3.6. Zufallsvariable

Der Begriff der *Zufallsvariable* spielt eine ganz zentrale Rolle in der Statistik. Wir beginnen mit einem einfachen Beispiel. Ein Pack Jasskarten besteht aus 36 verschiedenen Karten. Damit wir beim Jassen einen Sieger ermitteln können, werden den Jasskarten Zahlenwerte zugeordnet. So hat ein König den Wert 4. Ohne diese Werte wären verschiedene Stiche beim Jass sehr schwierig miteinander zu vergleichen. Wir haben also eine *Funktion*, die jeder Jasskarte einen Zahlwert zuordnet. Dieselbe Situation haben wir in der Stochastik. Oft wird ein Zufallsexperiment mit Zahlenwerten verknüpft. Zu jedem Elementarereignis ω gehört demnach ein Zahlenwert $X(\omega) = x$. Dabei ist X eine *Funktion*, die jedem Elementarereignis ω den Zahlenwert x zuordnet.

Beispiel 3.6.1

Wir ziehen Karten aus einem Stapel Jasskarten. Dabei wird jeder Karte ein Wert zugeordnet.

$$
\begin{array}{lcl}
\omega = \text{As} & \mapsto & X(\omega) = 11 \\
\omega = \text{König} & \mapsto & X(\omega) = 4 \\
\quad\vdots & & \quad\quad\vdots \\
\omega = \text{Sechs} & \mapsto & X(\omega) = 0
\end{array}
$$

□

Wie man in diesem Beispiel sieht, ist X eine Funktion auf dem Grundraum Ω. Diese Funktion wird *Zufallsvariable* genannt. Sie ordnet jedem Element des Grundraumes *eine Zahl* zu. Der Vorteil in diesem Vorgehen liegt darin, dass wir mit den Werten der Zufallsvariable rechnen können.

So ist es im Beispiel oben möglich, mit den Zahlenwerten $X(\omega)$ den „Durchschnitt“ (den sogenannten *Erwartungswert*, der später in diesem Kapitel behandelt wird) der gezogenen Karten zu berechnen. Für die Elementareignisse „As“, „König“ etc. macht das Wort „Durchschnitt“ keinen Sinn.

Wir halten fest:

Zufallsvariable

Eine **Zufallsvariable** X ist eine **Funktion**:

$$
\begin{array}{llcl}
X: & \Omega & \to & \mathbb{R} \\
 & \omega & \mapsto & X(\omega)\,,
\end{array}
$$

Die Notation X (oder auch $Y, Z, \ldots$) ist eher ungewohnt für die Bezeichnung einer Funktion, ist aber üblich in der Wahrscheinlichkeitsrechnung.

Konventionen

Eine Zufallsvariable wird mit einem **Grossbuchstaben** X (oder Y, Z) bezeichnet. Der entsprechende **Kleinbuchstabe** x (oder y, z) stellt einen konkreten Wert dar, den die Zufallsvariable annehmen kann. Für das Ereignis, bei dem die Zufallsvariable X den Wert x annimmt, schreiben wir $X = x$.

Bei einer Zufallsvariable ist nicht die Funktion $X(\cdot)$ zufällig, sondern nur das Argument ω: Je nach Ausgang des Zufallsexperiments ω erhalten wir einen anderen Wert $x = X(\omega)$. Wir nennen dann x eine **Realisierung** der Zufallsvariablen X. Falls das Experiment zweimal durchgeführt wird, und wir zweimal das gleiche Ergebnis ω erhalten, dann sind auch die realisierten Werte von X gleich.

Beispiel 3.6.2

Im Jasskartenbeispiel entspricht die Realisierung $X = 11$ dem Ziehen eines Asses.

□

In diesem Kapitel sind die Zahlen, die X annehmen kann, *diskret*. D.h. die Anzahl dieser Werte ist endlich wie im Jasskartenbeispiel

$$\{0, 2, 3, 4, 10, 11\}$$

oder sie ist eine unendliche Liste wie

$$\{2.5, 4.5, 6.5, 8.5, \ldots\}$$

Wir sagen in diesem Fall, dass die Zufallsvariable X *diskret* ist. Insbesondere sind Anzahlen stets diskret, während Messungen besser als kontinuierliche Zahlen, also mit $\mathbb{R}$, modelliert werden, obwohl man praktisch nur mit endlicher Genauigkeit messen kann.

Wahrscheinlichkeitsverteilung einer Zufallsvariablen

Wir haben schon gesehen, wie wir die Wahrscheinlichkeit $P(E)$ eines Ereignisses E berechnen. Entsprechend können wir auch die Wahrscheinlichkeit einer allgemeinen Realisierung x einer Zufallsvariable X definieren. Dazu zunächst folgendes Beispiel.

Beispiel 3.6.3

Wir definieren die Zufallsvariable X als den Wert einer gezogenen Jasskarte und fragen nun, wie gross die Wahrscheinlichkeit ist, dass die gezogene Karte den Wert 4 hat. Die Realisierung ist $X = 4$. Die zugehörige Wahrscheinlichkeit bezeichnen wir mit

$$P(X = 4)$$

Die Realisation $X = 4$ entspricht dem Ziehen eines Königs. D.h., wir suchen die Wahrscheinlichkeit, dass ein König gezogen wird:

$$\begin{aligned} P(X = 4) &= P(\{\omega \mid \omega = \text{ ein König}\}) \\ &= P(\text{Eicheln-König}) + P(\text{Rosen-König}) + \\ &\quad + P(\text{Schellen-König}) + P(\text{Schilten-König}) \\ &= \frac{4}{36} = \frac{1}{9} \end{aligned}$$

Das Vorgehen hier ist ausführlicher ausgeführt als unbedingt notwendig, aber es lässt sich auf nicht so einfache Beispiele verallgemeinern.

□

Die Wahrscheinlichkeit, dass ein König gezogen wird, ist also gleich der Summe der Wahrscheinlichkeiten, die verschiedenen Könige zu ziehen. Diese Überlegung können wir verallgemeinern.

Die Werte einer Zufallsvariablen X (die möglichen Realisierungen von X) treten mit gewissen Wahrscheinlichkeiten auf. Die Wahrscheinlichkeit, dass X den Wert x annimmt, berechnet sich wie folgt:

$$P(X = x) = P(\{\omega \mid X(\omega) = x\}) = \sum_{\omega; X(\omega) = x} P(\omega)$$

Im Jasskartenbeispiel ist $x = 4$ und ω alle möglichen Könige, deren entsprechende Wahrscheinlichkeiten aufaddiert werden.

