PV data - Parameter estimation and sample generation

Martina Kuchlbauer

Tuesday, May 17, 2016

We consider the data given in PVdata2.csv. Let's generate a matrix where each line represents a day and each column represents one minute of this day:

```
PV1<-PVdata[1:1440,1] #24*60=1440
for (i in 1:30) {
  PV1<-cbind(PV1,PVdata[((i*1440)+1):((i+1)*1440),1])
}
for(j in 2:12){
  for(i in 1:31){
    PV1<-cbind(PV1,PVdata[(((i-1)*1440)+1):(i*1440),j])
  }
}
PV1<-t(PV1)
```

1 Normal distribution

1.1 Parameter estimation

1.1.1 Parameter estimation - multivariate (dependent) for 1h intervals

We estimate the values of expectation and the covariance matrix under the assumption of a **multivariate** normal distribution for intervals of 1h:

```
PV1h<-matrix(rep(0,8928),nrow=372) #hourly values -> take means, 24*372=8928
for(i in 1:372){
  for (j in 1:24){
    PV1h[i,j]<-mean(PV1[i,((j-1)*60+1):(j*60)])
  }
}

estimates_n_dep<-mlest(PV1h) #under assumption of no independence: hourly means and covariance matrix
```

```
## Warning: NA/Inf durch größte positive Zahl ersetzt
```

```
estimates_n_dep$muhat #estimate of mean (mu)
```

```
##  [1] -3.911e-10 -3.911e-10 -3.911e-10 -3.911e-10 -3.911e-10  4.570e-01
##  [7]  1.288e+02  6.744e+02  1.279e+03  1.791e+03  2.132e+03  2.261e+03
## [13]  2.285e+03  2.118e+03  1.832e+03  1.383e+03  8.038e+02  2.254e+02
## [19]  4.466e+00 -3.911e-10 -3.911e-10 -3.911e-10 -3.911e-10 -3.911e-10
```

