# Die Probabilistische Methode

Wladimir Fridman
233827

Hauptseminar im Sommersemester 2004

Extremal Combinatorics

**Zusammenfassung**

Die Probabilistische Methode ist ein mächtiges Werkzeug zum Führen von Existenzbeweisen. Dieser Seminarbeitrag stellt zunächst die Idee dieser Methode vor, sowie die stochastischen Grundlagen und drei wichtige Ungleichungen, die sich bei der Anwendung der Probabilistische Methode als sehr nützlich erweisen. Abschließend wird ein Satz mittels der Probabilistischen Methode bewiesen. Als Grundlage dieses Beitrags diente das Kapitel 17 des Buches „Extremal Combinatorics“ von Stasys Jukna.

# Inhaltsverzeichnis

# 1 Einleitung: Probabilistische Methode

Die Probabilistische Methode wird dazu benutzt, Existenzbeweise zu führen. Diese Methode findet insbesondere Anwendung in der Kombinatorik und der Graphentheorie. Als ein sehr nützliches Mittel erweist sie sich auch in der Zahlentheorie und der kombinatorischen Geometrie. In der Informatik wird diese Methode zum Entwickeln effizienter Algorithmen und zur Problemanalyse angewandt.

Man will also nachweisen, dass ein Objekt mit bestimmten Eigenschaften existiert. Dazu definiert man einen geeigneten Wahrscheinlichkeitsraum und zeigt, dass ein aus dem Wahrscheinlichkeitsraum zufällig gewähltes Objekt mit einer positiven Wahrscheinlichkeit die gewünschten Eigenschaften hat.

Die Idee stützt sich auf das folgende Argument:

Seien $x_1, \ldots, x_n \in \mathbb{R}$ und

$$\frac{x_1 + \ldots + x_n}{n} \geq a$$

dann existiert mindestens ein $i \in \{1, \ldots, n\}$, so dass $x_i \geq a$.

Die zwei Thesen, auf denen die Probabilistische Methode basiert, sind also:

**These 1** Eine Zufallsvariable $X$ nimmt mindesten einen Wert $X = x$ an, so dass $x \geq E[X]$, wobei $E[X]$ der Erwartungswert von $X$ ist.

**These 2** Wenn ein aus einem Universum zufällig gewähltes Objekt mit einer positiven Wahrscheinlichkeit bestimmte Eigenschaften hat, dann muss in diesem Universum auch ein Objekt mit diesen Eigenschaften existieren.

Die Probabilistische Methode ist insofern mächtig, dass es oft einfacher ist, den Durchschnitt bzw. den Erwartungswert zu berechnen, als ein bestimmtes Objekt $x_i$ vorzuzeigen, um den Existenzbeweis zu führen.

# 2 Stochastische Grundlagen

**Definition 1** (*Wahrscheinlichkeitsraum*)
Ein *diskreter Wahrscheinlichkeitsraum* wird durch eine endliche Menge $\Omega$ und eine Funktion $Prob : \Omega \to [0, 1]$ mit der Eigenschaft $\sum_{x \in \Omega} Prob(x) = 1$ beschrieben. $\Omega$ heißt *Ergebnismenge*, $Pot(\Omega)$ *Ereignismenge*, wobei die Teilmengen $A \subseteq \Omega$ *Ereignisse* heißen. Die Wahrscheinlichkeit eines Ereignisses A ist definiert durch $Prob(A) = \sum_{x \in A} Prob(x)$. Man nennt $Prob$ eine *Wahrscheinlichkeitsverteilung*.

Die folgenden Eigenschaften lassen sich leicht aus der Definition ableiten. Seien $A$, $B$ und $C_1, \ldots, C_n$ Ereignisse und $C_1, \ldots, C_n$ eine Partition von $\Omega$, dann gilt:

- $Prob(A \cup B) = Prob(A) + Prob(B) - Prob(A \cap B)$
- $Prob(\overline{A}) = 1 - Prob(A)$
- $Prob(A \cap B) \geq Prob(A) - Prob(\overline{B})$
- $Prob(A) = \sum_{i=1}^{n} Prob(A \cap B_i)$

$\overline{A}$ bezeichne hier das Komplement von $A$ (also $\overline{A} = \Omega - A$).

