

Kapitel 2

Beschreibende Statistik

Die Aufgabe der beschreibenden oder deskriptiven Statistik besteht darin, empirisch gewonnene Daten

- zu **ordnen**;
- tabellarisch und grafisch **darzustellen**;
- durch Kennzahlen wie z. B. Mittelwert, Gesamtsumme, Minimum, Maximum oder den Streubereich **zusammenzufassen**.

Die beschreibende Statistik spielt in vielen Lebensbereichen eine wichtige Rolle, insbesondere in der Wirtschaft und in der Verwaltung, wo statistische Erhebungen Grundlage für Planungen und Entscheidungen sind.

Der überwiegende Teil des medizinischen Wissens gründet sich nicht auf theoretische Überlegungen oder Experimente, sondern auf Erfahrungen am kranken Menschen. Dieser in Jahrhunderten gewachsene ärztliche Erfahrungsschatz lässt sich als statistische Zusammenstellung betrachten über das Auftreten der verschiedenen Erkrankungen, die Bedeutung der Symptome, die Wirksamkeit der Therapien, die Häufigkeit von Komplikationen usw.

Die ärztliche Erfahrung

Jeder Arzt, der durch die Behandlung seiner Patienten ärztliche Erfahrung erwirbt, indem er z. B. auf die Wirkung und Nebenwirkungen seiner Therapie achtet, betreibt Statistik. Dies geschieht zwar nicht in formalisierter Form, sondern in intuitiver Art und Weise: Als Ergebnis der gesammelten Erfahrungen bildet sich nach einiger Zeit eine Meinung über seine therapeutischen Maßnahmen heraus.

Diese Meinung beruht auf dem in der Vergangenheit behandelten Patientenkollektiv. Dieses Patientenkollektiv ist die *Stichprobe*, anhand derer das medizinische Wissen gewonnen wurde. Die Übertragung dieses Wissens auf die *Grundgesamtheit aller Patienten mit der Krankheit XYZ* ist nur mit Einschränkung möglich.

Ebenso unsicher ist die Übertragung dieser Daten auf einen konkreten Einzelfall, etwa einen zukünftig zu behandelnden Patienten. Der Arzt spricht von der *Individualität des einzelnen Patienten*, der Statistiker von *zufallsbedingter Streuung, von Häufigkeiten und Wahrscheinlichkeiten*.

2.1 Begriffsbestimmungen

Merkmal

Eine statistische Untersuchung ist nicht in der Lage, die untersuchten Personen oder Objekte in ihrer „Gesamtheit“ zu erfassen, sondern beschränkt sich auf einzelne Merkmale wie beispielsweise die Symptome und Laborwerte eines Patienten.

Beobachtungseinheit – Merkmalsträger

Das Objekt der Untersuchung wird als Beobachtungseinheit oder Merkmalsträger bezeichnet, wobei die Untersuchung darin besteht, bestimmte Merkmale dieser Beobachtungseinheit zu erfassen.

Merkmalsausprägungen

Die Werte, die ein Merkmal annehmen kann, heißen Ausprägungen des Merkmals. Zum Beispiel werden die Therapie und der Heilungsverlauf bei einer bestimmten Krankheit untersucht.

Beobachtungseinheit bzw. **Merkmalsträger** ist jeder Patient, der wegen einer bestimmten Krankheit behandelt wird. Die beobachteten **Merkmale** des Patienten sind die bei ihm angewendete Therapie, sein Heilungsverlauf und weitere in diesem Zusammenhang wichtige Daten wie z. B. Blutuntersuchungen.

Unter **Merkmalsausprägungen** versteht man die beobachteten Werte wie z. B. die Art der Operation, Name und Dosierung der Medikamente, die Höhe der Laborwerte usw. Was umgangssprachlich als Merkmal bezeichnet wird, heißt in der Sprache der Statistik immer dann Merkmalsausprägung, wenn es um die einzelnen Werte geht, die dieses Merkmal annehmen kann. Beispielsweise ist „175 cm“ eine Merkmalsausprägung des Merkmals „Körpergröße“.

qualitativ – quantitativ

Wir können quantitative und qualitative Merkmale unterscheiden: **Qualitative Merkmale** sind Merkmale, die sich nicht zahlenmäßig erfassen, sondern **nur benennen** lassen: Haarfarbe, Blutgruppe, Medikamente A, B, C, Religionszugehörigkeit. Bei diesen Merkmalen lässt sich die Merkmalsausprägung nur qualitativ beschreiben, aber nicht wiegen, messen, abzählen oder auf andere Weise quantifizieren. Ein Beispiel für ein qualitatives Merkmal ist der Name des verabreichten Medikaments.

Quantitative Merkmale (quantum = Menge) sind **zahlenmäßig erfassbar**. Ein Beispiel für ein quantitatives Merkmal ist die Menge des verabreichten Medikaments.

diskret – stetig

Quantitative Merkmale lassen sich weiter unterteilen in diskrete und stetige Merkmale. **Diskrete Merkmale können nur bestimmte Werte annehmen.** Beispiele sind die Zahl der Kinder in einer Familie oder die Zahl der im letzten Jahr zugelassenen Personenwagen. Zwischenwerte, etwa 2,71, sind im konkreten Einzelfall nicht möglich, sondern nur als Mittelwert. Ergebnisse von Zählvorgängen führen stets zu diskreten Werten.

Messungen und Wägungen hingegen führen in der Regel zu vielstelligen, sog. krummen Werten. Es handelt sich hierbei um **stetige Merkmale, die in einem bestimmten Bereich jeden Zwischenwert annehmen können.** Das Gewicht eines Menschen beispielsweise kann beliebig genau angegeben werden; die Anzahl der Stellen hängt nur von der Genauigkeit der Waage ab.

Klassenbildung

Weil eine zu große Genauigkeit belanglos ist, wird das Gewicht in der Regel in vollen kg-Beträgen angegeben. Hierbei wird das im Prinzip stetige Merkmal Gewicht **klassiert**, d.h. in Klassen eingeteilt. Die Klasse 71 kg beispielsweise umfasst alle Werte zwischen 70,500 kg und 71,499 kg. Die Klassenbreite ist willkürlich und richtet sich nach dem Zweck der Untersuchung. Klassenbreiten von jeweils 10 kg sind ebenso möglich wie eine ganz grobe Klassierung: 0–49 kg, 50–99 kg und 100 kg und mehr. Eine Vergrößerung der Klassenbreite bedeutet eine Zusammenfassung und damit eine Vereinfachung des Datenmaterials, wodurch die weitere Auswertung rechentechnisch erleichtert wird. Allerdings geht bei einer Vergrößerung der Klassenbreite Information verloren, die in der weiteren Auswertung nicht mehr berücksichtigt werden kann.

Neben stetigen Merkmalen können auch diskrete und qualitative Merkmale in Klassen zusammengefasst werden, wenn dies für die Untersuchung sinnvoll ist. Zusammenfassend ergibt sich das umseitige Diagramm (Abbildung 2.1).

Skalenniveau

Üblicherweise werden Merkmalsausprägungen als Zahlenwerte angegeben, die Temperatur z.B. als 10 °C. Es wird eine Messung, Zählung oder Verschlüsselung vorgenommen, bei der einer bestimmten Merkmalsausprägung ein Wert auf einer Zahlenskala zugeordnet wird. Beispielsweise kann das qualitative Merkmal Geschlecht folgendermaßen verschlüsselt werden:

weiblich = 1
männlich = 2
Geschlecht unbekannt = 3.

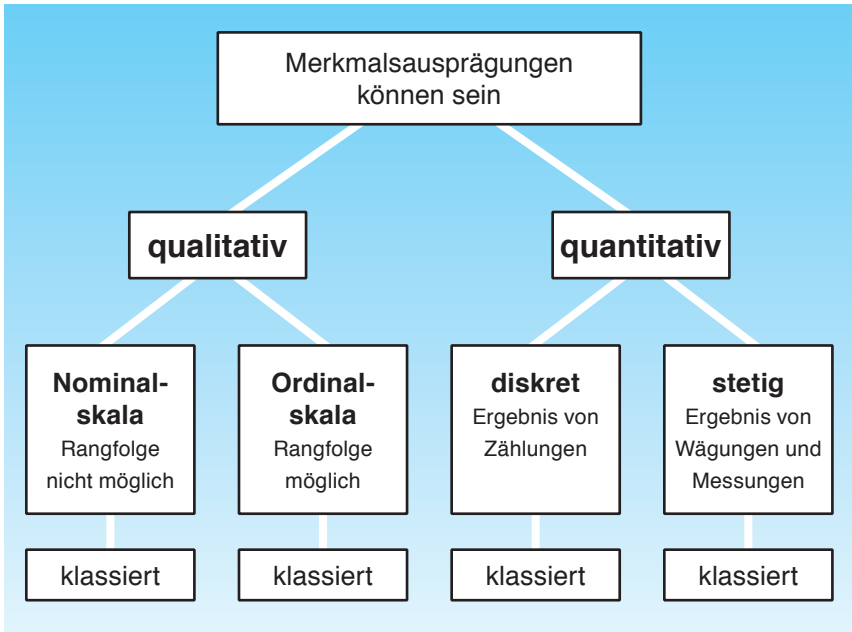


Abbildung 2.1: Darstellung der verschiedenen Skalenniveaus, in denen Merkmalsausprägungen vorliegen können. Diskrete und stetige Merkmale können sowohl auf dem Niveau einer Intervall- als auch auf dem Niveau einer Verhältnisskala vorliegen (siehe unten).

