Weitere Eigenschaften und Anwendungen diskreter Zufallsvariablen

Binomialverteilung

Bsp. 2.7 (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. 2.8 (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. 2.9 (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:

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

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 2.3 *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 $(1-p)^2 = (1-p-p_2)$; $p_2 = p(1-p)$.
$s = 1, t = 2$ liefert $(1-p)(1-p-p_2) = (1-p-p_2-p_3)$;
$(1 - p - p_2)(1 - p - 1) = -p_3$; also $p_3 = p(1 - p)^2$ usw.

$\square$

Bsp. 2.10 (Qualitätskontrolle) *Wkt., daß das i-te Item das erste defekte ist.*

Bsp. 2.11 (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. 2.12 (Repeat-Schleife) A: *aussagenlogischer Ausdruck,* $A = true$ *mit Wkt.* p.

`repeat` S `until` A.

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

Poisson-Verteilung

Sei $\{N_t\}_{t\in T}$ eine Menge von Zufallsvariablen (ein stochastischer Prozeß) mit folgenden Eigenschaften:

V1: Zuwächse sind unabhängig, dh. die Zufallsvariablen

$N_{t+h} - N_t$ und $N_t - N_{t-h}$ sind unabhängig [1]

V2: es ist egal wo wir Zeitintervall betrachten, dh.

N_{t+h} und N_t haben dieselbe Verteilung

V3: Wkt., daß mindestens ein Ereignis in der Zeit h

eintritt, z.B. ein Kunde ankommt.

$p(h) = a \cdot h + o(h), \quad a > 0, h \to 0$

V4: Wkt. für ≥ 2 Ereignisse in der Zeit h: $o(h)$

Anmerkung: Zwei Zufallsvariablen X und Y heißen unabhängig, falls

$$\forall A, B \in \mathcal{B}; \quad P(X \in A, Y \in B) = P(X \in A) \cdot P(X \in B)$$

Frage: Wkt. daß bis zum Zeitpunkt t k Ereignisse eintreten? (eingetroffene Kunden, zerfallene Teilchen)

$$P_k(t) := P(N_t = k), \qquad P_k(t) = 0 \quad \text{für} \quad k < 0$$

$$p(h) = \sum_{k=1}^{\infty} P_k(h) \qquad \geq 1\text{Ereignis tritt ein}$$

Offenbar:

$$1 = \sum_{k=0}^{\infty} P_k(t)$$

$$V3 \Rightarrow P_0(h) = 1 - p(h) = 1 - ah + o(h)$$

$$V4 \Rightarrow \sum_{k=2}^{\infty} P_k(h) = o(h), \qquad (h \to 0)$$

1. Schritt: Bestimmen $P_0(t)$.

$$\begin{aligned}
P_0(t+h) &= P(N_t = 0, N_{t+h} - N_t = 0) \\
&= P_0(t)P(N_{t+h} - N_t = 0) \quad \text{wegen V1} \\
&= P_0(t)P(N_h - N_0 = 0) \quad \text{wegen V2} \\
&= P_0(t)P_0(h) \quad \text{wegen} \quad N_0 = 0 \\
&= P_0(t)(1 - p(h)) \\
&= P_0(t)(1 - ah + o(h)) \quad \text{wegen V4}
\end{aligned}$$

Nacheinander folgt:

$$\frac{P_0(t+h) - P_0(t)}{h} = P_0(t)(-a + \frac{o(h)}{h})$$

$$P_0'(t) = -aP_0(t)$$

$$P_0(t) = ce^{-at}$$

Wegen $P_0(0) = 1$ folgt: $c = 1$ und

$$P_0(t) = e^{-at}$$

2. Schritt: Bestimmen $P_k(t)$.

Zerlegen das Ereignis $\{N_{t+h} = k\}$ in disjunkte Teilereignisse.

$$\begin{aligned}
\{N_{t+h} = k\} &= \{N_t = 0, N_{t+h} - N_t = k\} \cup \\
&= \{N_t = 1, N_{t+h} - N_t = k-1\} \cup \\
&= \{N_t = 2, N_{t+h} - N_t = k-2\} \cup \ldots \cup \\
&= \{N_t = k, N_{t+h} - N_t = 0\}
\end{aligned}$$

$$\begin{aligned}
P_k(t+h) &= \sum_{j=0}^{k} P(N_t = k-j, N_{t+h} - N_t = j) \\
&= \sum_{j=0}^{k} P_{k-j}(t) \underbrace{P(N_{t+h} - N_t = j)}_{=P(N_h - N_0 = j)} \quad \text{wegen V1} \\
&= \sum_{j=0}^{k} P_{k-j}(t) P_j(h) \quad \text{wegen V2} \\
&= P_k(t)P_0(h) + P_{k-1}(t)P_1(h) + \sum_{j=2}^{k} P_{k-j}(t)P_j(h)
\end{aligned}$$

$$
\begin{aligned}
P_1(h) &= \sum_{j=1}^{\infty} P_j(h) - \sum_{j=2}^{\infty} P_j(h) \\
&= p(h) + o(h) \\
&= ah + o(h)
\end{aligned}
$$

$$
\sum_{j=2}^{\infty} P_{k-j}(t) P_j(h) \leq \sum_{j=2}^{\infty} P_j(h) = o(h) \quad \text{wegen V2}
$$

Nacheinander folgt:

$$
\begin{aligned}
P_k(t+h) - P_k(t) &= (P_0(h) - 1) P_k(t) + P_{k-1}(t) P_1(h) + o(h) \\
&= -ahP_k(t) + ahP_{k-1}(t) + o(h) \\
\frac{P_k(t+h) - P_k(t)}{h} &= -aP_k(t) + aP_{k-1}(t) + \frac{o(h)}{h}
\end{aligned}
$$

$$
P_k'(t) = -aP_k(t) + aP_{k-1}(t), \quad P_k(0) = 0
$$

$$Q_k(t) := P_k(t)e^{at}$$

$\Rightarrow$

$$\begin{aligned}
Q_k'(t) &= P_k'(t)e^{at} + P_k(t)ae^{at} \\
Q_k'(t) &= e^{at}(\underbrace{-aP_k(t) + aP_{k-1}(t)}_{P_k'(t)} + aP_k(t)) \\
&= aQ_{k-1}(t) \\
Q_1'(t) &= aQ_0(t) = ae^{-at}e^{at} = a \Rightarrow Q_1(t) = at \\
Q_2'(t) &= aQ_1(t) = a^2t \Rightarrow Q_2(t) = \frac{a^2t^2}{2}
\end{aligned}$$

Durch vollständige Induktion:

$$Q_k(t) = \frac{a^kt^k}{k!}$$

$$P_k(t) = \frac{a^kt^k}{k!}e^{-at}$$

Poisson-Verteilung mit Parameter $\lambda = at$.

Programme:

`Descr_Binomial_neu.sas`

`Descr_Poisson.sas`

`Descr_Geometr.sas`

`Descr_Hypergeom.sas`

Bem: In den Wahrscheinlichkeiten können Parameter auftreten, die in der Regel unbekannt sind.

Die Parameter sind anhand der Beobachtungen (der Daten) zu bestimmen/zu <u>schätzen</u>!
$\longrightarrow$ Aufgabe der Statistik