# Kapitel 10

# VERTEILUNGEN

Fassung vom 18. Januar 2001

## 10.1 Zufallsvariable.

Häufig wird statt des Ergebnisses $\omega \in \Omega$ eines Zufalls-Experiments eine zugeordnete Zahl $X(\omega) \in \mathbb{R}$ angegeben, z.B. sei

$$\Omega : \text{ Stichprobenraum einer Personengruppe}$$

und für $\omega \in \Omega$

$$X(\omega) : \text{ Alter oder Körpergröße.}$$

**BEISPIEL** Wir betrachten nochmals das unabhängiges Dreistufen-Experiment aus Beispiel 9.4.3. Auf dem Stichprobenraum

$$\Omega = \{+++, ++-, +-+, +--, -++, -+-, --+, ---\}$$

definiert man für $\omega \in \Omega$

$$X(\omega) := k \text{ falls } Rh_- \ k - mal \text{ auftritt.}$$

Diese Funktion erzeugt eine Einteilung von $\Omega$ in 4 Klassen

$$(X = k) := \{\omega \in \Omega \mid X(\omega) = k\} \ ,$$

für $k = 0, 1, 2, 3$ , nämlich die Ereignisse, dass $Rh_-$ gerade $k-$mal auftritt.

Die Wahrscheinlichkeiten $P(X = k)$ sind für $k = 3$ und $k = 2$ schon berechnet. Für $k = 1$ erhält man

$$P(X = 1) = P(++-) + P(+-+) + P(-++) = 3 \cdot (0.85)^2 \cdot 0.15 \simeq 0.325 \ ,$$

und insgesamt

| $k$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = k)$ | 0.614 | 0.325 | 0.057 | 0.003 |

.

Diese Zahlen kann man auch in einem *Histogramm* (oder *Balkendiagramm* ) veranschaulichen.

**DEFINITION** Sei $\Omega$ ein Wahrscheinlichkeitsraum mit Wahrscheinlichkeitsfunktion $P$ . Eine Funktion

$$X : \Omega \longrightarrow \mathbb{R} : \omega \mapsto X(\omega)$$

heißt *Zufallsvariable* . Die Funktion

$$P_X : X(\Omega) \longrightarrow \mathbb{R} : x \mapsto P(X = x) \ ,$$

wobei

$$(X = x) := \{\omega \in \Omega \mid X(\omega) = x\} \ ,$$

heißt *Verteilung* von $X$ . Die Zahl $P(X = x)$ ist die Wahrscheinlichkeit, daß $X$ den Wert $x$ annimmt.

**BEMERKUNG 1** Bei vielen Berechnungen wird nur die Verteilung von $X$ und nicht die Wahrscheinlichkeiten auf $\Omega$ benötigt.

**BEMERKUNG 2** Die Menge $X(\omega)$ ist mit der Verteilung $P_x$ wieder ein Wahrscheinlichkeitsraum.

## 10.2 Erwartungswert und Varianz.

**DEFINITION** Sei $X$ eine Zufallsvariable auf dem Wahrscheinlichkeitsraum $\Omega$ mit der Verteilung $P_X$ und den Werten $x_1, x_2, \cdots, x_n$ . Man definiert

$$p_j := P_X(x_j) = P(X = x_j) \quad \text{für } j = 1, 2, \cdots, n \ .$$

Die Zahlen

$$\mu := E(X) := \sum_{j=1}^{n} p_j \cdot x_j \ ,$$

$$Var(X) := \sum_{j=1}^{n} p_j \cdot (x_j - \mu)^2$$

und

$$\sigma(X) := \sqrt{Var(X)}$$

heißen *Erwartungswert*, *Varianz* und *Standard-Abweichung* von $X$ .

Nimmt die Zufallsvariable unendlich viele Werte an, so werden diese Größen entsprechend durch unendliche Reihen definiert.

Diese Zahlen kann man folgendermaßen interpretieren : Der Erwartungswert gibt den Mittelwert der Werte von $X$ bei häufiger Wiederholung des Experiments. Die Varianz oder die Standard-Abweichung ist ein Maß für die Streuung um den Erwartungswert.