**Definition 2** (*Bedingte Wahrscheinlichkeit*)
Seien $A$ und $B$ Ereignisse und $Prob(B) \neq 0$.

$$Prob(A|B) \rightleftharpoons \frac{Prob(A \cap B)}{Prob(B)}$$

heißt *bedingte Wahrscheinlichkeit* von $A$ unter der Bedingung $B$.

Bedingte Wahrscheinlichkeit kann als Wahrscheinlichkeit für das Eintreten des Ereignisses A interpretiert werden, unter der Annahme, dass das Ereignis B bereits eingetreten ist. Sei A das Ereignis, dass bei einen fairen Würfel die Nummer 2 gewürfelt wird und B das Ereignis, dass die gewürfelte Nummer gerade ist, dann ist $Prob(A|B) = \frac{1}{3}$ und $Prob(B|A) = 1$.

**Definition 3** (*Stochastische Unabhängigkeit*)
Zwei Ereignisse $A$ und $B$ heißen *stochastisch unabhängig*, falls

$$Prob(A|B) = Prob(A). \qquad (\Leftrightarrow Prob(A \cap B) = Prob(A) \cdot Prob(B))$$

Ereignisse $A_1, \ldots, A_n$ heißen *gemeinsam stochastisch unabhängig*, falls

$$Prob(A_{i_1} \cap \ldots \cap A_{i_k}) = Prob(A_{i_1}) \cdot \ldots \cdot Prob(A_{i_k})$$

für beliebige $1 \leq i_1 < \ldots < i_k \leq n$, d.h. die Wahrscheinlichkeit eines beliebigen Durchschnitts lässt sich als das Produkt der Einzelwahrscheinlichkeiten bestimmen. Zu beachten ist, dass aus paarweiser stochastischer Unabhängigkeit nicht die gemeinsame stochastische Unabhängigkeit folgt.

**Definition 4** (*Zufallsvariable*)
Eine *Zufallsvariable* ist eine auf dem Wahrscheinlichkeitsraum definierte Funktion $X : \Omega \rightarrow S$, wobei $S \subseteq \mathbb{R}$. Die *Verteilung* einer Zufallsvariablen ist eine Funktion $f : S \rightarrow [0, 1]$, definiert als $f(i) \rightleftharpoons Prob(X = i)$, wobei $Prob(X = i)$ die Wahrscheinlichkeit des Ereignisses $A = \{x \in \Omega : X(x) = i\}$ ist.

Beispiel *(Indikatorvariable, Binomialverteilung)*: Betrachte n-fachen Münzwurf, sei $p$ die Wahrscheinlichkeit für das Auftreten von Kopf. Die Würfe sind unabhängig voneinander. $\Omega = \{(x_1, \ldots, x_n) | x_i \in \{0, 1\}\}$, wobei $x_i = 1$, falls Kopf fällt, $x_i = 0$ andernfalls. Die Zufallsvariable $X$ ist die Anzahl des Auftretens von Kopf, also $X((x_1, \ldots, x_n)) = \sum_{i=1}^{n} x_i$. Die Verteilung von $X$ ist $Prob(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Man sagt die Zufallsvariable $X$ ist ***binomialverteilt*** mit Parametern $n \in \mathbb{N}$ und $p \in [0, 1]$. Hier kann man die $x_i$ als ***Indikatorvariablen*** ansehen, denn eine Indikatorvariable für ein Ereignis $A$ ist eine Zufallsvariable $X_A : \Omega \to [0, 1]$, mit $X_A(\omega) = \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{falls } \omega \notin A. \end{cases}$

**Definition 5** (*Erwartungswert*)
Der *Erwartungswert* einer Zufallsvariable $X$ ist definiert durch:

$$E[X] \rightleftharpoons \sum_{i=1}^{\infty} x_i \cdot Prob(X = x_i)$$

Seien $X_1, \ldots, X_n$ Zufallsvariablen und $a \in \mathbb{R}$, dann gilt:

- $E[aX] = aE[X]$
- $E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$ (Linearität)
- $E[X_1 \cdot X_2 \cdot \ldots \cdot X_n] = E[X_1] \cdot E[X_2] \cdot \ldots \cdot E[X_n]$, falls $X_1, \ldots, X_n$ auch gemeinsam unabhängig.

