Script generated by TTT

Title: Esparza: DWT (24.05.2012)

Date: Thu May 24 14:17:32 CEST 2012

Duration: 88:50 min

Pages: 55

8. Abschätzen von Wahrscheinlichkeiten

8.1 Die Ungleichungen von Markov und Chebyshev

Satz 39 (Markov-Ungleichung)

Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$

Äquivalent dazu:

$$\Pr[X \geq t \cdot \mathbb{E}[X]] \leq 1/t\,.$$

8. Abschätzen von Wahrscheinlichkeiten

8.1 Die Ungleichungen von Markov und Chebyshev

Satz 39 (Markov-Ungleichung)

Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$

Äquivalent dazu:

$$\Pr[X \geq t \cdot \mathbb{E}[X]] \leq 1/t\,.$$

Beispiel 41

Wir werfen 1000-mal eine ideale Münze und ermitteln die Anzahl X der Würfe, in denen „Kopf“ fällt.

X ist binomialverteilt mit $X \sim \mathrm{Bin}(1000, p = \frac{1}{2})$,also gilt

$$\mathbb{E}[X] = \frac{1}{2}n = 500 \text{ und } \mathrm{Var}[X] = \frac{1}{4}n = 250.$$

Wie groß ist die Wahrscheinlichkeit, dass mehr als 550-mal „Kopf“ fällt?

Beispiel 41

Chebyshev-Ungleichung:

$$\Pr[X \geq 550] \leq \Pr[|X - 500| \geq 50] \leq \frac{250}{50^2} = 0{,}1\,.$$

Setze nun $n = 10000$ und betrachte wieder eine maximal 10%-ige Abweichung vom Erwartungswert:

$$\mathbb{E}[X] = 5000 \text{ und } \mathrm{Var}[X] = 2500, \text{ und damit}$$

$$\Pr[X \geq 5500] \leq \Pr[|X - 5000| \geq 500] \leq \frac{2500}{500^2} = 0{,}01\,.$$

Beweis:

Für $t > 0$ gilt

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}]\,.$$

Mit der Markov-Ungleichung folgt

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}] \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}}.$$

Wegen der Unabhängigkeit der Zufallsvariablen $X_1, \ldots, X_n$ gilt

$$\mathbb{E}[e^{tX}] = \mathbb{E}\left[\exp\left(\sum_{i=1}^n tX_i\right)\right] = \mathbb{E}\left[\prod_{i=1}^n e^{tX_i}\right] = \prod_{i=1}^n \mathbb{E}[e^{tX_i}].$$

Weiter ist für $i \in \{1, \ldots, n\}$:

$$\mathbb{E}[e^{tX_i}] = e^{t\cdot 1}p_i + e^{t\cdot 0}(1-p_i) = e^t p_i + 1 - p_i = 1 + p_i(e^t - 1)\,,$$

$x \geq y \iff e^x \geq e^y$

$\Pr[tX \geq t(1+\delta)\mu]$

gdw

ω

$e^{tX}(\omega)$

$= e^{tX(\omega)}$

Beweis:

Für $t > 0$ gilt

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}]\,.$$

Mit der Markov-Ungleichung folgt

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}] \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}}.$$

Wegen der Unabhängigkeit der Zufallsvariablen $X_1, \ldots, X_n$ gilt

$$\mathbb{E}[e^{tX}] = \mathbb{E}\left[\exp\left(\sum_{i=1}^n tX_i\right)\right] = \mathbb{E}\left[\prod_{i=1}^n e^{tX_i}\right] = \prod_{i=1}^n \mathbb{E}[e^{tX_i}].$$

Weiter ist für $i \in \{1, \ldots, n\}$:

$$\mathbb{E}[e^{tX_i}] = e^{t\cdot 1}p_i + e^{t\cdot 0}(1-p_i) = e^t p_i + 1 - p_i = 1 + p_i(e^t - 1)\,,$$

Beweis (Forts.):

und damit

$$\begin{aligned}
\Pr[X \geq (1+\delta)\mu] &\leq \frac{\prod_{i=1}^n (1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\
&\leq \frac{\prod_{i=1}^n \exp(p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\
&= \frac{\exp(\sum_{i=1}^n p_i(e^t - 1))}{e^{t(1+\delta)\mu}} = \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} =: f(t)\,.
\end{aligned}$$

Wir wählen nun t so, dass $f(t)$ minimiert wird, nämlich

$$t = \ln(1+\delta)\,.$$

Damit wird

$$f(t) = \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \frac{e^{\delta\mu}}{(1+\delta)^{(1+\delta)\mu}}\,.$$

$x \geq y \iff e^x \geq e^y$

Beweis:

$\Pr[tX \geq t(1+\delta)\mu]$

ω

Für $t > 0$ gilt

gdw

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}] .$$

$e^{tX}(\omega) = e^{tX(\omega)}$

Mit der Markov-Ungleichung folgt

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}] \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} .$$

Wegen der Unabhängigkeit der Zufallsvariablen $X_1, \ldots, X_n$ gilt

$$\mathbb{E}[e^{tX}] = \mathbb{E}\left[\exp\left(\sum_{i=1}^{n} tX_i\right)\right] = \mathbb{E}\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} \mathbb{E}[e^{tX_i}].$$

Weiter ist für $i \in \{1, \ldots, n\}$:

$$\mathbb{E}[e^{tX_i}] = e^{t \cdot 1} p_i + e^{t \cdot 0}(1 - p_i) = e^t p_i + 1 - p_i = 1 + p_i(e^t - 1) ,$$

Beispiel 44

Wir betrachten wieder das Beispiel, dass wir eine faire Münze n-mal werfen und abschätzen wollen, mit welcher Wahrscheinlichkeit „Kopf“

$$\frac{n}{2}(1 + 10\%)$$

oder öfter fällt.

n	Chebyshev	Chernoff
1000	0,1	0,0889
10000	0,01	$0{,}308 \cdot 10^{-10}$
n	$\frac{\frac{1}{4}n}{(0{,}1 \cdot \frac{1}{2}n)^2} = \frac{100}{n}$	$\left(\frac{e^{0,1}}{(1+0{,}1)^{1+0,1}}\right)^{n/2} \approx 0.9975^n$

$\Pr[|X - \mathbb{E}[X]| \geq t]$

$= \Pr\left[\frac{|X - \mathbb{E}[X]|}{\mathbb{E}[X]} \geq \frac{t}{\mathbb{E}[X]}\right]$

"

$1{,}1$

$$\Pr[|X-\mathbb{E}[X]| \geq t]$$

$$= \Pr\left[\frac{|X-\mathbb{E}(X)|}{\mathbb{E}(X)} \geq \frac{t}{\mathbb{E}(X)}\right]$$

$$= \Pr\left[0{,}1 \geq \frac{t}{\frac{1}{2}n}\right] = \Pr\left[\frac{n}{20} \geq t\right]$$