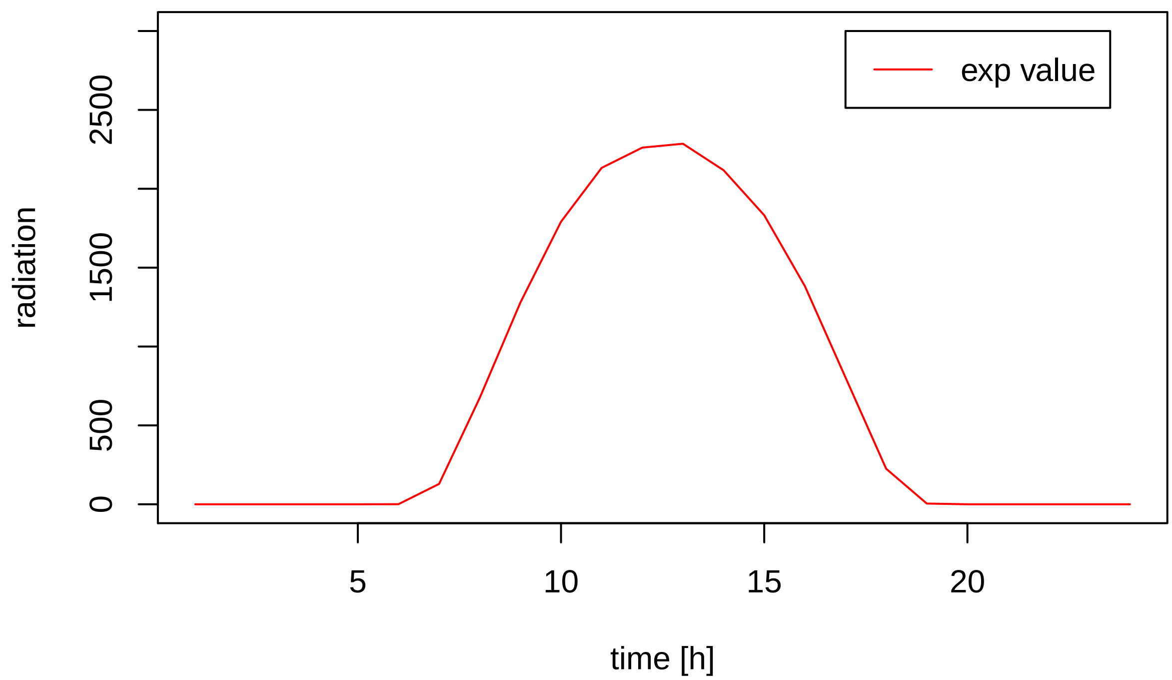
```r
estimates_n_dep$sigmahat #estimate of covariance matrix (sigma)
```

```
##            [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
##  [1,] 0.0009987 0.0000000 0.0000000 0.0000000 0.0000000 0.000e+00
##  [2,] 0.0000000 0.0009987 0.0000000 0.0000000 0.0000000 0.000e+00
##  [3,] 0.0000000 0.0000000 0.0009987 0.0000000 0.0000000 0.000e+00
##  [4,] 0.0000000 0.0000000 0.0000000 0.0009987 0.0000000 0.000e+00
##  [5,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0009987 0.000e+00
##  [6,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 1.541e+00
##  [7,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 8.303e+01
##  [8,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 1.083e+02
##  [9,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 8.063e+01
## [10,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 7.331e+01
## [11,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 4.836e+01
## [12,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 1.875e+01
## [13,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 4.147e+01
## [14,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 8.259e+01
## [15,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 8.120e+01
## [16,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 2.553e+01
## [17,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 5.759e+01
## [18,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 6.893e+01
## [19,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 3.182e+00
## [20,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 2.960e-06
## [21,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 2.960e-06
## [22,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 2.960e-06
## [23,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 2.960e-06
## [24,] 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 2.960e-06
##            [,7]      [,8]       [,9]      [,10]      [,11]      [,12]
##  [1,] 0.000e+00 0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [2,] 0.000e+00 0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [3,] 0.000e+00 0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [4,] 0.000e+00 0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [5,] 0.000e+00 0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [6,] 8.303e+01 1.083e+02  8.063e+01  7.331e+01  4.836e+01  1.875e+01
##  [7,] 1.532e+04 2.636e+04  3.182e+04  3.389e+04  3.457e+04  2.143e+04
##  [8,] 2.636e+04 9.441e+04  1.436e+05  1.717e+05  1.886e+05  1.546e+05
##  [9,] 3.182e+04 1.436e+05  2.726e+05  3.240e+05  3.641e+05  3.079e+05
## [10,] 3.389e+04 1.717e+05  3.240e+05  4.488e+05  5.008e+05  4.393e+05
## [11,] 3.457e+04 1.886e+05  3.641e+05  5.008e+05  6.195e+05  5.405e+05
## [12,] 2.143e+04 1.546e+05  3.079e+05  4.393e+05  5.405e+05  6.407e+05
## [13,] 2.289e+04 1.451e+05  2.833e+05  3.963e+05  4.854e+05  5.971e+05
## [14,] 2.320e+04 1.392e+05  2.620e+05  3.497e+05  4.282e+05  5.272e+05
## [15,] 2.045e+04 1.193e+05  2.199e+05  2.725e+05  3.439e+05  4.228e+05
## [16,] 1.622e+04 8.955e+04  1.695e+05  2.114e+05  2.654e+05  3.225e+05
## [17,] 1.523e+04 5.826e+04  9.817e+04  1.245e+05  1.569e+05  1.792e+05
## [18,] 1.248e+04 2.244e+04  2.960e+04  3.643e+04  4.377e+04  4.455e+04
## [19,] 4.035e+02 2.813e+02 -1.824e+02 -2.871e+02 -5.557e+02 -4.369e+02
## [20,] 1.322e-03 7.512e-03  1.435e-02  1.937e-02  2.347e-02  2.542e-02
## [21,] 1.322e-03 7.512e-03  1.435e-02  1.937e-02  2.347e-02  2.542e-02
## [22,] 1.322e-03 7.512e-03  1.435e-02  1.937e-02  2.347e-02  2.542e-02
## [23,] 1.322e-03 7.512e-03  1.435e-02  1.937e-02  2.347e-02  2.542e-02
## [24,] 1.322e-03 7.512e-03  1.435e-02  1.937e-02  2.347e-02  2.542e-02
##           [,13]     [,14]     [,15]     [,16]     [,17]     [,18]
```