Wir haben im Beispiel vorher die Wahrscheinlichkeit *einer* Realisierung berechnet. Natürlich können wir die Wahrscheinlichkeiten *aller* Realisierungen berechnen. Dies führt uns auf den *sehr wichtigen* Begriff einer Wahrscheinlichkeitsverteilung.

Wahrscheinlichkeitsverteilung

Berechnen wir für *jede* Realisierung einer Zufallsvariable die zugehörige Wahrscheinlichkeit, so sprechen wir von einer *Wahrscheinlichkeitsverteilung* dieser Zufallsvariablen.

Beispiel 3.6.4

Die Zufallsvariable X ist wieder der Wert einer gezogenen Jasskarte. Die Wahrscheinlichkeit

$$P(X = 4) = \frac{1}{9}$$

haben wir schon berechnet.

Die Wahrscheinlichkeit $P(X = 0)$ können wir mit der Laplace-Wahrscheinlichkeit berechnen. Es hat unter den 36 Karten genau 16 „leere" Karten. Somit gilt für die Wahrscheinlichkeit $P(X = 0)$:

$$P(X = 0) = \frac{16}{36} = \frac{4}{9}$$

Die Realisierung $X = 2$ entspricht dem Ziehen eines Unders. Da es 4 von denen gibt, gilt für die Wahrscheinlichkeit $P(X = 2)$:

$$P(X = 2) = \frac{4}{36} = \frac{1}{9}$$

Die Wahrscheinlichkeiten für die anderen Realisierungen berechnen wir analog. Wir haben dann *jeder* Realisierung einen Wahrscheinlichkeitswert zugeordnet. Wir sprechen dann eben von einer *Wahrscheinlichkeitsverteilung*. Die Wahrscheinlichkeitsverteilung von X ist in der Tabelle 3.5 zu aufgeführt. Die Werte für $P(X = 1)$ oder

x	0	2	3	4	10	11
$P(X = x)$	4/9	1/9	1/9	1/9	1/9	1/9

Tabelle 3.5.: Wahrscheinlichkeitsverteilung von gezogenen Jasskarten.

$P(X = 178)$ sind in Tabelle 3.5 nicht aufgeführt. Der Grund dafür ist natürlich, dass diese Werte nicht gezogen werden können. Wir können ihnen aber trotzdem eine Wahrscheinlichkeit zuordnen, nämlich die Zahl 0:

$$P(X = 1) = 0 \quad \text{oder} \quad P(X = 178) = 0$$

Addieren wir alle Werte in unserer Wahrscheinlichkeitsverteilung, so müssen wir 1 erhalten, da wir eine Realisierung ziehen *müssen*.

$$P(X = 0) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 10) + P(X = 11) = 1$$

□

Wahrscheinlichkeitsverteilung

Die „Liste“ von $P(X = x)$ für alle möglichen Werte x heisst diskrete **(Wahrscheinlichkeits-) Verteilung** der diskreten Zufallsvariablen X. Dabei gilt immer

$$\sum_{\text{alle möglichen } x} P(X = x) = 1.$$

Wir werden später auch nichtdiskrete Verteilungen kennenlernen.

3.7. Binomialverteilung

In diesem Kapitel betrachten wir die Situation, wo es um das Zählen der Anzahl Erfolge (oder Misserfolge) geht. Solche Anwendungen treten z.B. auf bei der Qualitätskontrolle, Erfolg/Misserfolg bei Behandlungen (medizinisch, biologisch) oder auch bei Glücksspielen. Wir wollen an einem Beispiel auf die Problemstellung eingehen.

Beispiel 3.7.1

Nehmen wir an, Sie werfen eine faire Münze viermal hintereinander. Die Versuchsreihe hat also einen Umfang von $n = 4$. Die Wahrscheinlichkeit, „Kopf“ (K) zu werfen, ist $\pi = 0.5$; entsprechend ist $1 - \pi = 1 - 0.5$ die Wahrscheinlichkeit, „Zahl“ (Z) zu werfen. Wir belassen die Schreibweise $1 - 0.5$, um das Beispiel später zu verallgemeinern.

Mit X bezeichnen wir die Anzahl Würfe mit „Kopf“. Die möglichen Ausgänge dieses Experiments können wir als eine Liste mit „Wörtern“ bestehend aus 4 Buchstaben darstellen (siehe Tabelle 3.6).

Dabei ist z.B.

$$X(ZZZZ) = 0$$

da kein K in $ZZZZ$ vorkommt. Entsprechend ist

$$X(KKZK) = 3$$

da 3 K's in $KKZK$ vorkommen. Das Ereignis mit $X = 2$ besteht aus allen Elementarereignissen, bei denen zwei K's (2 von 4 Würfen sind „Kopf“) in den Worten vorkommen. Dieses Ereignis besteht aus folgenden 6 Elementarereignissen

$$\{ZZKK, ZKZK, ZKKZ, KZZK, KZKZ, KKZZ\}$$

Elementarereignis	x	Elementarereignis	x
ZZZZ	0	KZZZ	1
ZZZK	1	KZZK	2
ZZKZ	1	KZKZ	2
ZZKK	2	KZKK	3
ZKZZ	1	KKZZ	2
ZKZK	2	KKZK	3
ZKKZ	2	KKKZ	3
ZKKK	3	KKKK	4

Tabelle 3.6.: Wir werfen eine faire Münze 4 mal und die Zufallsvariable X soll die Anzahl Würfe mit „Kopf“ sein.

Wir nehmen an, dass jeder Münzwurf unabhängig vom vorhergehenden Münzwurf ist. Dann gilt für die Wahrscheinlichkeit, dass das Elementarereignis $KKZZ$ eintritt

$$P(KKZZ) = P(K) \cdot P(K) \cdot P(Z) \cdot P(Z) = (0.5)^2 \cdot (1 - 0.5)^2 = 0.0625$$

Jedes der 6 Elementarereignisse vom Ereignis $X = 2$ kommt mit derselben Wahrscheinlichkeit vor. Darum gilt

$$P(X = 2) = 6 \cdot 0.0625 = 0.375$$

Die Wahrscheinlichkeit, dass wir bei 4 Würfen zweimal „Kopf“ werfen, ist 0.375 oder 37.5 %.

