

Math. II, Statistik, Uebung2, Termin 11.04.2013 (AI), 15.04.2013 (BI)
1-Logistischer Trend, 2-Faktoren, 3-Expon. Glätten, 4-Häufigkeit

ALLGEMEINES

- 1d) 94,2% der Zunahme der Sättigung mit Smartphones ist auf die Zeitentwicklung zurückzuführen. 5,8% der Ursachen sind durch die Regression nicht erklärbar.
- 3a) "eine Nachkommastelle" "2 können auch nicht schaden"
doch, 2 schaden:
1. die Eingabewerte haben null Nachkommastellen, mehr als eine in den Ergebnissen würde eine höhere Genauigkeit vorspiegeln, das grenzt an Betrug!
 2. Ganze Zahlen sind am besten abzulesen und zu erinnern.
Die Kennzahlen im Controlling sollen mit einem Blick erfasst werden können...
Jede überflüssige Nachkommastelle beeinträchtigt die Aussagekraft.
 3. Die Mitarbeiter in der Produktion müssen vor Ort ständig kürzen, um die Kennzahl vergleichen zu können (Soll-Ist-Vergleiche).
- 3e) "durch Normung geometrische Folgen vergleichbar". o.ä.
das hat keinen Sinn!
Eine passende Gewichtung könnte auch ohne eine geometrische Folge erreicht werden.
Aber die einfache fortlaufende Rechnung geht nur mit geometrischer Folge,
nur bei geometrischen Folgen kann stets rekursiv der Folgewert berechnet werden.

AUFGABE 1 10 PUNKTE

- a) Transformieren [2 P.]
- b) Gleichung der Trendfunktion. (Zwischenwerte) [4 P.]
- c) Prognosewerte für die Jahre 2013 und 2014. [1 P.]
- d) Bestimmtheitsmaß und interpretieren [2 P.]
- e) Skizzieren Sie den Funktionsgraphen bis Jahr 2020 [1 P]

$$y = \frac{S}{1 + e^{mx+b}} \Rightarrow y \cdot (1 + e^{mx+b}) = S \Rightarrow \frac{S}{y} = 1 + e^{mx+b} \Rightarrow \frac{S}{y} - 1 = e^{mx+b}$$

a) logarithmieren: $\ln\left(\frac{S}{y} - 1\right) = mx + b.$ $y_{\text{transformiert}}^* = \ln\left(\frac{S}{y} - 1\right) \Rightarrow y^* = mx + b.$

b) Ausstattung der Privathaushalte mit Smartphones y [%]

Jahr	Zeitraum x	y [%]	y _{transf}	y _{tr} * x	x ²	(ŷ _i - ȳ) ²	(y _i - ȳ) ²
2005	1	4,00	3,157	3,157	1	310,049	302,760
2006	2	8,00	2,420	4,841	4	178,088	179,560
2007	3	15,00	1,711	5,132	9	26,182	40,960
2008	4	30,00	0,818	3,273	16	75,745	73,960
2009	5	50,00	-0,041	-0,204	25	743,493	817,960
Summen:	15	107,00	8,066	16,199	55	1333,558	1415,200

Sättigungsniveau S = 98,00 %
 Koeffizient m = -0,800
 Koeffizient b = 4,012
 Trendfunktion:

b)
$$\hat{y} = \frac{98}{1 + e^{-0,8x + 4,012}}$$

c) 2013 entspricht x = 9
 2014 entspricht x = 10

$$\hat{y}(9) = \frac{98}{1 + e^{-0,8 \cdot 9 + 4,012}} = 94,1\%$$

$$\hat{y}(10) = \frac{98}{1 + e^{-0,8 \cdot 10 + 4,012}} = 96,2\%$$

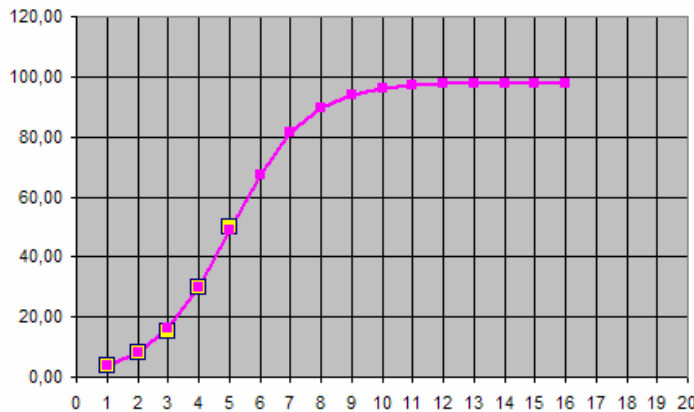
e) Mittelwert ȳ = 21,4 %
 Bestimmtheitsmaß r² = 0,942

$$m = \frac{n \sum x \cdot y_{tr} - \sum x \sum y_{tr}}{n \sum x^2 - (\sum x)^2} \quad b = \frac{1}{n} \sum y_{tr} - \frac{m}{n} \sum x$$

$$m = \frac{5 \cdot 16,197 - 15 \cdot 8,065}{5 \cdot 55 - 15^2} = \frac{-39,99}{50} = -0,8$$

$$b = \frac{1}{5} \cdot 8,065 - \frac{-0,8}{5} \cdot 15 = 1,613 + 2,4 = 4,01$$

94% des Anstiegs von y ist auf die Zeitentwicklung zurückzuführen, 6% unerklärte Ursachen.



e)

AUFGABE 2 10 PUNKTE

Eine betriebliche Kennzahl wurde für Quartale (I, II, III, IV) ermittelt.

Quartal i	Kennzahl y(i)	Wachstums- Faktoren	Zuwachs- raten [%]	Index für die Basis 1.Qu. 2011 = 100	Gleitende Mittelwerte Zyklus- länge = 4
I. Qu. 2011	16	--	--	100,0	--
II. Qu. 2011	21	1,313	31,3	131,3	--
III. Qu. 2011	24	1,143	14,3	150,0	20,75
IV. Qu. 2011	19	0,792	-20,8	118,8	22
I. Qu. 2012	22	1,158	15,8	137,5	--
II. Qu. 2012	25	1,136	13,6	156,3	--

- a) Bestimmen Sie die fehlenden Werte und tragen Sie diese in die Tabelle ein.
Bestimmen Sie die Wachstumsfaktoren auf drei Nachkommastellen genau. [4 × 1,5 P.]
- b) Bestimmen Sie den mittleren Wachstumsfaktor und die mittlere Zuwachsrate
(in Prozent) für den untersuchten Zeitraum. [2 P.]
- c) Beschreiben Sie mindestens zwei Nachteile der Glättung mittels gleitenden
Mittelwerten gegenüber der Glättung mittels Regressionsrechnung. [2 P.]

a) *Gleitende Mittelwerte:* $(8+21+24+19+11) / 4 = 20,75$. $(10,5+24+19+22+12,5) / 4 = 22$

b) *mittl. Wachstumsfaktor* $q = \sqrt[5]{\frac{25}{16}} = \sqrt[5]{1,5625} = 1,093$

Mittl. Zuwachsrate $= (1 - 1,093 \cdot 100) = 9,3 \%$

- c) (1) *Am Anfang und am Ende der Zeitreihe kommt es zu Informationsverlusten.*
 (2) *Es gibt keine Funktionsgleichung wie $\hat{y} = m x + b$ mit der man saisonbereinigte Prognosewerte $\hat{y}(x_{n+z})$ berechnen könnte.*
 (3) *Die gleitenden Mittelwerte $\hat{y}(i)$ lassen sich nur dann eindeutig dem Zeitraum i zuordnen, wenn die Zykluslänge k ungerade ist.*

AUFGABE 3 10 PUNKTE

In einem Unternehmen wird Controlling mit Balanced Scorecards organisiert. Dabei wurde eine bestimmte Kennzahl y [%] über 10 Wochen aufgezeichnet. Die Werte y_i sind nicht trendbehaftet.

