

# Generation zufälliger Partitionen

Stefan Volz

12. Juni 2021

## Zusammenfassung

Wir bezeichnen ein  $k$ -Tupel  $(c_1, \dots, c_k) \in \mathbb{N}_0^k$  mit Elementen  $c_i \leq r \in \mathbb{N}$  als geordnete  $r$ -beschränkte  $k$ -Partition von  $n \in \mathbb{N}$ , wenn  $\sum_{j=1}^k c_j = n$ . Bezeichne mit  $\mathcal{P}_k^r(n)$  die Menge dieser Partitionen. Gesucht wird ein effizienter Algorithmus um gleichverteilt ein zufälliges Element aus  $\mathcal{P}_k^r(n)$  zu generieren.

## Inhaltsverzeichnis

<b>1</b>	<b>Unbeschränkte Partitionen</b>	<b>2</b>
1.1	Motivation . . . . .	2
1.2	Geschlossene Form von $\psi_k(n)$ . . . . .	3
<b>2</b>	<b>Beschränkte Partitionen</b>	<b>5</b>
2.1	Motivation . . . . .	5
2.2	Wahrscheinlichkeitsverteilung der beschränkten Partitionen . . .	6
2.3	Effiziente Bestimmung von $P(X_j = m)$ . . . . .	6
2.4	Implementierung . . . . .	7
2.5	Bla . . . . .	8
<b>3</b>	<b>Alternativer Ansatz</b>	<b>8</b>

# 1 Unbeschränkte Partitionen

Betrachten wir zunächst die Menge der unbeschränkten Partitionen  $\mathcal{P}_k^n(n) =: \mathcal{P}_k(n)$ .

**Notation 1.** Wir bezeichnen die Anzahl an unbeschränkten Partitionen mit  $\psi_k(n)$  und nennen  $k$  die Dimension dieser Partitionen.

## 1.1 Motivation

Betrachten wir  $\mathcal{P}_k(n)$  als die Menge der Lösungen  $(x_1, \dots, x_k)$  der Gleichung  $x_1 + \dots + x_k = n$ , so liegen diese geometrisch betrachtet alle auf der durch selbige Gleichung definierten Hyperebene.

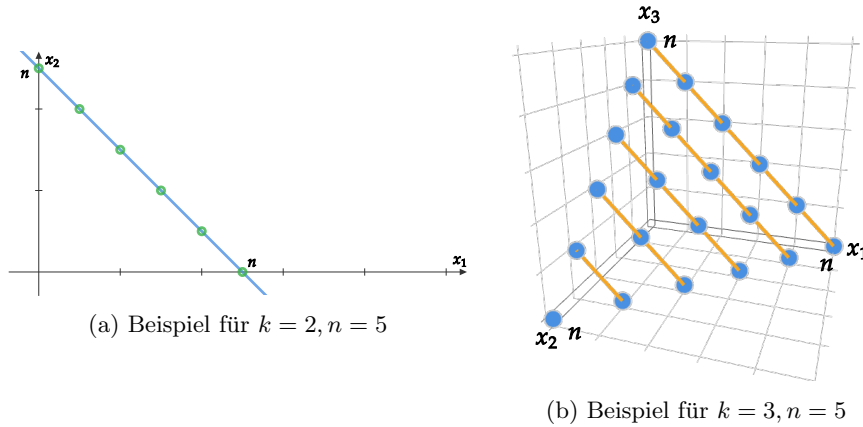


Abbildung 1: Geometrische Veranschaulichung der unbeschränkten Partitionen

Dieses Bild legt nahe, dass sich die Anzahl an 3-dimensionalen Partitionen aus der Anzahl an 2-dimensionalen Partitionen zusammensetzt - also eine rekursive Beziehung zwischen  $\psi_3$  und  $\psi_2$  besteht. Setzen wir z.B.  $x_3$  auf einen festen Wert, so folgt

$$x_1 + x_2 = n - x_3,$$

also ist  $(x_1, x_2)$  eine 2-Partition von  $n - x_3$  wenn  $(x_1, x_2, x_3)$  eine 3-Partition von  $n$  ist. Da  $x_3$  alle Werte in  $\{0, \dots, n\}$  annehmen kann ergibt sich somit

$$\psi_3(n) = \sum_{x_3=0}^n \psi_2(n - x_3) \underset{r:=n-x_3}{=} = \sum_{r=0}^n \psi_2(r). \quad (1)$$

