Vorlesung 2b

Diskrete Zufallsvariable

und ihre Verteilungen

Eine Zufallsvariable heißt **diskret**,

Eine Zufallsvariable heißt **diskret**,

falls ihr Zielbereich endlich oder abzählbar unendlich ist

Eine Zufallsvariable heißt **diskret**,

falls ihr Zielbereich endlich oder abzählbar unendlich ist

oder (allgemeiner) eine abzählbare Menge S enthält mit

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

Wir erinnern an die beiden Grundregeln

für Wahrscheinlichkeiten:

Wir erinnern an die beiden Grundregeln

für Wahrscheinlichkeiten:

Additivität:

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

Wir erinnern an die beiden Grundregeln

für Wahrscheinlichkeiten:

Additivität:

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

Normiertheit auf Eins:

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

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

heißt die **Verteilung** von X.

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

heißt die **Verteilung** von X.

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

sind die **Verteilungsgewichte**.

Im Fall

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

nennt man die Verteilung ρ von X auch

die *gemeinsame Verteilung* von $X_1, \ldots, X_n$.

Im Fall

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

nennt man die Verteilung ρ von X auch

die *gemeinsame Verteilung* von $X_1, \ldots, X_n$.

Die Verteilungen $\rho_1, \ldots, \rho_n$ von $X_1, \ldots, X_n$ nennt man

die *Rand*- oder *Marginalverteilungen* von ρ.

Beispiel: $S = S_1 \times S_2$

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

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

Der Übergang von $X = (X_1, X_2)$

zu einer Komponente X_1

ist ein Beispiel einer

Vergröberung (*Weiterverarbeitung*) einer Zufallsvariablen:

Der Übergang von $X = (X_1, X_2)$

zu einer Komponente X_1

ist ein Beispiel einer

Vergröberung (*Weiterverarbeitung*) einer Zufallsvariablen:

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

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

Der Übergang von $X = (X_1, X_2)$

zu einer Komponente X_1

ist ein Beispiel einer

Vergröberung (*Weiterverarbeitung*) einer Zufallsvariablen:

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

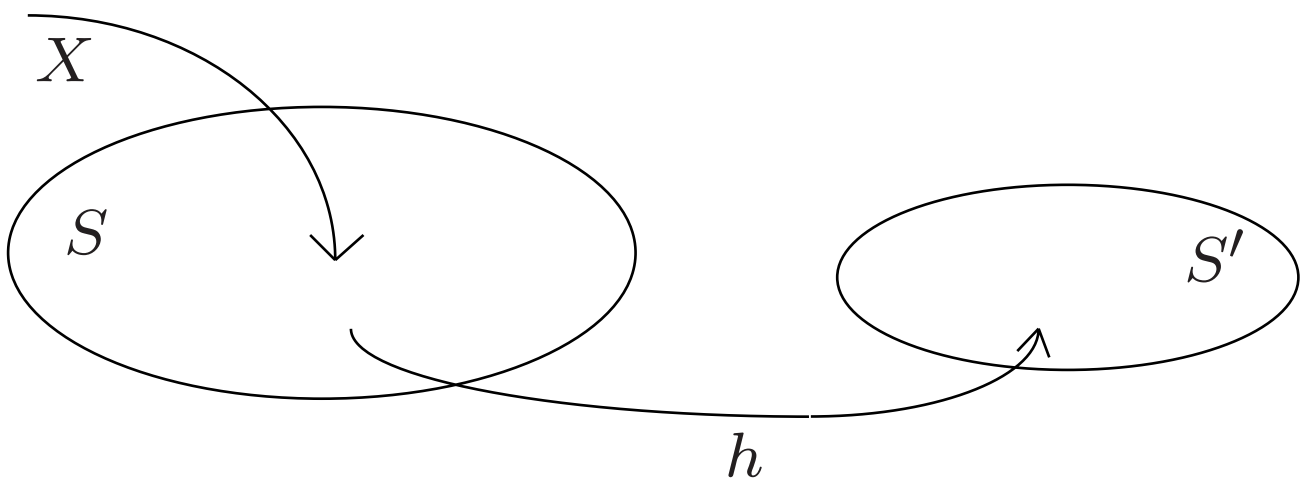
mit $h((a_1, a_2)) := a_1$

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',



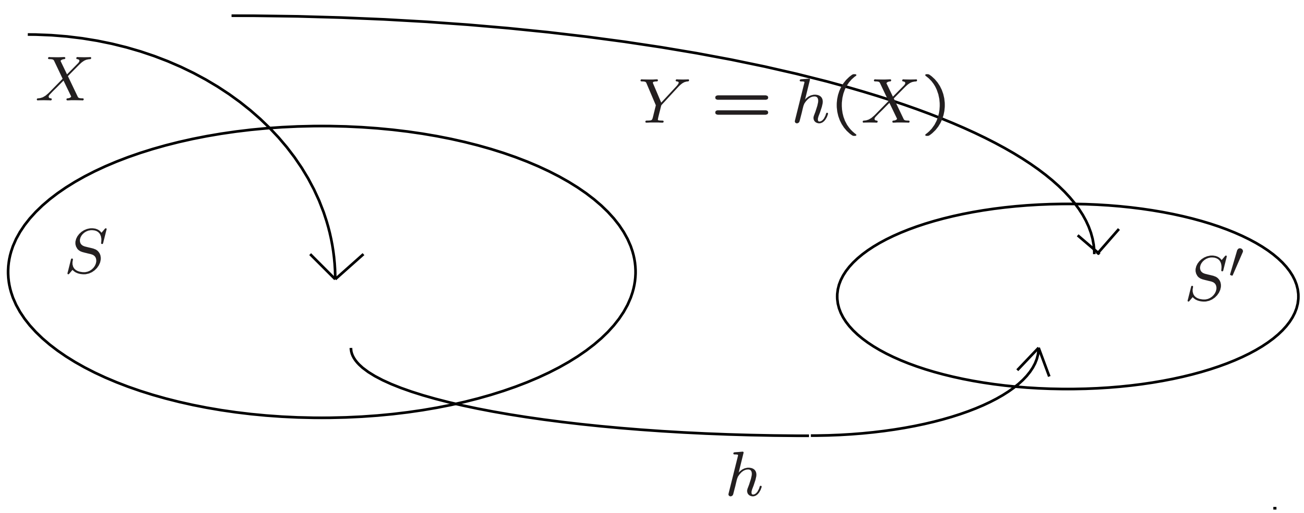
Sind S und S' zwei Mengen,

X eine Zufallsvariable mit Zielbereich S,

h eine Abbildung von S nach S',

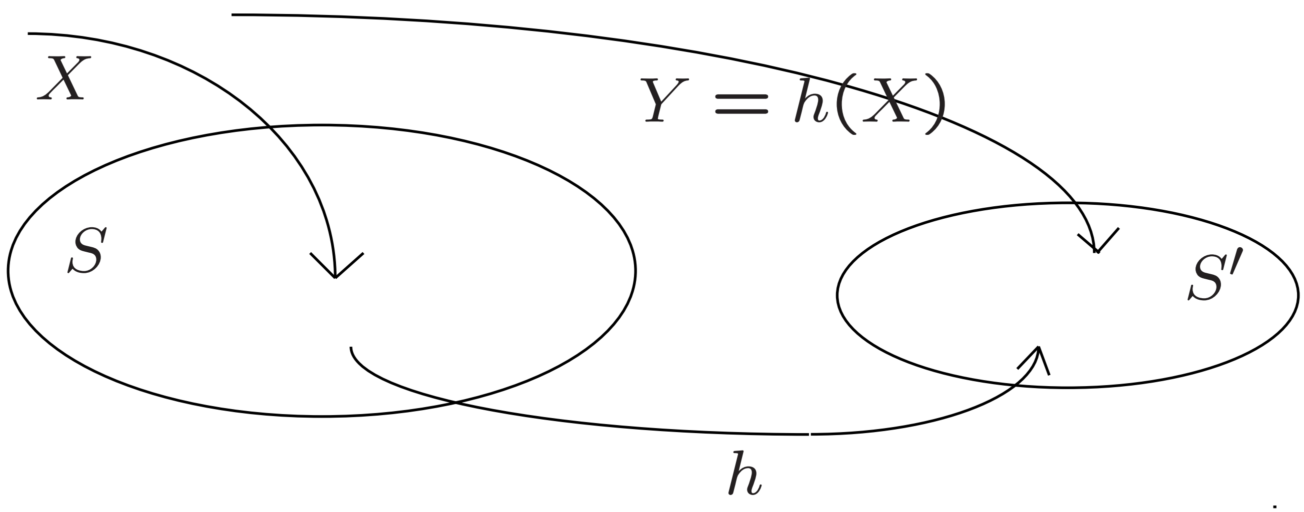
und nimmt man X als zufällige Eingabe von h,

dann bekommt man eine Zufallsvariable Y mit Zielbereich S':



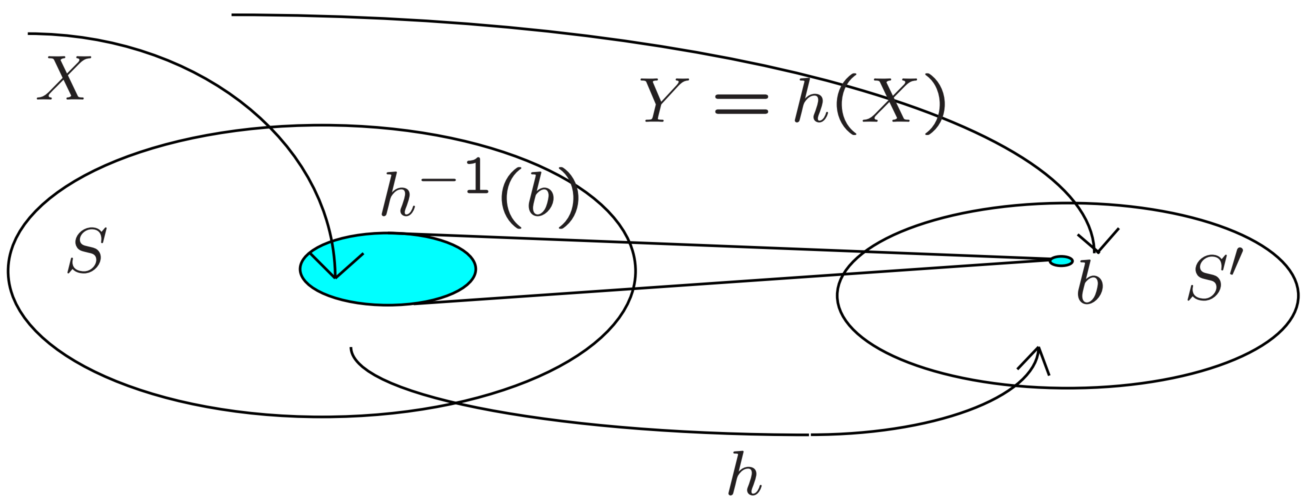
Für jedes $b \in S'$ gilt:

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



Für jedes $b \in S'$ gilt:

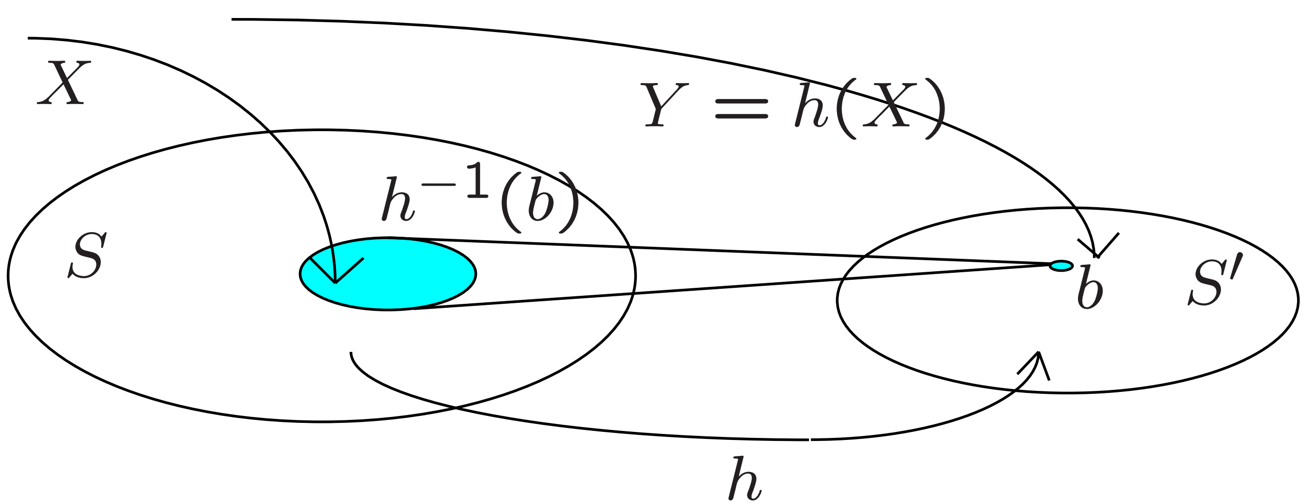
$$\{h(X) = b\} = \{X \in h^{-1}(b)\}$$



Für jedes $b \in S'$ gilt:

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

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

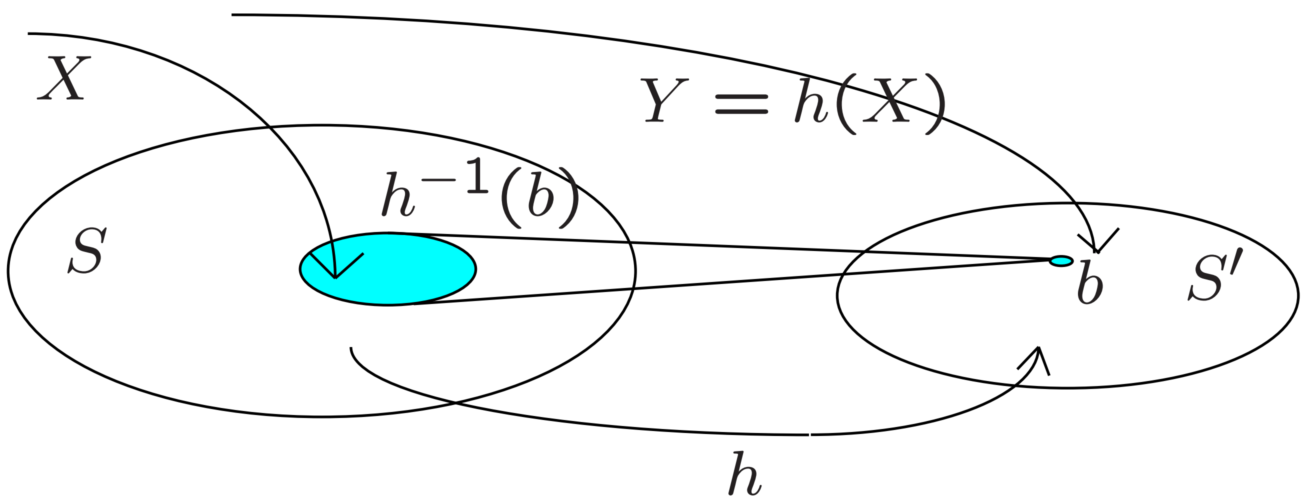


Für jedes $b \in S'$ gilt:

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

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

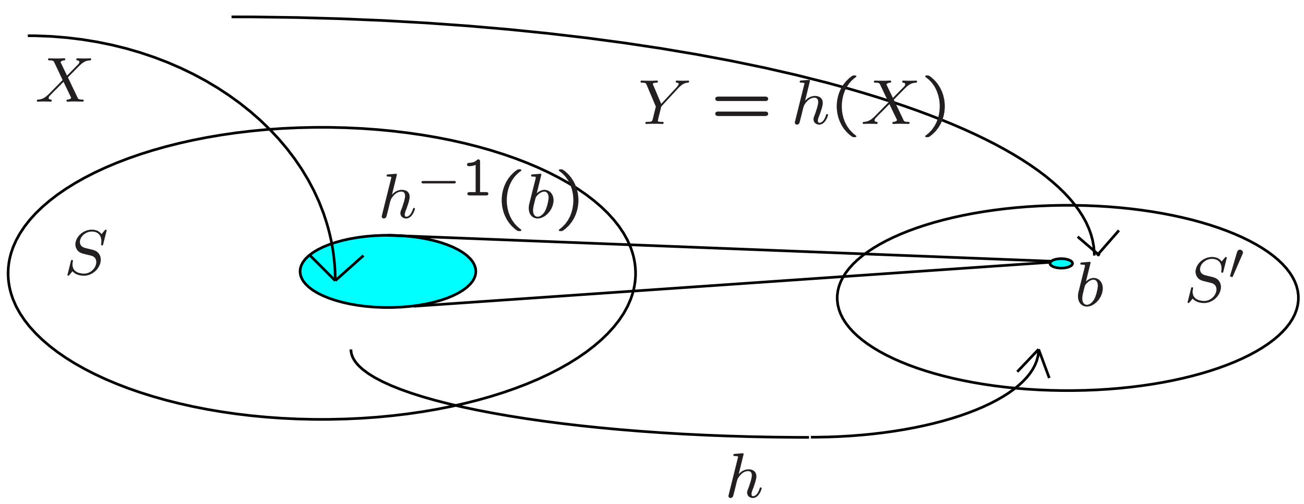
$$\mathbf{P}(Y = b) = \mathbf{P}(X \in h^{-1}(b)) = \sum_{a \in h^{-1}(b)} \mathbf{P}(X = a).$$



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

dann ist

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

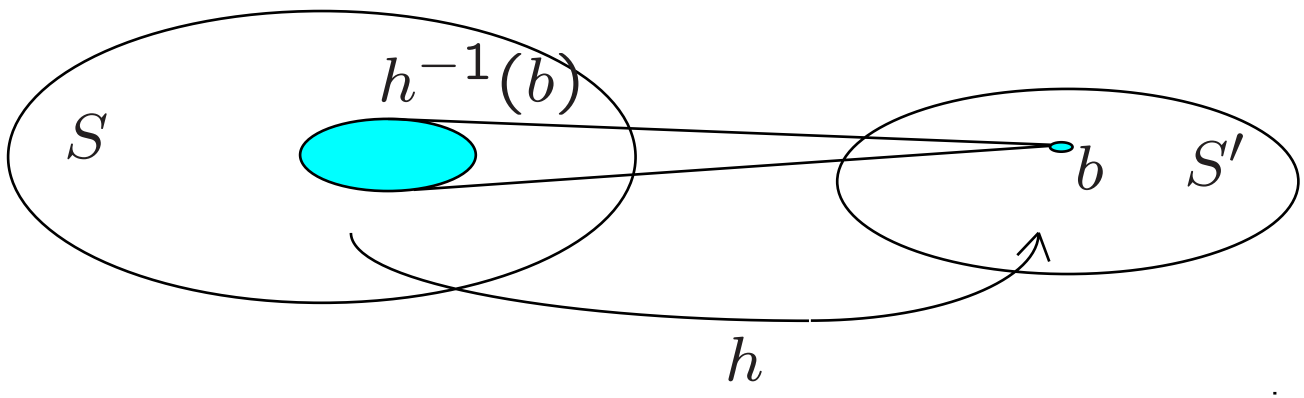


Bezeichnet ρ die Verteilung von X und $\tilde{\rho}$ die von Y,

dann ist

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

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



Beispiel:

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

die Menge der 01-Folgen der Länge n

Beispiel:

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

die Menge der 01-Folgen der Länge n

X sei uniform verteilt auf S

ein “fairer Münzwurf”: jeder Ausgang hat das Gewicht

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

Beispiel:

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

die Menge der 01-Folgen der Länge n

X sei uniform verteilt auf S

ein “fairer Münzwurf”: jeder Ausgang hat das Gewicht

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

$Y :=$ 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?

Antwort:

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

hat Gewicht

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

Wieviele derartige a gibt es?

Antwort: $\binom{n}{k}$

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

hat Gewicht

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

Wieviele derartige a gibt es?

Antwort: $\binom{n}{k}$

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

Beispiel

n-faches Würfeln:

$X = (X_1, \ldots, X_n)$ uniform verteilt auf

$$S := \{1, \ldots, 6\}^{\{1,\ldots,n\}}.$$

$X = (X_1, \ldots, X_n)$ uniform verteilt auf

$$S := \{1, \ldots, 6\}^{\{1,\ldots,n\}}.$$

$Z := (Z_1, \ldots, Z_n)$, mit

$$Z_i := \mathbf{1}_{\{6\}}(X_i)$$

$X = (X_1, \dots, X_n)$ uniform verteilt auf

$S := \{1, \dots, 6\}^{\{1,\dots,n\}}$.

$Z := (Z_1, \dots, Z_n)$, 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.

$X = (X_1, \ldots, X_n)$ uniform verteilt auf

$S := \{1, \ldots, 6\}^{\{1,\ldots,n\}}$.

$Z := (Z_1, \ldots, Z_n)$, 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?

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

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

$$\mathbf{P}(Z_1 = 1, \ldots, Z_k = 1,\ Z_{k+1} = 0, \ldots, Z_n = 0)$$

$$= \mathbf{P}(X_1 = 6, \ldots, X_k = 6,\ X_{k+1} \neq 6, \ldots, X_n \neq 6)$$

$$= \frac{1^k \cdot 5^{n-k}}{6^n}$$

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

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

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

Beispiel:

n-maliges Ziehen *mit Zurücklegen*

aus einer ideal durchmischten Urne.

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.

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 $Z = (Z_1, \ldots, Z_n)$:

$Z_i = 1$ 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 vorgegebe 0-1 Folge der Länge n mit k Einsen,

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

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

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

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

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

Sei g die Gesamtanzahl der Kugeln in der Urne.

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

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

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

Sei g die Gesamtanzahl der Kugeln in der Urne.

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

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

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

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

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 vorgegebe 0-1 Folge der Länge n mit k Einsen,

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

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

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.

