A&W 2026

G-13 Timo Stucki, LFW E 13

Week 7

Minitest 4

Distributions Recap (optional)

Multiple Random Variables (condensed)

Important Inequalities with Exercise

Infos

Feel free to contact me:

- tistucki@student.ethz.ch
- Discord: timostucki

Material:

- timostucki.com

Distributions

Recap (optional)

Distributions

Important Distributions

Name	Notation	Support	Density	Expectation	Variance
Bernoulli	$\mathrm{Bernoulli}(p)$	$\{0, 1\}$	$f_X(i) = \begin{cases} p & \text{for } i = 1, \\ 1-p & \text{for } i = 0. \end{cases}$	p	$p(1-p)$
Binomial	$\mathrm{Bin}(n, p)$	$\{0, 1, \ldots, n\}$	$f_X(i) = \binom{n}{i} p^i (1-p)^{n-i}$	np	$np(1-p)$
Geometric	$\mathrm{Geo}(p)$	$\{1, 2, 3, \ldots\}$	$f_X(i) = p(1-p)^{i-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson	$\mathrm{Po}(\lambda)$	$\{0, 1, 2, \ldots\}$	$f_X(i) = \frac{e^{-\lambda}\lambda^i}{i!}$	λ	λ

Bernoulli Distribution

2.5.1 Bernoulli-Verteilung

Eine Zufallsvariable X mit $W_X = \{0, 1\}$ und der Dichte

$$f_X(x) = \begin{cases} p & \text{für } x = 1, \\ 1 - p & \text{für } x = 0, \\ 0 & \text{sonst} \end{cases}$$

$\mathbb{E}[X] = p \quad \text{und} \quad \mathrm{Var}[X] = p(1 - p)$

Bernoulli Distribution

We prove $\mathrm{Var}[X] = p(1-p)$ for $X \sim \mathrm{Bernoulli}(p)$ using the identity:

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

$$\mathbb{E}[X] = 1(p) + 0(1-p) = p \implies (\mathbb{E}[X])^2 = p^2$$

$$\mathbb{E}[X^2] = 1^2(p) + 0^2(1-p) = p$$

Plugging these into the variance formula:

$$\begin{aligned} \mathrm{Var}[X] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= p - p^2 \\ &= \mathbf{p(1 - p)} \end{aligned}$$

Binomial Distribution

2.5.2 Binomialverteilung

Eine Bernoulli-verteilte Zufallsvariable erhalten wir zum Beispiel als Indikator für 'Kopf', wenn wir ein Münze einmal werfen. Werfen wir die Münze statt dessen n-mal und fragen wie oft wir Kopf erhalten haben, so ist die entsprechende Zufallsvariable *binomialverteilt*.

"If I throw $n = 25$ coins, what is the probability that $i = 10$ land on heads"

Binomial Distribution

“If I throw n = 25 coins, what is the probability that i = 10 land on heads”

$$\Pr[X=i] = \sum_{\omega\in\{K,Z\}^n, X(\omega)=i} \Pr[\omega] = \sum_{\omega\in\{K,Z\}^n, X(\omega)=i} p^i(1-p)^{n-i}$$

$$= p^i(1-p)^{n-i} \cdot |\{\omega \in \{K,Z\}^n,\ \omega \text{ enthält genau } i\text{-mal Kopf}\}|,$$

“(Anzahl Abfolgen mit x-mal Kopf) *
P[Spezifische Abfolge {K,Z,...,K} trifft ein mit genau x-mal Kopf]”

→ $\binom{n}{x}p^x(1-p)^{n-x}$

Binomial Distribution

$$X \sim \mathrm{Bin}(n, p)$$

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x \in \{0, 1, \dots, n\} \\ 0, & \text{sonst.} \end{cases}$$

$$\mathbb{E}[X] = np \quad \text{und} \quad \mathrm{Var}[X] = np(1-p)$$

Poisson Distribution

"What is the probability that exactly $i = 10$ people suffer from a heart attack in an hour?"