$p = 1/2$

Beispiel 44

Wir betrachten wieder das Beispiel, dass wir eine faire Münze n-mal werfen und abschätzen wollen, mit welcher Wahrscheinlichkeit „Kopf"

$$\underline{\frac{n}{2}(1+10\%)}$$

oder öfter fällt.

$$\Pr[|X-\mathbb{E}[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$$

$$\mathrm{Var}[X] = n/4$$

n	Chebyshev	Chernoff
1000	0,1	0,0889
10000	0,01	$0{,}308 \cdot 10^{-10}$
n	$\frac{\frac{1}{4}n}{(0{,}1\cdot\frac{1}{2}n)^2} = \frac{100}{n}$	$\left(\frac{e^{0,1}}{(1+0{,}1)^{1+0,1}}\right)^{n/2} \approx 0.9975^n$

Satz 45

Seien $X_1,\ldots,X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i=1]=p_i$ und $\Pr[X_i=0]=1-p_i$. Dann gilt für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$, sowie jedes $0<\delta<1$, dass

$$\Pr[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}.$$

Beweis:

Analog zum Beweis von Satz 43. □

Bemerkung: Abschätzungen, wie sie in Satz 43 und Satz 45 angegeben sind, nennt man auch tail bounds, da sie Schranken für die tails, also die vom Erwartungswert weit entfernten Bereiche angeben. Man spricht hierbei vom upper tail (vergleiche Satz 43) und vom lower tail (vergleiche Satz 45).

Die Chernoff-Schranken hängen exponentiell von μ ab!

8.3 Chernoff-Schranken

8.3.1 Chernoff-Schranken für Summen von 0–1–Zufallsvariablen

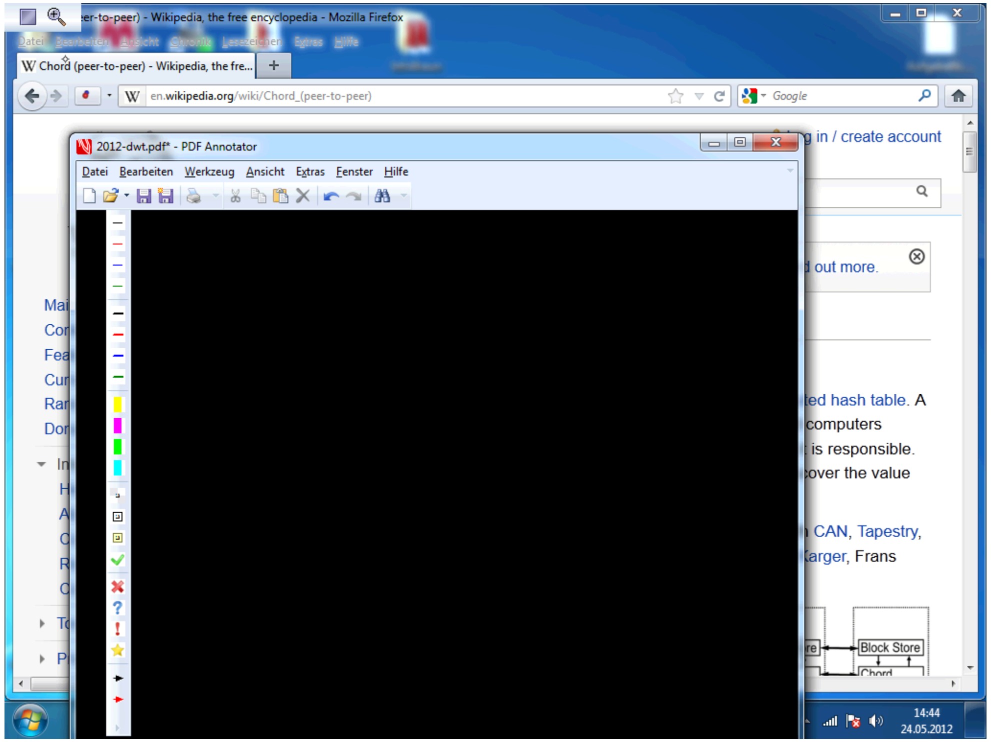
Die hier betrachtete Art von Schranken ist nach *Herman Chernoff* (*1923) benannt. Sie finden in der komplexitätstheoretischen Analyse von Algorithmen eine sehr häufige Verwendung.

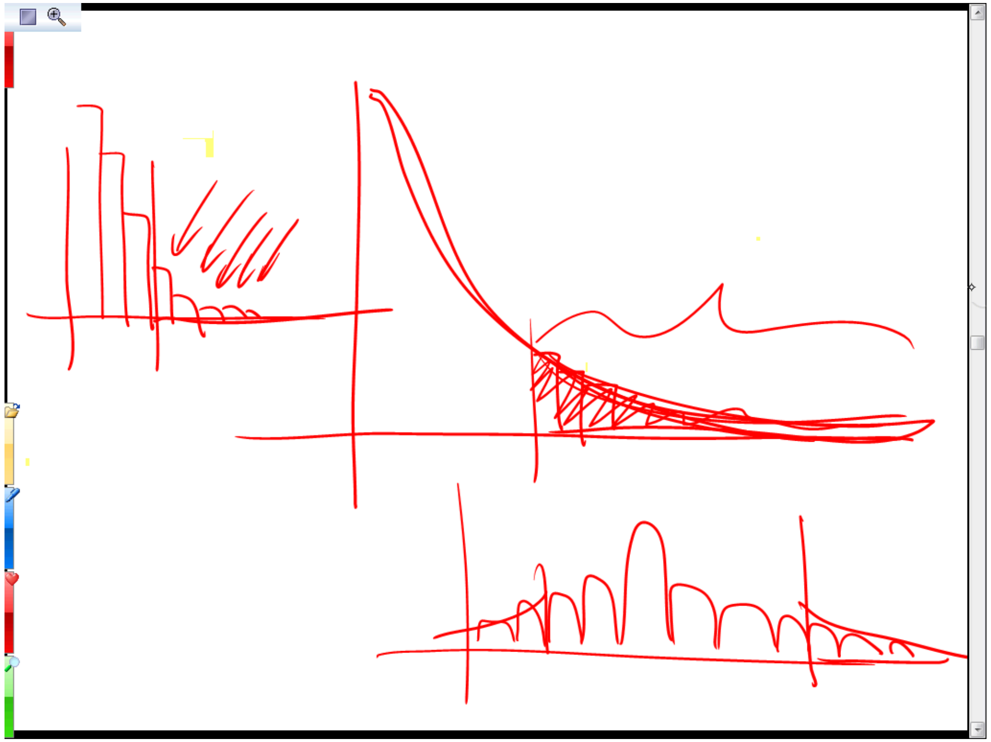
Satz 43

Seien $X_1,\ldots,X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i=1]=p_i$ und $\Pr[X_i=0]=1-p_i$. Dann gilt für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$, sowie jedes $\delta>0$, dass

$$\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}.$$

$\delta = 0{,}1$





Lemma 46

Für $0 \leq \delta < 1$ *gilt*

$$(1-\delta)^{1-\delta} \geq e^{-\delta+\delta^2/2} \quad \textit{und} \quad (1+\delta)^{1+\delta} \geq e^{\delta+\delta^2/3}.$$