Führen wir dieselbe Überlegung für $X = 3$ durch, so gibt es 4 Elementarereignisse mit der Wahrscheinlichkeit

$$P(KKKZ) = P(K) \cdot P(K) \cdot P(K) \cdot P(Z) = (0.5)^3 \cdot (1 - 0.5)^1 = 0.0625$$

Also

$$P(X = 3) = 4 \cdot 0.0625 = 0.25$$

Werfen wir allgemein x-mal K, so werfen wir $(4 - x)$-mal Z. Somit gilt analog zu oben:

$$P(X = x) = (\text{Anzahl Elementarereignisse mit } x \text{ mal Kopf}) \times (0.5)^x \cdot (1 - 0.5)^{4-x}$$

Jetzt stellt sich bloss noch die Frage, wie man im Allgemeinen die Anzahl Elementarereignisse mit x mal Kopf bei 4 Münzwürfen berechnen kann. Ist $x = 2$, so ist dies die Anzahl der Möglichkeiten (Anzahl Elementarereignisse), bei denen zweimal K vorkommt. Dies berechnen wir mit dem sogenannten Binomialkoeffizienten

$$\binom{4}{x} = \frac{4!}{x!(4-x)!}$$

wobei $4! = 4 \cdot 3 \cdot 2 \cdot 1$ („4 Fakultät").

Mit R berechnen wir den Binomialkoeffizienten mit Hilfe des Befehls

```
choose(4, 2)

## [1] 6
```

Somit lautet die Wahrscheinlichkeit $P(X = x)$ analog zu dem Beispiel oben:

$$P(X = x) = \binom{4}{x}(0.5)^x \cdot (1 - 0.5)^{4-x}$$

Im Falle von $x = 2$ haben wir $\binom{4}{2} = 4!/(2! \cdot 2!) = 6$.

$$P(X = 2) = \binom{4}{2}(0.5)^2 \cdot (1 - 0.5)^2 = \frac{4!}{2!2!}(0.5)^2 \cdot (1 - 0.5)^2 = 6 \cdot 0.5^4 = 0.375$$

Und für $X = 3$

$$P(X = 3) = \binom{4}{3}(0.5)^3 \cdot (1 - 0.5)^1 = \frac{4!}{3!1!}(0.5)^3 \cdot (1 - 0.5)^1 = 4 \cdot 0.5^4 = 0.25$$

Führen wir unser Experiment oftmals durch, so können wir in einem Viertel der Experimente dreimal K erwarten.

$\square$

Diese Überlegungen verallgemeinern wir nun.

Bernoulli-Verteilung

Hier interessiert uns jetzt nur, ob ein Ereignis eintrifft oder nicht. Wir haben es also ausschliesslich mit einer Ja-Nein-Entscheidung zu tun.

Beispiel 3.7.2

- Bei *einem* Münzwurf ist die Frage zum Beispiel, ob das Ergebnis „Kopf" eintritt. Falls die Münze fair ist, so trifft dieses Ereignis mit der Wahrscheinlichkeit $\pi = 1/2$ ein. Die Wahrscheinlichkeit, „Zahl" zu werfen, ist dann $1 - \pi = 1/2$.
- Beim Wurf eines fairen Würfels betrachten wir das Ereignis, die Zahl 4 zu werfen. Die Wahrscheinlichkeit, eine 4 zu werfen, ist $\pi = 1/6$. Die Wahrscheinlichkeit, keine 4 zu werfen, ist dann $1 - \pi = 5/6$.

□

Wir können dies mit Hilfe einer Zufallsvariablen formulieren. Dazu setzen wir die Zahl 1 für „Ereignis tritt ein" (oder „Erfolg haben") und die Zahl 0 für „Ereignis tritt nicht ein" (oder „kein Erfolg haben"). Damit haben wir eine Zufallsvariable definiert.

Wir können unsere Wahrscheinlichkeiten im Würfelwurfbeispiel mit Hilfe der Zufallsvariablen formulieren.

$$P(X = 1) = \frac{1}{6} \quad \text{und} \quad P(X = 0) = \frac{5}{6}$$

Wir können dies verallgemeinern.

Die Verteilung einer Zufallsvariable X mit Werten in $W = \{0, 1\}$ kann durch einen einzelnen Parameter π beschrieben werden:

$$P(X = 1) = \pi, \qquad P(X = 0) = 1 - \pi, \qquad 0 \leq \pi \leq 1.$$

Diese Verteilung heisst **Bernoulli(π)-Verteilung**.

Die Bernoulli-Verteilung ist für sich allein nicht sehr interessant.

Binomialverteilung

Interessanter wird es, wenn wir das Bernoulli-Experiment n Mal wiederholen, also z.B. die Münze n-mal werfen. Der Grundraum Ω besteht dann aus allen „Wörtern" der Länge n, welche man mit den Buchstaben K (für „Kopf") und Z (für „Zahl") schreiben kann. Der Grundraum Ω hat dann 2^n Elemente. Warum ist das so?

Nehmen wir $n = 4$. So gibt es für den 1. Wurf zwei Möglichkeiten, nämlich K und Z. Für jeden weiteren Wurf gibt es natürlich wieder je zwei Möglichkeiten. Die gesamte Anzahl Möglichkeiten besteht dann aus dem Produkt der Möglichkeiten pro Wurf:

$$2 \cdot 2 \cdot 2 \cdot 2 = 2^4 = 16$$

Für ein allgemeines n geht die Überlegung analog.

Wir betrachten die Zufallsvariablen

$$X_i = \begin{cases} 1 & \text{falls } K \text{ im } i\text{-ten Wurf} \\ 0 & \text{falls } Z \text{ im } i\text{-ten Wurf} \end{cases}$$

$$X = X_1 + X_2 + \ldots + X_n = \sum_{i=1}^{n} X_i \ : \ \text{Gesamtzahl von Würfen mit } K$$

Mit dem Wort $KKZK$ wollen wir die Verwendung dieser Zufallsvariablen veranschaulichen. Es ist dann

$$X_1(\boldsymbol{K}KZK) = 1$$

da beim 1. Wurf K geworfen wurde. Entsprechend gilt

$$X_2(K\boldsymbol{K}ZK) = 1; \quad X_3(KK\boldsymbol{K}K) = 0; \quad X_4(KKZ\boldsymbol{K}) = 1$$

Für die Zufallsvariable X erhalten wir

$$\begin{aligned} X(KKZK) &= X_1(\boldsymbol{K}KZK) + X_2(K\boldsymbol{K}ZK) + X_3(KK\boldsymbol{Z}K) + X_4(KKZ\boldsymbol{K}) \\ &= 1 + 1 + 0 + 1 \\ &= 3 \end{aligned}$$

Es kommt in diesem Wort also dreimal K vor.