- The expected number of heart attacks per hour is $\lambda = 5$ (not really)
- The number of time "instances" where a person could suffer a heart attack within an hour is essentially infinite $n \to \infty$ (informal)

Poisson Distribution

“What is the probability that exactly $i = 10$ people suffer from a heart attack in an hour?”

- The expected number of heart attacks per hour is $\lambda = 5$ (not really)
- The number of time “instances” where a person could suffer a heart attack within an hour is essentially infinite $n \rightarrow \infty$ (informal)

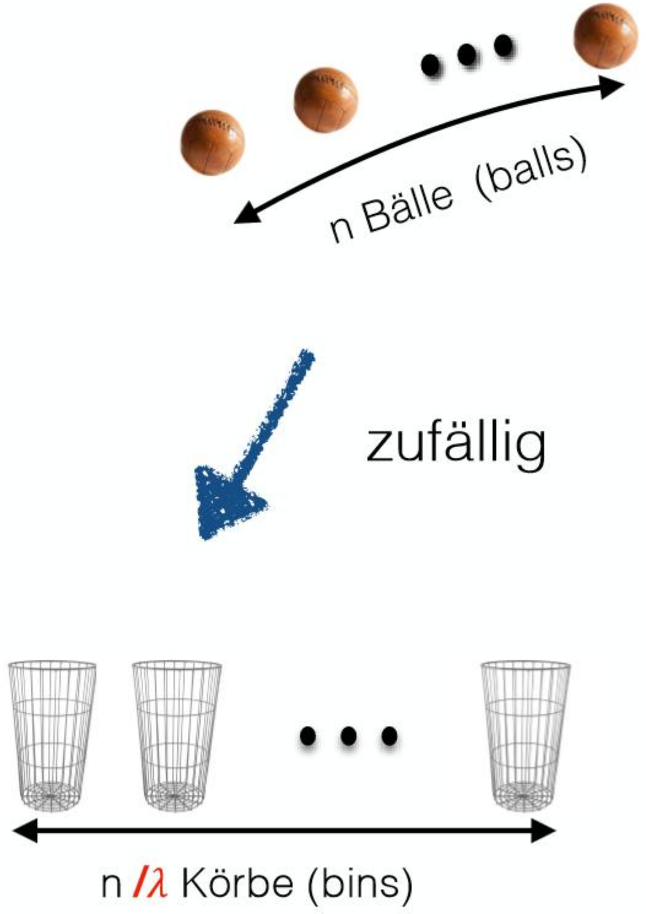
$$X \sim \mathrm{Po}(\lambda)$$

$$f_X(i) = \begin{cases} \frac{e^{-\lambda}\lambda^i}{i!} & \text{für } i \in \mathbb{N}_0 \\ 0 & \text{sonst.} \end{cases}$$

$$\mathbb{E}[X] = \mathrm{Var}[X] = \lambda$$

Poisson Distribution

Bin(n,λ/n) konvergiert für n → ∞ gegen Po(λ)



