

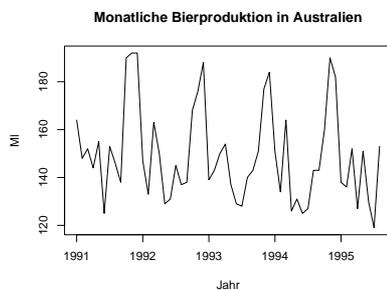
Basiswerkzeuge

Lernziele

- Beschreiben von Zeitreihen
- Graphische Darstellungen
- Univariate und bivariate Maßzahlen
- Transformationen von Zeitreihen
- Maßzahlen für die Prognosegüte

Zeitreihen-Plot

Zeitpfad wird visualisiert:



Muster und Charakteristika werden meist sichtbar:

- Trends
- Zyklen
- Saison
- Irreguläres Verhalten (keine Regelmäßigkeit)
- Ausreißer, atypische Beobachtungen

Muster

Wir können vier Muster unterscheiden:

- **Horizontales Muster.**
Daten fluktuieren um ein konstantes Mittel.
z.B.: Renditen, Bierproduktion
- **Saisonales Muster.**
Saisonbehafte Daten, fluktuieren mit fixer Periodizität.
z.B.: Stromverbrauch, Speiseeisproduktion
- **Zyklisches Muster.**
Daten fluktuieren in Zyklen unregelmäßiger Länge und Amplitude.
z.B.: Konjunkturzyklen
- **Trend.**
Daten steigen oder fallen über eine längere Periode.
z.B.: BIP-real, Stromproduktion

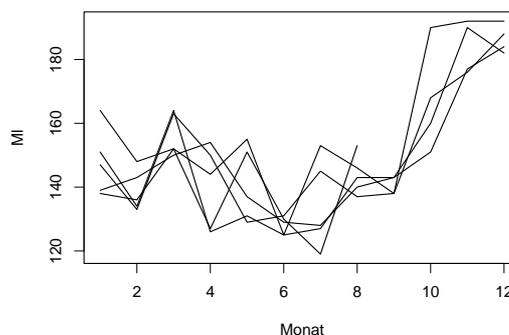
Überlagerung der Strukturen

- Ausgeprägte Strukturen machen die Wahl eines geeigneten Modells leicht.
- In kurzen Reihen sind einzelne Muster nicht zuverlässig erkennbar.
- Die Unterscheidung zwischen Trend und langen Zyklen ist allgemein ein schwieriges Problem.
- Zyklen, die kurz im Verhältnis zu Frequenz der Beobachtungen liegen sind schwer zu erkennen (oder erzeugen Scheinperiodizitäten).

Saisonale Plots

Saisonale Plots zeigen die Daten für einzelne „Saisonen“ (hier Monate).

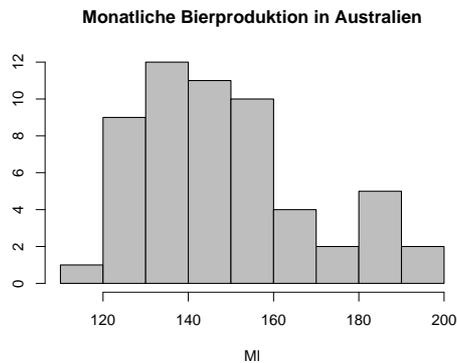
Monatliche Bierproduktion in Australien



Histogramm

Das **Histogramm** dient zur graphischen Beschreibung der Verteilung einer Variablen.

Nur sinnvoll bei horizontalem Muster (stationärer Struktur).

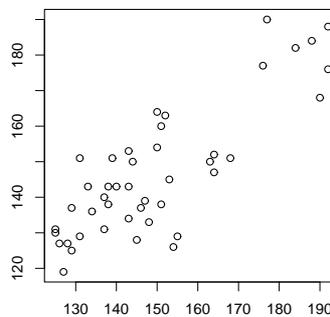


Streudiagramm (*scatter plot*)

Im **Streudiagramm** werden zwei Variable gegeneinander aufgetragen (Punktwolke).

Es erlaubt Korrelationen abzuschätzen.

(Korrelations-/Regressionsanalyse)



Univariate Maßzahlen

Maßzahlen für Zeitreihen $\{y_t\}, t = 1, \dots, T$.

