

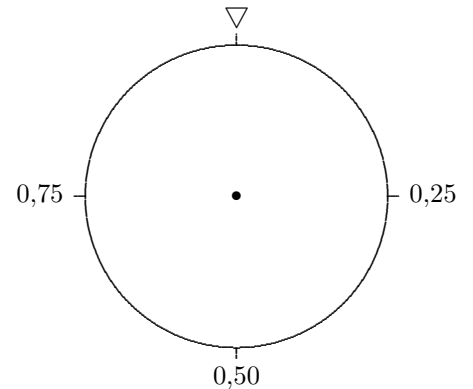
Optimales Stoppen

Eine junge Frau möchte unter den nächsten zehn Männern, die sie kennenlernen wird, einen Lebenspartner auswählen. Hierzu wird sie die Männer bewerten (in Prozent) und sofort eine Entscheidung treffen, ob sie den Mann akzeptiert und damit die Suche beendet oder ob sie den Mann ablehnt, auf ihn kann sie dann später nicht mehr zurückkommen. Die junge Frau wird also nicht erst alle Männer ausprobieren und sich dann entscheiden. Welche Strategie gibt ihr die größte Chance, den Besten auszuwählen?

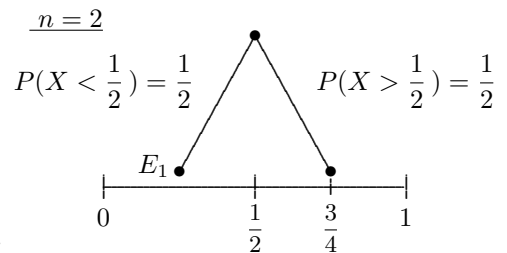
Als mathematisches Modell betrachten wir ein Glücksrad, das n -mal gedreht werden darf. Nach jedem Einzelversuch kann die erzeugte Zufallszahl X als Gewinn akzeptiert und damit das Experiment beendet oder der Versuch fortgesetzt werden.

Sei E_n die Gewinnerwartung für n Versuche.
Offenbar gilt $E_1 = \frac{1}{2}$.

Nehmen wir nun an, wir dürfen zweimal drehen.
Für die Zufallszahl X der ersten Drehung ist $X > \frac{1}{2}$ mit der Wahrscheinlichkeit $\frac{1}{2}$.
In diesem Fall breche ich das Experiment ab.
Meine Gewinnerwartung ist die Mitte des Intervalls $[\frac{1}{2}, 1]$.
Der Erwartungswert für $n = 2$ ist damit:

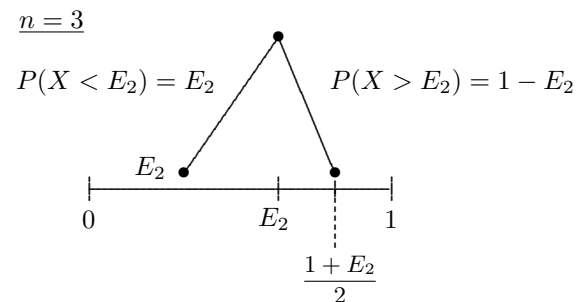


$$E_2 = \frac{1}{2} \cdot \frac{3}{4} + \frac{1}{2} \cdot E_1 = \frac{5}{8}$$



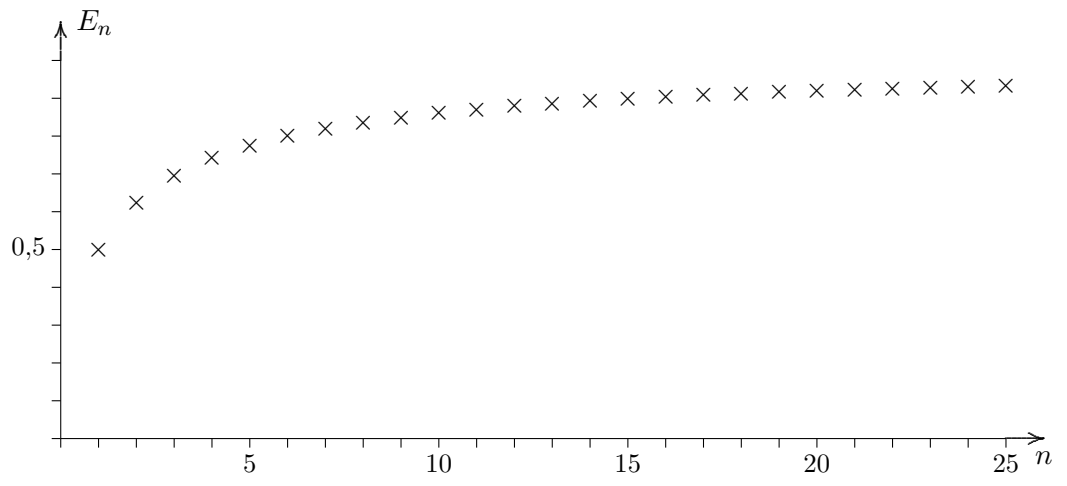
Für $n = 3$ stoppe ich nach der 1. Drehung, falls $X > \frac{5}{8}$.

$$\begin{aligned} E_3 &= \frac{1 + E_2}{2} \cdot (1 - E_2) + E_2^2 \\ &= \frac{1 - E_2^2}{2} + E_2^2 \\ &= \frac{1 + E_2^2}{2} \\ E_{n+1} &= \frac{1 + E_n^2}{2} \end{aligned}$$



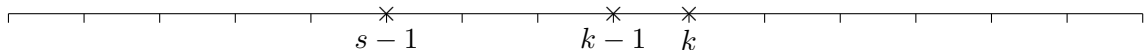
n	1	2	3	4	5	6	7	8	9	10
E_n	0,5	0,625	0,695	0,742	0,775	0,800	0,820	0,836	0,850	0,861

Für die junge Frau ist nun klar: Der n -te Kandidat (bei 10 beginnend) wird akzeptiert, falls die Bewertung über dem Erwartungswert E_{n-1} liegt. Für $X = E_{n-1}$ ist es gleichgültig, wie sie sich entscheidet.



Optimales Stoppen

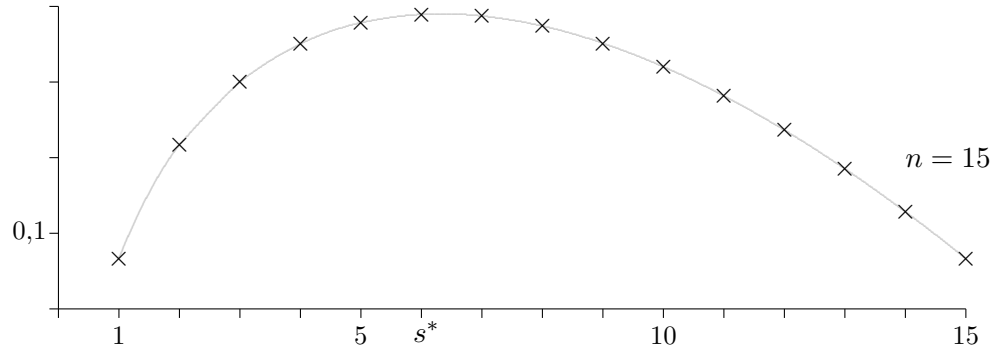
Das obige Problem soll noch auf eine zweite Weise (grob) modelliert werden. Die Warteschlange besteht aus n Kandidaten. Die ersten $s - 1$ werden durch Zahlen bewertet. Das hierbei erreichte Maximum M dient der nun beginnenden Auswahl. Sie endet, sobald ein Kandidat mit einer höheren Bewertung gefunden wird (alle Bewertungen sind verschieden). Auf abgelehnte Kandidaten kann später nicht wieder zurückgegriffen werden. Wie viele Kandidaten sind für den Vortest zur Ermittlung von M erforderlich, um mit möglichst großer Wahrscheinlichkeit den optimalen Kandidaten (mit potentieller maximaler Bewertung) mit dieser Strategie zu ermitteln? Befindet er sich unter den ersten $s - 1$, so wird er übersehen. Nehmen wir an, dass er die Stelle k einnimmt, $k \geq s$.



