

Vorlesung 2b

Diskrete Zufallsvariable und ihre Verteilungen

Eine Zufallsvariable heißt **diskret**,

falls ihr Zielbereich endlich oder abzählbar unendlich ist

oder (allgemeiner) eine abzählbare Menge S enthält mit

$$\mathbf{P}(X \in S) = 1.$$

Die zwei Grundregeln
für Wahrscheinlichkeiten:

Additivität:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a), \quad A \subset S$$

Normiertheit auf Eins:

$$\mathbf{P}(X \in S) = 1 .$$

Die Abbildung $A \mapsto \rho[A] := \mathbf{P}(X \in A)$, $A \subset S$,

heißt die **Verteilung** von X .

Die Zahlen $\rho(a) := \mathbf{P}(X = a)$, $a \in S$,

sind die **Verteilungsgewichte**.

Im Fall

$$X = (X_1, \dots, X_n)$$

nennt man die Verteilung ρ von X auch die *gemeinsame Verteilung* von X_1, \dots, X_n .

Die Verteilungen ρ_1, \dots, ρ_n von X_1, \dots, X_n
nennt man
die *Rand- oder Marginalverteilungen* von ρ .

Beispiel: Zufällige Paare

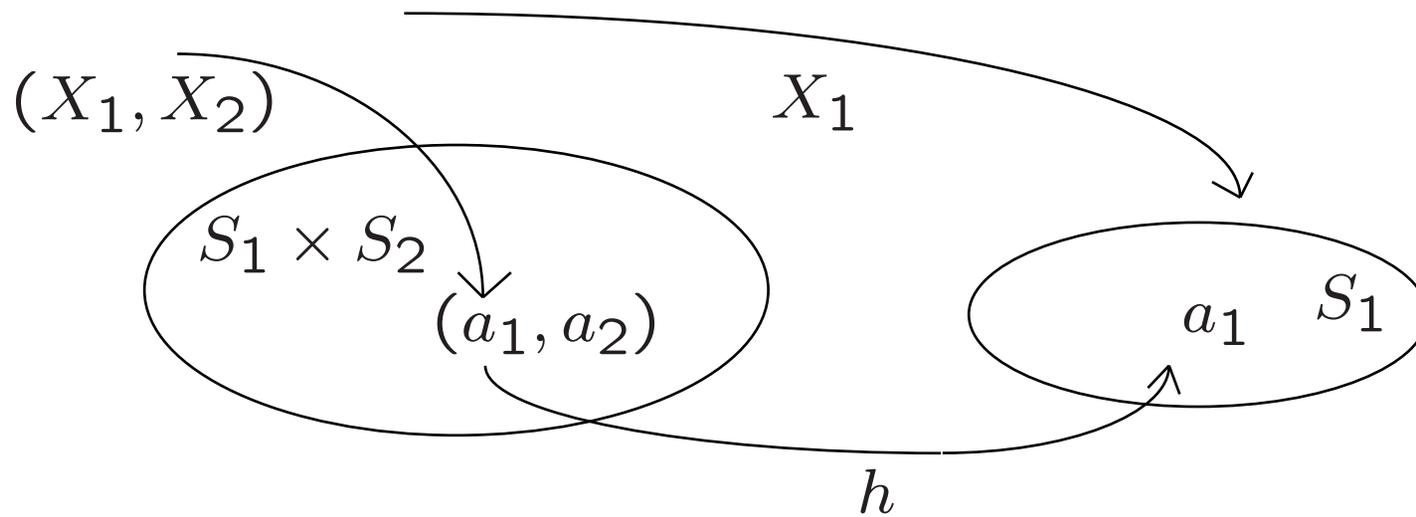
$$S = S_1 \times S_2$$

$\rho :=$ die Verteilung von (X_1, X_2)

$\rho_1 :=$ die Verteilung von X_1

$$\begin{aligned}\rho(a_1, a_2) &= \mathbf{P}((X_1, X_2) = (a_1, a_2)) \\ &= \mathbf{P}(X_1 = a_1, X_2 = a_2),\end{aligned}$$