Beweis:

Wir betrachten

$$f(x) = (1-x)\ln(1-x) \text{ und } g(x) = -x + \frac{1}{2}x^2 .$$

Es gilt für $0 \leq x < 1$:

$$g'(x) = x - 1 \leq -\ln(1-x) - 1 = f'(x)$$

sowie

$$f(0) = 0 = g(0)\,,$$

also im angegebenen Intervall $f(x) \geq g(x)$.

Die Ableitung der zweiten Ungleichung erfolgt analog. $\square$

Korollar 47

Seien $X_1, \ldots, X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gelten folgende Ungleichungen für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$:

1. $\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \leq 1{,}81$,
2. $\Pr[X \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2}$ *für alle* $0 < \delta \leq 1$,
3. $\Pr[|X - \mu| \geq \delta\mu] \leq 2e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \leq 1$,
4. $\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$ *und*
5. $\Pr[X \geq t] \leq 2^{-t}$ *für* $t \geq 2e\mu$.

Beispiel 48

Wir betrachten wieder balls into bins und werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe. Sei

$$Y_i := \text{Anzahl der Bälle im } i\text{-ten Korb}$$

für $i = 1, \ldots, n$, sowie $Y := \max_{1 \leq i \leq n} Y_i$.

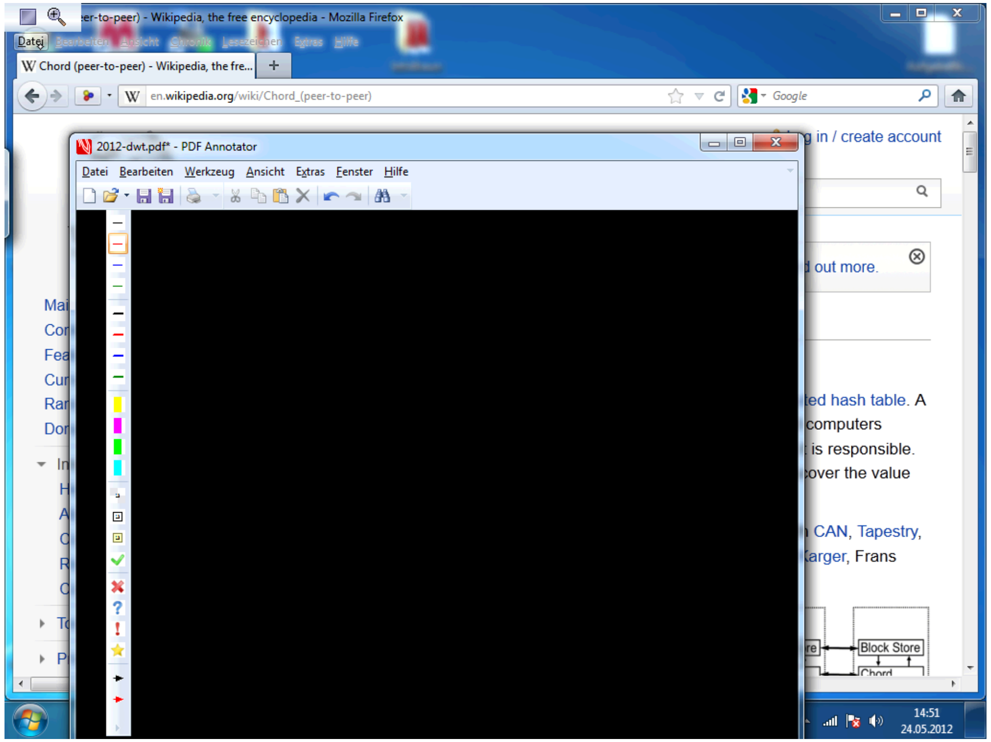
Für die Analyse von Y_i ($i \in \{1, \ldots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 47. Dabei ist X_j die Variable, die angibt, ob der j-te Ball in den i-ten Korb fällt. Wir haben $p_1 = \ldots = p_n = \frac{1}{n}$ und $\mu = 1$. Wir nehmen $t = 2\log n$. Es folgt

