1 Macht zweier Auswahlen

Klassisches Problem

Lemma 1.1 *Wenn n Bälle unabhängig und gleichmäßig zufällig in n Eimer geworfen werden, so ist die Wahrscheinlichkeit, dass die maximale Ladung* $> \frac{3 \ln n}{\ln \ln n}$ *ist, höchstens* $\frac{1}{n}$ *für ausreichend große n.*

Proof Die Wahrscheinlichkeit, dass ein ausgewählter Eimer wenigstens M Bälle erhält ist höchstens:

$$\binom{n}{M}(\frac{1}{n})^M$$

Jetzt kann man folgende Ungleichungen aufstellen:

$$\binom{n}{M}(\frac{1}{n})^M \leq \frac{1}{M!} \leq (\frac{e}{M})^M$$

Für $M \geq \frac{3 \ln n}{lnlnn}$ wird gezeigt, dass die Wahrscheinlichkeit, dass ein beliebiger Eimer wenigstens M Bälle enthält nach oben folgendermaßen beschränkt ist:

$$\begin{aligned} n(\frac{e}{M})^M &\leq n(\frac{e \ln \ln n}{3 \ln n})^{\frac{3 \ln n}{\ln \ln n}} \\ &\leq n(\frac{\ln \ln n}{\ln n})^{\frac{3 \ln n}{\ln \ln n}} \\ &= e^{\ln n}(e^{\ln \ln \ln n - \ln \ln n})^{\frac{3 \ln n}{\ln \ln n}} \\ &= e^{\frac{3 \ln n \ln \ln \ln n}{\ln \ln n} - 2 \ln n} \\ &\leq \frac{1}{n} \end{aligned}$$

für ausreichend große n. ∎

Macht der zwei Auswahlen

Obere Schranke

Theorem 1.2 *Angenommen n Bälle werden aufeinanderfolgend folgendermaßen in n Eimern verteilt. Für jeden Ball wähle d* $>= 2$ *Eimer unabhängig und gleichverteilt zufällig (mit Ersatz). Jeder Ball wird in dem Eimer plaziert, der die wenigsten Bälle enthält, bei Gleichstand wird zufällig entschieden. Nachdem alle Bälle plaziert wurden, ist die maximale Auslastung eines beliebigen Eimers mit Wahrscheinlichkeit* $1 - o(\frac{1}{n})$ *höchstens* $\frac{\ln \ln n}{\ln d} + O(1)$.

Wir benötigen eine Abschätzung der maximalen Auslastung eines Eimer für alle Werte i. Das erreicht man, indem man die Anzahl der Eimer abschätzt, die wenigstens i Bälle enthalten. Wir wollen ein β_i finden, das mit hoher Wahrscheinlichkeit die Anzahl der Eimer mit Ladung i nach oben beschränkt. Dafür bilden wir die Höhe eines Balles, welche angibt, an welcher Position im Eimer sich der Ball befindet.
Wenn es zu jeder Zeit höchstens β_i Eimer mit Auslastung i gibt, so ist die Wahrscheinlichkeit, dass ein Ball die Höhe $i + 1$ erreicht $(\frac{\beta_i}{n})^d$. Eine Chernoff-Ungleichung erlaubt uns eine Abschätzung bzgl. der

Anzahl der Bälle mit wenigstens Höhe $i+1$. Mit hoher Wahrscheinlichkeit ist diese höchstens $2n(\frac{\beta_i}{n})^d$. Daraus folgt

$$\frac{\beta_{i+1}}{n} \leq 2(\frac{\beta_i}{n})^d. \tag{1}$$

Wir definieren folgende Notation:

- Der Zustand zur Zeit t bezeichnet den Zustand im System nachdem der t-te Ball plaziert wurde
- $h(t)$ bezeichnet die Höhe der t-ten Balles
- $\nu_i(t)$ Anzahl der Eimer mit Ladung von wenigstens i Bällen zum Zeitpunkt t, $\nu_i = \nu_i(n)$
- $\mu_i(t)$ Anzahl der Bälle mit wenigstens Höhe i zum Zeitpunkt t, $\mu_i = \mu_i(n)$
- β_i obere Schranke fuer Anzahl der Eimer mit wenigstens i Bällen
- $\mathcal{E}_i$ Ereignis, dass $\nu_i(n) \leq \beta_i$
- ω_j Menge der Eimer, die der j-te Ball zu Auswahl hat