Poisson Distribution

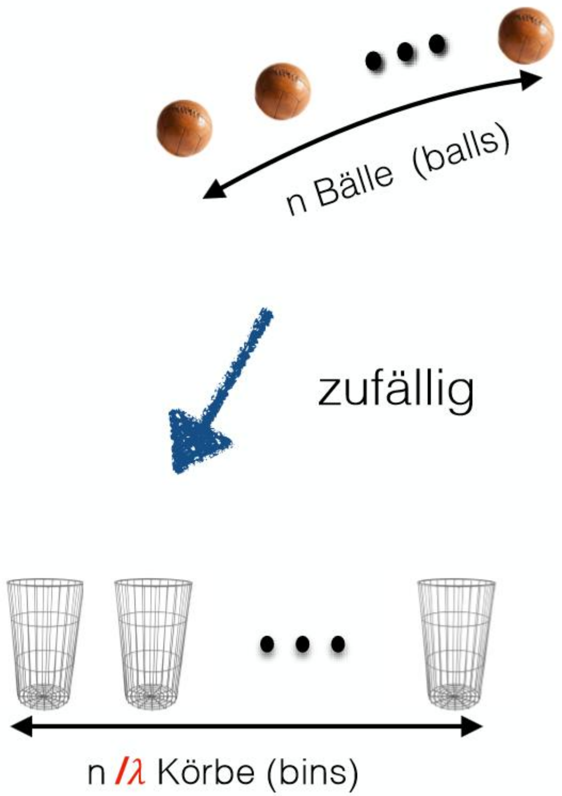
$X :=$ Anzahl Bälle im **ersten** Korb

$X \sim \text{Bin}(n, \lambda/n)$

$$\mathbb{E}[X] = n \cdot \frac{\lambda}{n} = \lambda$$

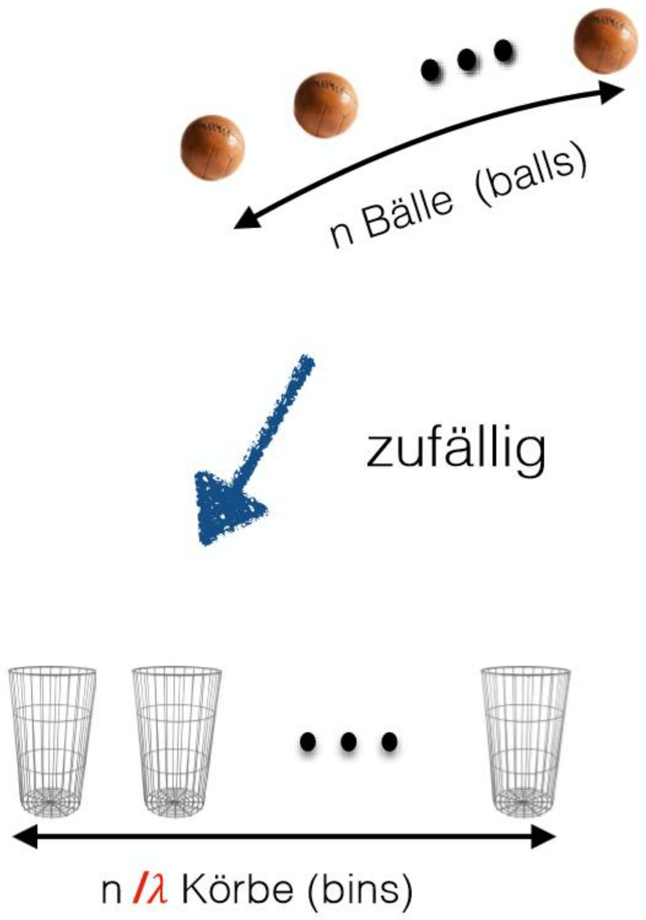
$$\lim_{n\to\infty} \Pr[X = i] = \lim_{n\to\infty} \underbrace{\binom{n}{i}\left(\frac{\lambda}{n}\right)^i}_{= \frac{n\cdot(n-1)\cdot\ldots\cdot(n-i+1)}{n\cdot n\cdot\ldots\cdot n}\cdot\frac{\lambda^i}{i!} \to \frac{\lambda^i}{i!}} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{n-i}}_{\approx\left(e^{-\frac{\lambda}{n}}\right)^{n-i} = e^{-\lambda\frac{n-i}{n}} \to e^{-\lambda}}$$

$$= \frac{\lambda^i}{i!}e^{-\lambda}$$



Poisson Distribution

$\boxed{\text{Bin}(n, \boldsymbol{\lambda}/n) \quad \text{konvergiert für } n \to \infty \quad \text{gegen} \quad \text{Po}(\boldsymbol{\lambda})}$

