Wahrscheinlichkeitstheorie und Statistik

3. Woche

Dávid Tóth (BME SZIT)

16., 19. September 2024

Diskrete Zufallsvariablen

Der Begriff des Wahrscheinlichkeitsraums ist nicht ausreichend zu unseren Zielen.

In den Anwerdungen arbeitet man typisch mit zufälligen Quantitäten. Z.B.:

- das Ergebnis eines Wurfes,
- das Lebensalter einer zufällig ausgewählten Person,
- die Anzahl der fehlerbehafteten Produkte bei einer Stichprobe,
- der Fehler einer Messung,
- die Anzahl der Bits, die bei einem Informationsaustausch durch einen Kanal falsch übertragen werden, usw.

Diskrete Zufallsvariablen

Die obigen Quantitäten könnten als die Elemente einer Ergebnismenge betrachtet werden, aber es ist im Allgemeinen nicht praktisch.

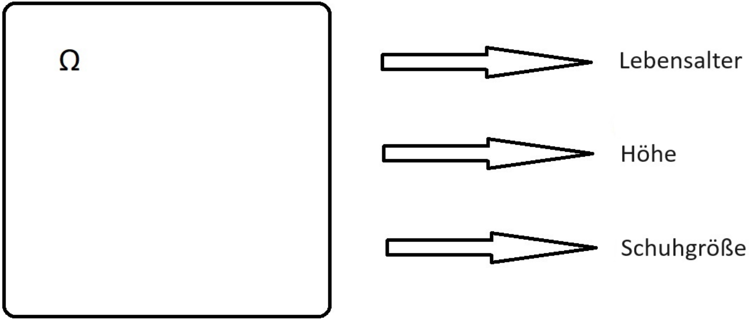
Beispiel: das Lebensalter einer zufällig ausgewählten Person.

- Hier wird keine Zahl, sondern eine Person aus einer Gruppe zufällig gewählt.
- Das Lebensalter ist ein *Merkmal* dieser Person.
- Neben dem Lebensalter können auch andere Merkmale interessant sein. Z.B.: die Höhe, die Schuhgröße, die Anzahl ihrer Kinder, usw.
- Vielleicht möchten wir die möglichen Zusammenhänge zwischen dieser Quantitäten analysieren.

Diskrete Zufallsvariablen

- Die obigen Attribute gehören zu einer Person, es wäre nicht praktisch, diese getrennt, in verschiedenen Ergebnismengen zu behandeln.
- Stattdessen ordnen wir sie der ausgewählten Person zu.
- Im Allgemeinen ordnen wir verschiede Werte oder Eigenschaften den Elementen einer Ergebnismenge zu, die danach zusammen behandelt werden können.
- Dieses Ziel wird durch sogenannte *Zufallsvariablen* erreicht, die auf der Ergebnismenge definierte Zuornungen, also Funktionen sind.

Diskrete Zufallsvariablen



Diskrete Zufallsvariablen

Zunächst betrachten wir Wahrscheinlichkeitsräume $(\Omega, \mathcal{F}, \mathbb{P})$ mit endlicher oder abzählbar unendlicher Ergebnismenge Ω, und wir nehmen auch an, dass $\mathcal{F} = \mathcal{P}(\Omega)$ gilt.

(Also betrachten wir zuerst diskerte Wahrscheinlichkeitsräume.)

Definition. Eine (diskrete) *Zufallsvariable* (valószínűségi változó) ist eine Funktion $X : \Omega \to \mathbb{R}$ mit Definitionsmenge Ω und Zielmenge $\mathbb{R}$.

Sie heißt *diskret*, weil ihr Bild in $\mathbb{R}$ abzählbar ist. Wenn das Bild endlich ist, dann nennt man die Zufallsvariable *einfach*.

Diskrete Zufallsvariablen

Beispiele.

- Ein Würfel wird geworfen. Sei X die Funktion, die jedem Ergebnis einfach den Wert des Ergebnisses zuordnet.
- Seien Ω und X wie oben. Dann ist die Funktion $Y = X^2$ eine Zufallsvariable mit Bildmenge $\{1, 4, 9, 16, 25, 36\}$.
- Eine Münze wird zweimal geworfen. Sei W die Anzahl der "Köpfe", dann ist W eine Zufallsvariable. Hier ist $\Omega = \{KK, KZ, ZK, ZZ\}$, weiter $W(KK) = 2$, $W(KZ) = W(ZK) = 1$, $W(ZZ) = 0$.
- Sei $A \subset \Omega$ ein Ereignis. Die *Indikatorvariable* von A ist eine Zufallsvariable $\mathbb{1}_A : \Omega \to \{0, 1\}$, die wie folgt definiert wird:

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \text{falls } \omega \in A, \\ 0, & \text{falls } \omega \notin A. \end{cases}$$

Diskrete Zufallsvariablen

Beispiele.

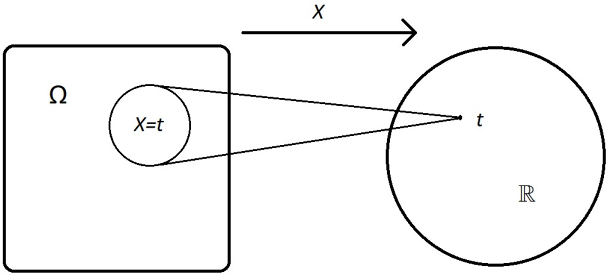
- Ein Experiment wird n-mal nacheinander unabhängig durchgeführt. Sei U die Anzahl der Eintritte einer bestimmten Ereignis. Z.B.: wir werfen eine Münze n-mal und zählen die Köpfe ab; wir werfen n-mal einen Würfel und betrachten die Anzalh der geraden Augenzahlen. Also ist diese Beispiel die Verallgemeinerung der Variablen W.
- Ein Experiment wird so lange (unabhängig) wiederholt, bis ein bestimmtes Ereignis eintritt. Sei V die Anzahl der Experimente bis zum "Erfolg".
- Außer der Variablen V sind die obigen Beispiele *einfache* Zufallsvariablen, weiter gilt $\operatorname{ran} V = \mathbb{N}^+$.

Diskrete Zufallsvariablen

- Das Verhalten einer Zufallsvariablen kann vermittels ihrer möglichen Werten beschrieben werden.
- Z.B.: Mit welcher Wahrscheinlichkeit wird der Wert eines Wurfes größer als 3? Mit welcher Wahrscheinlichkeit werden höchstens 10 Bits falsch durch einen Kanal übertragen?
- Damit man solche Fragen beanworten kann, es ist bei *diskreten Variablen* genügend, die Wahrscheinlichkeit der Annahme der einzelnen Werte zu kennen.

Diskrete Zufallsvariablen

Definition. Sei $X : \Omega \to \mathbb{R}$ eine diskrete Zufallsvariable, $t \in \mathbb{R}$, dann besteht die Menge $\{X = t\} \subset \Omega$ aus solchen Elementen ω von Ω, für die $X(\omega) = t$ gilt.



