

Geometrie und Münzwurf: das Modell zufälliger Graphen

PETER EICHELSBACHER, BOCHUM

Zusammenfassung: In diesem Artikel wird das Modell zufälliger Graphen betrachtet. Es entspricht einem Münzwurfmodell und liefert interessante sprunghaft auftretende Phänomene, die man Phasenübergänge nennt.

1 Einleitung

In diesem Aufsatz wollen wir das Modell eines Zufallsgraphen einführen. Das Modell kann für den Schulunterricht zum Beispiel durch die beiden folgenden Modellierungen motiviert werden:

1. Auf einem Stück Papier sei eine gewisse Anzahl von Punkten (Knoten) vorgegeben, sagen wir allgemein n Punkte. Je zwei dieser Punkte können mit einer gewissen vorgegebenen Wahrscheinlichkeit p miteinander durch ihre Verbindungsstrecke verbunden werden. Es entsteht ein zufälliges Gebilde von Kanten zwischen den n Knoten. Stellen wir uns zum Beispiel vor, dass die n Punkte sehr regelmäßig angeordnet sind und führen wir die Einschränkung ein, dass nur räumlich benachbarte Punkte durch eine Kante miteinander verbunden werden können. Dann kann das Modell dazu dienen, einen *Perkolationsvorgang* zu erklären: das Auftreten einer Kante entspricht dem Durchdringen einer Flüssigkeit innerhalb eines porösen Materials. Die zufällig erzeugte Figur entspricht den Wegen, die sich die Flüssigkeit ebnet. Es ergeben sich natürliche Fragen wie zum Beispiel die der Berechnung der Wahrscheinlichkeit, dass das vorgegebene Punktesystem von der Flüssigkeit durchdrungen wird, genauer: bei Eintritt in das Material an einem anderen Punkt das Material wieder verläßt.

2. Eine zweite Interpretation verzichtet auf den physikalischen Bezug. Man fragt ganz naiv nach der Wahrscheinlichkeit des Auftretens gewisser vertrauter Figuren der Elementargeometrie. Wie groß ist etwa die Wahrscheinlichkeit, dass ein Kreis der Länge 6 entsteht? Hierbei ist ein Kreis der Länge 6 ein System von 6 Punkten, das durch sechs Kanten kreisförmig verbunden ist. Mit welcher Wahrscheinlichkeit treten Dreiecke auf?

Die beiden beschriebenen Modelle tauchen üblicherweise in der Schulstochastik nicht auf. Dieser Aufsatz soll zeigen, dass die Modellierung einfach mit

tels eines *Münzwurfs* geschieht, also im Laufe eines Stochastikkurses ohnehin zur Verfügung steht. Wir wollen zeigen, dass im Falle der Wahl einer Kantenwahrscheinlichkeit p , die von der Knotenzahl n abhängt (wir schreiben dann $p(n)$), sehr interessante und überraschende Phänomene auftreten, die bei der Untersuchung des gewöhnlichen Münzwurfs gar nicht zu beobachten sind. Die Wahl einer Wahrscheinlichkeit $p(n)$ ist dabei ganz und gar nicht künstlich. Der Stochastiker denkt vermutlich direkt an den Grenzwertsatz von Poisson, bei dem die Wahrscheinlichkeit seltener Ereignisse studiert wird. In diesem Satz ist $\lim_{n \rightarrow \infty} p(n) = 0$ die Voraussetzung. In Bezug auf die Interpretation, ein poröses Medium auf seine Durchdringlichkeit hin zu untersuchen, ist n zunächst die Anzahl von möglichen Positionen der Kantenendpunkte in dem Medium, zum Beispiel in einem Stein. Natürlich gibt es unendlich viele, ja sogar überabzählbar viele Knotenpunkte. Bei dem Versuch, ein Modell zu betrachten, das immer genauer die Realität widerspiegelt, muß also die Zahl n der möglichen Knotenpunkte wachsen. Dann wird $p(n)$ klein gewählt. Dies entspricht der Vorstellung, in einem immer feineren Netz von möglichen Wegen den einzelnen Weg mit einer geringeren Wahrscheinlichkeit zu durchschreiten. Interessant wird sein, dass die Geschwindigkeit, in der $p(n)$ gegen null strebt, das Verhalten des Modells maßgeblich beeinflussen wird. Weiterhin wird sich zeigen, dass bei wachsendem n das Modell im Falle einer konstanten Kantenwahrscheinlichkeit p keine interessanten Resultate liefern kann.

Unser Modell ist die bekannte *Bernoulli-Kette*, die im Schulunterricht sowie in elementaren Einführungsvorlesungen an der Universität im Zentrum der Betrachtungen steht. Wir führen kurz ein wenig Notation ein.