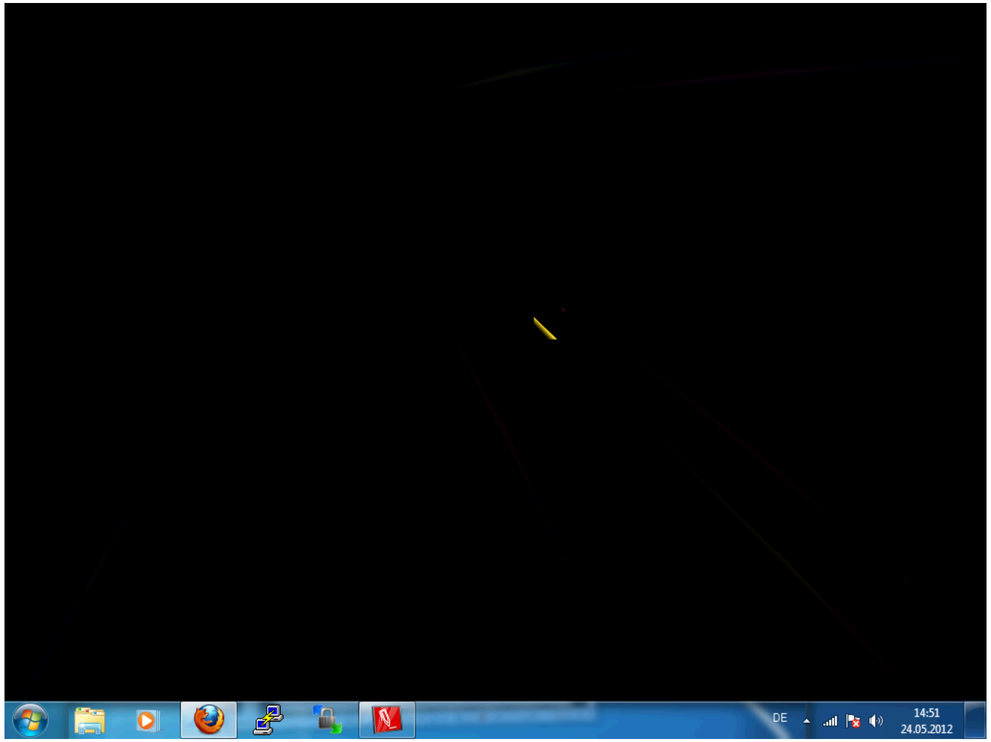
$$\Pr[X_i \geq 2\log n] \leq 1/n^2.$$

Daraus ergibt sich

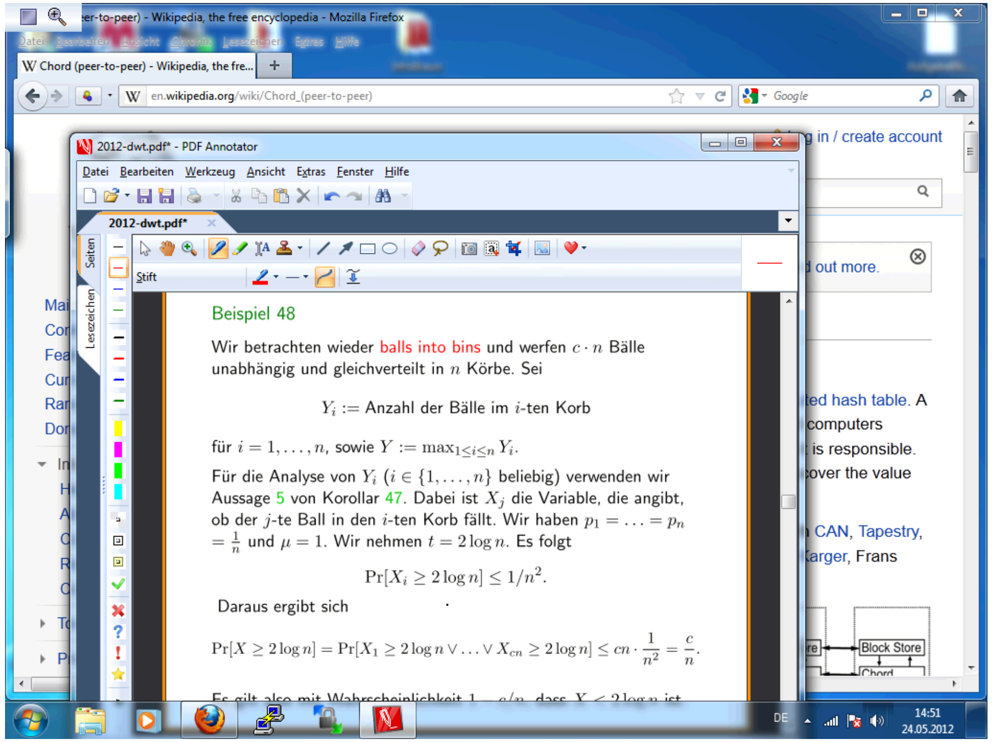
$$\Pr[X \geq 2\log n] = \Pr[X_1 \geq 2\log n \vee \ldots \vee X_{cn} \geq 2\log n] \leq cn \cdot \frac{1}{n^2} = \frac{c}{n}.$$

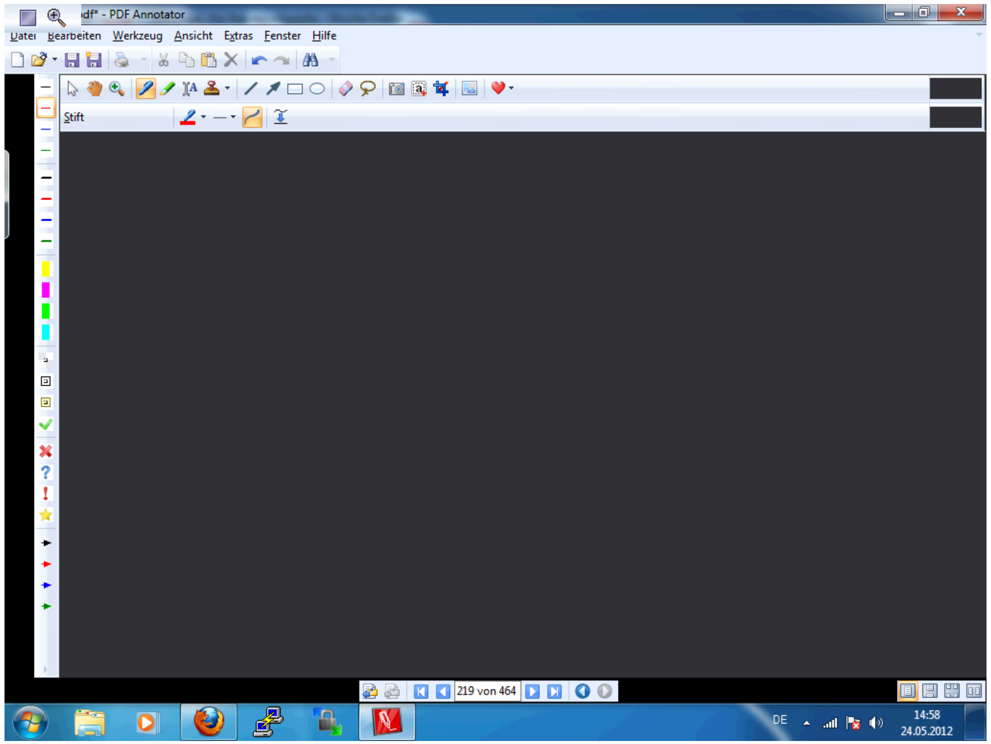
Es gilt also mit Wahrscheinlichkeit $1 - c/n$, dass $X < 2\log n$ ist.