Beispiel. Eine Münze wird zweimal geworfen, sei W die Anzahl der Köpfe. Dann gilt

$$\{W = 1\} = \{KZ, ZK\}.$$

Diskrete Zufallsvariablen

Definition. Sei $X : \Omega \to \mathbb{R}$ eine diskrete Zufallsvariable, $t \in \mathbb{R}$, dann besteht die Menge $\{X \leq t\}$ aus solchen Elementen $\omega \in \Omega$, für die $X(\omega) \leq t$ gilt.

Ähnlich werden die Teilmengen $\{X < t\}$, $\{X \geq t\}$, $\{X > t\}$, $\{s < X \leq t\}$, usw. definiert.

Im Allgeminen, falls $H \subset \mathbb{R}$ eine beliebige Zahlmenge ist, dann bezeichnet $X^{-1}(H) = \{X \in H\}$ das *Urbild* von H unter X, d.h. die Menge der Elementen $\omega \in \Omega$ mit $X(\omega) \in H$.

Beispiel. Eine Münze wird zweimal geworfen, sei W die Anzahl der Köpfe. Dann gilt

$$\{W < 2\} = \{W = 0\} \cup \{W = 1\} = \{ZZ, KZ, ZK\}.$$

Diskrete Zufallsvariablen

- Falls auch ein auf Ω definiertes Wahrscheilichkeitsmaß angegeben ist, dann können wir auch Wahrscheinlichkeiten wie $\mathbb{P}(X = 1) := \mathbb{P}(\{X = 1\})$ oder $\mathbb{P}(X < 2) := \mathbb{P}(\{X < 2\})$ betrachten.
- Um die Notation einfach zu halten, wir lassen die geschweiften Klammern "$\{$" und "$\}$" oft weg.

Diskrete Zufallsvariablen

- Der Wert einer Variablen X ist für ein Element $\omega \in \Omega$ eindeutig bestimmt, also werden die Ereignisse $\{X = s\}$ und $\{X = t\}$ für $s, t \in \mathbb{R}$, $s \neq t$ immer disjunkt.
- Folglich gilt

$$\mathbb{P}(X = s \text{ oder } X = t) = \mathbb{P}(X \in \{s, t\}) = \mathbb{P}(X = s) + \mathbb{P}(X = t).$$

$\Rightarrow$ Das Verhalten von (einer diskreten Variablen) X wird durch die Wahrscheinlichkeiten $\mathbb{P}(X = t)$ beschrieben, wo t die Wertemenge von X durchläuft.

- Nämlich gilt

$$\mathbb{P}(X \in H) = \sum_{t \in H \cap \operatorname{ran} X} \mathbb{P}(X = t)$$

für eine beliebige Menge $H \subset \mathbb{R}$.

Diskrete Zufallsvariablen

Beispiel. Sei X die Augenzahl bei einem Wurf mit einem Würfel. Dann gelten

$$\begin{aligned}\mathbb{P}(X=1) = \mathbb{P}(X=2) = \mathbb{P}(X=3) = \\ = \mathbb{P}(X=4) = \mathbb{P}(X=5) = \mathbb{P}(X=6) = \frac{1}{6},\end{aligned}$$

weiter (z.B.)

$$\begin{aligned}&\mathbb{P}(\text{die Augenzahl ist gerade}) = \\ &\quad = \mathbb{P}(X=2 \text{ oder } X=4 \text{ oder } X=6) \\ &\quad = \mathbb{P}(X=2)+\mathbb{P}(X=4)+\mathbb{P}(X=6) = \frac{3}{6} = \frac{1}{2}.\end{aligned}$$

Diskrete Zufallsvariablen

Definition. Sei $X : \Omega \to \mathbb{R}$ eine diskrete Zufallsvariable, und sei $f_X(t) := \mathbb{P}(X = t)$ für alle $t \in \operatorname{ran} X$. Die so definierte Funktion $f_X : \operatorname{ran} X \to [0; 1]$ heißt *die Wahrscheinlichkeitsfunktion* von X (súlyfüggvény).

Also bestimmt die Wahrscheinlichkeitsfunktion von X die Wahrscheinlichkeiten solcher Ereignisse, die durch die Werte von X ausgedrückt werden können.

Nämlich, für eine beliebige Menge $H \subset \mathbb{R}$ gilt

$$\mathbb{P}(X \in H) = \sum_{t \in H \cap \operatorname{ran} X} f_X(t).$$

Diskrete Zufallsvariablen

- Wir sagen, dass die Wahrscheinlichkeitsfunktion von X *die Verteilung* (eloszlás) von X bestimmt.
- Um die Funktion f_X anzugeben, muss man sowohl die Wertemenge von X (also die Definitionsmenge von f_X), als auch alle obigen Wahrscheinlichkeiten (also die Werte von f_X) bestimmen.
- Später werden wir sehen, dass die Verteilung von X auch anders angegeben werden kann, also ist die Wahrscheinlichkeitsfunktion nur ein mögliches Mittel.

Diskrete Zufallsvariablen

Beispiel. Sei Y das Quadrat der Augenzahl bei einem Wurf mit einem Würfel. Dann kann Y die Werte 1, 4, 9, 16, 25 és 36 annehmen, jeden Wert mit Wahtscheinlichkeit $\frac{1}{6}$, also wird die Wahrscheinlichkeitsfunktion durch

$$f_Y(1)\big(=\mathbb{P}(Y=1)=\big)f_Y(4)=f_Y(9)=$$
$$=f_Y(16)=f_Y(25)=f_Y(36)=\frac{1}{6}$$

angegeben.

Diskrete Zufallsvariablen

Definition. Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ heißt *gleichverteilt* (egyenletes eloszlású) oder *uniformverteilt*, oder *Laplace-verteilt* auf einer endlichen Menge $\{y_1, \ldots, y_n\} \subset \mathbb{R}$, falls

$$f_X(y_i) = \mathbb{P}(X = y_i) = \frac{1}{n}$$

für alle $i = 1, \ldots, n$ gilt.

Diskrete Zufallsvariablen

Beispiel. Sei A ein Ereignis mit $\mathbb{P}(A) = p$, und sei $\mathbb{1}_A$ die Indikatorvariable von A. Dann ist $\operatorname{ran} \mathbb{1}_A = \{0, 1\}$ und die Verteilung der Variablen wird durch

$$
\begin{aligned}
f_{\mathbb{1}_A}(1) &= \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A) = p, \\
f_{\mathbb{1}_A}(0) &= \mathbb{P}(\mathbb{1}_A = 0) = \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A) = 1 - p
\end{aligned}
$$

angegeben.