Es ist unsinnig, mit diesen willkürlich zugeordneten Zahlenwerten Rechenoperationen durchzuführen, etwa den Mittelwert zu errechnen. Welche Rechenoperationen sinnvoll sind, hängt vom sog. Skalenniveau der Zahlenwerte ab.

Das geringste Skalenniveau hat die **Nominalskala**, die durch **willkürliche Verschlüsselung eines qualitativen Merkmals** entsteht.

Eine **Rang- oder Ordinalskala** liegt vor, wenn **zwar die Rangfolge, nicht aber die Differenz** zwischen den Merkmalsausprägungen definiert ist. Ein gutes Beispiel ist der Karnofsky-Index, der in der Krebstherapie als Indikator für den Allgemeinzustand des Patienten dient. Hier bedeuten 100 % völlige Beschwerdefreiheit, 50 % Bettlägerigkeit während der Hälfte des Tages und 0 % Tod. Offensichtlich ist die Differenz zwischen 0 % und 50 % nicht mit der Differenz zwischen 50 % und 100 % identisch.

Eine **Intervallskala** ergibt sich, wenn die Differenzen so definiert sind, dass **gleiche Differenzen einem gleichen Merkmalsunterschied** entsprechen, aber **kein absoluter Nullpunkt vorliegt**. Ein Beispiel hierfür ist Grad Celsius: Der Unterschied zwischen 0 °C und 10 °C ist genau gleich der Differenz zwischen 10 °C und 20 °C, aber 20 °C ist nicht doppelt so warm wie 10 °C.

Eine **Verhältnisskala liegt bei definierten Differenzen und absolutem Nullpunkt** vor. Beispiele sind die Kelvinskala, die Anzahl, Längenmessungen oder das Ergebnis von Gewichtsbestimmungen.

Die Tabelle 2.1 zeigt, welche Berechnungen bei welchem Skalenniveau formal erlaubt sind:

Lage- und Streuungsmaße	Skalenniveau			
	Nominalskala z.B. Diagnose	Ordinalskala z.B. Karnofsky	Intervallskala z.B. °Celsius	Verhältnisskala z.B. Kelvin
Häufigkeit (s.S. 29)	+	+	+	+
Modalwert (s.S. 34)	+	+	+	+
Summenhäufigkeit (s.S. 29)	-	+	+	+
Quartilsabstand (s.S. 36)	-	+	+	+
Median (s.S. 35)	-	+	+	+
arith. Mittelwert (s.S. 34)	-	-	+	+
Standardabweichung (s.S. 38)	-	-	+	+
Variationskoeffizient (s.S. 39)	-	-	-	+

Tabelle 2.1: Die unterschiedliche Aussagekraft verschiedener Skalenniveaus. Die Tabelle gibt an, welches Skalenniveau Voraussetzung für die Berechnung der gängigen Lage- und Streuungsmaße ist.

2.2 Grafische Darstellung

Eine Datenanalyse besteht aus der tabellarischen und grafischen Darstellung des erhobenen Datenmaterials sowie der Zusammenfassung der vorhandenen Information durch wenige besonders charakteristische Kennzahlen wie Mittelwert, Varianz oder Standardabweichung.

Am Anfang einer Datenanalyse steht stets die genaue Eingrenzung und Defi-

nition des erhobenen Datenmaterials. Im folgenden Beispiel wollen wir untersuchen, am wievielten Tag nach der Schnittentbindung die Patientinnen im Krankenhaus XYZ entlassen wurden. Hierbei muss zunächst geklärt werden,

- auf welchen Zeitraum sich die Untersuchung beziehen soll;
- ob alle Patientinnen, die in dieser Zeit mit Kaiserschnitt entbunden wurden, in die Untersuchung einbezogen werden sollen oder nur eine Stichprobe. Falls nur eine Stichprobe analysiert werden soll, muss festgelegt werden, wie diese Stichprobe gezogen wird (Vielleicht haben Privatpatientinnen eine längere Liegedauer als Kassenpatientinnen?);
- wie Frauen gezählt werden sollen, die auf eigenen Wunsch entlassen oder die wegen Komplikationen in eine andere Klinik verlegt wurden.

In unserem Beispiel wollen wir uns auf alle Patientinnen beziehen, deren Schnittentbindung zwischen dem 1.1.2011 und dem 30.6.2011 stattgefunden hat. Die letzte Frage erübrigt sich, weil von diesen Frauen niemand auf eigenen Wunsch vorzeitig entlassen oder wegen Komplikationen in eine andere Klinik verlegt wurde.

In einer sog. Urliste sind alle Patientinnen (durch ihre Initialen abgekürzt) mit ihrer Aufenthaltsdauer aufgeführt:

M.K. 10;05	P.J. 14;02	T.E. 13;01	K.L. 11;04	L.K. 14;22
A.B. 9;13	ET. 11;23	R.H. 20;18	K.S. 11;09	L.W. 11;08
U.N. 14;01	V.H. 11;12	A.U. 28;00	E.R. 12;08	S.N. 12;06
B.E. 14;22	Z.N. 10;04	A.L. 10;17	J.W. 12;04	N.A. 13;23
P.A. 15;08	D.E. 17;19	H.T. 13;19	J.K. 8;20	U.C. 12;12
E.N. 12;08	L.B. 13;06	B.E. 15;20	N.N. 14;03	R.S. 11;16

Tabelle 2.2: Urliste aller 30 Patientinnen, die zwischen dem 1.1.2011 und dem 30.6.2011 durch Sectio entbunden wurden. Der Name ist durch die Initialen abgekürzt, die Liegedauer nach der Schnittentbindung ist in Tagen und Stunden angegeben, die durch ein Semikolon getrennt wurden. 10;05 bedeutet demnach 10 Tage und 5 Stunden.

Die Zeitdauer zwischen Kaiserschnitt und Entlassung ist eine stetige Größe, die auf die Minute genau angegeben werden könnte. Weil eine solche Genauigkeit belanglos ist, wurde diese Zeitdauer in vollen Tagen und vollen Stunden angegeben, die zwischen Sectio und Entlassung in der Klinik verbracht wurden.

Zur weiteren Auswertung sortieren wir die Urliste um, indem wir die Beobachtungseinheiten vom kleinsten zum größten Messwert in aufsteigender Rei-

henfolge notieren und als sog. **Stamm- und Blattdiagramm** anordnen. Unter dem Stamm werden die vollen Tage und unter dem Blatt die Stunden verstanden. Diese erst vor wenigen Jahren von dem amerikanischen Statistiker **Tukey** entwickelte Form der Tabelle vermittelt auch grafisch einen Eindruck von der Verteilung der Werte:

Stamm	Blatt	absolute Häufigkeit	relative Häufigkeit	Summen- häufigkeit
Tage	Stunden	$H(x)$	$f(x) = H(x)/30$	$F(x) = \sum f(x)$
8	20	1	0,033	0,033
9	13	1	0,033	0,067
10	04, 05, 17	3	0,100	0,167
11	04, 08, 09, 12, 16, 23	6	0,200	0,367
12	04, 06, 08, 08, 12	5	0,167	0,533
13	01, 06, 19, 23	4	0,133	0,667
14	01, 02, 03, 22, 22	5	0,167	0,833
15	08, 20	2	0,067	0,900
16		0	0,000	0,900
17	19	1	0,033	0,933
18		0	0,000	0,933
19		0	0,000	0,933
20	18	1	0,033	0,966
21		0	0,000	0,966
22		0	0,000	0,966
23		0	0,000	0,966
24		0	0,000	0,966
25		0	0,000	0,966
26		0	0,000	0,966
27		0	0,000	0,966
28	00	1	0,033	1,000

Tabelle 2.3: Stamm- und Blattdiagramm der Liegezeiten nach Kaiserschnitt. Aus der zweiten Spalte kann man einen optischen Eindruck von der Verteilung erhalten. Deshalb heißt diese Tabelle auch Diagramm. Die Spalten 3, 4 und 5 gehören nicht notwendigerweise zum Stamm- und Blattdiagramm, vervollständigen jedoch den Informationsgehalt. Die absolute und relative Häufigkeit $H(x)$ und $f(x)$ gibt die Anzahl bzw. den Anteil der Frauen an, die nach x Tagen entlassen wurden. Die Summenhäufigkeit $F(x)$ gibt den Anteil der Frauen an, die bis zum Tage x einschließlich entlassen wurden. (Näheres s. S. 40 ff.)

Stabdiagramm und Histogramm

Das **Stabdiagramm** dient zur grafischen Darstellung einer Häufigkeitsverteilung. Hierbei entspricht die Länge der Stäbe der Häufigkeit und die Position der Stäbe auf der x -Achse dem beobachteten x -Wert. In Abhängigkeit von der zeichnerischen Darstellung spricht man auch vom **Balken- oder Säulendiagramm** bzw. vom **Histogramm**, wobei sich beim Histogramm benachbarte Balken unmittelbar berühren. Wenn man die Enden der Balken miteinander verbindet, erhält man einen **Polygonzug**.