**Definition 6** (*Varianz*)
Die *Varianz* einer Zufallsvariable $X$ ist definiert durch:

$$Var[X] \rightleftharpoons E[(X - E[X])^2] \qquad / = E[X^2] - (E[X])^2 /$$

Seien $X$ und $Y$ Zufallsvariablen und $a \in \mathbb{R}$, dann gilt:

- $Var[aX] = a^2 Var[X]$
- $Var[X + Y] = Var[X] + Var[Y]$, falls $X$ und $Y$ stochastisch unabhängig.

Der Erwartungswert $E[X]$ gibt den erwarteten (durchschnittlichen) Wert von $X$ an, die Varianz $Var[X]$ die durchschnittliche Abweichung vom Erwartungswert. Sei $X$ binomialverteilt, dann ist $E[X] = np$ und $Var[X] = np(1-p)$, denn

$$E[X] \;=\; E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p = np$$

und

$$\begin{aligned} Var[X] &= Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i] = \sum_{i=1}^{n} (E[{X_i}^2] - (E[X_i])^2) \\ &= \sum_{i=1}^{n} (p - p^2) = \sum_{i=1}^{n} p(1-p) = np(1-p) \end{aligned}$$

# 3 Werkzeuge

Die folgenden drei Ungleichungen haben sich als sehr nützliche Werkzeuge bei der Anwendung der Probabilistischen Methode erwiesen.

## 3.1 Markov-Ungleichung

**Markov-Ungleichung**. Sei $X : \Omega \to \mathbb{R}^+$ eine nicht-negative Zufallsvariable und $\lambda \in \mathbb{R}^+$, dann gilt:

$$Prob(X \geq \lambda) \leq \frac{E[X]}{\lambda}.$$

Oder äquivalent

$$Prob(X \geq \lambda \cdot E[X]) \leq \frac{1}{\lambda}.$$

**Beweis**.

$$E[X] = \sum_{x} x \cdot Prob(X = x) \geq \sum_{x \geq \lambda} \lambda \cdot Prob(X = x) = \lambda \cdot Prob(X \geq \lambda)$$

## 3.2 Tschebyscheff-Ungleichung

**Tschebyscheff-Ungleichung**. Sei $X$ eine Zufallsvariable mit $Var[X] < \infty$ und $\lambda \in \mathbb{R}^+$, dann gilt

$$Prob(|X - E[X]| \geq \lambda) \leq \frac{Var[X]}{\lambda^2}$$

**Beweis**. Mit der Markov-Ungleichung folgt:

$$Prob(|X - E[X]| \geq \lambda) = Prob((X - E[X])^2 \geq \lambda^2) \leq \frac{E[(X - E[X])^2]}{\lambda^2} = \frac{Var[X]}{\lambda^2}$$

Beachte, dass hier die Markov-Ungleichung angewendet werden darf, denn $|X - E[X]|$ und damit auch $|X - E[X]|^2$ nicht negativ sind. Die Ungleichung gibt also eine obere Schranke für die Wahrscheinlichkeit, dass die Zufallsvariable von ihrem Erwartungswert um mehr als $\lambda$ abweicht, an.

## 3.3 Chernoff-Ungleichungen

Diese Ungleichungen kann mann als Spezialfälle der Markov-Ungleichung, daher mit einem größeren Informationsgehalt, ansehen, angewandt auf Summen von Zufallsvariablen $X_i$.