2 Die Bernoulli-Kette

Gegeben sei $\Omega = \{E, M\}$, wobei E für Erfolg und M für Mißerfolg steht. Die Wahrscheinlichkeit für Erfolg sei $0 \leq p \leq 1$. Dann setze $\Omega_n = \{E, M\}^n$, die Menge der E - M -Folgen der Länge n . Die Wahrscheinlichkeit für jedes $\omega = (\omega_1, \dots, \omega_n) \in \Omega_n$ sei gegeben durch $p(\omega) = p^k(1-p)^{n-k}$, wobei k die Anzahl der E 's in der Folge $\omega_1, \dots, \omega_n$ bezeichnet.

X sei die Anzahl der Erfolge in der Bernoulli-Kette der Länge n zum Parameter p . Setze $X_i = 1$, falls $\omega_i = 1$ und $X_i = 0$ sonst, $1 \leq i \leq n$, so folgt $X = X_1 + X_2 + \dots + X_n$. Es gilt

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} =: b(k; n, p)$$

für $k \in \{0, \dots, n\}$. Die Zufallsgröße mit obiger Verteilung heißt binomialverteilt mit Parametern n und p . Es gilt $\mathbb{E}(X) = np$ und $V(X) = np(1-p)$, wobei $\mathbb{E}(X)$ den Erwartungswert von X bezeichnet und $V(X)$ die Varianz von X .

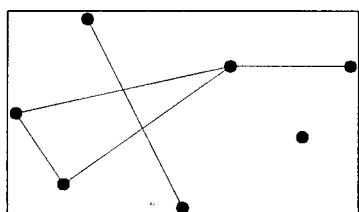
Die Tschebyschev-Ungleichung besagt ganz allgemein: Ist X eine Zufallsgröße, deren Erwartungswert $\mathbb{E}(X)$ und Varianz $V(X)$ existieren, so gilt für jedes $\varepsilon > 0$

$$P(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}.$$

Um nun das in der Einleitung angedeutete Modell genau beschreiben zu können, betrachten wir einen wichtigen Grundbegriff der Mathematik, der unbedeutenderweise in der Schulmathematik kaum betrachtet wird. Es ist der Begriff eines *Graphen*.

3 Elementare Begriffe der Graphentheorie

Ein Graph ist eine Ansammlung von *Punkten* (*Knoten*), wobei manche Punkte durch eine *Kante* verbunden sind. Etwa



Wir deuten hier die vielseitigen Anwendungen eines *Graphen* nur an: die Knoten können Orte, die Kanten Schienenverbindungen sein; der Graph beschreibt also ein *Schiennetz* oder ein *Straßennetz* oder ein *Wasserleitungsnetz*. Oder: Die Kanten mögen einen Widerstand r_i (Kante i) haben, die Knoten einer Kante eine Spannungsdifferenz p_i ; so fließt nach dem Ohmschen Gesetz durch Kante i ein Strom $w_i = p_i/r_i$. Der Graph beschreibt ein *elektrisches Netzwerk*. Diese Interpretationen eines Graphen spielen für uns weiter keine Rolle; im Gegenteil: Es soll nicht

der Eindruck erweckt werden, als bauen wir im Folgenden zufällige Straßen-Netze oder zufällige elektrische Schaltpläne. Dies wäre sicher unsinnig.

Formal lautet die Definition eines Graphen so:

Definition: Ein *Graph* ist ein Paar disjunkter Mengen (V, E) mit $E \subseteq [V]^2$, wobei $[V]^2$ die Menge aller 2 elementigen Teilmengen von V bezeichnet:

$$[V]^2 := \{\{x, y\} : x, y \in V\}.$$

Wir betrachten nur endliche Graphen, also V und E sind stets endliche Mengen. Elemente von V nennt man *Ecken* oder *Knoten* des Graphen, die Elemente von E heißen *Kanten*. Ist $G = (V, E)$ ein Graph, so heißt $V = V(G)$ Knotenmenge und $E = E(G)$ Kantenmenge des Graphen. Für die Anzahl der Elemente von V schreiben wir $|V|$. Dabei wird $|V|$ die *Ordnung* von G genannt. Zwei Knoten $x, y \in V$ von G heißen *benachbart* in G , wenn $\{x, y\} \in E$ ist, wenn sie also durch eine Kante miteinander verbunden sind. Wir bezeichnen im folgenden $xy := \{x, y\}$.

Bevor wir einfache Graphen kennenlernen, springen wir ins kalte Wasser und führen den Begriff des *Zufallsgraphen* ein.

