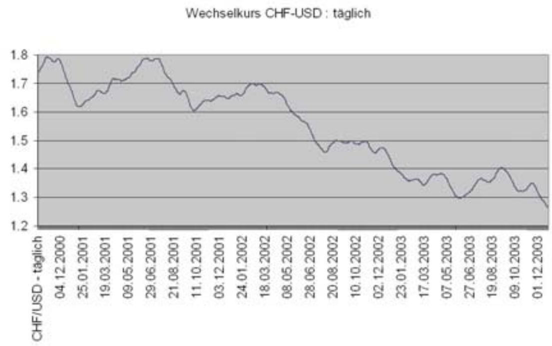
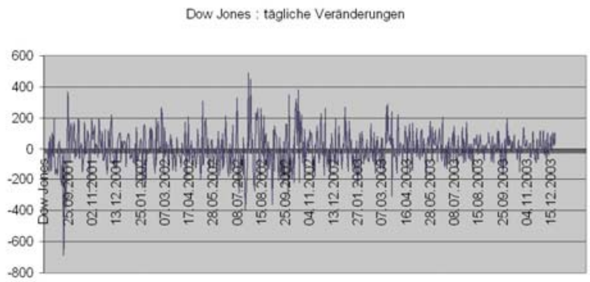
Teil 2: Zeitreihenanalyse

- In vielen Anwendungen in Finance interessiert die **zukünftige Entwicklung** gewisser „Phänomene". Als **Grundlage** für die **Analyse und für die Vorhersage** liegen deshalb oft Messungen **(Beobachtungen) dieser Phänomene im Zeitablauf** vor.
- **Eine Zeitreihe ist eine endliche Folge <u>zeitlich</u> angeordneter Beobachtungen, die in regelmässigen** (z.B. tägliche Erfassung der Meteo-Daten) **oder unregelmässigen** (Arztkonsultationen) **zeitlichen Abständen erhoben werden.**
  - o Die monatliche Erhebung von makroökonomischen Grössen ist oft nur eine **fast regelmässige Erhebung,** weil die Monatslängen unterschiedlich sind.
  - o Die **Einhaltung der zeitlichen Abfolge** legt den Beobachtungen eine **gewisse Struktur** auf.
    - Z.B. ergeben die SMI-Daten nur dann Sinn und daher eine Struktur, wenn sie zeitlich geordnet sind. Betrachten wir die SMI-Daten ohne zeitliche Ordnung, macht es für uns keinerlei Sinn, also kann es auch keine Struktur geben.
- Mit **Prognoseverfahren** wird versucht, diese Struktur zu erfassen, um daraus eine Vorschrift für die Bestimmung zukünftiger Beobachtungen abzuleiten (sog. **Punktprognose).**
  - o Nebst dem eigentlichen **Prognosewert** interessiert auch oft dessen **„Güte".** Die Güte kann in Verbindung mit einem bestimmten **„Risikobegriff"** gebracht werden.

1. Beispiele





- Im ersten Beispiel sind die **Wechselkurse in CHF für USD** für das Zeitfenster vom 12.12.2000 bis 04.12.2003 abgebildet.
  - o Der allgemeine **Verlauf** ist durch einen **starken Abwärtstrend (eindeutiger Trend)** gekennzeichnet. Der Abwärtstrend wird durch **kurz- und mittelfristige Schwankungen (stochastische „Zyklen") überlagert.**
- Im zweiten Beispiel sind die **UBS-Aktienquotierungen** vom 28.09.2001 bis 22.12.2003 abgebildet.
  - o In dieser Abbildung sind im Gegensatz zur vorherigen **keine eindeutigen Trends** noch **Zyklen** erkennbar.
  - o Vielmehr zeigt der Verlauf der Zeitreihe Ähnlichkeiten mit der Realisation eines **Random-Walk-Prozesses: Abschwung- und Aufschwungphasen lösen sich ab** ohne erkennbare Muster. (Weiter unten wird gezeigt, dass diese Vermutung hier nicht zutrifft.)
- Im dritten Beispiel werden die **täglichen Veränderungen** des **Dow Jones Industrial Index** vom 10.08.2001 bis 19.12.2003 abgebildet.
  - o Von blossem Auge sind **keine klaren Strukturen** (Trends, Zyklen, Auf- und Abschwungphasen) erkennbar.
  - o Die Beobachtungen scheinen **erratisch um den Wert 0 zu fluktuieren.**
  - o Überdies scheint sich die **Varianz der Beobachtungen** im Zeitablauf zu **verändern.** Man beschreibt dieses Phänomen mit dem Begriff **(bedingte) Heteroskedastizität.**
    - Dieses „stilisierte Faktum" spielt bei Financereihen eine wichtige Rolle und **widerspricht der klassischen Diffusionstheorie,** so wie sie beispielsweise in der Black-Scholes-Formel zum Ausdruck kommt.
  - o Statt im **Niveau** der Reihe ist hier also eine zeitliche Struktur in der **Varianz** beobachtbar. Auf dieses für Financereihen typische Merkmal (Risiko (Varianz) und Niveau der Returns sind verknüpft) wird weiter unten eingegangen.
- **Welche Punkte sind nun für die Prognose massgebend?**
  - o Bei der **Wechselkursreihe (Beispiel 1)** scheinen aufgrund des starken Trends vor allem die **Anfangs- und Endbeobachtungen** am Rande der Zeitreihe aufgrund des **„Hebelgesetzes"**

für eine **Prognose** von Bedeutung zu sein, denn diese Beobachtungen bestimmen weitgehend die Steigung des Trends (die mittigen Beobachtungen tragen eher zur Verschiebung und weniger zur Steigung des Trends bei).

- Bei der **UBS-Aktienzeitreihe (Beispiel 2)** ist optisch kein eindeutiger Trend auszumachen. Mittel- bis längerfristige Auf- und Abwärtsbewegungen prägen jedoch das Erscheinungsbild massgebend. Deshalb kann hier die **letzte Beobachtung** als **Prognose** für die zukünftige Entwicklung betrachtet werden.
- Bei den **täglichen Veränderungen des Dow Jones Index (Beispiel 3)** sind keine eindeutigen anhaltenden Niveauverschiebungen erkennbar. Das **arithmetische Mittel** – in dem **jede Beobachtung gleich gewichtet** zum Ausdruck kommt – scheint deshalb eine plausible **Prognosefunktion** darzustellen.

2. Prognoseverfahren: Exponentielle Glättung

2.0 Was ist eine Prognosefunktion?

- Eine **Prognosefunktion** ist eine quantitative **Aussage über zukünftige potentielle Ausprägungen,** die **aufgrund einer gewissen Informationsmenge** (Beobachtungen über gegenwärtige und vergangene Realisation) aufgestellt wird.
  - Wir verfügen also über die Daten bzw. Beobachtungen $x_1, \ldots x_T$.
  - Die Prognosefunktion beruht auf diesen vergangen Realisationen und ist also eine Funktion der Beobachtungen: $\hat{x}_{T+h} = f_h(x_T, x_{t-1}, \ldots, x_1)$ **(0)**

▪	h = 1	Prognose für Morgen
	h = 2	Prognose für Übermorgen
▪	x_T	Letzte Beobachtung
▪	T	Anzahl Beobachtungen
▪	$\hat{x}_{T+h}$	Schätzung (Prognosewert) für Tag in der Zukunft h

Das Ziel ist also die Konstruktion von $f_h(x_T, x_{t-1}, \ldots, x_1)$. Dabei wird klar, dass der Prognosewert x_{T+h} (z.b. h=1, also x_{T+1} der Prognosewert für Morgen) sich aus allen vergangenen Beobachtungen von x_T (letzte Beobachtung) bis zur ersten verfügbaren Beobachtung x_1 zusammen setzt.

  - Um eine solche Prognosefunktion zu erstellen gibt es verschiedene **Prognoseverfahren**:
    - **Exponentielle Glättung**
      - Einfach und robust
    - **Stationäre AR(1) Prozesse**
      - Modelle, aufwendiger, Intervallprognose
    - **GARCH-Prozesse**
      - Prognose der Volatilität (ergänzen AR(1)-Prozesse)

2.1 Aufbau der exponentiellen Glättung (Reihe ohne Trend)

- **Lineare Prognosefunktion (Ausgangsmodell):** $\hat{x}_{T+h} = \sum_{k=1}^{T} a_{kh} x_{T+1-k}$ **(1)**

▪	T	Anzahl Beobachtungen
▪	x_T	Letzte Beobachtung
▪	h = 1	Prognose für Morgen
	h = 2	Prognose für Übermorgen
▪	$\hat{x}_{T+h}$	Schätzung (Prognosewert) für Tag in der Zukunft h
▪	a_{kh}	Koeffizient/Gewichtung: Jede vergangene Realisation x_{T+1-k} wird durch einen Parameter a_{kh} gewichtet.

