DAS ZWEI-ZETTEL-SPIEL – EIN PARADOXON UND EINIGE SEINER VERWANDTEN

HANS HUMENBERGER, WIEN

Zusammenfassung: *In diesem Beitrag soll das so genannte „Zwei-Zettel-Spiel“ analysiert werden. Gibt es eine Strategie, mit der man bei der Entscheidung zwischen zwei Zahlen (man kennt nur eine der beiden) mit einer Wahrscheinlichkeit > ½ die größere erwischt? Wenn ja, welche? Überraschenderweise gibt es so eine. Es wird versucht dieses Phänomen mit möglichst elementaren Mitteln[1] zu beschreiben und zu erklären. Darüber hinaus werden einige „enge Verwandte“ dieses Zwei-Zettel-Spiels vorgestellt und diskutiert.*

1 Einleitung

In der Stochastik gibt es immer wieder Situationen, in denen man sich nicht sicher ist, ob man „richtig liegt“. Dies ist für viele Lehrkräfte ein Grund, Stochastik nicht gerne und nur im durch den jeweiligen Lehrplan vorgesehenen Mindestmaß zu unterrichten, für andere wiederum ein faszinierender Aspekt: Schon bei relativ einfach scheinenden Problemformulierungen kann es verschiedene plausible Ansätze geben, die aber zu unterschiedlichen Resultaten führen. Dann ist zu klären, welcher stimmt und wo genau der (Denk-) Fehler des anderen liegt. Da kann es viele Diskussionsmöglichkeiten und auch viel Diskussionsbedarf im Unterricht geben.

Es ist schon sehr viel über stochastische „Paradoxa“ geschrieben worden, auch über deren didaktischen Wert, so dass ich mich hier sehr kurz halten möchte: Sie können hohes Motivationspotential haben, das auch im Mathematikunterricht genutzt werden soll, nach dem Motto: „Dem muss ich jetzt auf den Grund gehen!“, „Was ist da los?“, „Das will ich jetzt wirklich wissen!“.

Bei Paradoxa geht es meist um ein der Intuition bzw. dem gesunden Menschenverstand widersprechendes Resultat. Sie lassen sich i. A. meist „aufklären“, d. h. nach Durchschauen der Situation ist diese dann gar nicht mehr paradox, die anfänglichen Widersprüche haben sich aufgelöst.

2 Das Problem (Phänomen) des Zwei-Zettel-Spiels

Einem „Spieler“[2] winkt folgende Chance auf einen Preis durch einen „Gönner“: Der Gönner hat zwei Zettel, schreibt je eine Zahl darauf (zwei verschiedene reelle Zahlen) und gibt diese in je einen Umschlag. Der Spieler gewinnt den Preis, wenn er die größere der beiden Zahlen erwischt. Er wählt zufällig einen Umschlag aus, öffnet ihn und sieht die Zahl auf dem Zettel: *U*. Er hat allerdings noch die Chance zu wechseln und sich für den anderen Umschlag zu entscheiden (natürlich ohne diesen aufzumachen und die andere Zahl zu sehen).

Soll er nun wechseln? Immer? Nie? Manchmal? Hat Wechseln hier überhaupt einen Einfluss? Kann man der „Ungewissheit ein Schnippchen schlagen“ (vgl. BRUSS 2000) und eine bessere Strategie als bloßes Raten entwickeln?

Auf den ersten Blick scheint eine solche Strategie nicht möglich zu sein, denn:
Wenn in einer Urne zwei verschiedene Zahlen liegen, so erwischt man bei zufälliger Ziehung einer Zahl mit Wahrscheinlichkeit ½ die kleinere bzw. größere. Dieses ist in keiner Weise zu beeinflussen, so dass es immer gleichgültig ist, ob man tauscht oder nicht, die Wahrscheinlichkeit, die größere Zahl zu erwischen, bleibt immer bei ½.

3 Das Paradoxon und einige seiner Facetten

Es gibt für den Spieler überraschender Weise doch eine Strategie, mit einer Wahrscheinlichkeit echt größer als ½ die größere Zahl zu erwischen. Dies ist natürlich gegen unsere erste Intuition und ist somit ein klassischer Fall für ein „Paradoxon“. Hier eine möglichst einfache Beschreibung der Lösung.

U sei die Zahl, die der Spieler auf dem Zettel des geöffneten Umschlages sieht. Dies ist natürlich eine Zufallsvariable, denn die Auswahl durch den Spieler – das Ziehen des Umschlages – geschieht ja zufällig.

Nach Besichtigung von *U* macht der Spieler ein Zufallsexperiment, das eine Zufallszahl *Z*

(Zufallsvariable) liefert. Diese Zufallsvariable soll zwei Bedingungen erfüllen:

1) Z ist unabhängig von U
2) Die Verteilung von Z hat eine *Dichte* (Z ist also insbesondere eine stetige Zufallsvariable), die auf $\mathbb{R}$ nirgendwo verschwindet. Solche Verteilungen gibt es ja viele, z. B. die Normalverteilung.

Strategie: Bei $Z \leq U$ bleibe bei U,
bei $Z > U$ wähle die andere Zahl

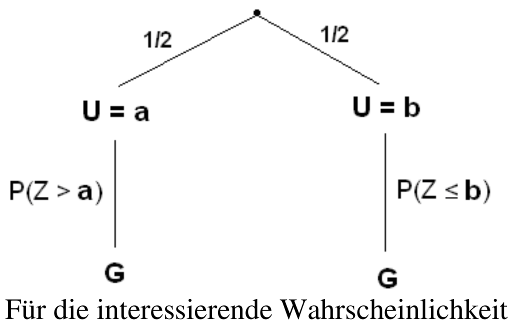
Diese garantiert ihm dann eine Gewinnwahrscheinlichkeit von mehr als ½. Dies klingt zunächst unglaublich: Wie kann und soll die Realisierung einer nahezu beliebigen Zufallsvariable Z (einzige Bedingungen: überall positive Dichte, Unabhängigkeit von U) Informationsgehalt über das in Rede stehende Problem haben: ob die noch unbekannte Zahl nun eher die kleinere oder die größere sein wird? Diese Zufallsvariable hat ja mit den geschriebenen Zahlen gar nichts zu tun, oder doch?

3.1 Zwei feste Zahlen

Bei unserer ersten Version wollen wir die beiden Zahlen in den Umschlägen als „fest“ betrachten (vgl. z. B. MEESTER 2003, S. 102f). Es gibt zwei Zahlen $a < b$, in den zwei Umschlägen liegt je eine von beiden. Nun kann es sein, dass die Ziehung des Spielers entweder a oder b bringt:

$$\mathrm{P}(U = a) = \frac{1}{2} = \mathrm{P}(U = b)$$

Für das Ereignis „Gewinnen“ (Gewinnsituation G) ergibt sich also ganz unabhängig von der konkreten Verteilung bzw. Dichte der Zufallsvariable Z folgendes Baumdiagramm:

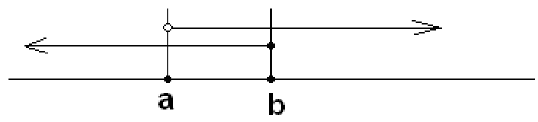


Für die interessierende Wahrscheinlichkeit ergibt sich:

$$\mathrm{P}(G) =$$
$$\mathrm{P}(Z > a \mid U = a) \cdot \underbrace{\mathrm{P}(U = a)}_{1/2} + \mathrm{P}(Z \leq b \mid U = b) \cdot \underbrace{\mathrm{P}(U = b)}_{1/2}$$

Wegen der Unabhängigkeit von Z und U können die Bedingungen weggelassen werden und es gilt:

$$\mathrm{P}(G) = \underbrace{\left(\mathrm{P}(Z > a) + \mathrm{P}(Z \leq b)\right)}_{>1} \cdot \frac{1}{2} \quad \boxed{> \frac{1}{2}}$$



Durch die „Überlappung“ der beiden Bereiche[3] zwischen a und b kommt es zu einer Wahrscheinlichkeitssumme $\mathrm{P}(Z > a) + \mathrm{P}(Z \leq b)$ von strikt größer als 1!