Binomialverteilung

Beispiel. Ein Würfel wird fünfmal geworfen, sei X die Anzahl der Sechsen. Was ist der Wert der Wahrscheinlichkeitsfunktion von X an der Stelle 2? D.h.: was ist die Wahrscheinlichkeit $f_X(2) = \mathbb{P}(X = 2)$?

Bei jedem Wurf: mit Wahrscheinlichkeit $\frac{1}{6}$ erhält man eine Sechs, mit Wahrscheinlichkeit $\frac{5}{6}$ eine andere Augenzahl.
$\Rightarrow$ Eine Serie von 5 Wurfen, wobei gerade zwei (fixe) Wurfe Sechsen geben, wird mit Wahrscheinlichkeit

$$\left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3$$

erhalten (denn die Wurfe sind unabhängig).

Die Stelle der zwei Sechsen in der Serie kann auf $\binom{5}{2} = 10$ verschiedene Weisen ausgewählt werden, also gilt

$$f_X(2) = \mathbb{P}(X = 2) = 10 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 = \frac{625}{3888} \approx 0{,}1608.$$

Binomialverteilung

Wir verallgemeinern das obige Beispiel:

- Ein Experiment wird n-mal nacheinander unabhängig durchgeführt. Sei X die Anzahl der Eintritte eines fixen Ereignisses mit Wahrscheinlichkeit p während der Experimente.
- $\operatorname{ran} X = \{0, 1, \ldots, n\}$
- Das Ereignis $\{X = k\}$ besteht aus solchen möglichen Serien von Experimenten, wobei das betrachtete Ereignis gerade k Mal eintritt.
- Folglich tritt dieses Ereignis in die andren $n - k$ Fällen nicht ein, d.h.: in diesen Fällen tritt sein Komplement ein.
- Die Experimente sind unabhängig, also erhält man eine solche Serie mit Wahrscheinlichkeit $p^k(1-p)^{n-k}$.

Binomialverteilung

Wir verallgemeinern das obige Beispiel:

- Die Experimente, wobei das betrachtete Ereignis eintritt, werden "erfolgreich" oder "Erfolge" genannt. Die andere Ereignisse werden Misserfolge genannt.
- Die Stellen der k Erfolge kann auf $\binom{n}{k}$ verschiedene Weisen ausgewählt werden.
- Also zerfällt das Ereignis $\{X = k\}$ in $\binom{n}{k}$ paarweise dijunkte Ereignisse mit gleicher Wahrscheinlichkeit $p^k(1-p)^{n-k}$, folglich gilt

$$f_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (0 \leq k \leq n).$$

Binomialverteilung

Definition. Eine diskrete Zufallsvariable X heißt *binomialverteilt* (binomiális elolszlású) mit Parametern $n \in \mathbb{N}^+$ und $p \in [0;1]$, falls $\operatorname{ran} X = \{0, 1, \ldots, n\}$ und

$$f_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

für jede $0 \leq k \leq n$ gilt. Bezeichnung: $X \sim \mathrm{Bin}(n; p)$.

Im obigen Beispiel ist die Variable X binomialverteilt, also kann die Binomialverteilung die Anzahl der Erfolge in einer Serie von n gleichartigen und unabhängigen Versuchen (Bernoulli-Experimente - Bernoulli kísérletek) beschreiben, die jeweils genau zwei mögliche Ergebnisse haben:

- "Erfolg" mit Wahrscheinlichkeit p,
- oder "Misserfolg" mit Wahrscheinlichkeit $1 - p$.

Binomialverteilung

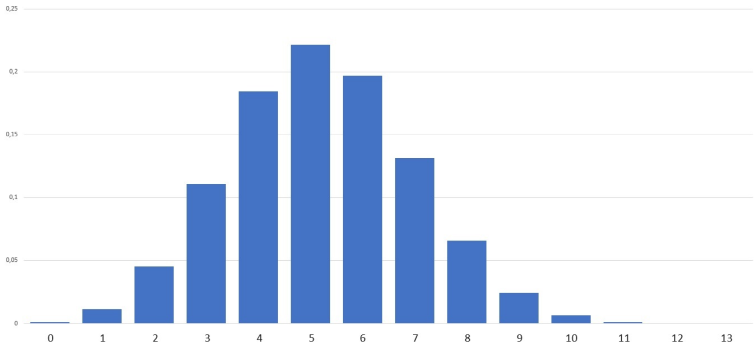
Beispiel. Für eine binomialverteilte Variable $X \sim \mathrm{Bin}(n; p)$ gilt immer

$$\begin{aligned} 1 = \mathbb{P}(\Omega) = \mathbb{P}(0 \leq X \leq n) &= \sum_{k=0}^{n} \mathbb{P}(X = k) \\ &= \sum_{k=0}^{n} f_X(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}. \end{aligned}$$

Diese Gleichung folgt auch aus dem binomischen Lehrsatz:

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = [p + (1-p)]^n = 1^n = 1.$$

Die Werte von f_X für $X \sim \mathrm{Bin}(13; 0{,}4)$



Binomialverteilung

Falls $X \sim \mathrm{Bin}(13; 0{,}4)$, dann

- zeigt die k-te Spalte im vorigen Diagramm den Wert $f_X(k) = \mathbb{P}(X = k)$.
- Diese Wahrscheinlichkeiten wachsen anfangs, dann sinken sie später ab.
- Diese Phänomene treten für eine beliebige binomialverteilte Zufallsvariable auf:

Binomialverteilung

Satz. Sei $X \sim \mathrm{Bin}(n; p)$ eine binomialverteilte Zufallsvariable. Dann ist die Funktion f_X

- monoton steigend in $[0; \lfloor (n+1)p \rfloor]$
- und monoton fallend in $[\lfloor (n+1)p \rfloor; n]$.
- Folglich nimmt f_X ihren maximalen Wert an der Stelle $m = \lfloor (n+1)p \rfloor$ an.
- Falls $(n+1)p$ eine ganze Zahl ist, und

$$m = \lfloor (n+1)p \rfloor = (n+1)p,$$

  dann gilt $f_X(m) = f_X(m-1)$, und f_X nimmt ihren maximalen Wert gerade an diesen zwei Stellen an.
- Sonst ist die Maximalstelle von f_X eindeutig.

Geometrische Verteilung

Beispiel. Ein Experiment wird so lange unabhängig nacheinander wiederholt, bis ein fixes Ereignis mit Wahrscheinlichkeit p eintritt.
Z.B.:

- eine Münze wird bis zum ersten Kopf geworfen (hier $p = \frac{1}{2}$),
- ein Würfel wird bis zur ersten Sechs geworfen (hier $p = \frac{1}{6}$).

Sei X die Anzahl der Experimente bis zum ersten Erfolg.

Geometrische Verteilung

Für eine $k \in \mathbb{N}^+$ ist der Wert $f_X(k) = \mathbb{P}(X = k)$ die Wahrscheinlichkeit dafür, dass das betrachtete Ereignis bei den ersten $k - 1$ Experimenten nicht eintritt (d.h.: sein Komplement tritt $k - 1$ Mal ein), aber das k-te Experiment schon erfolgreich wird.