```
##  [1,] 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00
##  [2,] 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00
##  [3,] 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00
##  [4,] 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00
##  [5,] 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00 0.000e+00
##  [6,] 4.147e+01 8.259e+01 8.120e+01 2.553e+01 5.759e+01 6.893e+01
##  [7,] 2.289e+04 2.320e+04 2.045e+04 1.622e+04 1.523e+04 1.248e+04
##  [8,] 1.451e+05 1.392e+05 1.193e+05 8.955e+04 5.826e+04 2.244e+04
##  [9,] 2.833e+05 2.620e+05 2.199e+05 1.695e+05 9.817e+04 2.960e+04
## [10,] 3.963e+05 3.497e+05 2.725e+05 2.114e+05 1.245e+05 3.643e+04
## [11,] 4.854e+05 4.282e+05 3.439e+05 2.654e+05 1.569e+05 4.377e+04
## [12,] 5.971e+05 5.272e+05 4.228e+05 3.225e+05 1.792e+05 4.455e+04
## [13,] 6.456e+05 5.756e+05 4.565e+05 3.565e+05 1.974e+05 5.327e+04
## [14,] 5.756e+05 5.959e+05 4.938e+05 3.697e+05 2.131e+05 5.644e+04
## [15,] 4.565e+05 4.938e+05 4.765e+05 3.477e+05 2.049e+05 5.634e+04
## [16,] 3.565e+05 3.697e+05 3.477e+05 3.023e+05 1.750e+05 5.464e+04
## [17,] 1.974e+05 2.131e+05 2.049e+05 1.750e+05 1.267e+05 4.951e+04
## [18,] 5.327e+04 5.644e+04 5.634e+04 5.464e+04 4.951e+04 3.376e+04
## [19,] 1.246e+02 1.272e+02 4.243e+02 4.549e+02 1.086e+03 1.014e+03
## [20,] 2.541e-02 2.398e-02 2.003e-02 1.542e-02 8.901e-03 2.477e-03
## [21,] 2.541e-02 2.398e-02 2.003e-02 1.542e-02 8.901e-03 2.477e-03
## [22,] 2.541e-02 2.398e-02 2.003e-02 1.542e-02 8.901e-03 2.477e-03
## [23,] 2.541e-02 2.398e-02 2.003e-02 1.542e-02 8.901e-03 2.477e-03
## [24,] 2.541e-02 2.398e-02 2.003e-02 1.542e-02 8.901e-03 2.477e-03
##            [,19]      [,20]      [,21]      [,22]      [,23]      [,24]
##  [1,]  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [2,]  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [3,]  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [4,]  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [5,]  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00  0.000e+00
##  [6,]  3.182e+00  2.960e-06  2.960e-06  2.960e-06  2.960e-06  2.960e-06
##  [7,]  4.035e+02  1.322e-03  1.322e-03  1.322e-03  1.322e-03  1.322e-03
##  [8,]  2.813e+02  7.512e-03  7.512e-03  7.512e-03  7.512e-03  7.512e-03
##  [9,] -1.824e+02  1.435e-02  1.435e-02  1.435e-02  1.435e-02  1.435e-02
## [10,] -2.871e+02  1.937e-02  1.937e-02  1.937e-02  1.937e-02  1.937e-02
## [11,] -5.557e+02  2.347e-02  2.347e-02  2.347e-02  2.347e-02  2.347e-02
## [12,] -4.369e+02  2.542e-02  2.542e-02  2.542e-02  2.542e-02  2.542e-02
## [13,]  1.246e+02  2.541e-02  2.541e-02  2.541e-02  2.541e-02  2.541e-02
## [14,]  1.272e+02  2.398e-02  2.398e-02  2.398e-02  2.398e-02  2.398e-02
## [15,]  4.243e+02  2.003e-02  2.003e-02  2.003e-02  2.003e-02  2.003e-02
## [16,]  4.549e+02  1.542e-02  1.542e-02  1.542e-02  1.542e-02  1.542e-02
## [17,]  1.086e+03  8.901e-03  8.901e-03  8.901e-03  8.901e-03  8.901e-03
## [18,]  1.014e+03  2.477e-03  2.477e-03  2.477e-03  2.477e-03  2.477e-03
## [19,]  9.792e+01 -1.271e-06 -1.271e-06 -1.271e-06 -1.271e-06 -1.271e-06
## [20,] -1.271e-06  9.987e-04  1.111e-09  1.111e-09  1.111e-09  1.111e-09
## [21,] -1.271e-06  1.111e-09  9.987e-04  1.111e-09  1.111e-09  1.111e-09
## [22,] -1.271e-06  1.111e-09  1.111e-09  9.987e-04  1.111e-09  1.111e-09
## [23,] -1.271e-06  1.111e-09  1.111e-09  1.111e-09  9.987e-04  1.111e-09
## [24,] -1.271e-06  1.111e-09  1.111e-09  1.111e-09  1.111e-09  9.987e-04
```