Die Dichte der Zufallsvariable Z verschwindet nach Voraussetzung auf $\mathbb{R}$ nirgendwo, so dass in Bezug auf Z dem Bereich zwischen a und b sicher eine positive Wahrscheinlichkeit zukommt. Es ist damit auch sofort klar, wie viel größer als ½ die Gewinnwahrscheinlichkeit ist:

$$\mathrm{P}(G) = \frac{1}{2} + \frac{1}{2}\mathrm{P}(a < Z \leq b)\,.$$

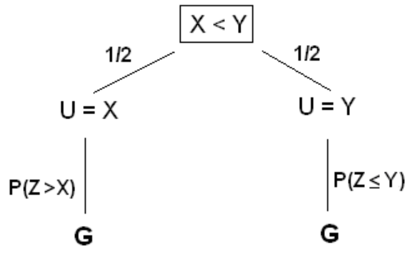
D. h. man kommt mit der „Z-Strategie“ so weit über ½ hinaus wie die halbe Wahrscheinlichkeit, dass Z zwischen a und b fällt.

3.2 Die Zahlen des Gönners als Zufallsvariablen

Wir stellen uns nun die Auswahl der beiden Zahlen durch den Gönner als Realisierung von Zufallsvariablen vor: X sei die erste Zahl, die der Gönner auf einen Zettel schreibt, Y die zweite. Wegen der vorausgesetzten Ungleichheit der beiden Zahlen gibt es nur die beiden Fälle $X < Y$ und $Y < X$.
In beiden Fällen ist es möglich, dass die Ziehung des Spielers X oder Y bringt, d. h. $U = X$ oder $U = Y$.

Die Zufallsvariable Z ist unabhängig von U, X, Y.

Wir betrachten zunächst den Fall $X < Y$:



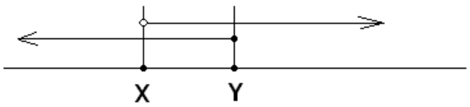
Für die interessierende Gewinnwahrscheinlichkeit ergibt sich bei der Z-Strategie:

$$\mathrm{P}(G \mid X < Y) = \frac{1}{2} \cdot \mathrm{P}(Z > X \mid U = X,\ X < Y) + \frac{1}{2} \cdot \mathrm{P}(Z \le Y \mid U = Y,\ X < Y)$$

Die Realisierung von U (also ob das Ereignis $U = X$ oder $U = Y$ eintritt) wird ausschließlich durch die Zufallsauswahl des Spielers festgelegt, ist also sicher unabhängig von den Ereignissen $Z > X$ und $Z \le Y$. Daher darf man die Bedingungen $U = X$ und $U = Y$ oben auch weglassen und es ergibt sich ganz analog zu 3.1 wieder eine Gewinnwahrscheinlichkeit von echt größer als ½ :

$$\mathrm{P}(G \mid X < Y) = \frac{1}{2} \cdot \underbrace{\big(\mathrm{P}(Z > X \mid X < Y) + \mathrm{P}(Z \le Y \mid X < Y)\big)}_{1+\mathrm{P}(X<Z\le Y \mid X<Y)}$$

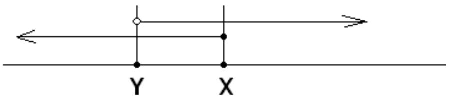
$$= \frac{1}{2} + \frac{1}{2} \cdot \mathrm{P}(X < Z \le Y \mid X < Y)$$



Ganz analog kommt man auch im Fall $Y < X$ zu dieser Überlappung im Bereich zwischen *X* und *Y*:

$$\mathrm{P}(G \mid Y < X) = \frac{1}{2} \cdot \underbrace{\big(\mathrm{P}(Z > Y \mid Y < X) + \mathrm{P}(Z \le X \mid Y < X)\big)}_{1+\mathrm{P}(Y<Z\le X \mid Y<X)}$$

$$= \frac{1}{2} + \frac{1}{2} \cdot \mathrm{P}(Y < Z \le X \mid Y < X)$$



Damit ergibt sich insgesamt für die Gewinnwahrscheinlichkeit bei der *Z*-Strategie:

$$\boxed{\mathrm{P}(G)} = \underbrace{\mathrm{P}(G \mid X < Y)}_{>1/2} \cdot \mathrm{P}(X < Y) + \underbrace{\mathrm{P}(G \mid Y < X)}_{>1/2} \cdot \mathrm{P}(Y < X) \ \boxed{>}$$

$$> \frac{1}{2} \underbrace{\big(\mathrm{P}(X < Y) + \mathrm{P}(Y < X)\big)}_{1} = \boxed{\frac{1}{2}}$$

Bemerkungen:

- Hier ist es gleichgültig, ob die Zufallsvariablen *X* und *Y* (Zahlen, die der Gönner schreibt) als abhängig oder unabhängig voneinander betrachtet werden. Es spielt auch keine Rolle, ob sie als identisch oder verschieden verteilt angenommen werden. Man braucht hier keine Voraussetzung der Art $\mathrm{P}(X < Y) = 1/2 = \mathrm{P}(Y < X)$.
- Durch Vergleich von 3.1 und 3.2 sieht man: Es macht keinen Unterschied, ob man die beiden Zahlen in den Umschlägen als fest oder als Ausprägungen von Zufallsvariablen ansieht, die „*Z*-Strategie" funktioniert in beiden Fällen.

Eine mögliche **Computersimulation** in diesem Zusammenhang kann die Theorie in konkreten Situationen bestätigen. Das heißt natürlich noch nicht, dass damit ein allgemeiner Beweis erbracht sei, dass die Z-Strategie „immer" so gut (bzw. schlecht) funktionieren würde. Einer Anregung von Dr. Jörg Meyer (Hameln) folgend habe ich mit EXCEL folgende Simulation durchgeführt (10.000-mal). Hier wird *Z* als standardnormal-verteilte Zufallsvariable betrachtet und die Zufallszahlen des Gönners als unabhängige und gleichverteilte Zufallszahlen aus dem Intervall $[-10^6;10^6]$:

Wir arbeiten in Zeile 2, so dass Zeile 1 für Überschriften frei bleibt. Wähle in Spalte A und B zwei gleichverteilte Zufallszahlen aus $[-10^6;10^6]$. Eine Zufallszahl aus [0;1] in Spalte C entscheidet, welche der beiden gewählt wird (*U*), diese wird in Spalte D geschrieben, die nicht gewählte in Spalte E. Nun kommt in Spalte F wieder eine Zufallszahl aus [0;1], die als Argument für Spalte G dient: =NORMINV(F2;0;1), dies ergibt die gewünschte *N*(0;1)-verteilte Zufallszahl *Z* in Spalte G. Nun kommt in Spalte H die eventuelle Umwahl, nämlich wenn die gewählte Zahl in Spalte D kleiner ist als *Z* in Spalte G: =WENN(D2<G2;E2;D2). In I schreibt man die aktuell (d. h. nach evtl. Umwahl) nicht gewählte Zahl. In J wird eine 1 gesetzt, wenn H2 > I2, sonst 0 (=WENN(H2>I2;1;0)). Dann braucht man nur mehr irgendwo die Einsen in Spalte J zu zählen (=SUMME(J2:J10001)). Es ergibt sich eine Erfolgsquote von ca. 75%. Dies wundert einen auch nicht wirklich, denn „im Durchschnitt" liefert die Normalverteilung ja den Wert 0, und wenn man immer mit dem festen Wert $z = 0$ vergleichen würde (statt mit der normalverteilten Zufallsvariable *Z*), so kann man leicht ausrechnen, dass sich dabei eine Erfolgswahrscheinlichkeit von genau 75% ergibt (vgl. Abschnitt 4).