Um die Verteilung von X zu bestimmen, müssen wir Wahrscheinlichkeiten auf Ω festlegen. Dazu sollen die Ereignisse $X_i = 1$ (also „K im i-ten Wurf") alle die gleiche Wahrscheinlichkeit π haben und unabhängig sein.

Bemerkungen:

i. Wir müssen hier jeweils klar zwischen den beiden Zufallsvariablen X_i und X unterscheiden. Die Zufallsvariable X_i kann nur die Werte 0 und 1 annehmen, während X Werte zwischen 0 und n annehmen kann. So gilt für $n = 5$ und das Wort $KKZZK$

$$X_4 = 0 \quad \text{und} \quad X = 3$$

das heisst, im 4. Wurf wurde Z geworfen und insgesamt trifft dreimal K ein.

Um $P(X = x)$ zu berechnen, brauchen wir zwei Überlegungen

- Wie gross ist die Wahrscheinlichkeit, dass in einem speziellen Fall (Wort) x mal die 1 und $n - x$ mal die 0 vorkommt?

  Nehmen wir wieder $n = 5$ und das Wort $KKKZZ$. Dann ist

$$P(KKKZZ) = \pi^3(1 - \pi)^2$$

- Auf wie viele Arten kann man x K's auf n Plätze anordnen?

  Nehmen wir wieder $n = 5$ und $x = 3$:

$$KKKZZ,\ KKZKZ,\ KZKKZ,\ ZKKKZ,\ ZKKZK,$$
$$ZKZKK,\ ZZKKK,\ KZZKK,\ KKZZK,\ KZKZK$$

  Dies sind 10 Anordnungen. Allerdings ist diese Möglichkeit des Abzählens nur für sehr kleine n durchführbar. Es gibt aber eine allgemeine Formel. Die Antwort ist gegeben durch den Binomialkoeffizienten[2]

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}.$$

  So ist für $n = 5$ und $x = 3$

$$\binom{5}{3} = \frac{5!}{3! \cdot 2!} = \frac{120}{6 \cdot 2} = 10$$

So ist dann für $x = 3$ und $n = 5$ der Ausdruck $P(X = 3)$ die Wahrscheinlichkeit, dass (in diesem Fall) genau dreimal K geworfen wird. Sie berechnet sich durch

$$P(X = 3) = \binom{5}{3} \pi^3 (1-\pi)^2$$

Bestimmen wir für *alle* Realisierungen von X, also für *jede* mögliche Anzahl K im Wort der Länge n, die Wahrscheinlichkeit $P(X = x)$, so erhalten wir die allgemeine *Binomialverteilung*. Die möglichen Realisierungen sind

$$\{0, 1, 2, \ldots, n\}$$

Binomial(n,π)-Verteilung

Eine Zufallsvariable X mit Werten in $\{0, 1, \ldots, n\}$ heisst Binomial(n, π)-verteilt, falls

$$P(X = x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}, \quad x = 0, 1, \ldots, n$$

Dabei ist $0 \leq \pi \leq 1$ der Erfolgsparameter der Verteilung.

[2] Der Binomialkoeffizient wird im Anhang A.1 hergeleitet.

Bemerkungen:

i. Damit sind wir in der Lage, für jeden Wert von $X = x$ die zugehörige Wahrscheinlichkeit für gegebene n und π zu berechnen.

Wir können diese Wahrscheinlichkeit auch mit R berechnen. Ist beispielsweise $x = 3$, $n = 5$ und $\pi = 0.5$, so berechnen wir $P(X = 3)$ durch

```
dbinom(3, 5, 0.5)

## [1] 0.3125
```

Wir können auch alle Wahrscheinlichkeiten der Verteilung auf einmal berechnen

```
dbinom(0:5, 5, 0.5)

## [1] 0.03125 0.15625 0.31250 0.31250 0.15625 0.03125
```

Analog wie in unserem Beispiel ist die Binomialverteilung angebracht für die Zufallsvariable „Anzahl Erfolge/Misserfolge" (Eintreten eines bestimmten Ereignis) bei n **unabhängigen** Versuchen. Die Bedingung „unabhängig" ist wesentlich für die Korrektheit der Binomialverteilung.

Konvention

Wenn man notieren will, dass die Zufallsvariable X einer gewissen Wahrscheinlichkeitsverteilung F folgt, schreibt man abgekürzt: $X \sim F$. Dabei kann F von Parametern abhängen, also z.B. $X \sim F(\theta)$. Wenn also X einer Binomialverteilung mit Parametern n und π folgt, schreibt man abgekürzt $X \sim \text{Binomial}(n, \pi)$ oder einfach nur $X \sim \text{Bin}(n, \pi)$.

Beispiel 3.7.3

Bei einer Losbude steht: „Jedes 5. Los gewinnt!", d.h., die Gewinnwahrscheinlichkeit ist bei jedem Los $\pi = 0.2$. Nehmen wir weiter an, dass das Ziehen von einem Los keinen Einfluss auf das Ziehen des nächsten Loses hat (z.B. gibt es eine riesige Anzahl Lose und die Lostrommel wird nach jedem Verkauf eines Loses gut gemischt). Wir kaufen 100 Lose und bezeichnen mit X die Anzahl Gewinne unter den 100 Losen. Dann ist X Binomial($n = 100$, $\pi = 0.2$) verteilt. Abgekürzt: $X \sim \text{Bin}(100, 0.2)$.

