Vorlesung 2a

Vorlesung 2a

Diskret uniform verteilte Zufallsvariable

(Buch S. 6-11)

0. Erinnerung und Auftakt

Sei S eine endliche Menge.

Eine Zufallsvariable X heißt *uniform verteilt auf S*,

wenn

$$\mathbf{P}(X = a) = \frac{1}{\#S} \quad \text{für alle } a \in S.$$

Damit beschreibt X eine rein zufällige Wahl aus S.

Beispiel aus Vorlesung 1b:

$$S = \{1, 2, \ldots, g\}^n$$

X := rein zufällige $1, \ldots, g$ - Folge der Länge n.

Eine auf einem endlichen Wertebereich

uniform verteilte Zufallsvariable nennt man auch

diskret uniform verteilt.

Heute lernen wir drei weitere Beispiele

von diskret uniform verteilten Zufallsvariablen kennen:

1. Rein zufällige Permutation
2. Rein zufällige k-elementige Teilmenge
3. Uniform verteilte Besetzung

Bei der Gelegenheit erarbeiten wir auch ein paar

Hilfen fürs Abzählen.

1. Rein zufällge Permutation

(Buch S. 6-8)

1a. Elementares

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?

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 $1, \ldots, n$,

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:

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

1b. Zufällige Permutation und zufälliges Ziehen

Wie kann man sich eine rein zufällige Permutation entstanden denken?

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

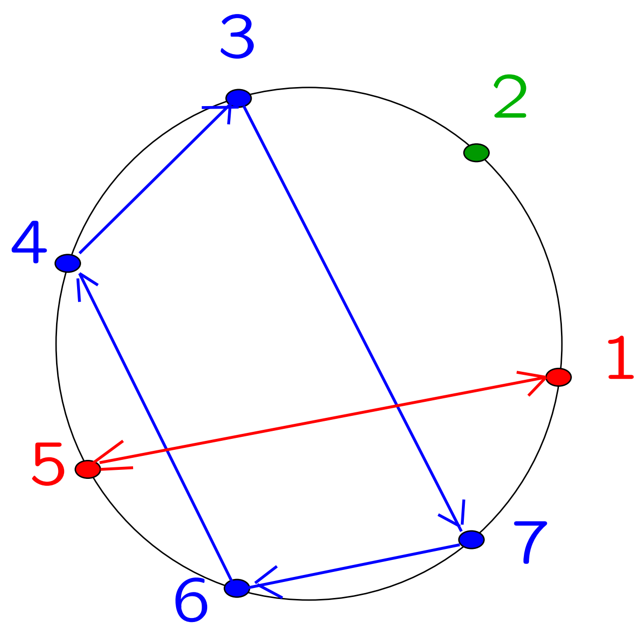
Szenario: eine stets ideal durchmischte Urne mit anfangs n Kugeln, beschriftet mit den Nummern $1, \ldots, n$. Ziehe sukzessive **ohne Zurücklegen** alle n Kugeln und notiere die gezogenen Nummern in Reihenfolge.

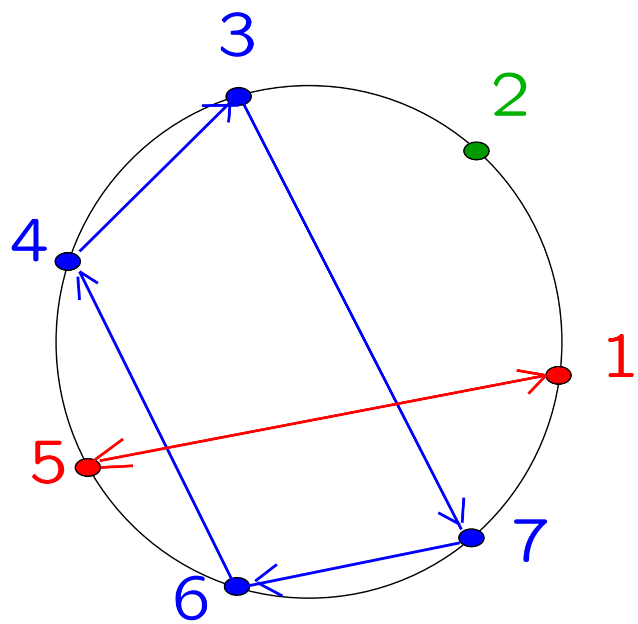
1c. Zyklendarstellung einer Permutation

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

zwei.

