Zufallsvariablen

Silke Bott

Sommersemester 2023

Zufallsvariablen

Die Ereignisse, die im Rahmen der Wahrscheinlichkeitstheorie betrachtet werden, können sehr unterschiedlicher Natur sein. Insbesondere sind sie in vielen Fällen keine Zahlen sondern abstrakte Ereignisse. Damit wir damit rechnen können, ist es in der Regel erforderlich, dass wir die Ereignisse in Zahlen (oder Zahlentupel) übertragen.

Definition

Eine *Zufallsvariable* oder *Messvariable* X ist eine Abbildung

$$X : \Omega \longrightarrow \mathbb{R}$$

so dass für jedes $a \in \mathbb{R}$ gilt: $\{\omega \in \Omega | X(\omega) \leq a\} \in \mathcal{A}$.

Zufallsvariablen

Beispiel

Wird ein Laplacewürfel geworfen, so ist $\Omega = \{1, 2, 3, 4, 5, 6\}$, und $X : \Omega \longrightarrow \mathbb{R}$ gegeben durch $X(\omega) = \omega$ ist eine Zufallsvariable.
Ebenso ist aber auch $Y : \Omega \longrightarrow \mathbb{R}$ mit

$$Y(\omega) = \begin{cases} 1 & \text{falls } \omega = 6 \\ 0 & \text{sonst} \end{cases}$$

eine Zufallsvariable.

Bemerkung

Für ein $x \in \mathbb{R}$ und eine Zufallsvariable X setzen wir $\{X \leq x\} = \{\omega \in \Omega | \, X(\omega) \leq x\}$ und allgemeiner für eine beliebige Teilmenge $B \subseteq \mathbb{R}$: $\{X \in B\} = \{\omega \in \Omega | \, X(\omega) \in B\}$

Zufallsvariablen

Die Definition einer Zufallsvariable stellt sicher, dass für jedes $B \subseteq \mathbb{R}$ wobei B ein (offenes, halboffenes oder abgeschlossenes) Intervall ist oder durch Bildung abzählbarerer Durchschnitt bzw. abzählbarer Vereinigung und Komplemente hiervon entsteht, gilt

$$\{X \in B\} \in \mathcal{A}$$

Insbesondere also

$$\{X \leq x\} \in \mathcal{A}$$

Speziell ist also die Wahrscheinlichkeit für diese Mengen wohldefiniert, d.h.

$$p(\{X \leq x\}) \quad \text{existiert}$$

Wir schreiben kurz $p(X \leq x)$ für $p(\{X \leq x\})$.

Zufallsvariablen

Definition

Eine Menge $B \subseteq \mathbb{R}$ heißt *messbar*, wenn B ein (offenes, halboffenes oder abgeschlossenes) Intervall ist oder durch Bildung abzählbarerer Durchschnitt bzw. abzählbarer Vereinigung und Komplemente hiervon entsteht
Die Menge aller Wahrscheinlichkeiten $p(X \in B)$ wobei B messbar ist, nennen wir die *Wahrscheinlichkeitsverteilung* von X.

Bemerkung

Die Wahrscheinlichkeitsverteilung einer Zufallsvariable ist schon bestimmt durch die Wahrscheinlichkeiten $p(X \leq x)$, $(x \in \mathbb{R})$. Die Funktion $F : \mathbb{R} \longrightarrow \mathbb{R}$ mit

$$F(x) = p(X \leq x)$$

heißt die **Verteilungsfunktion** der Zufallsvariable X.

Zufallsvariablen

Beispiel

Ist X die Zufallszahl *Augenzahl eines Laplacewürfels* so hat X die Verteilungsfunktion

Übung

Wir betrachten einen Laplacewürfel und die Zufallsvariable X, die gegeben ist durch

$$X(\omega) = \begin{cases} 1 & \text{falls } \omega = 6 \\ 0 & \text{sonst} \end{cases}$$

Bestimmen Sie die Verteilungsfunktion von X.

Zufallsvariablen

Lösung:

$$F(x) = \begin{cases} 0 & \text{für } x < 0 \\ \frac{5}{6} & \text{für } 0 \leq x < 1 \\ 1 & \text{für } x \geq 1 \end{cases}$$

Dann ist die Verteilungsfunktion von X hat die folgende Gestalt

Zufallsvariablen

Definition

Zwei Zufallsvariablen X und Y heißen *stochastisch unabhängig*, wenn für je zwei messbare Mengen $A, B \subseteq \mathbb{R}$ gilt

$$p((X \in A) \cap (Y \in B)) = p(X \in A) \cdot p(Y \in B)$$

Zufallsvariablen $X_1, \ldots, X_n$ heißen stochastisch unabhängig, wenn für alle messbaren Mengen $A_1, \ldots, A_n$ gilt

$$p((X_1 \in A_1) \cap \cdots \cap (X_n \in A_n)) = p(X_1 \in A_1) \cdots p(X_n \in A_n)$$

Zufallsvariablen

Bemerkung

Die stochastische Unabhängigkeit von X und Y bedeutet, dass sich die beiden Zufallsvariablen nicht gegenseitig beeinflussen, dass also die Messungen, die dadurch durchgeführt werden, voneinander unabhängig sind.