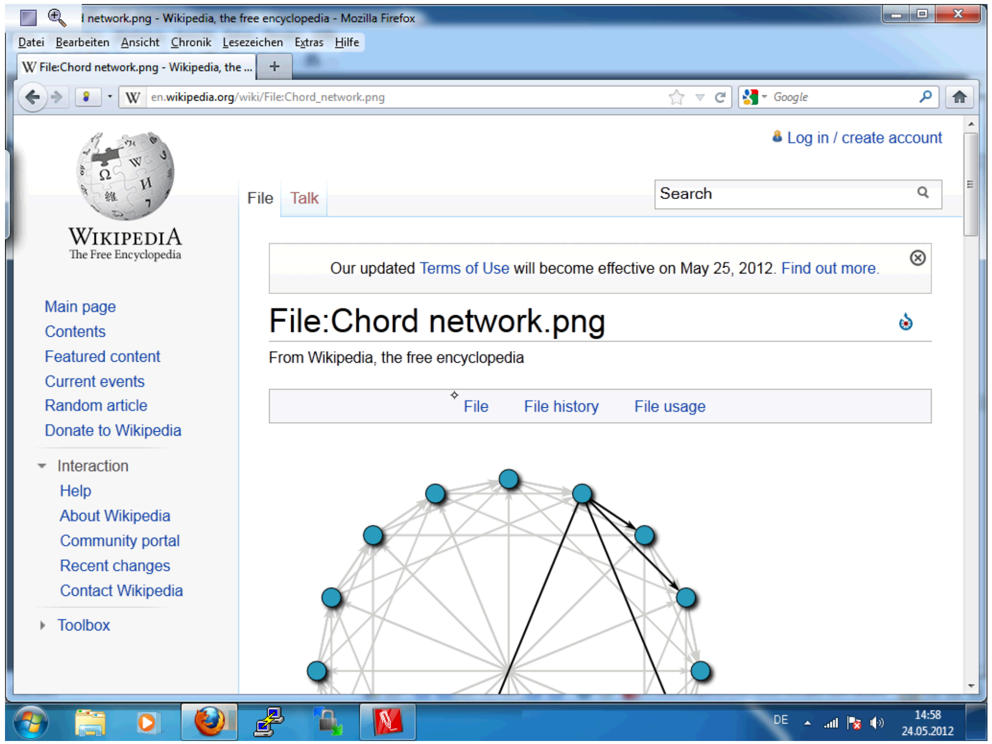


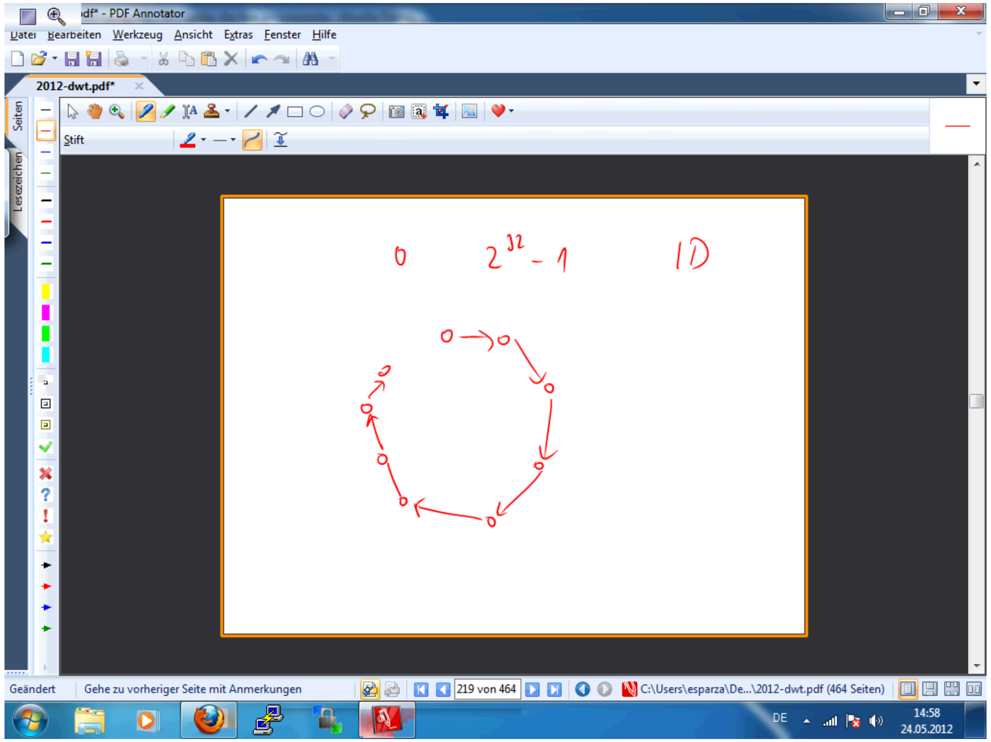



Chord (peer-to-peer) - Wikipedia, the free encyclopedia - Mozilla Firefox
Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
en.wikipedia.org/wiki/Chord_(peer-to-peer)
Log in / create account
Article Talk Read Edit View history Search
WIKIPEDIA
The Free Encyclopedia
Our updated Terms of Use will become effective on May 25, 2012. Find out more.
Chord (peer-to-peer)
From Wikipedia, the free encyclopedia
In computing, Chord is a protocol and algorithm for a peer-to-peer distributed hash table. A distributed hash table stores key-value pairs by assigning keys to different computers (known as "nodes"); a node will store the values for all the keys for which it is responsible. Chord specifies how keys are assigned to nodes, and how a node can discover the value for a given key by first locating the node responsible for that key.[1]
Chord is one of the four original distributed hash table protocols, along with CAN, Tapestry, and Pastry. It was introduced in 2001 by Ion Stoica, Robert Morris, David Karger, Frans Kaashoek, and Hari Balakrishnan, and was developed at MIT.[2]
Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Toolbox
Print/export
Contents [hide]
1 Overview
File System
Block Store
Block Store
Block Store
14:51
24.05.2012


2012-dwt.pdf* - PDF Annotator
Datei Bearbeiten Werkzeug Ansicht Extras Fenster Hilfe
Stift
Beispiel 48
Wir betrachten wieder balls into bins und werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe. Sei
$Y_i :=$ Anzahl der Bälle im i-ten Korb
für $i = 1, \ldots, n$, sowie $Y := \max_{1 \le i \le n} Y_i$.
Für die Analyse von Y_i ($i \in \{1, \ldots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 47. Dabei ist X_j die Variable, die angibt, ob der j-te Ball in den i-ten Korb fällt. Wir haben $p_1 = \ldots = p_n = \frac{1}{n}$ und $\mu = 1$. Wir nehmen $t = 2 \log n$. Es folgt
$\Pr[X_i \ge 2 \log n] \le 1/n^2$.
Daraus ergibt sich
$\Pr[X \ge 2 \log n] = \Pr[X_1 \ge 2 \log n \vee \ldots \vee X_{cn} \ge 2 \log n] \le cn \cdot \frac{1}{n^2} = \frac{c}{n}$.
14:51
24.05.2012


PDF Annotator
Datei Bearbeiten Werkzeug Ansicht Extras Fenster Hilfe
Stift
219 von 464
14:58
24.05.2012


File:Chord network.png - Wikipedia, the free encyclopedia - Mozilla Firefox
Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
en.wikipedia.org/wiki/File:Chord_network.png
Log in / create account
File Talk Search
WIKIPEDIA
The Free Encyclopedia
Our updated Terms of Use will become effective on May 25, 2012. Find out more.
File:Chord network.png
From Wikipedia, the free encyclopedia
File File history File usage
Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Toolbox
14:58
24.05.2012



Beispiel 48

Wir betrachten wieder balls into bins und werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe. Sei

$$Y_i := \text{Anzahl der Bälle im } i\text{-ten Korb}$$

für $i = 1, \ldots, n$, sowie $Y := \max_{1 \le i \le n} Y_i$.