Sei X eine rein zufällige Permutation von $1, \ldots, 7$,

Wie wahrscheinlich ist es, dass der die 1 enthaltende Zykel genau die Länge 2 hat?

Wieviele Permutationen von $1, \ldots, 7$ gibt es, bei denen der die 1 enthaltende Zykel genau die Länge 2 hat?

Es gibt davon $6 \cdot 5!$ Stück (warum?)

Also ist die gefragte W'keit: $\dfrac{6!}{7!} = \dfrac{1}{7}$.

Jetzt allgemein:

Für eine Permutation $a \in S$ bezeichne

$$h(a)$$

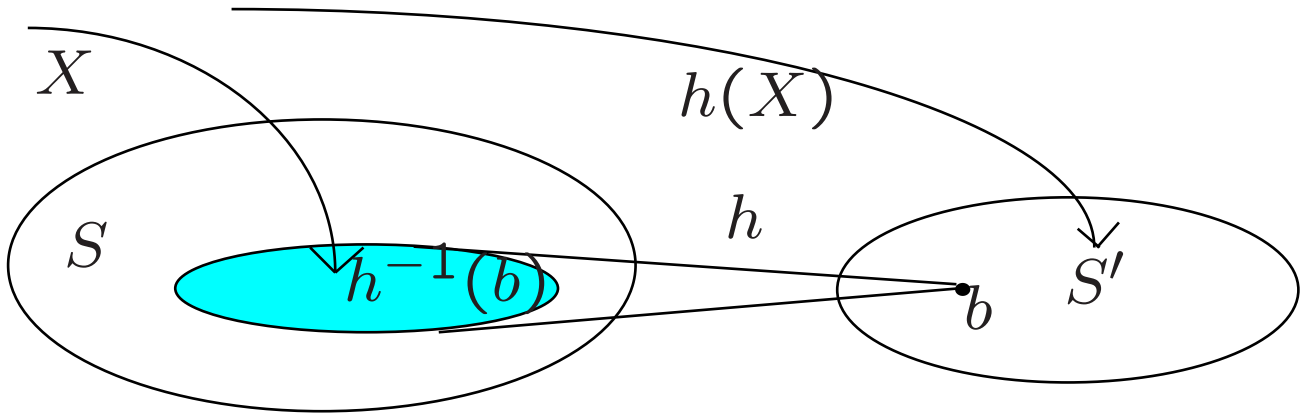
die Länge des Zyklus von a, der die Eins enthält.

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

also eine rein zufällige Wahl aus S,

und sei $b \in \{1, 2, \dots, n\}$.