Für den  $\psi_2$  finden wir mittels kombinatorischer Argumente (die Menge  $\{0, \dots, n\}$  hat  $n + 1$  Elemente, wir wählen hieraus eines für  $x_2$  aus und definieren somit

eindeutig eine Partition. Es gibt  $\binom{n+1}{1} = n+1$  Möglichkeiten diese Auswahl zu treffen.)  $\psi_2(n) = n+1$ . Mit Rückblick auf Gleichung 1 schreiben wir dies als

$$\psi_2(n) = n+1 = \sum_{r=0}^n 1 = \sum_{r=0}^n \psi_1(n), \quad (2)$$

wobei wir  $\psi_1(n) = 1$  als trivial wahr erkennen. Dies legt die generelle Formel

$$\psi_{k+1}(n) = \sum_{r=0}^k \psi_k(n), \quad \text{mit } k \in \mathbb{N}, n \in \mathbb{N}_0 \quad (3)$$

nahe, von deren Gültigkeit wir uns auch in höheren Dimensionen leicht überzeugen (Setzen wir einen Freiheitsgrad auf  $r \in \{0, \dots, n\}$  fest so müssen die restlichen eine beliebige  $k$ -Partition von  $n-r$  bilden.).

## 1.2 Geschlossene Form von $\psi_k(n)$

Berechnen wir nach obiger Formel  $\psi_k(n)$  für einige  $k$  erhalten wir:

$$\begin{aligned} \psi_1(n) &= 1 \\ \psi_2(n) &= (n+1) \\ \psi_3(n) &= \frac{1}{2}(n+1)(n+2)(n+3) \\ \psi_4(n) &= \frac{1}{6}(n+1)(n+2)(n+3)(n+4) \end{aligned}$$

was die Vermutung

$$\psi_k(n) = \frac{1}{(k-1)!} \prod_{r=1}^{k-1} (n+r) \quad (4)$$

nahelegt. Da es nicht völlig trivial ist die Gültigkeit dieser Formel anhand Gleichung 3 zu beweisen, wollen wir sie mittels generierenden Funktionen herleiten. Sei daher

$$G_n(x) := \sum_{k=1}^{\infty} \psi_k(n) x^k$$

die Familie der gewöhnlichen generierenden Funktionen den Folgen  $(\psi_k(n))_{k \in \mathbb{N}}$  für  $n \in \mathbb{N}_0$ . Multiplizieren wir Gleichung 3 beidseitig mit  $x^k$  und summieren über ihren Gültigkeitsbereich in  $k$  erhalten wir

$$\begin{aligned} \frac{1}{x} \sum_{k=1}^{\infty} \psi_k(n) x^k - \psi_1(n) &= \sum_{k=1}^{\infty} \sum_{r=0}^n x^k \psi_k(n) = \sum_{r=0}^n \sum_{k=1}^{\infty} x^k \psi_k(n) \\ \iff G_n(x) &= x \left( \sum_{r=0}^n G_r(x) - \psi_1(n) \right). \end{aligned}$$

Durch abspalten des letzten Terms der rechtsseitigen Summe und anschließendes Auflösen nach  $G_n(x)$  erhalten wir

$$G_n(x) = \frac{x}{1-x} \left( \sum_{r=0}^{n-1} G_r(x) + \psi_1(n) \right). \quad (5)$$

Wir berechnen die  $G_n$  mittels dieser Formel und

$$G_0(x) \underset{\psi_1(n)=1}{=} \sum_{k=1}^{\infty} x^k = \frac{1}{1-x} - 1 = \frac{x}{1-x}$$

für einige  $n$ :

$$\begin{aligned} G_1(x) &= \frac{x}{(1-x)^2} \\ G_2(x) &= \frac{x}{(1-x)^3} \\ G_3(x) &= \frac{x}{(1-x)^4}. \end{aligned}$$