$$\rho_1(a_1) = \sum_{a_2 \in S_2} \rho(a_1, a_2).$$



$$h((a_1, a_2)) := a_1$$

Projektion des Paares (a_1, a_2) auf seine erste Komponente

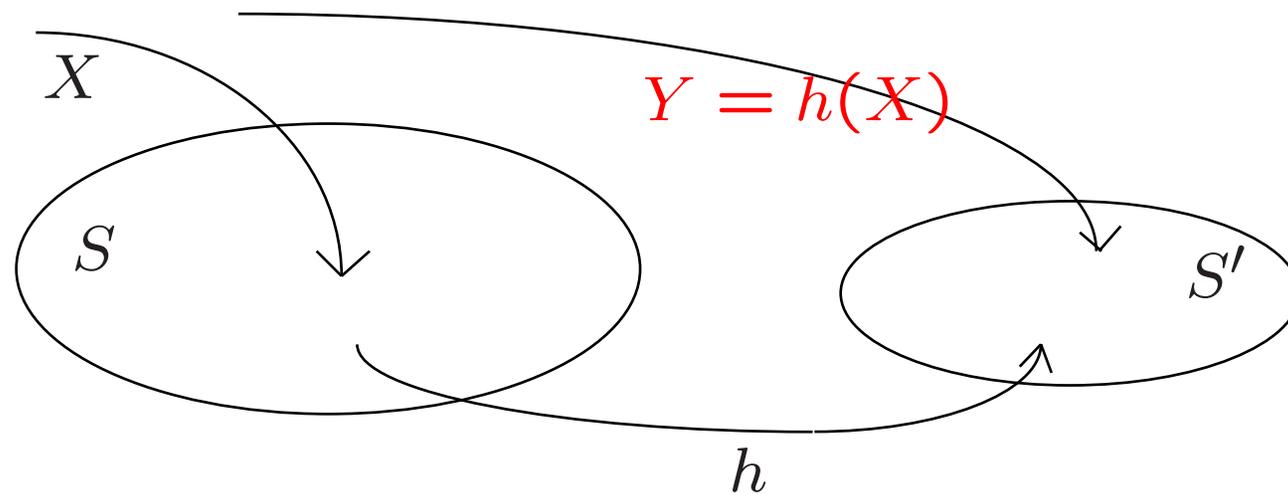
Der Übergang von $X = (X_1, X_2)$
zu einer Komponente X_1
ist ein Beispiel einer
Vergrößerung (Weiterverarbeitung) einer Zufallsvariablen:

$$Y := X_1 = h(X)$$

$$\text{mit } h((a_1, a_2)) := a_1$$

Kürzer schreiben wir: $h(a_1, a_2)$

Sind S und S' zwei Mengen,
 X eine Zufallsvariable mit Zielbereich S ,
 h eine Abbildung von S nach S' ,
und nimmt man X als zufällige Eingabe von h ,
dann bekommt man eine Zufallsvariable Y mit Zielbereich S' :

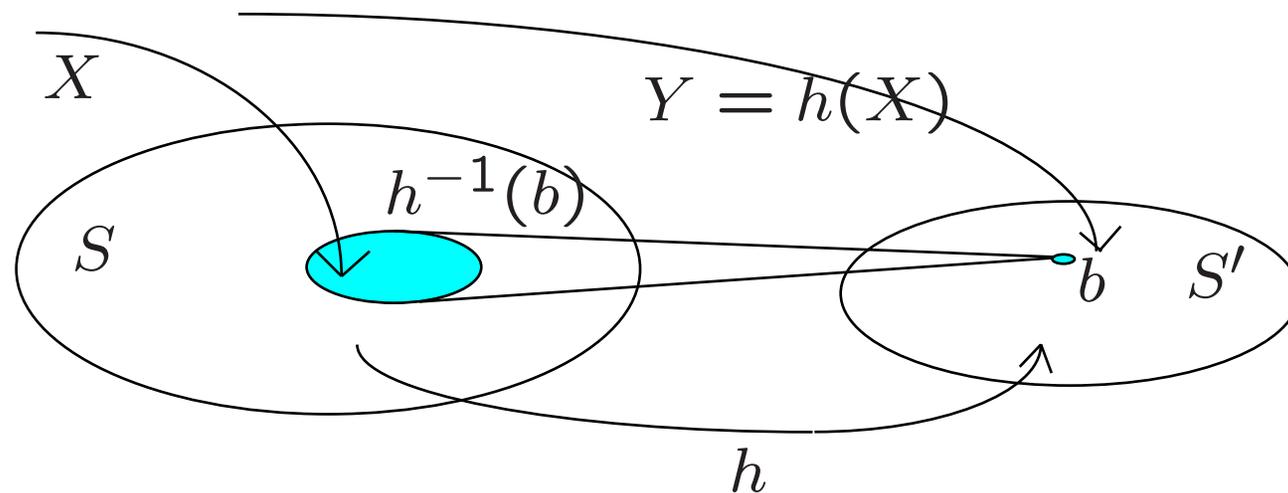


Für jedes $b \in S'$ gilt:

$$\{h(X) = b\} = \{X \in h^{-1}(b)\}$$

Für die Verteilungsgewichte von $Y = h(X)$ ergibt sich:

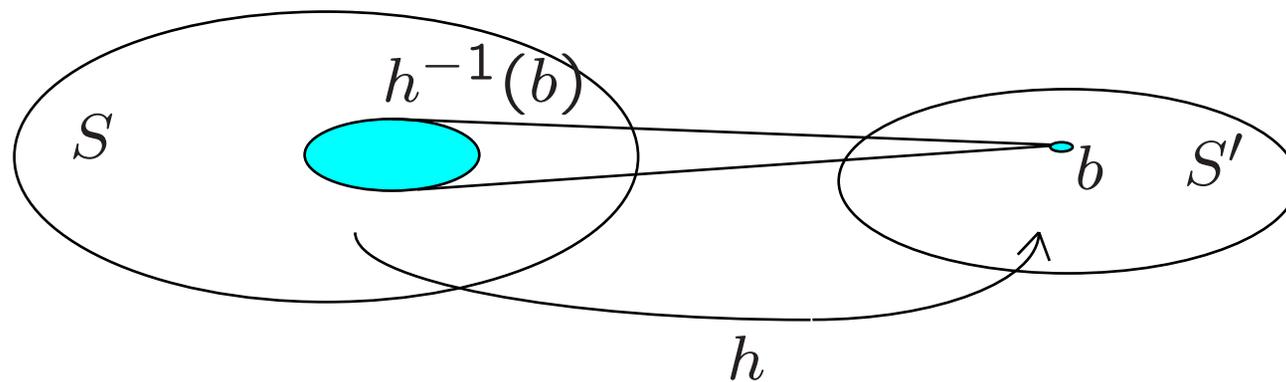
$$\mathbf{P}(Y = b) = \mathbf{P}(X \in h^{-1}(b)) = \sum_{a \in h^{-1}(b)} \mathbf{P}(X = a).$$