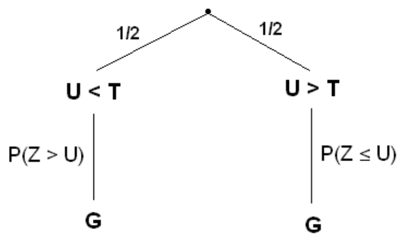
Natürlich könnte man bei einer Computersimulation auch andere („spannendere") Verteilungen als die „Gleichverteilung" für die Zahlen *X*, *Y* des Gönners betrachten, und auch andere Verteilungen als die *N*(0;1)-Verteilung für *Z*.

3.3 Gegendarstellung

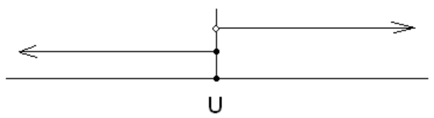
Von „Immer-noch-Zweiflern“ könnte relativ leicht eine Gegendarstellung konstruiert werden, die versucht, dieses Überlappungsargument geschickt zu entkräften, und zwar mit Hilfe eines ähnlichen Baumdiagramms wie in 3.1 oder 3.2, nur mit *anderen* Zufallsvariablen: nicht mit X und Y (erste und zweite Zahl, die der Gönner auf Zetteln schreibt), sondern mit U (Zahl im gewählten **U**mschlag) und T (Zahl im anderen Umschlag: „**T**auschen“).

Wir nehmen an, dass U und T *voneinander unabhängige* und *identisch verteilte* Zufallsvariablen sind mit $U \neq T$ (der Gönner wähle die beiden Zahlen unabhängig voneinander und nach demselben Zufallsmechanismus). Dann ist allein aus Symmetriegründen klar (ohne formalen Beweis): $\mathrm{P}(U < T) = \frac{1}{2} = \mathrm{P}(U > T)$. Die Zufallsvariable Z (hängt nur vom Spieler ab) kann sicher als unabhängig von U und T angesehen werden (diese hängen nur vom Gönner und von der Zufallsauswahl durch den Spieler ab).

Somit ergibt sich für die „Gewinnsituation“ G folgendes Baumdiagramm bei der Z-Strategie:



Hier kommt es zu keiner Überlappung mehr:



$$\mathrm{P}(G) = \frac{1}{2} \cdot \mathrm{P}(Z > U) + \frac{1}{2} \cdot \mathrm{P}(Z \leq U) =$$

$$\frac{1}{2} \cdot \underbrace{\left(\mathrm{P}(Z > U) + \mathrm{P}(Z \leq U)\right)}_{=1} = \frac{1}{2}$$

So gesehen nutzt die Z-Strategie also gar nichts im angenommenen Fall (U, T unabhängig und identisch verteilt, Z unabhängig von U und T), die Gewinnwahrscheinlichkeit bleibt bei ½.

Dies passt aber natürlich nicht zu obigem Resultat von 3.1 und 3.2. Wo ist der Fehler in dieser auf den ersten Blick ja auch plausiblen Gegendarstellung?

Auf den zweiten Blick sieht man, dass hier Wahrscheinlichkeiten addiert wurden, die sich auf *verschiedene Bedingungen* beziehen[4] (nämlich ob U die größere oder die kleinere der beiden Zahlen ist; die obigen Wahrscheinlichkeiten sind eigentlich *bedingte Wahrscheinlichkeiten*):

$$\mathrm{P}(G) = \frac{1}{2} \cdot \mathrm{P}(Z > U \mid U < T) + \frac{1}{2} \cdot \mathrm{P}(Z \leq U \mid U > T)$$

Gilt hier $\mathrm{P}(Z > U \mid U < T) + \mathrm{P}(Z \leq U \mid U > T) = 1$?

Oberflächlich betrachtet könnte man meinen, ja, da wir die *Unabhängigkeit* der Zufallsvariable Z von U und T vorausgesetzt haben, also spielen doch die Bedingungen $U < T$ bzw. $U > T$ bzgl. Z gar keine Rolle. Aber ein tieferer dritter Blick zeigt mittels Umschreiben:

$$\mathrm{P}(Z > U \mid U < T) + \mathrm{P}(Z \leq U \mid U > T) =$$

$$\mathrm{P}(Z - U > 0 \mid U - T < 0) + \mathrm{P}(Z - U \leq 0 \mid U - T > 0)$$

Obwohl Z, U und T voneinander unabhängig sind, sind aber $Z - U$ und $U - T$ nicht voneinander unabhängig (da U in beiden Zufallsvariablen vorkommt), so dass man die Bedingungen hier nicht einfach weglassen kann; die angegebenen bedingten Wahrscheinlichkeiten addieren sich doch nicht einfach zu 1. Hier muss man in Sachen „(Un)-Abhängigkeit“ schon sehr genau hinschauen, um den Fehler zu entdecken.

Die „Z-Strategie“ lebt geradezu von der positiven Wahrscheinlichkeit für das Ereignis „Z liegt zwischen X und Y“ bzw. „zwischen a und b“ (die Dichte von Z ist ja auf ganz $\mathbb{R}$ strikt positiv). Am besten wäre also eine Zufallsvariable Z, deren Dichte zwischen X (bzw. a) und Y (bzw. b) möglichst große Werte hat, so dass diesem Bereich in Bezug auf Z eine möglichst große Wahrscheinlichkeit zukommt. Dies ist natürlich nur eine theoretische Aussage, denn wie soll man das im Einzelfall denn anstellen, wenn man nur eine der beiden Zahlen kennt? Aber es ist ja schon verblüffend, dass es *theoretisch* überhaupt möglich ist, den Wahrscheinlichkeitswert über ½ hinaus zu bringen, der „Ungewissheit also ein Schnippchen zu schlagen“.

Möglicher Einsatz dieses Themas im Schulunterricht

Natürlich können Schülerinnen und Schüler nicht selbständig auf diese Idee mit der Zufallsvariable Z kommen, die für die Entscheidung hilfreich ist. Dies ist also sicher keine stochastische Situation, in der Schülerinnen und Schüler viele eigene Ideen einbringen werden. In einem möglichen Schulunterricht wird die Lehrkraft diese Lösung vorzustellen haben. Aber nach der Vorgabe der

Zufallsvariablen X, Y, Z und evtl. einem Baumdiagramm wie jenen aus 3.1 oder 3.2 könnten Schülerinnen und Schüler durchaus aufgefordert werden, in selbständiger Arbeit zu bestätigen, dass die Gewinnwahrscheinlichkeit P(G) bei dieser Strategie größer als ½ ist. Dabei müssten Schülerinnen und Schüler obiges „Überlappungsargument" selbständig finden. Auch wenn dies durch die Lehrkraft geschieht: Das Nachvollziehen der Lösung (eines Paradoxons, auf einem elementaren Niveau wie hier) kann aber auch durchaus ein interessantes Thema sein: Einsehen a posteriori, *wie* und *warum* genau eine ganz neue Zufallsvariable Z bei der Entscheidung, ob die „andere" Zahl die größere oder die kleinere sein wird, zumindest theoretisch einen kleinen Beitrag dazu leisten kann, mit einer Wahrscheinlichkeit von mehr als ½ die größere von beiden zu erwischen. Das technisch nötige Rüstzeug aus der Stochastik ist relativ bescheiden: es besteht in dieser elementaren Darstellung eigentlich nur in *Zufallsvariable, bedingte Wahrscheinlichkeiten (Bäume), Dichtefunktion bei stetigen Zufallsvariablen* und (*Un-*)*Abhängigkeit*.