- a) Bestimmen Sie die Schätzwerte \hat{y}_i mit Hilfe der exponentiellen Glättung. (Eine Nachkommastelle) Verwenden Sie eine Glättungskonstante von 0,2. [3,5 P.]
- b) Bestimmen Sie den Prognosewert für die 24. Kalenderwoche. [1 P.]
- c) Bestimmen Sie den THEIL'schen Ungleichheitskoeffizienten und interpretieren Sie das Ergebnis. Ergänzen Sie dazu die gegebene Arbeitstabelle. [3,5 P.]
- d) Beschreiben Sie den Hauptvorteil des exponentiellen Glättens gegenüber der Glättung mittels gleitender Mittelwerte. [1 P.]
- e) Zum Gewichten der Beobachtungswerte benutzt man die Werte einer geometrischen Folge. Welchen Hauptvorteil hat es, hier eine geometrische Folge zu verwenden? [1 P.]

a) b) c)

Kalender- woche	Index	Kenn- zahl	geglättet		
w	i	y_i	\hat{y}_i	$(y_i - \hat{y}_i)^2$	$(y_i - y_{i-1})^2$
14	1	27	27,0		
15	2	38	29,2	77,440	121
16	3	25	28,4	11,290	169
17	4	31	28,9	4,461	36
18	5	32	29,5	6,198	1
19	6	35	30,6	19,287	9
20	7	43	33,1	98,274	64
21	8	27	31,9	23,710	256
22	9	38	33,1	24,055	121
23	10	31	32,7	2,810	49
Summen:				267,525	826
b) Prognose:		26	31,3		
Glätt.Faktor α =		0,20			
c) U =		0,569			

Die Schätzung aus dem Prognoseverfahren ist signifikant besser als die naive Schätzung, weil $U < 1$

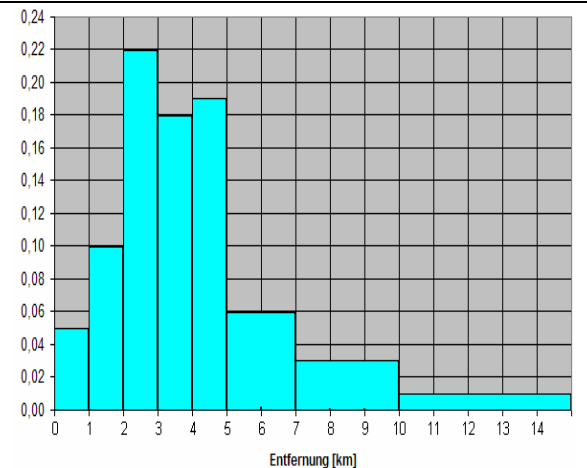
$$U = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{i-1})^2}}$$

- b) $\hat{y}(24) = \hat{y}_{11} = 0,2 \cdot 26 + 0,8 \cdot 32,7 = 31,3\%$
- d) Die aktuelleren Beobachtungswerte werden höher gewichtet als die älteren. Die Gewichtung der Beobachtungsdaten nimmt mit zunehmendem Alter der Beobachtungsdaten exponentiell ab.
- e) wegen der Gewichtung mit einer geometrischen Folge kann man die jeweiligen Glättungswerte \hat{y} nacheinander rekursiv berechnen. rekursiv: $\hat{y}_{i+1} = \alpha y_{i+1} + (1 - \alpha) \cdot \hat{y}_i$

AUFGABE 4 10 PUNKTE

Die nebenstehende Tabelle zeigt die Entfernungen vom Wohnort der Mitarbeiter zu ihrem Arbeitsplatz in km und die entsprechende Häufigkeiten.