Für die Analyse von Y_i ($i \in \{1, \ldots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 47. Dabei ist X_j die Variable, die angibt, ob der j-te Ball in den i-ten Korb fällt. Wir haben $p_1 = \ldots = p_n = \frac{1}{n}$ und $\mu = 1$. Wir nehmen $t = 2\log n$. Es folgt

$$\Pr[X_i \ge 2\log n] \le 1/n^2.$$

Daraus ergibt sich

$$\Pr[X \ge 2\log n] = \Pr[X_1 \ge 2\log n \vee \ldots \vee X_{cn} \ge 2\log n] \le cn \cdot \frac{1}{n^2} = \frac{c}{n}.$$

Es gilt also mit Wahrscheinlichkeit $1 - c/n$, dass $X < 2\log n$ ist.

Korollar 47

Seien $X_1, \ldots, X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gelten folgende Ungleichungen für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$:

1. $\Pr[X \ge (1+\delta)\mu] \le e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \le 1{,}81$,
2. $\Pr[X \le (1-\delta)\mu] \le e^{-\mu\delta^2/2}$ *für alle* $0 < \delta \le 1$,
3. $\Pr[|X - \mu| \ge \delta\mu] \le 2e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \le 1$,
4. $\Pr[X \ge (1+\delta)\mu] \le \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$ *und*
5. $\Pr[X \ge t] \le 2^{-t}$ *für* $t \ge 2e\mu$.

$\Pr[X \geq t] \leq 2^{-t}$

Beispiel 48

Wir betrachten wieder balls into bins und werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe. Sei

$$Y_i := \text{Anzahl der Bälle im } i\text{-ten Korb}$$

für $i = 1, \ldots, n$, sowie $Y := \max_{1 \le i \le n} Y_i$.

Für die Analyse von Y_i ($i \in \{1, \ldots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 47. Dabei ist X_j die Variable, die angibt, ob der j-te Ball in den i-ten Korb fällt. Wir haben $p_1 = \ldots = p_n = \frac{1}{n}$ und $\mu = 1$. Wir nehmen $t = 2\log n$. Es folgt

$$\Pr[X_i \ge 2\log n] \le 1/n^2.$$

Daraus ergibt sich

$$\Pr[X \ge 2\log n] = \Pr[X_1 \ge 2\log n \vee \ldots \vee X_{cn} \ge 2\log n] \le cn \cdot \frac{1}{n^2} = \frac{c}{n}.$$

Es gilt also mit Wahrscheinlichkeit $1 - c/n$, dass $X < 2\log n$ ist.

Korollar 47

Seien $X_1, \dots, X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gelten folgende Ungleichungen für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$:

1. $\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \leq 1{,}81$,
2. $\Pr[X \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2}$ *für alle* $0 < \delta \leq 1$,
3. $\Pr[|X - \mu| \geq \delta\mu] \leq 2e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \leq 1$,
4. $\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$ *und*
5. $\boxed{\Pr[X \geq t] \leq 2^{-t} \quad \textit{für } t \geq 2e\mu.}$

$X = Y_i$

$\Pr[X \geq t] \leq 2^{-t}$

Beispiel 48

Wir betrachten wieder balls into bins und werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe. Sei

X_{ij}

$$Y_i := \text{Anzahl der Bälle im } i\text{-ten Korb}$$

für $i = 1, \dots, n$, sowie $Y := \max_{1 \leq i \leq n} Y_i$.

Für die Analyse von Y_i ($i \in \{1, \dots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 47. Dabei ist X_j die Variable, die angibt, ob der j-te Ball in den i-ten Korb fällt. Wir haben $p_1 = \ldots = p_n = \frac{1}{n}$ und $\mu = 1$. Wir nehmen $t = 2\log n$. Es folgt

$$\Pr[X_i \geq 2\log n] \leq 1/n^2.$$

Daraus ergibt sich

$$\Pr[X \geq 2\log n] = \Pr[X_1 \geq 2\log n \vee \ldots \vee X_{cn} \geq 2\log n] \leq cn \cdot \frac{1}{n^2} = \frac{c}{n}.$$

Es gilt also mit Wahrscheinlichkeit $1 - c/n$, dass $X < 2\log n$ ist.

Beweis:

1 und 2 folgen direkt aus Satz 43 bzw. 45 und Lemma 46.

Aus 1 und 2 zusammen folgt 3.

Die Abschätzung 4 erhalten wir direkt aus Satz 43, da für den Zähler gilt

$$e \leq e^{(1+\delta)}.$$

5 folgt aus 4, indem man $t = (1+\delta)\mu$ setzt, $t \geq 2e\mu$:

$$\left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu} \leq \left(\frac{e}{t/\mu}\right)^t \leq \left(\frac{1}{2}\right)^t.$$

$\square$

Korollar 47

Seien $X_1, \dots, X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gelten folgende Ungleichungen für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$:

1. $\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \leq 1{,}81$,
2. $\Pr[X \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2}$ *für alle* $0 < \delta \leq 1$,
3. $\Pr[|X - \mu| \geq \delta\mu] \leq 2e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \leq 1$,
4. $\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$ *und*
5. $\Pr[X \geq t] \leq 2^{-t}$ *für* $t \geq 2e\mu$.

$\Pr[X \geq t] \leq 2^{-t}$ $\quad$ $\Pr[\cancel{X} \geq t] \leq 2^{-t}$ $\quad$ $t \geq \underline{2e}$

Beispiel 48

Wir betrachten wieder balls into bins und werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe. Sei

X_{ij}

$$Y_i := \text{Anzahl der Bälle im } i\text{-ten Korb}$$

2^{16} 32

für $i = 1, \dots, n$, sowie $Y := \max_{1 \le i \le n} Y_i$.

Für die Analyse von Y_i ($i \in \{1, \dots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 47. Dabei ist X_j die Variable, die angibt, ob der j-te Ball in den i-ten Korb fällt. Wir haben $p_1 = \dots = p_n = \frac{1}{n}$ und $\mu = 1$. Wir nehmen $t = 2\log n$. Es folgt

$$\Pr[\cancel{X}_i \ge 2\log n] \le 1/n^2.$$

Daraus ergibt sich

$$\Pr[X \ge 2\log n] = \Pr[X_1 \ge 2\log n \vee \ldots \vee X_{cn} \ge 2\log n] \le cn \cdot \frac{1}{n^2} = \frac{c}{n}.$$

Es gilt also mit Wahrscheinlichkeit $1 - c/n$, dass $X < 2\log n$ ist.