Let's visualize the expected value we estimated:

```r
plot(estimates_n_dep$muhat,xlab="time [h]", ylab="radiation", type="l", col="red", ylim=c(0,3000))
legend(17,3000, c("exp value"), col=c("red"), lty=c(1))
```



1.1.2 Parameter estimation - independent for 15min intervals

Using the package above it is not possible to analyse a multivariate normal distribution with more than 50 variables, hence we continue the analysis with the assumption of **independently** distributed variables for every 15 minutes.

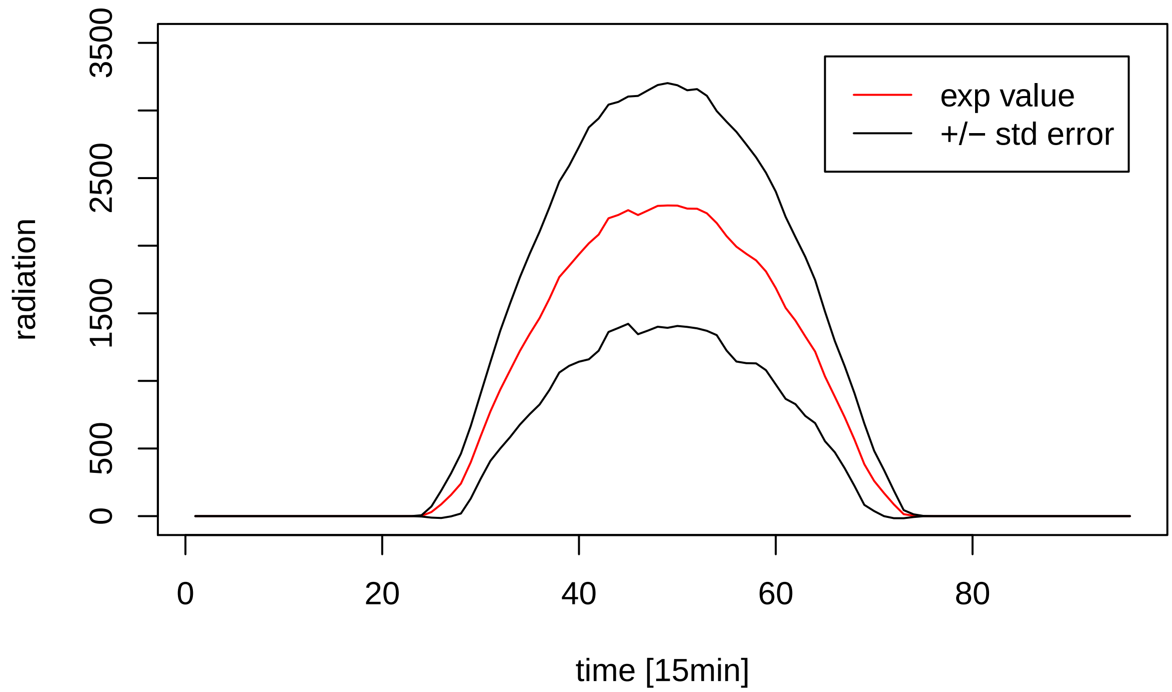
```r
PV1quh<-matrix(rep(0,(8928*4)),nrow=372) #quarter hourly values -> take means, 24*372=8928
for(i in 1:372){
  for (j in 1:(24*4)){
    PV1quh[i,j]<-mean(PV1[i,((j-1)*15+1):(j*15)])
  }
}

estimates_n_ind<-matrix(rep(0,2*96),nrow=2) #under assumption of independence, 4*24=96=T
for (i in 1:(24*4)){
  estimates_n_ind[1,i]<-fitdistr(PV1quh[,i],"normal")$estimate[1] #estimate of mean (mu)
  estimates_n_ind[2,i]<-fitdistr(PV1quh[,i],"normal")$estimate[2] #estimate of std error (sigma)
  }
estimates_n_ind
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13]
## [1,]    0    0    0    0    0    0    0    0    0     0     0     0     0
## [2,]    0    0    0    0    0    0    0    0    0     0     0     0     0
##      [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22]  [,23] [,24]
## [1,]     0     0     0     0     0     0     0     0     0 0.0717 1.756
## [2,]     0     0     0     0     0     0     0     0     0 0.3018 4.772
##      [,25]  [,26] [,27] [,28] [,29] [,30] [,31] [,32]  [,33]  [,34]  [,35]
## [1,] 30.15  87.52 157.3 240.3 398.0 589.5 774.5 935.5 1079.0 1222.2 1348.3
## [2,] 41.20 101.59 159.1 221.2 268.2 314.9 365.6 435.5  494.4  545.2  593.4
##       [,36]  [,37]  [,38]  [,39]  [,40]  [,41] [,42]  [,43]  [,44]  [,45]
## [1,] 1465.0 1608.6 1768.1 1851.2 1936.3 2017.4  