  - Beispiel:
    - k=1: Also wird als erstes x_{T+1-1} mit einem Koeffizient gewichtet, also zuerst die letzte Beobachtung x_T.
    - k=2: Als zweites wird x_{T+1-2} mit einem Koeffizient gewichtet, also die zweitletzte Beobachtung x_{T-1}.
    - In dieser Reihenfolge werden alle vergangenen Beobachtungen gewichtet.
    - Am Schluss werden die gewichteten Beobachtungen zusammenaddiert und ergeben die Schätzung/den Prognosewert $\hat{x}_{T+h}$.
  - Die **Probleme** bestehen darin, dass die **Koeffizienten** a_{kh} **einerseits unbekannt** sind und andererseits für jede vergangene Beobachtung ein Koeffizient bestimmt werden müsste, es also auch **zu viele solche Koeffizienten** gibt.

- **Problem: Unbekannte Koeffizienten** a_{kh}
    - Die Koeffizienten a_{kh} sind **unbekannt** und müssen **aus den Daten abgeleitet** werden.
    - Die Beispiele haben verdeutlicht, dass die **Gewichte** a_{kh} in Abhängigkeit der untersuchten Zeitreihe sehr unterschiedlich sein können:
        - **Wechselkursreihe:** Die **Anfangs- und Endbeobachtungen** müssten am stärksten gewichtet werden. Die dazwischen liegenden Beobachtungen gar nicht, oder nur sehr schwach.
        - **UBS-Aktienzeitreihe:** Die **letzte Beobachtung** müsste am stärksten gewichtet werden.
        - **Täglichen Veränderungen des Dow Jones Index: Alle Beobachtungen sind gleich wichtig** und müssen gleich gewichtet werden.
        - Zwischen diesen **Extremsituationen** eröffnet sich ein Spektrum möglicher **Gewichtungsschemata.**
- **Problem: Zu viele Koeffizienten** a_{kh}
    - Es muss vermieden werden, dass unnötigerweise **zu viele Parameter** in (1) geschätzt werden, weil sonst die Güte der Prognose schlechter wird.

- **Lösung dieser Probleme**
    - Aus diesen Problemen ergeben sich gegensätzliche **Ansprüche an die Prognosefunktion:**
        - Die Funktion (1) sollte möglichst flexibel sein.
        - Es sollten möglichst wenig Koeffizienten a_{kh} geschätzt werden.
    - Die sogenannte **exponentielle Glättung** im nächsten Abschnitt bietet eine flexible und robuste Lösung, um dieses Spektrum an möglichen **Gewichtungsvorschriften** auf eine gewisse Art umfassend **abzudecken. Die unbekannten Koeffizienten werden durch einen Parameter** anhand eines **einzigen Parameters** α**.**

2.2 Einfache exponentielle Glättung (Voraussetzung: Reihe ohne Trend)

- Aus der Formel (1) wird folgende Vorschrift erstellt: $\boxed{\hat{x}_{T+1} = \alpha x_T + \alpha(1-\alpha)x_{T-1} + \alpha(1-\alpha)^2 x_{T-2} + \alpha(1-\alpha)^3 x_{T-3} + \ldots}$ **(2)**

▪	$\hat{x}_{T+h}$	(h=1) Schätzung/Prognosewert für Morgen
▪	x_T	Letzte Beobachtung
▪	α	Parameter/Gewichtung α = erste Gewichtung für x_T (letzte Beob.) $\alpha(1-\alpha)^1$ = zweite Gewichtung für x_{T-1} (vorletzte Beob.)
▪	$\hat{x}_{T+h} = \hat{x}_t$	$\hat{x}_t$ ist also wie $\hat{x}_{T+h}$ Schätzung/Prognosewert für den Tag T + h. Ist h=1, so ist $\hat{x}_t$ der Prognose wert für Morgen.
▪	$\hat{x}_t \neq x_t$	x_t ist der Wert zum Zeitpunkt t, $\hat{x}_t$ aber ist die Schätzung für den Zeitpunkt t + h.

    - Die Parameter/Gewichtung hat folgende Vorschrift zu erfüllen: $0 < \alpha < 1$, denn die Gewichtungen der vergangenen Beobachtungen müssen sich logischerweise auf 1 aufsummieren weil nur 1 Prognosewert in der Zukunft gesucht wird.
    - Der Begriff „exponentielles Glätten" hängt mit der exponentiellen Abnahme der **Gewichte** $\alpha(1-\alpha)^k$ zusammen.
    - Die Koeffizienten a_{kh} wurden also durch Parameter α ersetzt: $a_{k1} = \alpha(1-\alpha)^{k-1},\ 0<\alpha<1$ (Annahme: h=1)
- Der **Parameter** α erfüllt definitionsgemäss folgende Bedingung $0 < \alpha < 1$, weshalb die resultierenden **Gewichte** a_{kh} **mit zunehmendem *k* monoton gegen Null konvergieren.**
    - Wenn α **sehr gross ist, also gegen 1,** dann wird nur die letzte Beobachtung x_T einbezogen, die anderen älteren Beobachtungen werden mit Null oder fast Null gewichtet. In diesem Fall wird die **letzte Beobachtung** am stärksten gewichtet.
    - Wenn α **sehr klein ist, also gegen 0.001,** dann erhalten alle Beobachtungen das gleiche Gewicht.
    - Das **grösste Gewicht** wird immer der **aktuellen Beobachtung** zugeordnet. **Je weiter eine Beobachtung in der Vergangenheit liegt, desto kleiner wird ihr Gewicht bei der Bestimmung der Prognose für zukünftige Beobachtungen.**

2.3 Prognosefunktion Einfache exponentielle Glättung

- Die Formel (2) kann noch weiter vereinfacht werden:

$$\boxed{\begin{aligned}\hat{x}_t &= \alpha x_t + (1-\alpha)\hat{x}_{t-1} \\ &= \hat{x}_{t-1} + \alpha(x_t - \hat{x}_{t-1}),\quad 0<\alpha<1\end{aligned}} \quad \textbf{(3)}$$

o	$\hat{x}_t$	Schätzung/Prognosewert für Tag t + h (entspricht also $\hat{x}_{T+h}$) (Wenn h=1 dann Prognose für Morgen)
o	x_t	Wert zum Zeitpunkt t (effektiver Wert heute)
o	$\hat{x}_{t-1}$	Schätzung für t - 1 → Alte Schätzung für x_t (Gestrige Schätzung/Prognose für heute)

  - **Die beste Prognose für Morgen $\hat{x}_t$, ist der Wert von heute x_t, weil wir Reihen ohne Trend betrachten.**
  - **Die Wahl von α**
    - **Mit abnehmenden α** nimmt die Bedeutung der Beobachtung x_t ab bzw. die Bedeutung des „alten“ Schätzers $\hat{x}_{t-1}$ nimmt vergleichsweise zu.
      - Oft gilt die Regel: **je unruhiger der Verlauf einer Zeitreihe, desto kleiner α .**
    - Für $\alpha \approx 1$ reduziert sich (3) auf $\hat{x}_t = x_t$ (UBS-Aktienkurs-Beispiel: Die letzte Beobachtung ist am wichtigsten) und für $\alpha = 1/T$ ist $1-\alpha \approx 1$ in (2), so dass $\hat{x}_T \approx 1/T\sum_{t=1}^{T} x_t$ (tägliche Veränderungen Dow-Jones Beispiel: Alle Beobachtungen sind gleich wichtig).
    - Für Werte von α dazwischen wird ein kontinuierliches Spektrum an Gewichtungsschemata abgebildet.
  - Die **zweite Gleichung in** (3) ist in der sogenannten **„Korrekturform“** geschrieben: Wenn die Prognose $\hat{x}_{t-1}$ für x_t zu klein war dann führt $\alpha(x_t - \hat{x}_{t-1})$ zu einer positiven Korrektur der „Prognose“ $\hat{x}_t$ für x_{t+1} (und umgekehrt, wenn $\hat{x}_{t-1}$ > x_t).
  - **Bestimmung des „Startwertes“**
    - $\hat{x}_t = \alpha x_t + (1 - \alpha)\hat{x}_{t-1}$
    - Annahme t = 2
    - $\hat{x}_2 = \alpha x_2 + (1 - \alpha)\hat{x}_1$
      - $\hat{x}_2$: Gesucht ist also ein Prognosewert (Schätzung) für den nächsten Tag 3
      - x_2 ist der effektive Wert am Tag 2
      - $\hat{x}_1$ ist der Prognosewert (Schätzung) für den Tag 2 („alte“ Schätzung)
    - **Für α und $\hat{x}_1$ muss also ein „Startwert“ definiert werden.**
      - Dies wird mit folgender Formel gemacht: $\boxed{\sum_{t=1}^{T-1}(\hat{x}_t(\alpha) - x_{t+1})^2}$ **(4)**
        (für 1-Schritt-Prognose)
        - o $\hat{x}_t$ = Schätzung/Prognosewert für Tag t + 1
          $\hat{x}_t(\alpha)$ bedeutet einfach dass $\hat{x}_t$ von α abhängig ist.
          x_{t+1} = Effektiver Wert für Tag t + 1
        - o Die **Differenz** ($\hat{x}_t(\alpha) - x_{t+1}$) entspricht gerade dem **Prognosefehler im Zeitpunkt t + 1.**
        - o Betrachtet man die zweite Gleichung in (3) $\hat{x}_t = \hat{x}_{t-1} + \alpha(x_t - \hat{x}_{t-1})$, so erkennt man, dass auch dort ($x_t - \hat{x}_{t-1}$) einem Prognosefehler entspricht, nämlich dem **Prognosefehler im Zeitpunkt t.**
        - o **Anschliessend soll der resultierende Wert aus (4), der sogenannte durchschnittliche quadratische Prognosefehler möglichst klein gemacht werden.** Dadurch wird ein α ermittelt, bei dem die Prognosefehler in der Vergangenheit möglichst klein waren, womit α auch für die **Zukunft ein guter Parameter** ist.
      - Für die 4-Schritt-Prognose (h=4) muss diese Formel abgeändert werden:

$$\boxed{\sum_{t=1}^{T-4}(\hat{x}_t(\alpha) - x_{t+4})^2}$$

2.3.1 Aufgabenbeispiel: 1_L_uebungen exp glättung.xls (Spalte „Einführungsbeispiel (h=1)")

Schritt 1: Einrichten

	A	B	C
1		α	0.500000000
2		$\hat{x}_1$	2.000000000
3			
4			
5		Wert	Prognose
6	t	x	$\hat{x}_t$
7	2	3.12	2.000000000
8	3	1.91	2.560000000
9	4	1.89	2.235000000
10	5	1.77	2.062500000
11	6	1.73	1.916250000
12	7	1.56	1.823125000
13	8	1.63	1.691562500

- In dieser Aufgabe geht es um eine **1-Schritt-Prognose** $\hat{x}_t = \alpha x_t + (1 - \alpha)\hat{x}_{t-1}$ (3).
  - Spalte A: **t** (Tage 2 – 149)
  - Spalte B: **x** (Beobachtungen Tage 2 – 149)
  - Als **erstes** muss ein beliebiger Startwert α sowie ein beliebiger Startwert $\hat{x}_1$ (Prognosewert für Tag t=1) gewählt werden.
    - α = 0.5 (zwischen 0 und 1)
    - $\hat{x}_1 = 2$
  - Spalte C: In der Spalte C werden die $\hat{x}_t$ (Schätzung/Prognosewert für Tag t + 1) aufgeführt.
    - $\hat{x}_1$ (Schätzung für Tag t=2) haben wir bereits als 2 definiert.
    - $\hat{x}_2$ (Schätzung für Tag t=3) können wir nun mit der Formel berechnen:

      $\hat{x}_t = \alpha x_t + (1 - \alpha)\hat{x}_{t-1}$

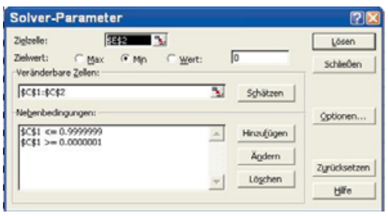
      $\hat{x}_2 = 0.5 \cdot 3.12\ (=x_2) + (1 - 0.5) \cdot 2$

      $\hat{x}_2 = 2.56$
    - Diese Formel kann nun weitergezogen werden bis am Ende, wobei logischerweise bereits ein Prognosewert für den Tag 150 resultiert.

Schritt 2: Optimale Bestimmung von α und $\hat{x}_1$ (bisher haben wir nur eine beliebige Annahme getroffen)

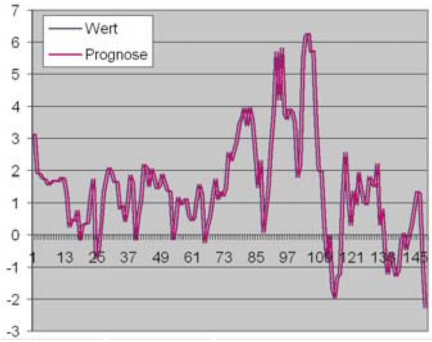
	A	B	C	D	E
1		α	0.9999999		
2		$\hat{x}_1$	3.1200000	Summe der quadr. Prognosefelder:	117.45650006
3					
4					
5		Wert	Prognose	Prognosefehler	quadr. Prognosefehler
6	t	x	$\hat{x}_t$	$x_{t+1} - \hat{x}_t$	$(x_{t+1} - \hat{x}_t)^2$
7	2	3.12	3.119999992	0.000000008	0.00000000
8	3	1.91	3.120000000	-1.210000000	1.46410000
9	4	1.89	1.910000121	-0.020000121	0.00040000
10	5	1.77	1.890000002	-0.120000002	0.01440000
11	6	1.73	1.770000012	-0.040000012	0.00160000
12	7	1.56	1.730000004	-0.170000004	0.02890000
13	8	1.63	1.560000017	0.069999983	0.00490000

- Als nächstes werden wir versuchen α und $\hat{x}_1$ mit Hilfe der Formel (4) optimal zu bestimmen:
  - $$\boxed{\sum_{t=1}^{T-1} (\hat{x}_t(\alpha) - x_{t+1})^2}$$
  - Spalte D: **Prognosefehler im Zeitpunkt t →** $x_{t+1} - \hat{x}_t$
    - Z.B. für die erste Zeile: $x_2 - \hat{x}_1$ (Wert bei t=2 abzüglich Schätzung für t=2)
    - Diese Formel kann nun weitergezogen werden bis am Ende.
  - Spalte E: **Quadrierte Prognosefehler im Zeitpunkt t →** $(x_{t+1} - \hat{x}_t)^2$
    - Diese Formel kann nun weitergezogen werden bis am Ende.
  - Die Zelle E2 zeigt nun die **Summe der quadrierten Prognosefelder.**
    - Diese Zelle gibt das Resultat der Formel (4) wieder.
    - **Jetzt muss α und $\hat{x}_1$ so gewählt werden, dass (4) möglichst klein wird.** Dies wäre mit der Differentialrechnung sehr aufwendig, weshalb uns Excel zu Hilfe kommt.
    - Dies kann im Excel mit dem „Solver" (Extras/Solver) gelöst werden:
      - Wähle die Zelle E2 die möglichst klein gemacht werden soll an, dann Extras/Solver.
      - Zielwert: Min
      - Veränderbare Zellen: C1 und C2 (α und $\hat{x}_1$, die wir bisher beliebig angenommen haben)
      - Nebenbedingung: $0 < \alpha < 1$



    - Der Solver wählt nun α = 0.9999 und $\hat{x}_1$ = 3.12. α wird also fast 1 und $\hat{x}_1$ wird so gross wie die erste Beobachtung.

- Wir haben nun also ein optimales α und $\hat{x}_1$ gewählt. Mit diesen Werten ist es uns nun möglich eine gute 1-Schritt-Prognose zu machen.
- Weiter können wir die Beobachtungen mit unseren 1-Tages-Schätzern vergleichen indem wir diese in einer Grafik darstellen.



- In der Datei 1_L_uebungen exp glättung.xls (Spalte „Einführungsbeispiel (h=4)") findet sich die gleiche Aufgabe nochmals, aber diesmal ist eine 4-Schritt-Prognose vorgesehen.

  In dieser Aufgabe geht es um eine **4-Schritt-Prognose** $\hat{x}_t$ = αx_t + (1 - α) $\hat{x}_{t-1}$ (3).
    - $\hat{x}_{t+3}$ = $\hat{x}_{2+3}$ = $\hat{x}_5$ = „neue" Schätzung/Prognose für t=6 (4-Schritt-Prognose)
    - x_t = x_2 = effektiver Wert für t = 2 (erster Wert)
    - $\hat{x}_{t-1}$ = $\hat{x}_{5+1}$ = $\hat{x}_4$ = „alter" Schätzer für t = 5
    - $\hat{x}_5$ = α · x_2 + (1 - α) · $\hat{x}_4$
  - Die Prognosewerte t = 3, 4 und 5 müssen auch mit einem Startwert gefüllt werden, da sie für die Prognose von t=7, 8 und 9 verwendet werden! Hierzu kann einfach der Startwert eingesetzt werden.