Auf der nächsten Seite sind die Entlassungshäufigkeiten $f(x)$ am Tage x als Stabdiagramm und als Histogramm dargestellt. Hierbei ist auf der x -Achse die Verweildauer in Tagen und auf der y -Achse die absolute Häufigkeit $H(x)$ und die relative Häufigkeit $f(h) = H(x)/30$ der Entlassung am Tag x dargestellt. Die grafische Darstellung berücksichtigt nur die vollen Tage, die Stunden bleiben unberücksichtigt.

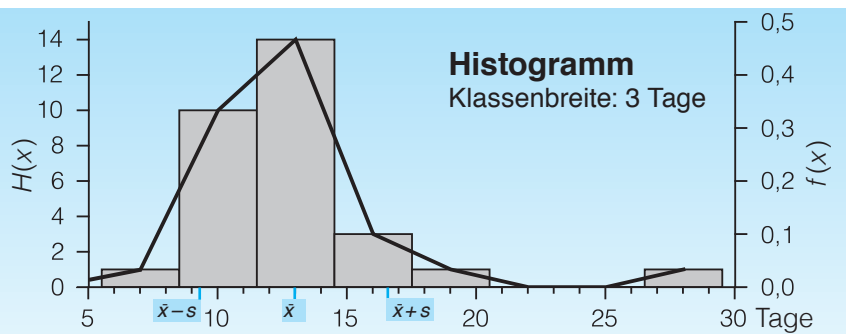
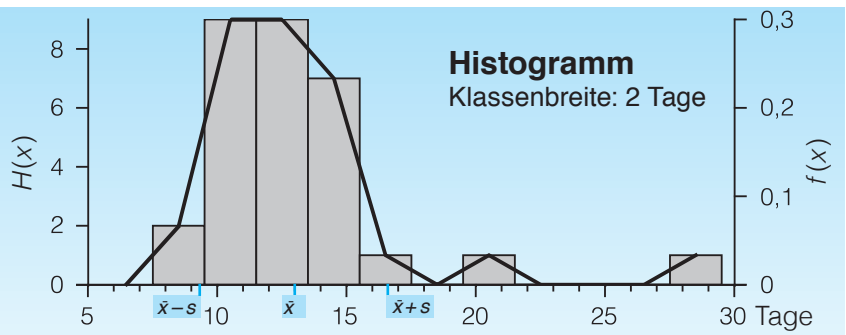
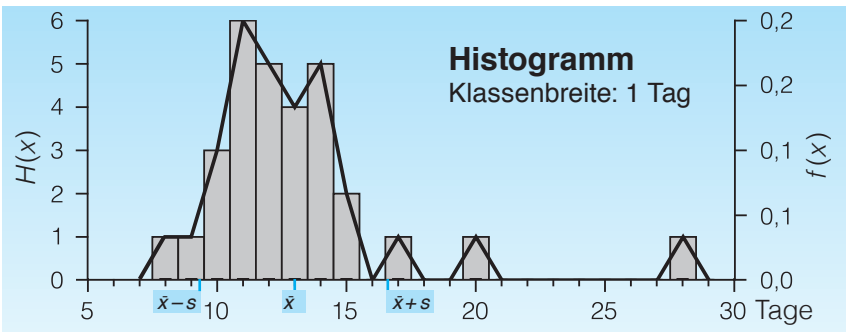
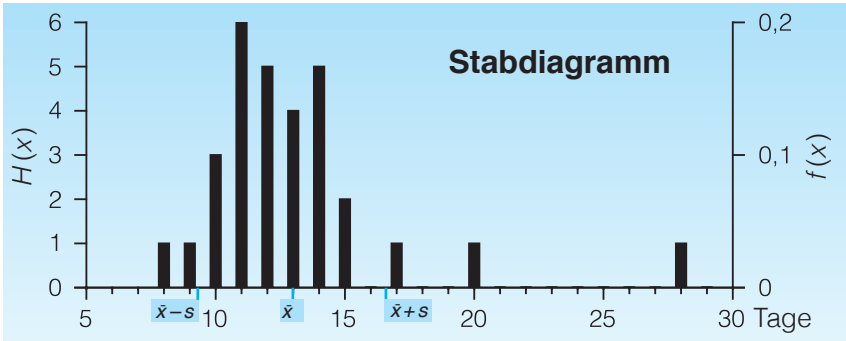
Die grafische Darstellung lässt sich weiter vereinfachen, indem die Verweildauer nicht tageweise, sondern zwei- oder dreitageweise klassiert wird. Hierbei geht Information verloren, und die Diagramme erhalten ein anderes Aussehen. Während bei eintägiger Klassierung der 11. Tag am häufigsten auftritt, weist bei dreitägiger Klassierung das Intervall zwischen dem 12. und 14. Tag die größte Häufigkeit auf.

Man beachte, wie sich bei Veränderung der Klassenbreite der Maßstab der $f(x)$ -Achse ändert. Pro Flächeneinheit wird stets dieselbe Zahl von Patienten dargestellt.

Das Stamm- und Blattdiagramm aus Tabelle 2.3 und das Stabdiagramm bzw. Histogramm in Abbildung 2.2 entsprechen sich, lediglich die Achsen sind vertauscht. Anhand der grafischen Darstellung kann man beurteilen, welche Form die Häufigkeitsverteilung hat:

- Ist die Verteilung ein- oder mehrgipflig?
- Liegt eventuell eine Gaußverteilung (s. S. 144 ff.) vor?
- Wie weit streuen die Werte? Ist die Streuung symmetrisch?
- Kann man Ausreißer identifizieren?

Abbildung 2.2 bis 2.5: Darstellung der Liegezeiten nach Kaiserschnitt als Stabdiagramm, als Histogramm und als Polygonzug, wobei unterschiedliche Klassenbreiten gewählt wurden. Auf der Abszisse ist die Liegezeit x und auf der Ordinate die absolute Häufigkeit $H(x)$ und die relative Häufigkeit $f(x)$ abgetragen. In allen Diagrammen sind der arithmetische Mittelwert \bar{x} und die Standardabweichung s eingetragen (s. S. 34 u. 36 ff. und Abbildung 2.6–2.8)



Ausreißer

Ausreißer sind Extremwerte, die oft nicht zum untersuchten Kollektiv gehören oder bei denen grobe Messfehler vorliegen. In unserem Beispiel könnte dies eventuell Frau A.U. mit 28 Tagen Liegezeit sein, falls sich z.B. herausstellen sollte, dass sie nur deshalb so lange im Krankenhaus war, weil sie sich kurz vor der geplanten Entlassung das Bein gebrochen hat.

Normalerweise wird vor der Erhebung des Datenmaterials ein Katalog von Ein- und Ausschlusskriterien aufgestellt, anhand derer festgelegt wird, wie in untypischen Fällen verfahren werden soll. Wenn man Ausreißer nachträglich von der Datenanalyse ausschließen will, muss man dies im Einzelfall begründen, um nicht den Verdacht der Manipulation zu wecken.

2.3 Beschreibung durch Maßzahlen

In diesem Abschnitt soll die Beschreibung eines Datenmaterials durch Maßzahlen wie Mittelwert, Varianz, Standardabweichung usw. besprochen werden. Es

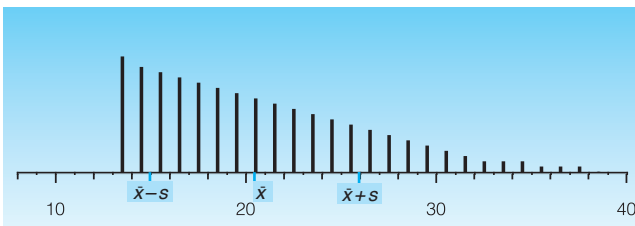
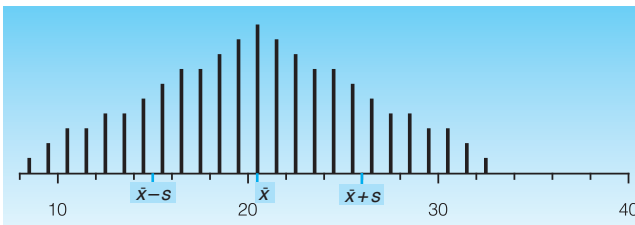
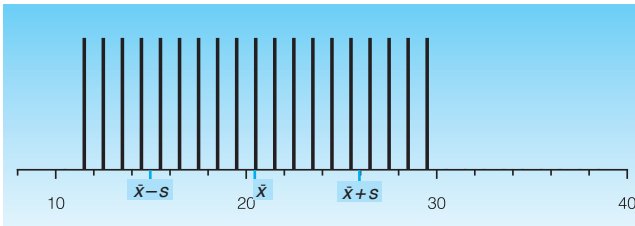


Abbildung 2.6 bis 2.8: Darstellung verschiedener Verteilungen mit identischem Mittelwert $\bar{x} = 20,5$ und identischer Standardabweichung $s = 5,5$. Die Unregelmäßigkeiten in der Abbildung 2.7 sind symmetrisch und wurden eingearbeitet, damit die Standardabweichung wie in den Abb. 2.6 und 2.8 genau $s = 5,5$ beträgt.

handelt sich hierbei um eine Zusammenfassung der vorhandenen Information durch wenige besonders charakteristische Kennzahlen.

