6 Kombinatorik: Einschluß-Ausschluß Formel

6.1 Indikatorfunktionen

$$I_A(\omega) = \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{falls } \omega \notin A \end{cases}$$

I_A ist eine Zufallsvariable

$$E[I_A] = P(A)$$

$$I_{\bar{A}} = 1 - I_A$$

$$I_{A \cap B} = I_A I_B$$

$$I_A^2 = I_A$$

$$Var[I_A] = P(A)(1 - P(A))$$

$$A \subset B \Rightarrow I_A \leq I_B$$

$$A \cap B = \emptyset \Rightarrow I_{A \cup B} = I_A + I_B$$

6.2 Wahrscheinlichkeiten von m Ereignissen

Gegeben m beliebige Ereignisse, $\{A_i\}$,
was ist die Wahrscheinlichkeit, dass:

- keine stattfinden?
- genau n stattfinden?
- mindestens n stattfinden?

Beliebig bedeutet,
dass wir Unabhängigkeit NICHT annehmen dürfen.

6.3 Einschluß-Ausschluß Formel

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

folgt aus den Axiomen von Kolmogoroff (§1.5.1).
Weiter können wir zeigen, dass

$$P(A_1 \cup A_2 \cup A_3) =$$

$$\sum_i P(A_i) - \sum_{j>i} P(A_i \cap A_j) + P(A_1 \cap A_2 \cap A_3)$$

Für den allgemeinen Fall mit m beliebigen Ereignissen $\{A_i\}$ brauchen wir

$$S_k = \sum_{\{i_1, i_2, \ldots i_k\} \subset \{1,2 \ldots m\}} P(A_{i_1} \cap A_{i_2} \ldots \cap A_{i_k})$$

die Summe der Wahrscheinlichkeiten von allen Durchschnitten, die aus genau k der m Ereignissen gebildet werden können. Es lässt sich für alle m zeigen:

$$P(\bigcup_{i=1}^{m} A_i) = S_1 - S_2 + S_3 - .... + (-1)^{m+1} S_m \quad (1)$$

$1 - P(\bigcup_{i=1}^{m} A_i)$ ist die Wahrscheinlichkeit, dass keins von den m Ereignissen stattfindet.

Beweis von der Einschluß-Ausschluß Formel

Sei B das Ereignis $\bigcup_{i=1}^{m} A_i$ dann ist

$$I_B = 1 - \prod_{i=1}^{m} (1 - I_{A_i})$$

Die rechte Seite wird ausmultipliziert und dann werden Erwartungswerte auf beiden Seiten gebildet.

6.3.1 Anwendung der Einschluß-Ausschluß Formel

k Kugeln werden auf n Fächer verteilt. Jede Kugel kann mit gleicher Wahrscheinlichkeit in jedes Fach kommen. Was ist $P(\text{mindestens 1 Fach leer bleibt})$?
Sei A_i das Ereignis, dass Fach i leer bleibt.

$$P(A_i) = \left(\frac{n-1}{n}\right)^k$$

$$P(A_{i_1} \cap A_{i_2} \ldots A_{i_m}) = \left(\frac{n-m}{n}\right)^k$$

$$S_m = \left(\frac{n-m}{n}\right)^k \binom{n}{m}$$

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{m=1}^{n} (-1)^{1+m} S_m$$

z.B.

$$n = 10, \quad k = 20 \quad \Rightarrow p =$$

$$n = 20, \quad k = 50 \quad \Rightarrow p =$$

6.4 Die Formeln von Waring

Sei (Ω, F, P) ein Wahrscheinlichkeitsraum und $A_1, \ldots A_N$ Ereignisse. Nehmen wir an, dass für beliebige $\{i_1, i_2, \ldots, i_k\}$ sei $P(A_{i_1} \cap A_{i_2} \ldots \cap A_{i_k})$ bekannt.

$$B_n = \{\omega \in \Omega; \omega \in A_k \text{ für genau } n \text{ Werte von } N\}$$

$$C_n = \{\omega \in \Omega; \omega \in A_k \text{ für mindestens } n \text{ Werte von } N\}$$

$$\begin{aligned} P(B_n) &= \sum_{k=n}^{N} (-1)^{k-n} \binom{k}{n} S_k && (1) \\ P(C_n) &= \sum_{k=n}^{N} (-1)^{k-n} \binom{k-1}{n-1} S_k && (2) \end{aligned}$$

$P(C_1)$ ist die Einschluß-Ausschluß Formel.