Bezeichnet ρ die Verteilung von X und $\tilde{\rho}$ die von Y ,
dann ist

$$\tilde{\rho}(b) = \sum_{a \in h^{-1}(b)} \rho(a).$$

Man sagt: Die Verteilung ρ wird durch die Abbildung h
in die Verteilung $\tilde{\rho}$ transportiert.



Beispiel:

$$S := \{0, 1\}^n$$

die Menge der 01-Folgen der Länge n

X sei uniform verteilt auf S

jeder Ausgang hat das Gewicht

$$\frac{1}{2^n} = \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2}$$

(ein "fairer Münzwurf")

$Y :=$ die Anzahl der Einsen in X .

Wie ist Y verteilt?

Jede einzelne 01-Folge a der Länge n mit genau k Einsen

hat Gewicht

$$\frac{1}{2^n}$$

Wieviele derartige a gibt es?

$$\binom{n}{k}$$

$$\mathbf{P}(Y = k) = \binom{n}{k} \frac{1}{2^n}, \quad k = 0, \dots, n.$$

Beispiel

n -faches Würfeln:

$X = (X_1, \dots, X_n)$ uniform verteilt auf
 $S := \{1, \dots, 6\}^n$.

$Z := (Z_1, \dots, Z_n)$, mit
 $Z_i := \mathbf{1}_{\{6\}}(X_i)$

Z ist also eine zufällige 01-Folge, mit
 $Z_i = 1$ falls der i -te Wurf eine Sechs ergibt
und $Z_i = 0$ sonst.
Wie ist Z verteilt?

$$\begin{aligned} & \mathbf{P}(Z_1 = 1, \dots, Z_k = 1, Z_{k+1} = 0, \dots, Z_n = 0) \\ &= \mathbf{P}(X_1 = 6, \dots, X_k = 6, X_{k+1} \neq 6, \dots, X_n \neq 6) \\ &= \frac{1^k \cdot 5^{n-k}}{6^n} \end{aligned}$$

$$= p^k q^{n-k},$$

$$\text{mit } p := \frac{1}{6} \text{ und } q := \frac{5}{6}.$$

Auch für jede andere Reihung von **genau k Treffern**
der Augenzahl 6 aus n Würfeln ergibt sich diese W'keit.

Beispiel:

n -maliges Ziehen *mit Zurücklegen*
aus einer ideal durchmischten Urne.

Ein Anteil p der Kugeln ist **rot**,
der restliche Anteil $q = 1 - p$ ist **blau**.

Zufällige 0-1 Folge $Z = (Z_1, \dots, Z_n)$:

$Z_i = 1$ wenn beim i -ten Zug eine rote Kugel kommt,
und $Z_i = 0$ wenn beim i -ten Zug eine blaue Kugel kommt.

Sei a eine vorgegebene 0-1 Folge der Länge n mit k Einsen,

$$\text{z. B.: } a := (\underbrace{1, \dots, 1}_{k\text{-mal}}, \underbrace{0, \dots, 0}_{(n-k)\text{-mal}})$$

$$\mathbf{P}(Z = a) = ?$$

Sei g die Gesamtanzahl der Kugeln in der Urne.

$$\mathbf{P}(Z = a) = \frac{(pg)^k (qg)^{n-k}}{g^n} = p^k q^{n-k}$$

Das ist so **für jede 0-1 Folge a mit k Einsen und $n - k$ Nullen.**

Sei a eine vorgegebene 0-1 Folge der Länge n mit k Einsen,

$$\text{z. B.: } a := (\underbrace{1, \dots, 1}_{k\text{-mal}}, \underbrace{0, \dots, 0}_{(n-k)\text{-mal}})$$

$$P(Z = a) = ?$$

Sei g die Gesamtanzahl der Kugeln in der Urne.

$$P(Z = a) = \frac{(pg)^k (qg)^{n-k}}{g^n} = p^k q^{n-k}$$

Das ist so für jede 0-1 Folge a mit k Einsen und $n - k$ Nullen.

Definition (p -Münzwurf):

Sei $p \in [0, 1]$, $q := 1 - p$.