Korollar 47

Seien $X_1, \dots, X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ und $\Pr[X_i = 0] = 1 - p_i$. Dann gelten folgende Ungleichungen für $X := \sum_{i=1}^n X_i$ und $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$:

1. $\Pr[X \ge (1+\delta)\mu] \le e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \le 1,81$,
2. $\Pr[X \le (1-\delta)\mu] \le e^{-\mu\delta^2/2}$ *für alle* $0 < \delta \le 1$,
3. $\Pr[|X - \mu| \ge \delta\mu] \le 2e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \le 1$,
4. $\Pr[X \ge (1+\delta)\mu] \le \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$ *und*
5. $\Pr[X \ge t] \le 2^{-t}$ *für* $t \ge 2e\mu$.

$2^{-2\log n} = \frac{1}{n^2}$

$\Pr[X \ge t] \le 2^{-t}$ $\quad$ $\Pr[\cancel{X} \ge t] \le 2^{-t}$ $\quad t \ge 2e$

Beispiel 48

Wir betrachten wieder balls into bins und werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe. Sei

X_{ij}

$$Y_i := \text{Anzahl der Bälle im } i\text{-ten Korb}$$

2^{16} 32

für $i = 1, \dots, n$, sowie $Y := \max_{1 \le i \le n} Y_i$.

Für die Analyse von Y_i ($i \in \{1, \dots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 47. Dabei ist X_j die Variable, die angibt, ob der j-te Ball in den i-ten Korb fällt. Wir haben $p_1 = \dots = p_n = \frac{1}{n}$ und $\mu = 1$. Wir nehmen $t = 2\log n$. Es folgt

$$\Pr[\cancel{X}_i \ge 2\log n] \le 1/n^2.$$

$\Pr(A_1 \cup \ldots \cup A_n) \le \Pr(A_1) + \ldots + \Pr(A_n)$

Daraus ergibt sich

$$\Pr[\cancel{X} \ge 2\log n] = \Pr[\cancel{X}_1 \ge 2\log n \cup \ldots \cup \cancel{X}_{cn} \ge 2\log n] \le cn \cdot \frac{1}{n^2} = \frac{c}{n}.$$

Es gilt also mit Wahrscheinlichkeit $1 - c/n$, dass $X < \boxed{2\log n}$ ist.

Beispiel 48

Wir betrachten wieder balls into bins und werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe. Sei

X_{ij}

X_i

$$Y_i := \text{Anzahl der Bälle im } i\text{-ten Korb}$$

2^{16} 32

für $i = 1, \dots, n$, sowie $Y := \max_{1 \le i \le n} Y_i$.

Für die Analyse von Y_i ($i \in \{1, \dots, n\}$ beliebig) verwenden wir Aussage 5 von Korollar 47. Dabei ist X_j die Variable, die angibt, ob der j-te Ball in den i-ten Korb fällt. Wir haben $p_1 = \dots = p_n = \frac{1}{n}$ und $\mu = 1$. Wir nehmen $t = 2\log n$. Es folgt

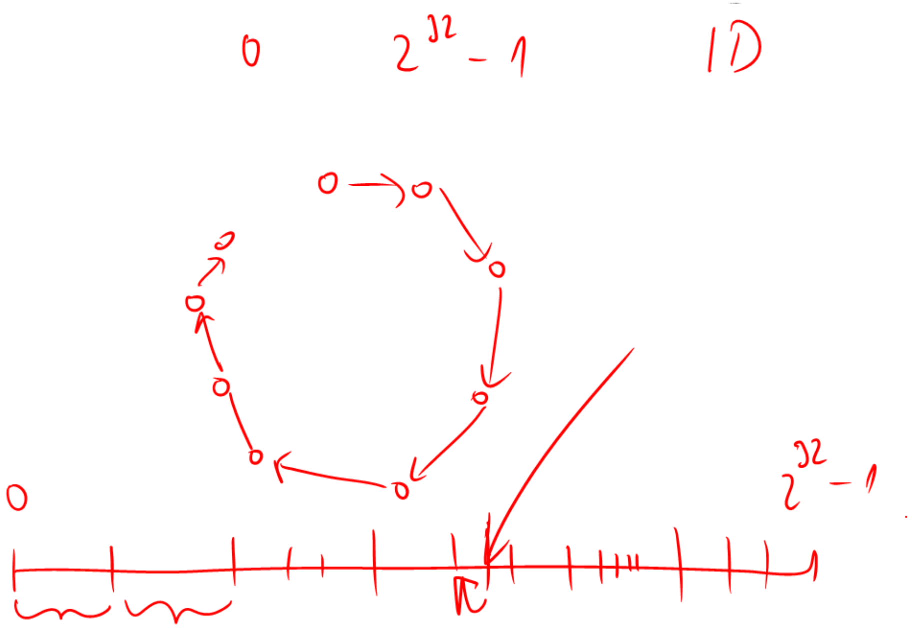
$$\Pr[\cancel{X}_i \ge 2\log n] \le 1/n^2.$$

$\Pr(A_1 \cup \ldots \cup A_n) \le \Pr(A_1) + \ldots + \Pr(A_n)$

Daraus ergibt sich

$$\Pr[\cancel{X} \ge 2\log n] = \Pr[\cancel{X}_1 \ge 2\log n \cup \ldots \cup \cancel{X}_{cn} \ge 2\log n] \le cn \cdot \frac{1}{n^2} = \frac{c}{n}.$$

Es gilt also mit Wahrscheinlichkeit $1 - c/n$, dass $X < \boxed{2\log n}$ ist.



Bei wahrscheinlichkeitserzeugenden Funktionen haben wir kein Problem mit der Konvergenz, da für $|s| < 1$ gilt

$$\begin{aligned} |G_X(s)| &= \left| \sum_{k=0}^{\infty} \Pr[X=k] \cdot s^k \right| \\ &\le \sum_{k=0}^{\infty} \Pr[X=k] \cdot |s^k| \le \sum_{k=0}^{\infty} \Pr[X=k] = 1\,. \end{aligned}$$

Bernoulli-Verteilung

Sei X eine Bernoulli-verteilte Zufallsvariable mit $\Pr[X=0] = 1-p$ und $\Pr[X=1] = p$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = (1-p) \cdot s^0 + p \cdot s^1 = 1 - p + ps\,.$$

Gleichverteilung auf $\{0,\ldots,n\}$

Sei X auf $\{0,\ldots,n\}$ gleichverteilt, d.h. für $0 \le k \le n$ ist $\Pr[X=k] = 1/(n+1)$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{n} \frac{1}{n+1} \cdot s^k = \frac{s^{n+1}-1}{(n+1)(s-1)}\,.$$

