Vorlesung 2a

Kollision von Kennzeichen

Kollision von Kennzeichen

Ein Beispiel zum Bekanntschaftmachen

mit den grundlegenden Begriffen

Zufallsvariable, Zielbereich,

Ereignis,

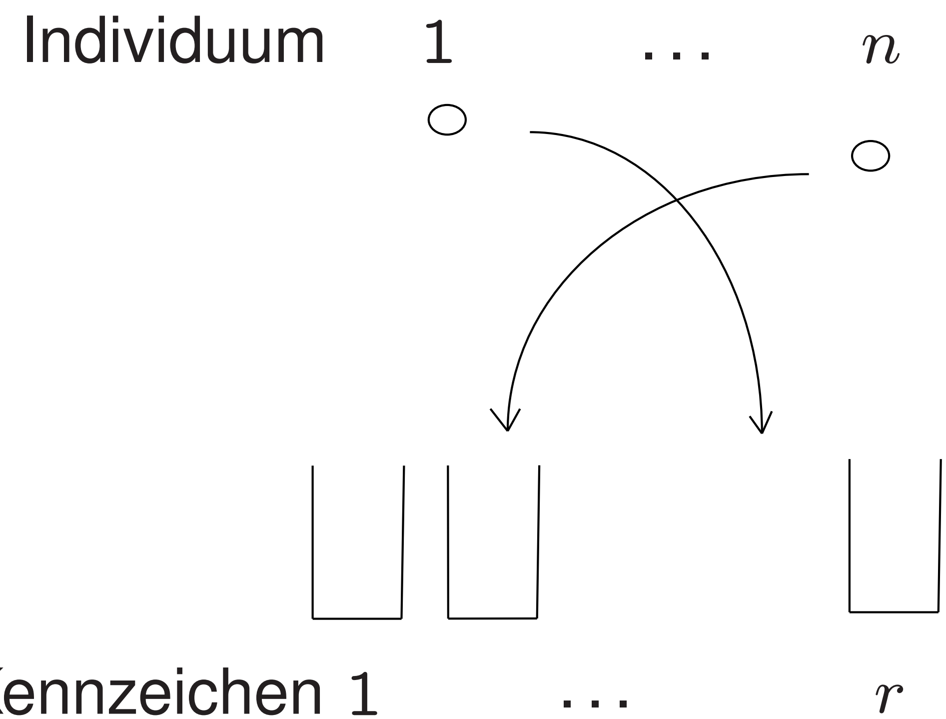
Wahrscheinlichkeit

Jedes von n Individuen ist mit je einem

von r Kennzeichen versehen, das vom Zufall bestimmt ist.

Jedes von n Individuen ist mit je einem

von r Kennzeichen versehen, das vom Zufall bestimmt ist.



Etwa:

$n = 25$ Leute auf einer Party

Etwa:

$n = 25$ Leute auf einer Party

Kennzeichen ... Geburtstag $r \in \{1, 2, \ldots, 365\}$

Wie stehen die Chancen,

dass es dabei zu keiner Kollision kommt, Wie stehen die Chancen,

dass es dabei zu keiner Kollision kommt,

dass also keine zwei Individuen gleich gekennzeichnet sind?

Fragen zur Orientierung:

Fragen zur Orientierung:

Wie beschreibt man die möglichen Ausgänge
der Kennzeichnung?

Fragen zur Orientierung:

Wie beschreibt man die möglichen Ausgänge

der Kennzeichnung?

Wie fasst man das Ereignis

„Keine zwei Individuen haben dasselbe Kennzeichen“?

Fragen zur Orientierung:

Wie beschreibt man die möglichen Ausgänge
der Kennzeichnung?

Wie fasst man das Ereignis
„Keine zwei Individuen haben dasselbe Kennzeichen“?

Wie kommt man zur Wahrscheinlichkeit dieses Ereignisses?

Fragen zur Orientierung:

Wie beschreibt man die möglichen Ausgänge
der Kennzeichnung?

Wie fasst man das Ereignis
„Keine zwei Individuen haben dasselbe Kennzeichen“?

Wie kommt man zur Wahrscheinlichkeit dieses Ereignisses?

Unter welchen Bedingungen kommt es
mit merklicher Wahrscheinlichkeit zu Kollisionen?

Die Individuen denken wir uns mit 1 bis n

und die Kennzeichen mit 1 bis r nummeriert.

Ein Ausgang der Kennzeichnung lässt sich beschreiben

durch das n-tupel

$$a = (a_1, \ldots, a_n) \, ,$$

wobei a_i das Kennzeichen des i-ten Individuums bezeichnet

Die Individuen denken wir uns mit 1 bis n

und die Kennzeichen mit 1 bis r nummeriert.

Ein Ausgang der Kennzeichnung lässt sich beschreiben

durch das n-tupel

$$a = (a_1, \ldots, a_n) ,$$

wobei a_i das Kennzeichen des i-ten Individuums bezeichnet

$(1 \leq a_i \leq r)$.

Die Individuen denken wir uns mit 1 bis n

und die Kennzeichen mit 1 bis r nummeriert.

Ein Ausgang der Kennzeichnung lässt sich beschreiben

durch das n-tupel

$$a = (a_1, \ldots, a_n) \ ,$$

wobei a_i das Kennzeichen des i-ten Individuums bezeichnet

$(1 \leq a_i \leq r)$.

Die Menge der möglichen Ausgänge der Kennzeichnung ist

Die Menge der möglichen Ausgänge der Kennzeichnung ist

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

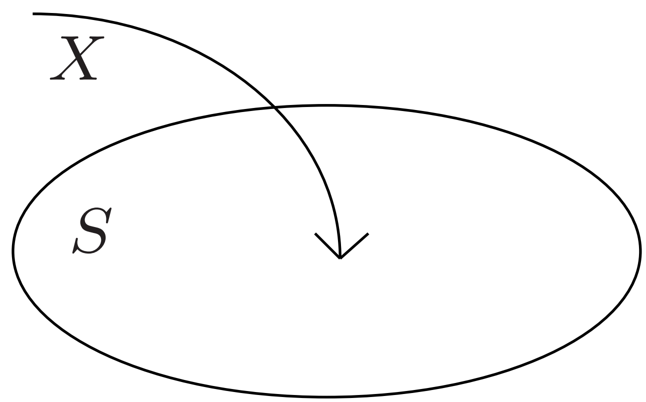
Die Menge der möglichen Ausgänge der Kennzeichnung ist

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

die Menge aller n-tupel $(a_1, \ldots, a_n)$ mit $1 \leq a_i \leq r$.

Den zufälligen Ausgang der Kennzeichnung

beschreiben wir durch eine *Zufallsvariable* X.



X kommt durch zufällige Wahl

eines Elementes aus S zustande.

Die Menge S heißt *Zielbereich* der Zufallsvariable X.

Wie jedes Element $(a_1, \ldots, a_n)$ unserer Menge S

Wie jedes Element $(a_1, \ldots, a_n)$ unserer Menge S

besteht auch die Zufallsvariable X aus n Komponenten:

Wie jedes Element $(a_1, \ldots, a_n)$ unserer Menge S

besteht auch die Zufallsvariable X aus n Komponenten:

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

Wir interessieren uns für das *Ereignis*,

dass keine zwei Komponenten von X gleich sind.

Wir interessieren uns für das *Ereignis*,

dass keine zwei Komponenten von X gleich sind.

Dieses Ereignis schreiben wir als

$$\{X_i \neq X_j \text{ für alle } i \neq j\}$$

Wir interessieren uns für das *Ereignis*,

dass keine zwei Komponenten von X gleich sind.

Dieses Ereignis schreiben wir als

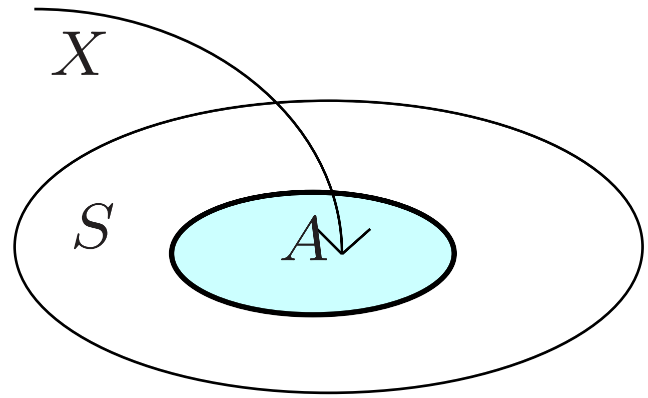
$$\{X_i \neq X_j \text{ für alle } i \neq j\}$$