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



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

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

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

$$\#A = ?$$

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

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

$$= (n-1)!$$

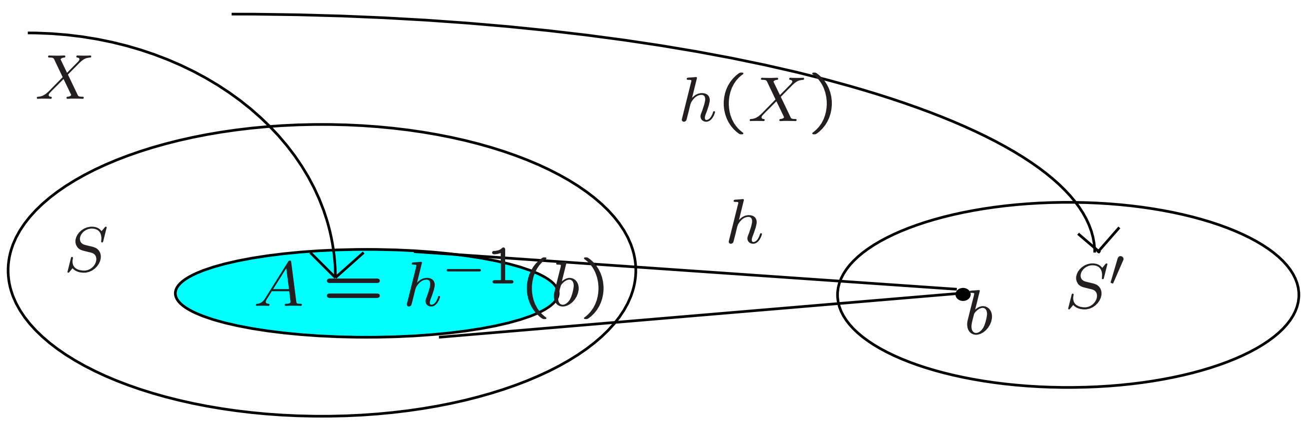
$$\#A = (n-1)!, \qquad \#S = n!$$

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

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

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

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



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

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

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

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

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

Fazit:

Die Länge desjenigen Zyklus

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

der die Eins enthält,

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

2. Rein zufällige Teilmenge einer festen Größe

(Buch S. 9-10)

2a. Die Anzahl der k-elementigen Teilmengen von $\{1, \ldots, n\}$: der Binomialkoeffizient.

Sei $k \leq n$.

Jetzt sei S

die Menge aller k-elementigen Teilmengen von $\{1, \ldots, n\}$.

Wieviele k-elementige Teilmengen von $\{1, \ldots, n\}$ gibt es?

Wieviele Möglichkeiten gibt es, aus n Personen

ein k-köpfiges Komitee ohne Reihung zu bilden?

Wird “nach der Reihe” ausgewählt, dann gibt es

$n(n-1)\cdots(n-(k-1))$ mögliche Wahlprotokolle.

Auf die Reihenfolge kommt es nicht an,

somit führen jeweils $k!$ dieser Wahlprotokolle

auf dieselbe k-elementige Teilmenge.

Also:

$$\#S = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

$$\#S = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

$$= \frac{n!}{k!(n-k)!} =: \binom{n}{k}$$

Binomialkoeffizient „ n über k “

die Anzahl der Möglichkeiten für „ k aus n “

Beispiel: Binomischer Lehrsatz:

$$(x+y)^n = (x+y)(x+y) = \cdots (x+y) = ?$$

Multipliziert man aus, dann ergeben sich 2^n Summanden

(entsprechend den 2^n xy-Folgen der Länge n).

Jeder dieser Summand ist von der Form $x^k y^{n-k}$.

Die Zahl k gibt an, wie oft der Faktor x

im jeweiligen Summanden zum Zug kommt.

Also $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

Pascal'sches Dreieck

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
 & & & & & \vdots & & & & &
\end{array}
$$

Rekursion: $$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

Rekursion: $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$.

Interpretation: Anzahl der Möglichkeiten,

aus n Hessen und einem Bayern

ein $k+1$ köpfiges Komitee auszuwählen.

Entweder der Bayer ist im Komitee...

oder er ist nicht im Komitee...

2b. Rein zufällige Teilmengen

Sei $0 \leq k \leq n$

und sei Y eine rein zufällige k-elementige Teilmenge von $\{1, \ldots, n\}$.

Wie wahrscheinlich ist das Ereignis $\{Y = \{1, \ldots, k\}\}$?

Der Zielbereich von Y ist

$$S := \{t : t \subset \{1, \ldots, n\},\ \#t = k\}\ ,$$

die Menge der k-elementigen Teilmengen von $\{1, \ldots, n\}$.

Wir haben gesehen:

$$\#S = \binom{n}{k}$$

Fazit:

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

2c. Ein Zusammenhang mit rein zufälligen Permutationen

Wie bekommt man eine rein zufällige k-elementige Teilmenge aus einer rein zufälligen Permutation?