RANG =$C$1*B7+(1-$C$1)*C10

	A	B	C	D	E
1		α	0.1146214		
2		$\hat{x}_1$	1.5653222	Summe der quadr. Prognosefelder:	372.31046658
3					
4					
5		Wert	Prognose	Prognosefehler	quadr. Prognosefehler
6	t	x	$\hat{x}_t$	$x_{t+4} - \hat{x}_t$	$(x_{t+1} - \hat{x}_t)^2$
7	2	3.12	1.565322242	1.554677758	2.41702293
8	3	1.91	1.565322242	0.344677758	0.11880276
9	4	1.89	1.565322242	0.324677758	0.10541565
10	5	1.77	1.565322242	0.204677758	0.04189298
11	6	1.73	=$C$1*B7+(1-	-0.013521659	0.00018284
12	7	1.56	1.762603648	-0.202603648	0.04104824
13	8	1.63	1.777206002	-0.147206002	0.02166961
14	9	1.7	1.776380040	-0.076380040	0.00583391

2.4 Prognoseverfahren: Exponentielle Glättung nach Holt-Winters (Reihen mit Trend)

- Die Prognosevorschrift (4) ist für trendbehaftete Reihen ungeeignet. Holt-Winters hingegen ist für trendbehaftete Reihen geeignet.
- **Trend**
  - Der **Begriff „Trend"** beschränkt sich im Folgenden nicht ausschliesslich auf klassische **deterministische** (lineare oder polynominale) Trends, sondern schliesst auch **variable „stochastische" Trends** ein (vgl. Abbildung: Tägliche Wechselkurse USD/CHF), die in vielen Anwendungen in der Finance die Reihenerscheinungsbilder prägen.
- **Exponentielle Glättungsvorschrift nach Holt-Winters**
  - **Niveaugleichung:** $\boxed{\hat{x}_t := Ax_t + (1-A)(\hat{x}_{t-1} + Tr_{t-1})}$ **(5.1)**

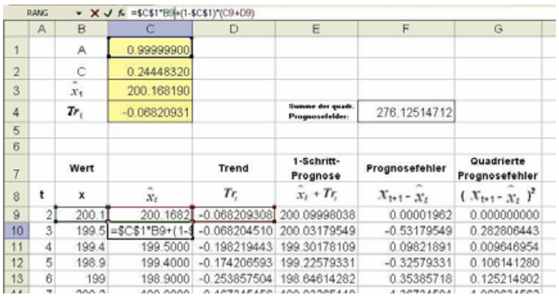
    - $\hat{x}_t$ Schätzung/Prognosewert für Tag t + h (entspricht also $\hat{x}_{T+h}$) (Wenn h=1 dann Prognose für Morgen) **= Neues Niveau**
    - x_t Wert zum Zeitpunkt t (effektiver Wert heute)
    - $\hat{x}_{t-1}$ Schätzung für t - 1 → Alte Schätzung für x_t (Gestrige Schätzung/Prognose für heute)
    - Tr_{t-1} Trend für den Zeitpunkt t-1 → Alter Trend für x_t Positiv bei steigendem Trend. Negativ bei sinkendem Trend.
    - A Gewichtung
    - **0 < A < 1**
      - Geht A gegen 1 (z.B. 0.99), so wird nur die letzte Beobachtung gewichtet (es zählt nur noch x_t) während der Trend ausser Acht gelassen wird. Man sagt dann, der Glättungseffekt sei gross.
    - **Andere Schreibweise:** $\hat{x}_t$ = ($\hat{x}_{t-1}$ + Tr_{t-1}) + A[x_t - ($\hat{x}_{t-1}$ + Tr_{t-1})]
      - [x_t - ($\hat{x}_{t-1}$ + Tr_{t-1})] entspricht dem **Prognosefehler.**
  - **Trendgleichung:** $\boxed{Tr_t := C(\hat{x}_t - \hat{x}_{t-1}) + (1-C)Tr_{t-1}}$ **(5.2)**

    - Tr_t **Trend** bzw. **Trendsteigung** im Zeitpunkt t + h (Wenn h=1 dann Trend für Morgen)
    - C Gewichtung
    - **0 < C < 1**
      - Geht C gegen 1 (z.B. 0.99), so wird $\hat{x}_t$ - $\hat{x}_{t-1}$ (=neue Niveauverschiebung) stärker gewichtet als die alte Trendprognose Tr_{t-1}.
    - **Andere Schreibweise:** Tr_t = Tr_{t-1} + C[($\hat{x}_t$ - $\hat{x}_{t-1}$) - Tr_{t-1}]

- **Erläuterungen**
  - Die Glättung nach Holt-Winters beruht auf **zwei Parametern** (zweiparametrisches Verfahren). Damit ist eine **grössere Flexibilität** gewährleistet als bei der einfachen exponentiellen Glättung. Allerdings birgt diese zusätzliche Flexibilität die **Gefahr von „Scheinanpassungen“** für Zeitreihen, die keinen nachhaltigen Trend aufweisen.
  - Die Gleichung (5.1) verdeutlicht, dass das neue Niveau $\hat{x}_t$ eine bestimmte Kombination aus der alten Information x_t und dem „erwarteten“ neuen Niveau $\hat{x}_{t-1}$ + Tr_{t-1} ist. Überraschungen bzw. Abweichungen vom erwarteten Niveau werden mit A gewichtet.
  - Die Gleichung (5.2) definiert die neue Trendsteigung Tr_t als Kombination zwischen der alten Trendsteigung und der neuen Niveauverschiebung $\hat{x}_t$ - $\hat{x}_{t-1}$.
- **Prognose:** Eine Prognose zukünftiger Beobachtungen x_{T+h} beruht auf dem letzten „geglätteten“ Wert $\hat{x}_T$ und der zuletzt ermittelten Trendsteigung Tr_t: $\boxed{x_{T+h} := \hat{x}_T + hTr_T}$
- **Startwerte**
  - Wiederum stellt sich das Problem **unbekannter Startwerte:** $\hat{x}_t$, Tr_1, A und C
  - Diese werden so bestimmt, das die **Summe der quadrierten Prognosefehler** möglichst klein ist: $\boxed{\sum_{t=1}^{T-1} [x_{t+1} - (\hat{x}_t + Tr_t)]^2}$

• x_{t+1}	Effektiver Wert bei t + 1
• $(\hat{x}_t + Tr_t)$	Schätzung für Zeitpunkt t+1

2.4.1 Aufgabenbeispiel: 1_L_uebungen exp glättung.xls (Spalte „Holt-Winters (h=1)“)

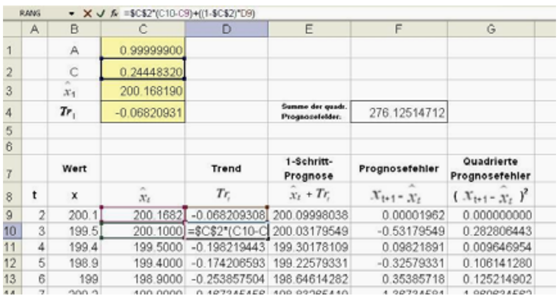


RANG =$C$1*B9+(1-$C$1)*(C9+D9)

	A	B	C	D	E	F	G
1		A	0.99999900				
2		C	0.24448320				
3		$\hat{x}_1$	200.168190				
4		Tr_1	-0.06820931		Summe der quadr. Prognosefelder:	276.12514712	
5							
6							
7		Wert		Trend	1-Schritt-Prognose	Prognosefehler	Quadrierte Prognosefehler
8	t	x	$\hat{x}_t$	Tr_t	$\hat{x}_t + Tr_t$	$x_{t+1} - \hat{x}_t$	$(x_{t+1} - \hat{x}_t)^2$
9	2	200.1	200.1682	-0.068209308	200.09998038	0.00001962	0.000000000
10	3	199.5	=$C$1*B9+(1-$	-0.068204510	200.03179549	-0.53179549	0.282806443
11	4	199.4	199.5000	-0.198219443	199.30178109	0.09821891	0.009646954
12	5	198.9	199.4000	-0.174206593	199.22579331	-0.32579331	0.106141280
13	6	199	198.9000	-0.253857504	198.64614282	0.35385718	0.125214902

Schritt 1: Einrichten

- Spalte A: Zeitpunkt **t** (mit 2 beginnend)
- Spalte B: **Beobachtungen** von t=2 bis t=151
- Spalte C: **Prognose** $\hat{x}_t := Ax_t + (1-A)(\hat{x}_{t-1} + Tr_{t-1})$
  - **Beispiel für Zeile t=3**
    - $\hat{x}_2$ = A · x_2 + (1 - A)($\hat{x}_1$ + Tr_1)
      - $\hat{x}_2$ „Neue“ Prognose für t=3
      - x_2 Wert von t=2
      - $\hat{x}_1$ „Alte“ Prognose für t=2
      - Tr_1 „Alter“ Trend für t=2



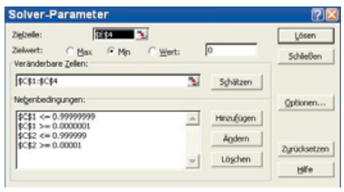
RANG =$C$2*(C10-C9)+((1-$C$2)*D9)

	A	B	C	D	E	F	G
1		A	0.99999900				
2		C	0.24448320				
3		$\hat{x}_1$	200.168190				
4		Tr_1	-0.06820931		Summe der quadr. Prognosefelder:	276.12514712	
5							
6							
7		Wert		Trend	1-Schritt-Prognose	Prognosefehler	Quadrierte Prognosefehler
8	t	x	$\hat{x}_t$	Tr_t	$\hat{x}_t + Tr_t$	$x_{t+1} - \hat{x}_t$	$(x_{t+1} - \hat{x}_t)^2$
9	2	200.1	200.1682	-0.068209308	200.09998038	0.00001962	0.000000000
10	3	199.5	200.1000	=$C$2*(C10-C	200.03179549	-0.53179549	0.282806443
11	4	199.4	199.5000	-0.198219443	199.30178109	0.09821891	0.009646954
12	5	198.9	199.4000	-0.174206593	199.22579331	-0.32579331	0.106141280
13	6	199	198.9000	-0.253857504	198.64614282	0.35385718	0.125214902

- Spalte D: **Trend** $Tr_t := C(\hat{x}_t - \hat{x}_{t-1}) + (1-C)Tr_{t-1}$
  - **Beispiel für Zeile t=3**
    - Tr_2 = C($\hat{x}_2$ - $\hat{x}_1$) + (1 - C) · Tr_1
      - Tr_2 „Neuer“ Trend für t=3
      - $\hat{x}_2$ „Neue“ Prognose für t=3
      - $\hat{x}_1$ „Alte“ Prognose für t=2
      - Tr_1 „Alter“ Trend für t=2
- Spalte E: **Prognose + Trend = 1-Schritt-Prognose** $\boxed{x_{T+h} := \hat{x}_T + hTr_T}$
- Diese beiden Formel können nun weitergezogen werden bis am Ende, wobei logischerweise bereits ein Prognosewert für den Tag t=152 resultiert.