Ob es sinnvoll ist, diese Maßzahlen zu berechnen und wie groß ihre Aussagekraft ist, kann am besten anhand der grafischen Darstellung beurteilt werden. Beispielsweise ist der Mittelwert wenig aussagekräftig, wenn die Verteilung mehrgipflig oder wenn die Streuung groß ist.

Die Maßzahlen fassen die im Datenmaterial vorhandenen Informationen zusammen. Hierbei gehen jedoch Einzelheiten der Datenstruktur verloren. Messreihen können trotz gleichen Mittelwertes und gleicher Varianz (s.S. 36) eine unterschiedliche Verteilung der Werte aufweisen, wie in den Abbildungen 2.6 bis 2.8 gezeigt wird. Deshalb reicht die Angabe von Mittelwert und Standardabweichung keineswegs aus, um eine Verteilung vollständig zu charakterisieren.

2.3.1 Stichprobe – Grundgesamtheit

Nur in seltenen Fällen wie z.B. bei einer Volkszählung wird die interessierende Grundgesamtheit vollständig erfasst. In der Regel werden die Daten anhand einer Stichprobe erhoben und die Ergebnisse werden dann auf die Grundgesamtheit übertragen. In unserem Beispiel sind wir an den Liegezeiten nach Schnittentbindung im Krankenhaus XYZ interessiert, um Voraussagen über zukünftige Patientinnen machen zu können. Die interessierende Grundgesamtheit aller Patientinnen mit Schnittentbindung liegt demnach zumindest zum Teil in der Zukunft und kann deshalb nicht in die Erhebung einbezogen werden. Die dreißig Patientinnen des ersten Halbjahres 2011 dienen lediglich als Stichprobe. Die Ergebnisse der Stichprobe sind die Grundlage, um die *Werte der Grundgesamtheit zu schätzen*.

Bei der tabellarischen und grafischen Darstellung ist es offensichtlich, dass die Werte aus der Stichprobe stammen. Wenn daraufhin Maßzahlen wie Mittelwert oder Varianz errechnet werden, muss deutlich gemacht werden, ob es sich um

- **Maßzahlen der Stichprobe** handelt oder um
- **Schätzwerte für die Grundgesamtheit**, die aus den Daten einer Stichprobe berechnet wurden, oder um die tatsächlichen (meist unbekannt)
- **Maßzahlen der Grundgesamtheit**.

Es liegt hier eine ähnliche Beziehung vor wie zwischen einer Meinungsumfrage, der Hochrechnung einer Wahl und dem amtlichen Wahlergebnis. Die Beziehung zwischen Stichprobe und Grundgesamtheit hat für die Statistik fundamentale Bedeutung.

2.3.2 Lage- oder Lokalisationsmaße

Die Lokalisationsmaße geben die Lage der Werte auf der Zahlengeraden an. Hierzu gehören Kennzahlen zur Charakterisierung des Durchschnittswertes und die sog. Quantile und Perzentile, wobei wir Letztere im Zusammenhang mit der Verteilungsfunktion besprechen werden (s.S. 40).

Arithmetischer Mittelwert \bar{x}

Durch Addition der Verweildauern erfahren wir, dass die Patientinnen zusammen insgesamt 390 Tage im Krankenhaus verbracht haben, im arithmetischen Mittel \bar{x} also $\bar{x} = \frac{390}{30} = 13$ Tage

$$\bar{x} = \frac{\text{Summe aller } x\text{-Werte}}{\text{Anzahl aller } x\text{-Werte}} = \frac{\bullet x_i}{n}$$

Hier und im Folgenden lassen wir aus Gründen der Übersichtlichkeit den Summationsindex unter dem Summenzeichen weg und schreiben:

$$\sum x_i \text{ statt } \sum_{i=1}^n x_i$$

Median \tilde{x} (Zentralwert)

Der Median liegt **in der Mitte aller beobachteten Werte**. Er wird von höchstens der Hälfte aller Werte unterschritten und von höchstens der Hälfte aller Werte überschritten. Nach dieser Definition muss der Median in unserem Beispiel so liegen, dass höchstens 15 Patientinnen einen kürzeren und höchstens 15 Patientinnen einen längeren Krankenhausaufenthalt haben. Dies ist, wie wir im Stamm- und Blattdiagramm leicht abzählen können, der 12. Tag.

Wenn z.B. Frau E.N. am Tag 13 (statt Tag 12) entlassen worden wäre, so läge der Median laut obiger Definition irgendwo zwischen dem 12. und 13. Tag. Aufgrund einer Konvention würde man den Median dann mit 12,5 Tagen beziffern.

Modalwert

Unter Modalwert versteht man den **am häufigsten auftretenden Wert** (leicht zu merken am Wort „Mode“). Wir erkennen aus dem Stamm- und Blattdiagramm, dass der 11. Tag der Tag mit der größten relativen Häufigkeit $f(x)$ ist, also der Tag, an dem die meisten Patientinnen das Krankenhaus verließen.

Vergleich der Lagemaße

Zur Kennzeichnung des „Durchschnitts“ gibt es drei Lokalisationsmaße:

- den arithmetischen Mittelwert \bar{x} (in unserem Beispiel 13);
- den Median \tilde{x} (in unserem Beispiel 12);
- den Modalwert (in unserem Beispiel 11).

Am gebräuchlichsten ist der arithmetische **Mittelwert** \bar{x} . Sein Nachteil besteht darin, dass er **von Ausreißern stark beeinflusst** wird. In unserem Beispiel könnte Frau A.U. ein Ausreißer sein. Wenn Frau A.U. als Ausreißer eingestuft und von der Berechnung des Mittelwertes ausgeschlossen wird, ergibt sich ein \bar{x} für die übrigen 29 Patientinnen von $\bar{x} = \frac{362}{29} = 12,48$.

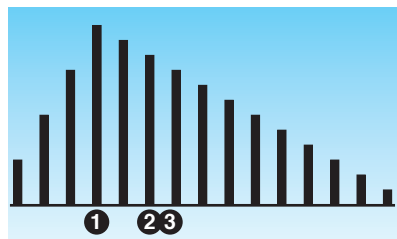
Der **Median** wird ebenfalls häufig verwendet. Er ist gegen Ausreißer in der Regel unempfindlich. Der Median \tilde{x} hat gegenüber dem arithmetischen Mittelwert \bar{x} gelegentlich noch den Vorteil, früher bestimmbar zu sein. Wenn man z.B. die durchschnittliche Überlebenszeit nach einer neuen Krebstherapie ermitteln will, kann man den arithmetischen Mittelwert \bar{x} erst errechnen, nachdem alle Patienten verstorben sind, während man die mediane Überlebenszeit kennt, sobald 50% der Patienten nicht mehr leben.

Der **Modalwert** ist nur dann sinnvoll, wenn die Klassenbesetzung groß ist. Wenn in jede Klasse nur wenige Werte fallen, hängt es sonst zu stark vom Zufall ab, in welche Klasse die meisten Beobachtungen fallen.

Unterschiede bestehen auch in Bezug auf das Skalenniveau, das Voraussetzung zur Berechnung der Lokalisationsmaße ist: Der arithmetische Mittelwert setzt eine intervallskalierte Größe voraus, der Median eine ordinalskalierte Größe, während für den Modalwert bereits eine Nominalskala ausreichend ist. Beispielsweise kann man fragen, welches das häufigste in Deutschland gefahrene Auto ist. Die Frage nach dem Median oder arithmetischen Mittelwert der benutzten Kraftfahrzeuge ist jedoch sinnlos.

Bei unsymmetrischen, sog. „schiefen“ Verteilungen nehmen Modalwert (1), Median (2) und arithmetischer Mittelwert (3) in der Regel unterschiedliche Werte an, wobei der Median zwischen den beiden anderen Lagemaßen liegt.

Abbildung 2.9: Die Beziehung zwischen dem Modalwert (1), dem Median (2) und dem arithmetischen Mittelwert (3) bei schiefen Verteilungen. Auf der Abszisse sind die Messwerte und auf der Ordinate die absolute oder relative Häufigkeit der jeweiligen Messwerte abgetragen. Der arithmetische Mittelwert wird von den extremen Werten in besonderem Maße beeinflusst.



2.3.3 Streuungs- oder Dispersionsmaße

Dispersions- oder Streuungsmaße geben die Streuung der Messwerte an. Hierzu gehören Spannweite, Varianz und Standardabweichung, die in diesem Abschnitt besprochen werden, sowie die Perzentile und Quartile, die im Zusammenhang mit der Verteilungsfunktion erläutert werden.