Die Experimente sind unabhängig, also

$$f_X(k) = (1 - p)^{k-1} p.$$

Geometrische Verteilung

Definition. Eine Zufallsvariable X heißt *geometrisch verteilt* (geometriai eloszlású) mit Parameter $p \in (0;1]$, falls $\operatorname{ran} X = \mathbb{N}^+$, und

$$f_X(k) = \mathbb{P}(X = k) = (1-p)^{k-1}p$$

für alle $k \in \operatorname{ran} X$ gilt. Bezeichnung: $X \sim \mathrm{Geo}(p)$.

Die Wartezeit auf den ersten "Erfolg" in einem "unendlich oft wiederholten Bernoulli-Experiment" mit Erfolgswahrscheinlichkeit p ist geometrisch verteilt.

Geometrische Verteilung

Bemerkung. Wenn X geometrisch verteilt ist, dann muss $\mathbb{P}(X \in \mathbb{N}^+) = \mathbb{P}(\Omega) = 1$ gelten. Die Wahrscheinlichkeiten $(1-p)^{k-1}p$ summieren sich entsprechend zu 1, denn

$$\sum_{k=1}^{\infty}(1-p)^{k-1}p = p\sum_{k=0}^{\infty}(1-p)^k = \frac{p}{1-(1-p)} = 1.$$

Oben wurde eine geometrische Reihe summiert, daher die Bezeichnung "geometrische Verteilung".

Gedächtnislosigkeit im diskreten Fall

- Wenn eine Münze mehrmals geworfen wird, dann beeinflusst das Ergebnis der ersten paar Würfe die Chancen bei den folgenden Würfen *nicht*.
- Z.B.: falls jede von 10 nacheinanderfolgenden Würfen eine Zahl gibt, so wird beim nächsten Wurf die Wahrscheinlichkeit einer Zahl wieder $\frac{1}{2}$ (wie am Anfang).
- Im Allgemeinen: wenn ein Experiment bis zum ersten Erfolg unabhängig wiederholt wird, und das Experiment k Mal schon erfolglos war, das beeinflusst nicht, wie lange wir danach das Experiment bis zum Erfolg wiederholen müssen.

Gedächtnislosigkeit im diskreten Fall

Definition. Ein Zufallsvariable heißt *gedächtnislos* (örökifjú) auf der Menge $G \subset \mathbb{R}$, falls

$$\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t)$$

für jede $s, t \in G$ gilt und $\mathbb{P}(X \in G) = 1$.

Behauptung. Sei $X \sim \mathrm{Geo}(p)$, dann ist X gedächtnislos auf der Menge $\mathbb{N}^+$.

Die obige Behauptung kann folgenderweise interpretiert werden: die Wahrscheinlichkeit dafür, dass man mehr als t Experimenten bis zum ersten Erfolg machen muss, nachdem man zuvor s Experimenten erfolglos gemacht hat, ist für beliebige s gleich.

Gedächtnislosigkeit im diskreten Fall

Beweis. Seien $s, t \in \mathbb{N}^+$, dann gilt

$$\mathbb{P}(X > s + t \mid X > s) = \frac{\mathbb{P}(X > s + t \text{ und } X > s)}{\mathbb{P}(X > s)}$$

$$= \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} = \frac{\sum_{k=s+t+1}^{\infty} \mathbb{P}(X = k)}{\sum_{l=s+1}^{\infty} \mathbb{P}(X = l)}$$

$$= \frac{\sum_{k=s+t+1}^{\infty} f_X(k)}{\sum_{l=s+1}^{\infty} f_X(l)}.$$

Diese Formel zeigt, dass die obige Wahrscheinlichkeit nur von der Verteilung, und nicht von der Variablen abhängt.
$\Longrightarrow$ Es genügt, den Wert für eine beliebige geometrisch verteilte Variable zu berechnen.

Gedächtnislosigkeit im diskreten Fall

Sei X die Wartezeit auf den ersten Erfolg in einer Serie von wiederholten Experimenten mit Erfolgswahrscheinlichkeit p.

Dann ist X geometrisch verteilt mit Parameter p und $\{X > s\}$ ist das Ereignis, dass die ersten s Experimente erfolglos sind, also

$$\begin{aligned}\mathbb{P}(X > s + t \mid X > s) &= \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} = \frac{(1-p)^{s+t}}{(1-p)^s} \\ &= (1-p)^t = \mathbb{P}(X > t).\end{aligned}$$

Gedächtnislosigkeit im diskreten Fall

Die umgekehrte Behauptung gilt auch:

Satz. Sei X eine nicht konstante Zufallsvariable, die auf der Menge $\mathbb{N}^+$ gedächtnislos ist. Dann ist X geometrisch verteilt.

Diskrete Zufallsvariablen

Beispiele. Wenn $X, Y : \Omega \to \mathbb{R}$ diskrete Zufallsvariablen sind, dann sind auch

- $X \pm Y : \Omega \to \mathbb{R}$, $(X \pm Y)(\omega) := X(\omega) \pm Y(\omega)$ $(\forall \omega \in \Omega)$
- $X \cdot Y : \Omega \to \mathbb{R}$, $(X \cdot Y)(\omega) := X(\omega) \cdot Y(\omega)$ $(\forall \omega \in \Omega)$

diskrete Zufallsvariablen. Wenn zusätzlich $Y \neq 0$ gilt, dann ist auch

- $\dfrac{X}{Y} : \Omega \to \mathbb{R}$, $\dfrac{X}{Y}(\omega) := \dfrac{X(\omega)}{Y(\omega)}$ $(\forall \omega \in \Omega)$

eine diskrete Zufallsvariable.

Diskrete Zufallsvariablen

Beispiel. Ein Würfel wird zweimal geworfen. Seien X das Ergebnis des esten und Y das Ergebnis des zweiten Wurfes. Dann gibt $X + Y$ die Summe der zwei Augenzahlen. Was ist die Verteilung von $X + Y$?

Die möglichen Werte der Summe sind die ganzen Zahlen zwischen 2 und 12.

Der Wert 2 (bzw. 12) ergibt sich nur dann, wenn zwei Einsen (bzw. zwei Sechsen) geworfen werden. Die Wahrscheinlichkeit dieser Ereignisse ist $\frac{1}{36}$, also

$$f_{X+Y}(2) = \mathbb{P}(X + Y = 2) = \frac{1}{36} = f_{X+Y}(12) = \mathbb{P}(X + Y = 12).$$

Diskrete Zufallsvariablen

Den Wert 3 geben auch zwei verschiedene Summen:
$3 = 1 + 2 = 2 + 1$ (die Reiheinfolge ist hier betrachtet).