Fakt (für $k \leq n$):

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

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

dann ist $\{X_1, \ldots, X_k\}$

eine rein zufällige k-elementige Teilmenge von $\{1, \ldots, n\}$.

Anders gesagt:

Beim Ziehen ohne Zurücklegen

führen die ersten k Züge auf eine

rein zufällige Teilmenge des anfänglichen Reservoirs.

Noch eine Möglichkeit zum Erzeugen

einer rein zufälligen k-elementigen Teilmenge:

Ziehe sukzessive ohne Zurücklegen aus einer Urne

mit k roten und $n - k$ blauen Kugeln.

Notiere die Nummern $X_1, \dots, X_k$ der Züge,

bei denen eine rote Kugel gezogen wird.

Dann ist $\{X_1, \dots, X_k\}$

eine rein zufällige k-elementige Teilmenge von $\{1, \dots, n\}$.

3. Besetzungszahlen

(vgl. Buch S. 10-11)

3a. Begriffsbildung

Sei $a = (a_1, \dots, a_n)$ eine $1, \dots, g$ - Folge der Länge n.

Vorstellung: Objekt Nr. i kommt auf Platz a_i.

Wie oft wird laut Protokoll a der Platz j besetzt?

Anders gesagt: Für wieviele i ist $a_i = j$?

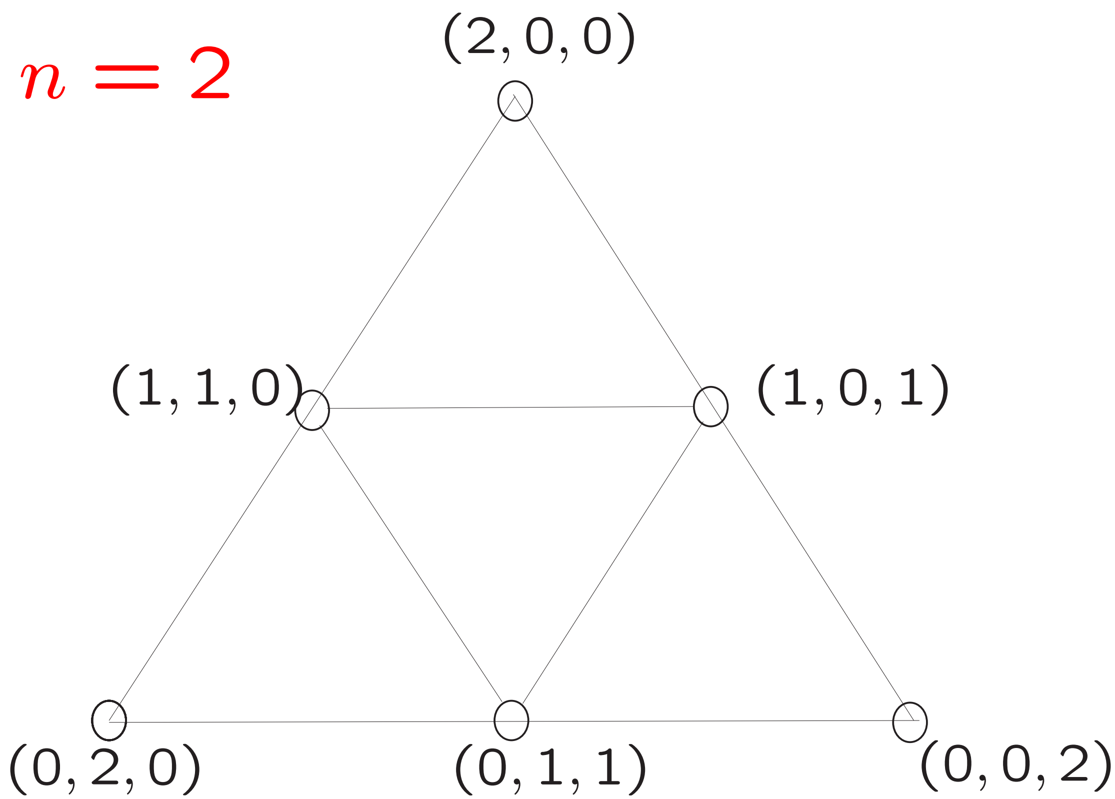
$$b_j(a) := \#\{i : a_i = j, 1 \le i \le n\}$$

Das g-tupel der Besetzungszahlen $b_j(a)$ nennen wir kurz
“die (durch a induzierte) Besetzung”.

