K3 (Diskrete) Zufallsvariablen

3.1 Basis

$\Omega = \{\omega\}$, $X(\omega)$ ist eine Größe die durch ω bestimmt ist. Bei der zufälligen Auswahl von ω bekommen wir den Wert, $X(\omega)$.

Definition: Ist (Ω, F, P) ein Wahrscheinlichkeitsraum und $\mathcal{X}$ eine beliebige Menge, so nennen wir eine Abbildung
$X: \Omega \rightarrow \mathcal{X}$ eine $\mathcal{X}$-wertige Zufallsvariable

Falls Ω abzählbar ist, wird der Wertebereich von X, $\{X(\omega): \omega \in \Omega\}$ auch abzählbar sein. Die Verteilung von X ist das Wahrscheinlichkeitsmaß

$$P_X(x) = P(\{\omega \in \Omega: X(\omega) = x\})$$

3.2 Erwartungswerte

Die Erwartung einer reellwertigen Zufallsvariablen X existiert wenn $\sum |X(\omega)| P(\omega)$ konvergiert. Wir definieren dann

$E[X] = \sum X(\omega)P(\omega)$
als den Erwartungswert von X

Ist $x_1, x_2, x_3, \ldots$ eine Abzählung des Wertebereichs von X, so ist

$E[X] = \sum x_i P(X = x_i)$

Eigenschaften:
(1) $E[X+a] = E[X] + a$
(2) $E[X+Y] = E[X] + E[Y]$
(3) $E[aX] = aE[X]$
(4) sind X, Y unabhängig:
$E[X*Y] = E[X]*E[Y]$
(5) $E[f(X)] \neq f(E[X])$

3.2.1 Warum brauchen wir Konvergenz?

(Das Paradox von St. Petersburg)

Man wirft eine Münze bis Kopf auftritt. Wenn das zum ersten Mal mit dem r.ten Wurf geschieht, gewinnt man

$$€\ 2^r$$

Nehmen wir an, daß P(Kopf) = p

$$\begin{aligned} E[\text{Gewinn}] &= \textstyle\sum_{r=1}^{\infty} 2^r p(1-p)^{r-1} \\ &= 2p \textstyle\sum_{r=1}^{\infty} (2q)^{r-1} \\ &= 2p \textstyle\sum_{r=0}^{\infty} (2q)^{r} \\ &= 2p/(1-2q) \quad \text{für } 2q < 1 \\ &\quad \text{sonst} \quad \infty \end{aligned}$$

3.3 Verteilungen auf den Ganzzahlen

$\Omega = \{0,1,2, ...\}$

$P(X = k) = p_k$

Man interessiert sich insbesondere für

$P(X = 0) = p_0$

$P(X \geq k) = \sum_{i \geq k} p_i$

$P(X = r \mid X \geq k) = p_r \,/ \sum_{i \geq k} p_i$

$P(j \leq X \leq k) = \sum^{k}_{i=j} p_i$

$E[\,X\,] = \sum i p_i$ (arithmetischer Mittel)

$V[\,X\,] = \sum (i - E[\,X\,])^2 \, p_i$

Standardabweichung von $X = \sqrt{V[X]}$

z.B.

k	$\mathbf{p}_k$
0	0.2
1	0.4
2	0.3
3	0.1
≥ 4	0.0

$P(X = 0) =$

$P(X \geq 2) =$

$P(X = 3 \mid X \geq 2) =$

$P(1 \leq X \leq 2) =$

$E[X] =$

$V[X] =$

3.4 Hypergeometrische Verteilung

Eine Urne enthält
S schwarze und W weiße Kugeln.

Es werden $n \leq S + W$ Kugeln ohne Rücklegen gezogen.

Die Wahrscheinlichkeit, daß die Stichprobe genau s schwarze Kugel enthält ist

$h(s; n, N, S) =$

wo
$N = S + W$, $n = s + w$ und $0 \leq s \leq \min(n, S)$

Für eine Urne mit mehreren Farben, z. B.
S schwarze, W weiße, R rote, B blaue Kugeln gilt:

$P(s, w, r, b) =$

3.5 Produktexperimente

Wir haben eine Reihe von unabhängigen Experimenten {1, 2, 3, ..., i, ...} mit Modellen (Ω_i, F_i, P_i). Das Gesamtergebnis hat einen Wahrscheinlichkeitsraum (Ω, F, P) mit

$$\Omega = \prod \Omega_i \, , \; F = \prod F_i \, , \; P = \prod P_i$$

das Produkt der Wahrscheinlichkeitsräume

Man zieht z.B. 4 Kugeln hintereinander aus einer Urne mit 10 schwarze Kugeln und 20 weiße Kugeln.

(a) mit Rücklegen
   alle 4 Experimente haben denselben Wahrscheinlichkeitsraum (Ω_i, F_i, P_i)

(b) ohne Rücklegen
   das Resultat der zweiten Ziehung ist vom Resultat der ersten abhängig und somit ist (Ω_2, F_2, P_2) im voraus nicht bestimmbar.