Wir benutzen ausserdem noch 2 Hilfslemmata:

Lemma 1.3

$$\Pr(B(n,p) \geq 2np) \leq e^{-\frac{np}{3}} \tag{2}$$

Lemma 1.4 *Seien $X_1, X_2, \ldots, X_n$ eine Sequenz von Zufallsvariablen und seien $Y_1, Y_2, \ldots, Y_n$ eine Sequenz von binären Zufallsvariablen mit der Eigenschaft $Y_i = Y_i(X_1, \ldots, X_i)$. Wenn*

$$\Pr(Y_i = 1 \mid X_1, \ldots, X_{i-1}) \leq p$$

dann

$$\Pr(\sum_{i=1}^{n} Y_i > k) \leq \Pr(B(n,p) > k)$$

Beweis Theorem 1.2 Wir konstruieren uns also ein β_i, sodass mit hoher Wahrscheinlichkeit $\nu_i(n) \leq \beta_i$ für alle i gilt. Sei $\beta_4 = \frac{n}{4}$ und $\beta_{i+1} = \frac{2\beta_i^d}{n^{d-1}}$ für $4 \leq i \leq i^*$, wobei i^* noch zu bestimmen ist.
Wir müssen zeigen, dass mit hoher Wahrscheinlichkeit gilt, dass wenn $\mathcal{E}_i$ als Schranke gilt, so gilt auch $\mathcal{E}_{i+1}$ für $4 \leq i \leq i^*$. Bestimme ein i aus dem gegebenen Raum und sei Y_t eine binäre Zufallsvariable, sodass gilt:

$$Y_t = 1, \text{ genau dann wenn, } h(t) \geq i+1 \text{ und } \nu_i(t-1) \leq \beta_i$$

Dann gilt,

$$\Pr(Y_t = 1 \mid \omega_1, \ldots, \omega_{t-1}) \leq \frac{\beta_i^d}{n^d}$$

Die Wahrscheinlichkeit, dass $Y_t = 1$ liefert ist beschränkt durch $\frac{\beta_i^d}{n^d}$, da aufgrund der Definition unserer Schranke höchstens β_i Eimer Ladung i haben. Daraus folgt, dass die d Auswahlmöglichkeiten des t-ten Balles mit Wahrscheinlichkeit $\frac{\beta_i^d}{n^d}$ eine Ladung von wenigstens i haben.

Sei $p_i = \frac{\beta_i^d}{n^d}$, so folgt nach **Lemma 1.4**

$$Pr(\sum_{t=1}^{n} Y_t > k) \leq \Pr(B(n, p_i) > k)$$

Dies gilt unabhängig von allen Ereignissen $\mathcal{E}_i$, aufgrund der Definition von Y_t.
Bedingt auf $\mathcal{E}_i$, gilt $\sum_{t=1}^{n} Y_t = \mu_{i+1}$ und ausserdem gilt $\nu_{i+1} \leq \mu_{i+1}$, daraus schliesst sich

$$\begin{aligned}
\Pr(\nu_{i+1} > k \mid \mathcal{E}_i) &\leq \Pr(\mu_{i+1} > k \mid \mathcal{E}_i) \\
&= \Pr(\sum_{t=1}^{n} X_t > k \mid \mathcal{E}_i) \\
&\leq \frac{\Pr(\sum_{t=1}^{n})X_t > k)}{\mathcal{E}_i} \\
&\leq \frac{\Pr(B(n, p_i) > k)}{\mathcal{E}_i}
\end{aligned}$$