2082 2202.7 2227.5 2262.7
## [2,]  639.4  675.5  706.1  739.9  794.2  857.8   859  840.7  835.9  840.4
##       [,46]  [,47]  [,48] [,49]  [,50]  [,51]  [,52] [,53]  [,54]  [,55]
## [1,] 2226.6 2260.3 2294.4  2297 2296.5 2274.3 2273.5  2240 2167.3 2070.9
## [2,]  881.2  888.3  893.7   905  890.2  875.2  884.6   869  828.5  846.1
##       [,56]  [,57]  [,58]  [,59]  [,60]  [,61]  [,62]  [,63]  [,64]  [,65]
## [1,] 1992.7 1940.1 1891.6 1809.8 1686.9 1540.5 1445.9 1329.4 1217.2 1033.4
## [2,]  849.4  808.9  761.9  730.2  713.2  673.2  617.8  588.3  529.1  479.4
##      [,66] [,67] [,68] [,69] [,70] [,71]  [,72] [,73] [,74]  [,75] [,76]
## [1,] 883.6 732.2 566.0   384 260.1 170.4  86.98 14.34  3.26 0.2626     0
## [2,] 411.4 377.6 342.9   301 222.9 170.6 102.51 29.92 10.05 1.1552     0
##      [,77] [,78] [,79] [,80] [,81] [,82] [,83] [,84] [,85] [,86] [,87]
## [1,]     0     0     0     0     0     0     0     0     0     0     0
## [2,]     0     0     0     0     0     0     0     0     0     0     0
##      [,88] [,89] [,90] [,91] [,92] [,93] [,94] [,95] [,96]
## [1,]     0     0     0     0     0     0     0     0     0
## [2,]     0     0     0     0     0     0     0     0     0
```

Let’s visualize the expected value and standard errors we estimated:

```r
plot(estimates_n_ind[1,],xlab="time [15min]", ylab="radiation", type="l", col="red", ylim=c(0,3500))
lines(estimates_n_ind[1,]+estimates_n_ind[2,],type="l")
lines(estimates_n_ind[1,]-estimates_n_ind[2,],type="l")
legend(65,3400, c("exp value", "+/- std error"), col=c("red","black"), lty=c(1,1))
```



1.2 Sample generation

1.2.1 Sample generation - dependent for 1h intervals

We generate a sample of size N of a **multivariate** normal distribution with the parameters estimated above:

```
N<-1000
Nsample<-mvrnorm(n=N, estimates_n_dep$muhat, estimates_n_dep$sigmahat)
```