Bemerkung

Stochastische Unabhängigkeit können wir auch für unendlich viele Zufallsvariablen $X_1, \dots, X_n, \dots$ betrachten. In diesem Fall bedeutet es, dass jede endliche Teilmenge davon stochastisch unabhängig ist.

Zufallsvariablen

Beispiel

Wir würfeln zweimal mit einem Laplacewürfel. Mit X bezeichnen wir das Ergebnis des ersten Wurfes, und mit Y das des zweiten. Dann sind X und Y stochastisch unabhängig.

Übung

Wir würfeln zweimal hintereinander mit einem Laplacewürfel und gehen dabei vor wie folgt: Fällt im ersten Wurf eine gerade Zahl, so wird ganz normal mit dem zweiten Laplacewürfel geworfen. Fällt dagegen im ersten Wurf eine ungerade Zahl, so wird auf dem zweiten Laplacewürfel die 1 mit einer 2, die 3 mit einer 4 und die 5 mit einer 6 überklebt und erst dann geworfen (wobei aber immer noch jede Seite gleichwahrscheinlich sein soll). Mit X bezeichnen wir das Ergebnis des ersten Wurfes, und mit Y das des zweiten.
Sind X und Y stochastisch unabhängig?

Zufallsvariablen

Lösung:

X und Y sind nicht stochastisch unabhängig.
Es gilt nämlich

$$p(X=6)=\frac{6}{36},\quad p(Y=6)=\frac{9}{36}$$

(wie wir schon festgestellt haben), und (wie wir ebenfalls dort schon berechnet haben)

$$p(X=6,Y=6)=\frac{1}{36}\neq p(X=6)\cdot p(Y=6)$$

diskrete Zufallsvariablen

Definition

Eine Zufallsvariable $X : \Omega \longrightarrow \mathbb{R}$ heißt **diskrete Zufallsvariable**, falls sie nur endlich viele Werte $x_1, \ldots, x_k$ oder abzählbar unendlich viele Werte $x_1, x_2, \ldots, x_k, \ldots$ annehmen kann.

Bemerkung

Die Wahrscheinlichkeitsverteilung einer diskreten Zufallsvariable mit Werten $x_1, x_2, \ldots, x_k \ldots$ ist schon bestimmt durch die **Elementarwahrscheinlichkeiten** $p_i = p(X = x_i)$.

Bemerkung

Zwei diskrete Zufallsvariablen X und Y sind schon dann stochastisch unabhängig, wenn

$$p(X = x, Y = y) = p(X = x) \cdot p(Y = y) \qquad \text{für alle } x, y \in \mathbb{R}$$

diskrete Zufallsvariablen

Beispiel

Die Elementarwahrscheinlichkeiten des Laplacewürfelwurfes sind gegeben durch

$$p(X = k) = \frac{1}{6} \qquad \text{für alle } k \in \{1, \ldots, 6\}$$

Beispiel

Beim Laplacewürfel sind die Elementarwahrscheinlichkeiten für

$$X(\omega) = \begin{cases} 1 & \text{falls } \omega = 6 \\ 0 & \text{sonst} \end{cases}$$

sind gegeben durch

$$p(X = 0) = \frac{5}{6} \qquad p(X = 1) = \frac{1}{6}$$

diskrete Zufallsvariablen

Übung

Wir würfeln zweimal unabhängig mit einem Laplacwürfel und bezeichnen mit X die Zufallsvariable *Summe der beiden Augenzahlen*. Bestimmen Sie die Elementarwahrscheinlichkeiten für X.

diskrete Zufallsvariablen

Übung

Wir würfeln zweimal unabhängig mit einem Laplacwürfel und bezeichnen mit X die Zufallsvariable *Summe der beiden Augenzahlen*. Bestimmen Sie die Elementarwahrscheinlichkeiten für X.

Lösung:

$$
\begin{array}{lllllllll}
p(X=2) & = & \frac{1}{36} & p(X=3) & = & \frac{1}{18} & p(X=4) & = & \frac{1}{12} \\
p(X=5) & = & \frac{1}{9} & p(X=6) & = & \frac{5}{36} & p(X=7) & = & \frac{1}{6} \\
p(X=8) & = & \frac{5}{36} & p(X=9) & = & \frac{1}{9} & p(X=10) & = & \frac{1}{12} \\
p(X=11) & = & \frac{1}{18} & p(X=12) & = & \frac{1}{36} & & &
\end{array}
$$

diskrete Zufallsvariablen

Übung

Wir würfeln zweimal unabhängig mit einem Laplacwürfel und bezeichnen mit X die Zufallsvariable *Summe der beiden Augenzahlen*. Bestimmen Sie die Elementarwahrscheinlichkeiten für X.

