Der Bildchensammler

von Christopher K. Lazik

Der Bildchensammler

Viele kennen dieses beliebte “Spiel”, welches gerne zu Vermarktungszwecken bei großen Sportevents oder neuerdings auch von Supermarktketten genutzt wird, um die Menschen zum Kauf anzuregen. Es handelt sich um das Sammeln von Abziehbildchen, indem man undurchsichtige Umschläge mit einer gewissen Anzahl von diesen Bildchen kauft, um Stück für Stück alle verschiedenen Bildchen mindestens einmal zu besitzen. Nun stellt sich hier natürlich die Frage, wie viele Bildchen man insgesamt ungefähr kaufen müsste, um wahrscheinlich alle wenigstens einmal in seinem Besitz zu haben.

Im Folgenden wollen wir nun zeigen, wo der Erwartungswert an zu kaufenden Bildchen liegt.

Ein genauso funktionierendes Beispiel ist das Ziehen von n Kugeln, welche jeweils eine von insgesamt n Farben besitzen. Wobei hier die Kugeln nach jedem Ziehen zurückgelegt werden und gut gemischt werden. Angenommen wir haben bereits k verschiedene Kugeln gezogen, dann ist die Wahrscheinlichkeit <u>keine</u> neue Kugel zu ziehen in unserem nächsten Zug

$$\frac{k}{n}$$

Sollten wir also zum Beispiel 5 verschiedene Kugeln in einer undurchsichtigen Urne haben und bereits mehrfach Kugeln herausgezogen und nach anschließendem Zurücklegen gut gemischt haben, dann könnte es sein, dass wir bereits zwei verschiedene Kugelfarben mindestens einmal in der Hand halten konnten. Somit würden wir mit einer Wahrscheinlichkeit von $\frac{2}{5}$ erneut eine der beiden bereits gezogenen Farben ziehen und somit keine neue Kugel ziehen.

Die Wahrscheinlichkeit genau s Ziehungen zu brauchen, um eine neue Kugel zu ziehen ist genau:

$$\left(\frac{k}{n}\right)^{s-1} \cdot \left(1-\frac{k}{n}\right)$$

Wobei die erste Hälfte der Formel die Ziehungen beschreibt, in denen wir keine neue Kugel ziehen und die zweite Hälfte genau den Zug der neuen Kugel beschreibt (also genau das Gegenteil der ersten Hälfte).

Behaupten wir mit dem bereits verwendeten Beispiel, dass wir genau in drei Zügen die dritte Kugelfarbe und damit eine neue Kugel ziehen, liegt also die Wahrscheinlichkeit dafür bei:

$$\left(\frac{2}{5}\right)^{3-1} \cdot \left(1 - \frac{2}{5}\right)$$
$$= \left(\frac{2}{5}\right)^{2} \cdot \left(\frac{3}{5}\right)$$
$$= \left(\frac{4}{25}\right) \cdot \left(\frac{3}{5}\right)$$
$$= \left(\frac{12}{125}\right)$$

Sei V_s die Zufallsvariable, die die Anzahl der Ziehungen zur nächsten neuen Kugel darstellt, dann muss für unseren Erwartungswert $E[V_s]$ nach der Definition für Erwartungswerte gelten:

$$E[V_s] = \sum_{i \in I} s_i \cdot P(V_s = s_i)$$

Ergo die Summe aller möglichen Anzahlen der Ziehungen multipliziert mit der Wahrscheinlichkeit dieser jeweiligen Ziehung. Unsere abzählbare Indexmenge wird nun zu den natürlichen Zahlen ≥ 1 (wir können ja schließlich enorm oft Kugeln ziehen, müssen jedoch mindestens eine Kugel ziehen, um eine neue zu ziehen).

$$\sum_{s \geq 1} s \cdot P(V_s = s)$$

Jetzt zählen wir unsere möglichen Anzahlen nur Schritt für Schritt hoch. Nun müssen wir nur noch $P(V_s = s)$ durch die oben genannte Formel ersetzen.