Definition (p-Münzwurf):

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

Eine Zufallsvariable Z mit Zielbereich

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

heißt n-facher p-Münzwurf,

Definition (p-Münzwurf):

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

Eine Zufallsvariable Z mit Zielbereich

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

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 $Z = (Z_1, \ldots, Z_n)$ ein n-facher p-Münzwurf

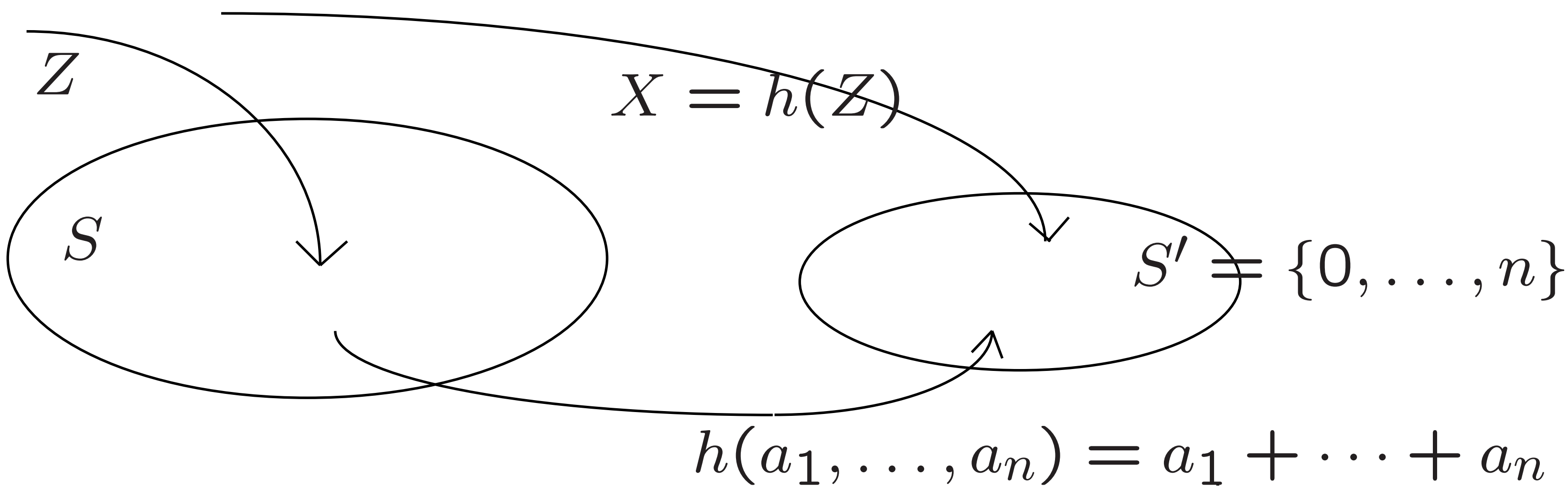
und $X = Z_1 + \cdots + Z_n$ die *Anzahl der Erfolge*

(die Anzahl der Einsen in der zufälligen 0-1 Folge Z)

Sei $Z = (Z_1, \ldots, Z_n)$ ein n-facher p-Münzwurf

und $X = Z_1 + \cdots + Z_n$ die *Anzahl der Erfolge*

(die Anzahl der Einsen in der zufälligen 0-1 Folge Z)

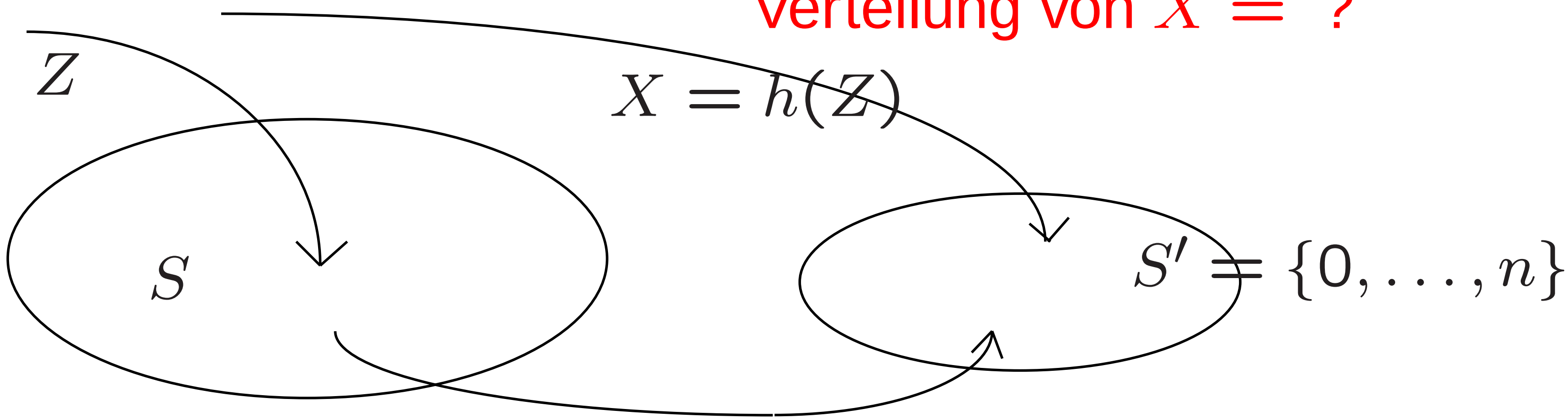


Sei $Z = (Z_1, \ldots, Z_n)$ ein n-facher p-Münzwurf

und $X = Z_1 + \cdots + Z_n$ die *Anzahl der Erfolge*

(die Anzahl der Einsen in der zufälligen 0-1 Folge Z)

Verteilung von $X =$?

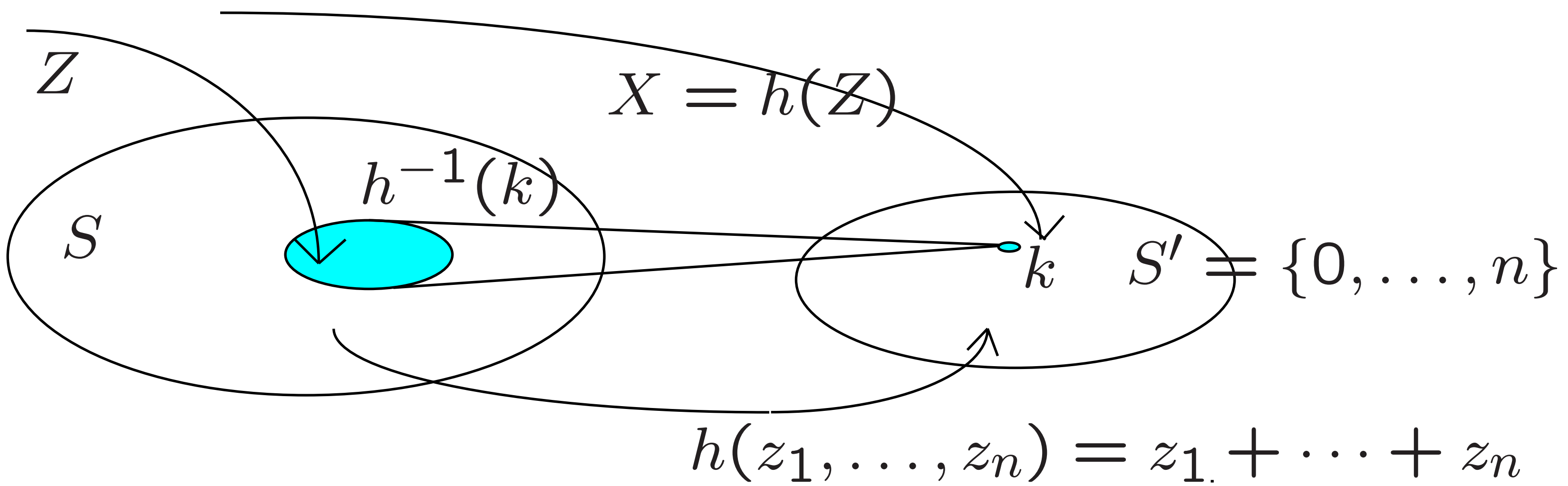


$h(a_1, \ldots, a_n) = a_1 + \cdots + a_n$

Jedes $a \in S$ mit $h(a) = k$

(d.h. mit k Einsen und $n - k$ Nullen)