4 Ein zufälliger Graph

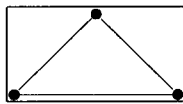
Wir wählen fortan die endliche Knotenmenge $V = \{0, 1, \dots, n-1\}$ von n Knoten. Wir verwandeln die Menge aller(!) Graphen G auf V in einen Wahrscheinlichkeitsraum: Für alle Eckenpaare $e \in [V]^2$ entscheidet man *unabhängig* voneinander, ob e eine Kante von G sein soll oder nicht. Die Wahrscheinlichkeit, dass e eine Kante sei, soll $0 < p < 1$ sein. Für dieses einfache Modell gibt es die folgende zeitliche Interpretation: Starte mit einem leeren Graphen auf n Knoten. Jede der möglichen $\binom{n}{2}$ Kanten wird mit Wahrscheinlichkeit p nacheinander eingefügt (oder mit Wahrscheinlichkeit $1-p$ nicht). Ist G_0 ein konkreter, durch diesen Prozeß entstandener Graph auf V mit m Kanten, so hat das Elementarereignis $\{G_0\}$ die Wahrscheinlichkeit

$$p^m (1-p)^{\binom{n}{2}-m}.$$

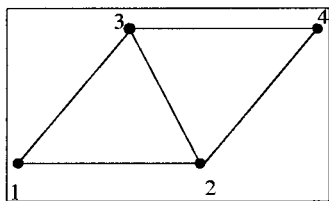
Man erkennt: Wir modellieren den $\binom{n}{2}$ -maligen unabhängigen Münzwurf mit Parameter p . Zu $e \in [V]^2$ sei $\Omega_e = \{0, 1\}$, $P_e(0) = 1-p$ und $P_e(1) = p$. Dann bezeichne $G(n, p)$ den Produktraum (Ω, P) mit $\Omega = \prod_{e \in [V]^2} \Omega_e$. P ist die entsprechende Produktwahrscheinlichkeit. Ein $\omega \in \Omega$ identifiziert man nun mit

einem Graphen G mit Knotenmenge $V = \{0, \dots, n-1\}$ und Kantenmenge $E = \{e : \omega(e) = 1\}$. Hierbei ist $\Omega \ni \omega = (\omega(e))_{e \in [V]^2}$. Man sagt, G ist ein *Zufallsgraph* auf V mit Kantenwahrscheinlichkeit p .

In einem Zufallsgraphen können wir nun (zum Beispiel) auszählen, wie häufig eine spezielle Kantenanordnung vorkommt. Wir wählen im folgenden als Beispiel ein Dreieck,



und wollen in einem wie oben beschriebenen Zufallsgraphen auszählen, wie häufig ein Dreieck vorkommt. Wir ordnen also jedem Element in $G(n, p)$ eine Zahl in $\mathbb{N} \cup \{0\}$ zu. Wir definieren eine *Zufallsgröße*! An dieser Stelle wollen wir mathematisch sehr präzise erklären, wie wir die Dreiecke zählen. Es soll eine Teilkonfiguration der Form



wie 2 Dreiecke zählen. Um dies zu klären geben wir die folgende Definition:

Definition: Es seien $G = (V, E)$ und $G' = (V', E')$ zwei Graphen. Gilt $V' \subseteq V$ und $E' \subseteq E$, so ist G' *Teilgraph* von G .

Auf $n \geq 3$ Knoten betrachten wir nun einen *Dreieck-Teilgraphen* D : Ein Dreieck ist ein Graph D mit Knotenmenge $V = \{x_0, x_1, x_2\}$ und Kantenmenge $E = \{x_0x_1, x_1x_2, x_2x_0\}$. Wir bestimmen nun die Anzahl der zu D *isomorphen* Teilgraphen D' in einem Zufallsgraphen $G \in G(n, p)$: Isomorph bedeutet hier:

Wir finden eine Eins-zu-Eins-Zuordnung f der Knoten $\{x_0, x_1, x_2\}$ von D auf die Knoten $\{y_0 = f(x_0), y_1 = f(x_1), y_2 = f(x_2)\}$ von D' , so dass

- x_0x_1 ist Kante von D impliziert, dass y_0y_1 Kante von D' ist,
- x_1x_2 ist Kante von D impliziert, dass y_1y_2 Kante von D' ist,

- x_2x_0 ist Kante von D impliziert, dass y_2y_0 Kante von D' ist.

Im obigen Bild sind die Dreiecke $\{1, 2, 3\}$ und $\{2, 3, 4\}$ isomorph: wähle $f(2) = 2, f(3) = 3$ und $f(1) = 4$.

Es sei nun $X : G(n, p) \rightarrow \mathbb{N} \cup \{0\}$ die Zufallsgröße, die jedem Zufallsgraphen $G \in G(n, p)$ die Anzahl seiner zu D isomorphen Teilgraphen zuordnet.

