4th Assignment: Title

David Ribberholt Ipsen (s164522), Kasper .... (s....), Michelle ... (s....)

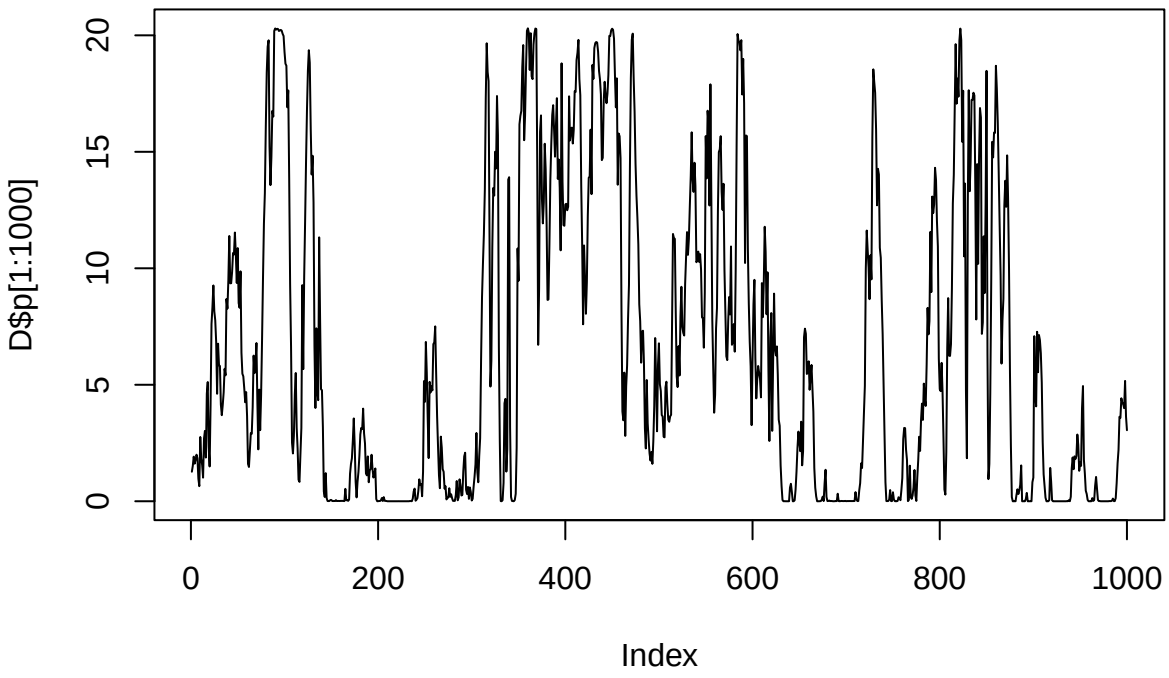
November 2021

Contents

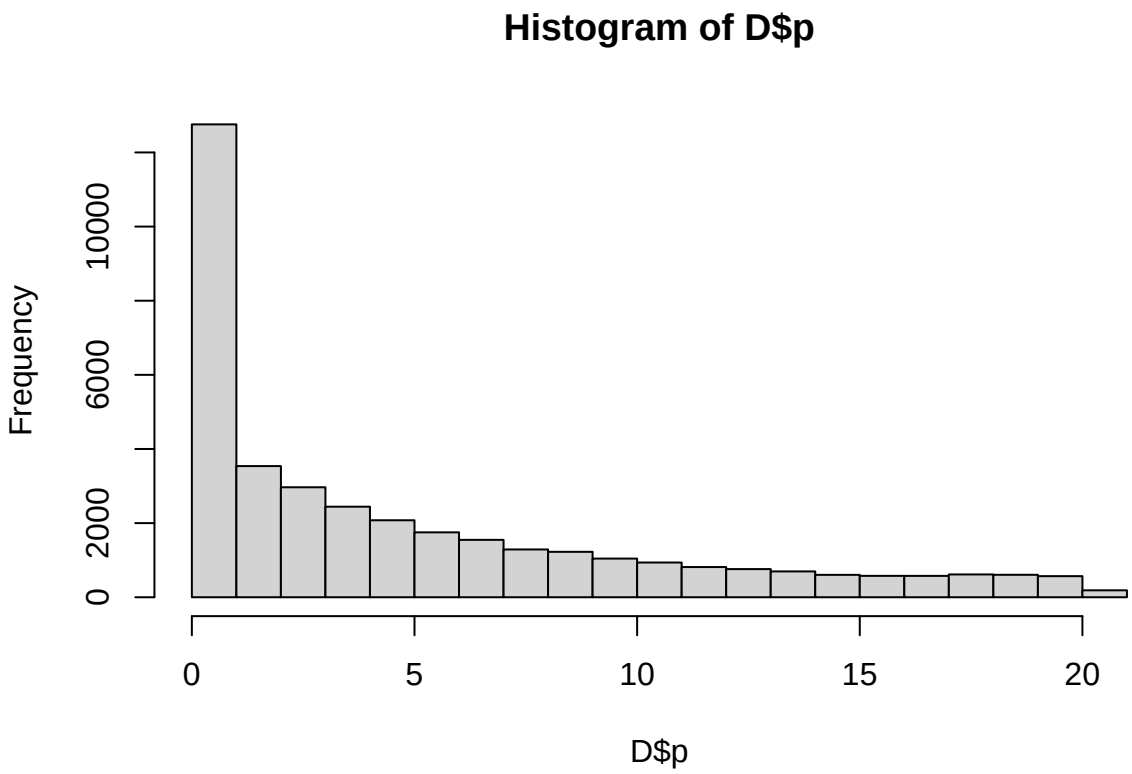
1. Introduction

1.1 Descriptive statistics

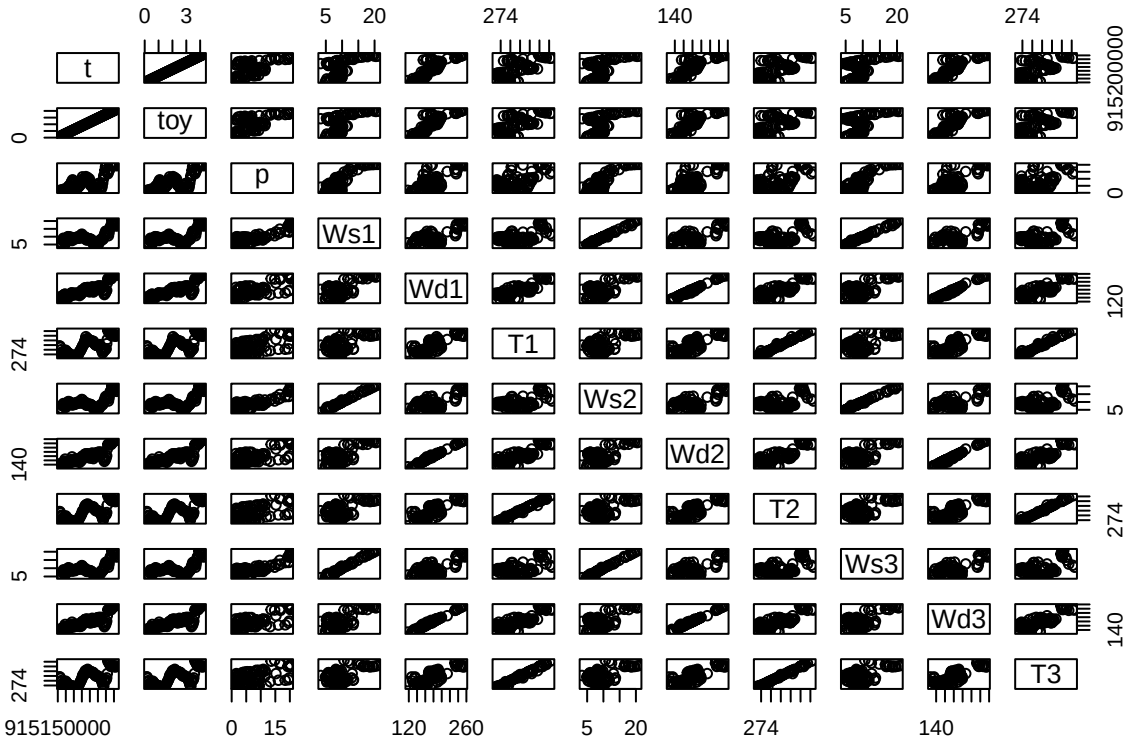
```
if (Sys.info()[7] == "davidipsen"){
  setwd("~/Documents/DTU/3. Semester (MSc)/Advanced Time Series/Assignments/4-Assignment")
} else if (Sys.info()[7] == "Hvad end Kaspers PC hedder-sikkert noget UNIX-hejs ^_^"){
}

D <- read.table("comp_ex_4_scripts_2011/data/cex4WindDataInterpolated.csv", sep=",",
                header=TRUE, stringsAsFactors=FALSE)
D$t <- as.POSIXct(D$t, tz="UTC")
n = dim(D)[1]

# Descriptive statistics
plot(D$p[1:1000],type='l')
```



