Statistik-Ü 4.KW (WS 2014/2015) Prof.Dr.L.Paditz
(Dokument als eActivity im ClassPad400 erstellt)

Liebe Studenten,
hier finden Sie noch einmal den vollständigen Lösungsweg für ausgewählte Ü-Aufg. mit dem ClassPad:

STAT F1:

==========

6⇒n

6.00

seq(t,t,1,n,1)⇒listt

{1.00,2.00,3.00,4.00,5.00,6.00}

{4943,5002,5069,5165,5168,5238}/10⇒listy

{494.30,500.20,506.90,516.50,516.80,523.80}

Punktwolke mit Trend

LinearReg listx,listy,1,y1,On

done

DispStat

done

==========

Lineare Regression

y=a+b·x

a = 489.06

b = 5.9114286

r = 0.9846951

r^2 = 0.9696244 = B (Bestimmtheitsmaß)

MSe = 4.7894286 (MeanSquareError = mittlere quadrat. Fehler)

==========

mean(listy)

509.75

y1(x)

5.91·x+489.06

Systemvariable residual (Residuen: $y_i - y(t_i)$):

residual

{-0.67,-0.68,0.11,3.79,-1.82,-0.73}

seq(y1(t),t,1,n,1)⇒listyd

{494.97,500.88,506.79,512.71,518.62,524.53}

listy-listyd⇒Resid

{-0.67,-0.68,0.11,3.79,-1.82,-0.73}

$\frac{\text{sum}(\text{Resid}^2)}{n-2}$⇒MSErr

4.79

Prognosen:

y1(7)

530.44

y1(8)

536.35

Bestimmtheitsmaß B=r^2, s.o.

stop

STAT F3:

=========

seq(t,t,1,16,1)⇒listt

{1.00,2.00,3.00,4.00,5.00,6.00,7.00,8.00,9.00,10.00▶

{842,948,1007,868,863,1005,1064,899,885,1021,1079,9▶

{84.20,94.80,100.70,86.80,86.30,100.50,106.40,89.90▶

mean(listy)

95.90

LinearReg listt,listy,1,y1,On

done

Punkteplot/Trend/Polygon

y1(x)

0.62·x+90.63

Trendbereinigung:

y1(listt)

{91.25,91.87,92.49,93.11,93.73,94.35,94.97,95.59,96▶

listy−y1(listt)⇒listysin

{−7.05,2.93,8.21,−6.31,−7.43,6.15,11.43,−5.69,−7.71▶

Sinus−Regression mit den trendbereinigten Daten:

sinReg listt,listysin,y2,On

done

Trendbereinigung/sin−Regression

linearer Trend:

y1(x)

$$0.62 \cdot x + 90.63$$

Saisonschwankungen, trendbereinigt:

Sinusregression:

y2(x)

$$9.74 \cdot \sin(1.58 \cdot x - 2.63) + 0.04$$

y3(x)=y1(x)+y2(x): period. Schwankungen mit lin. Trend:

Define y3(x)=y1(x)+y2(x)

done

y3(x)

$$0.62 \cdot x + 9.74 \cdot \sin(1.58 \cdot x - 2.63) + 90.67$$

Prognose (5.Jahr):

y3({17,18,19,20})

{93.23,107.47,110.36,97.34}

Saisonschwankungen eines Jahres gemäß sin-Regression:

y2({1,2,3,4})

{-8.42,4.92,8.44,-4.94}

stop

==========

Die oben durchgeführten Untersuchungen gehen von sin-förmigen Schwankungen aus und dienen nur als Näherung.

Man erkennt: Die Maxima der Zeitreihe sind nach rechts verschoben, so dass die Symmetrie der sin-Funktion nicht mehr gegeben ist.

Genauere Ergebnisse erhält man mit der Trendbereinigung, die auf gleitenden Mitteln beruht:

b) gleitende Mittel 4. Ordnung:

(Berechnung über die Listenarithmetik mit gleichlangen Listen, mit je 12 Elementen)

Listen um jeweils 4 Elemente kürzen:

subList(listy, 1, 12)⇒listym2

{84.20, 94.80, 100.70, 86.80, 86.30, 100.50, 106.40, 89.90▶

subList(listy, 2, 13)⇒listym1

{94.80, 100.70, 86.80, 86.30, 100.50, 106.40, 89.90, 88.50▶

subList(listy, 3, 14)⇒listy0

{100.70, 86.80, 86.30, 100.50, 106.40, 89.90, 88.50, 102.1▶

subList(listy, 4, 15)⇒listyp1

{86.80, 86.30, 100.50, 106.40, 89.90, 88.50, 102.10, 107.9▶

subList(listy, 5, 16)⇒listyp2

{86.30, 100.50, 106.40, 89.90, 88.50, 102.10, 107.90, 92.2▶

gleitende Mittel 4. Ordnung:

$\frac{1}{4}*\left(\frac{1}{2}*(\text{listym2}+\text{listyp2})+\text{listym1}+\text{listy0}+\text{listyp1}\right)\Rightarrow\text{listyg}$

