	Time Series Regression for time series trends/ Mixed models for TS regression (ARMA errors)	CM Ch 5
Supplement on nonparametric regression for time series (Additive Mixed Models)		
CC Ch 4	ARMA models/ Stationary Models	CM Ch 4,6
CC Ch 5	ARIMA models	CM Ch 7
CC Ch 6	ACF and PACF for model diagnostics	CM Ch 2
CC Ch 7	Parameter Estimation	
CC Ch 8	Diagnostics	
CC Ch 9	Forecasting using ARIMA models	
CC Ch 10	Seasonal ARIMA models	
CC Ch 11	Time Series Regression (Intervention/ Outlier) models	
Dynamic linear models/ State space models		CM Ch 12
CC Ch 13	Spectral Analysis	CM Ch 9
CC Ch 14	Estimating the Spectrum	CM Ch 9 (10?)
Additional (Supplemental) Topics (if time allows and depending on student interest):		
	Wavelets and dynamic Fourier Analysis	
	Multivariate Time Series (Spectral and/or VAR)	
	Generalized linear ARMA models	
CC Ch 12	ARCH and GARCH models	CM Ch 7
	Long Memory Processes (Fractional differencing)	CM Ch 8
CC Ch 15	Threshold Models	

```
> gas<-read.table("D:/usa.txt")
> colnames(gas)<-c("Date","Price")
> head(gas)
```

```
        Date Price
1 09/15/1981 1.471
2 10/15/1981 1.470
3 11/15/1981 1.470
4 12/15/1981 1.468
5 01/15/1982 1.466
6 02/15/1982 1.448
> tail(gas)
          Date Price
294 02/15/2006 2.519
295 03/15/2006 2.603
296 04/15/2006 2.967
297 05/15/2006 3.169
298 06/15/2006 3.139
299 07/15/2006 3.219
> require(car)
> some(gas)
          Date Price
8   04/15/1982 1.351
38  10/15/1984 1.365
66  02/15/1987 1.047
163 03/15/1995 1.306
168 08/15/1995 1.352
193 09/15/1997 1.458
213 05/15/1999 1.370
218 10/15/1999 1.464
232 12/15/2000 1.679
284 04/15/2005 2.468
> gasts2<-ts(gas$Price[245:299],frequency=12,start=c(2002,1),end=c(2006,7))
> par(mfrow=c(2,2))
> ts.plot(gasts2,ylab="Price")
> hist(gasts2,xlab="Price")
> require(mosaic)
> set.seed(97531)
> ts.plot(ts(shuffle(gas$Price[245:299]),frequency=12,start=c(2002,1),end=c(2006,7
)),ylab="Price",main="Permuted time")
> ts.plot(ts(shuffle(gas$Price[245:299]),frequency=12,start=c(2002,1),end=c(2006,7
)),ylab="Price",main="Permuted time")
```



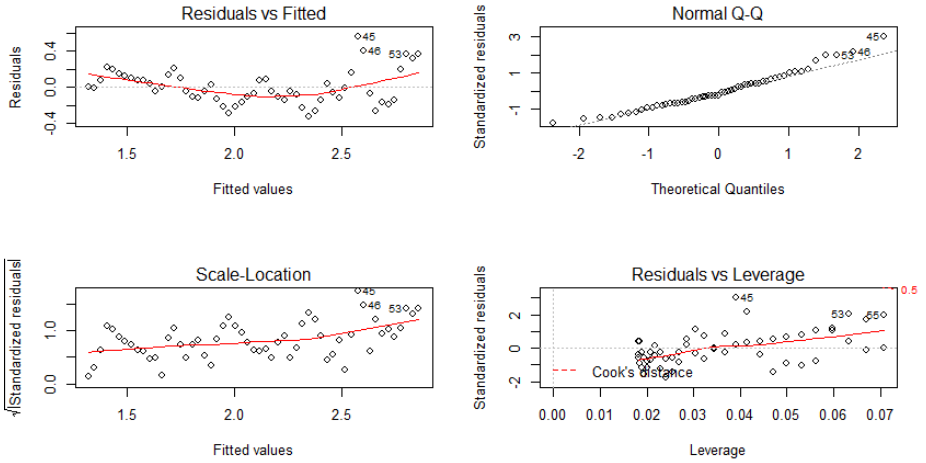
- How strong is the evidence for the trend?