Das einfachste Dispersionsmaß ist die **Spannweite**, auch **Variationsbreite** oder (engl.) **range** genannt, die sich aus der **Differenz zwischen dem höchsten und dem niedrigsten Wert** ergibt. In unserem Beispiel:

$$28 \text{ Tage} - 8 \text{ Tage} = 20 \text{ Tage.}$$

Der Nachteil dieses Streuungsmaßes besteht darin, dass es von Ausreißern sehr stark beeinflusst wird. Ohne Frau A.U. erhalten wir:

$$19 \text{ Tage} - 8 \text{ Tage} = 11 \text{ Tage.}$$

Quartilsabstand

Der Quartilsabstand umfasst den **Bereich mit den mittleren 50 % der Werte**: Der Quartilsabstand umfaßt die mittleren beiden Quartile (s.S. 43).

Im Box- und Whisker-Diagramm von Tukey (s.S. 43) wird im Bereich vom ersten bis zum dritten Quartil, d.h. zwischen dem 25. und 75. Perzentil, ein Kasten gezeichnet.

Der Quartilsabstand wird häufig verwendet, weil er von Ausreißern praktisch nicht beeinflusst wird und weil die Berechnung einfach ist.

Varianz s^2

Die empirische Varianz ist das am häufigsten verwendete Streuungsmaß. Sie ergibt sich als **durchschnittliches Abweichungsquadrat** nach folgender Formel:

$$\text{Varianz } s^2 = \frac{\text{Summe aller Abweichungsquadrate}}{\text{Zahl aller Messwerte} - 1}$$

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

(Rechenschema siehe Tabelle 2.4)

Wenn in der Formel zur Errechnung der Varianz anstelle des Mittelwertes \bar{x} ein anderer Wert, z. B. der Median oder Modalwert, eingesetzt wird, ergibt sich ein höherer Wert. Der arithmetische Mittelwert ist deshalb das Lokalisationsmaß, für das die Summe der Abweichungsquadrate am geringsten ist.

Patienten	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$H(x_i)$	$H(x_i) \cdot (x_i - \bar{x})^2$
J.K.	8	-5	+25	1	25
A,B,	9	-4	+16	1	16
Z.N., A.L., M.K.	10	-3	+9	3	27
R.S., L.W, D.S., K.L., VH., F.T.	11	-2	+4	6	24
U.C., S.N., J.W., ER., EN.	12	-1	+1	5	5
N.A., T.E., L.B., H.T.	13	0	0	4	0
U.N., BE., P.J., L.K., N.N.	14	+1	+1	5	5
B.E., P.A.	15	+2	+4	2	8
DE.	17	+4	+ 16	1	16
R.H.	20	+7	+ 49	1	49
A.U.	28	+15	+ 225	1	225
Summe				30	400

Tabelle 2.4: Berechnungsschema für die Varianz und Standardabweichung: In den Spalten 2, 3, 4, 5 und 6 sind die jeweilige Liegezeit x_i , die jeweilige Abweichung vom Mittelwert $(x_i - \bar{x})$, das Abweichungsquadrat $(x_i - \bar{x})^2$, die absolute Häufigkeit der jeweiligen Liegezeit $H(x_i)$ und schließlich das Produkt aus Häufigkeit und Abweichungsquadrat $H(x_i) \cdot (x_i - \bar{x})^2$ aufgeführt.

Division durch n oder durch $n - 1$?

Gelegentlich wird eine geringfügig differierende Formel verwendet, indem im Nenner n statt $n - 1$ steht. Die Formel mit n gibt die *Varianz innerhalb der Stichprobe* an, also die Varianz für die Liegezeiten der 30 Frauen, deren Initialen in der Urliste aufgeführt sind. In unserem Beispiel sind wir im Grunde aber nicht an den 30 Frauen interessiert, die im Behandlungszeitraum in der Klinik entbunden haben, sondern an der Behandlungssituation in der Klinik im Allgemeinen. Wie sich durch theoretische Überlegungen zeigen lässt, ist der aus einer Stichprobe (mit n) errechnete Wert in der Regel etwas kleiner als die in der Grundgesamtheit vorliegende Varianz. Deshalb wird zur *Schätzung der Varianz in der Grundgesamtheit* die Formel mit $n - 1$ verwendet.

Zusammenfassung:

- Varianz innerhalb der Stichprobe: Nenner n ;
- Schätzwert für die Varianz innerhalb der Grundgesamtheit: Nenner $n - 1$.

Der Unterschied zwischen beiden Formeln macht sich hauptsächlich bei kleinen Stichproben bemerkbar, bei unserer Stichprobe von 30 Patienten beträgt die Differenz ca. 3%.

Standardabweichung s

Die Varianz s^2 ist zwar ein häufig verwendetes Streuungsmaß, sie ist jedoch ohne anschauliche Bedeutung. Im Gegensatz hierzu hat die **Quadratwurzel der Varianz**, die Standardabweichung s , eine konkrete anschauliche Bedeutung: Bei einer Gauß- oder Normalverteilung (s.S. 144) liegen ca. 68 % der Werte im Intervall von $\bar{x} \pm s$, also in der Entfernung von höchstens einer Standardabweichung s vom Mittelwert \bar{x} . Etwa 95 % der Messwerte liegen im Bereich $\bar{x} \pm 2s$ und ca. 99,7 % im Intervall von $\bar{x} - 3s$ bis $\bar{x} + 3s$. Die Standardabweichung s besitzt dieselbe Maßeinheit wie die Messwerte.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{400}{29} = 13,79$$

$$s = \sqrt{s^2} = \sqrt{13,79} = 3,71$$

Die Varianz s^2 ergibt sich als 13,79, die Standardabweichung s als 3,71. Der Bereich $\bar{x} \pm s$ liegt also in den Grenzen von $\bar{x} - s = 13 - 3,71 = 9,29$ und $\bar{x} + s = 13 + 3,71 = 16,71$. 25 Patienten aus unserer Stichprobe, das sind 84,3 %, haben Liegezeiten innerhalb dieses Intervalls.

Alternativer Rechenweg

Die Summe der Abweichungsquadrate beträgt 400, wovon 255 durch den Ausreißer, Frau A.U., bedingt sind.

Wir wollen die Varianz ohne Frau A.U. ausrechnen. Nach obiger Formel müssen wir zunächst den Mittelwert \bar{x} neu berechnen: $\bar{x} = \frac{362}{29} = 12,483$. Als nächster Schritt muss die Differenz zwischen jedem x_i und dem Mittelwert \bar{x} bestimmt und quadriert werden. Hierbei entstehen „krumme“ Zahlen, die den Rechenaufwand gegenüber dem oben durchgerechneten Beispiel vergrößern.

Es gibt eine rechentechnisch einfachere Formel, um die Summe der Abweichungsquadrate zu bestimmen:

$$\begin{aligned} \sum (x_i - \bar{x})^2 &= \sum x_i^2 - n\bar{x}^2 \\ &= 4686 - 4518,93 \\ &= 167,07 \\ s^2 &= \frac{167,07}{28} = 5,97 \end{aligned}$$

Die Varianz ohne Frau A.U. beträgt $s^2 = 5,97$, sodass $s = 2,44$. Der Bereich $\bar{x} \pm s$

liegt in den Grenzen von $\bar{x} - s = 12,48 - 2,44 = 10,04$ und $\bar{x} + s = 12,48 + 2,44 = 14,92$ und umfasst 20 von 29 Werten, dies sind ca. 69 %. Dieser Bereich entspricht dem Wert (68 %), den wir bei einer Gaußverteilung erwarten würden.

Falls alle x -Werte sehr groß sind, lässt sich eine weitere rechentechnische Vereinfachung erzielen, indem man von jedem x -Wert einen konstanten Wert a abzieht. Das Ergebnis wird hiervon nicht beeinflusst, weil sich beide Summanden um denselben Betrag verkleinern.

Es gibt heute für ca. 15 Euro Taschenrechner mit statistischen Funktionen, die solche Berechnungen nach Eingabe der Messwerte automatisch durchführen.

Variationskoeffizient

Der Variations- oder Variabilitätskoeffizient gibt an, wie viel Prozent vom arithmetischen Mittelwert die Standardabweichung beträgt:

$$\text{Variabilitätskoeffizient} = \frac{s}{\bar{x}} \cdot 100 \%$$

In unserem Beispiel beträgt der Variationskoeffizient mit Frau A.U.: $\frac{3,71}{13} 100 \% = 28,5 \%$, ohne Frau A.U.: $\frac{2,44}{13,48} 100 \% = 19,5 \%$.

Der Vorteil des Variationskoeffizienten liegt darin, dass er von der gewählten Einheit unabhängig ist. Zum Beispiel können biochemische Messungen in mg, mval, mmol oder I.E. erfolgen, mit jeweils unterschiedlichen Mittelwerten und Standardabweichungen, aber gleichen Variationskoeffizienten.

Voraussetzung für die sinnvolle Verwendung des Variationskoeffizienten ist, dass die Variable rationalskaliert ist, also als Verhältnisskala vorliegt. Beispielsweise ergibt der Variationskoeffizient für Grad Celsius oder Grad Fahrenheit wenig Sinn, denn er würde dort zu unterschiedlichen Ergebnissen führen, je nachdem, in welcher Einheit die Temperatur angegeben wird. Wenn die Temperatur in Kelvin, also in einer rationalskalierten Größe angegeben wird, ist die Berechnung des Variationskoeffizienten sinnvoll.