Produktexperimente — Beispiele

Geometrische Verteilung

In jedem Experiment gibt es zwei mögliche Ausgänge, {1, 0} mit Wahrscheinlichkeiten p bzw. q (= 1-p).

Wie groß ist die Wahrscheinlichkeit, daß k Nullen vor dem ersten Einser auftreten?

$$p_k = pq^k \quad k = 0, 1, 2, \ldots$$

Binomiale Verteilung

Es gibt genau N der obigen Experimente. Wie groß ist die Wahrscheinlichkeit, daß man k Einsen bekommt?

$$[N!/(k!\,(N - k)!)]\; p^k(1 - p)^{N-k}$$

Negative Binomiale Verteilung

Es gibt keine Begrenzung der Zahl der Experimente. Wie groß ist die Wahrscheinlichkeit, daß man k Nullen vor dem r.ten Einser bekommt?

$$[(k + r - 1)!/(k!\,(r - 1)!)]\; p^r(1 - p)^k$$

3.6 Geometrische Verteilung

3.6.1 Formeln

$$p_k = pq^k \quad (q = 1 - p) \tag{3.6.1}$$

$$p_0 = p$$

$$P(X \geq k) = \sum_{i\geq k} p_i = pq^k \sum_{i\geq 0} q^i = q^k$$

$$P(X = r \mid X \geq k) = p_r / \sum_{i\geq k} p_i = pq^{r-k}$$

$$P(j \leq X \leq k) = \sum_{i=j}^{k} p_i = q^j (1 - q^{k+1-j})$$

$$E[\,X\,] = \sum_{i\geq 0} ip_i = pq\sum_{i\geq 1} iq^{i-1} = q/p$$

$$V[\,X\,] = \sum_{i\geq 0} (i - E[\,X\,])^2 p_i = q/p^2$$

3.6.2 Anwendungen — Geometrische Verteilung

1) Zuverläßigkeitsmodelle

Wieviele Backups?

P(Absturz) = p

P(k stürzen ab) = p^k

Unabhängigkeit ist eine wichtige Annahme

2) „Coupons“ Problem

P(man bekommt den seltenen Fall) = p

P(k + 1 kaufen müssen) = $q^k p$

3) Zeit zum ersten Erfolg

3.7 Binomialverteilung $B(N, p)$

3.7.1 Formeln

$$p_k = N!/(k!\,(N-k)!)\, p^k q^{N-k} \qquad (3.7.1)$$

$$p_0 = q^N$$

$$P(X \geq k) = \sum_{i \geq k} p_i$$

$$P(X = r \mid X \geq k) = p_r \,/ \sum_{i \geq k} p_i$$

$$P(j \leq X \leq k) = \sum_{i=j}^{k} p_i$$

$$E[\,X\,] = \sum_{i \geq 0} i p_i = Np$$

$$V[\,X\,] = \sum_{i \geq 0} (i - E[\,X\,])^2\, p_i = Npq$$

(Bernoulli Experimente)
Für N unabhängige Zufallsvariablen $\{X_i\}$, wo $P(X_i = 1) = p$ und $P(X_i = 0) = 1 - p$, hat $X = \sum X_i$ eine Binomialverteilung.

3.7.2 Anwendungen — Binomialverteilung

1) Qualitätskontrolle
Stichproben der Größe N. Jedes Produkt hat die gleiche Wahrscheinlichkeit, p, fehlerhaft zu sein.

2) „Multiple Choice“
N Fragen
r Antworte pro Frage P(richtig) = 1/r
Die Verteilung der #(richtige Antworte)

$$B(N, 1/r)$$

3) „ESP“
N Versuche p = P(richtig)

$$P(X \geq k) = \sum_{i=k}^{N} N!/[i!(N-i)!]p^i q^{N-i} = g_k$$

M Personen $p = g_k$

$$P(\text{mehr als } K \text{ Personen} \geq k \text{ richtig haben}) = \sum_{j>K}^{M} M!/[j!(M-j)!]\, g_k^{\,j}(1 - g_k)^{M-j}$$

3.8 Negative Binomialverteilung

3.8.1 Formeln

$p_k = (r + k - 1)! / (k!\,(r - 1)!)\ p^r q^k$

$k = 0, 1, 2, \ldots$

$p_0 = p^r$

$P(X \geq k) = \sum_{i\geq k} p_i$

$P(X = r \mid X \geq k) = p_r / \sum_{i\geq k} p_i$

$P(j \leq X \leq k) = \sum^k_{i=j} p_i$

$E[\,X\,] = \sum_{i\geq 0} i p_i = rq/p$

$V[\,X\,] = \sum_{i\geq 0} (i - E[\,X\,])^2\, p_i = rq/p^2$

(Die Summe von r geometrischverteilte Zufallsvariablen)

3.8.2 Anwendungen — Negative Binomialverteilung

1) Zählvariablen
   - Tore
   - Einkäufe
   - Polygynie in Afrika
   - …

2) Versuche / Anrufe / Einkäufe ... bis
   - alle Objekte verkauft sind;
   - genug Erfolge erzielt werden;
   - eine Stichprobengruppe gefunden wird.