

2. Grundlagen

2.1 Zustandsmaße, Ereignismaße, Kennziffern

Verschiedene Maße und Begriff

Rein methodisch unterscheidet man zwischen **Zuständen** (Bevölkerungsstruktur) und **Ereignissen** (Bevölkerungsdynamik), aus denen Kennziffern abgeleitet werden und die mit verschiedenen Messkonzepten beschrieben werden können.

Auf der einen Seite haben wir also die **Strukturmaße** (Bestandsgrößen, Ist-Daten, Zustandsmaße, Grundgesamtheiten, Prävalenzen, Daten per Stichtag).

Sie beschreiben den **Ist-Zustand** der Bevölkerung **eines definierten Gebietes zu einem bestimmten Zeitpunkt** hinsichtlich der Größe und Gliederung der Bevölkerung (Alter, Geschlecht, Erwerbstätigkeit, Nationalität etc.). Diese werden oft auf den Anfang eines Jahres bezogen oder als Jahresmittel dargestellt, aber im Prinzip ist jeder Zeitpunkt möglich.

Diese Maße können **absolut** (z.B. Größe einer Bevölkerung an einem Stichtag) oder **relativ** sein (s.u.).

Bei den relativen Maßen unterscheidet man zwischen

- **Quoten** (engl. *proportions*) beschreiben das Verhältnis einer Teilmasse zu einer Gesamtmasse an (z.B. Arbeitslosenquote, Anteil Singles)
- **Quotienten** (enlg. *ratios*) beschreiben das Verhältnis zwischen zwei Teilmassen (z.B. der Abhängigkeitsquotient oder das Geschlechterverhältnis)
- **Entsprechungszahlen** (engl. *rates*) beschreiben das Verhältnis zweier verschiedener Massen, die nicht Teilmassen einer Gesamtmasse bilden können (z.B. Einkommen per Einwohner)

Auf der anderen Seite haben wir die **Ereignisgrößen** (Inzidenzen) (engl. *incidences*). Sie beschreiben Ereignisse eines Zeitraumes, meist eines Kalender- oder Lebensjahres.

- **Absolute Ereignismaße** geben die absolute Häufigkeit eines Geschehens einer Bevölkerung innerhalb eines Zeitraums an (z.B. Anzahl der Geburten in Schweden 2020).
- **Ereignisquoten** geben Auskunft über die Verteilung von Ereignissen und setzen sie in Relation zu einander (z.B. Anteil der Uni-Absolventen eines Jahres an der Gesamtheit der Studenten).
- **Ereignisraten** setzen die Anzahl eines Ereignisses innerhalb eines Zeitraums in Relation zur Bevölkerung (z.B. Rohe Sterbeziffer, also die Anzahl der Verstorbenen je 1.000 Bewohner).
- Die **Wahrscheinlichkeit** (engl. *probability*) unterscheidet sich von der Ereignisrate dadurch, dass sie sich auf die Chance (oder das Risiko) eines Ereignisses bezieht (z.B. Überlebenswahrscheinlichkeit).

2.2 Längsschnitt- und Querschnittsanalyse

Es gibt zwei grundlegend verschiedene Konzepte bei der Betrachtung demografischer Ereignisse: die Längsschnitts- und die Periodenanalyse.

Bei der **Längsschnittanalyse** (Kohortenanalyse, Generationsanalyse) werden die Ereignisse innerhalb einer Kohorte analysiert. Im demografischen Sinne besteht eine Kohorte in aller Regel aus Mitgliedern eines Jahrgangs oder einer Gruppe von Jahrgängen (mit einer regionalen Gemeinsamkeit). Die Untersuchung von Kohorten ist immer dann interessant, wenn bereits abgeschlossene Ereignisse analysiert werden sollen. Voraussetzung dafür ist, dass die Daten der Kohorte möglichst vollständig und genau sind. Das wichtigste Kohortenmaß ist die durchschnittliche Kinderzahl pro Frau (abgeschlossene Fruchtbarkeit, endgültige Fertilität, completed fertility rate, CFR, siehe Abschnitt 4.3.3). Dieses Maß ist sinnvollerweise nur für Frauen, die älter als 49 Jahre sind (Ende der fertilen Periode), anwendbar. Aber auch das tatsächlich erreichte Lebensalter

einer Gruppe ist oft Gegenstand von Längsschnittanalysen (sinnvollerweise für Kohorten, deren Geburtsjahr mindestens 100 Jahre zurückliegt).