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

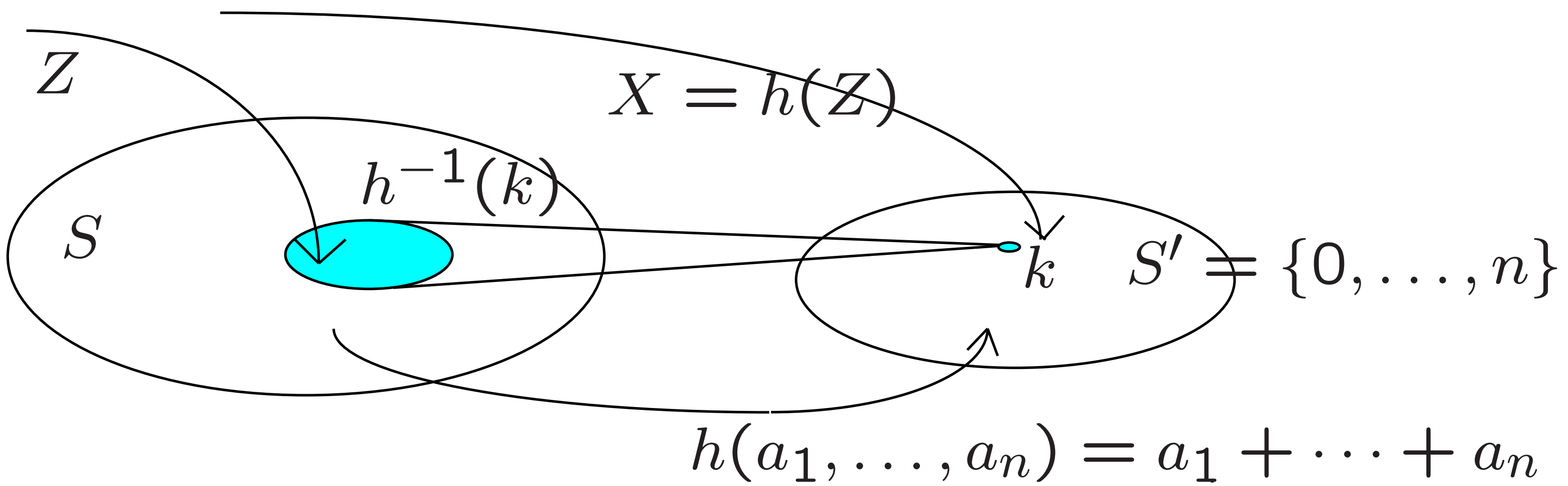


Jedes $a \in S$ mit $h(a) = k$

(d.h. mit k Einsen und $n - k$ Nullen)

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

Wieviele solche a gibt es?

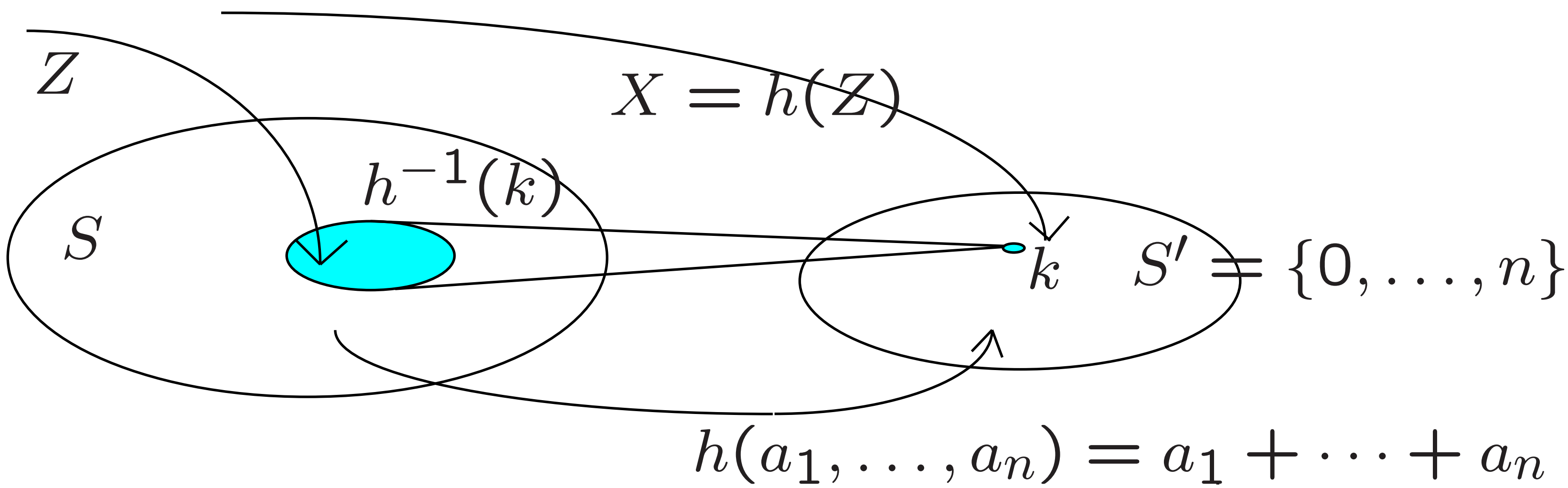


Jedes $a \in S$ mit $h(a) = k$

(d.h. mit k Einsen und $n - k$ Nullen)

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

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



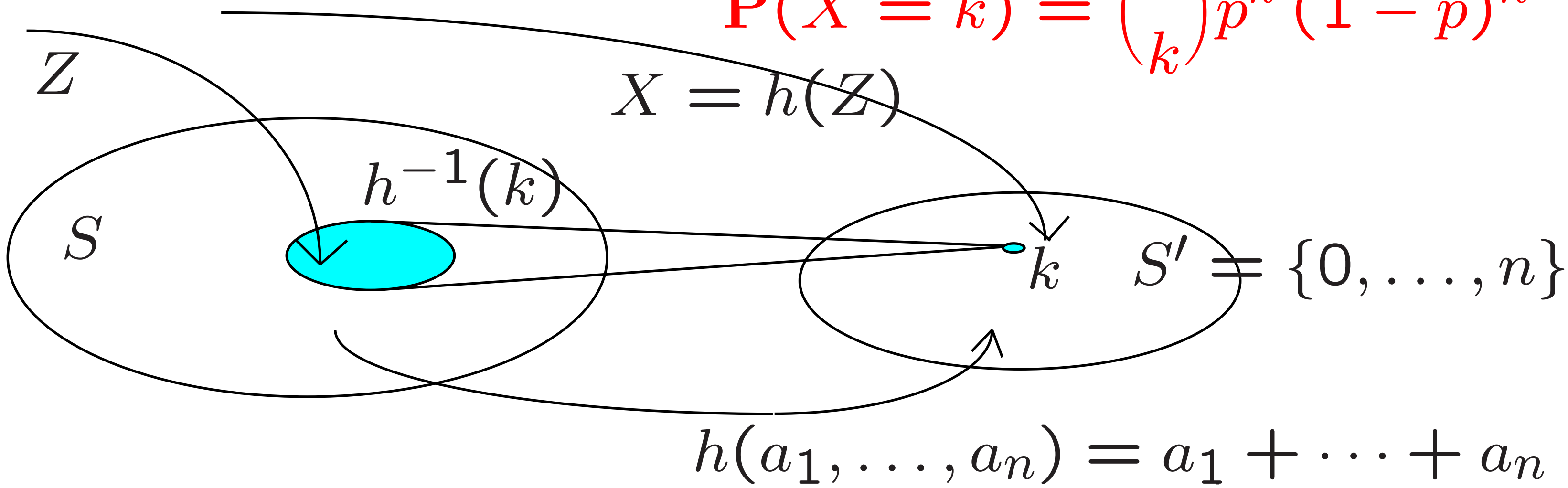
Jedes $a \in S$ mit $h(a) = k$

(d.h. mit k Einsen und $n - k$ Nullen)