Das gilt auch für 11: $11 = 5 + 6 = 6 + 5$, und wir erhalten

$$f_{X+Y}(3) = f_{X+Y}(11) = \frac{2}{36} = \frac{1}{18}.$$

Diskrete Zufallsvariablen

In ähnlicher Weise erhält man

$$f_{X+Y}(4) = f_{X+Y}(10) = \frac{3}{36} = \frac{1}{12},$$

$$f_{X+Y}(5) = f_{X+Y}(9) = \frac{4}{36} = \frac{1}{9},$$

$$f_{X+Y}(6) = f_{X+Y}(8) = \frac{5}{36},$$

$$f_{X+Y}(7) = \frac{6}{36} = \frac{1}{6}.$$

Diskrete Zufallsvariablen

Beispiel. Sei X die Anzahl der Erfolge in einer Serie von n unabhängig wiederholten Experimenten mit Erfolgswahrscheinlichkeit p. Dann gilt $X \sim \mathrm{Bin}(n; p)$.

Sei A_i das Ereignis, dass das i-te Experiment erfolgreich ist. Dann kann X als die Summe der Intdikatoren der Ereignisse A_i ausgedrückt werden:

$$X = \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n},$$

weil $X = k$ genau dann gilt, falls gerade k von den Ereignissen A_i eintritt, also falls die zugehörigen k Indikatorvariablen den Wert 1 und die anderen den Wert 0 annehmen.

Diskrete Zufallsvariablen

Beispiel. Im Allgemeinen, falls $A_1, \ldots, A_n \subset \Omega$ gemeinsam unabhängige Ereignisse mit Wahrscheinlichkeit $\mathbb{P}(A_i) = p$ sind, dann kann man leicht zeigen, dass die Summe

$$\mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n}$$

gibt die Anzahl der eintgetretenen Ereignisse A_i, und folglich ist sie binomialverteilt mit Parametern n und p.

Erwartungswert

Damit man eine Verteilung beschreibt, es ist oft genügend, ihre verschiedenen Parameter zu bestimmen. Zunächst betrachten wir den Erwartungswert, der den durchschnittlichen Wert einer Zufallsvariablen kennzeichnet.

Erinnerung. Eine Zufallsvariable ist *einfach*, wenn sie nur endlich viele Werte annimmt.

Definition. Sei X eine einfache Zufallsvariable mit $\operatorname{ran} X = \{x_1, \ldots, x_n\}$. Der *Erwartungswert* von X (az X várható értéke) ist die Zahl

$$\mathbb{E}X := \sum_{k=1}^{n} x_k \cdot f_X(x_k) = \sum_{k=1}^{n} x_k \cdot \mathbb{P}(X = x_k).$$

Der Erwartungswert ist also eine gewichtete Summe der Werte der Zufallsvariablen, in der jeder Wert mit seiner Wahrscheinlichkeit gewichtet wird. Je “wahrscheinlicher” ein Wert ist, desto stärker wird er berücksichtigt.

Erwartungswert

Beispiel.

Ein Würfel wird geworfen, $X :=$ "Augenzahl". Dann ist $\operatorname{ran} X = \{1,2,3,4,5,6\}$, und X nimmt diese Werte mit Wahrscheinlichkeit $1/6$ an, also

$$\mathbb{E}X = \sum_{k=1}^{6} k \cdot \mathbb{P}(X = k) = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3,5.$$

Erwartungswert

Beispiel.

Sei X wie oben, und sei $Y = X^2$. Dann ist $\operatorname{ran} Y = \{1, 4, 9, 16, 25, 36\}$, und Y nimmt diese Werte mit Wahrscheinlichkeit $1/6$ an, also

$$\mathbb{E}Y = \sum_{k=1}^{6} k^2 \cdot \mathbb{P}(Y = k^2) = \frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6}.$$

Erwartungswert

Beispiel.

Sei X eine gleichverteilte Zufallsvariable mit $\operatorname{ran} X = \{x_1, \ldots, x_n\}$, dann gilt

$$\mathbb{E}X = \frac{x_1 + \cdots + x_n}{n}.$$

Erwartungswert

Beispiel.

Sei $A \subset \Omega$ ein Ereignis, und sei $\mathbb{1}_A$ die Indikatorvariable von A. Dann gilt

$$\mathbb{E}\mathbb{1}_A = 0 \cdot \mathbb{P}(\mathbb{1}_A = 0) + 1 \cdot \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A).$$

Erwartungswert

Satz. (Linearität des Erwartungswertes) Seien $X, Y : \Omega \to \mathbb{R}$ einfache Zufallsvariablen, und sei $c \in \mathbb{R}$. Dann sind $X + Y$ und cX auch einfache Zufallsvariablen, und

$$\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y, \qquad \mathbb{E}(cX) = c \cdot \mathbb{E}X$$

gelten.

Beweis. Wenn $\operatorname{ran} X = \{x_1, \ldots, x_n\}$ ist, dann erhalten wir

$$\operatorname{ran} cX = \{cx_1, \ldots, cx_n\}.$$

Für $c \neq 0$ ergibt sich

$$\mathbb{E}(cX) = \sum_{k=1}^{n} cx_k \cdot \mathbb{P}(cX = cx_k) = c \cdot \sum_{k=1}^{n} x_k \cdot \mathbb{P}(X = x_k) = c \cdot \mathbb{E}X.$$

Erwartungswert

Für $c = 0$ ist die Funktion cX die Nullfunktion, also

$$\mathbb{E}(cX) = \mathbb{E}(0) = 0 = 0 \cdot \mathbb{E}X = c \cdot \mathbb{E}X.$$

Erwartungswert

$$
\begin{aligned}
\mathbb{E}(X+Y) &= \sum_{m \in \operatorname{ran} X+Y} m \cdot \mathbb{P}(X+Y=m) \\
&= \sum_{m \in \operatorname{ran} X+Y} m \cdot \mathbb{P}(\cup_{k+l=m}\{X=k, Y=l\}) \\
&= \sum_{m \in \operatorname{ran} X+Y} \sum_{k+l=m} (k+l) \cdot \mathbb{P}(X=k, Y=l) \\
&= \sum_{k \in \operatorname{ran} X} \sum_{l \in \operatorname{ran} Y} (k+l) \cdot \mathbb{P}(X=k, Y=l).
\end{aligned}
$$

Erwartungswert

Das ist

$$\sum_{k \in \operatorname{ran} X} \sum_{l \in \operatorname{ran} Y} k \cdot \mathbb{P}(X = k, Y = l) + \\ + \sum_{k \in \operatorname{ran} X} \sum_{l \in \operatorname{ran} Y} l \cdot \mathbb{P}(X = k, Y = l)$$

$$= \sum_{k \in \operatorname{ran} X} k \sum_{l \in \operatorname{ran} Y} \mathbb{P}(X = k, Y = l) + \\ + \sum_{l \in \operatorname{ran} Y} l \sum_{k \in \operatorname{ran} X} \mathbb{P}(X = k, Y = l)$$