Der Erwartungswert der erwarteten Ziehungen bis zur nächsten neuen Kugel liegt dann bei:

$$E[V_s] = \sum_{s \geq 1} \left(\frac{k}{n}\right)^{s-1} \cdot \left(1 - \frac{k}{n}\right) \cdot s$$
$$= \frac{1}{1 - \frac{k}{n}}$$

Für diese Umformung entwickeln wir folgende Reihe.

$$\sum_{s\geq 1} (1-x)\cdot x^{s-1}\cdot s$$
$$=\sum_{s\geq 1} x^{s-1}\cdot s-\sum_{s\geq 1} x^{s}\cdot s$$
$$=\sum_{s\geq 0} x^{s}\cdot (s+1)-\sum_{s\geq 0} x^{s}\cdot s$$
$$=\sum_{s\geq 0} x^{s}$$
$$=\frac{1}{1-x}\text{ (geometrische Reihe)}$$

Nun kennen wir zwar die erwartete Anzahl Ziehungen zur nächsten neuen Kugel doch wollten wir ja ursprünglich ermitteln wie viele Bildchen wir insgesamt kaufen müssen, um alle mindestens einmal zu besitzen. Also in unserem Beispiel die erwartete Anzahl Ziehungen, um jede der Farben einmal gezogen zu haben.

Somit berechnen wir nun die erwartete Anzahl Ziehungen, um alle Bildchen mindestens einmal zu besitzen.

$$\sum_{k=0}^{n-1} \frac{1}{1-\frac{k}{n}}$$

Wir addieren hierfür die erwarteten Werte jede nächste neue Kugel zu ziehen. Dies führt uns zur gesamten Summe aller jeweiligen erwarteten Ziehungen und dadurch zur erwarteten Anzahl an Ziehungen, um alle Kugeln als neue Kugel gezogen zu haben. Haben wir also tatsächlich fünf Kugeln ist die erste Kugel ganz sicher eine neue, also $\frac{1}{\left(1-\frac{0}{5}\right)}=1$ wir brauchen also eine Ziehung für die erste neue Kugel. Für die zweite neue Kugel addieren wir die erwartete Anzahl hinzu usw., bis wir bei $k=n-1$ angekommen sind. Ist $k=n$ haben wir alle Kugeln bereits gezogen. Wir ziehen also nicht weiter. Wir addieren also keine weiteren Werte.

Man kann leicht sehen, dass

$$\sum_{k=0}^{n-1} \frac{1}{1-\frac{k}{n}}$$

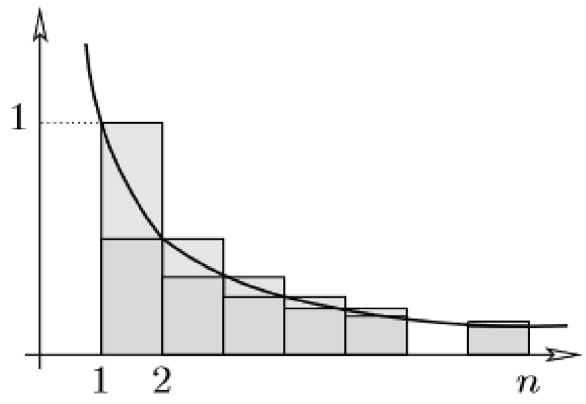
$$= \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{1}$$

$$= n \cdot H_n$$

Dies ist n -mal die harmonische Reihe und es gilt:

$$n \cdot H_n \approx n \cdot \ln(n)$$

Diese Abschätzung der harmonischen Reihe kann wie folgt gezeigt werden. Sie befindet sich auf Seite 13 im "Buch der Beweise" (Auflage 5).



$H_n = \sum_{k=1}^{n} \frac{1}{k}$

Zunächst betrachten wir nur die dunkel schraffierte Fläche.

$$H_n - 1 = \sum_{k=2}^{n} \frac{1}{k} < \int_1^n \frac{1}{t}\,\mathrm{dt} = \ln(n)$$

Nun betrachten wir die gesamte schraffierte Fläche (also helle und dunkle Flächen).

$$H_n - \frac{1}{n} = \sum_{k=1}^{n-1} \frac{1}{k} > \int_1^n \frac{1}{t}\,\mathrm{dt} = \ln(n)$$

Es folgt also.

$$H_n \approx \ln(n)$$

Somit ist der Erwartungswert, also die Antwort auf die Frage des Bildchensammlers, bewiesenermaßen:

$$n \cdot \ln(n)$$

Nun ist das Problem des Bildchensammlers zwar gelöst, jedoch ist eventuell noch etwas fern in welcher Relation diese mathematische Antwort zu unserer tatsächlichen Welt steht.