Mit der Chernoff-Ungleichung aus **Lemma 1.3** kann man nun den Rest der Binomialverteilung abschätzen. Sei $k = \beta_{i+1} = 2np_i$, dann

$$\Pr(\nu_{i+1} > \beta_{i+1} \mid \mathcal{E}_i) \leq \frac{\Pr(B(n,p) > 2np_i)}{\Pr(\mathcal{E}_i)} \leq \frac{1}{e^{\frac{np_i}{3}} \Pr(\mathcal{E}_i)}$$

daraus folgt,

$$\Pr(\neg\mathcal{E}_{i+1} \mid \mathcal{E}_i) \leq \frac{1}{n^2 \Pr(\mathcal{E}_i)}, \text{ falls } np_i \geq 6 \ln n$$

$$\Pr(\neg\mathcal{E}_{i+1}) \leq \Pr(\neg\mathcal{E}_i) + \frac{1}{n^2} \tag{3}$$

Daraus folgt, dass solange $np_i \geq 6 \ln n$ und $\mathcal{E}_i$ mit hoher Wahrscheinlichkeit gelten, solange gilt auch $\mathcal{E}_{i+1}$.
Es bleibt zu zeigen, dass $np_i < 6 \ln n$ gilt, wenn i $\frac{\ln \ln n}{\ln d}$ ist, da das unsere gewünschte Schranke für die maximale Auslastung ist. Dieser Fall muss außerdem separat gehandhabt werden, da die Chernoff Ungleichung nicht stark genug ist, um den Fall zu beschränken, wenn p_i so klein wird. Sei i^* der kleinste Wert von i, sodass $p_i = \frac{b_i^d}{n^d} < \frac{6 \ln n}{n}$. Wir zeigen per Induktion, dass i^* $\frac{\ln \ln n}{\ln d}$ $+O(1)$ ist indem wir folgende Schranke beweisen:
Induktionsanker:

$$\beta_{i+4} = \frac{n}{2^{2d^i - \sum_{j=0}^{i-1} d^j}}$$

Induktionsschritt:

$$\begin{aligned}
\beta_{(i+1)+4} &= \frac{2\beta_{i+4}^d}{n^{d-1}} \\
&= \frac{2\left(\frac{n}{2^{2d^i - \sum_{j=0}^{i-1} d^j}}\right)^d}{n^{d-1}} \\
&= \frac{n}{2d^{i+1} - \sum_{j=0}^{i} d^j}
\end{aligned}$$

Daraus folgt, dass $\beta_{i+4} \leq \frac{n}{2^{d^i}}$ und demzufolge $i^* = \frac{\ln \ln n}{\ln d} + O(1)$. Wenn wir jetzt induktiv Gleichung (3) anwenden, dann führt dies zu

$$\Pr(\neg \mathcal{E}_{i^*}) \leq \frac{i^*}{n^2}$$

Nun bestimmen wir noch die Wahrscheinlichkeit mit der unsere Schranke für die maximale Ladung eines beliebigen Eimers gilt. Da $np_i < 6 \ln n$

$$\begin{aligned}
\Pr(\nu_{i^*+1} > 18 \ln n \mid \mathcal{E}_{i^*}) &\leq \Pr(\mu_{i^*+1} > 18 \ln n \mid \mathcal{E}_{i^*}) \\
&\leq \frac{\Pr(B(n, \frac{6 \ln n}{n}) \geq 18 \ln n)}{\Pr(\mathcal{E}_{i^*})} \\
&\leq \frac{1}{n^2 \Pr(\mathcal{E}_{i^*})}
\end{aligned}$$

Wenn wir die Bedingung wie vorher entfernen, dann erhalten wir

$$\Pr(\nu_{i^*+1} > 18 \ln n) \leq \Pr(\neg \mathcal{E}_{i^*}) + \frac{1}{n^2} \leq \frac{i^* + 1}{n^2} \tag{4}$$

Wir sehen, dass

$$\Pr(\nu_{i^*+3} \geq 1) \leq \Pr(\mu_{i^*+3} \geq 1) \leq \Pr(\nu_{i^*+2} \geq 2).$$

Die letzte Wahrscheinlichkeit können wir wie folgt beschränken

$$\Pr(\mu_{i^*+2} \geq 2 \mid \nu_{i^*+1} \leq 18 \ln n) \leq \frac{\Pr(B(n, (\frac{18 \ln n}{n})^d) \geq 2)}{\Pr(\nu_{i^*+1} \leq 18 \ln n)} \leq \frac{\binom{n}{2}(\frac{18 \ln n}{n})^{2d}}{\Pr(\nu_{i^*+1} \leq 18 \ln n)}$$