Eine Zufallsvariable Z mit Zielbereich

$$S = \{0, 1\}^n = \{a = (a_1, \dots, a_n) : a_i \in \{0, 1\}\}$$

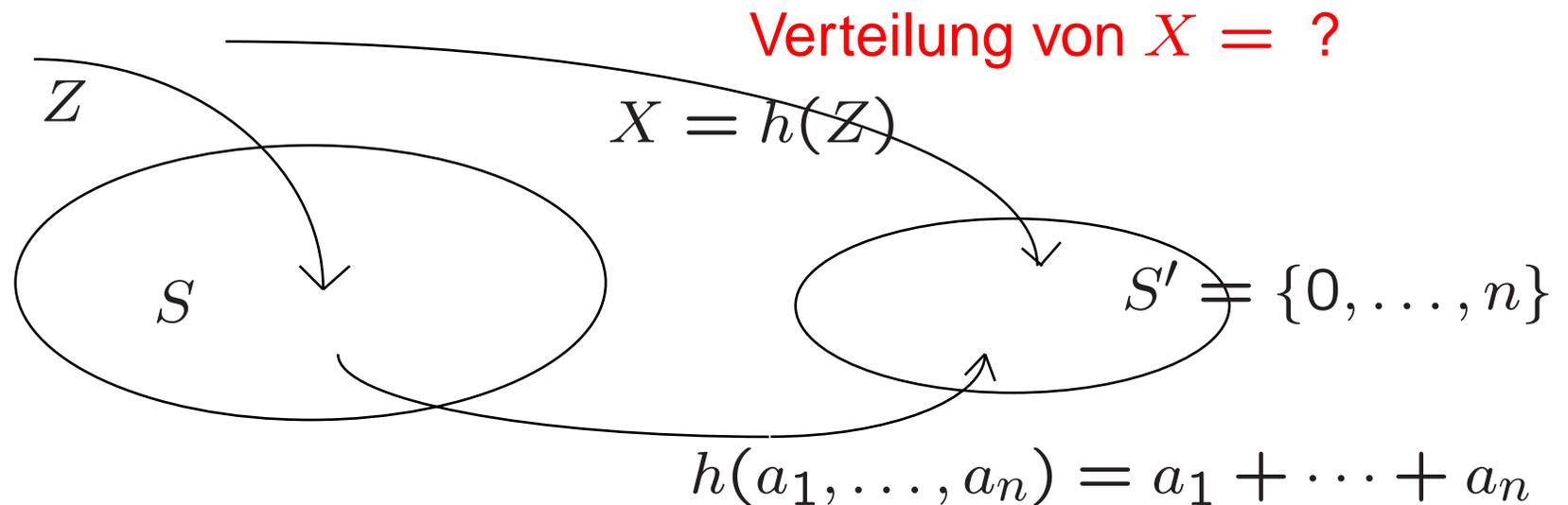
heißt **n -facher p -Münzwurf**,

wenn für alle $a \in S$ mit k Einsen und $n - k$ Nullen gilt:

$$\mathbf{P}(Z = a) = p^k q^{n-k}.$$

Ein Paradebeispiel für die
Weiterverarbeitung einer Zufallsvariablen ist die
Anzahl der Erfolge beim n -fachen p -Münzwurf:

Sei $Z = (Z_1, \dots, Z_n)$ ein n -facher p -Münzwurf
und $X = Z_1 + \dots + Z_n$ die *Anzahl der Erfolge*
(die Anzahl der Einsen in der zufälligen 0-1 Folge Z)

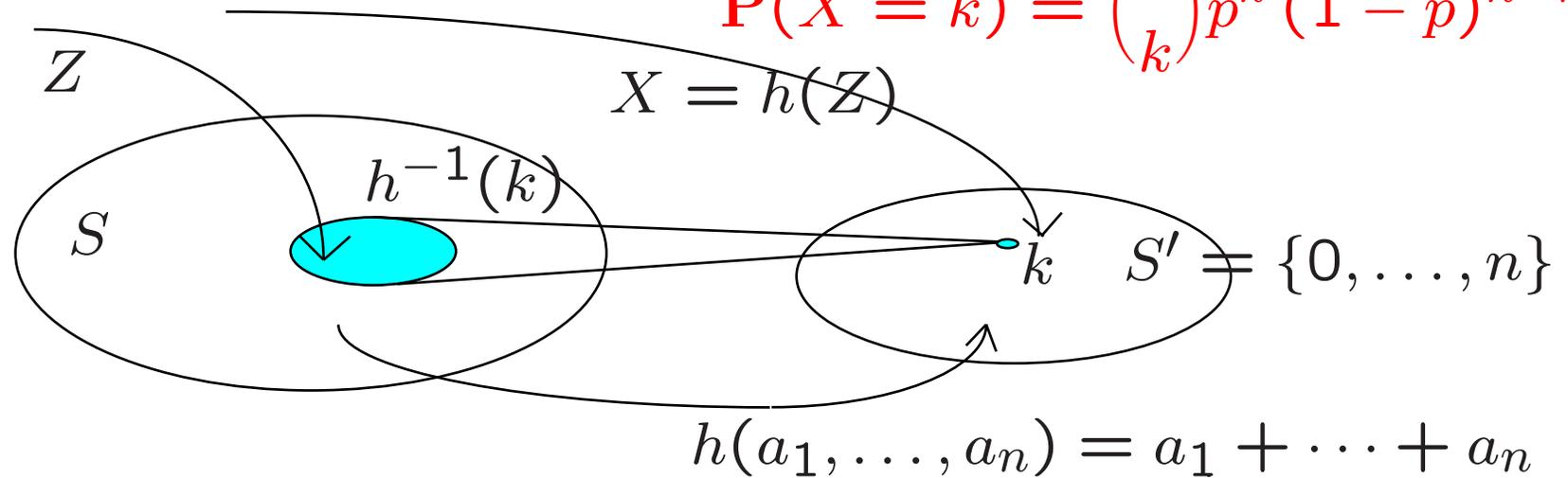


Jedes $a \in S$ mit $h(a) = k$
 (d.h. mit k Einsen und $n - k$ Nullen)

hat Gewicht $p^k (1 - p)^{n-k}$.

