5.3 Diskrete zufällige Variablen

$$X(\omega) \in \{x_1, x_2, x_3, \ldots\}$$

$$X : \begin{pmatrix} x_1 & x_2 & x_3 & \cdots x_n & \cdots \\ p_1 & p_2 & p_3 & \cdots p_n & \cdots \end{pmatrix}$$

$$p_i = P(X = x_i) > 0, \quad i = 1, 2, 3, \ldots \quad \sum_{i=1}^{\infty} p_i = 1$$

Die Funktion

$$f(x_i) = p_i$$

heißt Wahrscheinlichkeitsfunktion.

Bem: $f(x_i)$ nennt man manchmal auch Zähldichte.

Beispiele

a) Zweimaliges Werfen einer Münze

$$\Omega = \{ZZ, ZB, BZ, BB\}$$

$X :=$ Anzahl von Blatt

$$X : \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

b) Erfolge bei n Versuchen

X: Anzahl der “Erfolge” bei n Versuchen, wobei jeder der n Versuche eine Erfolgswahrscheinlichkeit p hat.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{Binomialwkt.}$$

$$F_X(k) = P(X < k) = \sum_{i=0}^{k-1} \binom{n}{i} p^i (1-p)^{n-i}$$

Verteilungsfunktion von X.

Beispiele: Es seien $p = \frac{1}{2}$ und $n = 5$. Für $x = 2.5$ gilt:

$$\begin{aligned}
F(2.5) &= \sum_{i:\, i<2,5} p_i \\
&= p_0 + p_1 + p_2 \\
&= \binom{5}{0}\left(\frac{1}{2}\right)^5 + \binom{5}{1}\left(\frac{1}{2}\right)^5 + \binom{5}{2}\left(\frac{1}{2}\right)^5 \\
&= \frac{1}{32} + \frac{5}{32} + \frac{10}{32} \\
&= 0.5
\end{aligned}$$

2. Würfeln 20 mal. Wie groß ist die Wkt. für

mindestens 4 Sechsen?

X: Anzahl der Sechsen.

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) = 1 - F_X(4) \\ &= 1 - \sum_{i=0}^{3} P(X = i) = \end{aligned}$$

$$1 - (\frac{5}{6})^{20} - 20(\frac{1}{6})(\frac{5}{6})^{19} - \frac{20 \cdot 19}{2}(\frac{1}{6})^2(\frac{5}{6})^{18} -$$

$$- \frac{20 \cdot 19 \cdot 18}{6}(\frac{1}{6})^3(\frac{5}{6})^{17}$$

$\approx 0.43.$

c) Telefonzentrale

X: Anzahl der Anrufe, die pro Zeiteinheit von einer Telefonzentrale vermittelt werden.

$$X: \begin{pmatrix} 0 & 1 & 2 & 3 & \cdots \\ \\ p_0 & p_1 & p_2 & p_3 & \cdots \end{pmatrix}$$

$$P(X=i) = p_i = \frac{\lambda^i}{i!} e^{-\lambda}, \quad \lambda > 0$$

Poisson- Wahrscheinlichkeiten.

$$\sum_{i=0}^{\infty} p_i = \underbrace{\sum_{i=0}^{\infty} \frac{\lambda^i}{i!}} e^{-\lambda} = 1.$$

Bez.: $X \sim Poi(\lambda)$.

Satz 9 *Seien* $X_n \sim Bi(n,p)$, $Y \sim Poi(\lambda)$

Für $\quad n \cdot p = \lambda$ *gilt:*

$$P(X_n = k) \longrightarrow_{n\to\infty} P(Y = k).$$

Beweis: Es gilt:

$$
\begin{aligned}
P(X_n = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
&= \frac{n(n-1)\cdots(n-k+1)}{k!} (\frac{\lambda}{n})^k (1 - \frac{\lambda}{n})^{n-k} \\
&= \frac{n(n-1)\cdots(n-k+1)}{k!(n-\lambda)^k} \frac{(n-\lambda)^k \lambda^k}{n^k} \frac{(n-\lambda)^{n-k}}{n^{n-k}} \\
&= \frac{1}{k!} \underbrace{\frac{n(n-1)\cdots(n-k+1)}{(n-\lambda)^k}}_{\longrightarrow 1} \lambda^k \underbrace{(1 - \frac{\lambda}{n})^n}_{\longrightarrow e^{-\lambda}}
\end{aligned}
$$