Wird die Bedingung wie vorher entfernt und Gleichung (4) angewendet, dann folgt

$$\begin{aligned}
\Pr(\nu_{i^*+3} \geq 1) &\leq Pr(\mu_{i^*+2} \geq 2) \\
&\leq \frac{(18 \ln n)^{2d}}{n^{2d-2}} + \frac{i^* + 1}{n^2}
\end{aligned}$$

Dies zeigt, dass $\Pr(\nu_{i^*+3} \geq 1)$ für $d \geq 2$ $o(1/n)$ ist und dadurch die Wahrscheinlichkeit, dass die maximale Ladung eines Eimers größer als $i^* + 3 = \frac{lnlnn}{lnd} + O(1)$ ist, $o(1/n)$ ist. ∎

Untere Schranke

Theorem 1.5 *Angenommen n Bälle werden aufeinanderfolgend folgendermaßen in n Eimern verteilt. Für jeden Ball wähle $d >= 2$ Eimer unabhängig und gleichverteilt zufällig (mit Ersatz). Jeder Ball wird in* *dem Eimer plaziert, der die wenigsten Bälle enthält, bei Gleichstand wird zufällig entschieden. Nachdem alle Bälle plaziert wurden, ist die maximale Auslastung eines beliebigen Eimers mit Wahrscheinlichkeit* $1 - o(\frac{1}{n})$ *wenigstens* $\frac{\ln \ln n}{\ln d} - O(1)$.

Wir definieren eine untere Schranke γ_i, die die Anzahl der Eimer mit wenigstens Ladung i zum Zeitpunkt $n(1 - \frac{1}{2^i})$ beschränkt und beschränken dadurch die Anzahl der der Bälle mit Höhe $i + 1$ die im Interval $n(1 - \frac{1}{2^i}), n(1 - \frac{1}{2^{i+1}})$ auftreten.
Die benötigten Hilfslemmata sind ähnlich denen, die bei der oberen Schranke verwendet wurden.

Lemma 1.6

$$\Pr(B(n,p) \leq \frac{np}{2}) \leq e^{-\frac{np}{8}}$$

Lemma 1.7 *Seien* $X_1, X_2, \ldots, X_n$ *eine Sequenz von Zufallsvariablen und seinen* $Y_1, Y_2, \ldots, Y_n$ *eine Sequenz von binären Zufallsvariablen mit der Eigenschaft* $Y_i = Y_i(X_1, \ldots, X_i)$. *Wenn*

$$\Pr(Y_i = 1 \mid X_1, \ldots, X_{i-1}) \geq p$$

dann

$$\Pr(\sum_{i=1}^{n} Y_i > k) \geq \Pr(B(n,p) > k)$$

Beweis Theorem 1.5 Sei $\mathcal{F}_i$ das Ereignis, dass $\nu_i(n(1 - \frac{1}{2^i})) \geq \gamma_i$ mit γ_i folgendermaßen definiert

$$\begin{aligned} \gamma_0 &= n; \\ \gamma_{i+1} &= \frac{n}{2^{i+3}} (\frac{\gamma_i}{n})^d \end{aligned}$$

$\mathcal{F}_0$ gilt mit Wahrscheinlichkeit 1, wir müssen nun induktiv zeigen, dass folgende $\mathcal{F}_i$ mit ausreichend hoher Wahrscheinlichkeit gelten, damit unsere angestrebte untere Schranke auch gilt.
Wir definieren uns eine binäre Zufallsvariable für t im Raum $R = n(1 - \frac{1}{2^i}), n(1 - \frac{1}{2^{i+1}})$

$$Z_t = 1 \text{ genau dann wenn, } h(t) = i + 1 \text{ oder } \nu_{i+1}(t - 1) \geq \gamma_{i+1}.$$

Die Wahrscheinlichkeit, dass der t-te Ball exakt die Höhe $i + 1$ hat ist also

$$(\frac{\nu_i(t-1)}{n})^d - (\frac{\nu_{i+1}(t-1)}{n})^d.$$

ω_i repräsentiert wieder die Eimer, die der j-te Ball zur Auswahl hat, dann können wir schliessen

$$\Pr(Z_t = 1 \mid \omega_1, \ldots, \omega_{t-1}, \mathcal{F}_i) \geq (\frac{\gamma_i}{n})^d - (\frac{\gamma_{i+1}}{n})^d.$$