- **Mittelwert**

$$\bar{y} = \frac{1}{T} \sum y_t$$

- **Median**

Me = „mittlerer Wert“

- **mittlerer absoluter Abstand**

$$\text{MAD} = \frac{1}{T} \sum |y_t - \bar{y}|$$

- **mittlerer quadratischer Abstand**

$$\text{MSD} = \frac{1}{T} \sum (y_t - \bar{y})^2$$

- **Varianz**

$$s_y^2 = \frac{1}{T-1} \sum (y_t - \bar{y})^2$$

- **Standardabweichung**

$$s_y = \sqrt{s^2} = \sqrt{\frac{1}{T-1} \sum (y_t - \bar{y})^2}$$

Summation über alle $t = 1, \dots, T$.

Bivariate Maßzahlen

Zwei Zeitreihen $\{x_t\}$ und $\{y_t\}$, $t = 1, \dots, T$.

- **Kovarianz** $\text{Cov}_{xy} = \frac{1}{T-1} \sum (x_t - \bar{x})(y_t - \bar{y})$

- **Korrelation** $r_{xy} = \frac{\text{Cov}_{xy}}{s_x s_y} = \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum (x_t - \bar{x})^2} \sqrt{\sum (y_t - \bar{y})^2}}$

Der Korrelationskoeffizient liegt immer zwischen -1 und 1 :

$$-1 \leq r_{xy} \leq 1$$

Wir verwenden hier die Stichprobenvarianz bzw. -kovarianz.

Autokorrelation

Die **Autokorrelation k -ter Ordnung** misst die Korrelation zwischen Werten einer Reihe, die k Perioden voneinander entfernt sind:

$$r_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}$$

Die Werte y_{t-k} und y_t sind k Perioden voneinander entfernt. Wir sagen y_{t-k} ist der Wert von y_t um k Perioden **verzögert**.

Durch Zusammenfassen der Autokorrelationen der Ordnung $0, 1, 2, 3, \dots$ erhält man die **Autokorrelationsfunktion** oder das **Korrelogramm**: $r_0 = 1, r_1, r_2, r_3, \dots$

Nur sinnvoll für stationäre Zeitreihen (horizontales Muster).

Autokorrelation erster Ordnung für die Australische Bierproduktion:

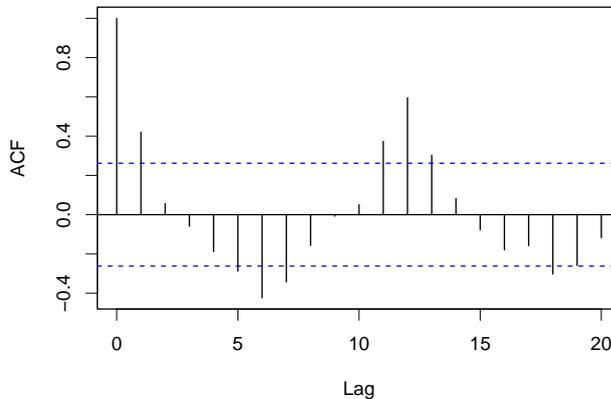
(1)	(2)	(3)	(4)	(5)	(6)	(7)
t	y_t	y_{t-1}	$(y_t - \bar{y})$	$(y_{t-1} - \bar{y})$	$(y_t - \bar{y})^2$	$(y_t - \bar{y}) \cdot (y_{t-1} - \bar{y})$
1	164	—	14.70	—	215.99	—
2	148	164	-1.30	14.70	1.70	-19.16
3	152	148	2.70	-1.30	7.27	-3.51
4	144	152	-5.30	2.70	28.13	-14.30
⋮						⋮
55	119	130	-30.30	-19.30	918.31	584.97
56	153	119	3.70	-30.30	13.66	-112.01
Σ	8361				21135.84	8893.51

$$T = 56, \bar{y} = 8361/56 = 149.30,$$

$$r_1 = \frac{\sum_{t=2}^T (y_t - \bar{y})(y_{t-1} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} = \frac{8893.51}{21135.84} = 0.421$$

Korrelogramm

Korrelogramm der australischen Bierproduktion



Stationärer Prozess

Ein **stationärer Prozess** hat konstanten

- Erwartungswert $E(y_t) = \mu$ für alle t
- Varianz $V(y_t) = \sigma$ für alle t
- Autokovarianzen $\text{Cov}(y_t, y_{t+k}) = \gamma(k)$ für alle t
- Autokorrelationen $\rho_k = \gamma(k) / \gamma(0)$ für alle t

Der Zeitreihenplot eines stationären Prozesses hat:

- horizontaler Verlauf des Mittels,
- Streuung um Mittel gleichmäßig über die Zeit.

Die Autokorrelationsfunktion eines stationären Prozesses

- geht geometrisch gegen Null,
- macht saisonale Muster sichtbar.