Es liegt nahe, dass

$$G_n(x) = \frac{x}{(1-x)^{n+1}}. \quad (6)$$

Sei  $g(x) := \frac{1}{1-x}$ . Nach der Cauchy-Produktformel gilt

$$g(x)^2 = \frac{1}{1-x} \cdot \frac{1}{1-x} = \sum_{k=0}^{\infty} \sum_{r=0}^k x^r x^{k-r} = \sum_{k=0}^{\infty} (k+1)x^k = Dg(x).$$

Somit gilt  $D(g^n(x)) = ng^{n-1}(x)Dg(x) = ng^{n+1}(x)$  woraus wir

$$\begin{aligned} g^{n+1}(x) &= \frac{1}{n} Dg^n(x) = \frac{1}{n(n-1)} D^2 g^{n-1}(x) = \dots = \frac{1}{n!} D^n g(x) \\ \iff D^k g^n(x) &= (n)^{\bar{k}} g^{n+k}(x) \end{aligned}$$

erhalten, wobei  $(n)^{\bar{k}}$  die steigende Faktorielle bezeichnet. Somit folgt für die Taylorentwicklung von  $g^n$  um 0

$$g^n(x) = \sum_{k=0}^{\infty} \frac{D^k g^n(0)}{k!} x^k = \sum_{k=0}^{\infty} \frac{(n)^{\bar{k}} g^{n+k}(0)}{k!} x^k = \sum_{k=0}^{\infty} \frac{(n)^{\bar{k}}}{k!} x^k,$$

womit wir für  $G_n$  folgende Reihendarstellung erhalten

$$G_n(x) = \sum_{k=0}^{\infty} \frac{(n+1)^{\bar{k}}}{k!} x^{k+1} = \sum_{k=1}^{\infty} \frac{(n+1)^{\bar{k-1}}}{(k-1)!} x^k.$$

Durch Koeffizientenvergleich mit der ursprünglichen Definition der generierenden Funktion erhalten wir die in Gleichung 4 bereits vermutete Darstellung

$$\psi_k(n) = \frac{(n+1)^{\overline{k-1}}}{(k-1)!} = \frac{1}{(k-1)!} \prod_{r=0}^{k-2} (n+1+r) = \frac{1}{(k-1)!} \prod_{r=1}^{k-1} (n+r) \quad (7)$$

welche für  $k \geq 1, n \geq 0$  gültig ist.

## 2 Beschränkte Partitionen

Wenden wir uns nun den beschränkten Partitionen zu.

**Notation 2.** Wir bezeichnen die Anzahl an  $r$ -beschränkten  $k$ -Partitionen mit  $\varphi_k^r(n)$ .

### 2.1 Motivation

Wie auch zuvor können wir uns die beschränkten Partitionen geometrisch veranschaulichen.

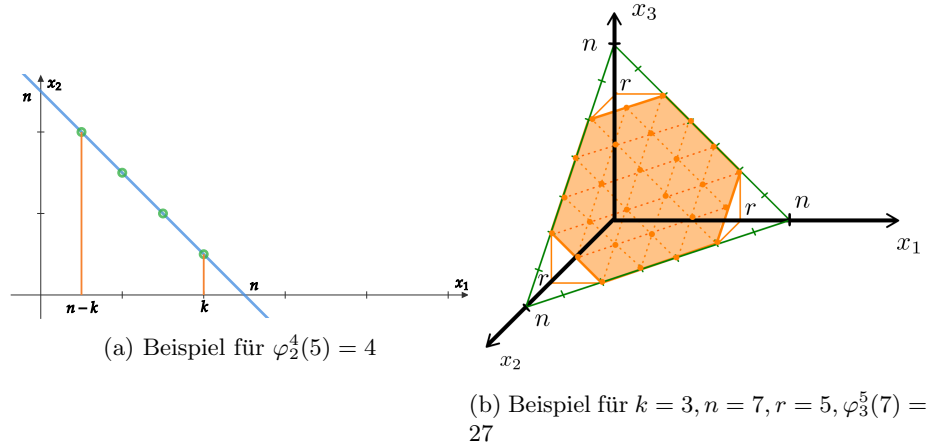


Abbildung 2: Geometrische Veranschaulichung der beschränkten Partitionen

Wir nehmen sozusagen alle unbeschränkten Partitionen, und schneiden dann die kleineren Teile weg, bei denen einzelne Komponenten die Schranke  $r$  überschreiten würden.