oder auch als

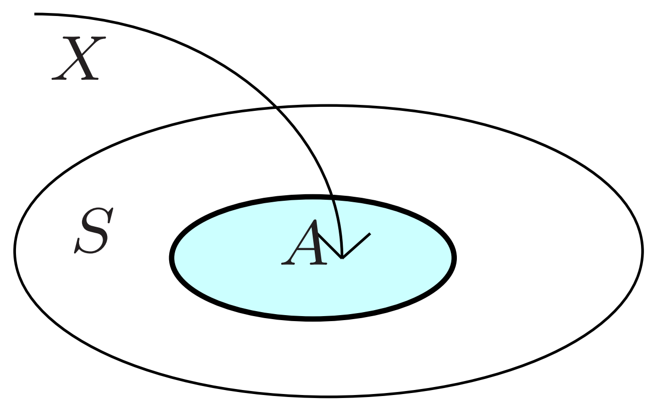
$$\{X \in A\}$$

mit der Teilmenge

$$A := \{a \in S \;:\; a_i \neq a_j \text{ für alle } i \neq j\} \;.$$



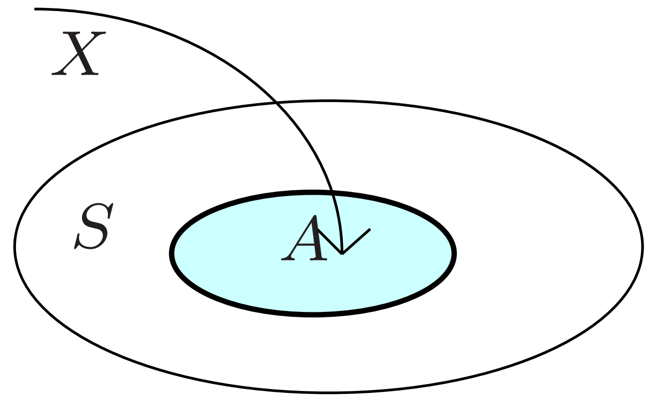
X
S
A



Ereignisse werden (wie Mengen)

in geschweiften Klammern notiert:

$$\{X \in A\}$$



Ereignisse werden (wie Mengen)

in geschweiften Klammern notiert:

$$\{X \in A\}$$

Lies:

Das Ereignis "X fällt in A".

Zurück zu unseren ersten beiden Fragen:

(i) Die möglichen Ausgänge des Kennzeichnens

beschreiben wir durch die Menge

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

Zurück zu unseren ersten beiden Fragen:

(i) Die möglichen Ausgänge des Kennzeichnens

beschreiben wir durch die Menge

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

(ii) Das Ereignis

„Keine zwei Individuen haben dasselbe Kennzeichen“

schreiben wir als

$\{X_i \neq X_j \text{ für alle } i \neq j\}$,

wobei $X = (X_1, \dots, X_n)$ eine *zufällige Wahl* aus S

bezeichnet.

Die weiteren Fragen waren:

(iii) Wie kommt man zur Wahrscheinlichkeit des Ereignisses „Keine zwei Individuen haben dasselbe Kennzeichen“?

Die weiteren Fragen waren:

(iii) Wie kommt man zur Wahrscheinlichkeit des Ereignisses „Keine zwei Individuen haben dasselbe Kennzeichen“?

(iv) Unter welchen Bedingungen kommt es mit merklicher Wahrscheinlichkeit zu Kollisionen?

Bei Frage (iii) kommen *Wahrscheinlichkeiten* ins Spiel.

Bei Frage (iii) kommen *Wahrscheinlichkeiten* ins Spiel.

Wahrscheinlichkeiten gehören zu Ereignissen

Bei Frage (iii) kommen *Wahrscheinlichkeiten* ins Spiel.

Wahrscheinlichkeiten gehören zu Ereignissen

und messen deren Chance einzutreten

Bei Frage (iii) kommen *Wahrscheinlichkeiten* ins Spiel.

Wahrscheinlichkeiten gehören zu Ereignissen

und messen deren Chance einzutreten

mit einer Zahl zwischen 0 und 1:

Bei Frage (iii) kommen *Wahrscheinlichkeiten* ins Spiel.

Wahrscheinlichkeiten gehören zu Ereignissen

und messen deren Chance einzutreten

mit einer Zahl zwischen 0 und 1:

$$\mathbf{P}(X \in A)$$

Zwei einleuchtende Regeln für das Rechnen mit Wahrscheinlichkeiten:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) \qquad (1)$$

Zwei einleuchtende Regeln für das Rechnen mit Wahrscheinlichkeiten:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) \tag{1}$$

d.h. die Wahrscheinlichkeit des Ereignisses, dass X in A fällt,

Zwei einleuchtende Regeln für das Rechnen mit Wahrscheinlichkeiten:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) \tag{1}$$

d.h. die Wahrscheinlichkeit des Ereignisses, dass X in A fällt, ist die Summe über die Wahrscheinlichkeiten der Ereignisse,

Zwei einleuchtende Regeln für das Rechnen mit Wahrscheinlichkeiten:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) \tag{1}$$

d.h. die Wahrscheinlichkeit des Ereignisses, dass X in A fällt, ist die Summe über die Wahrscheinlichkeiten der Ereignisse, dass X den Ausgang a hat, summiert über $a \in A$.

Zwei einleuchtende Regeln für das Rechnen mit Wahrscheinlichkeiten:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) \tag{1}$$

d.h. die Wahrscheinlichkeit des Ereignisses, dass X in A fällt, ist die Summe über die Wahrscheinlichkeiten der Ereignisse, dass X den Ausgang a hat, summiert über $a \in A$.

$$\mathbf{P}(X \in S) = 1 \tag{2}$$

Zwei einleuchtende Regeln für das Rechnen mit Wahrscheinlichkeiten:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) \tag{1}$$

d.h. die Wahrscheinlichkeit des Ereignisses, dass X in A fällt, ist die Summe über die Wahrscheinlichkeiten der Ereignisse, dass X den Ausgang a hat, summiert über $a \in A$.

$$\mathbf{P}(X \in S) = 1 \tag{2}$$

d.h. das *sichere Ereignis* $\{X \in S\}$ hat Wahrscheinlichkeit 1.

Um die Wahrscheinlichkeit $\mathbf{P}(X \in A)$ berechnen zu können,

muss man eine Modellannahme treffen.

Unsere Modellannahme:

Unsere Modellannahme:

rein zufällige Wahl.

Unsere Modellannahme:

rein zufällige Wahl.

Damit ist gemeint, dass für je zwei $a, a' \in S$

$$\mathbf{P}(X = a) = \mathbf{P}(X = a').$$

Unsere Modellannahme:

rein zufällige Wahl.

Damit ist gemeint, dass für je zwei $a, a' \in S$

$$\mathbf{P}(X = a) = \mathbf{P}(X = a').$$

Das heißt: kein Ausgang ist bevorzugt.

Aus

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

und der Gleichheit aller $\mathbf{P}(X = a)$ folgt

$$\mathbb{P}(X = a) = \frac{1}{\#S}\,, \quad a \in S\,.$$

Aus

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

und der Gleichheit aller $\mathbf{P}(X = a)$ folgt

$$\mathbb{P}(X = a) = \frac{1}{\#S}\,, \quad a \in S\,.$$

Also:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) = \sum_{a \in A} \frac{1}{\#S} = \frac{\#A}{\#S}\,.$$

Aus

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

und der Gleichheit aller $\mathbf{P}(X = a)$ folgt

$$\mathbb{P}(X = a) = \frac{1}{\# S}\,, \quad a \in S\,.$$

Also:

$$\mathbf{P}(X \in A) = \sum_{a \in A} \mathbf{P}(X = a) = \sum_{a \in A} \frac{1}{\# S} = \frac{\# A}{\# S}\,.$$

$$\mathbf{P}(X \in A) = \frac{\#A}{\#S} \, .$$

Wir haben nun die Aufgabe des Abzählens der Mengen

$$S := \{1, \ldots, r\}^n = \{(a_1, \ldots, a_n) : 1 \leq a_i \leq r\}$$

Wir haben nun die Aufgabe des Abzählens der Mengen

$$S := \{1, \ldots, r\}^n = \{(a_1, \ldots, a_n) : 1 \leq a_i \leq r\}$$

und

$$A := \{a \in S \,:\, a_i \neq a_j \text{ für alle } i \neq j\}$$

Wir haben nun die Aufgabe des Abzählens der Mengen

$$S := \{1, \ldots, r\}^n = \{(a_1, \ldots, a_n) : 1 \leq a_i \leq r\}$$

und

$$A := \{a \in S \; : \; a_i \neq a_j \text{ für alle } i \neq j\}$$

$$\#S = r^n$$

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

$$A := \{a \in S \,:\, a_i \neq a_j \text{ für alle } i \neq j\}$$

$$\#A = ?$$

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

$$A := \{a \in S \,:\, a_i \neq a_j \text{ für alle } i \neq j\}$$

$$\#A = ?$$

Für a_1 gibt es r mögliche Werte, für a_2 dann noch $r - 1$, usw.

Also:

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

$$A := \{a \in S \ : \ a_i \neq a_j \text{ für alle } i \neq j\}$$

$$\#A = ?$$

Für a_1 gibt es r mögliche Werte, für a_2 dann noch $r-1$, usw.