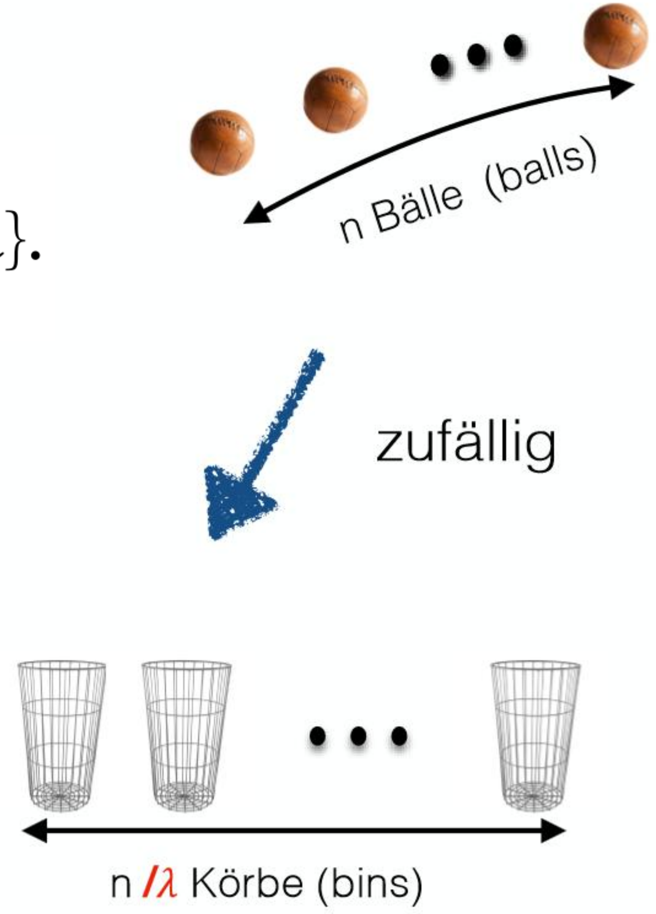


Poisson Distribution

- Case: $\lambda = 1$

$$X \sim \mathrm{Bin}(n, 1/n)$$

$$f_X(i) = \binom{n}{i} \cdot \frac{1}{n^i} \cdot \left(1 - \frac{1}{n}\right)^{n-i} \qquad \text{für alle } i \in \{0, 1, \ldots, n\}.$$



Poisson Distribution

- Case: $\lambda = 1$

$$X \sim \mathrm{Bin}(n, 1/n)$$

$$f_X(i) = \binom{n}{i} \cdot \frac{1}{n^i} \cdot \left(1 - \frac{1}{n}\right)^{n-i} \qquad \text{für alle } i \in \{0, 1, \ldots, n\}.$$

$$\lim_{n\to\infty} f_X(i) = \lim_{n\to\infty} \frac{n(n-1)\ldots(n-i+1)}{i!} \cdot \frac{1}{n^i} \cdot \left(1 - \frac{1}{n}\right)^{n-i} = \frac{1}{i!} \cdot e^{-1},$$

das heisst, für $n \to \infty$ konvergiert X, bzw. die Binomialverteilung $\mathrm{Bin}(n, 1/n)$, gegen die Poisson-Verteilung $\mathrm{Po}(1)$.

Geometric Distribution

"I flip an unfair coin (p = 0.4) until it lands on heads. What is the probability that I need to wait exactly i = 10 flips until I succeed?"

Geometric Distribution

"I flip an unfair coin (p = 0.4) until it lands on heads. What is the probability that I need to wait exactly i = 10 flips until I succeed?"

$$X \sim \mathrm{Geo}(p)$$

$$f_X(i) = \begin{cases} p(1-p)^{i-1} & \text{für } i \in \mathbb{N} \\ 0 & \text{sonst.} \end{cases}$$

$$\mathbb{E}[X] = \frac{1}{p} \quad \text{und} \quad \mathrm{Var}[X] = \frac{1-p}{p^2}$$

$$F_X(n) = \Pr[X \leq n] = \sum_{i=1}^{n} \Pr[X = i] = \sum_{i=1}^{n} p(1-p)^{i-1} = 1 - (1-p)^n$$

Geometric Distribution

Gedächtnislosigkeit.

Satz 2.45. Ist $X \sim \mathrm{Geo}(p)$, so gilt für alle $s, t \in \mathbb{N}$:

$$\Pr[X \geq s + t \mid X > s] = \Pr[X \geq t].$$

$$\Pr[X \geq s+t \mid X > s] = \frac{\Pr[X \geq s + t]}{\Pr[X > s]} = \frac{(1-p)^{s+t-1}}{(1-p)^{s}} = (1-p)^{t-1} = \Pr[X \geq t],$$

Geometric Distribution

Gedächtnislosigkeit.