4 Die Darstellung des Zwei-Zettel-Spiels nach BRUSS 2000

In der Version des Zwei-Zettel-Spiels von Abschnitt 3 war Z eine Zufallszahl, die der Spieler als Resultat eines Zufallsexperiments erhalten hatte. Aber die Z-Strategie klappt auch dann, wenn Z vorab festgelegt wird. In der folgenden Darstellung (vgl. BRUSS 2000) ist z eine willkürlich festgelegte Zahl (z. B. $z = 0$, klein geschrieben wegen: keine Zufallsvariable, sondern feste Zahl), eine Grenze, bis zu welchem Wert der zuerst gesehenen Zahl U man diese verwerfen und lieber die andere T ($\neq U$) als die größere ansehen wird. Hier stellen wir uns vor, dass wir U noch nicht kennen, wir wollen a priori eine Zahl z für die „z-Strategie" festlegen.

Es ist nach dem Obigen zwar eigentlich klar, dass die z-Strategie auch im Fall der Vorab-Festlegung die Gewinnwahrscheinlichkeit über ½ hinaus hebt, denn der Zeitpunkt der Festlegung von Z war in obiger Diskussion ja irrelevant. Trotzdem sei hier eine weitere Begründungsvariante angeführt:

Man überlegt sich leicht, dass genau eine der drei folgenden Möglichkeiten eintreten muss:

A) z liegt „links" von beiden, genauer: $z \le U, T$

B) z liegt zwischen beiden,
genauer: $U < z \le T$ oder $T < z \le U$

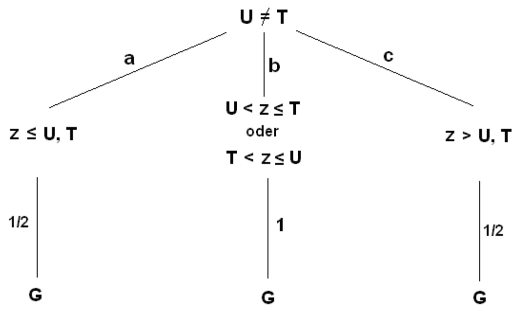
C) z liegt „rechts" von beiden, genauer: $z > U, T$

Im Fall **A)** führt unsere Strategie zu U, im Fall **C)** zu T. Welche Zahl (U oder T) in diesen beiden Fällen jeweils die größere ist, ist „rein zufällig", d. h. man kommt in diesen beiden Fällen mit Wahrscheinlichkeit ½ zur größeren Zahl. Genauer: Man glaubt an $\mathrm{P}(U < T) = \frac{1}{2} = \mathrm{P}(U > T)$; dies ist z. B. so, wenn U und T *unabhängige* und *identisch verteilte* Zufallsvariablen sind (klar aus Symmetriegründen). Dies soll auch unter zusätzlichen Bedingungen gelten, die für U und T gleichermaßen erfüllt sind, wie z. B. in den Fällen **A)** und **C)**:

$$\mathrm{P}(U < T \mid z \le U, T) = \frac{1}{2} = \mathrm{P}(U > T \mid z \le U, T) \text{ bzw.}$$

$$\mathrm{P}(U < T \mid z > U, T) = \frac{1}{2} = \mathrm{P}(U > T \mid z > U, T)$$

Im Fall **B)** erwischen wir mit unserer Strategie immer die größere Zahl (Gewinnsituation G), gleichgültig, ob in Wirklichkeit $U < z \le T$ oder $T < z \le U$ ist. Die Wahrscheinlichkeiten, mit denen die Fälle **A)**, **B)**, **C)** eintreten, seien mit a, b, c bezeichnet:



Daher ist $\mathrm{P}(G) = a \cdot \frac{1}{2} + b \cdot 1 + c \cdot \frac{1}{2}$ und wegen $a + b + c = 1$ ergibt sich $\mathrm{P}(G) = \frac{1}{2} + \frac{b}{2}$.

Wenn z. B. U und T *unabhängig* und *gleichverteilt* sind auf einem Intervall $[-R; R]$ (z. B. $[-10^6; 10^6]$) und man für die kritische Grenze dann sinnvoller Weise $z = 0$ wählt, so ist $b = \frac{1}{2}$ und somit $\mathrm{P}(G) = 3/4 = 75\%$ (vgl. die Computersimulation in Abschnitt 3.2, bei der sich auch eine Erfolgsquote von ¾ ergeben hat). Mit

$z = R/2$ ergibt sich $b = 2 \cdot \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{8}$ und somit immerhin noch $\mathrm{P}(G) = \frac{11}{16} = 68,75\%$. Allgemein erhält man mit der festen kritischen Grenze $g \in [-R; R]$: $b = 2 \cdot \frac{R+g}{2R} \cdot \frac{R-g}{2R} = \frac{1}{2} - \frac{g^2}{2R^2}$ und somit $\mathrm{P}(G) = \frac{3}{4} - \frac{g^2}{4R^2}$. Ein analoges Ergebnis ergäbe auch eine entsprechende Computersimulation (z. B. mit EXCEL).

Bemerkung: Obige Argumentation lässt sich auch auf den Fall übertragen, dass Z als Zufallsvariable mit nirgendwo verschwindender Dichte aufgefasst wird (vgl. 3.2). Man muss dabei aber eben zusätzlich voraussetzen, dass immer $\mathrm{P}(U < T) = 1/2 = \mathrm{P}(U > T)$ gilt (auch unter zusätzlichen Bedingungen – siehe oben). In 3.2 brauchten wir keine Voraussetzung der Art $\mathrm{P}(X < Y) = 1/2 = \mathrm{P}(Y < X)$.

Eine salopp formulierte Sichtweise (Erklärung), warum dieser Trick mit der Grenze z funktioniert: Bei $z \leq U$ hat U schon mal eine erste Probe (gegenüber z) erfolgreich bestanden, dies macht U zum Favoriten, vom anderen Wert weiß man nichts, also bleibe doch lieber bei U!

BRUSS (2000, S. 107) schreibt dazu: „Ihre Gewinnwahrscheinlichkeit ist um $b/2$ höher als bei der schlichten Zufallswahl, und b ist größer als null: Es kann ja vorkommen, dass Ihr z zwischen U und T fällt. Wie spielt man diese Strategie am geschicktesten? Offensichtlich so, dass Fall **B)** eine möglichst hohe Wahrscheinlichkeit b für sich hat. Also sollte man z so wählen, dass es mit höchster Wahrscheinlichkeit zwischen U und T liegt. Diese beiden Zahlen kennt man aber gerade nicht. Im Allgemeinen gibt es also keine eindeutige Empfehlung für die Wahl von z. Aber in konkreten Fällen findet man oft gute Anhaltspunkte." Mit dem letzten Satz sind vermutlich Situationen gemeint wie unten beschrieben (konkretes Beispiel: Hausverkauf).

Er berichtet auch von einem Experiment in seiner Vorlesung mit Studierenden: Diese sollten auf zwei Zettel je eine beliebige Zahl aufschreiben (ganz $\mathbb{R}$ erlaubt). Nach Bekanntgabe der ersten sollte er jeweils erraten, ob die andere doch die größere sei. „Bei 41 oder 42 Teilnehmern konnte ich 32 Erfolge verbuchen! Bei einer Zufallsstrategie wären 20 oder 21 zu erwarten, mit etwas Glück auch 3 oder 4 mehr. Aber 32 kann man nicht mehr mit der Begründung Glück abtun. Selbst die besten Studenten standen vor einem Rätsel. Es ist schwer, etwas zu sehen, wo man nichts erwartet. Aber Sie, liebe Leser, können es erraten. Ich wandte eine z-Strategie an, sogar eine besonders naive. Ich wählte $z = 0$." (BRUSS 2000, S. 107).