Die Kohortenanalyse ist besonders in der **Historischen Demografie** wichtig, wenn es beispielsweise Kirchenbücher und damit Geburtenregister gibt, aber keine Bevölkerungszählungen oder -schätzungen. Sie wird dort auch (wie oben erwähnt) oft auf **kleinere Populationen** (z.B. dörfliche Verbände) mit großem Nutzen angewandt.

Bei hoher Vollständigkeit und Genauigkeit des Ausgangsmaterials kann die Längsschnittanalyse die Wirklichkeit korrekt widerspiegeln. Erfüllt das Ausgangsmaterial die beiden Voraussetzungen nicht, kann sie auch in der Historischen Demografie nicht verwendet werden.

Auch können Längsschnittanalysen lediglich bedingt auf aktuelle Lagen bezogen werden. Sie sind also nicht für die Betrachtungen von Perioden (Jahren) geeignet.

Sie sind das Mittel der Wahl, wenn es um die **Analyse von Jahrgängen** geht, also von Kohortenbetrachtungen.

Beispiel: Die 1955 in Schweden geborenen Frauen gebaren im Durchschnitt 2,2 Kinder.

		Ereignisjahre							
		1970-1974	1975-1980	1981-1985	1986-1990	1991-1995	1996-2000	2001-2005	Summe
Kohorte	1954	...	...	...	...	...	...	...	...
	1955	0.03	0.30	0.56	0.78	0.42	0.10	0.01	2.20
	1956	...	...	...	...	...	...	...	...

Tabelle 1: Schematische Darstellung CFR

Bei der **Periodenanalyse** (Querschnittsanalyse) werden die Ereignisse eines bestimmten **Zeitraumes**, meist eines Kalenderjahres, analysiert. Sie stellen also gewissermaßen eine Momentaufnahme dar. Die wichtigsten Periodenmaße sind die **Zusammengefasste Geburtenziffer** oder auch Totale Fertilitätsrate, TFR (siehe Abschnitt 4.3.2) und die (verbleibende) **Lebenserwartung** bezogen auf ein Kalenderjahr (siehe Abschnitt 4.2.3), aber auch andere Ereignisse wie Eheschließungen und Scheidungen können Gegenstand von Periodenanalysen sein. Natürlich kann man die Betrachtung von Geburten, Todesfällen usw. auch anhand von Kohortenanalysen durchführen, da solche Untersuchungen aber meistens den Zweck haben, die aktuelle Lage zu beschreiben und nicht die Entwicklung gewisser Jahrgänge, kommen hier überwiegend Periodenanalysen zur Anwendung.

Bei diesen Berechnungen werden die **Ereignisse aller relevanten Altersklassen zusammengefasst**, so dass sich fiktive Maße mit einem hohen **Prognosewert** ergeben. Klingt komplizierter als es ist. Bei der TFR beispielsweise werden die Geburtenraten aller Altersklassen zwischen 15 und 49 Jahren zusammengefasst. Man erhält also eine „**fiktive Frau**“², die im Jahre der Messung ihre gesamte fertile Periode durchlebt. Anders ausgedrückt entspricht der TFR der Fruchtbarkeit der Frauen der untersuchten Gruppen, **wenn sich diese ihr Leben lang genauso verhalten würden, wie es die Frauen des untersuchten Jahres im entsprechenden Alter getan haben**.³

Das ist in der Realität nicht immer der Fall, gerade wenn durch außergewöhnliche Umstände Ereignisse verschoben oder vorgezogen werden (ohne dass sich ihre letztendliche Größe verändern muss).