Satz 2.45. Ist $X \sim \mathrm{Geo}(p)$, so gilt für alle $s, t \in \mathbb{N}$:

$$\Pr[X \geq s + t \mid X > s] = \Pr[X \geq t].$$

Beweis. Wir wissen bereits, dass für die Verteilungsfunktion gilt: $F_X(n) = 1-(1-p)^n$ für alle $n \in \mathbb{N}$. Somit gilt $\Pr[X \geq n] = (1-p)^{n-1}$ für alle $n \in \mathbb{N}$. Verwenden wir nun die Definition der bedingten Wahrscheinlichkeit, so erhalten wir daher

$$\Pr[X \geq s+t \mid X > s] = \frac{\Pr[X \geq s+t]}{\Pr[X > s]} = \frac{(1-p)^{s+t-1}}{(1-p)^s} = (1-p)^{t-1} = \Pr[X \geq t],$$

wie behauptet. $\square$

Negative Binomial Distribution

The negative binomial distribution describes the probability of needing **exactly z independent trials** (each with success probability p) to achieve **exactly n successes**.

$$X \sim \mathrm{NB}(n, p)$$

"I flip an unfair coin (p = 0.4) until it has landed on heads n = 10 times. What is the probability I need to wait exactly z = 20 flips until I succeed?"

The **Dichte** $f_X(z)$ is defined for the number of trials $z \in \{n, n+1, n+2, \dots\}$:

$$f_X(z) = \Pr[X = z] = \binom{z-1}{n-1} p^n (1-p)^{z-n}$$

Intuition: To have the n-th success at trial z, there must have been exactly $n-1$ successes in the previous $z-1$ trials.

Negative Binomial Distribution

Key Properties

The properties scale with the required number of successes n:

- **Expected Value:** $\mathbb{E}[X] = \frac{n}{p}$
- **Variance:** $\mathrm{Var}[X] = \frac{n(1-p)}{p^2}$

---

*Note: If $n = 1$, the distribution simplifies to the **Geometric Distribution** $X \sim Geo(p)$, representing the wait time for the first success.*

Distributions

Important Distributions

Name	Notation	Support	Density	Expectation	Variance
Bernoulli	Bernoulli(p)	$\{0,1\}$	$f_X(i) = \begin{cases} p & \text{for } i = 1, \\ 1-p & \text{for } i = 0. \end{cases}$	p	$p(1-p)$
Binomial	Bin(n,p)	$\{0,1,\ldots,n\}$	$f_X(i) = \binom{n}{i} p^i (1-p)^{n-i}$	np	$np(1-p)$
Geometric	Geo(p)	$\{1,2,3,\ldots\}$	$f_X(i) = p(1-p)^{i-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson	Po(λ)	$\{0,1,2,\ldots\}$	$f_X(i) = \frac{e^{-\lambda}\lambda^i}{i!}$	λ	λ

Multiple Random Variables

(condensed)

Multiple Random Variables

Satz 2.58. Für zwei unabhängige Zufallsvariablen X und Y sei $Z := X + Y$. Es gilt

$$f_Z(z) = \sum_{x \in W_X} f_X(x) \cdot f_Y(z - x).$$

Multiple Random Variables

Satz 2.58. Für zwei unabhängige Zufallsvariablen X und Y sei $Z := X + Y$. Es gilt

$$f_Z(z) = \sum_{x \in W_X} f_X(x) \cdot f_Y(z - x).$$

Beweis. Mit Hilfe des Satzes von der totalen Wahrscheinlichkeit folgt, dass

$$\begin{aligned} f_Z(z) &= \Pr[Z = z] = \sum_{x \in W_X} \Pr[X + Y = z \mid X = x] \cdot \Pr[X = x] \\ &= \sum_{x \in W_X} \Pr[Y = z - x] \cdot \Pr[X = x] = \sum_{x \in W_X} f_X(x) \cdot f_Y(z - x). \end{aligned}$$