## 2.2 Wahrscheinlichkeitsverteilung der beschränkten Partitionen

Diese Überlegungen können wir direkt in eine Gleichung übersetzen und erhalten somit

$$\varphi_k^{r-1}(n) = \psi_k(n) - k\psi_k(n-r) \text{ mit } k \geq 1, n, n-r \geq 0.$$

Hierbei erklärt sich der Index  $(r-1)$  dadurch, dass wir die Elemente der "Schnittgrenze" mitnehmen wollen. Nehmen wir nun an, dass alle Partitionen gleichwahrscheinlich sind, so erhalten wir ein Laplaceexperiment. Es sei  $\Omega := \mathcal{P}_k^r(n)$  die Ergebnismenge, sowie  $\omega := (X_1, \dots, X_k) \in \Omega$  eine Zufällige Partition. Da  $X_j = m$  genau dann wenn die restlichen Komponenten eine  $(k-1)$ -Partition bilden (also  $(X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_k) \in \mathcal{P}_k^r(n-m)$ ), gibt es  $\varphi_{k-1}^r$  Möglichkeiten für die Konfiguration von  $\omega$ . Es gilt demnach

$$p_k^r(n, m) := P(X_j = m) = \frac{\varphi_{k-1}^r(n-m)}{\varphi_k^r(n)}. \quad (8)$$

## 2.3 Effiziente Bestimmung von $P(X_j = m)$

Da in 7 und somit auch in 8 viele Fakultäten vorkommen ist diese Darstellung nicht praktikabel umsetzbar. Wir beginnen damit die Situation erstmal noch etwas schlimmer zu machen indem wir die fallende Faktorielle mittels Fakultäten ausdrücken und erhalten

$$\varphi_k^{r-1}(n) = \frac{(n+k-1)!}{n!(k-1)!} - k \frac{(n-r+k-2)!}{(n-(r+1))!(k-1)!} = \frac{\Gamma(n+k)}{n!\Gamma(k)} - k \frac{\Gamma(n-r+k-1)}{\Gamma(n-r)\Gamma(k)}$$

Wenden wir auf die beiden Terme jeweils  $\exp \circ \ln$  an, ergibt sich mit  $\ln F := \ln \circ (!)$

$$\begin{aligned} \varphi_k^{r-1}(n) &= \exp(\ln \psi_k(n)) - \exp(\ln(k\psi_k(n-r))) \\ &= \exp(\ln \Gamma(n+k) - \ln F(n) - \ln \Gamma(k)) \\ &\quad - \exp(\ln(k) + \ln \Gamma(n+k-(r+1)) - \ln \Gamma(n-r) - \ln \Gamma(k)). \end{aligned}$$

Die Funktion  $\ln \Gamma$  lässt sich über eine asymptotische Entwicklung auf Basis der verallgemeinerten Stirlingformel

$$\Gamma(x) = \sqrt{\frac{2\pi}{x}} \left(\frac{x}{e}\right)^x e^{\mu(x)} \text{ mit } x > 0, 0 < \mu(x) < \frac{1}{12x}$$

gut und effizient approximieren, es gilt

$$\ln \Gamma(x) = (x - \frac{1}{2}) \ln x - x + \frac{1}{2} \ln 2\pi + \frac{1}{12x} + \frac{1}{360x^3} + \mathcal{O}(x^{-4}).$$

Da selbst eine Taylorentwicklung höherer Ordnung von  $f : (x, y) \mapsto e^x - e^y$  (z.B.  $T_2 f(x, y, u, v) = e^u(1 + x - u + \frac{(x-u)^2}{2}) - e^v(1 + y - v + \frac{(y-v)^2}{2})$  mit  $u, v = \frac{|a-b|}{2}$ :

$T_2 f(x, y) = e^{\frac{|b-a|}{2}} (a-b)(1 + \max\{a, b\})$ , uns für ausreichend große Argumente in praktisch unbrauchbaren Größenordnungen lässt, ersetzen wir  $\exp$  in diesen Ausdrücken durch die Taylorentwicklung erster Ordnung und erhalten somit

$$\begin{aligned} \varphi_k^{r-1}(n) &\doteq \ln \Gamma(n+k) - \ln F(n) - \ln \Gamma(k) \\ &\quad - \ln k + \ln \Gamma(n+k-(r+1)) - \ln \Gamma(n-r) - \ln \Gamma(k) \end{aligned}$$

womit wir  $P(X_j = m)$  effizient berechnen können. Hiermit können wir uns eine Zufallsverteilung generieren<sup>1</sup> und auf Basis dieser Zufallszahlen erzeugen.