Schritt 2: Optimale Bestimmung der unbekannten Startwerte: $\hat{x}_t$, Tr_1, A und C

- Spalte F: **Prognosefehler** [x_{t+1} – ($\hat{x}_t$ + Tr_t)]
  - Zeile B (Wert) – Zeile C (Prognose)
- Spalte G: **Quadrierte Prognosefehler**
- Die Zelle E4 zeigt nun die **Summe der quadrierten Prognosefelder.**
  - Diese muss nun mit dem Solver minimiert werden (Differentialrechnung).



3. Stationäre stochastische Prozesse

3.1 Zufallsvariablen und Stochastische Prozesse

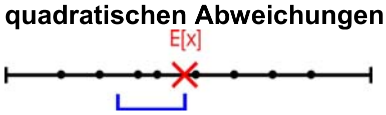
- Eine **Zufallsvariable** ist ein Potential an möglichen Ereignissen. Das Ergebnis kenne ich vor der **Realisation** noch nicht.
  - Z.B. besteht bei einem einzigen Würfelwurf die Zufallsvariable aus den Augenzahlen 1 – 6, wobei die Realisation 6 beträgt, wenn ich eine 6 würfle.
  - Der morgige Wert ist eine Zufallsvariable. Er wird irgendwo im Bereich des schwarzen Balkens realisiert werden.

- **Ein stochastischer Prozess ist eine Betrachtung einer Menge von zeitlich angeordneten Zufallsvariabeln x_t (t=1, … T).**
  - **Beispiel** eines stochastischen Prozesses: zukünftige Entwicklung eines Aktienkurses

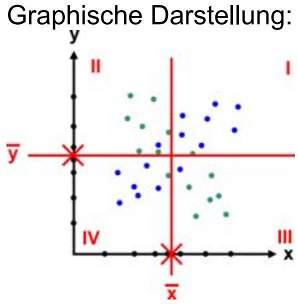
3.2 Masszahlen: Erwartungswert, Varianz, Kovarianz und Korrelation

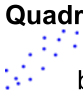
- **Erwartungswert E[x]**
  - Der Erwartungswert E[x] einer Zufallsvariablen x (ZV) ist derjenige Wert, der **im „Zentrum" der möglichen Ausprägungen** liegt. Er beschreibt also die **durchschnittliche Lage.**

  - Als **Schätzer** für den Erwartungswert wird das **arithmetische Mittel** verwendet: $\bar{x} := \frac{1}{T}\sum_{t=1}^{T} x_t$
- **Varianz Var[x]**
  - Die Varianz einer Zufallsvariablen x (ZV) ist ein **Streumass. Es misst die zu erwartenden quadratischen Abweichungen der Realisationen um ihr Zentrum, den Erwartungswert.**



  - Als **Schätzer** für den Erwartungswert wird die **empirische Varianz** verwendet: $\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2$
    - $x_t - \bar{x}$ entspricht dabei der jeweiligen Abweichung zum Erwartungswert
  - Je grösser die Varianz ist, desto schwieriger ist es den Zufall vorherzusagen.
- **Kovarianz**
  - Die Kovarianz ist ein (unnormiertes) **Mass für den Zusammenhang zwischen zwei Zufallsvariablen x und y.** Als eigentliches Abhängigkeitsmass ist diese Grösse jedoch oft unbrauchbar, weil's sie von der Skalierung der Zufallsvariablen abhängt (g, kg, t, CHF, USD…).
  - Als **Schätzer** für die Kovarianz wird die **empirische Kovarianz** verwendet: $\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y})$
  - Graphische Darstellung:

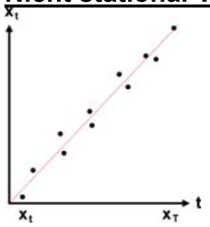


    - **Quadrant IV: $(x_t - \bar{x})(y_t - \bar{y}) > 0$**
      - $(x_t - \bar{x})$ ist negativ und auch $(y_t - \bar{y})$ ist negativ. Daher ist deren Produkt $(x_t - \bar{x})(y_t - \bar{y})$ also **Positiv.**
    - **Quadrant I: $(x_t - \bar{x})(y_t - \bar{y}) > 0$**
    - **Quadrant II: $(x_t - \bar{x})(y_t - \bar{y}) < 0$**
    - **Quadrant III: $(x_t - \bar{x})(y_t - \bar{y}) < 0$**
    -  befindet sich mehrheitlich in den Quadranten I und IV. **Die Kovarianz wird positiv sein. Es besteht also ein positiver Zusammenhang (Korrelation).**
    -  befindet sich mehrheitlich in den Quadranten II und III. **Die Kovarianz wird negativ sein. Es besteht also ein negativer Zusammenhang (Korrelation).**
    - Wäre die **Kovarianz = 0,** so müssten alle Werte in den Quadranten I – IV gleich verteilt sein. Alsdann bestünde **kein Zusammenhang.**
    - Das Problem bei der Kovarianz ist, dass die Zahl der Kovarianz an sich nichts aussagt. Nur das Vorzeichen hat eine Bedeutung.

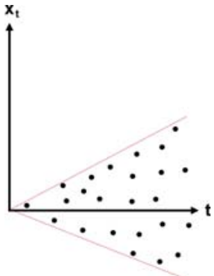
- **Korrelation ρ (rho) = (x_t, x_{t+1})**
  - Anstelle der Kovarianz wird oft die Korrelation als (normiertes) **Abhängigkeitsmass verwendet.**
  - Die Korrelation zeigt den Zusammenhang (Abhängigkeit) zwischen zwei Zufallsvariabeln.
  - Die Korrelation ist deshalb ein normiertes Mass, weil sie unabhängig von der Skalierung der Zufallsvariabeln ist.
  - **Schätzer** für die Kovarianz: $\boxed{\rho = \frac{Kov(x,y)}{\sqrt{Var(x)Var(y)}}}$
    - $-1 \leq \rho \leq 1$
      - **p = 0** Kein linearer Zusammenhang zwischen x und y
      - **p = 1** perfekter linearer positiver Zusammenhang
      - **p = -1** perfekter linearer negativer Zusammenhang

3.3 Stationarität (Stationäre Prozesse)

- Ein stochastischer Prozess heisst dann **stationär,** wenn der Erwartungswert, die Varianz, die Kovarianz und die Korrelation unabhängig von der Zeit sind.
  - Es handelt sich um eine Einschränkung aller möglichen Prozesse.
- **Definition von stationären Prozessen:**
  - **Stätionär wäre:** E[x_t] = μ (mü)
    Egal über welche Zeitperiode gemessen wird, der Erwartungswert bleibt immer derselbe. Der Erwartungswert ist also unabhängig von der Zeit. **Dies bedeutet, dass kein Trend möglich ist.**
    Nicht stationär wäre:



    Dieser stochastische Prozess ist nicht stationär, weil sich der Erwartungswert verändert, es also einen Trend gibt.
  - **Stationär wäre:** Var[x_t] = σ_x^2 (Sigma)
    Egal über welche Zeitperiode gemessen wird, die Varianz darf sich nicht verändern. Die Varianz gibt die Volatilität eines Prozesses wieder. **Dies bedeutet, die Volatilität muss konstant sein.**
    Nicht stationär wäre:



    Hier ist die Varianz rechts grösser als die Varianz links. Es besteht eine unterschiedliche Volatilität.

3.4 Transformation nichtstationärer Prozesse auf stationäre Prozesse

- Das **Problem** besteht darin, dass viele **Finance-Reihen nicht-stationär sind.**
  - z.B. weist der Wechselkurs in der Regel einen starken Trend auf, womit der Erwartungswert sich immer verändert.
- Die **Lösung** dieses Problems stellt die **Transformation** dar.
  - **Bei der Transformation werden nicht-stationäre Prozesse (trendbehaftet, variable Volatilität) in stationäre Prozesse (unabhängig von der Zeit) transformiert.**
  - Diese Transformation ist deshalb wichtig, weil es für stationäre Prozesse gute Prognoseverfahren gibt. Die Prognose von nicht-stationären Prozessen ist dagegen schwieriger.
  - Das **Ziel** besteht den auch darin, eine bestmögliche Prognosefunktion zu finden. Eine solche effiziente Prognosefunktion weist einen sehr kleinen Prognosefehler auf. Wäre die Varianz gar gleich 0, dann hätten wir eine Prognosefunktion ohne Prognosefehler.
  - **Drei Methoden der Transformation in stationäre Prozesse**
    - **Methode 1: Differenzenbildung**
      - **$y_t - y_{t-1} = x_t$**
      - Mit der Differenzenbildung wird jeglicher **Trend** (Niveau) **eliminiert.**

- Man kann so viele Male Differenzen bilden, bis wirklich jeder Trend verschwunden ist. Z.B. $x_t - x_{t-1} = z_t$
- Der Effekt einer Differenzbildung auf die Wechselkursreihe ist in der nächsten Abbildung ersichtlich:

  - o Die schwarze Linie zeigt den Wechselkurs. Dieser ist nicht-stationär, da er einen Trend hat.
  - o Die rote Linie zeigt den transformierten Wechselkurs mit der Methode Differenzbildung. Es handelt sich jetzt um einen stationären Prozess (unabhängig von der Zeit). Der Trend ist verschwunden.