**BEISPIEL** Wir betrachten nochmals das unabhängige Dreistufen-Experiment in Beispiel 10.1. Es gilt

$$E(X) = 0.614 \cdot 0 + 0.326 \cdot 1 + 0.057 \cdot 2 + 0.003 \cdot 3 = 0.449 \simeq 0.45 \ .$$

Die exakte Rechnung ergibt mit $p = 0,15$ und $q = 0,85$

$$\begin{aligned} E(X) &= q^3 \cdot 0 + 3 \cdot q^2 \cdot p \cdot 1 + 3 \cdot q \cdot p^2 \cdot 2 + p^3 \cdot 3 \\ &= 3 \cdot p\,(q^2 + 2pq + q^2) = 3 \cdot p\,(q + p)^2 \\ &= 3 \cdot p \cdot 1 = 0,45 \ . \end{aligned}$$

Für die Varianz erhält man

$$Var(X) = 0.614 \cdot (0 - 0.45)^2 + 0.326 \cdot (1 - 0.45)^2 + 0.057 \cdot (2 - 0.45)^2 + 0.003 \cdot (3 - 0.45)^2 \simeq 0.38 \ ,$$

und ähnlich wie oben die exakte Formel

$$Var(X) = 3pq = 0,3825 \ .$$

Viele Zufallsvariablen haben als Verteilung eine von vier *Standard-Verteilungen* :

Gleich- , Binomial- , Poisson- , Normal-Verteilung.

Beispiele zur Gleichverteilung finden sich in den Beispielen 9.2.1 und 9.2.3 , die übrigen Verteilungen werden in den nächsten Abschnitten behandelt.

## 10.3 Binomial-Verteilung.

**DEFINITION 1** Ein unabhängiges n-Stufen-Experiment, bei dem in jeder Stufe der gleiche Wahrscheinlichkeitsraum $\Omega = \{A, \bar{A}\}$ auftritt, heißt ein *$n-$stufiges Bernoulli-Experiment* .

Der Stichprobenraum wird mit $\Omega^n$ bezeichnet. Er besitzt $2^n$ Elemente. Jedes Element $\omega \in \Omega^n$ ist ein "Wort" der Länge $n$ mit den Buchstaben $A$ und $\bar{A}$ , das man im Baumdiagramm als Zweig darstellen kann.

Mit $p := P(A)$ und $q := P(\bar{A})$ $(= 1 - p)$ werden die elementaren Wahrscheinlichkeiten bezeichnet, und mit $X$ die Zufallsvariable auf $\Omega^n$ , die das Auftreten von $A$ in $\omega$ zählt, d.h.

$$X(\omega) = k \quad \text{falls } A \text{ in } \omega \ \ k - \text{mal auftritt.}$$

**DEFINITION 2** Die Verteilung

$$b_{n,p} : \{0, 1, 2, \ldots , n\} \longrightarrow \mathbb{R} : k \mapsto b_{n,p}(k) := P(X = k)$$

heißt die *Binomial-Verteilung* .

**SATZ** *Für die Binomialverteilung gilt*

$$b_{n,p}(k) = \binom{n}{k} \cdot p^k \cdot q^{n-k} \ .$$

*Ist $Y$ eine Zufallsvariable mit Werten in $\{1, 2, \ldots , n\}$ , die binomialverteilt ist, so gilt*

$$E(Y) = n \cdot p \quad \textit{und} \quad Var(Y) = n \cdot p \cdot q = n \cdot p \cdot (1 - p) \ .$$

Man vergleiche dazu die Berechnungen in Beispiel 10.2 .

**BEISPIEL** Eine Bevölkerungsgruppe habe zu 3 % die Erbanlage $E$ . Es werden 100 Personen zufällig gewählt. Wie groß ist die Wahrscheinlichkeit, daß genau 2 Personen die Erbanlage $E$ haben ? Es gilt

$$p = 0.03 \ , \quad q = 0,97$$

und

$$P(X = 2) = b_{n,p}(2) = \binom{100}{2} \cdot 0.03^2 \cdot 0.97^{98} \simeq 0.225 \ .$$

Wie groß ist die Wahrscheinlichkeit, daß mindestens 2 Personen die Erbanlage $E$ haben ? Man betrachte das Gegenereignis $(X < 2)$ : Es gilt

$$P(X < 2) = P(X = 0) + P(X = 1) = \binom{100}{0} \cdot 0.03^0 \cdot 0.97^{100} + \binom{100}{1} \cdot 0.03^1 \cdot 0.97^{99} \simeq 0.195 \ ,$$

also

$$P(X \geqslant 2) = 1 - P(X < 2) \simeq 0.805 \ .$$

Der Erwartungswert von $X$ ist

$$E(X) = 100 \cdot 0.03 = 3 \ ,$$

d.h. im Mittel gibt es 3 Personen, die die Erbanlage $E$ haben, wenn man mehrmals eine Stichprobe von 100 Personen auswählt.