Also:

$$\#A = r(r-1)\cdots(r-(n-1))$$

$$\mathbf{P}(X \in A) = \frac{r(r-1)\cdots(r-(n-1))}{r^n}$$

$$\mathbf{P}(X \in A) = \frac{r(r-1)\cdots(r-(n-1))}{r^n}$$

$$= \frac{r-1}{r}\,\frac{r-2}{r}\cdots\frac{r-(n-1)}{r}$$

$$\mathbf{P}(X \in A) = \frac{r(r-1)\cdots(r-(n-1))}{r^n}$$

$$= \frac{r-1}{r}\,\frac{r-2}{r}\cdots\frac{r-(n-1)}{r}$$

$$\mathbf{P}(X \in A) = \prod_{i=1}^{n-1}\left(1-\frac{i}{r}\right).$$

$$\mathbf{P}(X \in A) = \frac{r(r-1)\cdots(r-(n-1))}{r^n}$$

$$= \frac{r-1}{r}\,\frac{r-2}{r}\cdots\frac{r-(n-1)}{r}$$

$$\mathbf{P}(X \in A) = \prod_{i=1}^{n-1}\left(1 - \frac{i}{r}\right).$$

$$\mathbf{P}(X \in A) = \frac{r!}{r^n (r-n)!}$$

$$\mathbf{P}(X \in A) = \frac{r(r-1)\cdots(r-(n-1))}{r^n}$$

$$= \frac{r-1}{r}\,\frac{r-2}{r}\cdots\frac{r-(n-1)}{r}$$

$$\mathbf{P}(X \in A) = \prod_{i=1}^{n-1}\left(1-\frac{i}{r}\right).$$

$$\mathbf{P}(X \in A) = \frac{r!}{r^n(r-n)!}$$

mit $k! := 1 \cdot 2 \cdots k$, lies: k-Fakultät

Eine Formel aus der Analysis 1:

$$e^{-h} = 1 - h + o(h) \quad \text{mit} \quad \frac{o(h)}{h} \to 0 \quad \text{für } h \to 0.$$

Eine Formel aus der Analysis 1:

$$e^{-h} = 1 - h + o(h) \quad \text{mit} \quad \frac{o(h)}{h} \to 0 \quad \text{für } h \to 0.$$

Damit bekommen wir die “flotte” Approximation:

$$\prod_{i=1}^{n-1} \left(1 - \frac{i}{r}\right) \approx \prod_{i=1}^{n-1} \exp\left(-\frac{i}{r}\right)$$

Eine Formel aus der Analysis 1:

$$e^{-h} = 1 - h + o(h) \quad \text{mit} \quad \frac{o(h)}{h} \to 0 \quad \text{für } h \to 0.$$

Damit bekommen wir die “flotte” Approximation:

$$\prod_{i=1}^{n-1}\left(1 - \frac{i}{r}\right) \approx \prod_{i=1}^{n-1} \exp\left(-\frac{i}{r}\right)$$

$$= \exp\left(-\sum_{i=1}^{n-1}\frac{i}{r}\right) = \exp\left(-\frac{(n-1)n}{2r}\right).$$

Eine Formel aus der Analysis 1:

$$e^{-h} = 1 - h + o(h) \quad \text{mit} \quad \frac{o(h)}{h} \to 0 \quad \text{für } h \to 0.$$

Damit bekommen wir die “flotte” Approximation:

$$\prod_{i=1}^{n-1} \Big(1 - \frac{i}{r}\Big) \approx \prod_{i=1}^{n-1} \exp\Big(-\frac{i}{r}\Big)$$

$$= \exp\Big(-\sum_{i=1}^{n-1} \frac{i}{r}\Big) = \exp\Big(-\frac{(n-1)n}{2r}\Big) .$$

$$\mathbf{P}(X \in A) \approx \exp\Big(-\frac{(n-1)n}{2r}\Big) .$$

Eine Formel aus der Analysis 1:

$$e^{-h} = 1 - h + o(h) \quad \text{mit} \quad \frac{o(h)}{h} \to 0 \quad \text{für } h \to 0.$$

Damit bekommen wir die “flotte” Approximation:

$$\prod_{i=1}^{n-1} \left(1 - \frac{i}{r}\right) \approx \prod_{i=1}^{n-1} \exp\left(-\frac{i}{r}\right)$$

$$= \exp\left(-\sum_{i=1}^{n-1} \frac{i}{r}\right) = \exp\left(-\frac{(n-1)n}{2r}\right).$$

$$\mathbf{P}(X \in A) \approx \exp\left(-\frac{(n-1)n}{2r}\right).$$

Für welche n und r können wir dieser Approximation trauen?

Die Stirling-Formel:

Die Stirling-Formel:

$$k! \approx \sqrt{2\pi\, k}\, \left(\frac{k}{e}\right)^k$$

Abraham de Moivre (1667-1754)

$$k! \approx \sqrt{2\pi\, k}\, \left(\frac{k}{e}\right)^k$$

$$\frac{r!}{r^n (r-n)!} \approx \sqrt{\frac{r}{r-n}}\, \left(\frac{r}{r-n}\right)^{r-n} e^{-n}$$

$$k! \approx \sqrt{2\pi\, k}\, \left(\frac{k}{e}\right)^{k}$$

$$\frac{r!}{r^n (r-n)!} \approx \sqrt{\frac{r}{r-n}}\, \left(\frac{r}{r-n}\right)^{r-n} e^{-n}$$

$$= \frac{1}{\sqrt{1-\frac{n}{r}}} \left(1 - \frac{n}{r}\right)^{n-r} e^{-n}$$

$$k! \approx \sqrt{2\pi\, k}\, \left(\frac{k}{e}\right)^k$$

$$\frac{r!}{r^n (r-n)!} \approx \sqrt{\frac{r}{r-n}}\, \left(\frac{r}{r-n}\right)^{r-n} e^{-n}$$

$$= \frac{1}{\sqrt{1-\frac{n}{r}}} \left(1 - \frac{n}{r}\right)^{n-r} e^{-n}$$

$$k! \approx \sqrt{2\pi\, k}\, \left(\frac{k}{e}\right)^{k}$$

$$\frac{r!}{r^{n}(r-n)!} \approx \sqrt{\frac{r}{r-n}}\, \left(\frac{r}{r-n}\right)^{r-n} e^{-n}$$

$$= \frac{1}{\sqrt{1-\frac{n}{r}}} \left(1-\frac{n}{r}\right)^{n-r} e^{-n}$$

$$\left(1-\frac{n}{r}\right)^{n-r} = \exp\left((n-r)\ln\left(1-\frac{n}{r}\right)\right)$$

$$\left(1-\frac{n}{r}\right)^{n-r}e^{-n}=\exp\left(-n+(n-r)\ln\left(1-\frac{n}{r}\right)\right)$$

$$\left(1-\frac{n}{r}\right)^{n-r}e^{-n}=\exp\left(-n+(n-r)\ln\left(1-\frac{n}{r}\right)\right)$$

$$=\exp\left(-r\left(\frac{n}{r}+\left(1-\frac{n}{r}\right)\ln\left(1-\frac{n}{r}\right)\right)\right)$$

$$\left(1-\frac{n}{r}\right)^{n-r}e^{-n} = \exp\Big(-n+(n-r)\ln\Big(1-\frac{n}{r}\Big)\Big)$$

$$= \exp\left(-r\,\Big(\frac{n}{r}+\Big(1-\frac{n}{r}\Big)\ln\Big(1-\frac{n}{r}\Big)\Big)\right)$$

$$\left(1-\frac{n}{r}\right)^{n-r}e^{-n}=\exp\left(-n+(n-r)\ln\left(1-\frac{n}{r}\right)\right)$$

$$=\exp\left(-r\left(\frac{n}{r}+\left(1-\frac{n}{r}\right)\ln\left(1-\frac{n}{r}\right)\right)\right)$$

$$=\exp\left(-r\,\eta\left(\frac{n}{r}\right)\right)$$

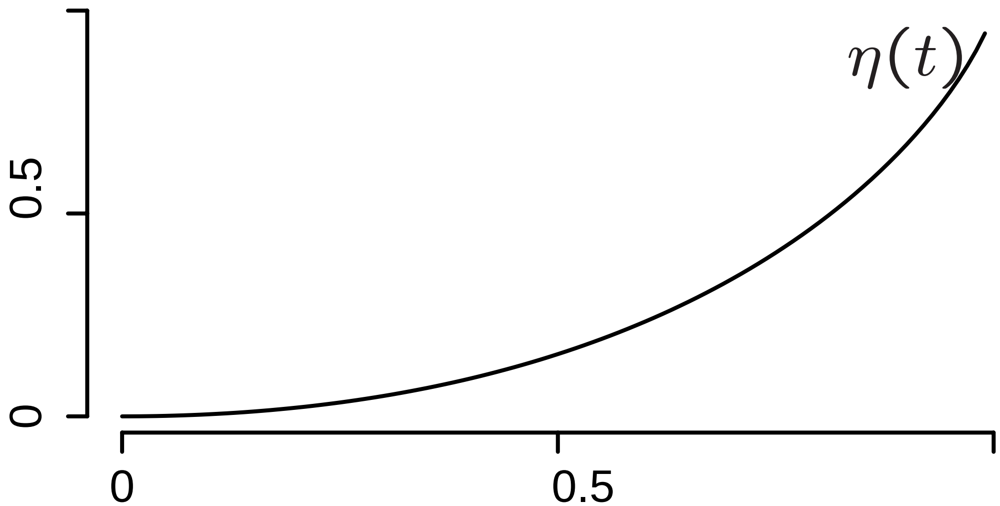
$$\Big(1-\frac{n}{r}\Big)^{n-r}e^{-n} = \exp\Big(-n+(n-r)\ln\Big(1-\frac{n}{r}\Big)\Big)$$