λ fest

k fest

$$
= \frac{\lambda^k}{k!} e^{-\lambda} = P(Y = k)
$$

$\square$

d) Münzwurf solange bis B(Blatt) kommt

Ω = {B, ZB, ZZB, ...}

X := Anzahl der Würfe bis zum ersten Blatt.

$$X = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n & \cdots \\ \frac{1}{2} & (\frac{1}{2})^2 & (\frac{1}{2})^3 & (\frac{1}{2})^4 & \cdots & (\frac{1}{2})^n & \cdots \end{pmatrix}$$

$$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} (1/2)^i = \frac{1}{1-\frac{1}{2}} - 1 = 1$$

geometrische Reihe

geometrische Verteilung mit p=1/2, $\quad p_i = (1/2)^i.$

Def. 17 *: Eine Zufallsvariable X mit*

$$P(X = i) = p(1-p)^{i-1}, \quad i = 1, 2, \ldots$$

heißt geometrisch verteilt.

Bez.: $X \sim Geo(p)$.

Anzahl der Schritte bis zum ersten “Erfolg”.

e) Qualitätskontrolle

Gegeben sei eine Grundgesamtheit (z.B. eine Warenlieferung) mit N Stücken, von denen genau n schlecht seien. Wie groß ist die Wkt., daß in einer Stichprobe vom Umfang m höchstens k Stück schlecht sind?

X: zufällige Anzahl der schlechten Stücke in der Stichprobe.

$$P(X = x) = \frac{\binom{n}{x} \cdot \binom{N-n}{m-x}}{\binom{N}{m}}$$

$\binom{N}{m}$: $\#$ möglichen Stichproben.

$\binom{n}{x}$: $\#$ Möglichkeiten, aus n schlechten Stücken in der Grundgesamtheit x schlechte Stücke zu ziehen.

$\binom{N-n}{m-x}$: $\#$ Möglichkeiten, aus $N-n$ guten Stücken in der Grundgesamtheit $m-x$ gute Stücke zu ziehen.

Offenbar:

$$0 \leq x \leq \min(n, m)$$
$$m - x \leq N - n.$$

Def. 18 *Eine Zufallsvariable mit der Wkt.funktion*

$$f(x|H_{N,n,m}) = P(X = x)$$

heißt hypergeometrisch verteilt.

Bez.: $X \sim H_{N,n,m}$. Verteilungsfunktion:

$$F(k|H_{N,n,m}) = \sum_{x=0}^{k-1} \frac{\binom{n}{x} \cdot \binom{N-n}{m-x}}{\binom{N}{m}}$$

Bemerkung: Für $N \to \infty$, $n \to \infty$, $\frac{n}{N} \to p$ gilt:

$$f(x|H_{N,n,m}) \to \binom{m}{x} p^x (1-p)^{m-x} = f(x|Bi(m,p))$$

ÜA.

Wkt. für höchstens k schlechte Stücke: $F(k+1|H_{N,n,m})$.

Weitere Eigenschaften und Anwendungen diskreter Zufallsvariablen

5.3.1 Binomialverteilung

Bsp. 37 (Kommunikationskanal) *Schicken Binärzahlen durch einen Kommunikationskanal.*

p*: Wkt. einer fehlerhaften Übertragung*

n*: Anzahl der übertragenen Zeichen*

Wkt. für genau i *Fehler:*

$$P(i) = \binom{\mathrm{N}}{i} p^i (1-p)^{n-i} =: b(i; n, p)$$

Bsp. 38 (Qualitätskontrolle) *Stichprobe von 10 Computerchips aus einer sehr großen Lieferung (Los). Wenn keine defekt, so wird die Lieferung angenommen, sonst nicht.*

p*: Wkt., ein zufällig ausgewählter Chip ist defekt.*

Wkt. für genau i *defekte Stücke =* $b(i; 10, p)$.