Ein konkretes Beispiel:

Sie wollen Ihr Haus verkaufen und schalten ein Inserat in den Zeitungen: „Verkaufe für das beste Angebot über € 400.000,–". Zwei Kaufinteressenten haben nun definitiv Interesse angemeldet und sind bereit mehr als € 400.000,– zu zahlen, nur wie viel mehr, haben sie noch nicht gesagt. Herr X kommt nächsten Samstag, will sich das Haus noch einmal ansehen und sein endgültiges Angebot abgeben. Er will dann aber auch die Entscheidung von Ihnen wissen und ist nicht bereit noch eine ganze Woche zu warten; Frau Y kann nämlich erst den darauf folgenden Samstag kommen, um ihr endgültiges Angebot abzugeben. Damit ergibt sich das Dilemma: Wenn Sie Herrn X zusagen, geht Ihnen das Angebot von Frau Y verloren und umgekehrt. Dies scheint ein Glücksspiel zu sein! Sie verpassen mit Wahrscheinlichkeit ½ das bessere der beiden Angebote – oder?

Die *Struktur*analogie zum Zwei-Zettel-Spiel ist offensichtlich: Der zuerst genannte Kaufpreis von Herrn X spielt die Rolle der Zahl im gewählten Umschlag, der noch unbekannte Kaufpreis von Frau Y entspricht der unbekannten Zahl (Unterschied: hier geht es um Preise > 400.000, beim 2-Zettel-Spiel auch negative Zahlen möglich).

Hier wäre es günstig, den Marktwert M des Hauses möglichst genau einschätzen zu können. Mit $z = M$ ist die Wahrscheinlichkeit, dass z zwischen den beiden kommenden Angeboten liegen wird, möglichst groß (die kommenden Angebote werden vermutlich ja um den Marktwert M streuen). Bei $z > X$ warte auf den zweiten Samstag und verkaufe das Haus an Frau Y; bei $z \leq X$ verkaufe das Haus Herrn X und sage den Termin mit Frau Y ab. Wo soll man allerdings die „Schmerzgrenze z" hier ansetzen? Das Minimum € 400.000,– ist schon festgelegt, es ist vermutlich auch sehr unwahrscheinlich, dass jemand über € 500.000,– bezahlt[5], d. h. man kann z schon einschränken auf den Bereich zwischen 400.000 und 500.000. Wenn man die „kritische Grenze“ bei € 401.000,– ansetzt (und dafür Herr *X* den Zuschlag erhält) und man dann nachher erfährt, dass Frau Y € 402.000,– geboten hätte, so ist dies kein Grund zur Verzweiflung, das ist keine Riesenchance, die einem da entgangen ist. Wenn allerdings Frau *Y* € 430.000,– bezahlt hätte, ist dies schon schmerzlicher. Man sollte also diese „Schmerzgrenze“ schon etwas höher ansetzen, aber auch nicht zu hoch, vielleicht bei € 415.000,–, wobei es hier sicher auch noch Alternativen gäbe: Z. B. zu würfeln und die (dann doch zufällige) Grenze Z bei $400.000 + W \cdot 5000$ anzusetzen, wobei $W \in \{1,\ldots,6\}$ das Würfelergebnis ist. Die obige Zahl € 415.000,– ergäbe sich dann bei einer gewürfelten „3“.

5 Verwandte Probleme

5.1 Das Paradoxon der zwei Umschläge

(siehe z. B. BRUSS 1996, RIEHL 2004, HÄGGSTRÖM 2006 S. 21ff; auch CHRISTENSEN&UTTS 1992, MEESTER 2003 S. 57f)

Ein Gönner stellt zwei Schecks mit Beträgen aus, wobei einer doppelt so groß wie der andere ist. Er steckt diese in je einen Umschlag und lässt einen Spieler einen der beiden Umschläge auswählen. Während es beim Zwei-Zettel-Spiel also um *irgendwelche* Zahlen geht, geht es hier um **positive** Zahlen (Geldbeträge), die sich wie **1 : 2** verhalten. Hier bekommt der Spieler den gewählten Geldbetrag als „Preis“, während beim Zwei-Zettel-Spiel das Erwischen der größeren Zahl über das Bekommen bzw. Nichtbekommen eines anderen Preises entschieden hat. Es gibt zwei verschiedene Versionen des Paradoxons.

5.1.1 Der Spieler darf den ersten Umschlag *nicht öffnen*

Der Spieler darf den gewählten Umschlag nicht öffnen, er kann aber nach dem Ziehen des Umschlages jetzt noch einmal überlegen: Entweder er bleibt bei diesem Umschlag, oder er nimmt den anderen. Was soll er tun? Ist hier nicht Tauschen immer die bessere Strategie, wie man durch Berechnen des Erwartungswertes für den eingetauschten Betrag zeigen kann?

Sei U der (unbekannte) Betrag im gewählten Umschlag (diesen Betrag bekommt man, wenn man nicht tauscht). Dann muss auf dem anderen Scheck entweder $U/2$ oder $2U$ stehen, jeweils mit Wahrscheinlichkeit ½. Als Erwartungswert für den Betrag T im eingetauschten Umschlag ergibt sich dann offenbar: $\mathrm{E}(T) = \frac{1}{2}\cdot 2U + \frac{1}{2}\cdot\frac{U}{2} = \frac{5}{4}U > U$.
Also lohnt sich Tauschen doch in jedem Fall, oder?

Dass dies nicht stimmen kann, ist klar, denn aus demselben Argument müsste sich dann ein erneuter Tausch lohnen (nochmalige Erhöhung des Erwartungswertes), was aber nicht sein kann, da man ja dann den ursprünglichen Umschlag in der Hand hält.

Aber *warum* stimmt obige Argumentation nicht? Worin liegt genau der Fehler?
Er liegt in der Erwartungswertbildung

$$\mathrm{E}(T) = \frac{1}{2}\cdot 2U + \frac{1}{2}\cdot\frac{U}{2}.$$

Ein erster Hinweis ist durch den verdächtigen Faktor 4 zwischen $2U$ und $U/2$ gegeben. Aufgrund der beschriebenen „Versuchsbedingungen“ ist ja klar, dass sich die Ausprägungen von T (und auch jene von U) nur um den Faktor 2 unterscheiden dürfen. Die „Ausprägungen“ von T darf man nicht mit $2U$ und $U/2$ bezeichnen, denn U hat darin eine *unterschiedliche Bedeutung*: in $2U$ ist U der kleinere der beiden möglichen Beträge, in $U/2$ ist U der größere[6].

Man kann dieses Phänomen auch mit bedingten Wahrscheinlichkeiten analysieren (vgl. z. B. HÄGGSTRÖM 2006 S. 21ff): Die beiden Zahlen des Gönners seien mit k bzw. mit $2k$ bezeichnet[7].
Dann ist:

$$\mathrm{P}(U = k\,, T = 2k) =$$

$$\underbrace{\mathrm{P}(U = k \mid T = 2k)}_{1}\cdot\underbrace{\mathrm{P}(T = 2k)}_{1/2} = \frac{1}{2}$$

und analog

$$\mathrm{P}(U = 2k, T = k) =$$

$$\underbrace{\mathrm{P}(U = 2k \mid T = k)}_{1}\cdot\underbrace{\mathrm{P}(T = k)}_{1/2} = \frac{1}{2}$$

D. h. insgesamt ergibt sich dadurch

$$\mathrm{P}(T = 2U) = 1/2 = \mathrm{P}(T = U/2)\ , \qquad (1)$$

also kann Tauschen nichts bringen.