- You should be able to note at least two issues in the following statement regarding the previous response variable:

  "The histogram suggests that my data is not normal, so I am going to log-transform it."

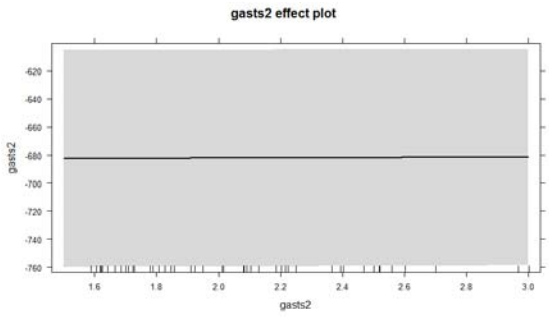
```
> time(gasts2)[1:25]
 [1] 2002.000 2002.083 2002.167 2002.250 2002.333 2002.417
 [7] 2002.500 2002.583 2002.667 2002.750 2002.833 2002.917
[13] 2003.000 2003.083 2003.167 2003.250 2003.333 2003.417
[19] 2003.500 2003.583 2003.667 2003.750 2003.833 2003.917
[25] 2004.000
> m1<-lm(gasts2~time(gasts2))
> summary(m1)
                Estimate Std. Error t value Pr(>|t|)
(Intercept)  -682.69393   38.55906  -17.70   <2e-16 ***
time(gasts2)    0.34166    0.01924   17.76   <2e-16 ***

Residual standard error: 0.1887 on 53 degrees of freedom
Multiple R-squared:  0.8561,   Adjusted R-squared:  0.8534
F-statistic: 315.4 on 1 and 53 DF,  p-value: < 2.2e-16
> par(mfrow=c(2,2))
> plot(m1)
```

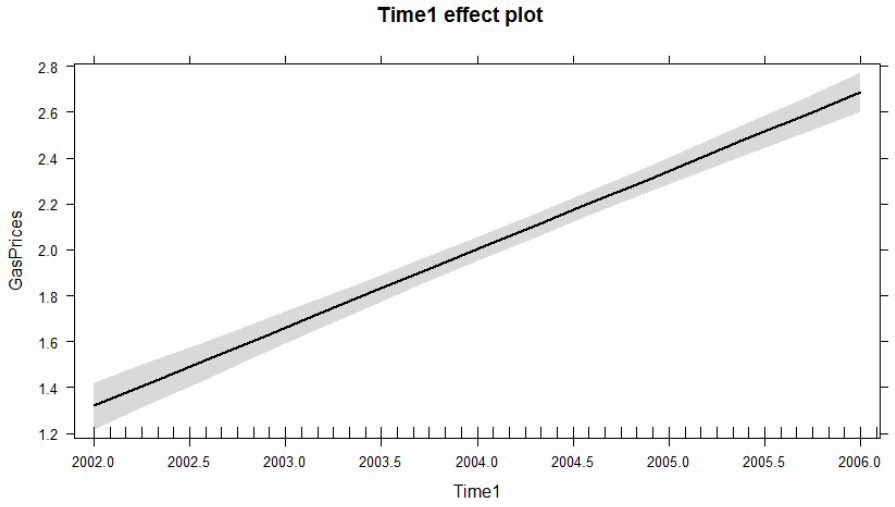



```
> require(effects)
> plot(allEffects(m1))

> data1<-data.frame(GasPrices=gasts2,Time1=time(gasts2
))
> head(data1)
  GasPrices    Time1
1     1.323 2002.000
2     1.330 2002.083
3     1.450 2002.167
4     1.622 2002.250
5     1.625 2002.333
```



```
6     1.606 2002.417
> m2<-lm(GasPrices~Time1,data=data1)
> plot(allEffects(m2))
```




```
> #Permutation test:
> B<- 1000
> Tstar<-matrix(NA,nrow=B)
> for (b in (1:B)){
+   Tstar[b]<-lm(GasPrices~shuffle(Time1),data=data1)$coefficients[2]
+ }
> par(mfrow=c(1,2))
> Tobs<-lm(GasPrices~Time1,data=data1)$coefficients[2]
> Tobs
    Time1
0.3416649
> hist(Tstar,labels=T)
> abline(v=Tobs,lwd=2,col="red")
> plot(density(Tstar),main="Density curve of Tstar")
> abline(v=Tobs,lwd=2,col="red")
> pdata(abs(Tstar),abs(Tobs),lower.tail=F)
Time1
    0
```