Die Wahrscheinlichkeit, dass Eugen und Henry in derselben Unterrichtsstunde anwesend waren

Henry S. Grasshorn Gebhardt

6. September 2008

Insgesamt gab es N Unterrichtsstunden. Die Wahrscheinlichkeit, dass Eugen anwesend war in einer Unterrichtsstunde ist p, für Henry ist sie q. Das heißt, Eugen war insgesamt Np Mal anwesend, Henry Nq Mal. Eugen war $N(1-p)$ Mal nicht da, Henry $N(1-q)$ Mal.

Die Anzahl von Eugens Anwesenheitsmöglichkeiten berechnet sich folgendermaßen. Man stelle sich alle N Unterrichtsstunden an einer Kette gereiht vor. Es gibt N Möglichkeiten Eugens erste Anwesenheitsstunde auf dieser Kette zu verteilen. Für die zweite Anwesenheitsstunde gibt es nur noch $N-1$ Möglichkeiten, für die dritte nur noch $N-2$. Für die letzte, die Np-te, gibt es nur noch $N-Np+1$ Möglichkeiten. Insgesamt gibt es also

$$N \cdot (N-1) \cdot (N-2) \ldots (N-Np+1) = \frac{N!}{(N-Np)!}$$

Möglichkeiten Eugens Anwesenheitsstunden auf die Unterrichsstunden zu verteilen.

Da jede Anwesenheitsstunde ununterscheidbar von jeder anderen Anwesenheitsstunde ist, muss durch die Anzahl der Permutationen $(Np)!$ geteilt werden. Die Anzahl der möglichen Anwesenheitsprofile Eugens ist also

$$\frac{N!}{(N-Np)!(Np)!}.$$

Für Henry gilt dasselbe, wenn p durch q ersetzt wird.

Es gibt also insgesamt

$$\frac{N!}{(N-Np)!(Np)!} \cdot \frac{N!}{(N-Nq)!(Nq)!}$$

verschiedene Möglichkeiten, wie Eugen und Henry die Unterrichtsstunden besucht haben könnten.

Um zu berechnen wie viele Möglichkeiten es gibt, dass sich Eugen und Henry in den l Unterrichtsstunden $k_1, \ldots, k_l$ trafen, aber in keinen anderen, stelle man sich wieder alle Unterrichtsstunden an einer Kette gereiht vor. Zunächst werden die l Unterrichtsstunden $k_1, \dots, k_l$ besetzt. Es bleiben noch $N - l$ Unterrichtsstunden übrig. Von Eugens Anwesenheitsstunden bleiben noch $Np - l$ übrig, von Henrys $Nq - l$.

Es gilt also $Np-l+Nq-l$ Anwesenheitsstunden auf $N-l$ Unterrichtsstunden zu verteilen. Dafür gibt es

$$\frac{(N-l)!}{(N-l-(Np-l+Nq-l))!}$$

Möglichkeiten. Eugens Anwesenheitsstunden sind unter sich ununterscheidbar, Henrys unter sich auch, aber eine Anwesenheitsstunde Eugens unterscheidet sich von einer Henrys. Also muss für die gerade verteilten Anwesenheitsstunden Eugens und Henrys separat durch ihre Anzahl von Permutationen geteilt werden. Es gibt also insgesamt

$$\frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}$$

Möglichkeiten, dass Eugen und Henry beide gemeinsam in den l Unterrichtsstunden $k_1, \dots, k_l$ saßen, aber keine weiteren gemeinsamen Unterrichtsstunden hatten.

Die Anzahl der Möglichkeiten diese l Unterrichtsstunden anzuordnen ist

$$\frac{N!}{(N-l)!l!}.$$

Es gibt also

$$\frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}$$

Möglichkeiten, dass Eugen und Henry genau l-mal in derselben Unterrichtsstunde saßen.

Eugen und Henry treffen sich maximal $l_{max} = \min(Np, Nq)$ mal, mindestens jedoch $l_{min} = \max(0, Np + Nq - N)$ mal. Das kleinste l bei dem Eugen und Henry sich mindestens einmal treffen, ist $\max(1, Np + Nq - N)$. Die Anzahl der Möglichkeiten, sodass Eugen und Henry mindestens einmal in derselben Unterrichtsstunde sitzen ist also

$$\sum_{l=\max(1,Np+Nq-N)}^{\min(Np,Nq)} \frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}.$$

Die Wahrscheinlichkeit, dass Eugen und Henry sich also getroffen haben ist

$$w = \frac{\sum_{l=\max(1,Np+Nq-N)}^{\min(Np,Nq)} \frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}}{\frac{N!}{(N-Np)!(Np)!} \cdot \frac{N!}{(N-Nq)!(Nq)!}}. \tag{1}$$

Beispiel: $N = 15$, $p = 0.4$, $q = 0.33$.
$Np = 6$, $N - Np = 9$.
$Nq = 5$, $N - Nq = 10$.
$l_{min} = 0$, $l_{max} = 5$.
$\frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}$ für $l = 0$ ist 630630. (unnötig)
$\frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}$ für $l = 1$ ist 3783780.
$\frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}$ für $l = 2$ ist 6306300.
$\frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}$ für $l = 3$ ist 3603600.
$\frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}$ für $l = 4$ ist 675675.
$\frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!}$ für $l = 5$ ist 30030.
$\sum_{l=1}^{5} \frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!} = 14399385$.
$\frac{N!}{(N-Np)!(Np)!} \cdot \frac{N!}{(N-Nq)!(Nq)!} = 15030015$.
Die Wahrscheinlichkeit ist also 0.958042.

Zusatzfrage: Gilt

$$\sum_{l=\max(0,Np+Nq-N)}^{\min(Np,Nq)} \frac{N!}{(N-l)!l!} \cdot \frac{(N-l)!}{(N-l-(Np-l+Nq-l))!(Np-l)!(Nq-l)!} = \frac{N!}{(N-Np)!(Np)!} \cdot \frac{N!}{(N-Nq)!(Nq)!}?$$

Vereinfachung: Gl. (1) kann folgendermaßen vereinfacht werden.

$$w = \sum_{l=\max(1,Np+Nq-N)}^{\min(Np,Nq)} \frac{(N-Np)!(Np)!(N-Nq)!(Nq)!}{l!(N+l-Np-Nq)!(Np-l)!(Nq-l)!}. \tag{2}$$