Beispiel: Kurz nach dem **Zusammenbruch der DDR** sank die TFR in den neuen Ländern, also auf dem Gebiet der ehemaligen DDR, auf die **tiefsten, jemals gemessenen Werte**, weil viele Frauen sich dazu entschlossen, Geburten aufzuschieben. In welchem Ausmaß die Geburten **nachgeholt** wurden, kann erst eine abschließende Kohortenanalyse im Jahre **2025** zeigen, wenn alle Frauen, die im Jahre 1990 im fertilen Alter waren, älter als 49 Jahre sein werden.

² Zum Thema „männliche TFR“ siehe Seite 56.

³ ...und wenn keine Frau vor Vollendung des fertilen Alters sterben würde.

Periodenmaße haben so immer eine Tendenz zu **stärkeren Ausschlägen** als Kohortenmaße (auf Dauer nähern sie sich jedoch immer den „endgültigen Zahlen“ an). Die Vorteile der Periodenmaße sind in erster Linie ihre zeitnahe Anwendbarkeit und ihr **Prognosewert**.⁴ Andererseits sind es eher fiktive Werte, die vielmehr Trends als endgültige Zahlen liefern.

Beispiel: 1975 lag der TFR in Schweden bei 2,00

		Ereignisjahr				
		1974	1975	1976	1977	1980
Kohorten	1926-1930	...	0,01	...	...	...
	1931-1935	...	0,08	...	...	...
	1936-1940	...	0,39	...	...	...
	1941-1945	...	0,77	...	...	...
	1946-1950	...	0,46	...	...	...
	1951-1955	...	0,27	...	...	...
	1956-1960	...	0,02	...	...	...
	Summe		2,00			

Tabelle 2: Schematische Darstellung TFR

⁴ Natürlich können auch Längsschnittsanalysen für Prognosen herangezogen werden. Diese jedoch eher indirekt, da sie methodisch auf das Beschreiben bereits Geschehens ausgerichtet sind, während Periodenmaßen von vorneherein einen prognostischen und damit fiktiven Ansatz haben (siehe auch die Konjunktive in der Definition vor der Tabelle).

2.3 Personenjahre im Lexisdiagramm

Die Grundlage vieler Raten bildet eine wie auch immer berechnete Bevölkerung. Aber eigentlich reden wir hier von Personenjahren, die dann in Relation zu einem Risiko gestellt werden. Dabei erfolgt die Berechnung oder Abschätzung der Bevölkerung nach sehr unterschiedlichen Methoden. Hier einige Beispiele.

- In der Stadt A lebten an jedem Tag im Schaltjahr 2020 genau 36.600 Menschen. Das ergibt 36.600 Personenjahre oder kurz: eine Bevölkerung von (durchschnittlich) 36.600.
- Am 1. Januar 2020 lebten im Ort B nur 100 Personen. Zur Freude der zuständigen Statistiker wuchs diese Zahl jeden Tag um ganz genau 100 Personen. Am Ende des Jahres waren es 36.600 Einwohner. Im Schnitt hatte jede Person genau ein halbes Jahr im Ort gelebt. Das ergibt 18.300 Personenjahre oder durchschnittlich 18.300 Einwohner.
- Im Nachbarort C lebten am 1. Juli 2020 genau 36.600 Menschen. Genauere Zahlen liegen nicht vor, aber da die Bevölkerung dort erfahrungsgemäß nicht stark schwankt, können wir von einer Durchschnittsbevölkerung von 36.600 (und damit 36.600 Personenjahren) ausgehen.
- Im Jahr 2020 ist in C an jedem Tag genau ein Kind geboren worden. Glücklicherweise ist – wie auch im Jahr darauf – kein Kind gestorben, es gab auch keine Wegzüge. Am 31. Dezember 2020 gab es dort 366 im Jahre 2020 Geborene. Diese hatten im Durchschnitt ein halbes Jahr im Ort gelebt. Das ergibt für 2020 183 Personenjahre oder Einwohner des Jahrgangs 2020.
- Am 31. Dezember des Jahres 2021 wird das letzte Kind des Jahrgangs 2020 ein Jahr alt. Die Unter-ein-jährigen des Jahrgangs 2020 haben dann bis zum Ende des Jahres 2021 (fast) genau 366 Personenjahre gelebt.⁵ Die Hälfte davon im Jahr 2020. Sie tragen damit im Jahr 2020 genau wie im Jahr 2021 mit je 183 zur Einwohnerzahl der Unter-1-jährigen bei.
- Wie hoch ist – wenn wir von den Kohortenstatistiken ausgehen – die Anzahl der Unter-1-jährigen im Jahre 2021 gewesen? Dankens-