$$= \exp\Big(-r\,\Big(\frac{n}{r}+\Big(1-\frac{n}{r}\Big)\ln\Big(1-\frac{n}{r}\Big)\Big)\Big)$$

$$= \exp\Big(-r\,\eta\Big(\frac{n}{r}\Big)\Big)$$

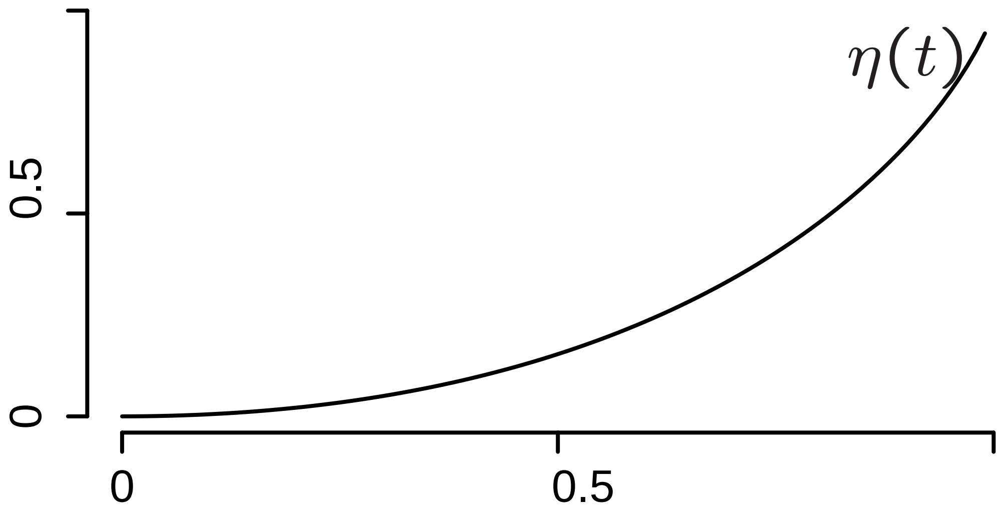
mit $\quad \eta(t) := t+(1-t)\ln(1-t), \qquad 0 \le t \le 1.$

$$\eta(t) := t + (1-t)\ln(1-t), \qquad 0 \le t \le 1.$$



$$\mathbf{P}(X \in A) \approx \frac{1}{\sqrt{1-\frac{n}{r}}} \exp\left(-r\,\eta\left(\frac{n}{r}\right)\right)$$

$$\eta(t) := t + (1-t)\ln(1-t), \qquad 0 \le t \le 1.$$

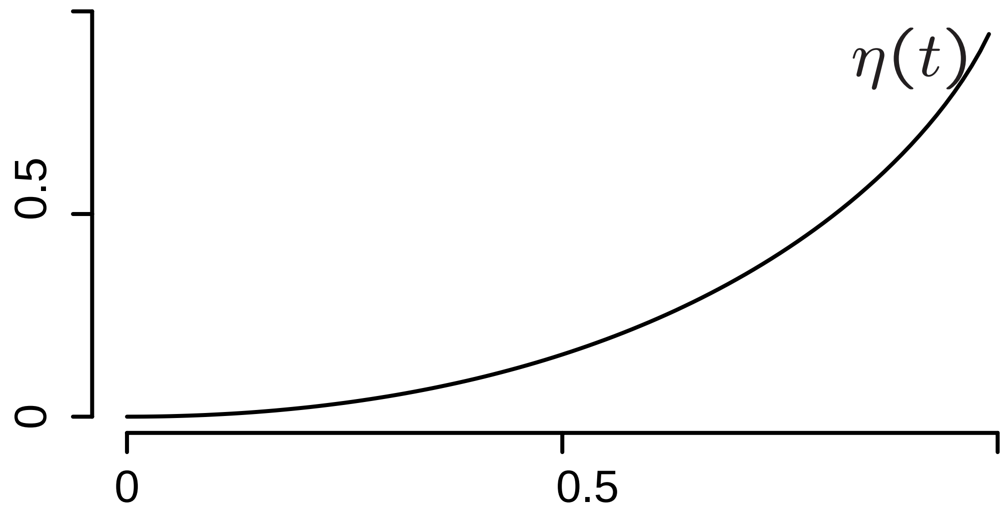


$$\mathbf{P}(X \in A) \approx \frac{1}{\sqrt{1-\frac{n}{r}}} \exp\left(-r\,\eta\left(\frac{n}{r}\right)\right)$$

Für $n \ll r$, also $t := \frac{n}{r} \ll 1$,

können wir $\eta(t)$ quadratisch approximieren:

$$\eta(t) := t + (1-t)\ln(1-t), \qquad 0 \le t \le 1.$$



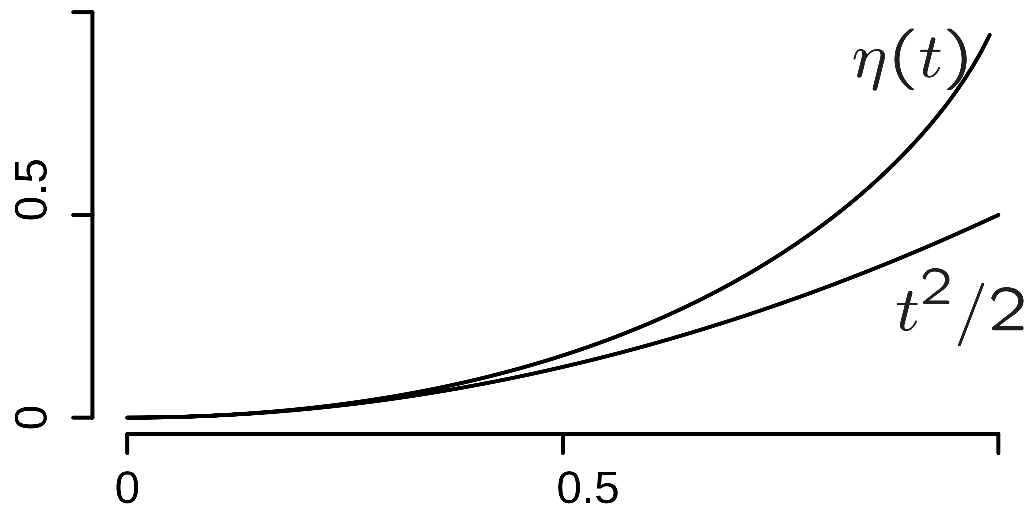
$$\eta(0) = \eta'(0) = 0, \quad \eta''(0) = 1$$

$$\eta(t) = t + (1 - t)\ln(1 - t), \qquad 0 \le t \le 1$$

$$\eta(0) = 0, \quad \eta'(0) = 0, \quad \eta''(0) = 1$$

$$\eta(t) = t + (1-t)\ln(1-t), \qquad 0 \le t \le 1$$

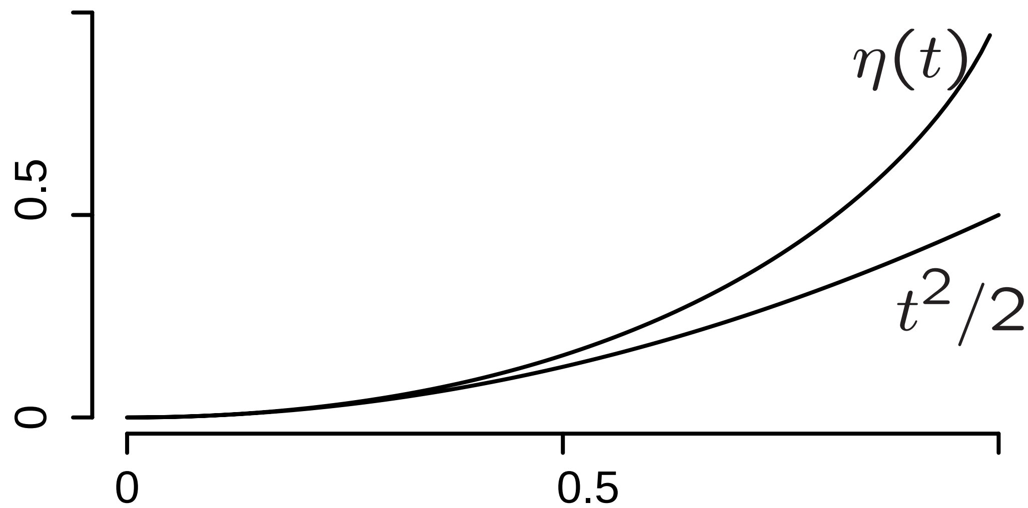
$$\eta(0) = 0, \quad \eta'(0) = 0, \quad \eta''(0) = 1$$



$t^2/2$ ist die quadratische Approximation von $\eta(t)$ um $t = 0$.