$\square$

Multiple Random Variables

Satz 2.60. *(Linearität des Erwartungswerts)* Für Zufallsvariablen $X_1, \dots, X_n$ und $X := a_1X_1 + \cdots + a_nX_n$ mit $a_1, \dots, a_n \in \mathbb{R}$ gilt

$$\mathbb{E}[X] = a_1\mathbb{E}[X_1] + \cdots + a_n\mathbb{E}[X_n].$$

Satz 2.61. *(Multiplikativität des Erwartungswerts)* Für unabhängige Zufallsvariablen $X_1, \dots, X_n$ gilt

$$\mathbb{E}[X_1 \cdot \dots \cdot X_n] = \mathbb{E}[X_1] \cdot \dots \cdot \mathbb{E}[X_n].$$

Satz 2.62. Für unabhängige Zufallsvariablen $X_1, \dots, X_n$ und $X := X_1 + \ldots + X_n$ gilt

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n].$$

Multiple Random Variables

Satz 2.65 (Waldsche Identität). N und X seien zwei unabhängige Zufallsvariable, wobei für den Wertebereich von N gilt: $W_N \subseteq \mathbb{N}$. Weiter sei

$$Z := \sum_{i=1}^{N} X_i,$$

wobei $X_1, X_2, \ldots$ unabhängige Kopien von X seien. Dann gilt:

$$\mathbb{E}[Z] = \mathbb{E}[N] \cdot \mathbb{E}[X].$$

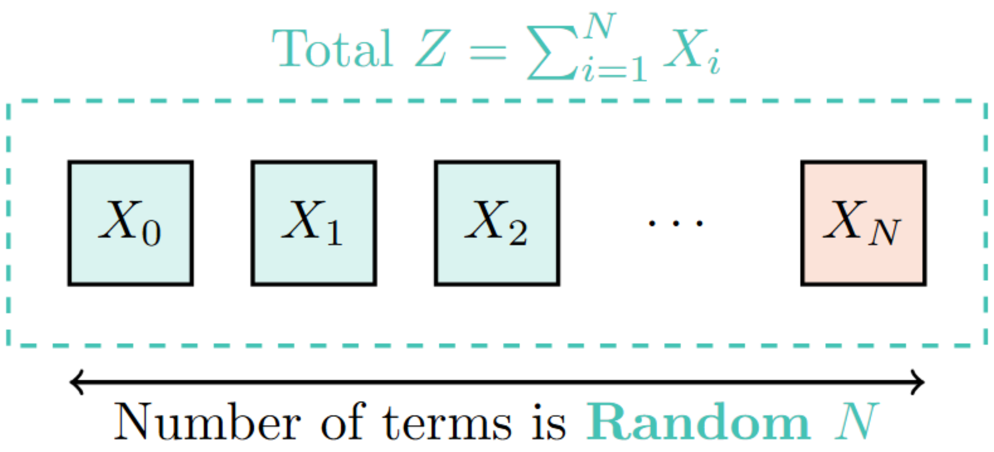
Multiple Random Variables

If N and X are **independent**, and Z is the sum of N independent copies of X, then:

$$\mathbb{E}[Z] = \mathbb{E}[N] \cdot \mathbb{E}[X]$$

Example

- **N (The Stopping Time):** A die roll deciding the number of bags of gummy bears we open. $\mathbb{E}[N] = 3.5$.
- **X (The Individual Value):** Red gummy bears per bag. Let's say $\mathbb{E}[X] = 10$.
- **Z (The Total):** The sum $X_1 + X_2 + \cdots + X_N$.



$$\mathbb{E}[Z] = \mathbb{E}[N] \cdot \mathbb{E}[X] = 3.5 \cdot 10 = 35$$

Approximating Probabilities – Inequalities

Inequalities

Satz 2.67. *(Ungleichung von Markov)* Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

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

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

Inequalities

Satz 2.67. *(Ungleichung von Markov)* Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

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

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

Beweis. Wir rechnen direkt nach, dass

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x \in W_X} x \cdot \Pr[X = x] \geq \sum_{x \in W_X,\ x \geq t} x \cdot \Pr[X = x] \\ &\geq t \cdot \sum_{x \in W_X,\ x \geq t} \Pr[X = x] = t \cdot \Pr[X \geq t]. \end{aligned}$$

Inequalities

Satz 2.68. *(Ungleichung von Chebyshev)* Sei X eine Zufallsvariable und $t \in \mathbb{R}$ mit $t > 0$. Dann gilt

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

oder äquivalent dazu $\Pr[|X - \mathbb{E}[X]| \geq t\sqrt{\mathrm{Var}[X]}] \ \leq \ 1/t^2$.