## 10.4 Poisson-Verteilung.

**BEISPIEL 1** Ein radioaktives Präparat sendet $\alpha$-Teilchen (Heliumionen) aus, wobei $m$ die mittlere Anzahl pro Sekunde ist. Diese Teilchen werden unabhängig voneinander ausgesandt. Das Aussenden von $\alpha$-Teilchen im Zeitintervall $[0,1]$ (eine Sekunde Beobachtungszeit) lässt sich dann als $n-$stufiges Bernoulli-Experiment interpretieren: Sei $p$ die Wahrscheinlichkeit, dass in einem Zeitintervall der Länge $1/n$ ein Teilchen ausgesandt wird, so gilt nach Satz 10.3 für die Wahrscheinlichkeit, dass $k$ Teilchen im Zeitintervall $[0,1]$ ausgesandt werden

$$P(X=k) = b_{n,p}(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \ .$$

Ist nun $n \gg m$ , so ist $p \simeq \frac{m}{n}$ , und demnach

$$P(X=k) \simeq b_{n,\frac{m}{n}}(k) \ ,$$

und zwar umso genauer, je größer $n \in \mathbb{N}$ wird.

S.D. Poisson hat für den Grenzwert $\lim_{n\to\infty} b_{n,\frac{m}{n}}(k)$ Folgendes gezeigt:

$$\lim_{n\to\infty} \binom{n}{k} \cdot \left(\frac{m}{n}\right)^k \cdot (1-\frac{m}{n})^{n-k} = \frac{m^k}{k!} \cdot e^{-m} \ . \qquad (*)$$

**DEFINITION** Die Verteilung

$$\mathbb{N} \longrightarrow \mathbb{R} : k \mapsto p_m(k) := \frac{m^k}{k!} \cdot e^{-m}$$

heißt *Poisson-Verteilung* .

In obigem Beispiel erhält man z.B. für $m = 10$

$$P(X=6) = p_{10}(6) = \frac{10^6}{6!} \cdot e^{-10} = 0.063055... \simeq 6,3\%$$

und

$$P(X=10) = p_{10}(10) = \frac{10^{10}}{10!} \cdot e^{-10} \simeq 12,5\% \ .$$

**SATZ** *Ist $Y$ eine Zufallsvariable mit Werten in $\mathbb{N}$ , die poissonverteilt ist, d.h.*

$$P(Y=k) = p_m(k) \ \textit{für alle} \ k \in \mathbb{N} \ ,$$

*so gilt*

$$E(Y) = m \quad \textit{und} \quad Var(Y) = m \ .$$

**BEMERKUNG** Mit denselben Überlegungen wie in dem Eingangsbeispiel erhält man, dass auch zufällig auf einer Fläche verteilten Partikeln -z.B. vom Wind verteilte Samenkörner oder Blutkörperchen auf einem Objektträger- poissonverteilt sind.

**BEISPIEL 2** Bei zufällig auf einer Fläche verteilten Partikeln soll die Wahrscheinlichkeit berechnet werden, dass auf einem Einheitsquadrat $Q$ mindestens ein Partikel liegt.

Statt $P(X \geq 1)$ ist es hier wesentlich günstiger, das Gegenereignis $P(X < 1)$ zu betrachten, man erhält damit

$$\begin{aligned} P(X &\geq 1) = 1 - P(X < 1) = 1 - P(X = 0) \\ &= 1 - \frac{m^0}{0!} e^{-m} = 1 - e^{-m} . \end{aligned}$$

Um $m$ zu ermitteln, muss ein Durschnittswert durch Auszählen aller Partikel auf einer größeren Fläche berechnet werden. Für $m = 1$ erhält man dann z.B.

$$P(X \geq 1) = 1 - e^{-1} \simeq 0,63 .$$

Ungefähr 63% aller Quadrate mit derselben Fläche wie $Q$ enthalten zumindest einen Partikel.

**BEMERKUNG** Für die obige Formel (*) gilt folgende präzisere Aussage

$$\sum_{k=0}^{\infty} | \, b_{n,p}(k) - p_{n \cdot p}(k) \, | \leq 2n \cdot p^2 .$$

Ist $2n \cdot p^2 \ll 1$, so kann in sehr guter Näherung die Binomial-Verteilung $b_{n,p}(k)$ durch die Poisson-Verteilung $p_{n \cdot p}(k)$ ersetzt werden. Letztere ist wesentlich einfacher zu berechnen.