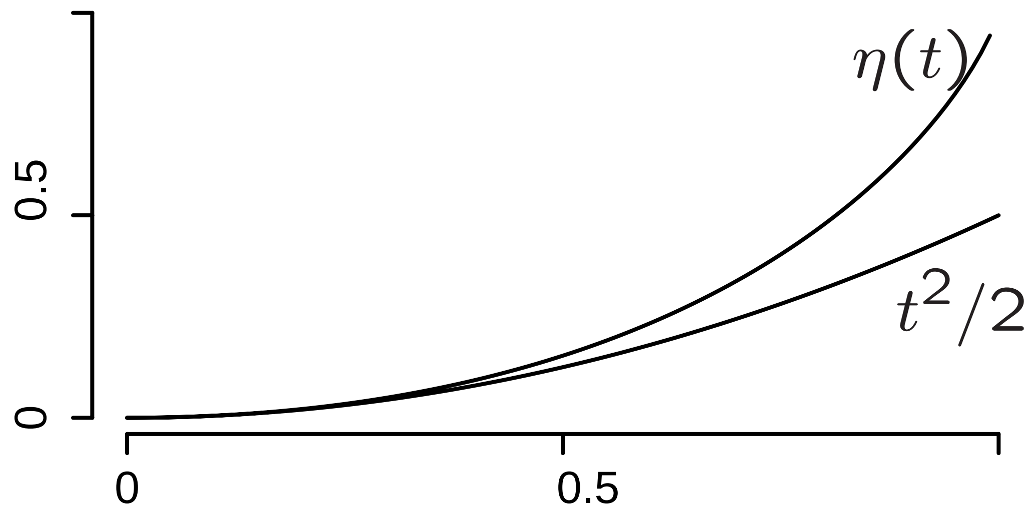
$$\eta(t) = t + (1-t)\ln(1-t), \qquad 0 \le t \le 1$$

$$\eta(0) = 0, \quad \eta'(0) = 0, \quad \eta''(0) = 1$$



$t^2/2$ ist die quadratische Approximation von $\eta(t)$ um $t = 0$.

(Taylor-Entwicklung der Ordnung 2.)



$t^2/2$ ist die quadratische Approximation von $\eta(t)$ um $t = 0$.

Für $n \ll r$ (wie z. B. für $n = 25$, $r = 365$) ist also

$$\eta\left(\frac{n}{r}\right) \approx \left(\frac{n}{r}\right)^2/2.$$

Fazit:

Fazit:

Für $n < r$ ist $\mathbf{P}(X \in A) \approx \dfrac{1}{\sqrt{1 - \frac{n}{r}}} \exp\left(-r\,\eta\left(\dfrac{n}{r}\right)\right)$

Fazit:

Für $n < r$ ist $\quad \mathbf{P}(X \in A) \approx \dfrac{1}{\sqrt{1-\frac{n}{r}}}\,\mathrm{exp}\left(-r\,\eta\Big(\dfrac{n}{r}\Big)\right)$

mit $\eta(t) = t + (1-t)\ln(1-t), \qquad 0 \leq t \leq 1$

(Stirling-Approximation).

Fazit:

Für $n < r$ ist $\quad \mathbf{P}(X \in A) \approx \dfrac{1}{\sqrt{1-\frac{n}{r}}} \exp\left(-r\,\eta\left(\dfrac{n}{r}\right)\right)$

mit $\eta(t) = t + (1-t)\ln(1-t), \qquad 0 \le t \le 1$

(Stirling-Approximation).

Für $n \ll r$ ist $\quad \mathbf{P}(X \in A) \approx \exp\left(-\dfrac{n^2}{2r}\right)$

Fazit:

Für $n < r$ ist $\quad \mathbf{P}(X \in A) \approx \dfrac{1}{\sqrt{1-\frac{n}{r}}} \exp\left(-r\,\eta\Big(\dfrac{n}{r}\Big)\right)$

mit $\eta(t) = t + (1-t)\ln(1-t), \qquad 0 \le t \le 1$

(Stirling-Approximation).

Für $n \ll r$ ist $\quad \mathbf{P}(X \in A) \approx \exp\Big(-\dfrac{n^2}{2r}\Big)$

(Stirling+Taylor-Approximation)

Fazit:

Für $n < r$ ist $\quad \mathbf{P}(X \in A) \approx \dfrac{1}{\sqrt{1-\frac{n}{r}}} \exp\left(-r\,\eta\Big(\dfrac{n}{r}\Big)\right)$

mit $\eta(t) = t + (1-t)\ln(1-t), \qquad 0 \le t \le 1$

(Stirling-Approximation).

Für $n \ll r$ ist $\quad \mathbf{P}(X \in A) \approx \exp\Big(-\dfrac{n^2}{2r}\Big)$

(Stirling+Taylor-Approximation)

Für $n^2 \ll r$ ist $\quad \mathbf{P}(X \in A) \approx 1.$

Man beachte:

Die Stirling+Taylor Approximation

$$\mathbf{P}(X \in A) \approx \exp\Big(-\frac{n^2}{2r}\Big)$$

ist fast identisch mit der “flotten” Approximation

$$\mathbf{P}(X \in A) \approx \exp\Big(-\frac{n(n-1)}{2r}\Big)$$

Beispiel: $n = 25, r = 365$:

$$\mathbf{P}(X \in A) = \frac{r(r-1)\cdots(r-n+1)}{r^n} = 0.431300$$

$$\frac{1}{\sqrt{1-\frac{n}{r}}}\exp\left(-r\,\eta\Big(\frac{n}{r}\Big)\right) = 0.431308$$

$$\exp\Big(-\frac{n^2}{2r}\Big) = 0.425$$

Von der Statik zur Dynamik:

Der Zeitpunkt der ersten Kollision.

Von der Statik zur Dynamik:

Der Zeitpunkt der ersten Kollision.

Wir denken uns r fest und lassen n laufen $(n = 1, 2, ..)$

Von der Statik zur Dynamik:

Der Zeitpunkt der ersten Kollision.

Wir denken uns r fest und lassen n laufen $(n = 1, 2, ..)$

Vorstellung: Ein Individuum nach dem anderen wird auf einen (immer wieder neu) rein zufällig gewählten Platz gesetzt.

Von der Statik zur Dynamik:

Der Zeitpunkt der ersten Kollision.

Wir denken uns r fest und lassen n laufen $(n = 1, 2, ..)$

Vorstellung: Ein Individuum nach dem anderen wird auf einen (immer wieder neu) rein zufällig gewählten Platz gesetzt.

Die Folge $(X_1, X_2, \ldots)$ der gewählten Plätze ist dann eine rein zufällige $1 \ldots r$-Folge

$A_n :=$ “ keine Kollision bis (einschließlich) n ”

Die Zufallsvariable T:= Zeitpunkt der ersten Kollision

(ist ablesbar aus $(X_1, X_2, \ldots)$). Es gilt

$A_n :=$ " keine Kollision bis (einschließlich) n "

Die Zufallsvariable T:= Zeitpunkt der ersten Kollision

(ist ablesbar aus $(X_1, X_2, \ldots)$). Es gilt

$$A_n = \{T > n\}$$

also insbesondere auch

$$\mathbf{P}(A_n) = \mathbf{P}(T > n).$$

$A_n :=$ " keine Kollision bis (einschließlich) n "

Die Zufallsvariable T:= Zeitpunkt der ersten Kollision

(ist ablesbar aus $(X_1, X_2, \ldots)$. Es gilt

$$A_n = \{T > n\}$$

also insbesondere auch

$$\mathbf{P}(A_n) = \mathbf{P}(T > n).$$

Wir haben somit AUCH die Verteilung der Zufallsvariablen T (exakt und näherungsweise) berechnet.

$A_n :=$ „ keine Kollision bis (einschließlich) n "

Die Zufallsvariable T:= Zeitpunkt der ersten Kollision

(ist ablesbar aus $(X_1, X_2, \ldots)$. Es gilt

$$A_n = \{T > n\}$$

also insbesondere auch

$$\mathbf{P}(A_n) = \mathbf{P}(T > n).$$

Wir haben somit AUCH die Verteilung der Zufallsvariablen T (exakt und näherungsweise) berechnet.

Eine erhellende Illustration findet man in den R-Programmen auf http://www.uni-frankfurt.de/fb/fb12/mathematik/st/Mitglieder/Franziska_Wandtner/R_Programme.html

Rein zufällge Permutation

Eine *Permutation* von $1, \ldots, n$

ist eine bijektive Abbildung der Menge $\{1, \ldots, n\}$ auf sich.

Eine *Permutation* von $1, \ldots, n$

ist eine bijektive Abbildung der Menge $\{1, \ldots, n\}$ auf sich.

Z. B. mit $n = 7$

Eine *Permutation* von $1, \ldots, n$

ist eine bijektive Abbildung der Menge $\{1, \ldots, n\}$ auf sich.

Z. B. mit $n = 7$

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Eine *Permutation* von $1, \ldots, n$

ist eine bijektive Abbildung der Menge $\{1, \ldots, n\}$ auf sich.

Z. B. mit $n = 7$

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Wie wahrscheinlich ist es,

dass eine rein zufällige Permutation

genau **so** ausfällt?

Wieviele Permutationen von $1, \ldots, n$ gibt es?