Interpretation von r_k

In der Regressionsanalyse wird ein lineares Modell betrachtet:

$$y_t = a + b x_t + \epsilon_t$$

Die Methode der kleinsten Quadrate liefert die Schätzer

$$\hat{b} = \frac{\text{Cov}(x, y)}{s_x^2} = \text{Corr}(x, y) \frac{s_y}{s_x} \quad \text{und} \quad \hat{a} = \bar{y} - \hat{b} \bar{x}$$

Im Fall $x_t = y_{t-k}$ gilt $\sigma_x = \sigma_y$ und $\text{Corr}(x, y) = r_k$ und somit

$$y_t = a + \rho_k y_{t-k} + \epsilon_t$$

Diese Regression macht Sinn, und y_{t-k} kann zur Prognose y_t verwendet werden.

Ziel der Transformation

Reihen können mit Standardverfahren modelliert werden, wenn sie

- einen linearen Trend, oder
- konstantes Mittel und
- konstante Varianz

besitzen.

Exponentielles Wachstum ist nicht linear kann aber durch Logarithmierung **linearisiert** werden:

$$x_t = \log(y_t)$$

(log ist hier der natürliche Logarithmus.)

z.B.: Aktienindizes, reales BIP

Varianzstabilisierung

Allgemeiner als Logarithmierung ist die **Box-Cox Transformation**.

$$y_t^{BC} = \begin{cases} -y_t^p & \text{für } p < 0 \\ \log(y_t) & \text{für } p = 0 \\ y_t^p & \text{für } p > 0 \end{cases}$$

Steigt (selten fällt) die Varianz einer Reihe mit steigendem Mittel, so kann eine konstante Varianz durch die geeignete Wahl von p erreicht werden (**Varianzstabilisierung**). **Rücktransformation:**

$$y_t = \begin{cases} (-y_t^{BC})^{1/p} & \text{für } p < 0 \\ \exp(y_t^{BC}) & \text{für } p = 0 \\ (y_t^{BC})^{1/p} & \text{für } p > 0 \end{cases}$$

Varianzstabilisierung / (2)

Bedeutung des Parameters p :

- $p > 1$: Die Varianz für große Werte wird größer.
- $p = 1$: Reihe bleibt unverändert.
- $p < 1$: Die Varianz für große Werte wird kleiner.
- $p = 0$: Geeignete Transformation für exponentielles Wachstum.
- $p < 0$: Die Varianz für große Werte wird noch stärker reduziert.

Prognose

Das **Prognoseproblem** lautet:

Wir beobachten Daten $y_i, i = 1, \dots, t$, und wollen wissen, wie sich die Reihe in der Zukunft, \hat{y}_j , für die Zeitpunkte $j = (t + 1), (t + 2), \dots$ entwickelt.

Allgemein unterscheidet man

- **Punkt-** und
- **Intervallprognose.**

Ähnlich wie Punkt- und Intervallschätzer.

Naive Prognose

Der letzte Wert wird als Prognose verwendet:

$$\hat{y}_{t+1} = y_t$$

z.B.: Wenn der Bierabsatz in der letzten Jahr 80 MI betragen hat, dann prognostizieren wir für heuer ebenfalls 80 MI.

Saisonale Effekte können dabei berücksichtigt werden:

$$\hat{y}_{t+1} = y_{t-i+1}$$

wobei i die Länge der Periode ist.

z.B.: Wenn der Bierabsatz im August des letzten Jahres 60 MI betragen hat, dann prognostizieren wir für den heuerigen August ebenfalls 60 MI.

Einfache Prognosen

Angenommen unsere Reihe weist

- ein **horizontales Muster** auf:
Das Mittel der Reihe wird fortgeschrieben:

$$\hat{y}_{t+1} = \bar{y}$$

- einen **linearen Trend** auf:
Der Trend wird zur Prognose fortgeschrieben:

$$\hat{y}_{t+1} = \hat{a} + \hat{b}(t + 1)$$

\hat{a} und \hat{b} sind die Schätzer für das lineare Modell $y_t = a + bt + \epsilon_t$.

τ -Schritt Prognose und Prognosefehler

Die Prognose für τ Schritte in die Zukunft heißt **τ -Schritt Prognose**:

$$\hat{y}_{t+\tau}$$

Der **Prognosefehler** für τ ($\tau = 1, 2, \dots$) ist

$$e_{t+\tau} = y_{t+\tau} - \hat{y}_{t+\tau}$$

$y_{t+\tau}$ sind die zukünftigen, noch nicht beobachteten Werte.