- **<u>Methode 2:</u> Log-Transformation**
  - **$lg(y_t)$**
  - Mit det Log-Transformation wird die Varianz beruhigt, also eine konstante Varianz geschaffen. Somit wird jegliche **Volatilität eliminiert.**
- **<u>Methode 3:</u> Kombination beider Varianten**
  - **$ln(y_t) - ln(y_{t-1}) = x_t$**
  - Ein Nebeneffekt dieser Kombination ist, dass das Transformationsresultat x_t eine Annäherung an den prozentualen Zuwachs zwischen y_{t-1} und y_t darstellt.
  - In der Finance wird sehr oft diese dritte Methode angewandt, obwohl es eigentlich gar nicht notwendig wäre.

3.5 Prognoseverfahren: AR(1)-Prozesse

- AR bedeutet Autoregressiv. Der Vorteil dieses Prognoseverfahrens ist es, das auch **Aussagen über die Unsicherheit der Prognose** gemacht werden können.
- Voraussetzung für diesen **stationären Prozess** ist, dass es sich um eine **stationäre Zeitreihe handelt.** Womöglich ist also eine Transformation notwendig.

1. Vorliegen einer Nicht-stationären Reihe
- o Eine nicht-stationäre Reihe wird an folgenden Eigenschaften erkannt:
  - $E[x_t] \neq \mu$ (mü) — Der Erwartungswert ist über die Zeit **nicht konstant.** Es gibt einen **Trend.**
  - $Var[x_t] \neq \sigma_x^2$ (Sigma) — Die Varianz ist über die Zeit **nicht konstant.** Es gibt eine **variable Volatilität.**

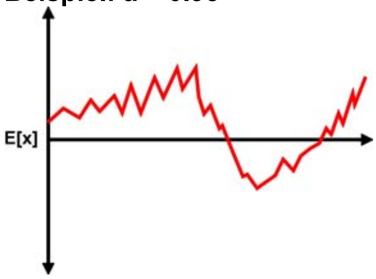
2. Transformation in eine stationäre Reihe
- o Vorgehen gemäss Methode 1, 2 oder 3.

3. Prognosemodell AR(1) vorbereiten
- o Gewöhnliche Regressionsgleichung AR(1): $\boxed{\mathbf{x_t = c + ax_{t-1} + \varepsilon_t}}$ **(9)**
- o Wäre ε_t nicht, würde es sich um eine klassische lineare Gleichung handeln.
- o **Eigenschaften:**
  - **x_t** — Prognose/Schätzer für morgen
  - **x_{t-1}** — Effektiver heutiger Wert
  - **ε_t** — Zu erwartender <u>Prognosefehler von morgen</u>/Störterm/Rauschen/
    - o ε_t = Epsilon
    - o Je grösser Epsilon ist, desto schlechter ist die Prognose.
    - o $E[\varepsilon_t] = 0$
      Im Schnitt werden die ε_t bei 0 sein. Das heisst ihr arithmetisches Mittel wird 0 betragen . Würde ε_t im Schnitt nicht 0 geben sondern 1, würde ich mit meiner Prognose systematisch zu tief liegen. Daher muss zwingen $E[\varepsilon_t] = 0$ gelten.
    - o $Varianz[\varepsilon_t] = \sigma^2$ (bestimmtes Sigma)

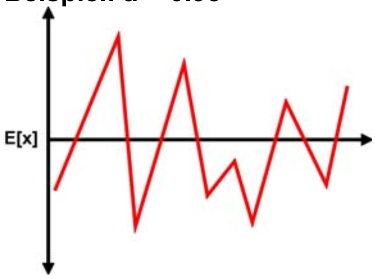
3 Parameter:
  - **c** — Niveau der Reihe (zweit-wichtigster Parameter)
    - o Ist c positiv, so wird die Reihe nach oben verschoben.
    - o Ist c negativ, so wird die Reihe nach unten verschoben.
  - **σ^2** — Varianz (Sigma-Quadrat) aller ε_t
    - o Die gesamte Reihe verfügt über die selbe Varianz, weil es eine stationäre Reihe ist.

- o Durch eine Veränderung der Varianz wird die Reihe aufgebläht oder zusammengedrückt. Es kann also die Skalierung angepasst werden.

- **a** Beschaffenheit der Reihe (wichtigster Parameter weil er Multiplikator ist)
  - o Mithilfe von a kann das Bild der Reihe anpasst werden.
  - o $-1 < a < 1$
  - o **Beispiel: a = 0.95**



    - Bei einem eher grossen a gilt folgendes Prinzip:
      Nach einem grossen Wert folgt grundsätzlich ein grosser Wert.
      Nach einem kleinen Wert folgt grundsätzlich ein kleiner Wert.
  - o **Beispiel: a= -0.95**



    - Bei einem negativen a gilt folgendes Prinzip:
      Nach einem grossen Wert folgt grundsätzlich ein kleiner Wert.
      Nach einem kleinen Wert folgt grundsätzlich ein grosser Wert.
- Mithilfe dieser drei Parameter lassen sich also die verschiedensten Reihen abbilden.

o **Die Parameter a, c und σ^2 müssen zuerst aus den Daten geschätzt werden.**

- **Schätzung von a und c**
  - $$\boxed{\sum_{t=1}^{T-1}\left(x_{t+1}-(\hat{c}+\hat{a}x_t)\right)^2=\sum_{t=1}^{T-1}\hat{\varepsilon}_{t+1}^2 \rightarrow Min}$$ **(10)** ^: Schätzer
    - o x_{t+1} Effektiver Wert von Morgen
    - o $(c + ax_t)$ Prognose für Wert von morgen
    - o $[x_{t+1} - (c + ax_t)]$ Prognosefehler von Morgen
  - **Die Summe aller quadrierten morgigen Prognosefehler wird also minimiert.** Somit werden diejenigen c und a ausgewählt, welche den kleinsten Prognosefehler in der Vergangenheit ergeben. Wenn man in der Vergangenheit die minimalsten Fehler hatte, wird man auch in der Zukunft kleine Fehler haben.
  - **Lösung dieser Formel**
    - o **1. Möglichkeit: Solver im Excel**
      - Es entsteht eine Annäherungslösung, nicht die exakte Lösung.
    - o **2. Möglichkeit: Weitere Formeln**
      - Es entsteht die exakte Lösung.
      - $$\hat{a} = \frac{\sum_{t=1}^{T-1}(x_{t+1}-\bar{x})(x_t-\bar{x})}{\sum_{t=1}^{T}(x_t-\bar{x})^2}$$ $\bar{x}$: Arith. Mittel
      - $\hat{c} = \bar{x}(1 - \hat{a})$
- **Schätzung von σ^2**
  - $$\boxed{\hat{\sigma}^2 = \frac{1}{T}\cdot\sum_{t=1}^{T}\varepsilon_t^2}$$
  - $$\boxed{\varepsilon_t = x_t - \hat{c} - \hat{a}\,(x_{t-1})}$$

4. Erstellen der Prognose für die stationäre Reihe

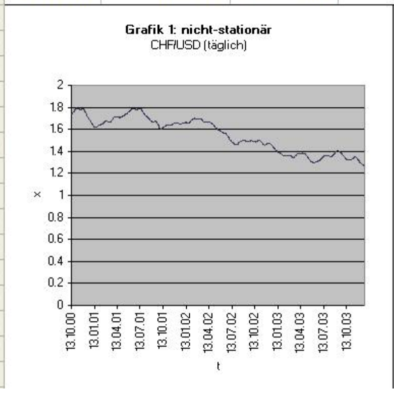
- Die Prognose wird zunächst für die **transformierten Daten** durchgeführt.

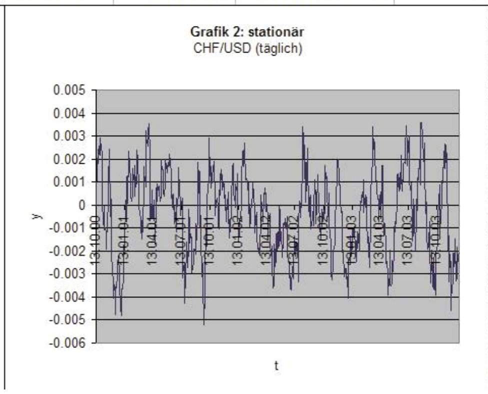
5. Rücktransformation

- Mithilfe der Rücktransformation werden die **prognostizierten stationären Daten** nun in **prognostizierte nicht-stationäre Daten** zurück-transferiert.