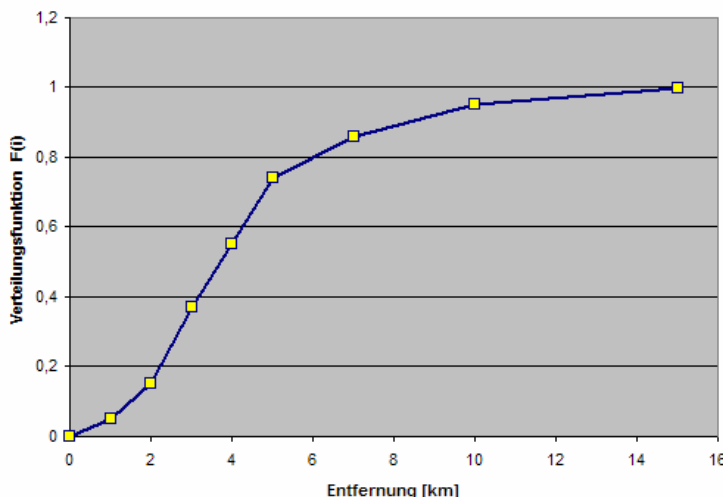
- a) Zeichnen Sie das Histogramm
x-Achse: 1 cm $\hat{=}$ 1 km [2 P.]
- b) Bestimmen Sie den Median (Zentralwert). [1 P.]
- c) Bestimmen Sie den Mittelwert. [1 P.]
- d) Bestimmen Sie Varianz und Standardabweichung.
Es möge sich hier um eine kleine Stichprobe handeln mit n-1 Freiheitsgraden. [2 P.]
- e) Bestimmen Sie den Variationskoeffizienten. [1 P.]
- f) Skizzieren den Funktionsgraphen der empirischen Verteilungsfunktion. [2 P.]
- g) Es soll die relative Häufigkeit dafür, dass ein Mitarbeiter höchstens 6 km vom Arbeitsplatz entfernt wohnt, bestimmt werden. Beschreiben Sie, wie man die aufgrund der Verteilungsfunktion bestimmen könnte. [1 P.]



Mittelwert	4,32	km	
Median	3 < x ≤ 4	(Zentralwert) auch 3,5 km	
Varianz	7,676		
s =	2,771	km	
v =	0,641		

Runden Sie sinnvoll und geben Sie die passenden Einheiten an.

x_i [km]	n_i	h_i	F_i	x^*_i	$x^*_i h_i$	$(x^*_i)^2 n_i$
$0 < x \leq 1$	10	0,05	0,05	0,5	0,025	2,500
$1 < x \leq 2$	20	0,1	0,15	1,5	0,15	45,000
$2 < x \leq 3$	44	0,22	0,37	2,5	0,55	275,000
$3 < x \leq 4$	36	0,18	0,55	3,5	0,63	441,000
$4 < x \leq 5$	38	0,19	0,74	4,5	0,855	769,500
$5 < x \leq 7$	24	0,12	0,86	6	0,72	864,000
$7 < x \leq 10$	18	0,09	0,95	8,5	0,765	1300,500
$10 < x \leq 15$	10	0,05	1	12,5	0,625	1562,500
	200				4,32	5260,000



g) Da die Verteilungsfunktion stetig ist, kann man die relative Häufigkeit, dafür dass ein Mitarbeiter höchstens 6 km vom Arbeitsplatz entfernt wohnt, schätzen: der Mittelwert von F_7 und F_5 ist zu bestimmen (oder Mittelpunkt der Strecke)

$(0,86 + 0,74) : 2 = 0,8$ auch: die Wahrscheinlichkeit, dass eine Mitarbeiter höchstens 6 km wohnt, ist 80 % (Rechnung ist hier nicht gefordert)

- natürlich könnte man aus das entsprechende halbe Rechteck aus dem Histogramm dazuzaddieren, aber es ist nach der "Verteilungsfunktion" gefragt