Wieviele Permutationen von $1, \ldots, n$ gibt es?

n Möglichkeiten für das Bild von 1

Wieviele Permutationen von $1, \ldots, n$ gibt es?

n Möglichkeiten für das Bild von 1

mal $(n-1)$ Möglichkeiten für das Bild von 2

Wieviele Permutationen von $1, \ldots, n$ gibt es?

n Möglichkeiten für das Bild von 1

mal $(n-1)$ Möglichkeiten für das Bild von 2

mal $(n-2)$ Möglichkeiten für das Bild von 3

Wieviele Permutationen von $1, \ldots, n$ gibt es?

n Möglichkeiten für das Bild von 1

mal $(n-1)$ Möglichkeiten für das Bild von 2

mal $(n-2)$ Möglichkeiten für das Bild von 3

$$\ldots$$

$$= n(n-1)(n-2)\cdots 2 \cdot 1$$

Wieviele Permutationen von $1, \ldots, n$ gibt es?

n Möglichkeiten für das Bild von 1

mal $(n-1)$ Möglichkeiten für das Bild von 2

mal $(n-2)$ Möglichkeiten für das Bild von 3

$$\ldots$$

$$= n(n-1)(n-2)\cdots 2 \cdot 1 =: n!$$

Sei X eine rein zufällige Permutation von $1, \ldots, n$,

Sei X eine rein zufällige Permutation von $1, \ldots, n$,

d.h. eine Zufallsvariable, deren Zielbereich

$S :=$ die Menge aller Permutationen von $1, \ldots, n$

ist, und die auf S uniform verteilt ist.

Sei X eine rein zufällige Permutation von $1, \ldots, n$,

d.h. eine Zufallsvariable, deren Zielbereich

$S :=$ die Menge aller Permutationen von $1, \ldots, n$

ist, und die auf S uniform verteilt ist.

Für alle Elemente $a \in S$ gilt also:

Sei X eine rein zufällige Permutation von $1, \ldots, n$,

d.h. eine Zufallsvariable, deren Zielbereich

$S :=$ die Menge aller Permutationen von $1, \ldots, n$

ist, und die auf S uniform verteilt ist.

Für alle Elemente $a \in S$ gilt also:

$$\mathbf{P}(X = a) =$$

Sei X eine rein zufällige Permutation von $1, \ldots, n$,

d.h. eine Zufallsvariable, deren Zielbereich

$S :=$ die Menge aller Permutationen von $1, \ldots, n$

ist, und die auf S uniform verteilt ist.

Für alle Elemente $a \in S$ gilt also:

$$\mathbf{P}(X = a) = \frac{1}{n!}$$

Wie kommen rein zufällige Permutationen zustande?

Wie kommen rein zufällige Permutationen zustande?

Zum Beispiel: als Folge der gezogenen Nummern

beim n-maligen rein zufälligen Ziehen ohne Zurücklegen

aus $\{1, 2, \ldots, n\}$.

Szenario: eine stets ideal durchmischte Urne

mit anfangs n Kugeln, beschriftet mit den Nummern $1, \ldots, n$.

Wie kommen rein zufällige Permutationen zustande?

Zum Beispiel: als Folge der gezogenen Nummern
beim n-maligen rein zufälligen Ziehen ohne Zurücklegen
aus $\{1, 2, \ldots, n\}$.

Szenario: eine stets ideal durchmischte Urne
mit anfangs n Kugeln, beschriftet mit den Nummern $1, \ldots, n$.
Ziehe sukzessive ohne Zurücklegen alle n Kugeln,

Wie kommen rein zufällige Permutationen zustande?

Zum Beispiel: als Folge der gezogenen Nummern

beim n-maligen rein zufälligen Ziehen ohne Zurücklegen

aus $\{1, 2, \ldots, n\}$.

Szenario: eine stets ideal durchmischte Urne

mit anfangs n Kugeln, beschriftet mit den Nummern $1, \ldots, n$.

Ziehe sukzessive ohne Zurücklegen alle n Kugeln,

notiere die gezogenen Nummern in Reihenfolge.

Jede Permutation zerfällt in **Zyklen**

Beispiel:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Jede Permutation zerfällt in **Zyklen**

Beispiel:

$$\begin{array}{ccccccc} \color{red}{1} & 2 & 3 & 4 & 5 & 6 & 7 \\ \color{red}{\downarrow} & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \color{red}{5} & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Jede Permutation zerfällt in **Zyklen**

Beispiel:

$$\begin{array}{ccccccc} \color{red}{1} & 2 & 3 & 4 & \color{red}{5} & 6 & 7 \\ \color{red}{\downarrow} & \downarrow & \downarrow & \downarrow & \color{red}{\downarrow} & \downarrow & \downarrow \\ \color{red}{5} & 2 & 7 & 3 & \color{red}{1} & 4 & 6 \end{array}$$

Jede Permutation zerfällt in **Zyklen**

Beispiel:

$$\begin{array}{ccccccc} \color{red}{1} & \color{green}{2} & 3 & 4 & \color{red}{5} & 6 & 7 \\ \color{red}{\downarrow} & \color{green}{\downarrow} & \downarrow & \downarrow & \color{red}{\downarrow} & \downarrow & \downarrow \\ \color{red}{5} & \color{green}{2} & 7 & 3 & \color{red}{1} & 4 & 6 \end{array}$$

Jede Permutation zerfällt in **Zyklen**

Beispiel:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Jede Permutation zerfällt in **Zyklen**

Beispiel:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Jede Permutation zerfällt in **Zyklen**

Beispiel:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Jede Permutation zerfällt in **Zyklen**

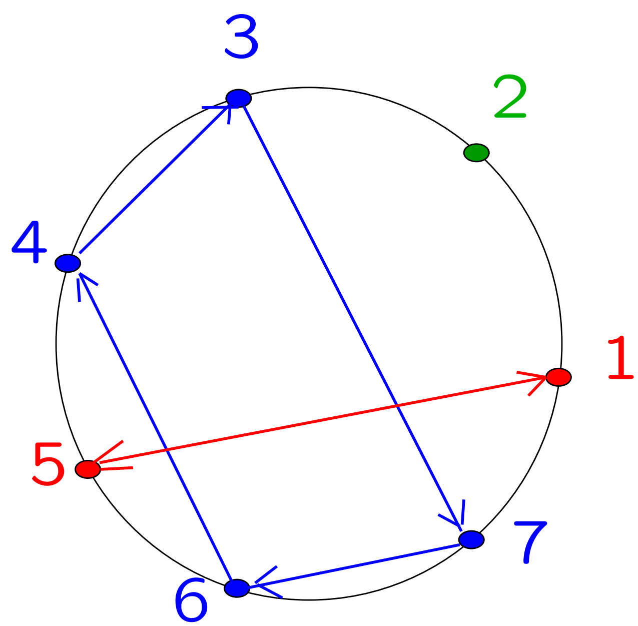
Beispiel:

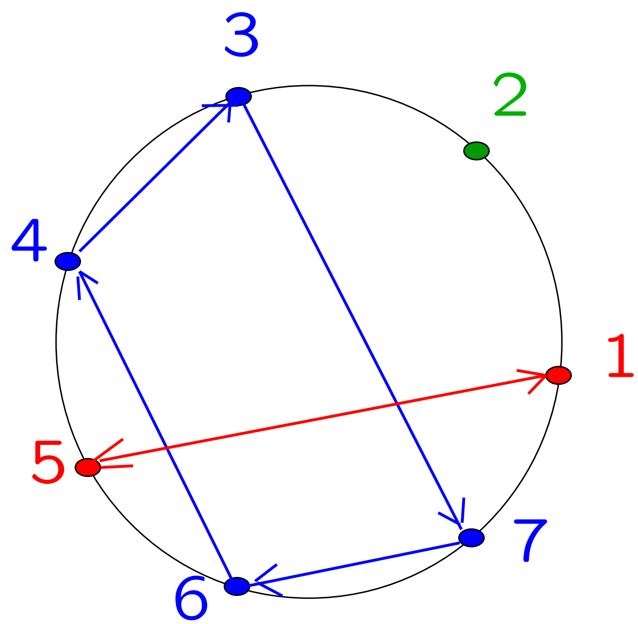
$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$

Jede Permutation zerfällt in **Zyklen**

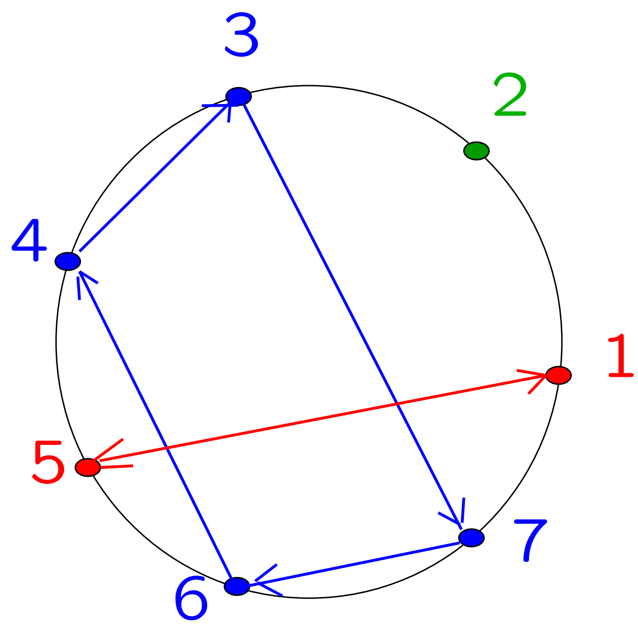
Beispiel:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 2 & 7 & 3 & 1 & 4 & 6 \end{array}$$





Die Länge des Zyklus, der die Eins enthält, ist hier



Die Länge des Zyklus, der die Eins enthält, ist hier

zwei.