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

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

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



Definition:

Eine Zufallsvariable X mit Zielbereich $\{0, 1, \ldots, n\}$

heißt *binomialverteilt* mit Parametern n und p,

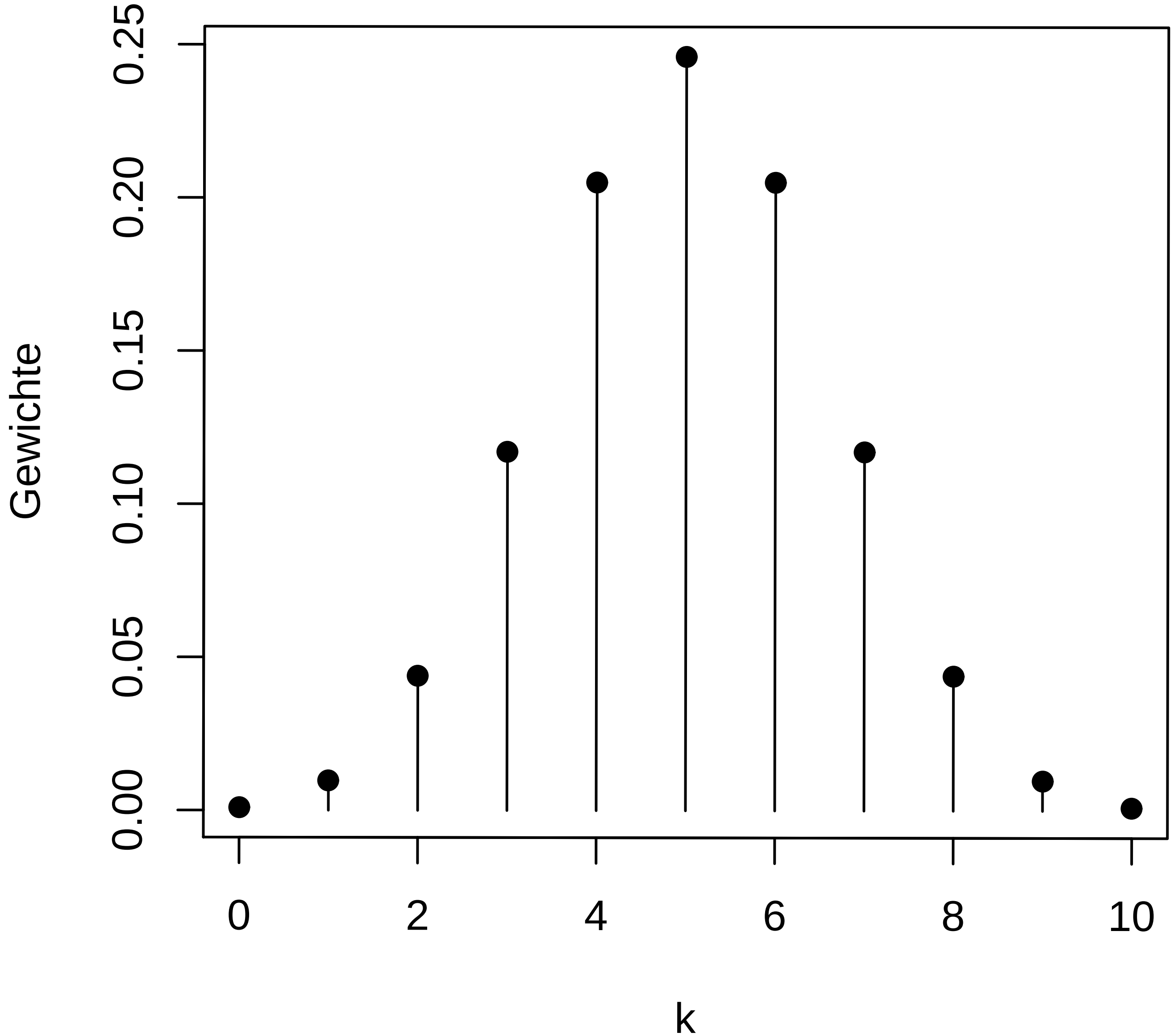
kurz

Bin(n, p)-verteilt,

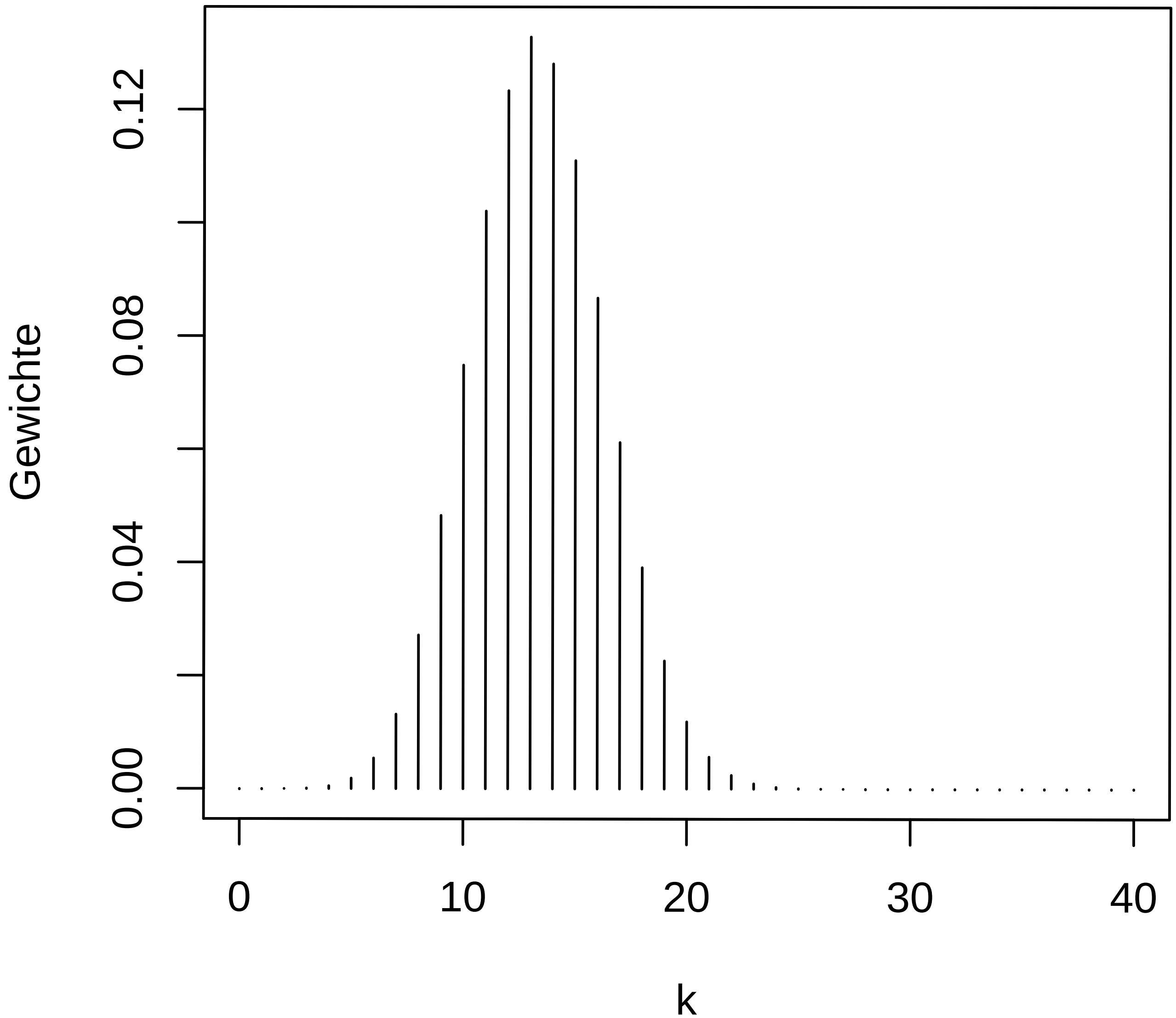
wenn

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

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 $p_1, \ldots, p_r \geq 0$ mit $p_1 + \ldots + p_r = 1$.

Definition ("verallgemeinertes Würfeln"):

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

Wir definieren Gewichte auf

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

durch

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

Definition ("verallgemeinertes Würfeln"):

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

Wir definieren Gewichte auf

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

durch

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

Eine Zufallsvariable Z mit diesem Zielbereich S

und diesen Verteilungsgewichten ρ nennen wir