Bedingt durch $\mathcal{F}_i$, gilt ausserdem $\nu_i(t - 1) \geq \gamma_i$. Aus der Definition von γ_i können wir schliessen, dass

$$\Pr(Z_t = 1 \mid \omega_1, \ldots, \omega_{t-1}, \mathcal{F}_i) \geq (\frac{\gamma_i}{n})^d - (\frac{\gamma_{i+1}}{n})^d \geq \frac{1}{2}(\frac{\gamma_i}{n})^d.$$

Sei $p_i = \frac{1}{2}(\frac{\gamma_i}{n})^d$, wenden wir **Lemma 1.7** an, so erhalten wir

$$\Pr(\sum_{t\in\mathbb{R}} Z_t < k \mid \mathcal{F}_i) \leq \Pr(B(\frac{n}{2^{i+1}, p_i}) < k).$$

Aus der Definition von γ_i koennen wir weiterhin schliessen

$$\gamma_{i+1} = \frac{1}{2}\frac{n}{2^{i+1}}p_i.$$

Mit **Lemma 1.6** gilt

$$\Pr(B(\frac{n}{2^{i+1}}, p_i) < \gamma_{i+1}) \leq e^{-\frac{np_i}{(8*2^{i+1})}}.$$

Falls $\frac{np_i}{2^{i+1}} \geq 17\ln n$ ergibt sich $o(\frac{1}{n^2})$. Sei nun i^* eine untere Grenze für die größte Ganzzahl, für die dies gilt. Dann können wir zeigen, dass i^* als $\frac{\ln\ln n}{\ln d} - O(1)$ gewählt werden kann. Nehmen wir an dies gilt, dann haben wir für alle $i \leq i^*$ gezeigt, dass

$$\Pr(\sum_{t\in\mathbb{R}} Z_t < \gamma_{i+1} \mid \mathcal{F}_i) \leq \Pr(B(\frac{n}{2^{i+1}, p_i}) < \gamma_{i+1}) = o(\frac{1}{n^2}).$$

Ausserdem nach Definition gilt, dass $\sum_{t\in\mathbb{R}} Z_t < \gamma_{i+1}$ bedeutet dass $\neg\mathcal{F}_{i+1}$. Daraus können wir schliessen

$$\Pr(\neg\mathcal{F}_{i+1} \mid \mathcal{F}_i) \leq \Pr(\sum_{t\in\mathbb{R}} Z_t < \gamma_{i+1} \mid \mathcal{F}_i) = o(\frac{1}{n^2}).$$

Für hinreichend große n gilt,

$$\begin{aligned}\Pr(\mathcal{F}_{i^*}) &\geq \Pr(\mathcal{F}_{i^*} \mid \mathcal{F}_{i^*-1}) * \Pr(\mathcal{F}_{i^*-1} \mid \mathcal{F}_{i^*-2}) * \ldots * \Pr(\mathcal{F}_0)\\ &\geq (1 - \frac{1}{n^2})^{i^*}\\ &= 1 - o(\frac{1}{n}).\end{aligned}$$

Es bleibt noch zu zeigen, dass $\frac{\ln\ln n}{lnd} - O(1)$ eine angemessene Schranke für i^* ist. Es reicht hier zu zeigen, dass $\gamma_i \geq 17\ln n$ ist wenn $i = \frac{\ln\ln n}{lnd} - O(1)$. Aus unserer Definition von γ_i ergibt sich

$$\gamma_i = \frac{n}{2^{\sum_{k=0}^{i-1}(i+2-k)d^k}}.$$
$$\gamma_i \geq \frac{n}{2^{10d^{i-1}}}.$$

Wir suchen also nach dem größtstem i, sodass gilt

$$\frac{n}{2^{10d^{i-1}}} \geq 17\ln n.$$ Nach i auflösen liefert

$$i \leq \frac{\ln\ln n}{lnd} - O(1).$$

∎