Wie gross ist die *mittlere Anzahl* von Dreiecken in $G \in G(n, p)$? Wir würfeln einen zufälligen Graphen aus und bestimmen die Anzahl der Dreiecke. Die mittlere Anzahl ist präzise der Erwartungswert der Zufallsgröße X . Sei D ein Dreieck mit Ecken aus $V = \{0, \dots, n-1\}$. Weiter sei

$$X_D : G \mapsto \begin{cases} 1, & \text{falls } D \subseteq G \\ 0, & \text{sonst.} \end{cases}$$

Dann ist $\mathbb{E}(X_D) = P(X_D = 1)$ und $\{X_D = 1\}$ ist die Menge aller Graphen aus $G(n, p)$ die D enthalten. Also ist

$$\mathbb{E}(X_D) = P(D \subseteq G) = p^3.$$

Wie viele dieser Dreiecke $D = \{x_0, x_1, x_2\}$ gibt es? Es gibt $n(n-1)(n-2)$ Folgen $x_0x_1x_2$ mit unterschiedlichen Ecken. Jedes Dreieck wird durch 2×3 dieser Folgen beschrieben, denn mit $x_0x_1x_2$ ist $x_1x_2x_0, x_2x_0x_1$ sowie $x_0x_2x_1, x_2x_1x_0$ und $x_1x_0x_2$ jeweils das gleiche Dreieck. So gibt es $\frac{n(n-1)(n-2)}{6}$ solcher Dreiecke. Da X die Summe aller X_D ist, wobei D die $\frac{n(n-1)(n-2)}{6}$ Dreiecke mit Ecken aus V durchläuft, folgt

$$\mathbb{E}(X) = \frac{n(n-1)(n-2)}{6} p^3 = \binom{n}{3} p^3.$$

Beispiel: Auf 20 Knoten gibt es im Mittel $\binom{20}{3} p^3$ Dreiecke.

Auch wenn im folgenden primär die Anzahl der Dreiecke genauer analysiert werden soll, betrachten wir zuvor einen kleinen *Exkurs*, der eine Zufallsvariable zum Zählen beliebiger Teilgraphen einführt. Man erkennt hieran besonders schön, wie einfach man in vielen Modellen von Spezialfällen aus abstrahieren kann.

5 Allgemeine Auszählung von Teilgraphen im $G(n, p)$ Modell

Man sagt, dass ein Graph G eine Eigenschaft $A \subseteq G(n, p)$ hat, wenn $G \in A \subseteq G(n, p)$ gilt. Zu einem beliebigen Graphen H bezeichne A_H die Grapheneigenschaft, eine Kopie (einen zu H isomorphen Teilgraphen) von H als Teilgraphen zu enthalten:

$$A_H := \{G \in G(n, p) : H \subseteq G\}.$$

Bisher betrachteten wir $H = D$, wobei D ein Dreieck war.

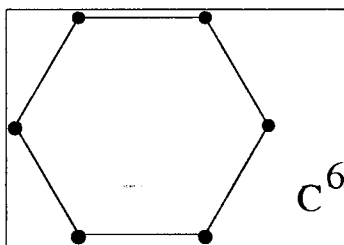
Es sei $X : G(n, p) \rightarrow \mathbb{N} \cup \{0\}$ die Anzahl der zu H isomorphen Teilgraphen in einem $G \in G(n, p)$. Wir müssen den Begriff der Isomorphie nochmals, nun allgemein, einführen:

Definition: Zwei Graphen $G = (V, E)$ und $G' = (V', E')$ heißen isomorph zueinander, wenn es eine Eins-zu-Eins-Abbildung (Bijektion) $f : V \rightarrow V'$ gibt mit

$$xy \in E \Leftrightarrow f(x)f(y) \in E' \quad \text{für alle } x, y \in V.$$

Es existiert also eine Eins-zu-Eins Umbenennung der Knoten von V , so dass jede Kante in G in eine Kante in G' übergeht, und umgekehrt!

Eine etwas speziellere Klasse von Teilgraphen H ist die Klasse der Kreise: Hier bezeichnet die Anzahl der Kanten die Länge des Kreises. Ein Kreis der Länge k wird mit C^k bezeichnet. Für $k = 3$ erhalten wir wieder ein Dreieck. C^6 ist der Kreis



6 Phasenübergänge

Treten beim Münzwurf Erscheinungen auf, die uns überraschen?