Welche Aussagen können wir machen? Zuerst betrachten wir den Plot der Wahrscheinlichkeitsverteilung (R-Befehl weiter unten). In Abbildung 3.10 sehen wir, dass

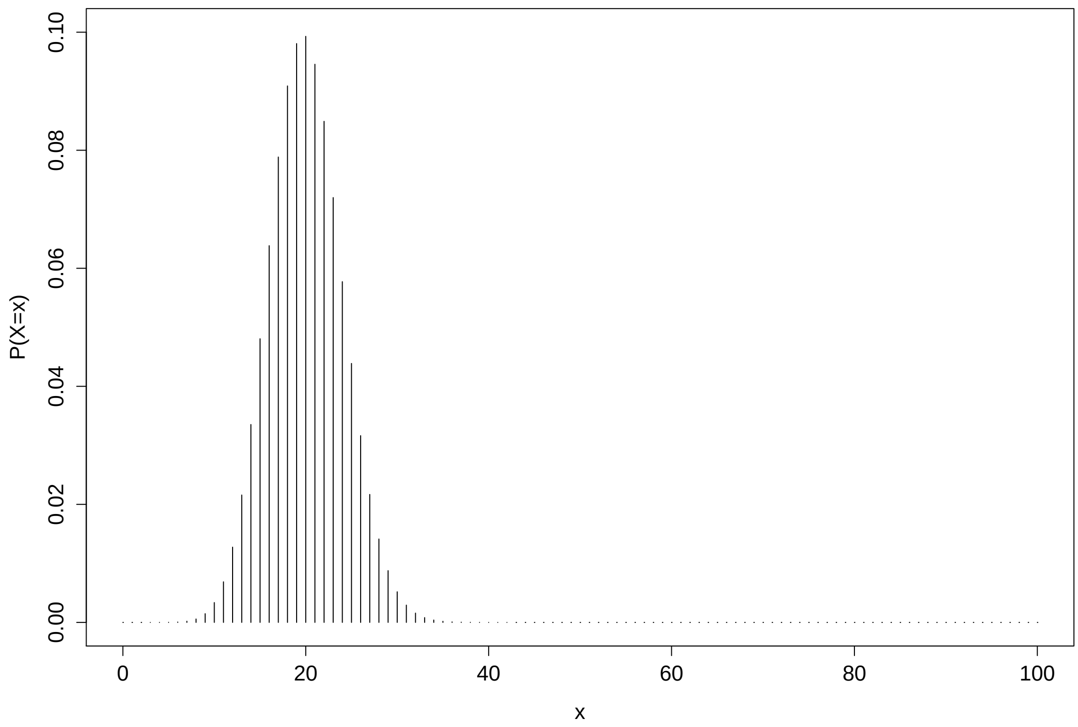

Abbildung 3.10.: Wahrscheinlichkeiten für die Anzahl Gewinne in 100 Losen

es am wahrscheinlichsten ist, etwa 20 Gewinne auf 100 Lose zu ziehen. Das war aber auch zu erwarten, da die Erfolgswahrscheinlichkeit 0.2 ist.

Für $X = 20$ erhalten wir die Wahrscheinlichkeit

$$P(X = 20) = \binom{100}{20} \cdot 0.2^{20} \cdot 0.8^{80} = 0.0993$$

Diesen Wert haben wir natürlich mit R berechnet. Also mit einer Wahrscheinlichkeit von 0.1 gewinnen wir 20-mal beim Kauf von 100 Losen.

Aus Abbildung 3.10 ist auch ersichtlich, dass mit einer Anzahl Gewinne zwischen 15 und 25 zu rechnen ist (was das genau bedeutet, werden wir im nächsten Kapitel sehen). Wir wollen demnach die Wahrscheinlichkeit von $15 \leq X \leq 25$ berechnen:

$$\begin{aligned} P(15 \leq X \leq 25) &= P(X = 15) + P(X = 16) + \ldots + P(X = 25) \\ &= \binom{100}{15} \cdot 0.2^{15} \cdot 0.8^{85} + \binom{100}{16} \cdot 0.2^{16} \cdot 0.8^{84} + \ldots + \binom{100}{25} \cdot 0.2^{25} \cdot 0.8^{75} \\ &= 0.83 \end{aligned}$$

Auch dies wurde mit R berechnet:

```
sum(dbinom(15:25, 100, 0.2))

## [1] 0.8320809
```

Mit über 80 % Wahrscheinlichkeit haben wir zwischen 15 und 25 Gewinnen beim Kauf von 100 Losen.

In Abbildung 3.10 sehen wir auch, dass die Wahrscheinlichkeit, 35 und mehr Gewinne zu ziehen, sehr klein ist. Die Rechnung bestätigt diesen Eindruck

$$P(35 \leq X \leq 100) = 0.000\,34$$

Wir wären ein ziemlich grosser Glückspilz, wenn wir 35 oder mehr Gewinne hätten. Andererseits ist man ein Pechvogel, wenn man weniger als 10 Gewinne zieht

$$P(0 \leq X \leq 9) = 0.0023$$

□

Eigenschaften der Binomialverteilung (siehe Abb. 3.11):

- $P(X = x)$ ist maximal, wenn x gleich dem ganzzahligen Teil von $(n+1)\pi$ ist, und auf beiden Seiten von diesem Wert nehmen die Wahrscheinlichkeiten monoton ab.
- Wenn $n\pi(1-\pi)$ nicht allzu klein ist, ist die Verteilung praktisch symmetrisch und hat die Form einer Glocke.
- Wenn n gross ist, sind die meisten Wahrscheinlichkeiten $P(X = x)$ verschwindend klein, d.h. grosse Abweichungen von $(n+1)\pi$ sind extrem unwahrscheinlich.

Mit R können die sogenannten *Stabdiagramme* aus Abbildung 3.11 wie folgt erzeugt werden:

```
plot(0:10, dbinom(0:10, size = 10, prob = 0.25), ylim = c(0,
    0.3), xlim = c(0, 100), type = "h", xlab = "x",
    ylab = "P(X=k)")

points(0:10, dbinom(0:10, size = 10, prob = 0.25), cex = 0.5,
    col = "dark red", xlim = c(0, 100), ylim = c(0,
        0.3))
```