Angenommen eine verbreitete Supermarktkette bietet in einem gewissen Zeitraum eine Sammelbildchenaktion an. Den Kunden wird die Möglichkeit geboten bei jedem Einkauf einige Bildchen zu erhalten, um ihre Sammlung zu vervollständigen. Die Aktion soll auf Tierschutz aufmerksam machen und eingenommenes Geld gespendet werden. Es gibt insgesamt 188 verschiedene Tiermotive.

Wir rechnen also:

$$188 \cdot \ln(188) \approx 984{,}45$$

Wir erwarten also 985 dieser Bildchen erhalten zu müssen, um eine vollständige Sammlung zu besitzen. Nehmen wir an es ist möglich pro 10€ Einkaufswert 5 Bildchen zu erhalten. Da $\frac{985}{5} = 197$ und $197 \cdot 10€ = 1970€$, müssen wir Einkäufe mit einem insgesamten Wert von 1970€ tätigen. Allerdings gilt dies auch nur, wenn wir immer Einkaufswerte, welche durch 10 teilbar sind bezahlen müssen. Andernfalls kommen natürlich noch einige Eurobeträge zu der bereits hohen Summe hinzu. *(Es wird deutlich warum Supermärkte solche Aktionen anbieten)*

Wollen wir nun noch aus zusätzlichem Interesse abschätzen wie wahrscheinlich es ist deutlich mehr Ziehungen als $n \cdot \ln(n)$ Ziehungen zu brauchen. Sei V_n die zufällige Variable, die die Anzahl der Ziehungen repräsentiert. Wobei der Erwartungswert$E[V_n] \approx n \cdot \ln(n)$ ist.

Somit ist für $n \geq 1$, $c \geq 0$ und $m := \lceil n \cdot \ln(n) + cn \rceil$die Wahrscheinlichkeit deutlich mehr als m Ziehungen zu brauchen:

$$\text{Prob}[V_n > m] \leq e^{-c}$$

Denn sei $A_i := \{\text{Ereignis Kugel } i \text{ wird in den ersten } m \text{ Ziehungen nicht erwischt}\}$, dann gilt

$$\text{Prob}[V_n > m] = \text{Prob}[\cup A_i] \leq \sum_i \quad \text{Prob}[A_i] = n\left(1 - \frac{1}{n}\right)^m < n \cdot e^{-\frac{m}{n}} \leq e^{-c}$$

Dies gilt, da $\left(1 - \frac{1}{n}\right)^n$ monoton steigend ist und für $n \geq 1$ gegen $\frac{1}{e}$ konvergiert.

$$\begin{aligned}
\left(1 - \frac{1}{n}\right)^n &= \left(\frac{n}{n} - \frac{1}{n}\right)^n = \left(\frac{n-1}{n}\right)^n \\
&= \left(\frac{n}{n-1}\right)^{-n} \\
&= \left(\frac{n+1-1}{n-1}\right)^{-n} = \left(\frac{n-1}{n-1} + \frac{1}{n-1}\right)^{-n} = \left(1 + \frac{1}{n-1}\right)^{-n} \\
&= \left(\left(1 + \frac{1}{n-1}\right)^n\right)^{-1}
\end{aligned}$$

Bekanntermaßen ist $\lim\limits_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e$, also gilt $\lim\limits_{n\to\infty} \left(\left(1 + \frac{1}{n}\right)^n\right)^{-1} = \frac{1}{e}$.

Quellen:

1. M.Aigner, G.M.Ziegler. “Der Bildchensammler” aus Buch der Beweise, 5.Auflage, Berlin,Heidelberg: Springer Verlag, 2018
2. M.Aigner, G.M.Ziegler. Buch der Beweise S. 13, 5.Auflage, Berlin,Heidelberg: Springer Verlag, 2018