Sicher kann der Schüler nicht durch eine so gestellte Frage in eine spezielle Richtung gelenkt werden. Es gibt keinerlei Anhaltspunkte für überraschende Phänomene beim unabhängigen Münzwurf. Dabei

stellt sich die Frage, was wir mit Überraschungen meinen könnten. Etwa ein Phänomen, was man in der ersten Begegnung überraschend findet, nur weil es einer teils subjektiv gelenkten Intuition widerspricht? Diese Art Phänomene sind intensiv in der Literatur diskutiert und werden häufig – aus meiner Sicht eher verwirrend – *Paradoxa* genannt. Eine reichhaltige Sammlung solcher Paradoxien in der Wahrscheinlichkeitsrechnung findet man zum Beispiel in [6]. Verschiedene klassische Beispiele aus diesem Buch waren bereits Gegenstand vieler didaktischer Ausarbeitungen für den Schulunterricht.

Wir werden beobachten, dass im Modell des unabhängigen Münzwurfes Ereignisse bei größer werdendem Stichprobenumfang n *sprunghaft* eintreten. In Anlehnung an den Sprachgebrauch in der Physik spricht man auch in der Mathematik von einem *Phasenübergang*.

Um sprunghaftes Verhalten beobachten zu können, lassen wir nicht nur n , die Anzahl der Knoten des Zufallsgraphen, groß werden, sondern wählen die Kantenwahrscheinlichkeit p in Abhängigkeit von n : dafür schreiben wir $p(n)$. Nun studieren wir $P(G \in A_D)$ für große n bei variierendem $p(n)$. Wie gross ist die Wahrscheinlichkeit, einen zufälligen Graphen G zu realisieren, der die Eigenschaft A_D hat, also die Eigenschaft, eine Kopie eines Dreiecks als Teilgraphen zu enthalten. Wie verändert sich diese Wahrscheinlichkeit mit wachsendem n ?

Zunächst beantworten wir die Frage für eine konstante Kantenwahrscheinlichkeit p : Es gilt $P(D \subseteq G) = p^3$. Nun hat $V(G)$ genau $\lfloor n/3 \rfloor$ disjunkte Teilmengen von jeweils 3 Knoten. Hier bezeichnet $\lfloor \cdot \rfloor$ die untere Gaußklammer. Damit ist die Wahrscheinlichkeit, dass *kein* Teilgraph von $G \in G(n, p)$ isomorph zu D ist, mindestens $(1 - p^3)^{\lfloor n/3 \rfloor}$, und dies konvergiert gegen null für $n \rightarrow \infty$. Also ist $\lim_{n \rightarrow \infty} P(G \in A_D) = 1$ für festes p . Die Analyse konstanter Kantenwahrscheinlichkeiten ist also nicht so interessant.

Es sei fortan $p = p(n)$ abhängig von n .

Wir wissen bereits $\mathbb{E}(X) = \binom{4}{3} p^3(n) \leq \text{const. } n^3 p^3(n)$. Weiter gilt

$$\mathbb{E}(X) = \sum_{i \geq 0} iP(X = i) \geq \sum_{i \geq 1} P(X = i) = P(X > 0),$$

also

$$P(X > 0) \leq \mathbb{E}(X).$$

Somit ist im Falle des Zählens von Dreiecken

$$P(X > 0) \leq \text{const. } n^3 p^3(n).$$

Ist $p(n)$ nun so gewählt, dass

$$\lim_{n \rightarrow \infty} np(n) = 0,$$

so folgt

$$\lim_{n \rightarrow \infty} P(X > 0) = 0, \text{ also } \lim_{n \rightarrow \infty} P(X = 0) = 1.$$

Wir halten fest:

Satz: Wählen wir eine Kantenwahrscheinlichkeit $p(n)$, die langsamer als $1/n$ wächst, genauer, für die $\lim_{n \rightarrow \infty} np(n) = 0$ gilt, so folgt

$$\lim_{n \rightarrow \infty} P(G \in A_D) = 0.$$

Man sagt: *Fast kein* G hat die Eigenschaft A_D . Kantenwahrscheinlichkeitsfolgen wie $p(n) = 1/n^2$, allgemeiner $p(n) = 1/n^{1+\varepsilon}$, $\varepsilon > 0$, erfüllen die Eigenschaft $\lim np(n) = 0$. Dies sind die Folgen, die langsamer als $1/n$ gegen null konvergieren. Kanten treten dann mit zu kleiner Wahrscheinlichkeit auf, als dass wir für große n mit einem Dreieck als Teilgraphen im Zufallsgraphen $G \in G(n, p)$ rechnen sollten. Diese Beobachtung legt die Frage nahe: Was passiert für größere Kantenwahrscheinlichkeiten $p(n)$, sagen wir für Wahlen von $p(n)$, so dass

$$\lim_{n \rightarrow \infty} np(n) = +\infty$$

($p(n) = 1/n^{1-\varepsilon}$, $\varepsilon > 0$ erfüllen dies, so z.B. $p(n) = 1/\sqrt{n}$)? Wir werden sehen, dass in dieser Situation

$$\lim_{n \rightarrow \infty} P(X = 0) = 0, \text{ also } P(G \in A_D) \rightarrow_{n \rightarrow \infty} 1$$

gilt:

Oberhalb von $p(n) = 1/n$ *treten Dreiecke sprunghaft auf.*

Ab genügend großem n treten Dreiecke fast sicher auf, wenn $p(n)$ in Ordnung oberhalb $1/n$ liegt.