Lösung:

$$
\begin{array}{llllll}
p(X=2) = \frac{1}{36} & p(X=3) = \frac{1}{18} & p(X=4) = \frac{1}{12} \\
p(X=5) = \frac{1}{9} & p(X=6) = \frac{5}{36} & p(X=7) = \frac{1}{6} \\
p(X=8) = \frac{5}{36} & p(X=9) = \frac{1}{9} & p(X=10) = \frac{1}{12} \\
p(X=11) = \frac{1}{18} & p(X=12) = \frac{1}{36} &
\end{array}
$$

diskrete Zufallsvariablen

Definition

Die **Wahrscheinlichkeitsfunktion** $f(x)$ einer diskreten Zufallsvariable X mit Werten $x_1, x_2, \ldots, x_k, \ldots$ ist definiert durch

$$f(x) = \begin{cases} p_i = p(X = x_i) & \text{für} \quad x = x_i \in \{x_1, x_2, \ldots, x_k, \ldots\} \\ 0 & \text{sonst} \end{cases}$$

Bemerkung

Die **Verteilungsfunktion** $F(x)$ einer diskreten Zufallsvariable X mit Werten $x_1, x_2, \ldots, x_k, \ldots$ ist gegeben durch

$$F(x) = p(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

diskrete Zufallsvariablen

Definition

Ist X eine diskrete Zufallsvariable mit endlich vielen Werten $x_1, \dots, x_k$, so heißt X **gleichverteilt**, wenn

$$p(X = x_i) = \frac{1}{k} \qquad \text{für alle } i \in \{1, \dots, k\}$$

Beispiel

Beim Werfen eines Laplacewürfels mit Zufallsvariable X mit

$$X(\omega) = \text{ Augenzahl des Wurfes } \omega$$

ist eine gleichverteilte Zufallsvariable mit

$$p(X = i) = \frac{1}{6} \quad \text{für alle } \; i \in \{1, \dots, 6\}$$

diskrete Zufallsvariablen

Definition

Eine diskrete Zufallsvariable X heißt **Bernoulli–Variable**, wenn X nur die beiden Werte 0 und 1 annimmt. Die zugehörige Bernoulli–Verteilung ist schon festgelegt durch

$$p := p(X = 1)$$

denn notwendig gilt dann $p(X = 0) = 1 - p$. Der Wert p heißt dann Parameter der Bernoulli–Variable

Beispiel

Eine Produktionsanlage produziert Teile, die mit einer Wahrscheinlichkeit von 20 % von minderer Qualität sind. Dann ist die Zufallsvariable X mit

$$X(\omega) = \begin{cases} 0 & \text{falls } \omega \text{ von minderer Qualität} \\ 1 & \text{sonst} \end{cases}$$

eine Bernoulli–Variable mit Parameter $p = 0.80$

diskrete Zufallsvariablen

Definition

Der *Erwartungswert* einer diskreten Zufallsvariable X mit Werten $x_1, \dots, x_k, \dots,$ und Elementarwahrscheinlichkeiten $p_k = p(X = x_k)$ ist

$$\mu = E(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + \cdots + x_k \cdot p_k + \cdots = \sum_{k \geq 1} x_k \cdot p_k$$

falls diese Summe existiert.
Die *Varianz* einer diskreten Zufallsvariable X mit Erwartungswert μ ist

$$\begin{aligned} \sigma^2 = \mathrm{Var}(X) &= (x_1 - \mu)^2 \cdot p_1 + (x_2 - \mu)^2 \cdot p_2 + \cdots \\ &= \sum_{k \geq 1} (x_k - \mu)^2 \cdot p_k \end{aligned}$$

falls diese Summe existiert. Die *Standardabweichung* von X ist

$$\sigma = +\sqrt{\mathrm{Var}(X)}$$

diskrete Zufallsvariablen

Regel

Für Erwartungswerte gilt:

1. *Ist* $Y = aX + b$, *so gilt*

$$E(Y) = aE(X) + b$$

2. *Sind* X *und* Y *zwei diskrete Zufallsvariablen, so gilt*

$$E(X + Y) = E(X) + E(Y)$$

3. *Sind* X *und* Y *zwei stochastisch unabhängige Zufallsvariablen, so gilt*

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

diskrete Zufallsvariablen

Regel

Für Varianzen gilt

1. *Ist* $\mu = E(X)$, *so ist*

$$\mathrm{Var}(X) = E(X-\mu)^2$$

2. *Ist* $Y = aX + b$, *so ist*

$$\mathrm{Var}(Y) = a^2 \cdot \mathrm{Var}(X)$$

3. *Sind* X *und* Y *zwei stochastisch unabhängige Zufallsvariablen, so gilt*

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

diskrete Zufallsvariablen

Beispiel

Für eine Bernoulli–Variable X mit $p(X = 1) = p$ gilt

$$\begin{aligned} E(X) &= p \\ \mathrm{Var}(X) &= p \cdot (1 - p) \end{aligned}$$

Beispiel

In dem Produktionsbeispiel (mit $p = 0.80$) ist

$$E(X) = 0.80, \quad \mathrm{Var}(X) = 0.8 \cdot 0.2 = 0.16, \quad \sigma = 0.40$$

diskrete Zufallsvariablen

Übung

Bestimmen Sie Erwartungswert und Varianz für den Laplacewürfel und die Zufallsvariable X die die Augenzahl des Wurfergebnisses misst.