(**Chernoff-Ungleichung 1**.) Seien $X_1, \ldots, X_n$ n unabhängige Zufallsvariablen, mit $Prob(X_i = 1) = Prob(X_i = -1) = \frac{1}{2}$ für $i = 1, \ldots, n$, und $X = \sum_{i=1}^{n} X_i$, dann gilt für jedes $\lambda > 0$

$$Prob(X \geq \lambda) \leq e^{-\lambda^2/2n}$$

**Beweis**. Es gilt:

$$Prob(X \geq \lambda) = Prob(e^{tX} \geq e^{t\lambda}) \leq \frac{E[e^{tX}]}{e^{t\lambda}}.$$

für ein beliebiges $t \geq 0$. Der erste Teil gilt, da $exp(\cdot)$ die Ordnung beibehält und der zweite folgt aus der Markov-Ungleichung. Wegen der Taylor-Entwicklung von $e$ folgt:

$$\begin{aligned} E[e^{tX_i}] &= \frac{1}{2}e^t + \frac{1}{2}e^{-t} \\ &= \frac{1}{2}\left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\right) + \frac{1}{2}\left(1 - \frac{t}{1!} + \frac{t^2}{2!} - \frac{t^3}{3!} + \cdots\right) \\ &= \left(1 + 0 + \frac{t^2}{2!} + 0 + \cdots + \frac{t^{2k}}{(2k)!} + \cdots\right) \end{aligned}$$

Wegen $(2k)! \geq (k!)2^k$ folgt die Abschätzung:

$$E[e^{tX_i}] = \sum_{i=0}^{\infty} \frac{t^{2i}}{(2i)!} \leq \sum_{i=0}^{\infty} \frac{t^{2i}}{2^i(i!)} = \sum_{i=0}^{\infty} \frac{1}{i!}\left(\frac{t^2}{2}\right)^i = e^{t^2/2}$$

Wegen der Unabhängigkeit der $X_i$:

$$\begin{aligned} E[e^{tX}] &= E\left[e^{\left(\sum_i tX_i\right)}\right] = E\left[\prod_i e^{tX_i}\right] = \prod_{i=1}^{n} E[e^{tX_i}] \\ &\leq \prod_{i=1}^{n} e^{t^2/2} = e^{nt^2/2} \end{aligned}$$

Dieses Ergebnis oben eingesetzt ergibt:

$$Prob(X \geq \lambda) \leq \frac{E[e^{tX}]}{e^{t\lambda}} \leq \frac{e^{nt^2/2}}{e^{t\lambda}} = e^{nt^2/2-t\lambda}$$

Für $t = \lambda/n$ nimmt $e^{nt^2/2-t\lambda}$ den kleinsten Wert an. Setze also $t = \lambda/n$:

$$Prob(X \geq \lambda) \leq e^{\left(\frac{n}{2}\left(\frac{\lambda}{n}\right)^2 - \frac{\lambda}{n}\lambda\right)} = e^{-\lambda^2/2n}$$

(**Chernoff-Ungleichung 2.**) Seien $X_1, \ldots, X_n$ n unabhängige Indikatorvariablen, mit $Prob(X_i = 1) = p$ und $Prob(X_i = 0) = 1 - p$ für $i = 1, \ldots, n$ und $0 < p < 1$, und $X = \sum_{i=1}^{n} X_i$. $X$ ist also binomialverteilt $X \sim B(n, p)$ mit $E[X] = np \rightleftharpoons \mu$. Dann gilt für jedes $0 < \lambda < 1$

$$Prob(X \geq (1 + \lambda)\mu) \leq e^{-\mu\lambda^2/3} \qquad (*)$$

und

$$Prob(X \leq (1 - \lambda)\mu) \leq e^{-\mu\lambda^2/2} \qquad (**)$$

**Beweis**. Es gilt:

$$Prob(X \geq m) = Prob(e^{tX} \geq e^{tm}) \leq \frac{E[e^{tX}]}{e^{tm}}.$$

und

$$Prob(X \leq m) = Prob(e^{-tX} \geq e^{-tm}) \leq \frac{E[e^{-tX}]}{e^{-tm}}.$$

für ein beliebiges $t \geq 0$. Wegen $1 + a \leq e^a$ und der Unabhängigkeit der $X_i$ folgen die Abschätzungen:

$$\begin{aligned} E[e^{tX}] &= E\left[e^{\left(\sum_{i=1}^{n} tX_i\right)}\right] = E\left[\prod_{i}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} E[e^{tX_i}] \\ &= (pe^t + 1 - p)^n \leq e^{pn(e^t-1)} \end{aligned}$$

und

$$\begin{aligned} E[e^{-tX}] &= E\left[e^{\left(\sum_{i=1}^{n} -tX_i\right)}\right] = E\left[\prod_{i}^{n} e^{-tX_i}\right] = \prod_{i=1}^{n} E[e^{-tX_i}] \\ &= (pe^{-t} + 1 - p)^n \leq e^{pn(e^{-t}-1)} \end{aligned}$$