⁵. Wir ignorieren, dass 2021 kein Schaltjahr ist

werter Weise war der Geburtenverlauf des Jahres 2021 genauso linear wie der des Vorjahres (bei gleichen Rahmenbedingungen). Die 2020 Geborenen tragen ja nur mit einem halben Jahr bei, weil sie im Laufe des Jahres ein Jahr alt wurden und damit die Gruppe verlassen haben. Genau wie die 2020 Geborenen tragen die Geborenen des Jahres 2021 mit 183 Personenjahren zur Einwohnerzahl bei. Das ergibt dann in der Summe 366 Personenjahre der Unter-ein-jährigen oder eine Durchschnittliche Größe von 366 Bewohnern unter einem Jahr.

Das Lexisdiagramm verdeutlicht die Tatsache, dass es bei der Erfassung von Ereignissen **keine unmittelbare Kongruenz zwischen** den beiden Zeitdimensionen **Kalenderzeit** und **Alter** gibt. Auf einem Zeitstrahl wird ein Träger des untersuchten Ereignisses abgebildet, dessen Bewegung in Relation zur Kalenderzeit t (Abszisse) und der Individualzeit x (Ordinate) steht. Diese „Trägerschaft“ kann das Leben einer Person oder mehrerer Personen sein.

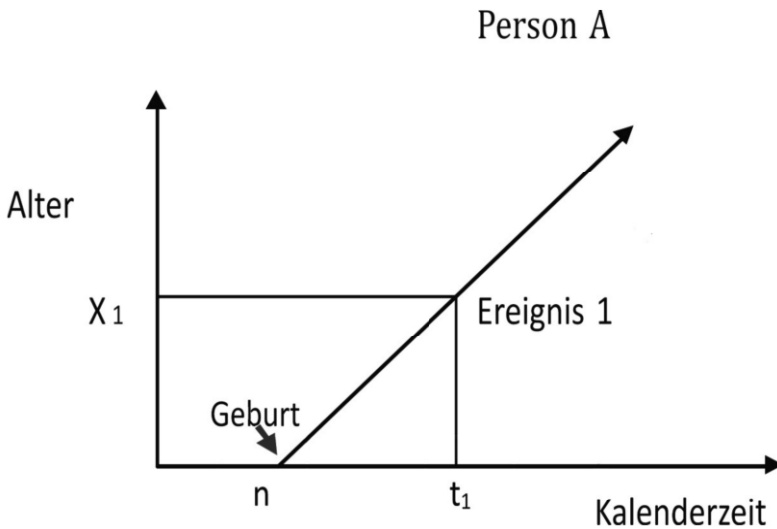


Abbildung 2: Ereignis in der Zeit

Dementsprechend ergeben beispielsweise die „**Lebenslinien**“ aller Individuen mit einem gemeinsamen Startereignis einen Schrägstreifen (links: Anfang des Geburtsjahres, rechts das Ende). Hier können z.B. alle in einem Jahr geborener Personen, also eine Kohorte, Gegenstand der Darstellung sein (Längsschnitt).