Lösung:

$$E(X) = 3.5, \quad \mathrm{Var}(X) = 2.916\overline{6}, \quad \sigma = 1.7078$$

diskrete Zufallsvariablen

Wir betrachten eine Experiment, dass durch eine Zufallsvariable X auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, p)$ beschrieben wird. Wird dieses Experiment n mal stochastich unabhängig voneinander wiederholt, so bezeichnen wir mit X_i die Zufallsvariable, die die i–te Wiederholung des Experiments beschreibt.

$$X_1, \ldots, X_n \quad \text{sind unabhängige Wiederholungen von } X$$

diskrete Zufallsvariablen

Wiederum sei bei einem Zufallsvorgang nur von Interesse, ob ein bestimmtes Ereignis A eintritt, wobei $P(A) = p$ mit $0 < p < 1$ gelte. Der Zufallsvorgang werde nun unabhängig voneinander sooft wiederholt, bis zum ersten Mal A eintritt. Als Zufallsvariable definieren wir

$$Y = \text{Anzahl der Versuche bis zum ersten Mal } A \text{ eintritt.}$$

Um die Wahrscheinlichkeitsfunktion zu bestimmen, betrachten wir das Ereignis $Y = n$. Es ist dadurch definiert, dass zunächst $(n-1)$-mal das Komplementärereignis $\overline{A}$ und beim n-ten Versuch A auftritt. Wegen der Unabhängigkeit der Versuche gilt

$$p_n = P(Y = n) = (1-p)^{n-1} \cdot p \text{ für } k = 1, 2, 3, \ldots.$$

Da die Wahrscheinlichkeiten p_n eine geometrische Reihe bilden, heißt die Verteilung geometrisch.

diskrete Zufallsvariablen

Beispiel

Wir betrachten unabhängige Wiederholungen $X_1, X_2, \dots, X_n, \dots$ einer Bernoulli–Variable X mit $p(X = 1) = p\,(> 0)$. Dann heißt die Zufallsvariable Y mit $Y(\omega) = \min\{n \in \mathbb{N} |\, X_n(\omega) = 1\}$ *geometrisch verteilte Zufallsvariable* mit Parameter p.
Hierfür gilt

$$p(Y = n) = (1 - p)^{n-1} \cdot p$$

und damit

$$E(Y) = \frac{1}{p} \qquad \mathrm{Var}(Y) = \frac{1 - p}{p^2}$$

Bemerkung

Für eine geometrisch verteilte Zufallsvariable X mit Parameter p schreiben wir kurz $X \sim G(p)$.

diskrete Zufallsvariablen

Beispiel

Wir betrachten einen Laplacewürfel und die Zufallsvariable X Augenzahl, sowie unabhängige Wiederholgungen $X_1, X_2, \ldots, X_n, \ldots$ von X und die Zufallsvariable Y mit

$$Y(\omega) = \min\{n \in \mathbb{N} | \, X_n(\omega) = 6\}$$

Dann ist $Y \sim G\left(\frac{1}{6}\right)$ mit

$$E(Y) = 6, \quad \mathrm{Var}(Y) = \frac{\frac{5}{6}}{\frac{1}{36}} = 30$$

diskrete Zufallsvariablen

Übung

Wir betrachten wieder das Beispiel mit der Produktionsanlage mit einer Wahrscheinlichkeit von 20 % für ein minderwertiges Teil.
Nach wie vielen zufällig ausgewählten Teilen erwarten Sie im Mittel zum ersten Mal ein fehlerhaftes Teil?

Lösung:

Nach 5 Teilen kommt im Mittel erstmals ein fehlerhaftes.

diskrete Zufallsvariablen

Wir betrachten eine Bernoulli-Kette von n-mal wiederholten Bernoulli-Experimenten mit $p = P(A)$ für das interessierende Ereignis und der Folge $A_1, A_2, ..., A_n$ von unabhängigen Ereignissen. Dabei steht A_i für das Eintreten von A im i-ten Versuch. Zur Bernoulli-Kette definieren wir die Zufallsvariable

$$Y = \text{Anzahl der Versuche, bei denen } A \text{ eintritt.}$$

Das Standardmodell für eine Bernoulli-Kette ist das *Ziehen mit Zurücklegen* aus einer Urne, die N Kugeln, darunter M schwarze, enthält. Daraus werden zufällig und mit Zurücklegen nacheinander n Kugeln gezogen. Die Ereignisse A_i Bei der i-ten Ziehung wird eine schwarze Kugel gezogen. sind dann unabhängig und es ist $P(A_i) = M/N = p$

diskrete Zufallsvariablen

Die Wahrscheinlichkeitsfunktion von Y lässt sich folgendermaßen ableiten. Das Ereignis $Y = k$ resultiert z.B. für die Ereignisfolge $A_1 A_2 ... A_k \overline{A}_{k+1} ... \overline{A}_n$, bei der in der Bernoulli-Kette zuerst k-mal das Ereignis A und anschließend $n - k$-mal $\overline{A}$ auftritt. Wegen der Unabhängigkeit der einzelnen Versuche gilt:
$P(A_1 A_2 ... A_k \overline{A}_{k+1} ... \overline{A}_n) = p^k (1 - p)^{n-k}$. Das Ereignis tritt aber auch ein, wenn n-mal A und $(n - k)$-mal $\overline{A}$ in irgendeiner Reihenfolge erscheinen.