Erwartungswert

$$\sum_{k \in \operatorname{ran} X} k \sum_{l \in \operatorname{ran} Y} \mathbb{P}(X = k, Y = l) + \\ + \sum_{l \in \operatorname{ran} Y} l \sum_{k \in \operatorname{ran} X} \mathbb{P}(X = k, Y = l)$$

$$= \sum_{k \in \operatorname{ran} X} k \cdot \mathbb{P}(\cup_{l \in \operatorname{ran} Y} \{X = k, Y = l\}) + \\ + \sum_{l \in \operatorname{ran} Y} l \cdot \mathbb{P}(\cup_{k \in \operatorname{ran} X} \{X = k, Y = l\})$$

$$= \sum_{k \in \operatorname{ran} X} k \cdot \mathbb{P}(X = k) + \sum_{l \in \operatorname{ran} Y} l \cdot \mathbb{P}(Y = l) = \mathbb{E}X + \mathbb{E}Y.$$

Erwartungswert

Anwendung: wir bestimmen den Erwartungswert einer binomialverteilten Zufallsvariablen.

Bemerkung. Der Erwartungswert hängt nur von der Verteilung ab. Es ist also ausreichend, den Erwartungswert für eine konkrete Variable zu berechnen.

Seien $A_1, \dots, A_n$ (gemeinsam) unabhängige Ereignisse mit $\mathbb{P}(A_1) = \cdots = \mathbb{P}(A_n) = p$, und sei $\mathbb{1}_{A_i}$ die Indikatorvariable von A_i. Dann gilt

$$\mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n} \sim \mathrm{Bin}(n; p),$$

und für eine Variable $X \sim \mathrm{Bin}(n; p)$ haben wir

$$\mathbb{E}X = \mathbb{E}(\mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n}) = \sum_{i=1}^{n} \mathbb{E}(\mathbb{1}_{A_i}) = \sum_{i=1}^{n} \mathbb{P}(A_i) = \sum_{i=1}^{n} p = np.$$

Erwartungswert

Aufgabe. Berechnen wir den Erwartungswert einer binomialverteilten Variablen anhand der Definition.

Erwartungswert

Beispiel. Wir beweisen die Siebformel für 3 Ereignisse:

- Aus dem de-morganschen Gesetz folgt

$$\mathbb{P}(A \cup B \cup C) = 1 - \mathbb{P}(\overline{A \cup B \cup C}) = 1 - \mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C}).$$

- Die Indikatorvariable des Ereignisses $\overline{A} \cap \overline{B} \cap \overline{C}$ ist $\mathbb{1}_{\overline{A}} \cdot \mathbb{1}_{\overline{B}} \cdot \mathbb{1}_{\overline{C}}$, also gilt

$$\mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C}) = \mathbb{E}(\mathbb{1}_{\overline{A}} \cdot \mathbb{1}_{\overline{B}} \cdot \mathbb{1}_{\overline{C}}).$$

- Aus der Definition der Indikatorvariablen folgt

$$\begin{aligned}
\mathbb{1}_{\overline{A}} \cdot \mathbb{1}_{\overline{B}} \cdot \mathbb{1}_{\overline{C}} &= (1 - \mathbb{1}_A)(1 - \mathbb{1}_B)(1 - \mathbb{1}_C) \\
&= 1 - \mathbb{1}_A - \mathbb{1}_B - \mathbb{1}_C + \mathbb{1}_A \cdot \mathbb{1}_B + \mathbb{1}_B \cdot \mathbb{1}_C + \mathbb{1}_A \cdot \mathbb{1}_C \\
&\qquad - \mathbb{1}_A \cdot \mathbb{1}_B \cdot \mathbb{1}_C \\
&= 1 - \mathbb{1}_A - \mathbb{1}_B - \mathbb{1}_C + \mathbb{1}_{A \cap B} + \mathbb{1}_{B \cap C} + \mathbb{1}_{A \cap C} - \mathbb{1}_{A \cap B \cap C}.
\end{aligned}$$

Erwartungswert

Folglich gilt

$$
\begin{aligned}
&\mathbb{P}(A \cup B \cup C) = \\
&\qquad = \mathbb{E}(\mathbb{1}_A + \mathbb{1}_B + \mathbb{1}_C - \mathbb{1}_{A\cap B} - \mathbb{1}_{B\cap C} - \mathbb{1}_{A\cap C} + \mathbb{1}_{A\cap B\cap C}) \\
&\qquad = \mathbb{E}(\mathbb{1}_A) + \mathbb{E}(\mathbb{1}_B) + \mathbb{E}(\mathbb{1}_C) \\
&\qquad\quad - \mathbb{E}(\mathbb{1}_{A\cap B}) - \mathbb{E}(\mathbb{1}_{B\cap C}) - \mathbb{E}(\mathbb{1}_{A\cap C}) + \mathbb{E}(\mathbb{1}_{A\cap B\cap C}) \\
&\qquad = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\
&\qquad\quad - \mathbb{P}(A \cap B) - \mathbb{P}(B \cap C) - \mathbb{P}(C \cap A) + \mathbb{P}(A \cap B \cap C)
\end{aligned}
$$

Erwartungswert

Definition. Sei X eine diskrete Zufallsvariable mit einer abzählbar unendlichen Bildmenge. Nehmen wir an, dass die Reihe

$$\sum_{k \in \operatorname{ran} X} |k| \cdot \mathbb{P}(X = k)$$

konvergent ist. Dann ist auch die Reihe

$$\sum_{k \in \operatorname{ran} X} k \cdot \mathbb{P}(X = k)$$

konvergent, und ihr Grenzwert definiert den Erwartungswert von X.

(Wegen der absoluten Konvergenz hängt der letzte Grenzwert von der Reihenfolge der Glieder nicht ab, also ist die obige Definition sinvoll.)

Erwartungswert

Bemerkung. Wenn $\operatorname{ran} X$ nur positive oder nur negative Zahlen enthält, und die Reihe

$$\sum_{k \in \operatorname{ran} X} |k| \cdot \mathbb{P}(X = k)$$

nicht konvergent ist, dann divergiert die Reihe

$$\sum_{k \in \operatorname{ran} X} k \cdot \mathbb{P}(X = k)$$

gegen plus oder minus Unendlich. In diesem Fall sagen wir, dass der Erwartungswert (definiert und) plus bzw. minus Unendlich ist. Den allgemeinen Fall betrachten wir später.

Erwartungswert

Satz. Wenn die Erwartungswerte der diskreten Zufallsvariablen X und Y endlich sind, dann gelten

$$\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y \text{ und } \mathbb{E}(cX) = c \cdot \mathbb{E}X$$

für jede reelle Zahl c.

(Hier ist es nicht angenommen, dass $\operatorname{ran} X$ oder $\operatorname{ran} Y$ endlich ist.)