Er wird genau dann entdeckt, wenn der beste unter den ersten $k - 1$ Kandidaten bereits unter den ersten $s - 1$ vorkommt. Die Wahrscheinlichkeit für dieses Ereignis (die Reihenfolge ist zufällig) beträgt $p_k = \frac{1}{n} \frac{s-1}{k-1}$ und die Wahrscheinlichkeit $P(s)$, an den optimalen Kandidaten zu gelangen, daher:

$$P(s) = \sum_{k=s}^n p_k = \frac{s-1}{n} \sum_{k=s}^n \frac{1}{k-1}, \quad s \geq 2, \quad P(1) = \frac{1}{n}$$

Für welches s ist $P(s)$ maximal?



Für das Maximum gilt: $P(s^* - 1) < P(s^*) > P(s^* + 1)$

oder umgeformt: $\frac{1}{s^*} + \frac{1}{s^*+1} + \dots + \frac{1}{n-1} < 1 < \frac{1}{s^*-1} + \frac{1}{s^*} + \dots + \frac{1}{n-1}$

$$\begin{aligned} & P(s^* - 1) < P(s^*) \\ \Leftrightarrow & (s^* - 2) \left(\frac{1}{s^*-2} + \frac{1}{s^*-1} + \dots + \frac{1}{n-1} \right) < (s^* - 1) \left(\frac{1}{s^*-1} + \frac{1}{s^*} + \dots + \frac{1}{n-1} \right) \\ \Leftrightarrow & (s^* - 2) \cdot \frac{1}{s^*-2} + (s^* - 2) \left(\frac{1}{s^*-1} + \frac{1}{s^*} + \dots + \frac{1}{n-1} \right) < \dots \\ \Leftrightarrow & 1 < \underbrace{((s^* - 1) - (s^* - 2))}_1 \left(\frac{1}{s^*-1} + \frac{1}{s^*} + \dots + \frac{1}{n-1} \right) \end{aligned}$$

Die zweite Ungleichung wird entsprechend nachgewiesen. Zu jedem n gibt es genau ein s^* , z.B. $n = 20$, $s^* = 8$, $P(s^*) = 0,384$. 7 Kandidaten muss man passieren lassen.

Roof's

Optimales Stoppen

Um das asymptotische Verhalten von $P(s^*)$ und s^* zu untersuchen, wird die Näherung

$$\sum_{k=s}^{n-1} \frac{1}{k} \approx \int_s^n \frac{1}{x} dx = \ln n - \ln s = \ln \frac{n}{s}$$

herangezogen.

$$P(s) = \frac{s-1}{n} \sum_{k=s}^n \frac{1}{k-1} = \frac{s-1}{n} \sum_{k=s-1}^{n-1} \frac{1}{k} \approx \frac{s}{n} \sum_{k=s}^{n-1} \frac{1}{k} \approx \frac{s}{n} \ln \frac{n}{s}$$

Mit der Differentialrechnung folgt: $s^* \approx \frac{n}{e}$, $P(s^*) = \frac{1}{e} = 0,368$

Faustregel: 37% der Kandidaten muss man passieren lassen.

Dasselbe Ergebnis kann auch mit der Doppelungleichung auf der vorigen Seite erzielt werden.

$$P(s^*) \approx \frac{s^*}{n} \sum_{k=s^*}^{n-1} \frac{1}{k} \approx \frac{s^*}{n} \cdot 1 \approx \frac{s^*}{n} \ln \frac{n}{s^*} \implies \ln \frac{n}{s^*} \approx 1$$

...

Die bearbeitete Fragestellung wird in der Literatur als Sekretärinnen-Problem bezeichnet.