Diese Ergebnisse oben eingesetzt:

$$Prob(X \geq m) \leq e^{-tm} \cdot e^{pn(e^t-1)}$$

und

$$Prob(X \leq m) \leq e^{tm} \cdot e^{pn(e^{-t}-1)}$$

Setze nun in $(*)$ $t = \ln(m/pn)$, da in $(*)$ $m \geq pn$ und in $(**)$ $t = \ln(pn/m)$, da in $(**)$ $m \leq pn$

Daraus folgt für $(*)$:

$$Prob(X \geq (1+\lambda)\mu) \leq \left(\frac{\mu}{(1+\lambda)\mu}\right)^{(1+\lambda)\mu} \cdot e^{(1+\lambda)\mu-\mu} = \left(\frac{e^{\lambda}}{(1+\lambda)^{(1+\lambda)}}\right)^{\mu}$$

und für $(**)$:

$$Prob(X \leq (1-\lambda)\mu) \leq \left(\frac{\mu}{(1-\lambda)\mu}\right)^{(1-\lambda)\mu} \cdot e^{(1-\lambda)\mu-\mu} = \left(\frac{e^{-\lambda}}{(1-\lambda)^{(1-\lambda)}}\right)^{\mu}$$

Da $\lambda - \ln((1+\lambda)^{(1+\lambda)}) \leq -\lambda^2/3$ für $0 < \lambda < 1$ folgt die Behauptung $(*)$:

$$Prob(X \geq (1+\lambda)\mu) \leq e^{-\mu\lambda^2/3}$$

Da $(1-\lambda)^{(1-\lambda)} \geq e^{-\lambda+\lambda^2/2}$ für $0 < \lambda < 1$ folgt die Behauptung $(**)$:

$$Prob(X \leq (1-\lambda)\mu) \leq e^{-\mu\lambda^2/2}$$

# 4 Beispiel: First Moment Method - $k$SAT

Für jede Zufallsvariable $X$ bezeichnet man $E[X^k]$ als das $k$-te Moment von $X$. So benutzt man in der First Moment Method die Größe $E[X^1]$, also den Erwartungswert.

Die *First Moment Method* besagt:
Wenn $E[X] \leq t$, dann $Prob(X \leq t) > 0$.

Wir wollen nun den folgenden Satz beweisen:

**Satz 1**. Jede Instanz des $k$-SAT mit weniger als $2^k$ Klauseln ist erfüllbar.

**Beweis**. Man betrachte eine zufällig generierte Belegung, wobei jede Variable der Formel unabhängig von den anderen mit gleicher Wahrscheinlichkeit auf *true* oder *false* gesetzt wird. Seien $X_i$ Indikatorvariablen definiert wie folgt:

$$X_i = \begin{cases} 1 & \text{falls } i\text{-te Klausel nicht erfüllt} \\ 0 & \text{falls } i\text{-te Klausel erfüllt} \end{cases}$$

Da es für jede Klausel $2^k$ Belegungen gibt und nur eine Belegung die Klausel nicht erfüllt, ist $Prob(X_i = 1) = \frac{1}{2^k}$. Ferner sei die Zufallsvariable $X = \sum_{i=1}^{n} X_i$ die Anzahl der unerfüllten Klauseln, wobei $n$ die Anzahl der Klauseln in der Formel sei.

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{2^k} = \frac{n}{2^k}$$

Da die Anzahl der Klauseln $n < 2^k$, folgt $E[X] < 1$. Dann ist $Prob(X < 1) > 0$ (First Moment) und daraus folgt sofort $Prob(X = 0) > 0$. Also existiert eine Belegung, so dass 0 Klauseln unerfüllt, d.h. alle Klauseln erfüllt sind.

# Literatur

[1] JUKNA, S.: *Extremal Combinatorics - With Applications in Computer Science*. Springer-Verlag, 2001.