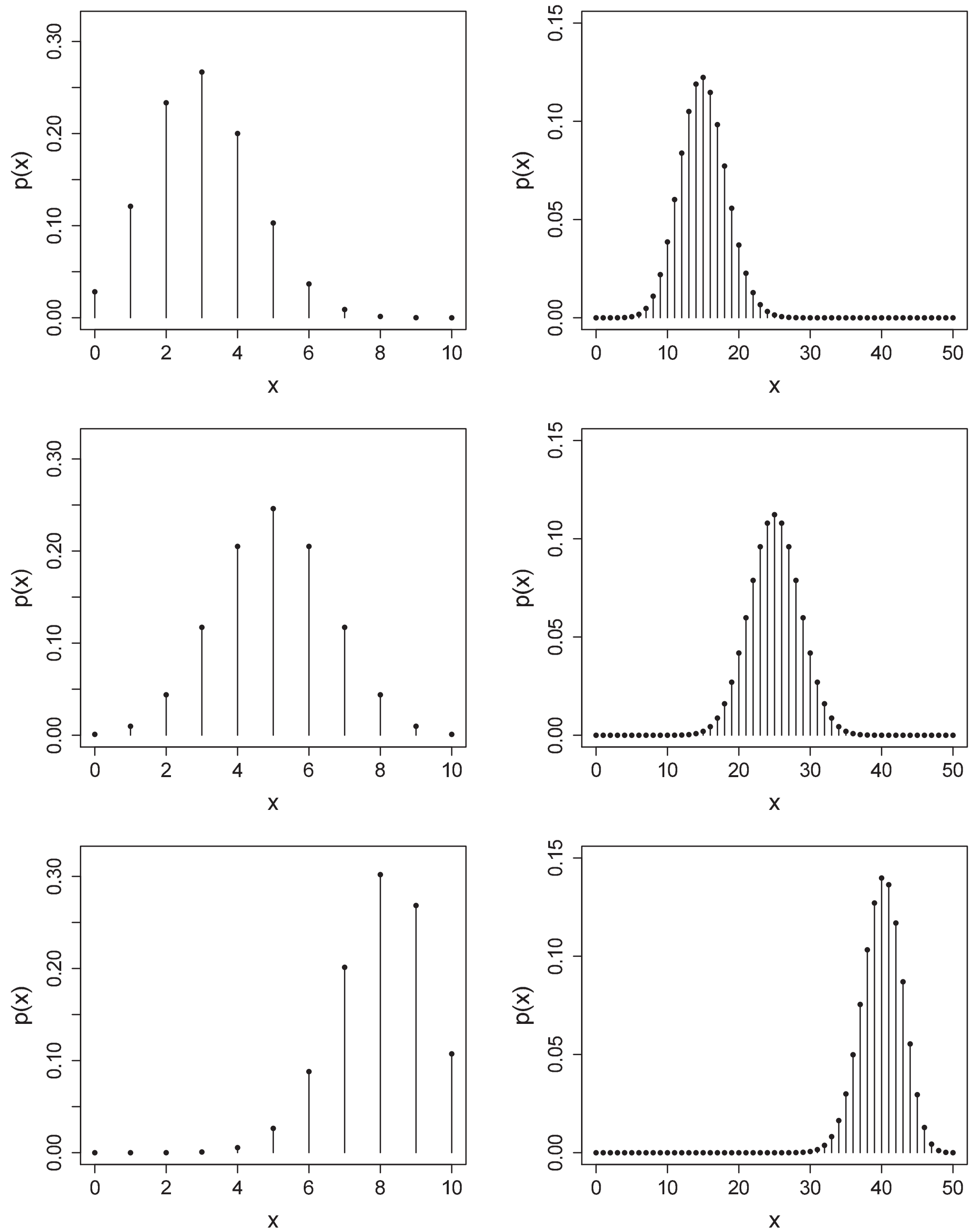

Abbildung 3.11.: Die Binomialwahrscheinlichkeiten $P(X = x)$ wird als Funktion von x für verschiedene Werte von n und π in Form eines Stabsdiagramms dargestellt. Links ist $n = 10$ und rechts ist $n = 50$ für jeweils $\pi = 0.3, 0.5, 0.8$ (von oben nach unten).

Beispiel 3.7.4

Ein Analog-Digital-Wandler oder *A/D-Wandler* ist ein elektronisches Gerät oder Bauteil zur Umsetzung analoger Eingangssignale in digitale Daten bzw. in einen Datenstrom, der dann weiterverarbeitet oder gespeichert werden kann. Analog-Digital-Wandler sind elementare Bestandteile fast aller Geräte der modernen Kommunikations- und Unterhaltungselektronik wie beispielsweise Mobiltelefonen, Digitalkameras, oder Camcordern. Zudem werden sie in der Messwerterfassung in industriellen Anwendungen, Maschinen und in technischen Alltagsgegenständen wie Autos oder Haushaltsgeräten eingesetzt.

Nach der A/D-Wandlung eines analogen Signals stehen die Daten als binäre Rechteckfolgen direkt für eine weitere Verarbeitung, Speicherung oder Übertragung zur Verfügung. Während in Verarbeitungskomponenten im allgemeinen alle Bits parallel vorliegen, wird insbesondere bei Übertragungen digitaler Signale eine serielle Darstellung der Bits hintereinander in verschachtelter Form bevorzugt.

Die Übertragung von digitalen Signalen ist häufig fehleranfällig. Wir nehmen an, dass die n Übertragungsversuche unabhängig voneinander sind. Wir bezeichnen mit X die Anzahl fehlerhaft übertragener Bits. Wir testen ein neues Material für den Übertragungskanal und stellen fest, dass von 100 übertragenen Bits 6 fehlerhaft empfangen werden. Ein vernünftiges Modell ist dann:

$$X \sim \text{ Binomial}(100, \pi),$$

wobei π unbekannt ist. Effektiv beobachtet wurden $x = 6$ fehlerhaft übertragene Bits d.h. $X = x = 6$ wurde tatsächlich **realisiert**. Fortsetzung folgt.

□

3.8. Kennzahlen einer Verteilung

Eine beliebige (diskrete) Verteilung kann vereinfachend zusammengefasst werden durch 2 Kennzahlen, den *Erwartungswert* $\mathrm{E}(X)$ und die *Standardabweichung* $\sigma(X)$.

Der **Erwartungswert** einer diskreten Zufallsvariable X beschreibt die mittlere Lage der Verteilung und ist wie folgt definiert:

$$\mathrm{E}(X) = \sum_x xP(X = x)$$

Die **Standardabweichung** beschreibt die Streuung der Verteilung. Rechnerisch ist das Quadrat der Standardabweichung, die sogenannte **Varianz** bequemer:

$$\begin{aligned}\mathrm{Var}(X) &= \sum_x (x - \mathrm{E}(X))^2 P(X = x) \\ \sigma(X) &= \sqrt{\mathrm{Var}(X)}.\end{aligned}$$

Der Erwartungswert einer diskreten Zufallsvariable ist das gewichtete arithmetische Mittel von allen möglichen Werten, wobei die Werte mit ihrer Wahrscheinlichkeit gewichtet werden.