Erwartungswert

Beispiel. $X \sim \mathrm{Geo}(p)$ $(p \in (0;1])$

$$\begin{aligned}
\mathbb{E}X &= \sum_{k=1}^{\infty} k \cdot \mathbb{P}(X = k) = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} p \\
&= p + \sum_{k=2}^{\infty} k \cdot (1-p)^{k-1} p = p + \sum_{i=1}^{\infty} (i+1) \cdot (1-p)^{i} p \\
&= p + \sum_{i=1}^{\infty} i(1-p)^{i} p + \sum_{i=1}^{\infty} (1-p)^{i} p \\
&= p + (1-p) \sum_{i=1}^{\infty} i(1-p)^{i-1} p + (1-p) \sum_{i=1}^{\infty} (1-p)^{i-1} p \\
&= p + (1-p)\mathbb{E}X + (1-p) \sum_{i=1}^{\infty} \mathbb{P}(X = i) \\
&= p + (1-p)\mathbb{E}X + (1-p) = 1 + (1-p)\mathbb{E}X
\end{aligned}$$

Erwartungswert

Also $p \cdot \mathbb{E}X = 1$, d.h. $\mathbb{E}X = \frac{1}{p}$.

Erwartungswert

Ein alternativer Beweis:

$$\begin{aligned}
\mathbb{E}X &= \sum_{k=1}^{\infty} k \cdot \mathbb{P}(X = k) = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} p \\
&= p \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} = p \sum_{k=1}^{\infty} \sum_{i=1}^{k} (1-p)^{k-1} \\
&= p \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} (1-p)^{k-1} = p \sum_{i=1}^{\infty} (1-p)^{i-1} \sum_{k=i}^{\infty} (1-p)^{k-i} \\
&= p \sum_{i=1}^{\infty} (1-p)^{i-1} \sum_{k=0}^{\infty} (1-p)^{k} = p \sum_{i=1}^{\infty} (1-p)^{i-1} \cdot \frac{1}{1-(1-p)} \\
&= \sum_{i=1}^{\infty} (1-p)^{i-1} = \sum_{i=0}^{\infty} (1-p)^{i} = \frac{1}{p}
\end{aligned}$$

Charakterisierung der Wahrscheinlichkeitsfunktion

Sei $X : \Omega \to \mathbb{R}$ eine beliebige diskrete Zufallsvariable.

- Falls t die Wertemenge von X durchläuft, dann gibt die Vereinigung der paarweise disjunkten Ereignisse $\{X = t\}$ die ganze Ergebnismenge Ω (denn jede Zahl $t \in \operatorname{ran} X$ wird von X für mindestens ein Element $\omega \in \Omega$ angenommen.)
- Also bilden die Ergebnisse ein volständiges Ereignissystem.
- Es folgt, dass die Summe der Werte $f_X(t) = \mathbb{P}(X = t)$, wo t die Menge $\operatorname{ran} X$ durchläuft, die Wahrscheinlichkeit von Ω, d.h. 1 ist.

Charakterisierung der Wahrscheinlichkeitsfunktion

Behauptung. Sei X eine diskrete Zufallsvariable, und sei f_X die zugehörige Wahrscheinlichkeitsfunktion. Dann gilt

$$\sum_{t \in \operatorname{ran} X} f_X(t) = 1.$$

(Wir zeigten/überprüften das schon in den Fällen der Binomial- und geometrischen Verteilung und es gilt auch im Allgemeinen.)

Charakterisierung der Wahrscheinlichkeitsfunktion

Die obige Gleichung

$$\sum_{t \in \mathrm{ran}\, X} f_X(t) = 1$$

charakterisiert die Wahrscheinlichkeitsfunktion:

Behauptung. Sei $p_1, \ldots, p_n, \ldots \in [0;1]$ eine endliche oder abzählbar unendliche Folge von Zahlen, deren Summe 1 ist. Dann existiert eine diskrete Zufallsvariable X, für die $f_X(k) = p_k$ für alle Indexe k gilt.

Poisson-Verteilung

Definition. Eine Zufallsvariable X mit $\operatorname{ran} X = \mathbb{N}$ ist Poisson-verteilt (Poisson-eloszlású) mit Parameter $\lambda > 0$, falls

$$f_X(k) = \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

für jede $k \in \mathbb{N}$ gilt. Wir schreiben dann $X \sim \mathrm{Poi}(\lambda)$.

Poisson-Verteilung

Bemerkung. Die obige Wahrscheinlichkeiten summieren sich zu 1, denn

$$\sum_{k=0}^{\infty} \mathbb{P}(X = k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

Poisson-Verteilung

Satz. Für $X \sim \mathrm{Poi}(\lambda)$ gilt $\mathbb{E}X = \lambda$.

Beweis.

$$\mathbb{E}X = \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{k!}$$

$$= \lambda \cdot e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda.$$

Poisson-Verteilung

Die Poisson-Verteilung ist oft ein gutes Modell, wenn wir viele unabhängige Ereignisse betrachten, die einzeln mit kleinen Wahrscheinlichkeiten eintreten.

Sei X die Anzahl der Ereignisse, die eintreten, dann können (und werden) wir oft $X \sim \mathrm{Poi}(\lambda)$ (mit einem passenden Parameter λ) annehmen.

Poisson-Verteilung

Beispiele von Zufallsvariablen, die approximativ Poisson-verteilt sind:

- Anzahl der Schäden, die einer Versicherung gemeldet werden (viele Versicherungsverträge, jeder Vertrag erzeugt mit einer sehr kleinen Wahrscheinlichkeit einen Schaden).
- Anzahl der Druckfehler in einem Buch (viele Buchstaben, jeder Buchstabe kann mit einer sehr kleinen Wahrscheinlichkeit ein Druckfehler sein).
- Anzahl der Zugriffe auf einen Webserver (viele User, jeder User greift mit einer sehr kleinen Wahrscheinlichkeit zu).

Poisson-Verteilung

Wie können wir den Parameter λ bei der Approximation bestimmen?

Beispiel. Wie ist die Anzahl der Tore in einem Fussballspiel verteilt?

- Stellen wir vor, dass ein Fussballspiel in (eine große Zahl) n kleiner Zeitabschnitte aufgeteilt ist.
- In jedem Zeitabschnitt kann ein Tor mit einer kleinen Wahrscheinlichkeit fallen.
- Nehmen wir an, dass verschiedene Zeitabschnitte unabhängig sind.
- Dann wird die Anzahl X der Tore im ganzen Spiel (approximativ) Poisson-verteilt. Also gilt $X \sim \mathrm{Poi}(\lambda)$.

Poisson-Verteilung

Der Parameter λ kann dabei (zum Beispiel) als mittlere Anzahl der Tore pro Spiel interpretiert werden.