Sei X eine rein zufällige Permutation von $1, \ldots, n$,

Sei X eine rein zufällige Permutation von $1, \dots, n$,

und b eine Zahl zwischen 1 und n.

Sei X eine rein zufällige Permutation von $1, \ldots, n$,

und b eine Zahl zwischen 1 und n.

Gefragt ist nach der Wahrscheinlichkeit des Ereignisses
“ der Zyklus von X, der die Eins enthält, hat die Länge b ”.

Für eine Permutation $a \in S$ bezeichne

$$h(a)$$

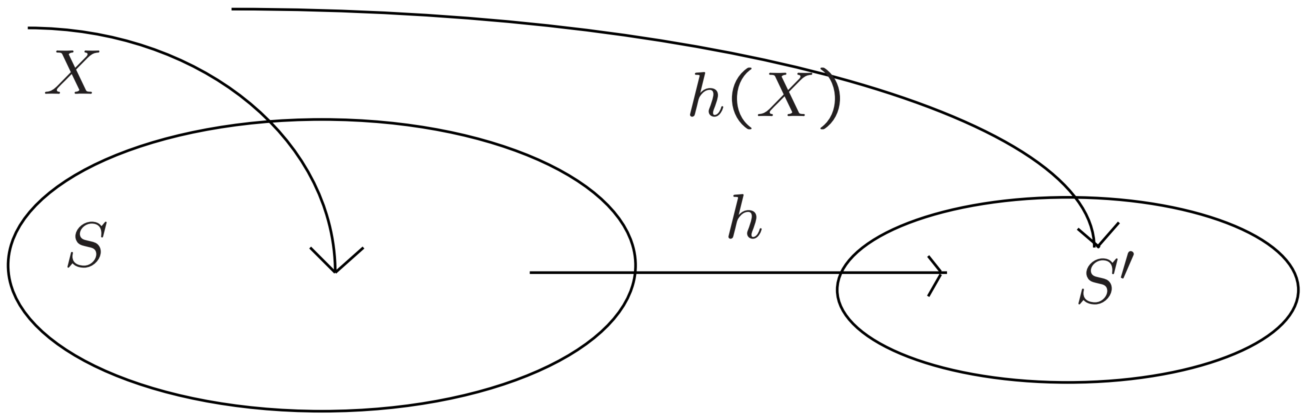
die Länge des Zyklus von a, der die Eins enthält.

Für eine Permutation $a \in S$ bezeichne

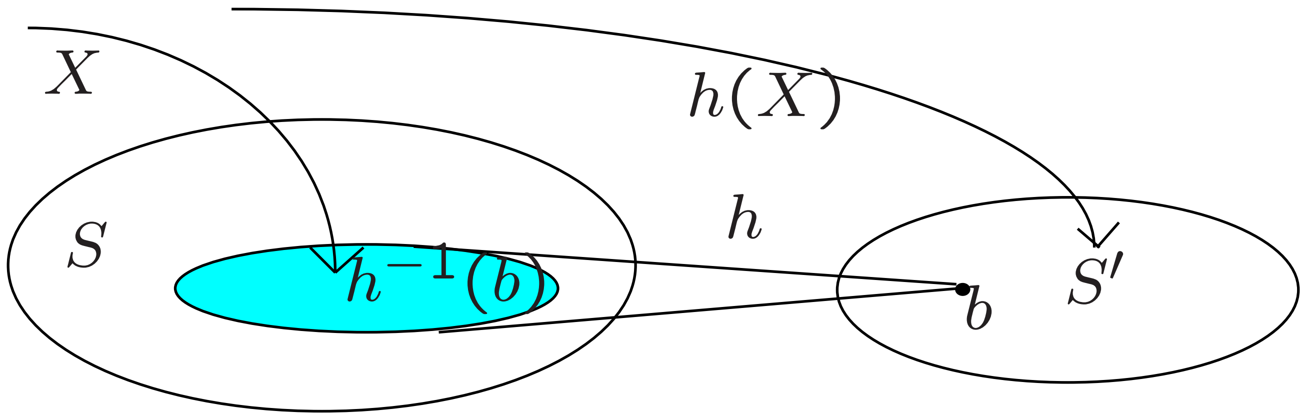
$$h(a)$$

die Länge des Zyklus von a, der die Eins enthält.

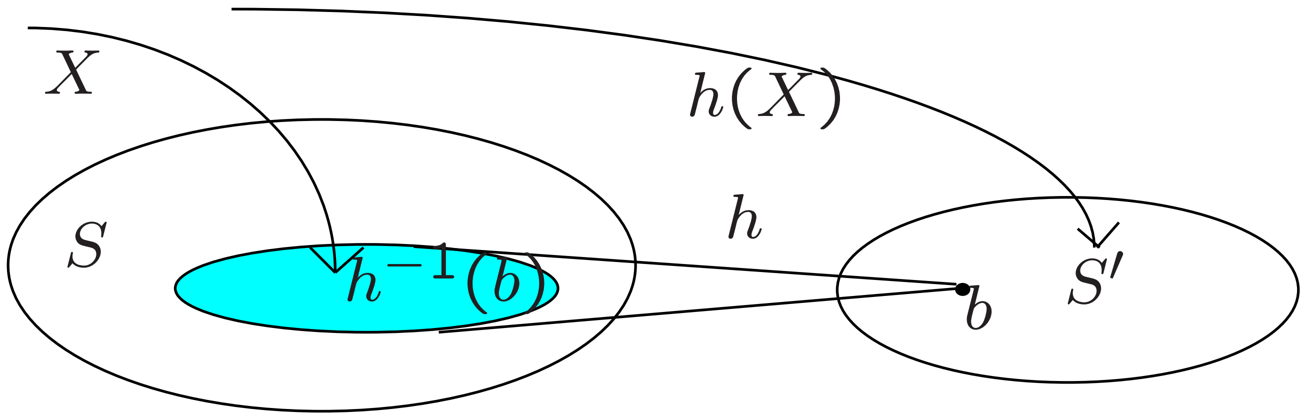
$$\mathbf{P}(h(X) = b) = ?$$



$$S' = \{1, \ldots, n\}$$

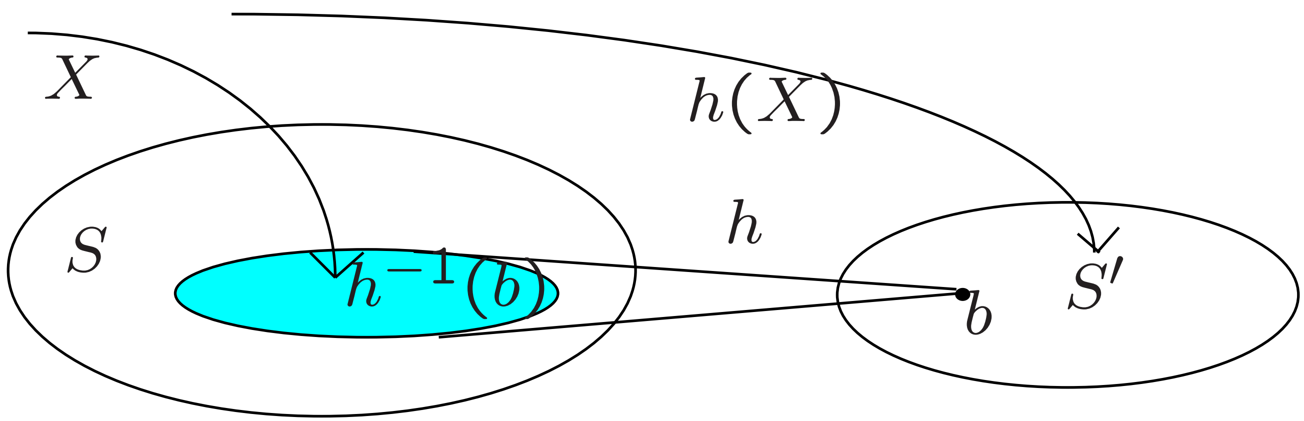


$$S' = \{1, \ldots, n\}$$



$S' = \{1, \ldots, n\}$

Wieviele Permutationen $a \in S$ gibt es mit $h(a) = b$?



$$S' = \{1, \ldots, n\}$$

Wieviele Permutationen $a \in S$ gibt es mit $h(a) = b$?