To validate this sample generation, we estimate its distribution and standard error and compare them with the parameters we used:

```
test<-mlest(Nsample)
diffmu<-estimates_n_dep$muhat-test$muhat
diffsigma<-estimates_n_dep$sigmahat-test$sigmahat
max(diffmu)
```

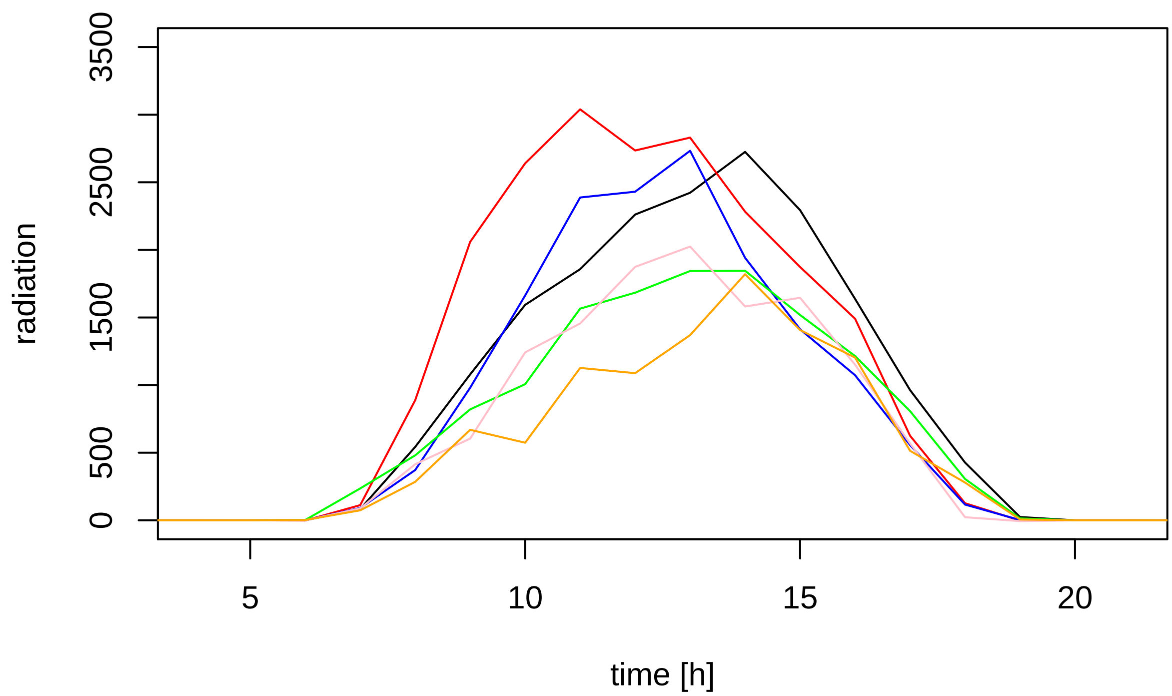
```
## [1] 34.37
```

```
max(diffsigma)
```

```
## [1] 2519
```

These values show us that a sample size of $N = 1000$ is probably not enough.
To visualize, the first three realizations that were generated, look like this:

```
plot(Nsample[1,], type="l", xlim=c(4,21), ylab="radiation", xlab="time [h]", ylim=c(0,3500))
lines(Nsample[2,], type="l", col="red")
lines(Nsample[3,], type="l", col="blue")
lines(Nsample[4,], type="l", col="green")
lines(Nsample[5,], type="l", col="pink")
lines(Nsample[6,], type="l", col="orange")
```