Eine etwas andere Klärung dieser Version des Paradoxons kann im weitesten Sinn angelehnt an BRUSS 1996 gegeben werden. x sei der erste Betrag, den der Gönner auf einen Scheck schreibt und in einen Umschlag steckt. Dann wirft er eine faire Münze, um zu entscheiden, ob sein zweiter Betrag Y doppelt oder halb so groß sein soll: $\mathrm{P}(Y=2x)=1/2=\mathrm{P}(Y=x/2)$. Dann gilt offenbar $\mathrm{E}(Y)=\frac{1}{2}\cdot 2x+\frac{1}{2}\cdot\frac{x}{2}=\frac{5}{4}x$ und damit: $\mathrm{E}(U)=\frac{1}{2}\cdot x+\frac{1}{2}\cdot \mathrm{E}(Y)=\frac{9}{8}x$. Dasselbe ergibt sich auch für $\mathrm{E}(T)$, so dass klar ist: Tauschen bringt nichts! Dies kann man auch noch ein wenig anders aufschreiben: Nach Werfen der Münze liegt ein Zahlenpaar vor, entweder $(x,2x)$ oder $(x,x/2)$ – jeweils mit Wahrscheinlichkeit ½. Der (bedingte) Erwartungswert ist im einen Fall $\mathrm{E}(U\,|\,(x,2x))=\frac{1}{2}\cdot x+\frac{1}{2}\cdot 2x=\frac{3}{2}x$ und im anderen $\mathrm{E}(U\,|\,(x,x/2))=\frac{1}{2}\cdot x+\frac{1}{2}\cdot x/2=\frac{3}{4}x$. Für den Gesamterwartungswert ergibt sich dann $\mathrm{E}(U)=\frac{1}{2}\cdot\frac{3}{2}x+\frac{1}{2}\cdot\frac{3}{4}x=\frac{9}{8}x$; analog für $\mathrm{E}(T)$.

5.1.2 Der Spieler *öffnet* den gewählten Umschlag

Der Spieler sieht den auf den Scheck geschriebenen Betrag u. Wenn T wieder die Zufallsvariable ist, die den Betrag des anderen (getauschten) Umschlages angibt, so erhalten wir für deren Erwartungswert $\mathrm{E}(T)=\frac{1}{2}\cdot 2u+\frac{1}{2}\cdot\frac{u}{2}=\frac{5}{4}u>u$. Analog ist auch hier klar, *dass* dies nicht stimmen kann, aber *warum* stimmt dies nicht, worin liegt genau der Fehler?
Wieder gibt es den verdächtigen Faktor 4 zwischen den Ausprägungen von T: $2u$ und $u/2$. Auch diese Erwartungswertbildung ist falsch.

Der mit u bezeichnete Inhalt des gewählten Umschlages wird hier wie eine Konstante behandelt, er ist aber eigentlich eine Zufallsvariable. Dadurch kommt es zu dem Effekt mit Faktor 4 statt Faktor 2 (vgl. RIEHL 2004):

Seien k der kleinere und $2k$ der größere der beiden in Rede stehenden Beträge[8]. Diese Beträge werden in zwei verschiedene Umschläge gesteckt. Die Zufallsvariable U bezeichne den Betrag im gewählten Umschlag, T bezeichne den Betrag im anderen Umschlag. Die beiden folgenden Fälle mit jeweils Wahrscheinlichkeit ½ sind möglich:

$$1)\, U=k \;\Rightarrow\; T=2k=2U$$
$$2)\, U=2k \;\Rightarrow\; T=k=U/2$$

Die beiden möglichen Ausprägungen von T unterscheiden sich also tatsächlich nur um den Faktor 2 ($2k$ und k), nur bezogen auf die Zufallsvariable U gibt es den Faktor 4, denn U ist im einen Fall die kleinere der beiden Zahlen, im anderen die größere. Oben hatte aber die vermeintliche Konstante u die Rolle der Zufallsvariable U, so dass es zu einem falschen Ergebnis kam.

Es gilt jedenfalls in beiden Fällen (ob man nun den Umschlag öffnet oder nicht):

$$\mathrm{E}(U)=\frac{1}{2}\cdot k+\frac{1}{2}\cdot 2k=1{,}5\cdot k \text{ und analog}$$
$$\mathrm{E}(T)=\frac{1}{2}\cdot k+\frac{1}{2}\cdot 2k=1{,}5\cdot k\,.$$

Da U und T denselben Erwartungswert haben, kann Tauschen also nichts bringen.

Wieder kann man auch mit bedingten Wahrscheinlichkeiten analysieren: Wenn man den Umschlag öffnet, so kennt man die Realisierung von U, nämlich $U=u$. Insbesondere folgt aus (1) aber **nicht allgemein**: $\mathrm{P}(T=u/2)=1/2=\mathrm{P}(T=2u)$, wie es bei der Schilderung des Paradoxons eingangs suggestiv dargestellt wird.

Um die Chancen dafür, dass ein Tausch zu einer Verbesserung führt, zu bestimmen, müsste man die *bedingte* Wahrscheinlichkeit $\mathrm{P}(T=2u\,|\,U=u)$ berechnen:
K bezeichne die kleinere der beiden Zahlen des Gönners und $p_k:=\mathrm{P}(K=k)$. Dann ist $\mathrm{P}(U=u)=p_u\cdot\frac{1}{2}+p_{u/2}\cdot\frac{1}{2}$ und es ergibt sich:

$$\mathrm{P}(T=2u\,|\,U=u)=\frac{\mathrm{P}(T=2u,U=u)}{\mathrm{P}(U=u)}=\frac{p_u\cdot\frac{1}{2}}{p_u\cdot\frac{1}{2}+p_{u/2}\cdot\frac{1}{2}}=\frac{p_u}{p_u+p_{u/2}}$$

Zur Berechnung solcher bedingten Wahrscheinlichkeiten braucht man also die Wahrscheinlichkeiten p_u und $p_{u/2}$, mit der der Gönner seine Beträge wählt. Aber diese kennt man i. A. nicht, so dass man solche Überlegungen im konkreten Fall $U=u$ gar nicht anstellen kann. Damit – losgelöst von konkreten Werten für u bzw. p_u und $p_{u/2}$ – *allgemein* $\mathrm{P}(T=2u\,|\,U=u)=1/2$ wäre, müsste also *allgemein* $p_k = p_{k/2}$ für alle möglichen kleineren Geldbeträge k des Gönners sein. Bei einer *endlichen* Menge ist dies a priori nicht möglich (für den kleinsten Wert $k_{\min}$ müsste gelten $p_{k_{\min}} = p_{k_{\min}/2}$; Widerspruch, dass $k_{\min}$ minimal ist; analog am „oberen Ende" bei $k_{\max}$). Bei einer *unendlichen* Trägermenge des kleineren Gönnerbetrages K (wegen $p_k = p_{k/2}$ für alle möglichen k käme vordergründig z. B. die Menge $\{k = 2^i, i \in \mathbb{Z}\}$ in Frage) kann es aber bekanntlich keine *Gleichverteilung* geben (es müsste aber eine Gleichverteilung sein, wenn allgemein $p_k = p_{k/2}$ gelten sollte). Eine nach oben unbeschränkte Trägermenge von K wäre natürlich auch inhaltlich absurd: der Gönner müsste über beliebig viel Geld verfügen.

Wenn man Informationen darüber hat, wie der Gönner seine Zahlen wählt (z. B. über die Verteilung der Zufallsvariable K – die kleinere der beiden Zahlen), dann kann es natürlich sein, dass man eine bessere Strategie findet als „tausche immer" oder „tausche nie", z. B. „tausche, wenn u einen bestimmten Mindestwert nicht erreicht".