6.4.1 Beweis für $P(B_n)$

$$I_{B_n} = \sum_{\{i_1,i_2,\ldots,i_n\}} I_{A_{i_1}} I_{A_{i_2}} \ldots I_{A_{i_n}} (1 - I_{A_{i_{n+1}}}) \ldots (1 - I_{A_{i_N}})$$

$$= \sum_{\{i_1,i_2,\ldots,i_n\}} I_{A_{i_1}} \ldots I_{A_{i_n}} \sum_{l=0}^{N-n} (-1)^l \sum_{\{j_1,j_2,\ldots,j_l\}} I_{A_{j_1}} \ldots I_{A_{j_l}}$$

Es wird über alle $\{j_1, j_2, \ldots, j_l\} \subset \{i_{n+1}, \ldots, i_N\}$ summiert; dabei ist für $l = 0$ $\{j_1, j_2, \ldots, j_l\}$ die leere Menge und $I_{A_{j_1}} \ldots I_{A_{j_l}} = 1$.

$$= \sum_{l=0}^{N-n} (-1)^l \sum_{\{i_1,\ldots,i_n\}} \sum_{\{j_1,\ldots,j_l\}} I_{A_{i_1}} \ldots I_{A_{i_n}} I_{A_{j_1}} \ldots I_{A_{j_l}}$$

Es gibt $\binom{n+l}{n}$ Möglichkeiten, eine Menge $\{h_1, \ldots h_{n+l}\}$ in disjunkte Teilmengen $\{i_1, \ldots, i_n\}$ und $\{j_1, j_2, \ldots, j_l\}$ zu zerlegen.

$$= \sum_{l=0}^{N-n} (-1)^l \sum_{\{h_1,\ldots,h_{n+l}\}} \binom{n+l}{n} I_{A_{h_1}} \ldots I_{A_{h_{n+l}}}$$

Aus $P(B_n)$=$E[I_{B_n}]$ und der Additivität des Erwartungswerts folgt das Resultat: Setzen wir $k = n + l$ und

$$P(B_n) = \sum_{k=n}^{N} (-1)^{k-n} \binom{k}{n} S_k$$

6.4.2 Beweis für $P(C_n)$

Die Formel für $P(C_n)$ folgt nach Induktion.
Für $n = N$ ist $P(C_n) = P(B_n) = S_N$
Für $n \leq N$ gilt

$$P(C_{n-1}) = P(B_{n-1}) + P(C_n)$$

$$= \sum_{k=n-1}^{N} (-1)^{k-n+1} \binom{k}{n-1} S_k + \sum_{k=n}^{N} (-1)^{k-n} \binom{k-1}{n-1} S_k$$

$$= S_{n-1} + \sum_{k=n}^{N} (-1)^{k-n+1} S_k [\binom{k}{n-1} - \binom{k-1}{n-1}]$$

Aus

$$\binom{k}{n-1} - \binom{k-1}{n-1} = \binom{k-1}{n-2}$$

folgt das Resultat.

6.4.3 Anwendung von Waring: Matching

z.B. N Briefe und N dazu gehörende Umschläge

Was ist die Wahrscheinlichkeit bei einer zufälligen Verteilung der Briefe an die Umschläge, dass genau n Briefe die richtigen Umschläge bekommen?

Sei A_{k_1} das Ereignis, dass Brief k_1 den richtigen Umschlag bekommt:

$$P(A_{i_1} \cap A_{i_2} \ldots \cap A_{i_k}) = \frac{(N-k)!}{N!}$$

Wir suchen $\quad P(B_n) = \sum_{k=n}^{N} (-1)^{k-n} \binom{k}{n} S_k$

$$S_k = \binom{N}{k} \frac{(N-k)!}{N!} = \frac{1}{k!}$$

$$P(B_n) = \sum_{k=n}^{N} (-1)^{k-n} \frac{1}{n!} \frac{1}{(k-n)!}$$

$$= \frac{1}{n!} \sum_{l=0}^{N-n} (-1)^l \frac{1}{l!}$$

$$\text{für} \quad N \to \infty, \quad P(B_0) \to e^{-1} \doteqdot 0.368$$

und die Anzahl Briefe mit dem richtigen Umschlag $\sim P(1)$

6.5 Diskrete Verteilungen — Rekapitulation

H: Hypergeometrische — ziehen ohne Rücklegen

B: Binomial — Wieviele Treffer aus N?