## 2.4 Implementierung

Sofern man wirklich formal korrekt gleichverteilte vollständige Partitionen erzeugen will müsste man in jedem Schritt  $j$  als Verteilung die Wahrscheinlichkeiten  $P(X_j = m | X_{j-1} = x_{j-1}, \dots, X_1 = x_1), 0 \leq m \leq r$  nutzen.

Der tatsächlich implementierte Algorithmus geht wie folgt vor

```

Data:  $n, m, k \in \mathbb{N}$ 
Result: zufällige Partition  $(x_1, \dots, x_m)$  von  $n$  sodass  $x_j \leq k$ 
Gewichte  $\leftarrow$  Feld für  $n+1$  natürliche Zahlen auf 0 initialisiert
for  $i \leftarrow 0$  to  $\min\{k, n\}$  do
    | Gewicht( $i$ )  $\leftarrow p_m^k(\text{Rest}, i)$ 
end
do
    Partition =  $(x_1, \dots, x_m) \leftarrow$  Feld für  $m$  natürliche Zahlen auf 0
    initialisiert
    Rest  $\leftarrow n$ 
    for  $i \leftarrow 1$  to  $m-1$  do
        while True do
             $r \leftarrow \text{sample}(\text{Gewichte})$  #Generiere Zufallszahl nach
            gegebener Verteilung
            if  $r \leq \text{Rest}$  then
                Rest  $\leftarrow \max\{0, \text{Rest} - r\}$ 
                 $x_i \leftarrow r$ 
                break
            end
        end
    end
     $x_m \leftarrow \text{Rest}$ 
while  $\max(\text{Partition}) \leq k$  #Die generierte Partition ist gültig;
return Partition

```

<sup>1</sup>Sofern es zu teuer ist die ganze Verteilung zu erzeugen könnte man ggf. durch lückenweises Sampling und anschließende Renormalisierung einzelne Zahlen erhalten. Man erzeuge also für  $\nu \in \mathbb{N}$  zunächst  $M := \{(\nu m, p_k^r(n, \nu m)) | 0 \leq \nu m \leq n\}$ ,  $N := \{p(\cdot, p) \in M\}$  und hieraus die Wahrscheinlichkeitsverteilung  $\{(m, \frac{p}{\sum_{\rho \in N} \rho}) | (m, p) \in M\}$ . Ggf. kann man nach Erzeugen einer Zahl iterativ mit einem kleineren  $\nu$  eine Auswahl auf dem durch die erste Zufallszahl bestimmten Teilintervall durchführen.

## 2.5 Bla

Es sei

$$\rho_k^r(n, m) := \begin{cases} \varphi_{k-1}^r(n-m) & k > 1 \\ \delta_{n,m} & k = 1 \end{cases}, 0 \leq m \leq n$$

die Anzahl an  $r$ -beschränkten  $k$ -Partitionen von  $n$  bei denen ein Element den Wert  $m$  besitzt (der Zähler von Gleichung 8).

## 3 Alternativer Ansatz

Stellen wir zunächst  $n \in \mathbb{N}$  als Summe  $\underbrace{1 + 1 + \dots + 1}_{n \text{ mal}}$  dar. Diese Summe enthält  $(n-1)$  *plus*-Symbole. Wählen wir aus diesen  $k-1$  aus und entfernen alle restlichen, so erhalten wir  $k \in \mathbb{N}$  Blöcke aus Einsen welche wir zusammenaddieren können um eine  $k$ -Partition von  $n$  zu erhalten - umgedreht können wir aus einer Partition auch immer so eine Folge konstruieren. Unser Auswahlargument bedeutet somit, dass

$$\psi(n) = \binom{n-1}{k-1}.$$

Wir rechnen leicht nach, dass dies mit Gleichung 7 übereinstimmt. Aus [1] erhalten wir weiterhin folgendes Ergebnis