Wir wollen zeigen:

Satz: Wählen wir $p(n)$ so, dass $\lim_{n \rightarrow \infty} np(n) = +\infty$ gilt, so folgt

$$\lim_{n \rightarrow \infty} P(G \in A_D) = 1.$$

Fast alle G haben die Eigenschaft A_D .

Ein wesentlicher Beweisschritt für die zweite Aussage ist die Beobachtung, dass allgemein gilt:

$$P(X = 0) \leq P(|X - \mathbb{E}(X)| \geq \mathbb{E}(X)) \leq \frac{V(X)}{\mathbb{E}(X)^2}$$

nach der Tschebyschev-Ungleichung. Wir wollen daher zeigen:

$$\lim_{n \rightarrow \infty} \frac{V(X)}{\mathbb{E}(X)^2} = 0.$$

Wir hatten bereits die Darstellung $X = \sum_D X_D$ betrachtet, wobei sich die Summe über alle $\frac{n(n-1)(n-2)}{6}$ Dreiecke mit Ecken aus $V = \{0, \dots, n-1\}$ erstreckt. Eine wichtige Beobachtung ist nun, dass die Zufallsgrößen (X_D) *nicht* stochastisch unabhängig sind:

Sind D, D' so gewählt, dass sie eine gemeinsame Kante haben, so ist $P(X_D = 1, X_{D'} = 1) = p^5(n)$, aber $P(X_D = 1) \cdot P(X_{D'} = 1) = p^6(n)$. Die Berechnung der Varianz von $X = \sum_D X_D$ wird bekanntermaßen komplizierter, wenn die einzelnen Summanden nicht stochastisch unabhängig sind.

Wir erinnern: Es seien X und Y Zufallsgrößen. Mit $V(X) = \mathbb{E}((X - \mathbb{E}X)^2)$ folgt

$$\begin{aligned} V(X+Y) &= \mathbb{E}(((X+Y) - \mathbb{E}(X+Y))^2) \\ &= \mathbb{E}((X - \mathbb{E}X)^2) + \mathbb{E}((Y - \mathbb{E}Y)^2) \\ &\quad + 2\mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y)) \\ &= V(X) + V(Y) + 2(\mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y). \end{aligned}$$

$\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) =: \text{cov}(X, Y)$ heißt die *Kovarianz* zwischen X und Y . Beachte $\text{cov}(X, Y) = \text{cov}(Y, X)$. Für n Zufallsgrößen X_1, \dots, X_n gilt analog zur obigen Rechnung:

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + \sum_{i \neq j} \text{cov}(X_i, X_j).$$

Hier gilt also

$$V(X) = V\left(\sum_D X_D\right) = \sum_D V(X_D) + \sum_{D \neq D'} \text{cov}(X_D, X_{D'}).$$

Da $\mathbb{E}(X_D) > 0$, folgt für $D \neq D'$

$$\begin{aligned} \text{cov}(X_D, X_{D'}) &\leq \mathbb{E}(X_D X_{D'}) = P(X_D = X_{D'} = 1) \\ &= P(X_D = 1, X_{D'} = 1) \\ &= P(X_D = 1)P(X_{D'} = 1 | X_D = 1) \end{aligned}$$

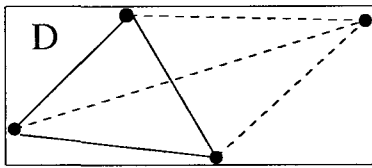
nach Definition der bedingten Wahrscheinlichkeit. Sind D, D' disjunkt, so ist $\text{cov}(X_D, X_{D'}) = 0$, denn für unabhängige $X_D, X_{D'}$ gilt $\mathbb{E}(X_D X_{D'}) = \mathbb{E}(X_D)\mathbb{E}(X_{D'})$. Zwei Dreiecke, die sich schneiden, aber nicht gleich

sind, haben 2 Knoten beziehungsweise 1 Kante gemeinsam: $|D \cap D'| = 2$. Also ist

$$\begin{aligned} \sum_{D \neq D'} \text{cov}(X_D, X_{D'}) &= \\ &= \sum_D \sum_{D': |D \cap D'|=2} \text{cov}(X_D, X_{D'}) \\ &\leq \sum_D P(X_D = 1) \sum_{D': |D \cap D'|=2} P(X_{D'} = 1 | X_D = 1). \end{aligned}$$

Beachte: $\sum_D P(X_D = 1) = \mathbb{E}(X)$. Es ist $P(X_{D'} = 1 | X_D = 1) = p^2(n)$ für D, D' mit $|D \cap D'| = 2$.