Ziel ist es, den Prognosefehler so klein wie möglich zu machen.

in-sample / *out-of-sample* Prognose

Wir unterscheiden:

- ***in-sample* Prognose (ex post)**:
Ein Modell wird für $t = 1, \dots, T$ geschätzt. Für die Zeitpunkte der Schätzperiode (i.e. $1 \leq t \leq T$) werden Prognosen berechnet.
- ***out-of-sample* Prognose (ex ante)**:
Die Prognose wird für Zeiten außerhalb der Schätzperiode berechnet. z.B. für $T + \tau$, $\tau = 1, 2, \dots$

Out-of-sample Prognosen werden oft so durchgeführt, daß die Länge der Schätzperiode festgehalten wird, und diese Periode über die Zeit hinweg verschoben wird (*moving window*).

Der **Prognosehorizont** (z.B. 5-Schritt Prognose) wird festgehalten und mit der Prognoseperiode mitbewegt.

Prognoseintervall

Sind die Fehler ϵ_t bzw. e_t normalverteilt, so kann man auch **Prognoseintervalle** angeben.

95% Prognoseintervall:

$$[\hat{y}_{t+\tau} - 1.96 \hat{\sigma}_e(\tau), \hat{y}_{t+\tau} + 1.96 \hat{\sigma}_e(\tau)]$$

$\hat{\sigma}_e(\tau)$ ist die Standardabweichung des τ -Schritt Prognosefehlers.

Wie man diese bestimmt, werden wir später sehen.

Maße für die Prognosegüte

- **Mittlerer Fehler, Mean Error, ME**

misst den *Bias* (Verzerrtheit) einer k -Schritt Prognose:

$$ME = \frac{1}{k} \sum_{j=1}^k (y_{t+j} - \hat{y}_{t+j})$$

- **Mittlere Fehlerquadrate, Mean Square Error, MSE**

misst $Bias^2 + \text{Fehlervarianz}$:

$$MSE = \frac{1}{k} \sum_{j=1}^k (y_{t+j} - \hat{y}_{t+j})^2$$

Der MSE ist analog zur Fehlerquadratsumme konstruiert.

Maße für die Prognosegüte / (2)

- **Root Mean Square Error, RMSE**

analog zur Standardabweichung:

$$RMSE = \sqrt{MSE}$$

- **Mittlerer absoluter Fehler, Mean Absolute Error, MAE**

misst die absoluten Abweichungen (nicht die quadrierten).

$$MAE = \frac{1}{k} \sum_{j=1}^k |y_{t+j} - \hat{y}_{t+j}|$$

Maße für die Prognosegüte / (3)

Alle vier Kriterien können auch als **relative Fehler** (*Percent Error*) definiert werden.

- **Mean Absolute Percent Error, MAPE**

$$MAPE = \frac{100}{k} \sum_{j=1}^k \left| \frac{y_{t+j} - \hat{y}_{t+j}}{y_{t+j}} \right|$$

Hier wird der Fehler in Relation zum realisierten Wert gesetzt. „Große“ Fehler bei „großen“ y_t -Werten werden genauso bewertet wie „kleine“ Fehler bei „kleinen“ Realisationen.

Relative Fehler sind nur sinnvoll, wenn die verwendete Skala einen sinnvollen Nullpunkt hat.

Maße für die Prognosegüte / (4)

Wenn ein Modell für mehrere Reihen, $y_t^{(i)}$, $i = 1, \dots, n$ auf dessen Prognosegüte überprüft werden sollen, hält man die Schrittlänge k fest und mittelt über die Werte $\hat{y}_{t+k}^{(i)}$, $i = 1, \dots, n$, z.B.:

$$\text{MSE}_k = \frac{1}{n} \sum_{i=1}^n \left(y_{t+k}^{(i)} - \hat{y}_{t+k}^{(i)} \right)^2$$

Beispiele:

- Der datengenerierende Prozess ist bekannt oder vorgegeben. Mittels Simulation werden künstliche Datensätze generiert und die Prognosen aus den ausgewählten Modellen getestet.
- Die Modelle sollen für eine Vielzahl von Datensätzen eingesetzt werden.

Theil's U -Statistik

Theil's U -Statistik vergleicht die verwendete Methode mit der Güte der naiven Methode.

$$U = \sqrt{\frac{\sum_{j=1}^{k-1} \left(\frac{y_{t+j+1} - \hat{y}_{t+j+1}}{y_t} \right)^2}{\sum_{j=1}^{k-1} \left(\frac{y_{t+j+1} - y_{t+j}}{y_t} \right)^2}}$$

Interpretation:

- $U = 1$: die naive Methode ist genauso gut.
- $U < 1$: die gewählte Prognosetechnik ist besser als die naive Methode.
- $U > 1$: die naive Methode produziert bessere Prognosen.