**Satz 1.** *Es sei  $A \subseteq \mathbb{N}_0$ ,  $k \in \mathbb{N}$  und  $(a_n)_{n \in \mathbb{N}_0}$  die Folge der Anzahl  $A$ -elementigen  $k$ -Partitionen von  $n$ . Dann ist die generierende Funktion  $F$  von  $(a_n)$  gegeben durch*

$$F(x) = \left( \sum_{\alpha \in A} x^\alpha \right)^k.$$

Hieraus folgt direkt

**Korollar 1.** *Die generierende Funktion  $\Phi_k^r$  von  $(\varphi_k^r(n))_{n \in \mathbb{N}_0}$  lautet*

$$\Phi_k^r(x) = \left( \sum_{i=0}^r x^i \right)^k = \sum_{\sum_{i=1}^k \kappa_i = k} \binom{k}{\kappa_1, \dots, \kappa_k} \cdot \prod_{i=1}^k x^{\kappa_i}.$$

*Beweis.* Die Gleichung

$$\Phi_k^r(x) = \left( \sum_{i=0}^r x^i \right)^k$$



folgt mit  $A = \{0, \dots, r\}$  direkt aus Satz 1. Anwenden des Multinomialsatzes liefert uns

$$\begin{aligned} \left( \sum_{i=0}^r x^i \right)^k &= \sum_{\sum_{i=0}^r \kappa_i = k} \binom{k}{\kappa_0, \dots, \kappa_r} \cdot \prod_{i=0}^r (x^i)^{\kappa_i} \\ &= \sum_{\sum_{i=0}^r \kappa_i = k} \binom{k}{\kappa_0, \dots, \kappa_r} \cdot x^{\sum_{i=0}^r i \kappa_i} \\ &= \sum_{\substack{(\kappa_0, \dots, \kappa_r) \\ \in \mathcal{P}_{r+1}(k)}} \binom{k}{\kappa_0, \dots, \kappa_r} \cdot x^{\sum_{i=0}^r i \kappa_i}. \end{aligned}$$

□

In der aktuellen Implementierung gewinnen wir Verteilungen durch explizites bestimmen der kompletten Polynome für kleine  $r$ . Für große wählen wir eine grobe Näherung mittels symmetrischer Dreiecksverteilung. Evtl. interessant wäre die z.B. in [3] beschriebene numerische Bestimmung der Polynomkoeffizienten mittels der aus der Cauchy'schen Integralformel folgenden Formel

$$f^{(n)}(a) = \frac{n!}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z-a)^{n+1}} dz$$

für eine auf einer offenen Teilmenge  $U \subseteq \mathbb{C}$  der komplexen Zahlen holomorphen Funktion  $f$  und einer Kreiskurve  $\gamma$  mit  $\text{Bild}(\gamma) \subset D$ , oder auch Lanczos's Formel [2]. Hierbei evtl. auch von Interesse ist [4].

## Literatur

- [1] “Composition (combinatorics)”. In: *Wikipedia* (02.05.2021). URL: [https://en.wikipedia.org/wiki/Composition\\_\(combinatorics\)](https://en.wikipedia.org/wiki/Composition_(combinatorics)).
- [2] C. W. Groetsch. “Lanczo’s Generalized Derivative”. In: *The American Mathematical Monthly* (1998). URL: <https://www.jstor.org/stable/2589707>.
- [3] Toshio Hosono. “Numerical algorithm for Taylor series expansion”. In: *Electronics and Communications in Japan Part 1: Communications* (1986). URL: <https://onlinelibrary.wiley.com/doi/epdf/10.1002/ecja.4410690602>.
- [4] J. N. Lyness und L. M. Delves. “On Numerical Contour Integration Round a Closed Contour”. In: *AMS* 21.100 (1967). URL: <https://www.ams.org/journals/mcom/1967-21-100/S0025-5718-1967-0229388-0/S0025-5718-1967-0229388-0.pdf>.