Es gibt $3(n-3)$ Dreiecke D' zu festem D mit $|D \cap D'| = 2$, denn zu jedem der $(n-3)$ Punkte außerhalb von D kann man auf 3 verschiedene Weisen 2 Kanten so legen, dass $|D \cap D'| = 2$ gilt:



Also ist $\sum_{D \neq D'} \text{cov}(X_D, X_{D'}) \leq 3(n-3)p^2(n)\mathbb{E}(X)$.

Weiter ist $V(X_D) = p^3(n)(1 - p^3(n)) \leq p^3(n) = \mathbb{E}(X_D)$. Also insgesamt

$$\begin{aligned} V(X) &\leq \mathbb{E}(X) + 3(n-3)p^2(n)\mathbb{E}(X) \\ \frac{V(X)}{\mathbb{E}(X)^2} &\leq \frac{1}{\mathbb{E}(X)} + \frac{3(n-3)p^2(n)}{\mathbb{E}(X)}. \end{aligned}$$

Da $\lim_{n \rightarrow \infty} n p(n) = +\infty$ gelten soll, folgt

$$\lim_{n \rightarrow \infty} \mathbb{E}(X) = \lim_{n \rightarrow \infty} \binom{n}{3} p^3(n) = +\infty.$$

Weiter ist

$$\lim_{n \rightarrow \infty} \frac{3(n-3)p^2(n)}{\binom{n}{3} p^3(n)} \leq \lim_{n \rightarrow \infty} \frac{3(3!)}{(n-1)(n-2)p(n)} = 0.$$

Also folgt insgesamt $\lim_{n \rightarrow \infty} \frac{V(X)}{\mathbb{E}(X)^2} = 0$, womit der Satz bewiesen ist.

Die Methode, die wir für den Beweis der beiden Sätze verwendet haben, nennt man die *Methode der ersten und der zweiten Momente*. Das erste Moment ist ein Name für $\mathbb{E}(X)$, da X^1 auftaucht. Die Methode der ersten Momente verwendet

$$P(X > 0) \leq \mathbb{E}(X).$$

Das zweite Moment ist ein Name für $\mathbb{E}(X^2)$. Die Methode der zweiten Momente verwendet

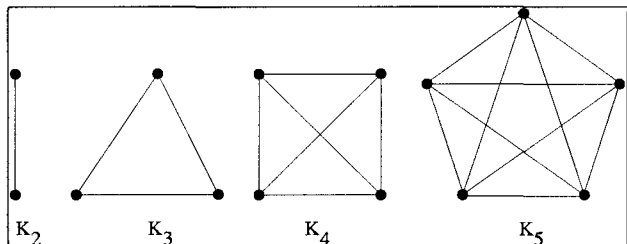
$$P(X = 0) \leq \frac{V(X)}{\mathbb{E}(X)^2} = \frac{\mathbb{E}(X^2) - \mathbb{E}(X)^2}{\mathbb{E}(X)^2}.$$

7 Allgemeine Betrachtungen und historische Bemerkungen

Wir wollen bemerken, dass nicht nur die Grapheneigenschaft, dass $G \in G(n, p)$ ein Dreieck als Teilgraphen enthält, sprunghaftes Verhalten aufweist. Viele Teilgraphen H treten sprunghaft auf ab einer Kantenvahrscheinlichkeit $p(n) = n^{-l/k}$, wobei k die Anzahl der Ecken und l die Anzahl der Kanten von H sei. Diese Universalität wurde 1960 in einer Arbeit von Erdős und Rényi [5] bewiesen.

Beispiele:

- Es sei H ein Dreieck, dann ist $l = k = 3$ und somit $k/l = 1$. $p(n) = 1/n$ ist die *Schwelle*. Dies haben wir vollständig bewiesen.
- *Kreise der Länge k*: Hier ist $l = k$, also $p(n) = 1/n$ erneut Schwelle, die die Länge des Kreises unberücksichtigt läßt. Ab $p(n) = 1/n$ tauchen Kreise beliebiger Länge auf. Das ist noch überraschender.
- Sind je zwei Ecken von H benachbart, so heißt H *vollständig* und wird mit K_k abgekürzt:



Hier ist $p(n) = n^{-2/k-1}$ die Schwelle. Oberhalb treten sprunghaft K_k 's auf.