Binomialverteilung

Für $X \sim \mathrm{Bin}(n,p)$ gilt nach der binomischen Formel

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot s^k = (1-p+ps)^n\,.$$

Geometrische Verteilung

Sei X eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit p. Dann gilt

$$\begin{aligned} G_X(s) = \mathbb{E}[s^X] &= \sum_{k=1}^{\infty} p(1-p)^{k-1} \cdot s^k \\ &= ps \cdot \sum_{k=1}^{\infty} ((1-p)s)^{k-1} = \frac{ps}{1-(1-p)s}\,. \end{aligned}$$

Bernoulli-Verteilung

Sei X eine Bernoulli-verteilte Zufallsvariable mit $\Pr[X = 0] = 1 - p$ und $\Pr[X = 1] = p$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = (1-p) \cdot s^0 + p \cdot s^1 = 1 - p + ps\,.$$

Gleichverteilung auf $\{0, \ldots, n\}$

Sei X auf $\{0, \ldots, n\}$ gleichverteilt, d.h. für $0 \le k \le n$ ist $\Pr[X = k] = 1/(n+1)$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{n} \frac{1}{n+1} \cdot s^k = \frac{s^{n+1} - 1}{(n+1)(s-1)}\,.$$

Beobachtung:

Sei $Y := X + t$ mit $t \in \mathbb{N}_0$. Dann gilt

$$G_Y(s) = \mathbb{E}[s^Y] = \mathbb{E}[s^{X+t}] = \mathbb{E}[s^t \cdot s^X] = s^t \cdot \mathbb{E}[s^X] = s^t \cdot G_X(s)\,.$$

Ebenso lässt sich leicht nachrechnen, dass

$$G'_X(s) = \sum_{k=1}^{\infty} k \cdot \Pr[X = k] \cdot s^{k-1}, \text{ also}$$
$$G'_X(0) = \Pr[X = 1], \text{ sowie}$$
$$G_X^{(i)}(0) = \Pr[X = i] \cdot i!, \text{ also}$$
$$G_X^{(i)}(0)/i! = \Pr[X = i]\,.$$

Beispiel 51

Sei X binomialverteilt mit $X \sim \text{Bin}(n, \lambda/n)$, Für $n \to \infty$ folgt

$$G_X(s) = \left(1 - \frac{\lambda}{n} + \frac{\lambda s}{n}\right)^n = \left(1 + \frac{\lambda(s-1)}{n}\right)^n \to e^{\lambda(s-1)}\,.$$

Man kann beweisen, dass aus der Konvergenz der wahrscheinlichkeitserzeugenden Funktion die Konvergenz der Verteilung folgt.

Beispiel 52

Sei X binomialverteilt mit $X \sim \text{Bin}(n, p)$, also

$$G_X(s) = (1 - p + ps)^n\,.$$

Dann gilt

$$G'_X(s) = n \cdot (1 - p + ps)^{n-1} \cdot p$$

und somit

$$\mathbb{E}[X] = G'_X(1) = np\,.$$

Beispiel 52

Sei X binomialverteilt mit $X \sim \mathrm{Bin}(n,p)$, also

$$G_X(s) = (1-p+ps)^n\,.$$

Dann gilt

$$G'_X(s) = n \cdot (1-p+ps)^{n-1} \cdot p$$

und somit

$$\mathbb{E}[X] = G'_X(1) = np\,.$$

Bernoulli-Verteilung

Sei X eine Bernoulli-verteilte Zufallsvariable mit $\Pr[X=0] = 1-p$ und $\Pr[X=1] = p$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = (1-p)\cdot s^0 + p \cdot s^1 = 1-p+ps\,.$$

Gleichverteilung auf $\{0,\ldots,n\}$

Sei X auf $\{0,\ldots,n\}$ gleichverteilt, d.h. für $0 \le k \le n$ ist $\Pr[X=k] = 1/(n+1)$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{n} \frac{1}{n+1}\cdot s^k = \frac{s^{n+1}-1}{(n+1)(s-1)}\,.$$

Binomialverteilung

Für $X \sim \mathrm{Bin}(n,p)$ gilt nach der binomischen Formel

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot s^k = (1-p+ps)^n\,.$$

Geometrische Verteilung

Sei X eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit p. Dann gilt

$x = k$

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=1}^{\infty} p(1-p)^{k-1} \cdot s^k$$
$$= ps \cdot \sum_{k=1}^{\infty} ((1-p)s)^{k-1} = \frac{ps}{1-(1-p)s}\,.$$

Poisson-Verteilung

Für $X \sim \mathrm{Po}(\lambda)$ gilt

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}\cdot s^k = e^{-\lambda+\lambda s} = e^{\lambda(s-1)}.$$

Beispiel 51

Sei X binomialverteilt mit $X \sim \mathrm{Bin}(n, \lambda/n)$, Für $n \to \infty$ folgt

$$G_X(s) = \left(1 - \frac{\lambda}{n} + \frac{\lambda s}{n}\right)^n = \left(1 + \frac{\lambda(s-1)}{n}\right)^n \to e^{\lambda(s-1)}.$$

Man kann beweisen, dass aus der Konvergenz der wahrscheinlichkeitserzeugenden Funktion die Konvergenz der Verteilung folgt.

9.0.2 Zusammenhang zwischen der w.e. Funktion und den Momenten

Da

$$G_X(s) := \sum_{k=0}^{\infty} \Pr[X = k] \cdot s^k = \mathbb{E}[s^X],$$

gilt

$$\boxed{G'_X(1)} = \sum_{k=1}^{\infty} k \cdot \Pr[X = k] = \mathbb{E}[X].$$

$\Pr[X=0] + \Pr[X=1]\, s + \Pr[X=2]\, s^2 + \Pr[X=3]\, s^3 + \ldots$

$\Pr[X=1] + 2\Pr[X=2]\, s + 3\Pr[X=3]\, s^2 + \ldots$

$\Pr[X=1] + 2 \cdot \Pr[X=2] + 3 \cdot \Pr[X=3]$

Beispiel 52

Ebenso ergibt sich

$$\mathbb{E}[X(X-1)\ldots(X-k+1)] = G_X^{(k)}(1),$$

also etwa

$$\begin{aligned}\underline{\mathrm{Var}[X]} &= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2\\ &= G''_X(1) + G'_X(1) - (G'_X(1))^2.\end{aligned}$$

Andere Momente von X kann man auf ähnliche Art und Weise berechnen.

$G''_X(1) \quad G'_X(1) \quad G'_X(1) \quad G_X^{(2)}(1)$

Beispiel 52

Ebenso ergibt sich

$$\underbrace{\mathbb{E}[X(X-1)\ldots(X-k+1)]} = \boxed{G_X^{(k)}(1)},$$

also etwa

$$\begin{aligned}\underline{\mathrm{Var}[X]} &= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2\\ &= G''_X(1) + G'_X(1) - (G'_X(1))^2.\end{aligned}$$

Andere Momente von X kann man auf ähnliche Art und Weise berechnen.