Insgesamt gibt es genau $\binom{n}{k}$ verschiedene derartige Reihenfolgen.

diskrete Zufallsvariablen

Beispiel

Wir betrachten n unabhängige Wiederholungen $X_1, X_2, \ldots, X_n$ einer Bernoulli–Variable X mit $p(X = 1) = p\,(> 0)$. Dann heißt die Zufallsvariable Y mit $Y(\omega) = |\{i \in \{1, \ldots, n\} \,|\, X_i(\omega) = 1\}|$ *binomial verteilte Zufallsvariable* mit Parametern n und p.
Hierfür gilt

$$p(Y = k) = \binom{n}{k} \cdot (1 - p)^{n-k} \cdot p^k$$

und damit

$$E(Y) = n \cdot p \qquad \mathrm{Var}(Y) = n \cdot p \cdot (1 - p)$$

Bemerkung

Ist X binomialverteilt mit Parametern n und p, so schreiben wir hierfür kurz $X \sim B(n, p)$.

diskrete Zufallsvariablen

Beispiel

Wir betrachten das viermalige unabhängige Würfeln eines Laplacewürfels, die Zufallsvariable X_k, die die Augenzahl im k–ten Wurf registriert und die Zufallsvariable

$$Y(\omega) = \text{Anzahl der } k \in \{1, \dots, 4 \text{ mit } X_k(\omega) = 6\}$$

Dann ist Y binomialverteilt mit Parametern $n = 4$ und $p = \frac{1}{6}$.
Es ist

$$p(Y = 2) = \binom{4}{2} \cdot \left(\frac{5}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2 = \frac{150}{1296}$$

und

$$p(Y \leq 2) = p(Y = 0) + p(Y = 1) + p(Y = 2) = \frac{1275}{1296}$$

diskrete Zufallsvariablen

Übung

Wir betrachten die Produktionsanlage mit einem Anteil von 20 % fehlerhafter Teile.
Mit welcher Wahrscheinlichkeit sind von 12 zufällig herausgegriffenen Teilen mindestens 4 fehlerhaft?

diskrete Zufallsvariablen

Übung

Wir betrachten die Produktionsanlage mit einem Anteil von 20 % fehlerhafter Teile.
Mit welcher Wahrscheinlichkeit sind von 12 zufällig herausgegriffenen Teilen mindestens 4 fehlerhaft?

Lösung:

Die Zufallsvariable $Y =$ *Anzahl der fehlerhaften Teile* ist binomialverteilt mit den Parametern $p = 0.2$ und $n = 12$. Damit berechnet sich diese Wahrscheinlichkeit zu

$$\begin{aligned} p(Y \geq 4) &= 1 - p(Y \leq 3) \\ &= 1 - \binom{12}{3} \cdot 0.2^3 \cdot 0.8^9 - \binom{12}{2} \cdot 0.2^2 \cdot 0.8^{10} \\ &\quad - \binom{12}{1} \cdot 0.2^1 \cdot 0.8^{11} - \binom{12}{0} \cdot 0.2^0 \cdot 0.8^{12} \\ &= 0.2054 \end{aligned}$$

diskrete Zufallsvariablen

Übung

Wir betrachten die Produktionsanlage mit einem Anteil von 20 % fehlerhafter Teile.
Mit welcher Wahrscheinlichkeit sind von 12 zufällig herausgegriffenen Teilen mindestens 4 fehlerhaft?

Lösung:

Die Zufallsvariable $Y = \textit{Anzahl der fehlerhaften Teile}$ ist binomialverteilt mit den Parametern $p = 0.2$ und $n = 12$. Damit berechnet sich diese Wahrscheinlichkeit zu

$$
\begin{aligned}
p(Y \geq 4) &= 1 - p(Y \leq 3) \\
&= 1 - \binom{12}{3} \cdot 0.2^3 \cdot 0.8^9 - \binom{12}{2} \cdot 0.2^2 \cdot 0.8^{10} \\
&\quad - \binom{12}{1} \cdot 0.2^1 \cdot 0.8^{11} - \binom{12}{0} \cdot 0.2^0 \cdot 0.8^{12} \\
&= 0.2054
\end{aligned}
$$

diskrete Zufallsvariablen

Aus einer endlichen Grundgesamtheit von N Einheiten, bei denen M eine Eigenschaft A besitzen, wird *n-mal rein zufällig, aber ohne Zurücklegen* gezogen. Im Urnenmodell mit N Kugeln, davon M schwarzen, entspricht dies dem zufälligen n-maligen Ziehen ohne Zurücklegen. Wir interessieren uns wieder für die Zufallsvariable

$$Y = \text{Anzahl der gezogenen Objekte mit der Eigenschaft } A.$$

Dies führt zur *hypergeometrischen Verteilung*.

diskrete Zufallsvariablen

Definition

Eine Zufallsvariable X heißt *hypergeometrisch verteilt* mit Parametern n, M und N wenn gilt

$$p(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$

diskrete Zufallsvariablen

Bemerkung

Ist X Hypergemetrisch-verteilt mit Parameter n, M und N so schreiben wir kurz $X \sim H(n, M, N)$.