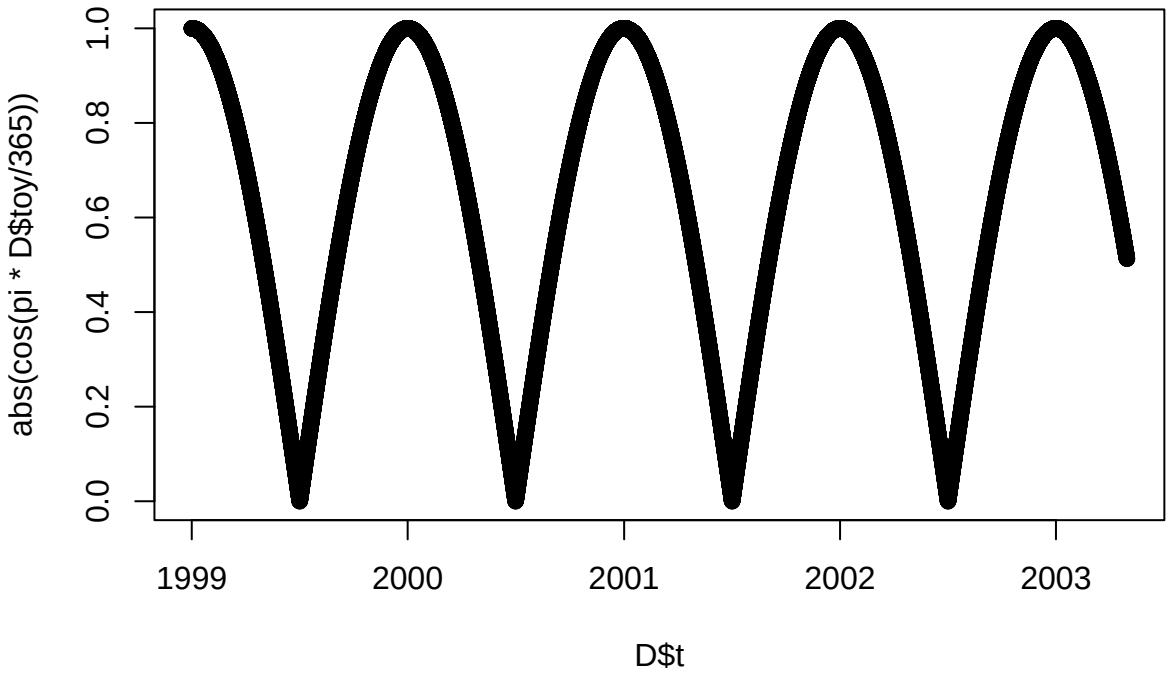

```
hist(D$p)
```




```
plot(D[1:100,]) # Plot subset of from all variables
```




```
plot(D$t, abs(cos(pi*D$toy/365))) # Utilise the time-of-year variable as follows, decsribing the degree
```