Wir haben mit diesen Beispielen quasi einen Teil der Entwicklungsgeschichte eines zufälligen Graphen untersucht. Lassen wir $p(n)$ gedanklich von null wachsen, so tauchen ab genau charakterisierten Schwellenwerten sprunghaft diverse Teilgraphen mit einer Wahrscheinlichkeit auf, die für große n gegen 1 konvergiert.

8 Historische Bemerkungen:

Historisch sind Zufallsgraphen 1959 von Paul Erdős [4] eingeführt worden, und zwar zur Beantwortung einiger Fragen der Graphentheorie. Mittlerweile ist die Theorie der Zufallsgraphen eine eigenständige Teildisziplin geworden. Es gibt seit 10 Jahren eine Fachzeitschrift zu diesem Gebiet. In einem Lehrbuch von Bollobás [2] mit dem Titel *Random graphs* findet man eine sehr reichhaltige Kollektion von interessanten Phänomenen bei zufälligen Graphen.

9 Abschließende Bemerkungen:

Wir wollen der Frage nachgehen, in wiefern das Vorge stellte geeignet sein kann für eine Unterrichtsreihe. Die technischen Begriffe stehen für dieses Thema zur Verfügung, sobald man entlang des ersten Kapitels die Grundlagen des Bernoulli-Modells vorgestellt hat. Die Methode der ersten und zweiten Momente sind unmittelbar ableitbar. Rein technisch gesehen stellt die Bestimmung der Varianz eine gewisse Hürde dar. Man kann, neben dem hypergeometrisch-verteilten Urnenmodell, dies als ein interessantes Beispiel nicht unabhängiger Zufallsgrößen betrachten, bei dem die Bestimmung der Varianz eine Herausforderung ist. Das Thema scheint daher geeignet für einen Leistungskurs. Das hergeleitete Phänomen des Phasenübergangs bedarf allerdings rein begrifflich einer ruhigen Einführung, um genau zu verstehen, was mit der Evolution eines zufälligen Graphen gemeint ist. Wir lassen die Knotenmenge mit n wachsen und stellen dabei fest, dass unterhalb einer Schwelle für die Kantenwahrscheinlichkeit $p(n)$ Dreiecke fast sicher nicht zu beobachten sind, man also zu wenig Kanten realisiert, oberhalb allerdings fast sicher immer Dreiecke, ja sogar Kreise beliebiger Größe, vorfindet. Der *fast sichere* Gedanke muß sehr sorgfältig herausgearbeitet werden. Er ist ein Limesgedanke und unterscheidet sich im übrigen vom Sprachgebrauch der Wahrscheinlichkeitstheorie, in dem man fast sichere Ereignisse definiert. Dies führt

aber im Rahmen des Schulunterrichts zu keinerlei Verwirrung, da man den maßtheoretischen Begriff der Nullmenge gar nicht behandelt.

Die Intention dieser Darstellung ist, Schülerinnen und Schüler im Rahmen des Stochastik-Unterrichts mit Phänomenen vertraut zu machen, die den Reiz der Stochastik erhöhen sollten. Dabei ist es ein Beispiel, was sich von allen klassischen Diskussionen um kombinatorische Aufgaben und deren Paradoxien unterscheidet. Das Thema bietet darüber hinaus reichlich Gelegenheit für die Erarbeitung von Computersimulationen. Dies kann auch eine heuristische Vorbereitung für eine genauere Analyse der Interpretation des Zufallsgraphenmodells als Perkolationsmodell dienen.

Literatur

- [1] Barth, F. und Haller, R. (1998): Stochastik, Leistungskurs. Oldenbourg Verlag, München.
- [2] Bollobás, B. (1985): Random Graphs. Academic Press, London.
- [3] Eichelsbacher, P. (1998): Steinsche Methode und Zufallsgraphen, Vorlesungsskript, Sommersemester 1998, Wintersemester 1998/99.
- [4] Erdős, P. (1959). Graph theory and probability. Canadian Journal of Mathematics, 11, 34-38.
- [5] Erdős, P. and Rényi, A. (1960): On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy, 5, 17-61.
- [6] G. Székely (1990): Paradoxa, Klassische und neue Überraschungen aus Wahrscheinlichkeitsrechnung und mathematischer Statistik. Verlag Harri Deutsch, Thun.

Anschrift des Verfassers
Peter Eichelsbacher
Fakultät für Mathematik
Ruhr-Universität Bochum
NA 3 / 68
44780 Bochum
peter.eichelsbacher@ruhr-uni-bochum.de