$$P(\textit{Los angenommen}) = (1-p)^{10}$$

Bsp. 39 (k aus n Systeme) *Jede Komponente habe die Intaktwkt. p.*

Wkt., daß genau i Komponenten ausfallen:

$$P(X = i) = \binom{n}{i} p^{n-i}(1-p)^i$$

Wkt., daß höchstens k Komponenten ausfallen:

$$\begin{aligned} P(X \le k) &= \sum_{i=0}^{k} \binom{n}{i} p^{n-i}(1-p)^i \\ &= \sum_{i=n-k}^{n} \binom{n}{i} p^{i}(1-p)^{n-i} \end{aligned}$$

5.3.2 Geometrische Verteilung

Bem. 2 *Sei $Y \sim Geo(p)$, d.h.*

$$
\begin{aligned}
P(Y > s) &= 1 - \sum_{i=1}^{s} (1-p)^{i-1} \cdot p = (1-p)^s \\
P(Y > t) &= 1 - \sum_{i=1}^{t} (1-p)^{i-1} \cdot p = (1-p)^t
\end{aligned}
$$

$$
\begin{aligned}
P(Y > s) \cdot P(Y > t) &= (1-p)^{s+t} \\
&= 1 - \sum_{i=1}^{s+t} (1-p)^{i-1} \cdot p \\
&= P(Y > s + t).
\end{aligned}
$$

also:

$$\begin{aligned} P(Y > s+t|Y > t) &= \frac{P(Y > s+t, Y > t)}{P(Y > t)} \\ &= \frac{P(Y > s+t)}{P(Y > t)} \\ &= P(Y > s) \end{aligned}$$

Bez. 5 *Wir sagen, Verteilungen mit*

$$P(Y > s+t|Y > t) = P(Y > s)$$

besitzen die sogenannte Markov-Eigenschaft oder sie sind gedächtnislos.

Satz 10 *Sei X diskrete Zufallsvariable mit Werten in $\{1, 2, 3, \ldots\}$ und X habe die Markov-Eigenschaft. Dann ist $X \sim Geo(p)$ für ein $p, p \in (0, 1)$*

Beweis: : Sei

$$X : \begin{pmatrix} 1 & 2 & 3 & \ldots \\ p_1 & p_2 & p_3 & \ldots \end{pmatrix}$$

Aus der Markov-Eigenschaft folgt:

$$\begin{aligned} P(X > s) \cdot P(X > t) &= P(X > s + t) \quad \forall s, t \\ (1 - \sum_{i=1}^{s} p_i)(1 - \sum_{i=1}^{t} p_i) &= 1 - \sum_{i=1}^{s+t} p_i \end{aligned}$$

Setzen $p := p_1$. Einsetzen von

$s = 1, t = 1$ liefert $(1-p)^2 = (1-p-p_2)$; $\quad p_2 = p(1-p)$.

$s = 1, t = 2$ liefert $(1-p)(1-p-p_2) = (1-p-p_2-p_3)$;

$$(1-p-p_2)(1-p-1) = -p_3; \quad \text{also}$$

$$p_3 = p(1-p)^2 \quad \text{usw.} \qquad \square$$

Bsp. 40 (Qualitätskontrolle) *Wkt., daß das i-te Item das erste defekte ist.*

Bsp. 41 (Time-sharing computer system) *mit festen Zeitscheiben.*

Programm wird in der Zeitscheibe vollständig abgearbeitet
mit Wkt. p

Wenn nicht, neuer Versuch in der neuen Zeitscheibe

X: # *benötigten Zeitscheiben*

$$X \sim Geo(p).$$

Bsp. 42 (Repeat-Schleife) A*: aussagenlogischer Ausdruck,* $A = true$ *mit Wkt.* p*.*

`repeat` S `until` A.

$\#$ *der Durchläufe von S:* $\sim Geo(p)$.