Beispiel

Ist X Hypergeometrisch–verteilt mit Parameter mit Parameter n, M und N, so gilt

$$\begin{aligned} E(X) &= n\frac{M}{N} \\ \mathrm{Var}(X) &= n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1} \end{aligned}$$

diskrete Zufallsvariablen

Binomial- und hypergeometrische verteilte Zufallsvariablen zählen, *wie oft* bei n-maligen Ziehen aus Urnen oder Grundgesamtheiten ein bestimmtes Ereignis A eintritt. Der Wertebereich ist nach oben durch n begrenzt und somit endlich. Die geometrische Verteilung zählt, *wie lange* man warten muss, bis ein Ereignis A *zum erstenmal* eintritt. Der Wertebereich ist die Menge $\mathbb{N}$ der natürlichen Zahlen und nicht nach oben begrenzt. Die *Poisson-Verteilung* eignet sich ebenfalls zur Modellierung von *Zählvorgängen*. Dabei werden bestimmte Ereignisse gezählt die innerhalb eines festen, vorgegebenen Zeitintervalls eintreten können. Die mögliche Anzahl der Ereignisse ist gleichfalls nicht nach oben begrenzt.

diskrete Zufallsvariablen

Definition

Eine diskrete Zufallsvariable X mit Wertebereich $\mathbb{N}$ heißt **Poisson–verteilt** mit Parameter λ, wenn gilt

$$P(X = n) = \frac{\lambda^n}{n!} \cdot e^{-\lambda} \quad \text{für alle } n \in \mathbb{N}$$

Bemerkung

Ist X Poisson–verteilt mit Parameter λ, so schreiben wir kurz $X \sim Po(\lambda)$.

Beispiel

Ist X Poisson–verteilt mit Parameter λ, so gilt

$$\begin{aligned} E(X) &= \lambda \\ \mathrm{Var}(X) &= \lambda \end{aligned}$$

diskrete Zufallsvariablen

Beispiel

Bei einem radioaktiven Element mit hoher Halbwertszeit ist die Zufallsvariable

$$X = \textit{Anzahl der Teilchen, die in einer Sekunde zerfällt}$$

Poisson–verteilt mit Parameter λ, wobei λ die Anzahl der Teilchen ist, die im Mittel in einer Sekunde zerfallen.

Beispiel

Zerfallen in einer Sekunde im Mittel 4 Teilchen, so gilt

$$p(X=0) = \frac{4^0}{0!} \cdot e^{-4} = e^{-4} = 0.0183$$

diskrete Zufallsvariablen

Beispiel

Die Anzahl X der Blitze die pro Jahr in einem Hektar Land in Deutschland einschlägt ist (in sehr guter Näherung) Poisson–verteilt mit Parameter $\lambda = 0.10$. Damit gilt

$$p(X = 0) = \frac{0.10^0}{0!} \cdot e^{-0.10} = 0.9048 \quad p(X = 1) = \frac{0.10^1}{1!} \cdot e^{-0.10} = 0.09048$$

und

$$p(X \geq 4) = 1 - p(X \leq 3) = 0.000003847$$

diskrete Zufallsvariablen

Übung

Die Anzahl der Verkehrsunfälle, die sich an einer unübersichtlichen Straßenkreuzung in einem Monat ereignen, ist (in guter Näherung) Poissson–verteilt. Im Mittel ereignen sich dabei an dieser Kreuzung 1.4 Unfälle pro Monat.

a) Wie groß ist die Wahrscheinlichkeit, dass sich an dieser Kreuzung in einem gegebenen Monat genau ein Unfall ereignet?
b) Wie groß ist die Wahrscheinlichkeit, dass sich an dieser Kreuzung in einem gegebenen Monat höchsten drei Unfälle ereignen?
c) Wie groß ist die Wahrscheinlichkeit, dass sich an dieser Kreuzung in einem gegebenen Monat mindestens zwei Unfälle ereignen?

diskrete Zufallsvariablen

Lösung:

Mit X bezeichnen wir die Zufallsvariable, die die Anzahl der Verkehrsunfälle an dieser Kreuzung in einem Monat zählt. Dann ist X eine Poisson–verteilte Zufallsvariable mit Parameter $\lambda = 1.4$ (denn 1.4 ist der Erwartungswert von X). Damit gilt

$$p(X = k) = \frac{1.4^k}{k!} \cdot e^{-1.4} \quad \text{für alle } k \geq 0$$

also

a) $p(X = 1) = \frac{1.4}{1} \cdot e^{-1.4} = 0.3452.$

b) $p(X \leq 3) = p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3) = 0.9463.$

c) $p(X \geq 2) = 1 - p(X \leq 1) = 1 - (p(X = 0) + p(X = 1)) =$
$1 - 0.5918 = 0.4082.$

diskrete Zufallsvariablen

Übung

Welche Verteilungen besitzen die folgenden Zufallsvariablen:

a) Die Anzahl der Richtigen beim Lotto 6 aus 49 (X_1).

b) Die Anzahl der Richtigen beim Fußballtoto, wenn alle Spiele wegen unbespielbarem Platz ausfallen und die Ergebnisse per Los ermittelt werden (X_2).