Inequalities

Satz 2.68. *(Ungleichung von Chebyshev)* Sei X eine Zufallsvariable und $t \in \mathbb{R}$ mit $t > 0$. Dann gilt

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

oder äquivalent dazu $\Pr[|X - \mathbb{E}[X]| \geq t\sqrt{\mathrm{Var}[X]}] \;\leq\; 1/t^2$.

Beweis. Es gilt

$$\Pr[|X - \mathbb{E}[X]| \geq t] \;=\; \Pr[(X - \mathbb{E}[X])^2 \geq t^2].$$

Die Zufallsvariable $Y := (X - \mathbb{E}[X])^2$ ist nicht-negativ und hat nach Definition der Varianz den Erwartungswert $\mathbb{E}[Y] = \mathrm{Var}[X]$. Damit folgt die Behauptung durch Anwendung der Markov-Ungleichung:

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

Inequalities

Satz 2.70 (Chernoff-Schranken). Seien $X_1, \ldots, X_n$ unabhängige Bernoulli-verteilte Zufallsvariablen mit $\Pr[X_i = 1] = p_i$ and $\Pr[X_i = 0] = 1 - p_i$. Dann gilt für $X := \sum_{i=1}^{n} X_i$:

(i) $\Pr[X \geq (1+\delta)\mathbb{E}[X]] \leq e^{-\frac{1}{3}\delta^2 \mathbb{E}[X]}$ für alle $0 < \delta \leq 1$,

(ii) $\Pr[X \leq (1-\delta)\mathbb{E}[X]] \leq e^{-\frac{1}{2}\delta^2 \mathbb{E}[X]}$ für alle $0 < \delta \leq 1$,

(iii) $\Pr[X \geq t] \leq 2^{-t}$ für $t \geq 2e\mathbb{E}[X]$.

Inequalities

Which inequality gives tighter bounds? Depends on the situation:

Beispiel 2.71. Wir betrachten wieder das Beispiel von vorher: Wir werfen n faire Münzen und bezeichnen mit X die Anzahl von „Kopf“-Ergebnissen. Dann gilt $\mathbb{E}[X] = n/2$. Wie gross ist nun die Wahrscheinlichkeit $\Pr[|X - n/2| > 0.1 \cdot n/2]$, dass die tatsächliche Anzahl um mindestens 10% vom Erwartungswert abweicht? Die folgende Tabelle vergleicht die Schranken, und die Chebyshev-Ungleichung und die Chernoff-Ungleichung liefern: Für

n	Chebyshev	Chernoff
1000	0.1	0.270961
2000	0.05	0.0424119
5000	0.02	0.000244096
10000	0.01	$5.77914 \cdot 10^{-8}$
100000	0.001	$4.14559 \cdot 10^{-73}$

Exercise S8

Inequalities

Exercise S8.1 – *Inequalities*

We throw 1000 fair coins and denote the results by $C_1, \ldots, C_{1000}$. We want to count the number of neighbouring coins that both show "heads". Two neighbouring coins are coins of the form C_i, C_{i+1} or C_{1000}, C_1 (imagine the coins being placed on a circle). Let X denote the number of neighbouring coins that both show "heads". We will use different inequalities to bound the probability that X is significantly larger than its expected value. Make sure to check all relevant conditions before applying an inequality.

(a) Show that $\mathbb{E}[X] = 250$.

(b) Use Markov's inequality to bound $\Pr[X \geq 300]$.

(c) Compute $\mathrm{Var}[X]$ and use Chebychev's inequality to bound $\Pr[X \geq 300]$.

(d) We define Y as the number of neighbouring coins that both show "heads" and for which the first coin has an odd index. That is, we only consider pairs of the form X_{2i-1}, X_{2i}, with $i = 1, \ldots, 500$. Show that $\mathbb{E}[Y] = 125$ and use Chernoff's bound for $\Pr[Y \geq 150]$.