Es gibt $\binom{n}{k}$ solche a .

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



Definition:

Eine Zufallsvariable X mit Zielbereich $\{0, 1, \dots, n\}$
heißt *binomialverteilt* mit Parametern n und p ,

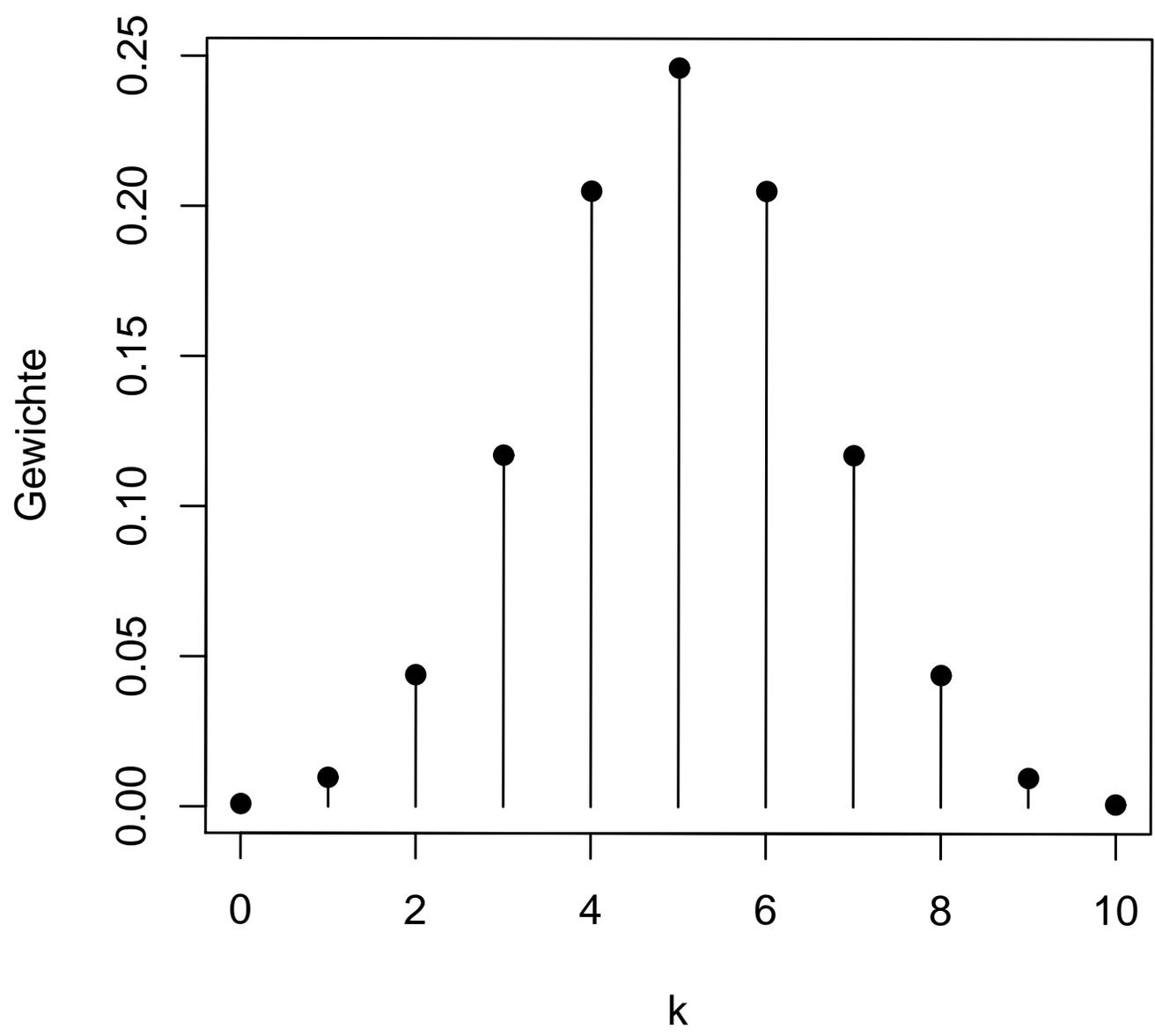
kurz

Bin(n, p)-verteilt,

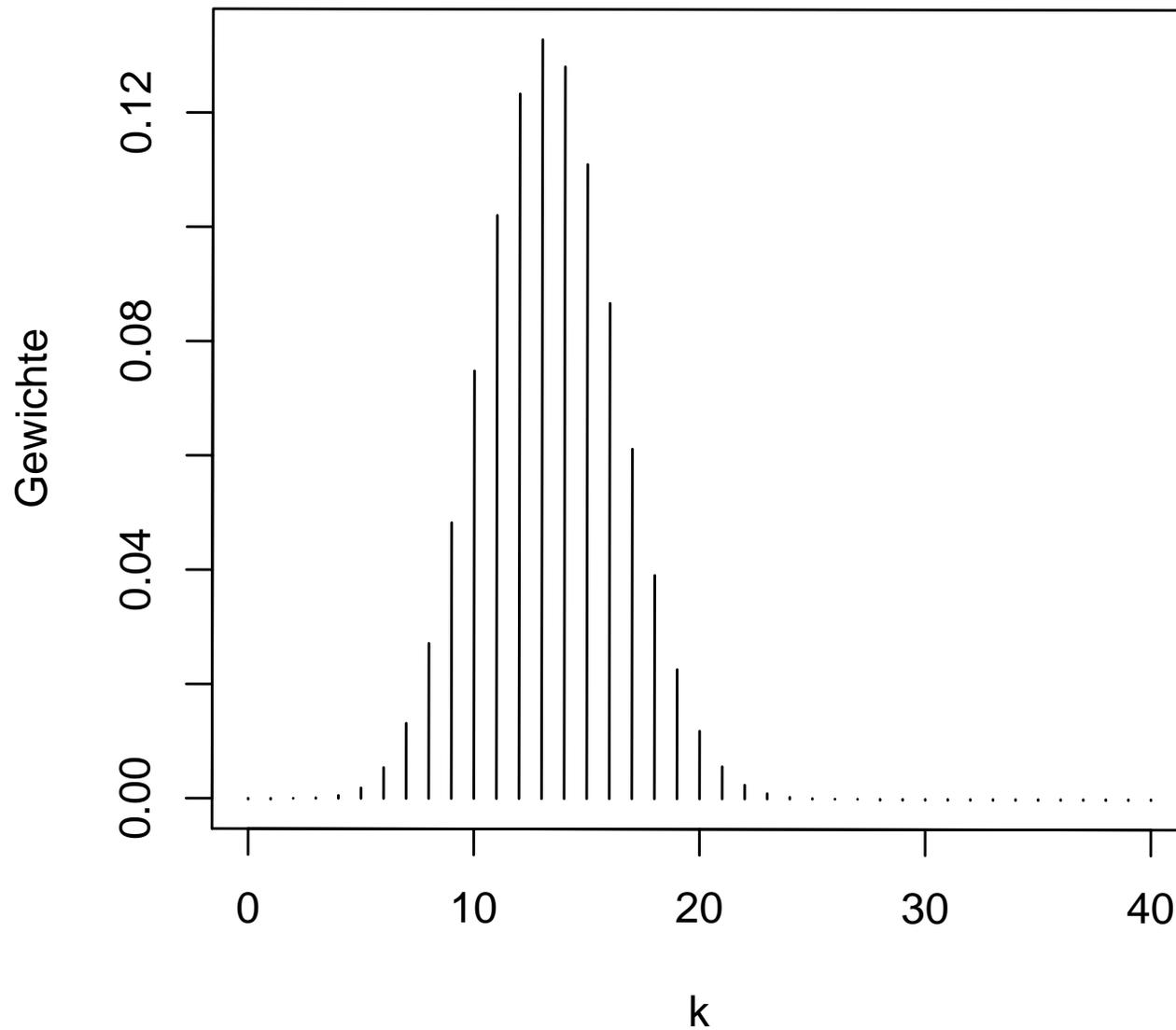
wenn

$$\mathbf{P}(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \dots, n,$$

mit $q = 1 - p$.



Gewichte der Bin(10, 1/2) Verteilung



Gewichte der Bin(40, 1/3) Verteilung

Definition (“verallgemeinertes Würfeln”):

Seien $r \in \mathbb{N}$ und $p_1, \dots, p_r \geq 0$ mit $p_1 + \dots + p_r = 1$.

Wir definieren **Gewichte** auf

$$S := \{a = (a_1, \dots, a_n) : a_i \in \{1, \dots, r\}\}$$

durch

$$\rho(a_1, \dots, a_n) := p_{a_1} \cdot p_{a_2} \cdots p_{a_n}.$$

Eine Zufallsvariable Z mit diesem Zielbereich S und diesen Verteilungsgewichten ρ nennen wir

n -faches (p_1, \dots, p_r) -Würfeln.

Für jedes $a \in S$ mit
 k_1 Komponenten gleich 1,
 k_2 Komponenten gleich 2,
...
 k_r Komponenten gleich r

ist dann

$$\mathbf{P}(Z = a) = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

Beispiel: Besetzung der Ergebnisse beim “Würfeln”:

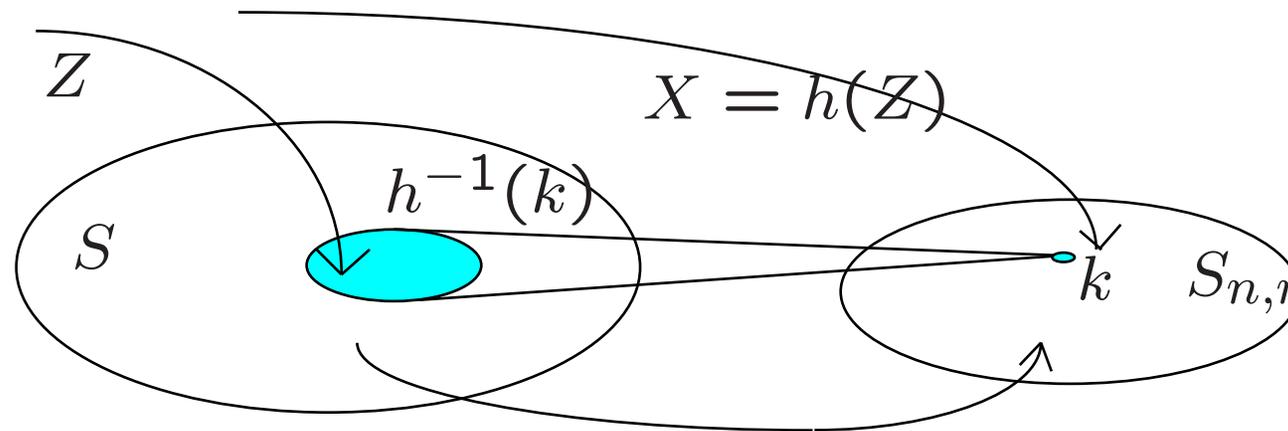
$$X_j := \#\{i : Z_i = j\}$$

(die Anzahl der Würfe mit Ergebnis j).

$X := (X_1, \dots, X_r)$ hat Zielbereich

$$S_{n,r} = \{(k_1, \dots, k_r) : k_1 + \dots + k_r = n\}.$$

Verteilung von $X = ?$



$$h(a_1, \dots, a_n) = (k_1, \dots, k_r) =: k$$

mit $k_j := \#\{i : a_i = j\}$, $j = 1, \dots, r$

Jedes $a \in S$ mit $h(a) = (k_1, \dots, k_r)$

hat Gewicht $p_1^{k_1} \dots p_r^{k_r}$

Wieviele solche a gibt es?

Dazu überlegen wir:

Auf wieviele Arten kann man
 n Objekte so auf r Fächer verteilen,
dass das j -te Fach genau k_j Objekte enthält?

Dabei ist $k_1 + \dots + k_r = n$.

$$\binom{n}{k_1} \cdot \binom{n - k_1}{k_2} \cdots \binom{n - k_1 - \dots - k_{r-1}}{k_r}$$

$$= \frac{n!}{k_1! k_2! \cdots k_r!}$$

Auf wieviele Arten kann man
 n Objekte so auf r Fächer verteilen,
dass das j -te Fach genau k_j Objekte enthält?

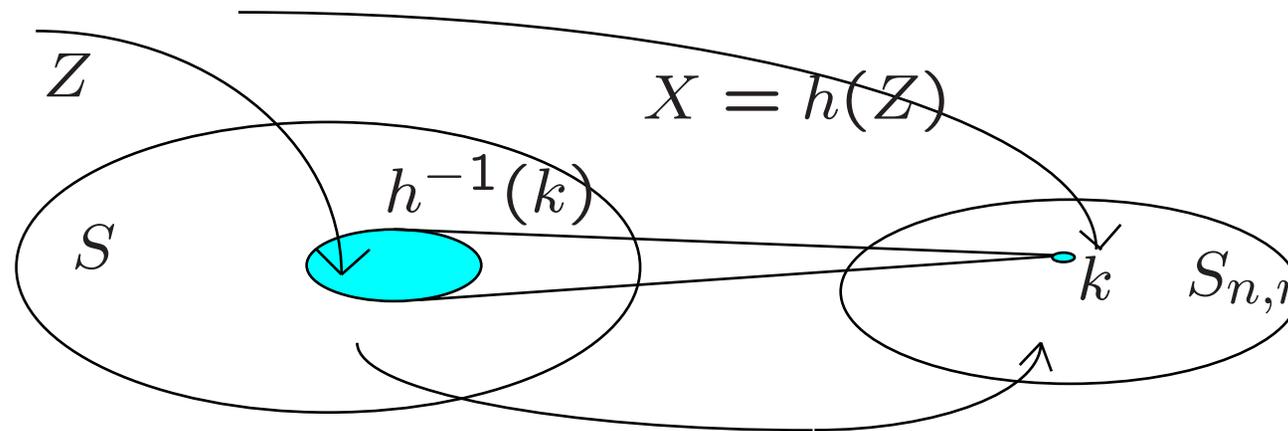
Dabei ist $k_1 + \dots + k_r = n$.