n-faches $(p_1, \ldots, p_r)$-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).

Beispiel: Besetzung der Ergebnisse beim “Würfeln”:

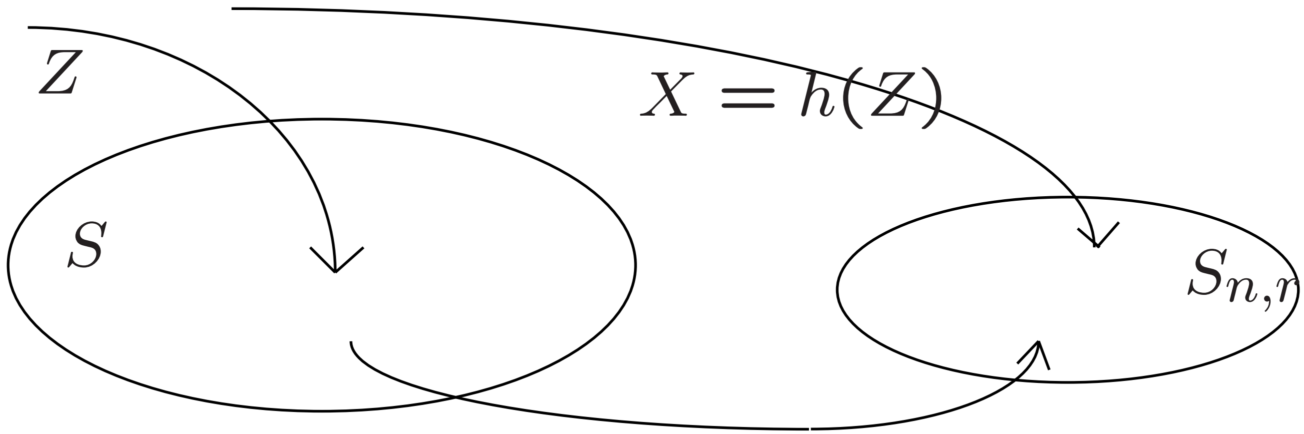
$$X_j := \#\{i \ : \ Z_i = j\}$$

(die Anzahl der Würfe mit Ergebnis j).

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

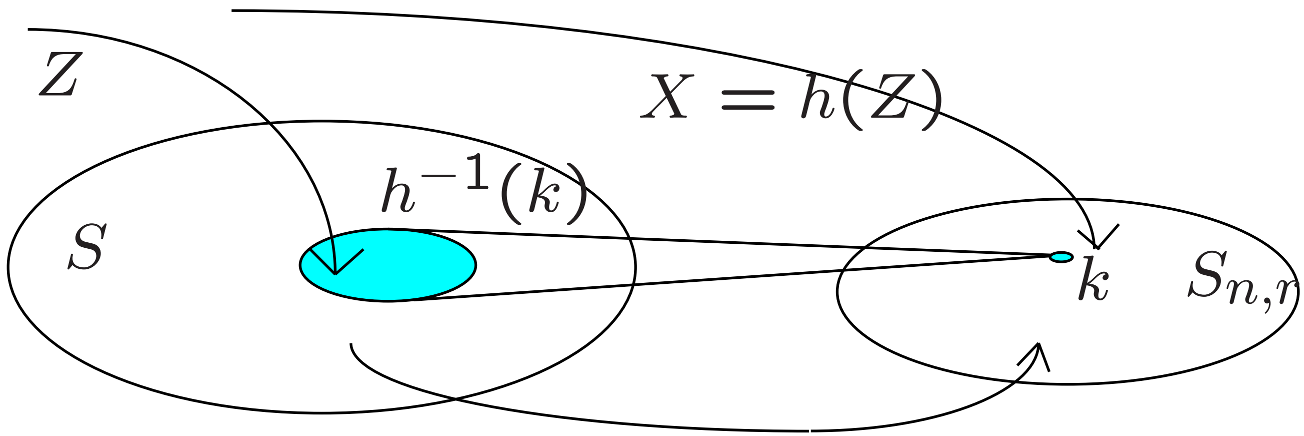
$$S_{n,r} = \{(k_1, \ldots, k_r) \ : k_1 + \ldots + k_r = n\}.$$

Verteilung von X = ?



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

mit $k_j := \#\{i : a_i = j\}), \; j = 1, \ldots, r$



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

mit $k_j := \#\{i : a_i = j\}), \; j = 1, \ldots, r$

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

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

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 $k_1 + \cdots + k_r = n$.

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

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

Auf wieviele Arten kann man

n Objekte so auf r Fächer verteilen,

dass das j-te Fach genau k_j Objekte enthält?

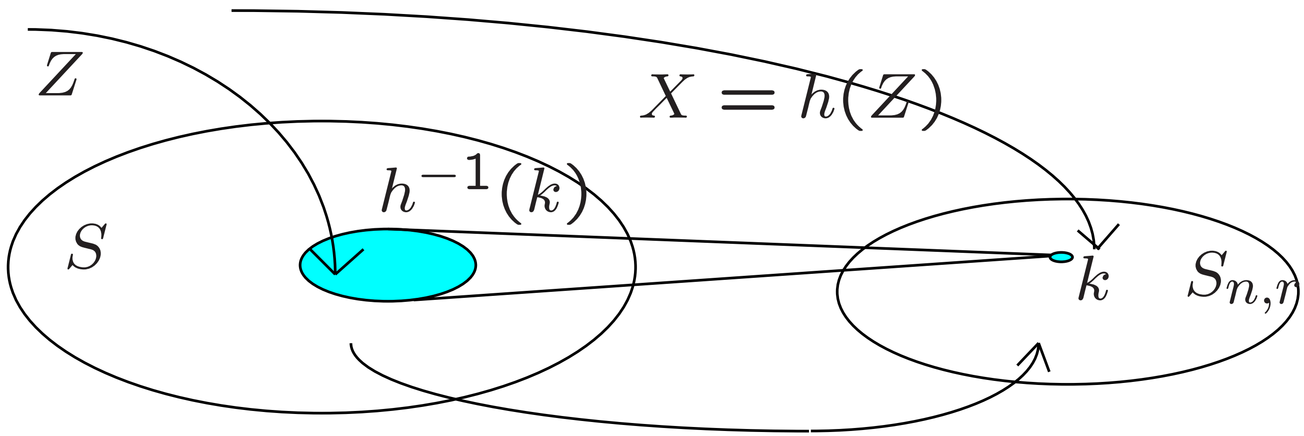
Dabei ist $k_1 + \cdots + k_r = n$.