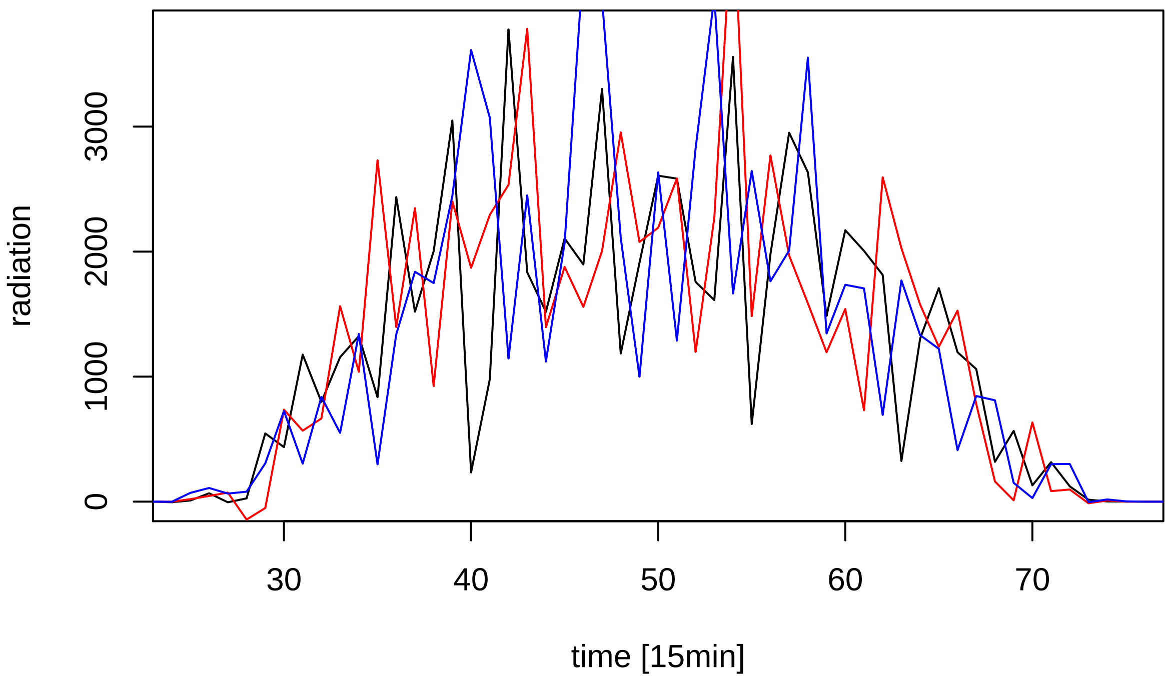
1.2.2 Sample generation - independent for 15min intervals

We generate a sample of N **independently** identically distributed random variables of the normal distribution with the parameters estimated above:

```
N<-10000
N_ind_sample<-matrix(rep(0,24*4*N), ncol=24*4) # 1 column for 1 time interval, 1 row for 1 realization
for(i in 1:N){
N_ind_sample[i,]<-rnorm(24*4,estimates_n_ind[1,], estimates_n_ind[2,])
}
```

To visualize, the first three realizations that were generated, look like this:

```r
plot(N_ind_sample[1,], type="l", xlim=c(25,75), ylab="radiation", xlab="time [15min]")
lines(N_ind_sample[2,], type="l", col="red")
lines(N_ind_sample[3,], type="l", col="blue")
```



2 Weibull distribution

2.1 Parameter estimation

2.1.1 Parameter estimation - independent for 15min intervals

We assume **independently** distributed values for every 15 minutes and estimate a and b called the *shape parameter* and the *scale parameter*.

```r
# estimates_w_ind<-matrix(rep(0,2*96),nrow=2) #under assumption of independence
# for (i in 1:(4*24)){
#   coli<-PV1quh[,i]
#   coli<-coli[coli>0]
#   if(length(coli)>0){
#   estimates_w_ind[1,i]<-fitdist(coli,"weibull", lower=c(0,0))$estimate[1] #estimate of shape paramter
```

```
#   estimates_w_ind[2,i]<-fitdist(coli,"weibull", lower=c(0,0))$estimate[2] #estimate of scale paramete
#   } else {
#     estimates_w_ind[1,i]<-NaN
#     estimates_w_ind[2,i]<-NaN
#   }
#   }
# estimates_w_ind
```

2.2 Sample generation

2.2.1 Sample generation - independent for 15min intervals

```
# N<-10000
# W_ind_sample<-matrix(rep(0,24*4*N), ncol=24*4) # 1 column for 1 time interval, 1 row for 1 realizatio
# for(i in 1:N){
# W_ind_sample[i,]<-rweibull(24*4,estimates_w_ind[1,i], estimates_w_ind[2,i])
# }
```