G: Geometrische — Wieviele Versuche bis Erfolg?

NB: Negative Binomial – Wieviele Versuche zum r. Erfolg?

P: Poisson — Wieviele Ereignisse?

B, G, NB, P — alle nehmen Unabhängigkeit an.

und Spezialfälle wie Benford's Gesetz:

Die Verteilung der ersten Ziffer in Zahlen ist

$$P(\text{erste Ziffer } = n) = \log_{10}(1 + \frac{1}{n}) \quad n = 1, \ldots 9$$

6.6 Anwendungen

Epidemiologie

In der Umgebung von 10 Kernkraftwerken werden je 100 (unabhängig ausgewählte) Personen auf eine bestimmte Krankheit hin untersucht, die im Bundesdurchschnitt bei 1% der Bevölkerung vorkommt.

Wenn die Erkrankungswahrscheinlichkeit in der Umgebung der Werke auch 1% ist, was ist:

$$P(> 3 \text{ kranke Personen bei keinem Werk}) =?$$

Wenn die Erkrankungswahrscheinlichkeit in der Umgebung der Werke 2% ist, was ist:

$$P(> 3 \text{ kranke Personen bei keinem Werk}) =?$$

(Was heißt in der Umgebung? Was heißt krank?
Könnte Alter ein Faktor sein?....)

Banachs Streichholzproblem

Ein Mann hat zwei Schachteln mit Streichhölzern. Er bedient sich mit gleicher Wahrscheinlichkeit links oder rechts. Wenn am Anfang beide N enthalten, wieviele übrig bleiben, wenn der erste Schachtel leer ist?

Gepäckverlust

Sie fliegen von Augsburg nach Frankfurt, von Frankfurt nach Toronto und von Toronto nach San Francisco. An jedem Flughafen muß Ihre Koffer verladen werden und wird mit Wahrscheinlichkeit p fehlgeleitet. Bei Ankunft in San Francisco stellen Sie fest, dass Ihr Koffer verloren ist. Berechnen Sie die Wahrscheinlichkeit dafür, dass er in Frankfurt fehlgeleitet wurde.

Entschlossene Minderheiten

An einer Wahl zwischen zwei Kandidaten A und B nehmen N Wähler teil. Davon stimmen m Wähler geschlossen für A, aber die übrigen sind unentschlossen und treffen ihre Meinung unabhängig voneinander, als ob sie eine faire Münze geworfen hatten. Was ist $P(A \text{ gewinnt })$?

Wieviele Patienten?

1. Wir suchen r Patienten mit einem bestimmten Krankheitsmuster für eine Studie. Wieviele Patienten müssen wir untersuchen, bis wir r finden?

2. Wir haben nur Zeit N Patienten zu untersuchen. Wieviele mit dem Krankheitsmuster werden wir finden?

3. Im Krankenhaus gibt es R Patienten mit dem Krankheitsmuster und S ohne, insgesamt $R + S$ Patienten. Wenn eine Stichprobe der Größe n aus dieser Grundgesamtheit gezogen wird, wieviele Patienten mit dem Krankheitsmuster werden darin enthalten sein?

4. Jede Woche kommen im Durchschnitt drei neue Patienten mit dem Krankheitsmuster an. Was ist die Wahrscheinlichkeit, dass mehr als N solche Patienten innerhalb von w Wochen ankommen?

Wieviele Fische?
n Fische werden gefangen und markiert. Eine Woche später werden m Fische gefangen. Was ist die Wahrscheinlichkeit, dass y davon markiert sind, wenn es N insgesamt gibt?

Wieviele Flugkarten sollen verkauft werden? Fluggesellschaften verkaufen mehr Karten als Sitze vorhanden sind, weil nicht alle Passagiere erscheinen. Nehmen Sie an, dass die Firma Einnahmen von a bei jedem mitfliegenden Fluggast und einen Verlust von b $(b > a)$ für jede überzählige Person hat. Wenn ein Flugzeug N Plätzte hat und jede Person, die gebucht hat, unabhängig von anderen mit Wahrscheinlichkeit p zum Flug erscheint, wieviele Karten sollen verkauft werden?

Essen im Flugzeug Es werden zwei Menüs im Flugzeug angeboten. Um Geld zu sparen werden für 200 Leute genau 100 von jedem Menü mitgenommen. Wieviele Fluggäste werden unzufrieden sein?