Die mittlere Abweichung (der Messwerte)

Es handelt sich hierbei um ein sehr anschauliches Streuungsmaß, das allerdings selten verwendet wird.

$$\text{Mittlere Abweichung} = \frac{\sum |x_i - \bar{x}|}{n}$$

Die Betragstriche sind notwendig, damit sich bei der Summation der Abweichungen ($x_i - \bar{x}$) positive und negative Werte nicht zu Null ergänzen, denn $\sum (x_i - \bar{x}) = 0$. Die mittlere Abweichung darf nicht mit dem mittleren Fehler des Mittelwertes (s. S. 175) verwechselt werden.

Für unser Beispiel erhalten wir mit Frau A.U. eine mittlere Abweichung von $\frac{70}{30} = 2,33$ und ohne Frau A.U. von $\frac{55}{29} = 1,89$. Dieses Streuungsmaß ist erheblich weniger ausreißerempfindlich als die Standardabweichung, weil keine Quadratzahlen zu bilden sind.

Vergleich der Streuungsmaße

Am gebräuchlichsten sind der **Quartilsabstand**, weil die Berechnung einfach ist und weil diese Größe von Ausreißern praktisch nicht beeinflusst wird, und die **Varianz** bzw. die **Standardabweichung**. Die zuletzt genannten Größen haben aus mathematischen Gründen eine große praktische Bedeutung, obwohl sie sehr ausreißerempfindlich sind. Im Kapitel 8 wird die Bedeutung der Standardabweichung bei normalverteilten Werten ausführlich besprochen.

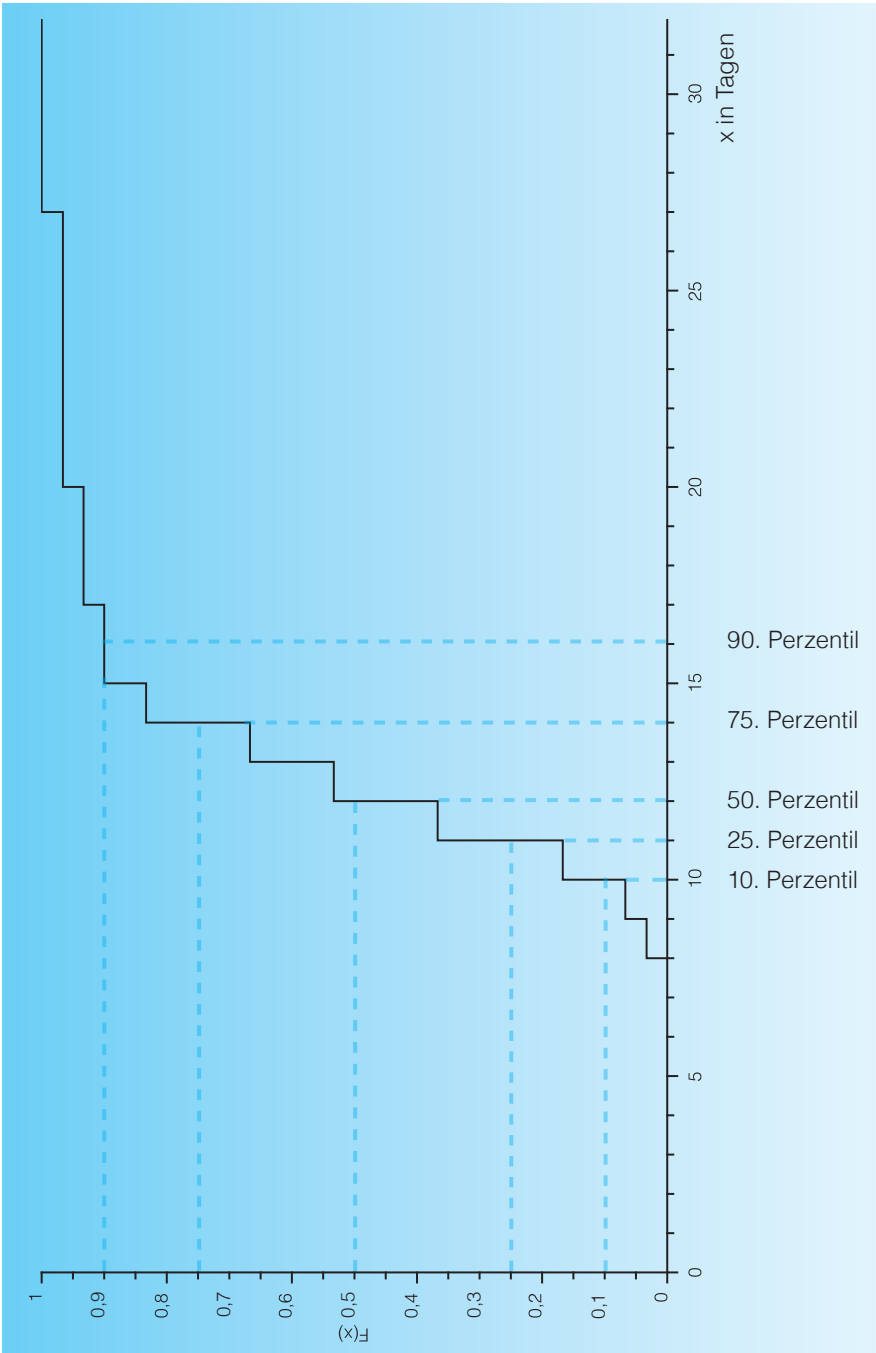
Die **mittlere Abweichung** wird in der Praxis kaum verwendet, obwohl sie eine anschauliche Größe ist. Dies gilt auch für den **Variationskoeffizienten**, der nur bei rationalskalierten Größen erlaubt, dann jedoch unabhängig von der Maßeinheit ist.

2.4 Die empirische Verteilungsfunktion

Die empirische Verteilungsfunktion $F(x)$ **gibt an, welcher Anteil aller Werte kleiner oder gleich dem Wert x ist**. Die Verteilungsfunktion wird deshalb auch **kumulative Häufigkeitsfunktion** oder **Summenhäufigkeitsfunktion** genannt. Man spricht in diesem Zusammenhang auch vom Prozentrang eines Wertes. Die Verteilungsfunktion für unser Beispiel mit den Liegezeiten nach Kaiserschnitt ist in der rechten Spalte von Tabelle 2.3 angegeben. In dieser Tabelle sind die Liegezeiten ihrer Größe nach angeordnet. Die zweite Spalte von rechts gibt die relative Häufigkeit $f(x)$ für die einzelnen Entlassungstage x an. Die kumulative Häufigkeit des Tages x_k ergibt sich als Summe der relativen Häufigkeiten $f(x)$ für alle Tage x , die kleiner oder gleich x_k sind:

$$F(x_k) = \sum_{i=0}^k f(x_i)$$

Abbildung 2.10: Summenhäufigkeitsfunktion der Liegezeiten nach Kaiserschnitt. Auf der Abszisse ist die Zeit x in Tagen abgetragen, auf der Ordinate die Summenhäufigkeit $F(x)$, deren Werte in Tabelle 2.3 angegeben sind. Die gestrichelten Linien stellen die Perzentile dar.



Für unser Beispiel können wir aus Tabelle 2.3 entnehmen:

$$F(7) = 0$$

$$F(8) = \frac{1}{30}$$

$$F(9) = \frac{2}{30}$$

...

...

$$F(28) = \frac{30}{30} = 1.$$

Die Verteilungsfunktion $F(x)$ kann stets nur Werte zwischen 0 und 1 annehmen, weil sie den relativen Anteil aller Werte angibt, die kleiner oder gleich x sind.

In der grafischen Darstellung ergibt sich eine Treppenkurve, wobei die Höhe der Stufen der Zahl der am jeweiligen Tag entlassenen Patienten entspricht. Die Darstellung als Polygonzug ist ebenfalls möglich.

Perzentiltabellen

Die kumulative Häufigkeitsfunktion $F(x)$ gibt für einen gegebenen x -Wert an, wie viel Prozent aller Werte kleiner oder gleich diesem x -Wert sind. Häufig ist die Fragestellung jedoch umgekehrt, wenn man wissen will, **welcher x -Wert von z.B. 10 % aller Werte unterschritten oder erreicht wird.**

Diese Frage wird durch Perzentiltabellen beantwortet, die dazu dienen, bei gegebenem $F(x)$ -Wert den dazugehörigen x -Wert zu suchen. Dies kann auch anhand der grafischen Darstellung der Verteilungsfunktion geschehen.

Unser Beispiel zeigt, dass mehrere Perzentile auf denselben x -Wert fallen können. Das 40. und das 50. Perzentil fallen z.B. beide auf den 12. Tag. Umgekehrt kann auch ein Perzentil auf mehrere x -Werte fallen: Das 90. Perzentil (90 % aller Werte sind kleiner oder gleich dem 90. Perzentil) liegt zwischen dem 15. Tag, an dem die 27. Patientin entlassen wurde, und dem 17. Tag, an dem die 28. Patientin entlassen wurde. Einer Konvention entsprechend gilt der 16. Tag als 90. Perzentil.