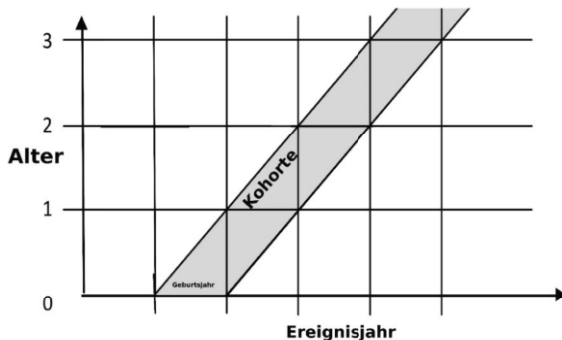


Abbildung 3: Längsschnittanalyse im Lexisdiagramm

Die Ereignisse eines Zeitabschnittes (z.B. eines Jahres) wie in der Abbildung 4 werden hingegen als eine senkrechte Fläche wahrgenommen (links: Anfang des Kalenderjahres, rechts das Ende).

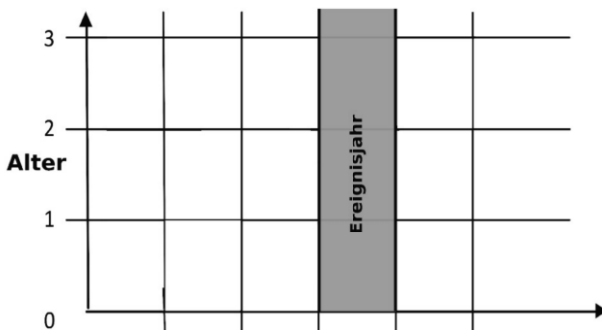


Abbildung 4: Querschnittsanalyse im Lexisdiagramm

Stellen wir uns die Frage, wie viele 40-jährige im Jahre 2020 einem Risiko (z.B. Tod oder Hochzeit) ausgesetzt sind.

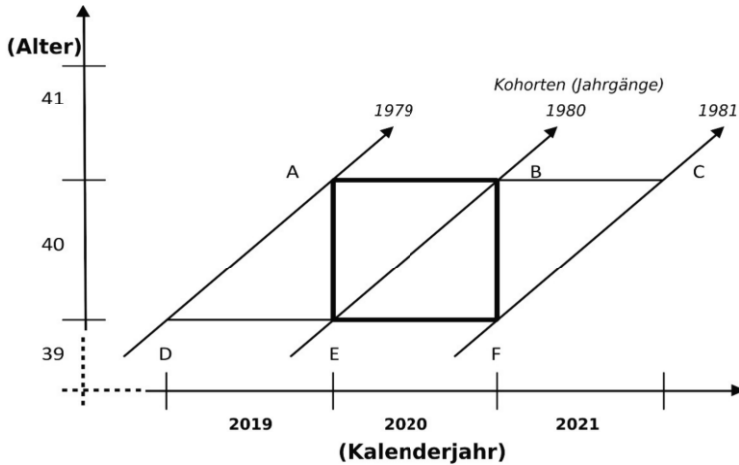


Abbildung 5: Kohorte und Kalenderjahr im Lexisdiagramm

Im Diagramm 5 sehen wir, dass die **Risikogruppe** aus **zwei Jahrgängen** besteht. In diesem wiederum realisiert sich das Risiko innerhalb zweier Jahre ($\square DABE$ und $\square EBCF$).

Bei der **Längsschnittanalyse** verteilt sich das Risiko auf $\square EBCF$. Das sind die zwölf Monate, in denen der Jahrgang 1980 das 40. Lebensjahr vollendet. Dieser Bereich liegt in den Jahren 2020 und 2021. Sie beschreibt also die **Entwicklung in der Kohorte** korrekt.

Hier haben wir bei der Beurteilung des Risikos eine gute Aussagekraft hinsichtlich der Kohorte, aber nicht in Bezug auf die Periode. Bei sehr gleichmäßigen Verläufen spielt das keine Rolle, da $\triangle ABE \approx \triangle BCF$.