- **Aufgabenbeispiel:** 2_L_uebungen ar(1).xls (Spalte „Einführungsbeispiel AR(1)“)

	A	B	C	D	E	F	G	H	I	J
1	c	-0.00003648								
2	a	0.91770000			Summe Prognosefehler	0.0003993				
3	σ^2									
4										
5										
6										
7		nicht-stationär	stationär			1-Schritt-Prognose (stationär)	nicht-stationär aber logarithmiert	nicht-stationär		
8	**Datum (t)**	**CHF/USD (x_t)**	**$y_t = \ln(x_t) - \ln(x_{t-1})$**	$\varepsilon_t = x_t - c - \hat{a}\,(x_{t-1})$	ε_t^2	**$y_t = c + ay_{t-1} + 0$**	**$\ln(x_t) = \ln(x_{t-1}) + y_t$**	**Prognose x_t**		
9	13.10.00	1.73991	0	-	-	-0.00205353477	-	-		
10	16.10.00	1.74346	0.0020383	0.0000000	0.0000000	-0.00192100882	0.2330009	1.2623826	= 1-Schritt-Prognose	
11	17.10.00	1.74658	0.0017879	-0.0000461	0.0000000	-0.00179938976	0.2312015	1.2601131	= 2-Schritt-Prognose	
12	18.10.00	1.75047	0.0022247	0.0006204	0.0000004	-0.00168777994	0.2295137	1.2579881	= 3-Schritt-Prognose	
13	19.10.00	1.75473	0.0024307	0.0004255	0.0000002	-0.00158535561	0.2279284	1.2559954		
14							2264370	1.2541236		
15							2250319	1.2523627		
16							2237060	1.2507032		
17							2224527	1.2491367	= 8-Schritt-Prognose	
18							2212660	1.2476553		
19							2201406	1.2462519		
20							2190713	1.2449200		
21							2180535	1.2436536		
22							2170830	1.2424472		
23							2161559	1.2412959		
24							2152686	1.2401950		
25							2144179	1.2391404		
26							2136007	1.2381281		
27							2128142	1.2371548		
28							2120561	1.2362172		





- **1. Vorliegen einer Nicht-stationären Reihe**
  - **Spalte A:** Datum (t)
  - **Spalte B:** x_t (CHF/USD)
  - Grafik 2 bildet die vergangenen Beobachtungen ab. Man erkennt einen Trend sowie eine variable Volatilität, somit handelt es sich um eine nicht-stationäre Reihe die in eine stationäre Reihe transferiert werden muss.
- **2. Transformation in eine stationäre Reihe**
  - **Spalte C:** $y_t = \ln(x_t) - \ln(x_{t-1})$
  - Für die erste Differenz wird einfach 0 eingetragen.
  - Somit verfügen wir nun mit y_t über eine stationäre Reihe.
  - Grafik 2 zeigt das neue Bild der stationären Reihe.
- **3. Prognosemodell AR(1) vorbereiten**
  - **Spalte D:** ε_t (Prognosefehler)
  - **Spalte F:** $y_t = c + ay_{t-1} + \varepsilon_t$ (Schätzung für die stationären Daten y)
    - Das Problem besteht darin, dass ich ε_t nicht kenne. Denn es handelt sich dabei um den Prognosefehler für den Tag, wo ich die Schätzung machen will. Ich kann über ε_t keine Aussage machen. Die Prognosefunktion wird deshalb folgendermassen abgeändert: $\mathbf{y_t = c + ay_{t-1} + 0}$
      Der Prognosefehler wird gar nicht berücksichtigt.
    - Ich setze deshalb auch für den ersten Prognosefehler 0 (Zelle D11) ein.
    - Für den letzten Wert y_{t-1} (Zelle F10) verwende ich den allerletzten Wert der stationären Reihe y_T.
    - Damit kann ich die erste Prognose für y_2 machen (Zelle F11). Diese Formel kann ich sodann nach unten weiter ziehen.
  - **Schätzung der drei Parameter**
    - **Spalte D:** $\varepsilon_t = x_t - c - \hat{a}\,(x_{t-1})$
      - Diese Formel kann erstmals für Zelle D12 verwendet werden.
      - Diese Formel kann sodann weiter gezogen werden.
    - **Spalte E:** ε_t^2 (quadrierte Prognosefehler)
    - **a und c**
      - Zelle F2: Summe aller quadrierten morgigen Prognosefehler, also die Summe aller ε_t^2

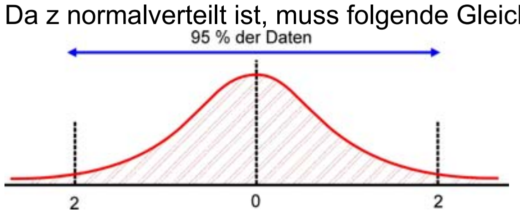
        - Mit Hilfe des Solvers (oder der Formeln) kann nun die Zelle F2 (Summe der quadrierten Prognosefehler) minimiert werden. Beachte, dass σ^2 nicht im Solver miteinbezogen werden soll (keine veränderbare Zelle).
      - σ^2 wird im folgenden Beispiel nicht angegangen, weil wir von einer vereinfachten Prognosefunktion $\mathbf{y_t = c + ay_{t-1} + 0}$ ausgehen. ε_t dient uns anfänglich nur zur Bestimmung des optimalen a und c. **Wir machen für ε_t noch keine Prognose.**
  - **4. Erstellen der Prognose für die stationäre Reihe**
    - In der Spalte F haben wir nun die 1-Schritt-Prgnose für die stationären Daten erstellt. Es fehlt uns nun noch die Transformation auf die nicht-stationären Daten, also auf jene Daten, die uns wirklich interessieren.
  - **5. Rücktransformation**
    - Zu Beginn haben wir folgende Transformation durchgeführt: $y_t = \ln(x_t) - \ln(x_{t-1})$
    - **Spalte G**
      - Jetzt müssen wir die umgekehrte Transformation vornehmen.
        $y_t = \ln(x_t) - \ln(x_{t-1}) \qquad / + \ln(x_{t-1})$
        $\boxed{\mathbf{\ln(x_t) = \ln(x_{t-1}) + y_t}}$
        - Ein Problem besteht darin, dass für $\ln(x_1) = \ln(x_{1-1}) + y_1$ ein $\ln(x_{1-1})$ nicht vorhanden ist. Es wird deshalb der letzte x-Wert der nicht-stationären Ausgangslage verwendet: $\ln(x_T)$
    - **Spalte H**
      - Von $\ln(x_t)$ müssen wir nun noch auf x_t zurück rechnen: $x_t = e^{\ln(xt)}$
        e = 2.718281828
- **Zusammenfassung**
  - 1-Schritt-Prognose: $x_{t+1} = c + ax_T + \cancel{\varepsilon_{t+1}}$ (wobei ε_{t+1} nicht bestimmbar, also 0 ist)
  - 2-Schritt-Prognose: $x_{t+2} = c + ax_{t+1} + \cancel{\varepsilon_{t+2}}$ (wobei ε_{t+2} nicht bestimmbar, also 0 ist)
  - k-Schritt-Prognose: $x_{t+k} = c + ax_{t+k-1} + \cancel{\varepsilon_{t+k-1}}$ (wobei ε_{t+k-1} nicht bestimmbar, also 0 ist)

3.6 Prognoseintervalle für AR(1)-Prozesse

- **Idee**
  - Konstruiere ein Intervall so, dass $\hat{x}_{t+h}$ mit vorgegebener Wahrscheinlichkeit im Intervall liegt.
    - x_T Effektiver Wert heute
    - $\hat{x}_{T+h}$ Prognose/Schätzer für h; h=1 morgen (1-Schritt-Prognose); h=2 übermorgen (2-Schritt-Prognose)
- **Prognosefehler**
  - $x_{T+1} - \hat{x}_{T+1}$ — **Prognosefehler für morgen** (Schreibweise 1)
    Effektiver Wert morgen – Prognose für Morgen = Prognosefehler
  - $(c + ax_T + \varepsilon_{t+1}) - (c + ax_T)$ — **Prognosefehler für morgen** (Schreibweise 2)
    Effektiver Wert morgen – Prognose für Morgen = Prognosefehler
  - ε_{t+1} — **Prognosefehler für morgen** (Schreibweise 3)
    - Im Durchschnitt muss der Prognosefehler ε_t Null geben, ansonsten wäre unsere Prognose systematisch zu tief oder zu hoch. Es gilt also: Erwartungswert $E[x_{T+1} - \hat{x}_{T+1}] = 0$
    - Weiter muss für die Varianz der Prognosefehler ε_t gelten: $Var[x_{T+1} - \hat{x}_{T+1}] = \sigma^2$
- **Herleitung**
  - Es wird die Annahme getroffen, die Prognosefehler ε_t seien (standard-)normalverteilt.