2. BENCHMARKS

Let's mark some benches, bitches

```
# Mean Model
fit1 = lm(p ~ 1, data=D); fit1; logLik(fit1)
```

```
##
## Call:
## lm(formula = p ~ 1, data = D)
##
## Coefficients:
## (Intercept)
##       4.912

## 'log Lik.' -117332.2 (df=2)
```

```
# AR(1)
fit2 = ar(D$p, na.action = na.pass, order.max=1); fit2;
```

```
##
## Call:
## ar(x = D$p, order.max = 1, na.action = na.pass)
```

```
##
## Coefficients:
##      1
## 0.9544
##
## Order selected 1  sigma^2 estimated as  2.685
```

```
# Linear Regression
fit3 = lm(p ~ ., data=D[,-1]) ; summary(fit3); logLik(fit3)
```

```
##
## Call:
## lm(formula = p ~ ., data = D[, -1])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -35.208  -1.920  -0.247   1.688  19.852
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -9.6948533  0.7912322 -12.253  < 2e-16 ***
## toy         -0.0033685  0.0001615 -20.853  < 2e-16 ***
## Ws1          0.5428779  0.0360215  15.071  < 2e-16 ***
## Wd1         -0.0002256  0.0008518  -0.265 0.791134
## T1           0.2599813  0.0755274   3.442 0.000578 ***
## Ws2          0.1069968  0.0510650   2.095 0.036151 *
## Wd2          0.0003192  0.0012187   0.262 0.793405
## T2          -0.0687244  0.1068469  -0.643 0.520095
## Ws3          0.4404737  0.0363361  12.122  < 2e-16 ***
## Wd3          0.0008769  0.0008821   0.994 0.320164
## T3          -0.1703619  0.0759947  -2.242 0.024983 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.049 on 36272 degrees of freedom
##   (1658 observations deleted due to missingness)
## Multiple R-squared:  0.6924, Adjusted R-squared:  0.6923
## F-statistic:  8164 on 10 and 36272 DF,  p-value: < 2.2e-16
```

```
## 'log Lik.' -91923.63 (df=12)
```

```
# AR(1)X
fit4 = lm(p ~ . + D$p[-n], data=D[2:n,-1]); summary(fit4); logLik(fit4)
```

```
##
## Call:
## lm(formula = p ~ . + D$p[-n], data = D[2:n, -1])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.4375  -0.6794  -0.0147   0.5726  13.1556
##
```

```
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.673e+00  4.148e-01  -4.032 5.53e-05 ***
## toy         -5.633e-04  8.502e-05  -6.626 3.49e-11 ***
## Ws1          1.277e-01  1.890e-02   6.757 1.43e-11 ***
## Wd1         -3.281e-04  4.457e-04  -0.736  0.46163
## T1           7.424e-02  3.954e-02   1.878  0.06044 .
## Ws2         -1.867e-02  2.672e-02  -0.699  0.48473
## Wd2          8.040e-05  6.377e-04   0.126  0.89966
## T2          -3.110e-03  5.592e-02  -0.056  0.95564
## Ws3          5.842e-02  1.906e-02   3.066  0.00217 **
## Wd3         -5.116e-05  4.615e-04  -0.111  0.91175
## T3          -6.699e-02  3.976e-02  -1.685  0.09209 .
## D$p[-n]      8.493e-01  2.738e-03 310.134  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.595 on 36265 degrees of freedom
##   (1663 observations deleted due to missingness)
## Multiple R-squared:  0.9158, Adjusted R-squared:  0.9157
## F-statistic: 3.585e+04 on 11 and 36265 DF,  p-value: < 2.2e-16

## 'log Lik.' -68410.06 (df=13)
```

```
# Thoughts:
# 1. Wind has a logistic effect for the power: ... + logit(wind)
# 2. toy = abs(cos(pi*D$toy/365))
```

Problem MANGLER: Fucking NaN

3. AUTOREGRESSIVE MARKOV-SWITCHING MODEL (MSM-AR)

"*A Markov-switching (also regime-switching) model is a general- ization both of Markov models and AR(p) processes. It can be seen as an autoregressive model with a state-dependent mean and variance where the states follow a Markov process. An AR(1) Markov- switching process is given by*" (REF JAN KLOP) (REF: https://cran.r-project.org/web/packages/MSwM/vignettes/examples.pdf)

- regime-dependant AR-coefficient(s).

```
#install.packages('MSwM')
library(MSwM)
```

```
## Loading required package: parallel
```

```
library(parallel)
Dsub = D[1:200,] # Subset data for dev

# Simple Linear Regression
fit2b = lm(p ~ Ws1, data=Dsub) ; summary(fit2b); logLik(fit2b)
```