Die oben geschilderten Probleme der eindeutigen Zuordnung zwischen den $F(x)$ - und x -Werten sind umso geringer, je höher die Anzahl und je kleiner die Höhe der einzelnen Stufen in der Treppenkurve ist, d.h. je näher die einzelnen Werte beieinander liegen. Bei einer empirischen Verteilungsfunktion einer stetigen Variablen, die auf sehr viele Stellen hinter dem Komma bestimmt worden ist, sodass niemals zwei Beobachtungen denselben Wert ergeben, erhalten wir eine Treppenkurve, bei der die Anzahl der Stufen gleich der Anzahl der Beobachtungen ist.

Wahrscheinlichkeitsverteilung

Wenn man eine Verteilungsfunktion ohne Stufen haben will, bei der sich jedem x -Wert genau ein $F(x)$ -Wert zuordnen lässt und umgekehrt, benötigt man unendlich viele Beobachtungen, was in der Praxis unmöglich ist.

Anstatt die Summenhäufigkeitsfunktion von unendlich vielen Beobachtungen zu bilden, kann man anhand eines mathematischen Modells die Wahrscheinlichkeiten berechnen, mit der die einzelnen x -Werte bzw. kumulativen $F(x)$ -Werte zu erwarten sind. Auf diese Weise erhält man eine glatte Kurve ohne Stufen, die Wahrscheinlichkeitsverteilung heißt. Näheres hierzu wird im Kapitel 7 *Binomialverteilung* erläutert. Für die Wahrscheinlichkeitsfunktion einer stetigen Variablen gilt, dass genau $p\%$ aller Werte kleiner und genau $(100 - p)\%$ aller Werte größer als das p -te Perzentil sind.

Quartile

Perzentile geben sowohl Auskunft über die Lage als auch über die Streuung der Variablen. Sie werden deshalb von einigen Autoren als Lagemaße und von anderen als Streuungsmaße bezeichnet. Das Perzentil wird deshalb häufig **Quantil** oder **Fraktil** genannt.

Unter **Quartilen** versteht man das 25., 50. und 75. Perzentil, also das Quantil 0,25, 0,5 und 0,75. Der **Median** entspricht dem 50. Perzentil.

Box- und Whisker-Diagramm

Das Kasten- und (Katzen-)Schnurrhaar-Diagramm von Tukey, oft auch abgekürzt als „Boxplot“ bezeichnet, stellt schematisch eine Häufigkeitsverteilung dar: Zwischen dem 1. und 3. Quartil wird ein Kasten aufgebaut. In diesen Be-

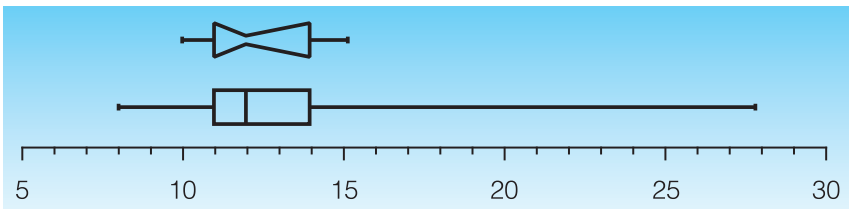


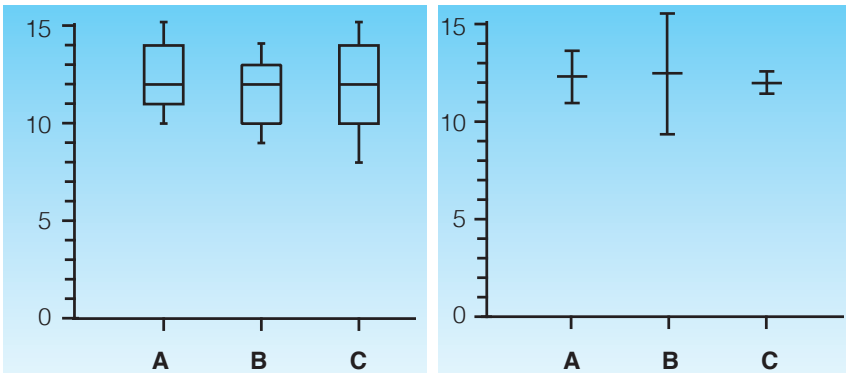
Abbildung 2.11 und 2.12: Das Box- und Whisker-Diagramm von Tukey am Beispiel der tageweise klassierten Daten aus Tabelle 2.4. Beide Abbildungen geben in komprimierter Form die Daten der Abb. 2.2 bis 2.5 wider. Die Whiskers sind in der oberen Abbildung vom 10. bis zum 90. Perzentil, unten bis zu den Extremwerten durchgezogen. Der Median ist oben durch eine Einkerbung und unten durch den senkrechten Strich angedeutet. Die beiden auf *identischen Daten* beruhenden Boxplots sollen deutlich machen, wie unterschiedlich Boxplots dargestellt werden können. Abbildung 2.13 stellt dieselben Daten dar. Dort ist das Boxplot um 90 Grad gedreht, die Darstellung dort ist eine Mischung aus 2.11 und 2.12.

reich fallen 50 % der Beobachtungen. Die seitlich angesetzten Schnurrhaare vermitteln einen Eindruck, wie weit die restlichen Werte streuen.

Wie weit die Schnurrhaare ausgezogen werden, wird unterschiedlich gehandhabt. Einige Autoren gehen bis zu den Extremwerten, andere bis zum 5. und 95. Perzentil oder auch nur bis zum 10. und 90. Perzentil, wodurch das Boxplot weniger ausreißerempfindlich wird.

Vergleich mehrerer Verteilungen

Das Box- und Whisker-Diagramm ist gut geeignet, um mehrere Verteilungen miteinander zu vergleichen. Jedes einzelne Diagramm nimmt weniger Platz in Anspruch als ein Histogramm, sodass man zahlreiche horizontale Diagramme übereinander oder vertikale Diagramme nebeneinander anordnen kann. Auf diese Weise lassen sich die Mediane und Quartile mit einem Blick vergleichen. Wenn der Median nicht an der Mitte des Kastens angeordnet ist, handelt es sich um eine schiefe Verteilung, unregelmäßig lange Schnurrhaare lassen auf Ausreißer schließen. Als Beispiel betrachten wir den Vergleich der Liegezeiten nach



Links: Abbildung 2.13 bis 2.15: Die drei Boxplots geben die **Quartile** tageweise klassierter Verweildauern nach Kaiserschnitt in den Krankenhäusern A, B und C an, wobei die Abb. 2.12 und 2.13 dieselben Daten darstellen wie Abb. 2.11. Lediglich die Art der Darstellung differiert. Die Tatsache, dass in allen Krankenhäusern der 12. Tag der Median ist, hängt auch damit zusammen, dass die Liegezeit tageweise klassiert ist, sodass als Median nur volle Tage infrage kommen. Lediglich für den unwahrscheinlichen Fall, dass der Median genau zwischen zwei Tagen liegt, wäre ein Zwischenwert möglich. Dieselben Überlegungen gelten für das erste und dritte Quartil. Wenn man die *stundenweise* klassierten Daten der Tab. 2.3 verwenden würde, gäbe es dieses Problem nicht.

Rechts: Abbildung 2.16 bis 2.18: In den Abb. 2.16 bis 18 sind der **arithmetische Mittelwert** zusammen mit dem sog. **Fehlerbalken** dargestellt. Der Fehlerbalken umfasst ein Intervall von zwei Standardfehlern zu jeder Seite des Mittelwertes. Im Krankenhaus B wurden nur 10 Patienten untersucht, in Klinik C jedoch 300. Dies erklärt die unterschiedliche Länge der Fehlerbalken. Zum Konzept des Standardfehlers s. S. 175.

Sectio in drei verschiedenen Krankenhäusern A, B und C. Zu diesem Zweck ordnen wir in den Abbildungen 2.13 bis 2.15 die Boxplots senkrecht an.

Mittelwert und Fehlerbalken

Die wahrscheinlich häufigste Darstellung wissenschaftlicher Daten ist in den Abbildungen 2.16 bis 2.18 wiedergegeben. Die Darstellung sieht ähnlich aus wie ein Boxplot, stellt aber einen vollkommen anderen Sachverhalt dar: Hier sind die Daten auf den **arithmetischen Mittelwert \bar{x}** und den **mittleren Fehler des Mittelwertes $s_{\bar{x}}$** komprimiert. Der sog. **Fehlerbalken** umfasst in der Regel den Bereich von zwei Standardfehlern zu beiden Seiten des Mittelwertes \bar{x} , in dem der wahre Wert μ mit einer Wahrscheinlichkeit von 95% liegt. Einzelheiten werden in den Abschnitten 9.2 *Zufällige Fehler* und 15.1 *Schätzen* erläutert. Der mittlere Fehler des Mittelwertes errechnet sich als:

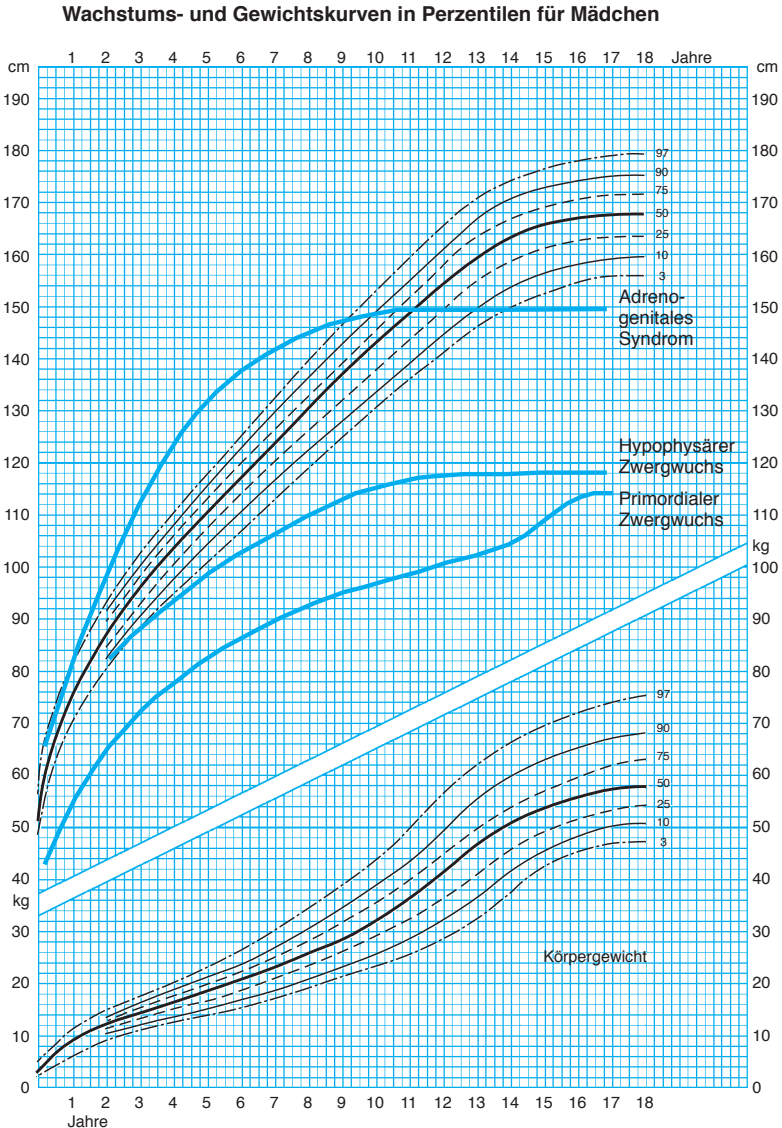
$$\text{Mittlerer Fehler des Mittelwertes } s_{\bar{x}} = \text{Standardfehler } s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

In unserem Fall hatten wir die Standardabweichung als $s = 3,71$ Tage ausgerechnet, sodass $s_{\bar{x}} = 3,71 / \sqrt{30} = 0,68$ Tage. Dies bedeutet, dass der Mittelwert $\bar{x} = 13$ Tage den tatsächlichen Mittelwert μ mit einer Genauigkeit von ungefähr $\pm 2 \cdot 0,68$ Tagen schätzt. Doch Vorsicht: Häufig täuscht der Standardfehler eine Genauigkeit vor, die gar nicht vorhanden ist. In unserem Beispiel haben wir in Tab. 2.4 nur mit vollen Tagen gerechnet. Hätten wir die Tab 2.3 mit Angabe der Stunden der Mittelwertberechnung zugrunde gelegt, wäre \bar{x} um etwa 0,5 Tage größer gewesen. Wenn man bei der Datenerhebung nur an den vollen Tagen interessiert ist und die Stunden ignoriert, aber später die Verweildauer auf zwei Stellen hinter dem Komma angibt, entsteht ein *systematischer Fehler!*

Somatogramme in der Kinderheilkunde

Perzentiltabellen spielen in der Kinderheilkunde eine große Rolle, wenn es darum geht, Wachstumsstörungen zu diagnostizieren. Es werden hauptsächlich Perzentiltabellen der Körpergröße, des Körpergewichts und des Kopfumfangs verwendet. Diese Somatogramm genannten Darstellungen werden erstellt, indem Tausende von Kindern einer jeden Altersstufe untersucht und dann für jeden Jahrgang die Summenhäufigkeiten für die jeweiligen Maße errechnet werden.

Vor der Anwendung eines Somatogramms muss man prüfen, ob das Kind derselben Population angehört, in der dieses erstellt worden ist. Für Kinder aus dem mediterranen Raum sind die in der deutschen oder nordamerikanischen Bevölkerung erhobenen Tabellen nur mit Vorbehalt und entsprechenden Korrekturen gültig. Auch haben Tabellen, die vor 70 oder 80 Jahren erarbeitet wurden, wegen der Akzeleration heute nur noch eingeschränkte Gültigkeit.



Lothar Reinken et al.: *Klin. Pädiatr.* **192**, 25-33 (1980).
 Ingeborg Brandt: *Der Kinderarzt* **11**, 45-51 (1980).
 Ingeborg Brandt: *Human Growth. A Comprehensive Treatise*. 2. Ed. Vol. 1. Hrsg. F. Falkner und J. M. Tanner, Plenum Press. New York 1986.
 Ingeborg Brandt und Lothar Reinken: *Klin. Pädiatr.* **200**, 451-456 (1988).
 Lothar Reinken und Gerta v. Oost: *Klin. Pädiatr.* **204**, 129-133 (1992).

Abbildung 2.19: Drei typische Beispiele pathologischen Wachstums im Vergleich mit einem Somatogramm, wie es in den pädiatrischen Vorsorgeuntersuchungen verwendet wird. Die mit Zahlen versehene Kurvenschar gibt den Bereich normalen Wachstums an, wobei die Linien die verschiedenen Perzentile darstellen. Eine Entwicklungsstörung hat häufig zur Folge, dass das Kind seine bisherigen Perzentile verläßt.

Überlebenskurven

Die Verteilungsfunktion findet unmittelbar praktische Anwendung in Form der Überlebenskurven. Zur Beurteilung der Effektivität einer bestimmten Therapie zeichnet man eine Kurve, die angibt, nach welcher Zeit noch wie viel Prozent der Patienten leben. Es handelt sich hierbei um „eine auf den Kopf gestellte“ Summenhäufigkeitsfunktion der aufgetretenen Todesfälle.

Die Analyse von Überlebenskurven spielt in der Medizin eine große Rolle. Weitere Einzelheiten werden im Zusammenhang mit dem Logranktest und der Kaplan-Meier-Methode auf Seite 351 besprochen.

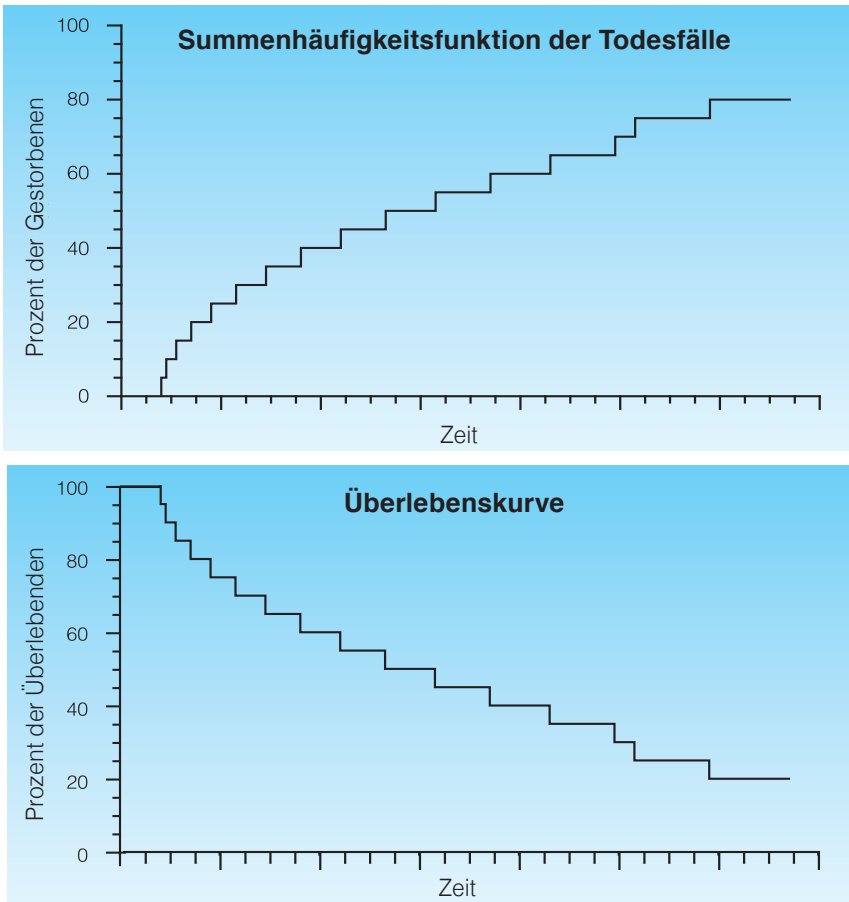


Abbildung 2.20 und 2.21: Oben: Summenhäufigkeitsfunktion von Todesfällen anhand eines fiktiven Beispiels. Unten: Dieselben Daten in anderer Darstellung. Im oberen Diagramm gibt die Ordinate den Prozentsatz der verstorbenen Patienten an, im unteren Diagramm den Prozentsatz der noch lebenden Patienten.