    In diesem Fall gilt: $\dfrac{x_{T+1} - \hat{x}_{T+1}}{\sigma} = z$

    Definitionsgemäss ist dann der resultierende Wert z normalverteilt mit E[z] = 0 und Var[z] = 1. Die Standardabweichung σ der Prognosefehler muss geschätzt werden.
  - Da z normalverteilt ist, muss folgende Gleichung gelten:

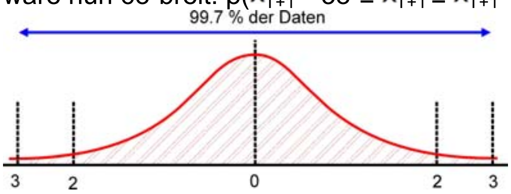


$\rho(-2 \leq z \geq 2) = 0.95$
Die Wahrscheinlichkeit dass z zwischen $-2\cdot\sigma$ und $2\cdot\sigma$ (wobei σ = 1 da normalverteilt) liegt beträgt 95 %.

- Damit muss auch gelten: $\rho(-2 \leq \frac{x_{T+1} - \hat{x}_{T+1}}{\sigma} \leq 2) = 0.95$

  Umgeformt gilt also: $\rho(-2\sigma \leq x_{T+1} - \hat{x}_{T+1} \leq 2\sigma) = 0.95\sigma$

  Weiter gilt also: $\rho(\hat{x}_{T+1} - 2\sigma \leq x_{T+1} \leq \hat{x}_{T+1} + 2\sigma) = 0.95$
  - Die linke Seite der Ungleichung, $\hat{x}_{T+1} - 2\sigma$ stellt die untere Grenze des Intervalls dar.
  - Die rechte Seite der Ungleichung, $\hat{x}_{T+1} + 2\sigma$ stellt die obere Grenze des Intervalls dar.
  - Im obigen Fall ist das Intervall 4σ breit und weist eine Wahrscheinlichkeit von 0.95 auf.
  - Wenn wir eine sicherere Prognose wollten, also beispielsweise mit einer Wahrscheinlichkeit von 99.7 %, so müssten wir $3 \cdot \sigma$ (wobei $\sigma=1$) wählen. Das Intervall wäre nun 6σ breit: $\rho(\hat{x}_{T+1} - 3\sigma \leq x_{T+1} \leq \hat{x}_{T+1} + 3\sigma) = 0.997$



- Aus dieser Herleitung gibt sich sodann die folgende (identische) Formel für die Prognoseintervalle.

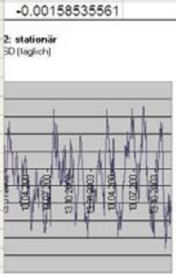
- **Formeln Prognoseintervalle**
  - $PI\ (1\text{-Schritt}) = \hat{x}_{T+1} \pm 2\sqrt{Var(x_{T+1} - \hat{x}_{T+1})}$
  - $PI\ (k\text{-Schritt}) = \hat{x}_{T+k} \pm 2\sqrt{Var(x_{T+k} - \hat{x}_{T+k})}$

  Das „2" gilt nur für den 95 %-Prognoseintervall.
  Das „3" gälte für den 99.7 %-Prognoseintervall.

  - Ein Problem besteht darin, dass die Varianz des Prognosefehlers, also $Var(x_{T+k} - \hat{x}_{T+k})$ unbekannt ist.
  - Es gilt jedoch: $E[\varepsilon_{th}^2] = Var[\varepsilon_t]$
  - $E[\varepsilon_{th}^2] = \sigma^2 + a^2 \cdot E[e_{t(h-1)}^2]$
  - **Beispiel**
    - **1-Schritt-Prognose**
      - $a = 0.5$
        $\sigma^2 = 10$ Varianz (Sigma-Quadrat) aller ε_t
      - Varianz = <u>10</u> identisch mit σ^2
    - **2-Schritt-Prognose**
      - $10 + 0.5^2 \cdot 10 = \underline{12.5}$
    - **3-Schritt-Prognose**
      - $10 + 0.5^2 \cdot 12.5 = \underline{13.125}$

- **Aufgabenbeispiel:** 2_L_uebungen ar(1).xls (Spalte „Prognoseintervalle AR(1)")

	F	G	H	I	J	K	L	M	N
1									
2	0.0003993					σ^2	0.0000004823		
3								Mit 95 % Wahrscheinlichkeit wird der Wert morgen/übermorgen	
4								überübermorgen zwischen diesen beiden Werten liegen	
5									
6								Prognoseintervall für **stationäre** Prognosen	
7	1-Schritt-Prognose (stationär)	nicht-stationär aber logarithmiert	nicht-stationär			Varianzen	2 * Wurzel	**Untere Grenze Prognoseintervall (-)**	**Obere Grenze Prognoseintervall (+)**
8	$y_t = c + ay_{t-1} + 0$	$\ln(x_t) = \ln(x_{t-1}) + y_t$	**Prognose** x_t			$Var(x_{T+k} - \hat{x}_{T+k})$	$2\sqrt{Var(x_{T+1} - \hat{x}_{T+1})}$	$\hat{x}_{T+1} \pm 2\sqrt{Var(x_{T+1} - \hat{x}_{T+1})}$	$\hat{x}_{T+1} \pm 2\sqrt{Var(x_{T+1} - \hat{x}_{T+1})}$
9	-0.00205353477	-	-	13.10.2000	**1-Schritt:**	0.0000004823041	0.001388962	-0.00344249718	-0.00066457237
10	-0.00192100882	0.2330009	1.2623826	16.10.2000	**2-Schritt:**	0.0000008884878	0.001885193	-0.00380620145	-0.00003581620
11	-0.00179938976	0.2312015	1.2601131	17.10.2000	**3-Schritt:**	0.0000012305648	0.002218617	-0.00401800630	0.00041922679
12	-0.00168777994	0.2295137	1.2579881	18.10.2000	**4-Schritt:**	0.0000015186530	0.002464673	-0.00415245273	0.00077689285
13	-0.00158535561	0.2279284	1.2559954	19.10.2000		0.0000017612731	0.002654259	-0.00423961492	0.00106890370
14		2264370	1.2541236	20.10.2000		0.0000019656013	0.002803998	-0.00429535890	0.00131263729
15		2250319	1.2523627	23.10.2000		0.0000021376811	0.002924162	-0.00432926393	0.00151906039
16		2237060	1.2507032	24.10.2000		0.0000022826021	0.003021657	-0.00434759839	0.00169571468
17		2224527	1.2491367	25.10.2000		0.0000024046506	0.003101387	-0.00435468399	0.00184809039
18		2212660	1.2476553	26.10.2000		0.0000025074367	0.003166978	-0.00435360796	0.00198034710
19		2201406	1.2462519	27.10.2000		0.0000025940003	0.00322118	-0.00434663082	0.00209572941
20		2190713	1.2449200	30.10.2000		0.0000026669019	0.00326613	-0.00433543647	0.00219682432
21		2180535	1.2436536	31.10.2000		0.0000027282977	0.003303512	-0.00432129405	0.00228572976
22		2170830	1.2424472	01.11.2000					
23		2161559	1.2412959	02.11.2000		**Die Varianz bleibt beschränkt. Sie wird nicht unendlich gross.**			
24		2152686	1.2401950	03.11.2000		Es handelt sich um eine Prognose für einen stationären Prozess. Die Varianz ist			
25		2144179	1.2391404	06.11.2000		endlich. Die Realisationen können nicht beliebig zu oder abnehmen. Der Prognose-			
26		2136007	1.2381281	07.11.2000		fehler kann nicht unendlich gross werden.			



- **Zelle K2**
  - **σ^2** Varianz (Sigma-Quadrat) aller ε_t (Prognosefehler)
    =Varianz(alle ε_t)
- **Spalte J**
  - $Var(x_{T+k} - \hat{x}_{T+k})$
  - Diese ist identisch mit $E[\varepsilon_{th}^2] = \sigma^2 + a^2 \cdot E[e_{t(h-1)}^2]$

- **1-Schritt (Zeile 10)**
  - Varianz = $E[e_{t(h-1)}^2]$ = 0.0000004823
- **2-Schritt (Zeile 11)**
  - $E[\varepsilon_{th}^2] = \sigma^2 + a^2 \cdot E[e_{t(h-1)}^2]$
  - $0.0000004823 + 0.91770000^2 \cdot 0.0000004823 = 0.0000008884878$
- **3-Schritt (Zeile 12)**
  - $0.0000004823 + 0.91770000^2 \cdot 0.0000008884878 = \ldots$

o **Spalte L**

- $2\sqrt{Var(x_{T+1} - \hat{x}_{T+1})}$

o **Spalte M und N**

- $\hat{x}_{T+1} \pm 2\sqrt{Var(x_{T+1} - \hat{x}_{T+1})}$
- Beachte, dass es sich bei $\hat{x}_{T+k}$ in obiger Formel um die stationären Prognosen (im Beispiel y_t) handelt und noch nicht um den Prognoseintervall für die nicht-stationären Prognosen (im Beispiel x_t).
- **Mit 95 % Wahrscheinlichkeit wird der Wert morgen/übermorgen/überüber-morgen zwischen diesen beiden Werten liegen.**