- Wir wissen, dass 973 Tore in 306 Spielen in einer Saison geschossen wurden.
- Die mittlere Anzahl der Tore pro Spiel war $\frac{973}{306} \approx 3{,}18$.
- Also $\mathbb{E}X = \lambda \approx 3{,}18$, d.h. wir können $\lambda = 3{,}18$ wählen.

Poisson-Verteilung

Können wir anhand dieser Daten die Anzahl der torlosen Spiele in einer Saison (mit 306 Spielen) schätzen?

- Die Wahrscheinlichkeit, dass ein Spiel torlos ausgeht, ist

$$p_0 = \mathbb{P}(X = 0) = e^{-\lambda} = e^{-3{,}18} \approx 0{,}0416.$$

- Die Anzahl Y der torlosen Spiele ist binomialverteilt mit Parametern 306 und p_0.
- Die erwartete Anzahl der torlosen Spiele ist dann

$$\mathbb{E}Y = 306 \cdot p_0 \approx 12{,}72.$$

- In der oben beschriebenen Saison (in Wirklichkeit) war diese Zahl 17, also scheint unsere Schätzung gar nicht so schlecht.

Poisson-Verteilung

Beispiel. Im Hörsaal befinden sich $n = 100$ Personen. Betrachten wir das Ereignis

$$A = \{\text{mindestens eine Person im Hörsaal hat heute Geburtstag}\}.$$

Bestimmen wir die Wahrscheinlichkeit von A.

Wir numerieren die Personen mit $1, \ldots, n$. Sei A_i das Ereignis, dass die Person i heute Geburtstag hat. Wir können die Ereignisse A_i als unabhängig betrachten. Nehmen wir noch der Einfachheit halber an, dass

$$\mathbb{P}(A_i) = \frac{1}{365}$$

gilt.

Poisson-Verteilung

Es handelt sich also um ein $n = 100$-faches Bernoulli-Experiment mit Erfolgswahrscheinlichkeit $p = \dfrac{1}{365}$, und die Anzahl der Personen im Hörsaal, die heute Geburtstag haben, ist binomialverteilt mit Parametern n und p. Also

$$\mathbb{P}(A) = 1 - \mathbb{P}(\overline{A}) = 1 - \left(\frac{364}{365}\right)^{100} \approx 0{,}2399.$$

Poisson-Verteilung

Hier zählen wir die Eintritte von vielen (100) unabhängigen Ereignissen, die einzeln mit kleiner Wahrscheinlichkeit (1/365) eintreten, also können wir auch mit Poisson-Verteilung approximieren!

Sei $X \sim \mathrm{Poi}(\lambda)$ die Anzahl der Personen, die heute Geburtstag haben. Der Parameter λ sollte der Erwartungswert der Binomialverteilung sein, also $\lambda = np = \frac{100}{365} = \frac{20}{73}$. Dann

$$\mathbb{P}(A) = 1 - \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(X = 0) = 1 - e^{-\frac{20}{73}} \approx 0{,}2396.$$

Poisson-Verteilung

D.h.: die Poisson-Verteilung $\mathrm{Poi}(\lambda)$ ist eine gute Approximation der Binomialverteilung $\mathrm{Bin}(n, p)$ für eine "große" n und für eine "kleine" p. Hierbei ist $\lambda = np$ eine gute Wahl für den Parameter, weil dann die Erwartungswerte der zwei Verteilungen gleich sind.

Diese Behauptung wird jetzt formalisiert und bewiesen:

Poisson-Verteilung

Satz. (Poisson-Grenzwertsatz) Sei $p_n \in [0,1]$ eine Folge mit

$$\lim_{n\to\infty} np_n = \lambda \in (0,\infty).$$

Sei noch S_n eine Zufallsvariable mit $S_n \sim \mathrm{Bin}(n; p_n)$. Für jedes $k \in \mathbb{N}$ gilt dann

$$\lim_{n\to\infty} \mathbb{P}(S_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Die Behauptung gilt insbesondere für $p_n = \frac{\lambda}{n}$.

Poisson-Verteilung

Beweis. Sei $k \in \mathbb{N}$ fest. Da S_n binomialverteilt ist, gilt

$$\mathbb{P}(S_n = k) = \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k}$$

$$= \frac{n(n-1)\dots(n-k+1)}{k!} \cdot p_n^k \cdot (1 - p_n)^{n-k}$$

$$= \frac{n(n-1)\dots(n-k+1)}{n^k} \cdot \frac{(np_n)^k}{k!} \left(1 - \frac{np_n}{n}\right)^n \cdot (1 - p_n)^{-k}$$

$$\longrightarrow 1 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1$$

Poisson-Verteilung

Warum konvergiert $\left(1 - \frac{np_n}{n}\right)^n$ gegen $e^{-\lambda}$?

- Wegen der Dreiecksungleichung gilt

$$\left|\left(1 - \frac{np_n}{n}\right)^n - e^{-\lambda}\right| \leq$$

$$\leq \left|\left(1 - \frac{np_n}{n}\right)^n - \left(1 - \frac{\lambda}{n}\right)^n\right| + \left|\left(1 - \frac{\lambda}{n}\right)^n - e^{-\lambda}\right|.$$

- Wir wissen, dass $\left|\left(1 - \frac{\lambda}{n}\right)^n - e^{-\lambda}\right| < \frac{\varepsilon}{2}$ für jede feste $\varepsilon > 0$ und für jede genügend große n (die nur von λ und ε abhängt) gilt.

Poisson-Verteilung

Zusätzlich, falls n genügen groß ist, dann haben wir auch

$$\left|\left(1-\frac{np_n}{n}\right)^n-\left(1-\frac{\lambda}{n}\right)^n\right|=$$

$$=\left|1-\frac{np_n}{n}-\left(1-\frac{\lambda}{n}\right)\right|\cdot\left|\left(1-\frac{np_n}{n}\right)^{n-1}+\dots\right|$$

$$=\frac{|\lambda-np_n|}{n}\cdot\left\{(1-p_n)^{n-1}+(1-p_n)^{n-2}\left(1-\frac{\lambda}{n}\right)+\dots\right\}$$

$$\leq\frac{\varepsilon/2}{n}\cdot n\cdot 1=\frac{\varepsilon}{2},$$

und die Behauptung folgt.

Poisson-Verteilung

Der Fehler der Approximation kann auch beschätzt werden:

Satz. Sei $S \sim \mathrm{Bin}(n, \lambda/n)$ eine binomialverteilte Variable mit $\lambda > 0$. Dann gilt die Abschätzung

$$\sum_{k=0}^{n} \left| \mathbb{P}(S = k) - e^{-\lambda} \frac{\lambda^k}{k!} \right| \leq \frac{2\lambda^2}{n}.$$

Im obigen Beispiel ist die rechte Seite

$$2 \cdot \left(\frac{100}{365} \right)^2 \cdot \frac{1}{100} = \frac{200}{365^2} \approx 0{,}0015.$$