Die Antwort ist:

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

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

Multinomialkoeffizient, lies: n über $k_1,\ldots,k_r$

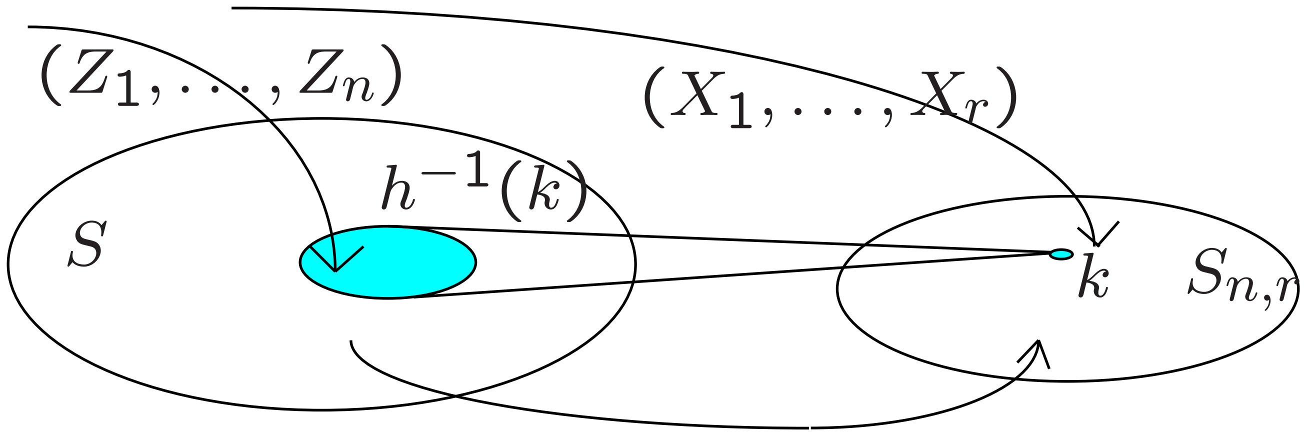


$$h(a_1, \ldots, a_n) = (k_1, \ldots, k_r) = k$$

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

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

Wieviele solche a gibt es?



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

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

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

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

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

Definition:

Eine Zufallsvariable X mit Zielbereich $S_{n,r}$

heißt *multinomialverteilt* mit Parametern $n; p_1, \ldots, p_r$,

Definition:

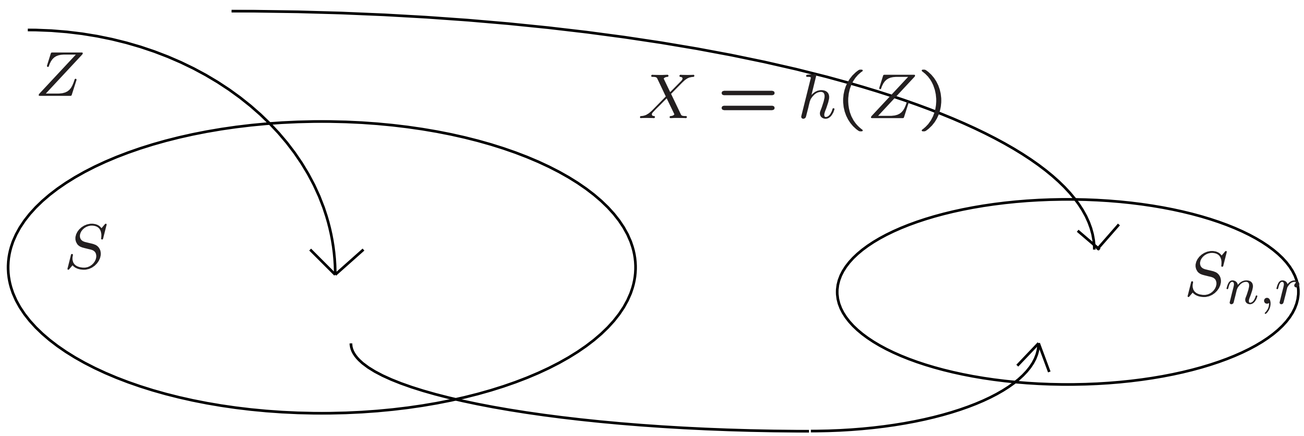
Eine Zufallsvariable X mit Zielbereich $S_{n,r}$

heißt *multinomialverteilt* mit Parametern $n; p_1, \ldots, p_r$,

wenn

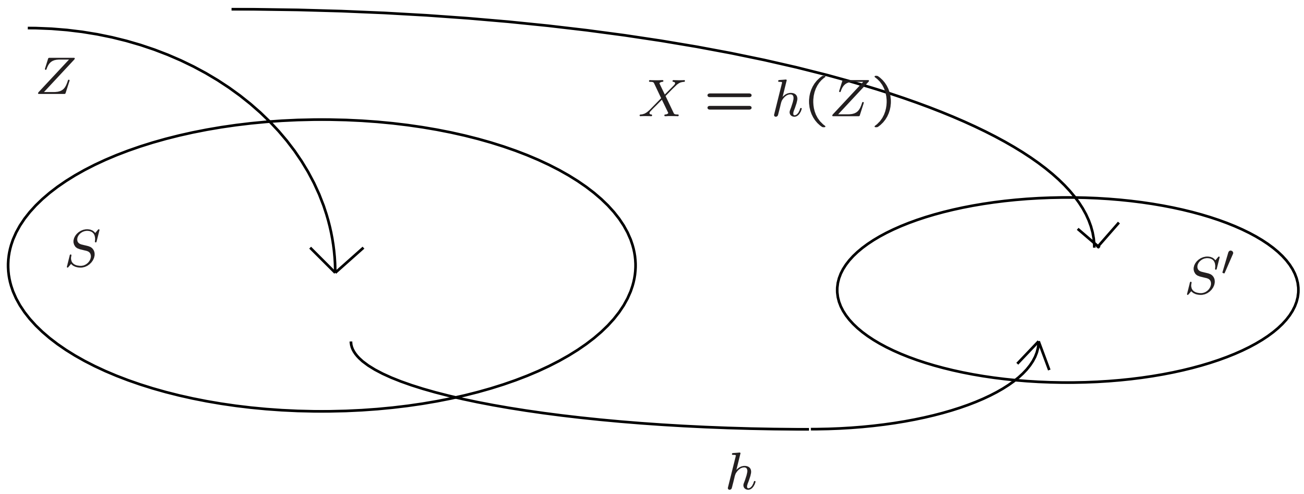
$$\mathbf{P}(X = (k_1, \ldots, k_r)) = \binom{n}{k_1, \ldots, k_r} p_1^{k_1} \ldots p_r^{k_r},$$

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



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

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


Z
$X = h(Z)$
S
S'
h