In der Vorstellung des
Setzens von n Objekten auf g mögliche Plätze
gibt sie an, wieviele Objekte auf welchem Platz landen
(und unterscheidet nicht, welche Objekte das sind).

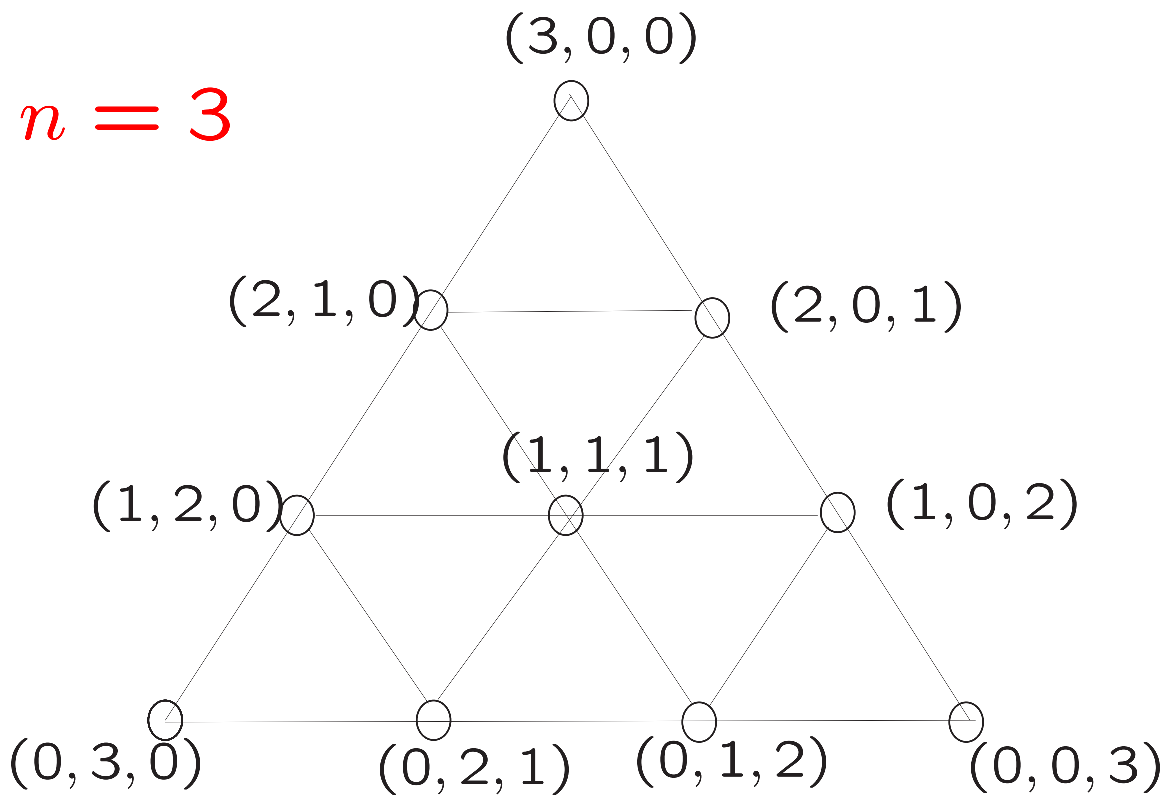
Im Fall $g = 3$ gibt es eine nette Darstellung der Menge der Besetzungen mittels des sogenannten de Finetti-Dreiecks:

$n = 2$



Im Fall $g = 3$ gibt es eine nette Darstellung der Menge der Besetzungen mittels des sogenannten de Finetti-Dreiecks:

$n = 3$



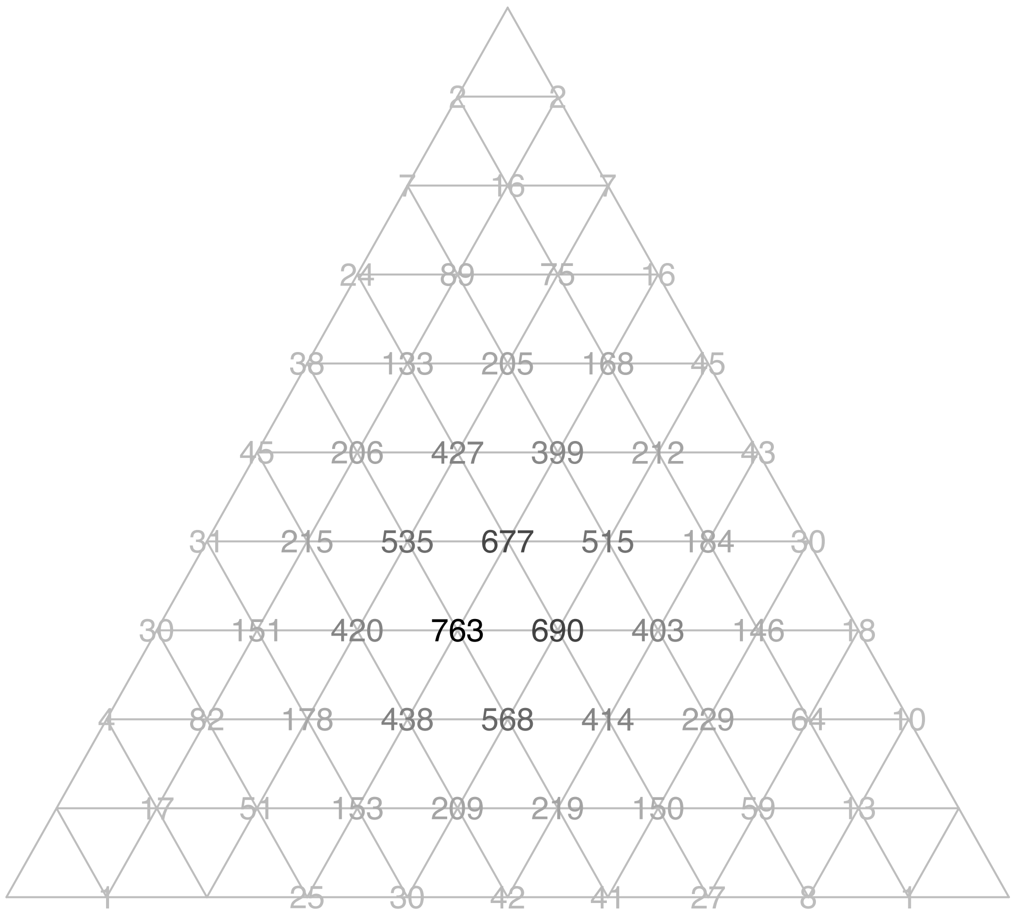
3b. Besetzungszahlen bei sukzessiver rein zufälliger Platzwahl

Machen wir uns ein Bild von der zufälligen Besetzung,
die aus einem auf $\{1,\ldots,g\}^n$ uniform verteilten X entsteht.

Das war das Szenario aus Vorlesung 1b:
n Individuen werden rein zufällig (und “unabhängig”)
auf g mögliche Plätze gesetzt.

Es folgt das Ergebis einer Simulation für $n = 10$ und $g = 3$.
Die Ecken des *de Finettii-Dreiecks* auf der nächsten Folie
entsprechen den Besetzungen
$(10,0,0)$ (oben), $(0,10,0)$ (links) und $(0,0,10)$ (rechts).