{91.89, 92.86, 94.29, 95.39, 96.05, 96.53, 96.91, 97.39, 98▶

fRound(ans, 1)

{91.90, 92.90, 94.30, 95.40, 96.10, 96.50, 96.90, 97.40, 98▶

c) Lineare Regression für die y^*-Daten:

subList(listt, 3, 14)⇒listtg

{3.00, 4.00, 5.00, 6.00, 7.00, 8.00, 9.00, 10.00, 11.00, 12.▶

LinearReg listtg, listyg, 1, y4, On

done

DispStat

done

==========

Lineare Regression

y=a+b·x

a = 91.283057

b = 0.5767483

r = 0.9502363

r^2 = 0.9029491

MSe = 0.5112636

==========

linearer Trend

Die ursprüngliche Regression und die über die gleitenden Mittel unterscheiden sich kaum:

Define y1(x)=0.62·x+90.63

done

Define y4(x)=0.5767482517·x+91.28305653

done

y1(x)=0.62·x+90.63 (ohne gleitende Mittel)

y4(x)=0.577·x+91.283 (über gleitende Mittel)

d) Saisonschätzung:

listy

{84.20,94.80,100.70,86.80,86.30,100.50,106.40,89.90

Trendschätzung mit Prognose 5.Saison:

seq(y4(t),t,1,20,1)⇒listygt

{91.86,92.44,93.01,93.59,94.17,94.74,95.32,95.90,96▶

Trendbereinigung: Saisonschwankungen

listy−sublist(listygt,1,16)⇒listd

{−7.66,2.36,7.69,−6.79,−7.87,5.76,11.08,−6.00,−7.97▶

zur Ansicht gerundet:

fRound(ans,3)

{−7.66,2.36,7.69,−6.79,−7.87,5.76,11.08,−6.00,−7.97▶

seq(listd[j],j,1,16,4)⇒lists1

{−7.66,−7.87,−7.97,−7.58}

seq(listd[j],j,2,16,4)⇒lists2

{2.36,5.76,5.05,3.14}

seq(listd[j],j,3,16,4)⇒lists3

{7.69,11.08,10.27,6.27}

seq(listd[j],j,4,16,4)⇒lists4

{−6.79,−6.00,−6.00,−6.31}

{mean(lists1),mean(lists2),mean(lists3),mean(lists4)}⇒lists

{−7.77,4.08,8.83,−6.28}

gemittelte Saisonanteile:

fRound(lists1,1)

{−7.70,−7.90,−8.00,−7.60}

augment(lists,lists)

{−7.77,4.08,8.83,−6.28,−7.77,4.08,8.83,−6.28}

augment(ans,ans)

{−7.77,4.08,8.83,−6.28,−7.77,4.08,8.83,−6.28,−7.77,▶

augment(ans,lists)⇒lists

{−7.77,4.08,8.83,−6.28,−7.77,4.08,8.83,−6.28,−7.77,▶

zur Ansicht gerundet:

fRound(lists, 1)

{-7.8, 4.1, 8.8, -6.3, -7.8, 4.1, 8.8, -6.3, -7.8, 4.1, 8.8, -▶

Schätzungen mit linearem Trend plus Saisonschwankungen:

seq(t, t, 1, 20, 1)⇒listts

{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20}

seq(y4(t), t, 1, 20)⇒listygt

{91.85980478, 92.43655303, 93.01330129, 93.59004954, 94▶

listygt+lists⇒listygs

{84.08951049, 96.51451049, 101.8395105, 87.31451049, 86▶

zur Ansicht gerundet:

geschätzte Zeitreihe (einschließlich Prognose) durch Überlagerung von linearem Trend (aus Gleitmitteldaten) und gemittelten Saisonschwankungen:

fRound(ans, 1)

{84.1, 96.5, 101.8, 87.3, 86.4, 98.8, 104.1, 89.6, 88.7, 101▶

zum Vergleich sin-Regress. mit lin. Trend:

Define y3(x)=0.62·x+90.63+9.744748062·sin(1.57654559·x-▶

done

y3(listts)

{82.82923348, 96.79449547, 100.9316097, 88.16590629, 85▶

fRound(ans, 1)

{82.8, 96.8, 100.9, 88.2, 85.4, 99.5, 103.3, 90.5, 88, 102.1▶

statist. Grafiken

Hinweis: Die Berechnungen können auch über eine Tabellenkalkulation ausgeführt werden

e) Bestimmtheitsmaß B:

$$B:=1-\frac{\text{sum}\left(\left(\text{listy}-\text{subList}(\text{listygs},1,16)\right)^2\right)}{\text{sum}\left(\left(\text{listy}-\text{mean}(\text{listy})\right)^2\right)}$$

0.9751171458

$$\text{sum}\left(\left(\text{listy}-\text{mean}(\text{listy})\right)^2\right)$$

929.04

$$\text{sum}\left(\left(\text{listy}-\text{subList}(\text{listygs},1,16)\right)^2\right)$$

23.11716685