c) Die Anzahl von Telefonanrufen in einer Auskunftsstelle während einer Stunde (X_3).

d) In einer Urne mit 100 Kugeln befinden sich 5 rote Kugeln. X_4 sei die Anzahl der roten Kugeln in der Stichprobe, wenn 10 Kugeln auf einen Schlag entnommen werden.

diskrete Zufallsvariablen

Übung

e) Die Anzahl Studenten, die den Unterscheid zwischen der Binomial- und der hypergeometrischen Verteilung verstanden haben, unter 10 zufällig ausgewählten Hörern einer Statistikveranstaltung, an der 50 Studierende teilnehmen (X_5).

f) Die Stückzahl eines selten gebrauchten Produkts, das bei der Lieferfirma an einem Tag nachgefragt wird (X_6).

diskrete Zufallsvariablen

Lösung:

a) Da die Lottozahlen ohne Zurücklegen gezogen werden, gilt $X_1 \sim H(6, 6, 49)$.

b) Da die Einzelergebnisse voneinander unabhängig sind und die Wahrscheinlichkeit, ein Einzelergebnis richtig zu tippen, jeweils $1/3$ beträgt gilt $X_2 \sim B(11, 1/3)$.

c) Falls eher selten angerufen wird, ist, da die einzelnen Anrufe als unabhängig angesehen werden können, $X_3 \sim Po(\lambda)$-verteilt. Dabei ist λ die mittlere Anzahl von Anrufen pro Stunde.

d) Ziehen auf einen Schlag entspricht dem Modell ohne Zurücklegen, d.h. $X_4 \sim H(10, 5, 100)$.

diskrete Zufallsvariablen

Lösung:

e) Befragungen entsprechen in der Regel dem Ziehen ohne Zurücklegen, d.h. $X_5 \sim H(10, M, 50)$, wobei M Hörer den Unterschied verstanden haben.

f) Ist λ die Anzahl, die im Mittel an einem Tag nachgefragt wird, dann gilt $X_6 \sim Po(\lambda)$.

diskrete Zufallsvariablen

Übung

Eine Teetrinkerin behauptet schmecken zu können, ob der Tee beim Eingießen auf die Milch gegeben wurde oder umgekehrt. Sie erklärt sich auch zu einem Experiment bereit. Eine Person füllt zehn Tassen mit Milch und Tee. Bei jeder Tasse entscheidet sie rein zufällig, ob zuerst die Milch oder zuerst der Tee in die Tasse gegeben wird. Nachdem alle Tassen gefüllt sind, wird die Teetrinkerin ins Zimmer gelassen und darf probieren. Nehmen Sie an, sie rät nur und tippt bei jeder Tasse (jeweils unabhängig von den anderen) mit Wahrscheinlichkeit 0,5 auf die richtige Reihenfolge von Tee und Milch. Wie groß ist die Wahrscheinlichkeit, dass sie mindestens achtmal richtig tippt?

diskrete Zufallsvariablen

Lösung:

Mindestens acht richtige Tipps sind gleichbedeutend mit höchstens zwei falschen Tipps. Die Anzahl X der falschen Tipps unter den zehn Versuchen ist hier aufgrund der Unabhängigkeit binomialverteilt mit den Parametern $p = 0,5$ (Wahrscheinlichkeit für einen falschen Tipp in einem Versuch) und $n = 10$ (Anzahl der Versuche insgesamt).
Damit ist die Wahrscheinlichkeit, dass höchstens zwei Tipps falsch sind, gegeben durch:

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2).$$

diskrete Zufallsvariablen

Lösung:

Mit Hilfe der Binomialverteilung ergeben sich diese Wahrscheinlichkeiten als

$$\begin{aligned} P(X=0) &= \binom{10}{0} \cdot 0,5^0 \cdot 0,5^{10} = 0,000977, \\ P(X=1) &= \binom{10}{1} \cdot 0,5^1 \cdot 0,5^9 = 0,009766, \\ P(X=2) &= \binom{10}{2} \cdot 0,5^2 \cdot 0,5^8 = 0,043945. \end{aligned}$$

Und damit ist schließlich

$$P(X \leq 2) = 0,054688.$$

diskrete Zufallsvariablen

Übung

In einer Tüte befinden sich zehn Pralinen: vier aus Nougat und sechs aus Marzipan. Hein, der absolut keine Nougat-Pralinen mag, darf nun drei Pralinen zufällig (ohne Zurücklegen) auswählen.

a) Wie ist die Anzahl X gezogener Marzipan-Pralinen verteilt? Wie viele Marzipan-Pralinen kann Hein erwarten?

b) Wie groß ist die Wahrscheinlichkeit, dass Hein genau 3 Marzipan-Pralinen zieht?

c) Wie groß ist die Wahrscheinlichkeit, dass Hein mindestens 1 Marzipan-Praline zieht?