Hierzu möge ein einfaches Beispiel genügen. Wir nehmen an, dass K nur die Werte 1000 (Beträge 1000 und 2000) und 500 (Beträge 500 und 1000) annehmen kann mit $p_{1000} =: p$ und $p_{500} = 1 - p$. Dann ergibt sich für den Erwartungswert des Gewinns bei „tausche nie" und bei „tausche immer":

$$(1000 \cdot \frac{1}{2} + 2000 \cdot \frac{1}{2}) \cdot p + (500 \cdot \frac{1}{2} + 1000 \cdot \frac{1}{2}) \cdot (1-p)$$
$$= 750 + 750p$$

Man rechnet leicht aus, dass die tauschfreudige Strategie „tausche immer bei $u < 2000$" eine Gewinnerwartung von $750 + 1250p$ hat, d. h. sie lohnt sich gegenüber „tausche nie (ímmer)" solange $p > 0$ ist, solange also die höheren Beträge (1000 , 2000) überhaupt in Frage kommen. Die tauschfeindliche Strategie „tausche nur bei $u = 500$" hat eine Gewinnerwartung von $1000 + 500p$. Diese ist den obigen beiden („tausche nie" und „tausche immer") überlegen solange $p < 1$, d. h. solange die niedrigeren Beträge (500 , 1000) überhaupt in Frage kommen; sie ist der tauschfreudigen überlegen, solange $p < 1/3$ ist.

5.1.3 Eine Situation, in der obige Argumentation richtig ist

Wie müsste die Situation beschaffen sein, damit die obige, das Paradoxon einleitende Argumentation richtig wäre? So eine Beschreibung kann man leicht angeben. Dem Spieler wird z. B. gesagt: „Sie halten einen Umschlag mit einem gewissen Betrag in der Hand[9]; der Inhalt des *anderen* Umschlages ist so zustande gekommen: es wurde eine faire Münze geworfen, die entschieden hat, ob der doppelte oder der halbe Betrag des Ihren dort hinein kommt. Wollen Sie tauschen?"

Dann lässt sich der Unterschied zwischen den beiden möglichen Ausprägungen der in Rede stehenden Zufallsvariable wirklich mit dem Faktor 4 beziffern und somit ist obige Argumentation korrekt. Man sollte in dieser Situation jedenfalls tauschen, wenn man den Erwartungswert des Gewinnes maximieren will.

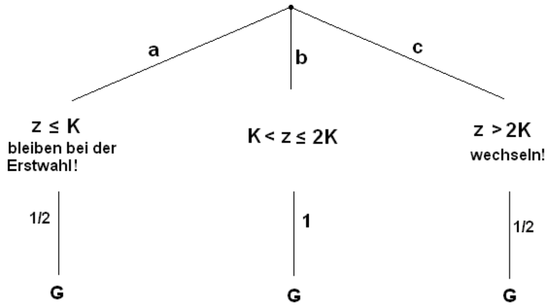
5.1.4 *Z*-Strategie beim Umschlagparadoxon

Mit der *Z*-Strategie kann man auch beim Umschlagparadoxon punkten. Dort wird sie üblicherweise nur nicht erwähnt, es geht dabei eben nur um die Frage: Kann *Tauschen* was bringen? – siehe oben.

- Die zwei positiven Zahlen k und $2k$ werden als fest angesehen. Die Erkenntnisse in 3.1 waren ja ganz unabhängig vom Verhältnis der beteiligten Zahlen. Hier beim Umtauschparadoxon liegt ein Spezialfall vor: Verhältnis 2 : 1. Da es sich hier um *positive* Zahlen (Geldbeträge) handelt, braucht man die Dichte nur auf $\mathbb{R}^+$ zu betrachten (es genügt hier, dass sie auf $\mathbb{R}^+$ nicht verschwindet – z. B. Exponentialverteilung). Ganz analog zu 3.1 käme man mit $a = k$ und $b = 2k$ zu der Erkenntnis: Mit der *Z*-Strategie ist $\mathrm{P}(G) > 1/2$, und zwar: $\mathrm{P}(G) = 1/2 + \mathrm{P}(k < Z \le 2k)/2$.

- In 3.2 waren die Erkenntnisse ebenso unabhängig vom konkreten Verhältnis der beteiligten Zufallsvariablen. D. h. auch wenn man die beiden (jetzt positiven) Zahlen des Gönners als Zufallsvariablen X und Y ansieht, ergibt sich mit der *Z*-Strategie $\mathrm{P}(G) > 1/2$.

- Auch wenn man z als vorab festgelegte Schmerzgrenze ansieht (vgl. Abschnitt 4), funktioniert die z-Strategie: K bzw. $2K$ seien hier die Zufallsvariablen der beiden in Rede stehenden Beträge:



Im Fall A) bei $z \le K$ bedeutet die *z*-Strategie, bei der Erstwahl zu bleiben: Je nachdem, ob man schon ursprünglich die größere Zahl erwischt hat (p = ½), wird dies *Gewinn* (d. h. größere Zahl) oder *nicht Gewinn* bedeuten.

Mit konkreten Annahmen über die Art und Weise, wie die beiden Geldbeträge des Gönners zustande kommen, und z. B. mit der Exponentialverteilung könnte man wieder ähnliche Computersimulationen wie oben machen.

5.2 Das „Sekretärinnenproblem"

Eine Firma hat eine Sekretärinnenstelle zu besetzen und möchte die beste unter den *n* Bewerberinnen einstellen (vgl. z. B. BRUSS 2004). Diese stehen alle vor der Tür und werden der Reihe nach zu einem Bewerbungsgespräch gebeten[10]. Anders als im wirklichen „Bewerbungsleben" soll hier aber die Regel sein: Es muss jeder Bewerberin am Ende des jeweiligen Gesprächs definitiv gesagt werden, ob sie den Job bekommt oder nicht (bei „ja" ohne die nachfolgenden Bewerberinnen anzuhören; bei „nein" steht diese Person nicht mehr zur Verfügung). Der Personalchef soll also dann „ja" zu einer Person sagen, wenn sie besser war als die vorangegangenen und keine Bessere mehr nachkommt. Aber wie soll er das wissen?

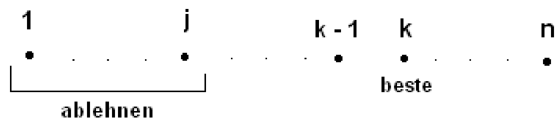
Eine mögliche Strategie ist die folgende: *n* Sekretärinnen bewerben sich. Der Personalchef lehnt die ersten *j* Sekretärinnen ab (in der Hoffnung, dass später noch bessere kommen werden) und nimmt dann die nächste, die besser ist als jede der ersten *j*. Das Problem dabei ist: Wie soll *j* gewählt werden, so dass die Wahrscheinlichkeit möglichst groß ist, wirklich die beste Sekretärin unter den *n* zu erwischen?

Der Situationszusammenhang zum obigen Zwei-Zettel-Spiel ist evident: Der Spieler im Zwei-Zettel-Spiel muss sich für einen der beiden Umschläge entscheiden (ohne den Inhalt des anderen zu kennen), der Personalchef muss sich für eine Sekretärin entscheiden, ohne die „nachfolgenden" zu kennen. So stehen gewissermaßen beide vor demselben Dilemma. Der obige Spieler fragt sich: Ist im anderen Umschlag nicht doch die höhere Zahl? Der Personalchef fragt sich: Ist unter den restlichen Sekretärinnen nicht doch noch eine bessere? Aus der Sichtweise des Zwei-Zettel-Spiels könnte man das Sekretärinnenproblem auch als *n*-Zettel-Spiel bezeichnen.