Ebenso wird bei der Varianz das Quadrat der Abweichung eines Wertes der Zufallsvariable vom Erwartungswert mit der Wahrscheinlichkeit des Wertes gewichtet. Die Standardabweichung hat dieselbe Einheit wie X, während die Einheit der Varianz deren Quadrat ist: Wird z. B. X in Metern (m) gemessen, so besitzt $\mathrm{Var}(X)$ die Dimension Quadratmeter (m^2) und $\sigma(X)$ wiederum die Dimension Meter (m).

Beispiel 3.8.1

Wir betrachten wiederum das Jasskartenbeispiel mit der Verteilung

x	0	2	3	4	10	11
$P(X = x)$	4/9	1/9	1/9	1/9	1/9	1/9

Wir ziehen aus dem Stapel eine Karte. Welches ist der durchschnittliche Wert der Karte, die wir ziehen?

Dazu berechnen wir den Erwartungswert $\mathrm{E}(X)$:

$$\mathrm{E}(X) = 0 \cdot \frac{4}{9} + 2 \cdot \frac{1}{9} + 3 \cdot \frac{1}{9} + 4 \cdot \frac{1}{9} + 10 \cdot \frac{1}{9} + 11 \cdot \frac{1}{9} = 3.33$$

Dies ist der durchschnittliche Wert, den wir erwarten können, wenn wir sehr oft eine Karte ziehen und diese wieder in den Stapel zurücklegen.

Wir berechnen noch die Varianz und die Standardabweichung:

$$\begin{aligned}\mathrm{Var}(X) &= (0 - 3.33)^2 \cdot \frac{4}{9} + (2 - 3.33)^2 \cdot \frac{1}{9} + (3 - 3.33)^2 \cdot \frac{1}{9} + (10 - 3.33)^2 \cdot \frac{1}{9} + (11 - 3.33)^2 \cdot \frac{1}{9} \\ &= 16.67\end{aligned}$$

und

$$\sigma(X) = \sqrt{16.67} = 4.08$$

Das ist die „mittlere“ Abweichung vom Mittelwert, wobei diese vor allem nach oben möglich ist.

□

Beispiel 3.8.2

Sei $X \sim$ Bernoulli(π). Dann gilt

$$\begin{aligned}
\mathrm{E}(X) &= 0 \cdot P(X=0) + 1 \cdot P(X=1) = \pi \\
\mathrm{Var}(X) &= (0-\mathrm{E}(X))^2 P(X=0) + (1-\mathrm{E}(X))^2 P(X=1) \\
&= (0-\pi)^2(1-\pi) + (1-\pi)^2\pi \\
&= \pi(1-\pi) \\
\sigma(X) &= \sqrt{\pi(1-\pi)}
\end{aligned}$$

□

Rechenregeln für Erwartungswerte und Varianz

$$\mathrm{E}(a+bX) = a + b \cdot \mathrm{E}(X) \quad a,b \in \mathbb{R}$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$$

$$\mathrm{Var}(a+bX) = b^2\,\mathrm{Var}(X) \quad a,b \in \mathbb{R}$$

$$\mathrm{Var}(a) = 0 \quad a \in \mathbb{R}$$

Für die Binomial(n,π)-Verteilung erhält man mit einigen Rechnungen[3]

$$\mathrm{E}(X) = \sum_{x \in W_X} xP(X=x) = \sum_{x \in W_X}^{n} x \cdot \frac{n!}{x!(n-x)!}\pi^x(1-\pi)^{n-x} = n\pi$$

und

$$\mathrm{Var}(X) = \sum_{x \in W_X} (x-\mathrm{E}(X))^2 P(X=x) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = n\pi(1-\pi)$$

und somit

$$\sigma(X) = \sqrt{n\pi(1-\pi)}$$

Beachten Sie, dass aufgrund von Bernoulli(π) = Binomial$(1,\pi)$ diese Formeln mit dem Resultat des obigen Beispiels übereinstimmen.

Die Kennzahlen fassen also sehr gut zusammen, was wir in der Abbildung 3.11 gesehen haben: Die Verteilung ist um den Erwartungswert konzentriert, die Streuung wächst mit n, aber langsamer als n. Für festes n ist die Streuung maximal, wenn $\pi = 1/2$.

[3]Die Rechnungen werden im Anhang A.2 ausgeführt.

Beispiel 3.8.3

Wir sind wieder bei der Losbude, bei der wir (nach dem vierten Bier) 100 Lose gekauft hatten. Um die Freundin zu beeindrucken, kramen wir unser Statistikwissen hervor und berechnen im Kopf den Erwartungswert und die Standardabweichung der Anzahl Gewinne unter 100 Losen.

$$\mathrm{E}(X) = n \cdot \pi = 100 \cdot 0.2 = 20$$

$$\sigma(X) = \sqrt{n\pi(1-\pi)} = \sqrt{100 \cdot 0.2 \cdot 0.8} = 4$$

Wir können also in 100 Losen im Mittel 20 Gewinne erwarten. Die mittlere Abweichung von diesen 20 Gewinnen ist 4. Das stimmt auch in etwa mit unserer Einschätzung aus Abbildung 3.10 auf Seite 85 überein. Wir erinnern uns, dass Beobachtungen typischerweise ein bis zwei Standardabweichungen vom Erwartungswert entfernt liegen und prophezeien der Freundin mit Stolz geschwellter Brust, dass wir wohl zwischen 16 und 24 Gewinnen zu erwarten haben. Sie solle sich schon mal einen Teddybär aussuchen.

Hundert nervenaufreibende Öffnungsversuche später stehen wir mit nur 8 Gewinnen und hängenden Schultern da. Kann das Zufall sein? Ja, aber wir beobachten ein sehr unwahrscheinliches Ereignis (später werden wir solche Überlegungen mit einem statistischen Test präzise formulieren können). Wir beschliessen, den hühnenhaften und grimmig drein schauenden Losbudenbesitzer nicht mit unserer Erkenntnis zu verärgern und denken uns: „Pech im Spiel, Glück in der Liebe!“

□

3.9. Kumulative Verteilungsfunktion

Manchmal ist es für Rechnungen nützlicher, statt der „Liste“ $P(X = x)$ (für alle x) die sukzessiven Summen

$$\sum_{y \le x} P(X = y) = P(X \le x)$$

anzugeben. Dabei läuft x ebenfalls über alle Werte von X.