Bei der **Querschnittsanalyse** würde der Bereich durch $\square ABFE$ dargestellt werden. Hier betrifft der Risikobereich **zwei Jahrgänge**: 1979 und 1980. Es wird also die Situation in der betreffenden Periode (hier Kalenderjahr) korrekt abgebildet.

Hier haben wir umgekehrt zu oben eine gute Beschreibung der Periode. Und hier gilt ebenso, dass bei gleichmäßigen Verläufen die Unkorrektheiten gering sind, da $\triangle DAE \approx \triangle ABE$.

2.4 Projektionen und Prognosen

Eine weitere ewige Baustelle innerhalb der in Deutschland oft mehr emotional als rational geführten Diskussionen über demografische Fragen ist die Unterscheidung von **Bevölkerungsprojektionen** und **Bevölkerungsprognosen**. Da ja bekanntlich Prognosen immer dann recht problematisch sind, wenn sie die Zukunft betreffen, wird sich kein seriöser Demograph auf Prognosen für längere Zeiträume und /oder größere Gebiete einlassen, auch wenn in den o.g. Diskussionen häufig (und mitunter wider besseren Wissens) das Gegenteil behauptet wird.

Eine **Prognose** geht von einer definierten **Ausgangslage** und einem angenommenen Entwicklungsverlauf aus und gibt mit einer (mitunter inkludiert) formulierten **Eintrittswahrscheinlichkeit** zwischen null und hundert Prozent ein daraus resultierendes **Ereignis**⁶ an. Sie liefert immer ein formal-logisch betrachtet **sinnvolles Ergebnis**.

Eine **Projektion** schreibt nur einen **Trend** von einer **vollständig definierten Ausgangs- und Entwicklungslage** betrachtet fort. Sie ist somit eine modellhafte Berechnung. Sie kann **mehrere** in Konkurrenz stehende Szenarien entwerfen. Das **Resultat muss nicht einmal zwingend sinnvoll** sein. Im Gegenteil: sie kann z.B. darlegen, dass eine bestimmte Entwicklung so nicht über einen gewissen Zeitpunkt hinweg fortlaufen kann. Als reine Fortschreibung kann sie natürlich **keine Wahrscheinlichkeiten** hantieren.

Beispiel: Sie haben vor einem Jahr 84 kg gewogen, haben dann Ihre Ernährungsgewohnheiten dauerhaft umgestellt und wiegen nun 72 kg.

- a) **Prognose:** wenn Sie so weiter machen und nichts Unerwartetes passiert, wird sich Ihr Gewicht in zwei Jahren mit hoher Wahrscheinlichkeit bei 60 kg einpendeln — ergo: weitermachen!
- b) **Projektion:** Wenn Sie weiterhin ein Kilogramm pro Monat abnehmen, werden Sie in sechs Jahren nichts mehr wiegen — ergo: trotz gleicher Ernährung wird der Gewichtsverlust abgebremst werden und zum Erliegen kommen müssen.

⁶. Es können auch mehrere Alternativen sein, doch darf da die Gesamtwahrscheinlichkeit nicht über 100 Prozent liegen

So veröffentlicht das deutsche Statistische Bundesamt jährlich seine **Bevölkerungsvorausberechnungen mit zwölf unterschiedlichen Szenarien**. (Statistisches Bundesamt 2009) Jedoch wird dort niemals behauptet, dass eines dieser Szenarien (oder gar alle gleichzeitig) Realität werden wird; es wird also **bewusst auf eine Prognose verzichtet!**

Insofern läuft die häufig geäußerte und an sich nicht unbegründete Kritik, eine Prognose über fünf Jahrzehnte sei praktisch unmöglich, man schaue sich nur ten fünf Jahr- d e u t s c h e r rungsentwick- ins Leere. Eine tion wird erst ter in den Hän-

Prognosen sind schwierig,
besonders wenn sie die
Zukunft betreffen.