Der Fokus bei der „Lösung" ist in beiden Fällen natürlich ein ganz anderer: Während man beim Zwei-Zettel-Spiel fragt, ob Tauschen (Kriterium: Zufallsvariable *Z*) zu einer Erfolgswahrscheinlichkeit von mehr als ½ führt, sucht man beim Sekretärinnenproblem nach einer „optimalen" „Ablehnungszahl *j*". Obwohl vielen sicher bekannt und auch an vielen anderen Stellen nachzulesen hier eine

Lösung in Kurzform: Wir nehmen zunächst einmal an, dass die beste der *n* Sekretärinnen in der Schlange an *k*-ter Stelle steht. Man erwischt sie mit obiger Strategie genau dann, wenn sich die beste der ersten $k-1$ (d. h. die zweitbeste der ersten *k*) bereits in den ersten abgelehnten *j* befindet: Wenn die beste der ersten $k-1$ erst nach *j* käme, würde man ja nach obiger Strategie diese insgesamt zweitbeste und nicht die beste nehmen!

Bei jeder zufälligen Anordnung von unbekannten Sekretärinnen geht man davon aus, dass keine Stelle bevorzugt ist bzgl. „beste Sekretärin" – es gibt keinen Grund, warum sich die beste eher am Anfang, eher in der Mitte oder eher am Ende befinden sollte.



Daher ist die Wahrscheinlichkeit, dass an der Stelle *k* die beste steht, gleich $1/n$. Analog ist die Wahrscheinlichkeit, dass die beste der ersten $k-1$ Sekretärinnen sich unter den ersten *j* befindet, gleich $\frac{j}{k-1}$ (Laplace). Daraus ergibt sich die Wahrscheinlichkeit P für einen Erfolg der obigen Strategie: $\mathrm{P} = \sum_{k=j+1}^{n} \frac{1}{n} \cdot \frac{j}{k-1} = \frac{j}{n} \cdot \sum_{k=j}^{n-1} \frac{1}{k}$. Hier könnte man mit einem CAS nun einfach probieren, für welches j der Wert von P maximal wird (diskretes Problem).

Um analytisch weiterarbeiten zu können, kann man diese Summe auch mit der Integral-Näherung approximieren: $P \approx \frac{j}{n} \cdot \int_j^n \frac{1}{k} \mathrm{d}k = \frac{j}{n} \ln \frac{n}{j}$. Nullsetzen der Ableitung nach j ergibt: $j_{\text{opt}} = n / e$ (das Maximum ist ziemlich flach, nicht sehr ausgeprägt). Der Wert an dieser Stelle ist $P_{\text{opt}} \approx 1/e$.

Man sollte nach diesem Prinzip etwas mehr als das erste Drittel (ca. 36,8%) ablehnen, und dann die nächste Sekretärin nehmen, die besser ist als jede der abgelehnten.

Bemerkung: In der Praxis hält sich die Bedeutung dieses Prinzips natürlich in engen Grenzen. Bei Bewerbungen muss man eben nicht am Ende des jeweiligen Gesprächs definitiv entscheiden, Personalchefs schauen sich alle Bewerber/innen an und entscheiden erst dann. Bei anderen Situationen kennt man i. A. die Anzahl n nicht, z. B. beim Verkauf einer Immobilie: Jemand möchte im nächsten Jahr sein Haus verkaufen; er/sie muss natürlich bei den ersten Angeboten reagieren (ja oder nein sagen) ohne zu wissen, welche weiteren Angebote in den nächsten Monaten noch kommen werden. Dies ist zwar im Prinzip die Situation beim Sekretärinnenproblem, aber man kennt ja die Zahl n der insgesamt eintrudelnden Angebote nicht. In der Literatur findet man auch das Sekretärinnen-problem gelegentlich im Zusammenhang mit der Wahl der Lebenspartnerin / des Lebenspartners: Soll man sich für diese Person entscheiden? Aber auch hier kennt man ja die Zahl der zukünftigen potentiellen Partner/innen leider (oder sollte man vielleicht besser sagen: Gott sei Dank!) nicht.

Danksagung: Ich danke den Kollegen Prof. Dr. Peter RAITH (Wien) und Dr. Jörg MEYER (Hameln) für viele wertvolle Hinweise.

Literatur

BRUSS, F. T. (1996): The Fallacy Of The Two Envelopes Problem. *The Mathematical Scientist* 21, 112 – 119.

BRUSS, F. T. (2000): Der Ungewissheit ein Schnippchen schlagen. *Spektrum der Wissenschaft* 06/2000, 106 – 108.

BRUSS, F. T. (2004): Strategien der besten Wahl. *Spektrum der Wissenschaft*, Mai 2004, 102 – 104.

CHRISTENSEN, R u. J. UTTS (1992): Bayesian Resolution of the „Exchange Paradox". *The American Statistician* 46 (4), 274 – 276.

ENGEL, J. (1996): Das Achensee-Paradoxon. *Stochastik in der Schule* 16 (1), 3 – 10.

HÄGGSTRÖM, O. (2006) : *Streifzüge durch die Wahrscheinlichkeitstheorie*. Berlin: Springer.

MEESTER, R. (2003): *A natural introduction to probability theory*. Basel: Birkhäuser.

RIEHL, G (2004): Wer tauscht gewinnt nicht. *Stochastik in der Schule* 24 (3), 45 – 51.

Anschrift des Verfassers

Hans HUMENBERGER

Universität Wien, Fakultät für Mathematik, Nordbergstraße 15 (UZA 4), A – 1090 Wien.
Mail: hans.humenberger@univie.ac.at

Anmerkungen:

[1] Eine andere Darstellung mittels Mehrfachintegralen findet man z. B. bei ENGEL 1996. In dieser Darstellung wollten wir hingegen möglichst elementar bleiben.

[2] Z. B. Sieger eines Preisausschreibens.

[3] Die Dichtefunktion selbst ist hier nicht dargestellt, nur die zugehörigen Argumentbereiche; auf diese Dichtefunktion kommt es aber auch gar nicht an.

[4] Ein ähnlich sorgloser Umgang (Nichtbeachtung, ob U die größere oder die kleinere der beiden Zahlen ist) wird uns im Prinzip auch wieder in Abschnitt 5 („Paradoxon der zwei Umschläge") begegnen – vgl. Anmerkung 6.

[5] Wenn ein Verkäufer selbst das Minimum bei € 400.000,– festlegt, so wird der Marktwert i. A. nicht weit davon entfernt sein, so dass ein Kaufpreis von mehr als € 500.000,– kaum zu erwarten ist.

[6] Auch hier wird also nicht unterschieden, ob U der kleinere oder der größere der beiden Beträge ist – vgl. Abschnitt 3.3 und Anmerkung 4.

[7] Nach welchen Kriterien bzw. Überlegungen diese beiden Beträge aus der Hand des Gönners gewählt wurden, spielt hier keine Rolle. Wir könnten hier auch an Zufallsvariablen K bzw. $2K$ denken statt an konkrete Zahlen. k bzw. K sollen jedenfalls an „**k**leinere Zahl" erinnern.

[8] Auch hier ist es gleichgültig, ob man an feste Zahlen oder an Zufallsvariablen K bzw. $2K$ denkt.

[9] Gleichgültig, ob der Spieler hineinschaut und die Zahl sieht oder nicht.

[10] Wir gehen dabei davon aus, dass nie zwei Sekretärinnen „gleich gut" sein können und dass der Personalchef dieser Firma nach kurzen Gesprächen immer eine eindeutige Rangfolge bzgl. der Eignung der befragten Personen aufstellen kann.