(e) Use (d) to bound $\Pr[X \geq 300]$.
Remark: Could we just apply Chernoff's bound right away?

Randomised Algorithms

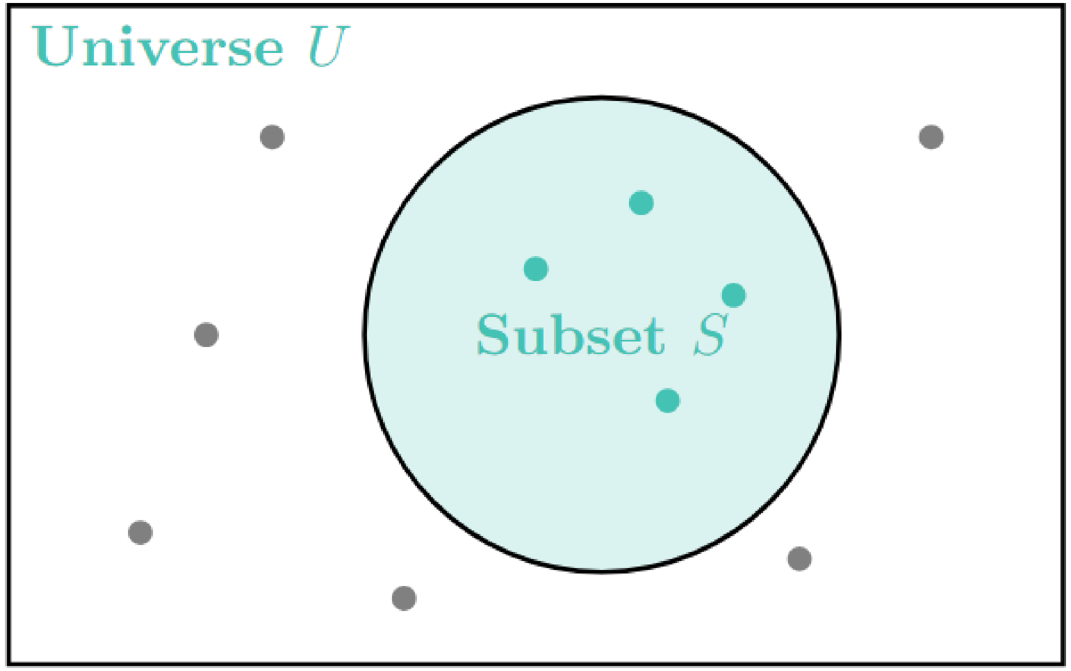
Target-Shooting

Preview

Target-Shooting

Example: How many people have a certain virus (i.e. are in S) ?

The goal is to estimate the ratio $\frac{|S|}{|U|}$ by sampling N points. Each point u_i is a "shot." If it hits the subset S, the indicator $I_S(u_i) = 1$; otherwise, it is 0.



Algorithm Output:

$$X = \frac{\text{Hits in } S}{\text{Total Shots } N}$$

Target-Shooting

TARGET-SHOOTING

1: Wähle $u_1, \ldots, u_N \in U$ zufällig, gleichverteilt und unabhängig

2: **return** $N^{-1} \cdot \sum_{i=1}^{N} I_S(u_i)$

Target-Shooting

Wir wollen den Target-Shooting-Algorithmus genauer analysieren. Sei dazu für $i \in [N]$

$$Y_i := I_S(u_i).$$

Nun sind wegen der unabhängigen und uniformen Wahl der u_i die Variablen $Y_1, \ldots, Y_N$ unabhängige Bernoulli-Variablen mit $\Pr[Y_i = 1] = |S|/|U|$ für jedes $i = 1, \ldots, N$. Für die Zufallsvariable

$$Y := \frac{1}{N} \sum_{i=1}^{N} Y_i = \frac{1}{N} \sum_{i=1}^{N} I_S(u_i)$$

erhalten wir demnach $\mathbb{E}[Y] = |S|/|U|$, unabhängig von der Wahl von N.