Kumulative Verteilungsfunktion

Man kann in dieser Definition aber auch beliebige reelle Werte x betrachten und erhält dann die sogenannte **kumulative Verteilungsfunktion**

$$F(x) = P(X \le x) = \sum_{y \le x} P(X = y),$$

Diese springt an den Stellen, wo X realisiert wird und ist dazwischen konstant. Siehe auch Abbildung 3.12.

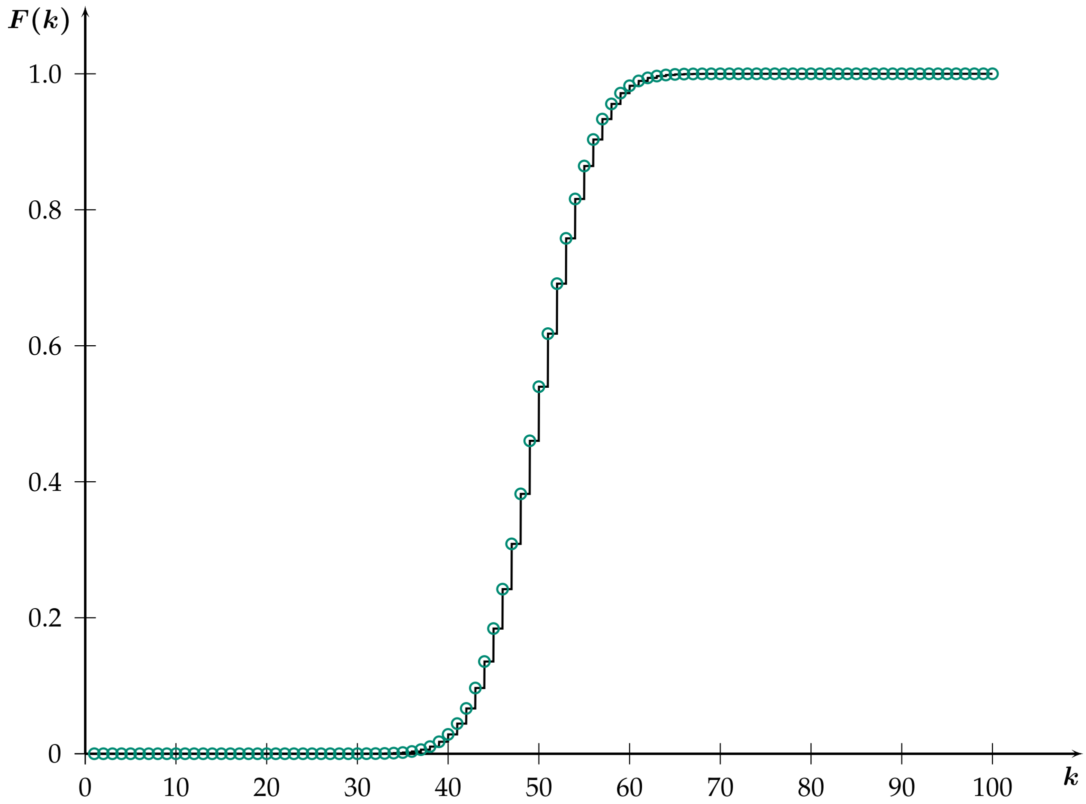

Abbildung 3.12.: Kumulative Verteilungsfunktion $F(\cdot)$ für $X \sim$ Binomial(100,0.5). Die Kreise zeigen an, dass an den Sprungstellen der obere Wert gilt.

Angenommen die Zufallsvariable X folgt der Verteilung Binomial(100,0.5), dann lässt sich die Wahrscheinlichkeit $P(X \leq 50)$ mit R wie folgt berechnen:

```
pbinom(50, 100, 0.5)

## [1] 0.5397946
```

Die kumulative Verteilungsfunktion in Abbildung 3.12 kann mit R wie folgt erzeugt werden

```
plot(pbinom(0:100, size = 101, prob = 0.5), type = "s",
    ylim = c(0, 1), xlab = "k", ylab = "F(k)",
    main = "Binom(100,0.5)")
```

```r
points(pbinom(0:100, size = 101, prob = 0.5), cex = 0.5,
    col = "forest green")
```

Die Wichtigkeit der kumulativen Verteilungsfunktion wird erst im nächsten Kapitel klar.

3.10. Poisson-Verteilung

Bei der Binomial(n, π)-Verteilung nimmt X den beschränkten Wertebereich

$$W = \{0, 1, \ldots, n\}$$

an. Oft ist aber nicht von vorneherein klar, wieviele Versuche/Beobachtungen eintreten. So ist die Anzahl Verkehrsunfälle pro Jahr in der Schweiz zumindest theoretisch jede Zahl möglich. Da wir hier das n nicht kennen, können wir auch nicht die Binomialverteilung anwenden.

Für solche Fälle gibt es aber trotzdem eine Verteilung, die sogenannte *Poisson-Verteilung*[4]

Poisson-Verteilung

Eine Zufallsvariable X mit Werten in $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ heisst Poisson(λ)-verteilt, falls

$$P(X = x) = \exp(-\lambda)\frac{\lambda^x}{x!} \quad (x = 0, 1, 2, \ldots),$$

wobei $\lambda > 0$ ein Parameter der Verteilung ist.

Die Poisson-Verteilung ist die Standardverteilung für unbeschränkte **Zähldaten**.

Die Poisson(λ)-Verteilung kann bei folgenden Anwendungen als Modell gebraucht werden. Bei jeder ist, zumindest theoretisch, eine beliebig grosse Anzahl der betrachteten Ereignisse möglich.

Beispiel 3.10.1

- Anzahl Schadensmeldungen eines Versicherten pro Jahr,
- Anzahl spontaner Ereignisse in einer Nervenzelle während einer Sekunde via Transmitterfreisetzung an einer Synapse.
- Anzahl Autounfälle pro Tag.

[4] Eine Herleitung der Poisson-Verteilung findet sich im Anhang A.3.