Häufigkeiten der Besetzungen bei 10000 Wiederholungen



Wir sehen aus der Simulation:

Die Verteilung dieser zufälligen Besetzung ist
(bei weitem) nicht uniform.

Wir kommen auf diese Verteilung
in der nächsten Vorlesung zurück.

Zum Kontrast betrachten wir jetzt die

3c. Uniform verteilte Besetzung

von g Plätzen mit n Objekten:

Der Zielbereich ist

$$S_{n,g} := \{k = (k_1, \ldots, k_g) : k_j \in \mathbb{N}_0,\ k_1 + \cdots + k_g = n\}$$

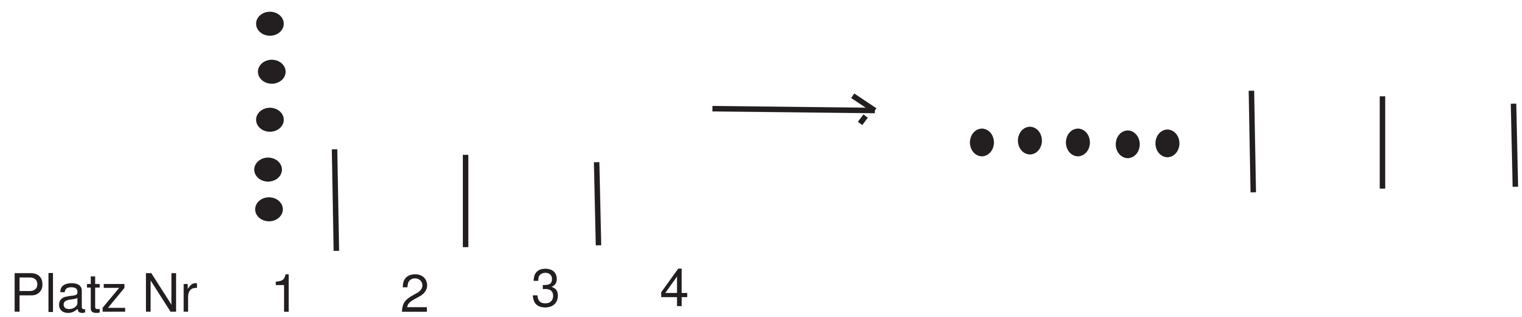
k ist ein g-tupel von Besetzungszahlen, kurz: eine Besetzung.

$$S_{n,g} := \{k = (k_1, \ldots, k_g) : k_j \in \mathbb{N}_0,\ k_1 + \cdots + k_g = n\}$$

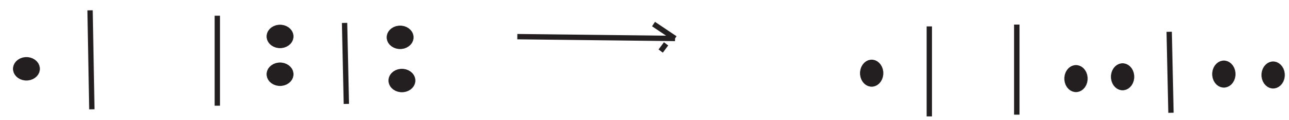
$$\# S_{n,g} = \text{?}$$

Hier hilft ein hübscher Trick, $S_{n,g}$ anders darzustellen.

Beispiele: $g = 4$ Plätze, $n = 5$ Objekte

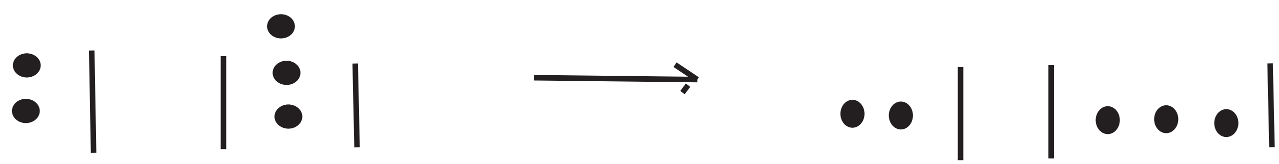


$$(5, 0, 0, 0) \qquad \longrightarrow \qquad (1, 1, 1, 1, 1, 0, 0, 0)$$



Platz Nr 1 2 3 4

$$(1,0,2,2) \longrightarrow (1,0,0,1,1,0,1,1)$$



Platz Nr 1 2 3 4

$$(2,0,3,0) \longrightarrow (1,1,0,0,1,1,1,0)$$

Fakt:

$$h(k_1, k_2, \ldots, k_g) := \underbrace{1\ldots1}_{k_1\text{-mal}}\,0\,\underbrace{1\ldots1}_{k_2\text{-mal}}\,0\ldots0\,\underbrace{1\ldots1}_{k_g\text{-mal}}$$

ist eine bijektive Abbildung von $S_{n,g}$ nach

$S :=$ Menge der 01-Folgen der Länge $n + g - 1$

mit genau n Einsen

$$h(k_1, k_2, \ldots, k_g) := \underbrace{1\ldots1}_{k_1\text{-mal}}0\underbrace{1\ldots1}_{k_2\text{-mal}}0\ldots0\underbrace{1\ldots1}_{k_g\text{-mal}}$$

Voriges Beispiel: $n = 5$, $g = 4$:

$$(k_1, k_2, \ldots, k_4) = (2, 0, 3, 0)$$

$$h(2, 0, 3, 0) = 11001110$$

Der zweite und der vierte Block aus Einsen sind hier leer.

$$h(k_1, k_2, \ldots, k_g) := \underbrace{1\ldots1}_{k_1\text{-mal}}0\underbrace{1\ldots1}_{k_2\text{-mal}}0\ldots0\underbrace{1\ldots1}_{k_g\text{-mal}}$$

Die Länge des j-ten Blocks aus Einsen steht für die Anzahl der Objekte auf Platz j.

Die Blöcke aus Einsen sind durch Nullen getrennt. Die Nullen fungieren als “Trennwände” zwischen den g Plätzen, insgesamt gibt es $g - 1$ solche Trennwände.

$S :=$ Menge der 01-Folgen der Länge $n + g - 1$

mit genau n Einsen

$$\#S = ?$$

$$\#S = \binom{n + g - 1}{n}$$

Also (wegen der Bijektion mittles h) auch:

$$\#S_{n,g} = \binom{n + g - 1}{n}$$

Eine Möglichkeit zum Erzeugen

einer uniform verteilten Besetzung:

Ziehe aus einer Urne mit n weißen und $g-1$ schwarzen

Kugeln sukzessive ohne Zurücklegen.

Notiere 0 beim Zug einer schwarzen

und 1 beim Zug einer weißen Kugel.

Erzeuge so ein rein zufälliges Element aus S.

Übersetze dieses (mit der Umkehrung von h)

in eine rein zufällige Besetzung.

In der Vorlesung haben wir noch eine weitere Möglichkeit durchgespielt, die auf eine uniform verteilte Besetzung $(Z_1, \ldots, Z_g)$ führt – dass das tatsächlich so ist, werden wir im Kapitel über mehrstufige Zufallsexperimente unter der Überschrift *Pólya-Urne* sehen (siehe Vorlesung 10b).

Das Experiment (unter dem Stichwort *Wo Tauben sind, fliegen Tauben zu*) läuft wie folgt: Anfangs sind g Personen auf der Bühne, nummeriert mit $1, \ldots, g$. Sukzessive kommen Ankömmlinge dazu. Ankömmling 1 wählt rein zufällig eine der g Personen und stellt sich zu ihr, damit steigt die Zahl der Personen auf der Bühne auf $g + 1$. Ankömmling 2 wählt rein zufällig eine der $g + 1$ Personen und stellt sich zu ihr, u.s.w. Betrachtet wird dann (bei insgesamt n Ankömmlingen) das g-Tupel $(Z_1, \ldots, Z_g)$, wobei Z_j die Anzahl der Ankömmlinge ist, die bei der Person Nr. j stehen.