die letz- z e h n t e Bevölke- lung an, Projek- (mitun- den von

Journalisten und Politikern) zur Prognose, wenn behauptet wird, dieser oder jener Fall werde aller Wahrscheinlichkeit nach (oder mit einer bestimmten, sehr hohen Wahrscheinlichkeit) eintreffen. Im Beispiel der Bevölkerungsvorausberechnungen hieße das dann logischer Weise, eines der Szenarien, die anderen elf nicht. Und über genau diese Wahrscheinlichkeit macht wie angeführt die Projektion keine Aussage, das unterscheidet sie von der Prognose. Vielmehr verdeutlicht die Projektion, welche Trends welche Folgen nach sich ziehen können, wenn sich wichtige Parameter in dieser (z.B. Beibehaltung der aktuellen Fertilität, Lebenserwartung oder Wanderungssaldo) oder jener (Abnahme oder Zunahme der genannten Faktoren) Weise entwickeln würden.

Projektionen sind also, wenn man so will **Planspiele – Prognosen** dagegen „**relative Vorhersagen**“.

Warum gibt es dann überhaupt Projektionen und Prognosen als Planungsgrundlage, wo doch jeder altkluge Gymnasiast sehen kann, dass die Wirklichkeit sich nicht selten anders entwickelt, als sie sich laut Planung entwickeln könnte? Selbst kurzfristige „Voraussagen“ treffen selten hundertprozentig ein wie formuliert. Ganz einfach: weil sie bei aller Unvollkommenheit immer noch besser sind als nichts, weil sie sich in den Händen derer, die sie hantieren können, ihren Charakter verstehen,

bewährt haben. Außerdem wissen alle, die schon mal irgendwo ein Proseminar an einer Hochschule oder Universität erfolgreich beendet haben, dass in der Welt der Wissenschaft so gut wie alle **Erkenntnisse** mehr oder weniger **vorläufig** sind, ohne dadurch sinnlos zu sein. Das unterscheidet die **Wissenschaft** von **Ideologien** und **Religionen**, welche von unabänderlichen Wahrheiten ausgehen. Trotzdem liefern die Wissenschaften in den meisten Lebensbereichen die besseren Werkzeuge und Ergebnisse. Man kann das mit einer Rathausuhr vergleichen. In einer eher nach Objektivität bemühten Gemeinschaft wird man versuchen, die Zeit immer wieder neu einzustellen und die Mechanismen so zu verfeinern, dass die Ungenauigkeiten immer geringer werden, wohlwissend, dass die Uhr niemals „richtig gehen“ wird, man sich der Realzeit nur nähern kann. In der eher dogmatischen Nachbargemeinde gibt man sich nicht die Mühe und lässt die Uhr irgendwann ganz stehen – und oh Wunder, nun geht sie zumindest zweimal am Tag genau richtig. Womit die Überlegenheit des dogmatischen Systems bewiesen wäre...

2.5 Die Bevölkerungspyramide

Ein sehr wichtiges, wenn auch nicht völlig unumstrittenes Symbol für jede demografische Analyse, ist die **Bevölkerungspyramide** (engl. *population pyramid*).

Hier wird nach Geschlechtern getrennt die altersgemäße Verteilung der Bevölkerung dargestellt. Auf der X-Achse wird entweder der pro-zentuale oder totale Anteil angegeben, auf der Y-Achse das Alter, meist in Ein- oder Fünf-JahresSchritten.



Der Begriff Pyramide leitet sich aus der Frühzeit der Demografie her, also dem europäischen 19. Jahrhundert, wo, u.a. bedingt durch eine hohe Kindersterblichkeit, die Form des Balkendiagramms dem einer ägyptischen Pyramide glich. Später wurde die Form – nicht zuletzt von nationalsozialistischen Bevölkerungswissenschaftlern wie Burgdörfer (Burgdörfer 1932) - ideologisiert. Die Dreiecksform der ägyptischen Pyra-