diskrete Zufallsvariablen

Lösung:

a) Da hier ohne Zurücklegen gezogen wird, ist die Anzahl X der gezogenen Marzipan-Pralinen hypergeometrisch verteilt mit den Parametern $n = 3$ (Anzahl der Züge), $M = 6$ (Anzahl der Marzipan-Pralinen in der Tüte) und $N = 10$ (Anzahl der Pralinen insgesamt).
Der Erwartungswert von X ist gegeben durch
$E(X) = n \cdot \frac{M}{N} = 3 \cdot \frac{6}{10} = 1,8$. Hein kann also im Schnitt mit 1,8 Marzipan-Pralinen rechnen.

diskrete Zufallsvariablen

Lösung:

b) Mit Hilfe der hypergeometrischen Verteilung ergibt sich die Wahrscheinlichkeit, genau drei Marzipan-Pralinen zu ziehen, als

$$P(X = 3) = \frac{\binom{M}{3}\binom{N-M}{0}}{\binom{N}{n}} = \frac{\binom{6}{3}\binom{4}{0}}{\binom{10}{3}} = 0,167.$$

c) Die Wahrscheinlichkeit, mindestens eine Marzipan-Praline zu ziehen, berechnet sich als:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \frac{\binom{6}{0}\binom{4}{3}}{\binom{10}{3}} = 0,967.$$

diskrete Zufallsvariablen

Übung

Aus Erfahrung weiß man, dass die Wahrscheinlichkeit dafür, dass bei einem Digitalcomputer eines bestimmten Typus während 12 Stunden kein Fehler auftritt, 0,7788 beträgt.

a) Welche Verteilung eignet sich zur näherungsweisen Beschreibung der Zufallsvariablen $X =$ Anzahl der Fehler, die während 12 Stunden auftreten?

b) Man bestimme die Wahrscheinlichkeiten dafür, dass während 12 Stunden mindestens zwei Fehler auftreten.

c) Wie groß ist die Wahrscheinlichkeiten, dass bei vier (voneinander unabhängigen) Digitalcomputern desselben Typus während 12 Stunden genau ein Fehler auftritt?

diskrete Zufallsvariablen

Lösung:

a) $X \sim Po(\lambda)$ mit Wahrscheinlichkeit

$$P(X = n) = \frac{\lambda^n}{n!} \cdot e^{-\lambda} \quad \text{für alle } x = 0, 1, 2, \ldots$$

Wegen $P(X = 0) = e^{-\lambda} = 0,7788$ gilt $\lambda = -\ln 0,7788 = 0,25$, also $X \sim Po(0,25)$.

b) Man berechnet

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - 0,7788 - \tfrac{0,25^1}{1!} 0,7788 = 0,0265 \end{aligned}$$

c) Sei Y = Anzahl der Fehler, die bei vier Computern während 12 Stunden auftreten. Dann ist Y die Summe von vier unabhängigen $Po(0,25)$-verteilten Zufallsvariablen, also $Y \sim Po(1)$.

diskrete Zufallsvariablen

Übung

Die Rückversicherung will die Prämien für Versicherungen gegen Großunfälle kalkulieren. Aus Erfahrung weiß sie, dass im Mittel 3,7 bzw. 5,9 Großunfälle im Winter- bzw. Sommerhalbjahr vorfallen.

a) Welche Verteilungsannahme erscheint für die Zufallsvariablen

$$\begin{aligned} X &= \text{Anzahl der Großunfälle im Winterhalbjahr} \\ Y &= \text{Anzahl der Großunfälle im Sommerhalbjahr} \end{aligned}$$

sinnvoll?

b) Wie wahrscheinlich ist es, dass im Winterhalbjahr nicht mehr als zwei Großunfälle vorfallen? Wie wahrscheinlich ist es im Sommerhalbjahr?

c) Wie wahrscheinlich ist es, dass sowohl im Winter- als auch im Sommerhalbjahr nicht mehr als zwei Großunfälle verfallen? Welche Annahme unterstellen Sie dabei?

diskrete Zufallsvariablen

Lösung:

a) X und Y sind Poisson-verteilt, d.h. $X \sim Po(\lambda)$ und $Y \sim Po(\mu)$.

b) Die Wahrscheinlichkeiten dafür, dass nicht mehr als zwei Großunfälle auftreten, berechnet sich jeweils als:

$$\begin{aligned} P(X \leq 2) &= P(X=0)+P(X=1)+P(X=2) \\ &= e^{-3{,}7} \cdot \left(\frac{3{,}7^0}{0!}+\frac{3{,}7^1}{1!}+\frac{3{,}7^2}{2!}\right) \\ &= e^{-3{,}7} \cdot (1+3,7+6,845) = 0,285, \end{aligned}$$

$$\begin{aligned} P(X \leq 2) &= P(X=0)+P(X=1)+P(X=2) \\ &= e^{-5{,}9} \cdot \left(\frac{5{,}9^0}{0!}+\frac{5{,}9^1}{1!}+\frac{5{,}9^2}{2!}\right) \\ &= e^{-5{,}9} \cdot (1+5,9+17,405) = 0,0666. \end{aligned}$$

c) Man darf annehmen, dass X und Y unabhängig sind. In diesem Fall folgt $P(X \leq 2, Y \leq 2) = P(X \leq 2) \cdot P(Y \leq 2) = 0,285 \cdot 0,0666$ $= 0,0188$.