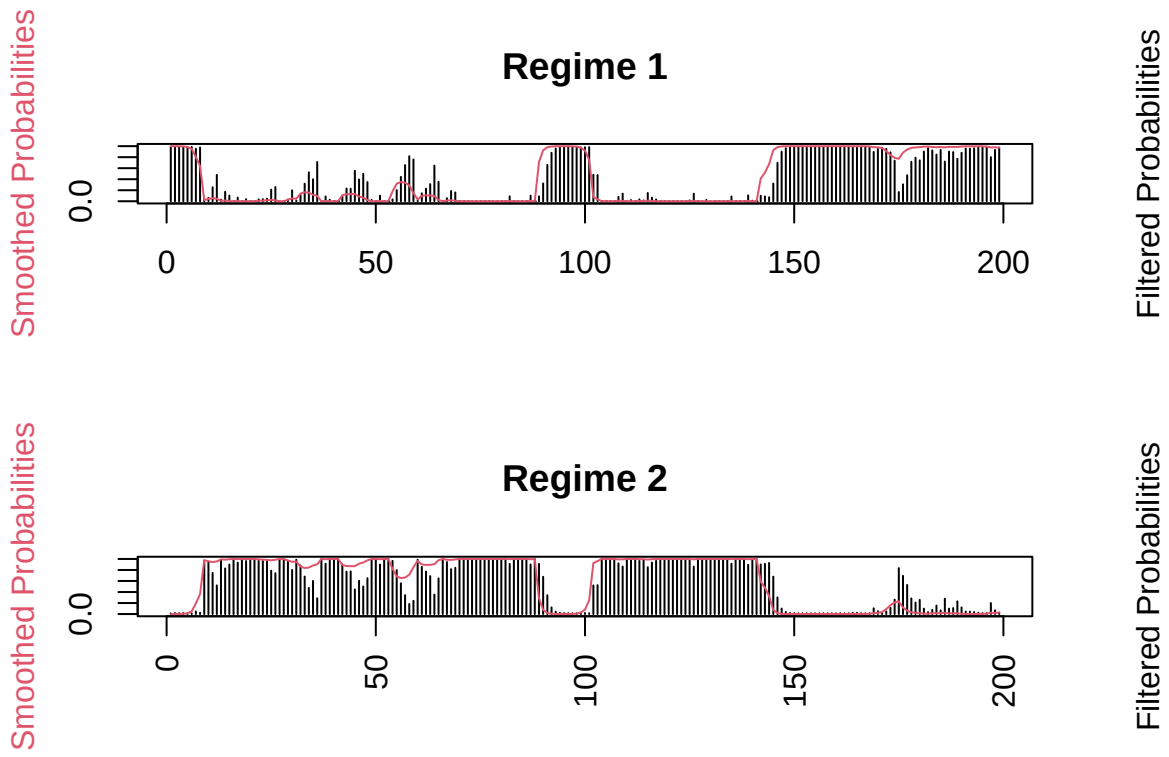
```
##
## Call:
## lm(formula = p ~ Ws1, data = Dsub)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.6299 -2.1934 -0.3328  2.4609  8.4385
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.55666    0.59614  -11.00   <2e-16 ***
## Ws1          1.47883    0.06343   23.32   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.236 on 198 degrees of freedom
## Multiple R-squared:  0.733,  Adjusted R-squared:  0.7317
## F-statistic: 543.6 on 1 and 198 DF,  p-value: < 2.2e-16

## 'log Lik.' -517.6688 (df=3)
```

```
# Fit MSM
fit_msm = msmFit(fit2b, k=2, p=1, sw=rep(TRUE,4), control = list(parallel = TRUE))
summary(fit_msm)
```

```
## Markov Switching Model
##
## Call: msmFit(object = fit2b, k = 2, sw = rep(TRUE, 4), p = 1, control = list(parallel = TRUE))
##
##       AIC      BIC    logLik
##   684.559 736.0786 -336.2795
##
## Coefficients:
##
## Regime 1
## ---------
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)(S)  -0.1274     0.2094 -0.6084   0.5429
## Ws1(S)           0.0252     0.0368  0.6848   0.4935
## p_1(S)           0.9756     0.0254 38.4094   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5255476
## Multiple R-squared: 0.9943
##
## Standardized Residuals:
##           Min            Q1           Med            Q3           Max
## -1.3696947746 -0.0458646867  0.0001212949  0.0794490057  1.0639433062
##
## Regime 2
## ---------
##                Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)(S)  -0.8326     0.8069 -1.0319  0.30212
## Ws1(S)           0.2303     0.1188  1.9386  0.05255 .
## p_1(S)           0.8354     0.0588 14.2075  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.076205
## Multiple R-squared: 0.846
##
## Standardized Residuals:
##         Min          Q1         Med          Q3         Max
## -5.93973524 -0.70323578 -0.01450194  0.47824011  6.48764380
##
## Transition probabilities:
##           Regime 1   Regime 2
## Regime 1 0.9577938 0.02797245
## Regime 2 0.0422062 0.97202755
```

```
plotProb(fit_msm, which=1)
```




```
#plotProb(fit_msm, which=2)
#plotProb(fit_msm, which=3)
```

DESCRETIZING (EULER-MARYAMA)

Continous-Descrete State Space Model (with adaptivity) Thoughts

1. + ... rX(1-X/CAP) | W = wind, X = True wind power
2. r = f()