$$A := \{a \in S : h(a) = b\}$$

$$\#A = ?$$

$$A = \{a \in S : a(1) \neq 1, a^2(1) \neq 1, \ldots, \\ a^{b-1}(1) \neq 1, a^b(1) = 1\}$$

$$A = \{a \in S : a(1) \neq 1, a^2(1) \neq 1, \ldots,$$
$$a^{b-1}(1) \neq 1, a^b(1) = 1\}$$

$$\#A = (n-1)(n-2)\cdots(n-b+1)\cdot 1 \cdot (n-b)\cdots 1$$

$$A = \{a \in S : a(1) \neq 1, a^2(1) \neq 1, \ldots, \\ a^{b-1}(1) \neq 1, a^b(1) = 1\}$$

$$\#A = (n-1)(n-2)\cdots(n-b+1)\cdot 1 \cdot (n-b)\cdots 1$$

$$= (n-1)!$$

$$\#A = (n-1)!, \qquad \#S = n!$$

$$\mathbf{P}(X \in A) = \frac{\#A}{\#S}$$

$$\#A = (n-1)!, \qquad \#S = n!$$

$$\mathbf{P}(X \in A) = \frac{\#A}{\#S}$$

$$= \frac{(n-1)!}{n!}$$

$$\#A = (n-1)!, \qquad \#S = n!$$

$$\mathbf{P}(X \in A) = \frac{\#A}{\#S}$$

$$= \frac{(n-1)!}{n!}$$

$$= \frac{1}{n}$$

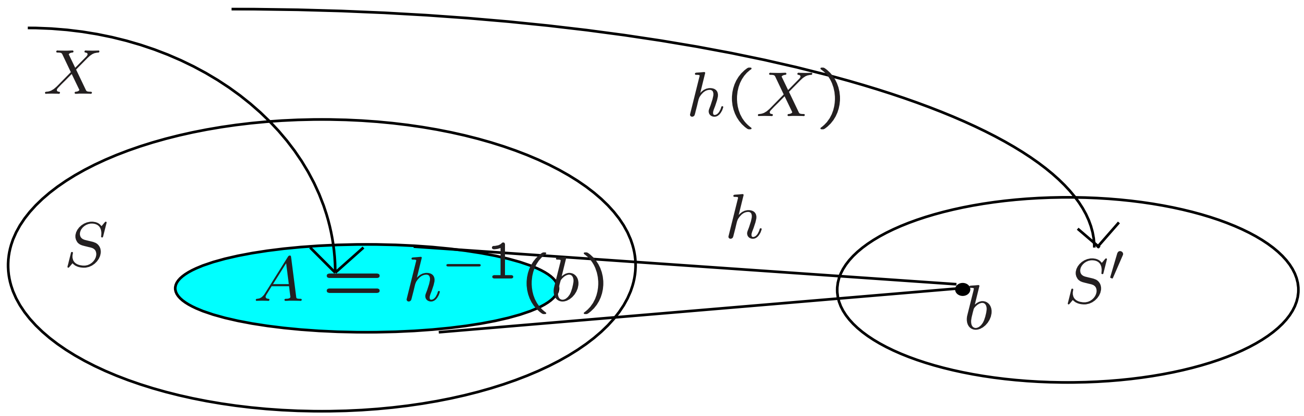
$$\#A = (n-1)!, \qquad \#S = n!$$

$$\mathbf{P}(X \in A) = \frac{\#A}{\#S}$$

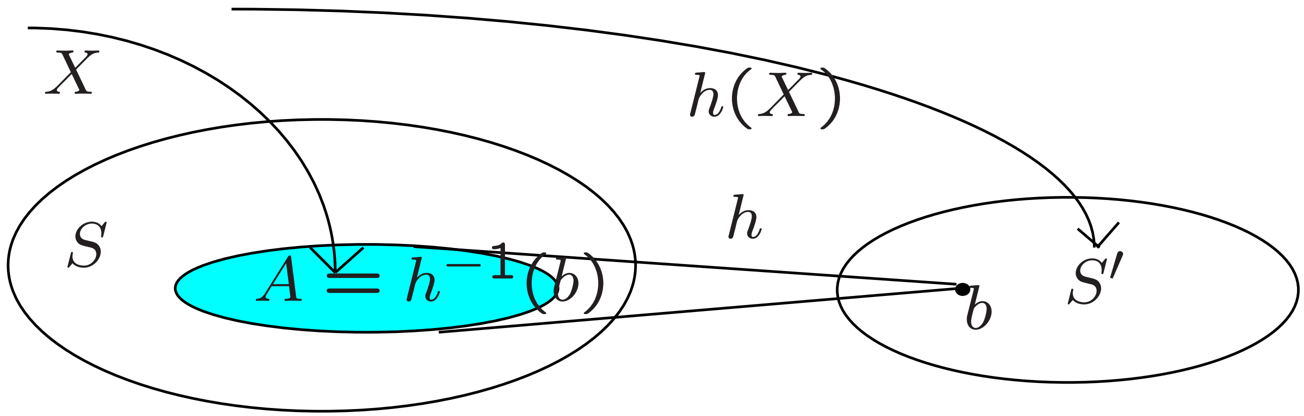
$$= \frac{(n-1)!}{n!}$$

$$= \frac{1}{n}$$

$$\mathbf{P}(X \in A) = \frac{1}{n}$$

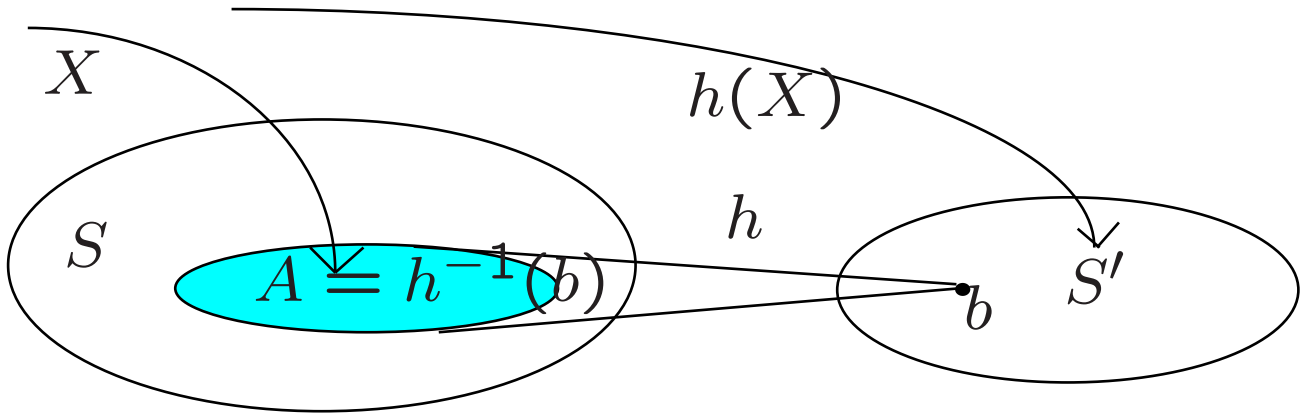


$$S' = \{1, \ldots, n\}$$



$$S' = \{1, \ldots, n\}$$

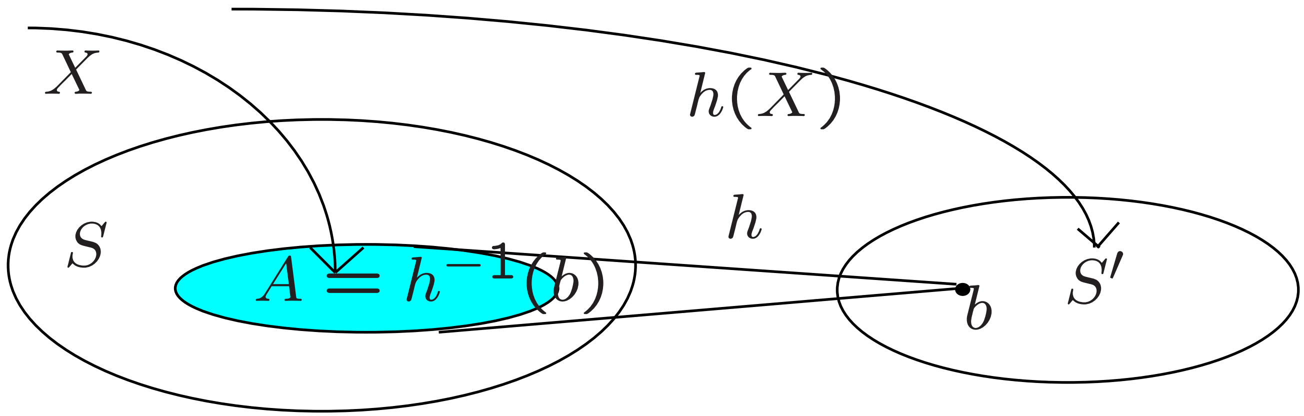
$$A = \{a \in S : h(a) = b\}$$



$$S' = \{1, \ldots, n\}$$

$$A = \{a \in S : h(a) = b\}$$

$$\{X \in A\} = \{h(X) = b\}$$



$S' = \{1, \ldots, n\}$

$$A = \{a \in S : h(a) = b\}$$

$$\{X \in A\} = \{h(X) = b\}$$

$$\mathbf{P}(h(X) = b) = \frac{1}{n}$$

$$\mathbf{P}(h(X) = b) = \frac{1}{n}, \quad b = 1, \ldots n.$$

$$\mathbf{P}(h(X) = b) = \frac{1}{n}, \quad b = 1, \ldots n.$$

Fazit:

$$\mathbf{P}(h(X) = b) = \frac{1}{n}, \quad b = 1, \ldots n.$$

Fazit:

Die Länge desjenigen Zyklus

einer rein zufälligen Permutation von $1, \ldots, n$,

der die Eins enthält,

ist uniform verteilt auf $\{1, \ldots, n\}$.