Die Antwort ist:

$$\binom{n}{k_1} \cdot \binom{n - k_1}{k_2} \cdots \binom{n - k_1 - \dots - k_{r-1}}{k_r}$$

$$= \frac{n!}{k_1! k_2! \cdots k_r!} =: \binom{n}{k_1, \dots, k_r}$$

Multinomialkoeffizient, lies: n über k_1, \dots, k_r

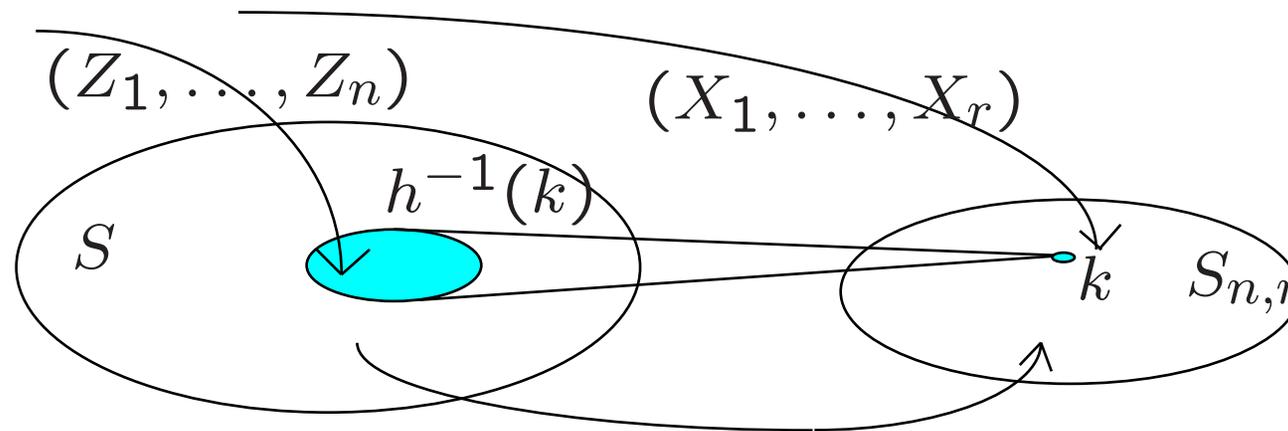


$$h(a_1, \dots, a_n) = (k_1, \dots, k_r) = k$$

Jedes $a \in S$ mit $h(a) = (k_1, \dots, k_r)$

hat Gewicht $p_1^{k_1} \dots p_r^{k_r}$

Wieviele solche a gibt es?



$$h(a_1, \dots, a_n) = (k_1, \dots, k_r)$$

Jedes $a \in S$ mit $h(a) = (k_1, \dots, k_r)$

hat Gewicht $p_1^{k_1} \dots p_r^{k_r}$

Es gibt $\binom{n}{k_1, \dots, k_r}$ solche a .

$$\mathbf{P}(X_1 = k_1, \dots, X_r = k_r) = \binom{n}{k_1, \dots, k_r} p_1^{k_1} \dots p_r^{k_r}$$

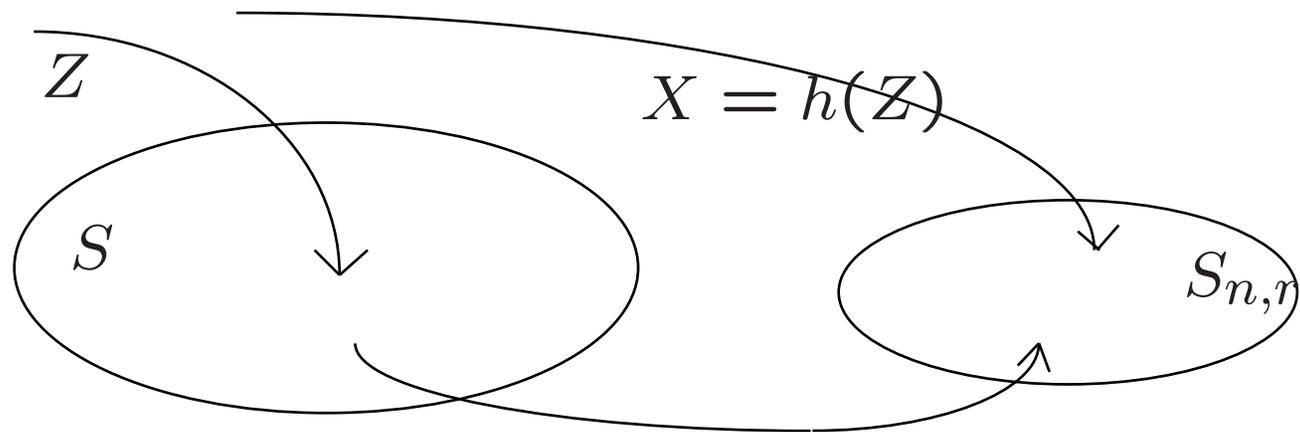
Definition:

Eine Zufallsvariable X mit Zielbereich $S_{n,r}$
heißt *multinomialverteilt* mit Parametern $n; p_1, \dots, p_r$,

wenn

$$\mathbf{P}(X = (k_1, \dots, k_r)) = \binom{n}{k_1, \dots, k_r} p_1^{k_1} \cdots p_r^{k_r},$$

$$(k_1, \dots, k_r) \in S_{n,r} .$$



$$h(a_1, \dots, a_n) = (k_1, \dots, k_r) =: k$$

mit $k_j := \#\{i : a_i = j\}$, $j = 1, \dots, r$

Zwei wichtige Beispiele aus der heutigen Vorlesung

Stichprobenziehen → Münzwurf → Binomialverteilung

$(r = 2)$

Stichprobenziehen → Würfeln → Multinomialverteilung

$(r = 6)$

werden illustriert durch die zwei Wandtner'schen
R-Programme zu VL 2b.