

Stochastik

—

Zusammenfassung & Fragenkatalog

Jan Wiefel

5. November 2007

Inhaltsverzeichnis

1	Zufallsexperiment	5
2	Ereignis	5
3	Zufallsvariablen	5
4	Relative Häufigkeiten	6
5	Grundbegriffe der deskriptiven Statistik	7
5.1	Empirische Häufigkeitsverteilung	7
5.2	Darstellungsformen	7
5.3	Relative Klassenhäufigkeiten	8
5.4	Kennzahlen für Lagemaße	8
5.4.1	Arithmetisches Mittel	8
5.4.2	Empirischer Median, Zentralwert	8
5.4.3	Getrimmtes Mittel	8
5.5	Streuungsmaße	9
5.5.1	empirische Varianz	9
5.5.2	Medianabweichung	9
5.5.3	Quartilsabstand	9
6	Endliche Wahrscheinlichkeitsräume	11
6.1	Verteilung einer Zufallsvariablen	11
6.2	Subjektive Wahrscheinlichkeiten	12
7	Laplace-Modelle	12
8	Grundlagen der Kombinatorik	13
8.1	Fundamentalprinzip des Zählens (Multiplikationsregel)	13
8.2	Permutationen und Kombinationen	13
9	Urnen und Teilchen/Fächer-Modelle	14
9.1	Urnenmodelle	14
9.2	Teilchen/Fächer-Modelle	14
9.3	Zusammenfassung	14
10	Das Paradoxon der ersten Kollision	15
11	Die Formel des Ein- und Ausschließens	16
11.1	Formel des Ein- und Ausschließens, Siebformel, Formel von Poincaré-Sylvester	16
11.2	Zählvariablen	16
11.3	Austauschbarkeit	16
11.4	Rencontre-Problem (min. ein Fixpunkt)	17

11.5	Rencontre-Zahl	17
12	Erwartungswert	18
12.1	Erwartungswert einer Zählvariablen	18
12.2	Tansformationsformel	18
13	Stichprobenentnahme: Die hypergeometrische Verteilung	19
14	Mehrstufige Experimente	20
14.1	Erste Pfadregel	20
14.2	Zweite Pfadregel	20
14.3	Produktexperimente	20
15	Bedingte Wahrscheinlichkeiten	21
15.1	Zusammenhang zwischen Übergangswahrscheinlichkeiten	21
15.2	Formeln	21
15.3	Subjektive Wahrscheinlichkeiten	22
15.4	Bewertung medizinischer Tests	22
16	Stochastische Unabhängigkeit	23
16.1	Produktexperimente	23
16.2	Blöcke-Lemma	23
17	Gemeinsame Verteilung von Zufallsvariablen	24
17.1	Funktionen von ZV	24
17.2	Unabhängigkeit von ZV	24
17.3	Summen unabhängiger ZV	24
18	Die Binomalverteilung und die Multinomialverteilung	25
18.1	Verallgemeinerung der Binomialverteilung	25
19	Varianz	27
19.1	Eigenschaften der Varianz	27
19.2	Standardisierung einer Zufallsvariablen	27
19.3	Elementare stochastische Ungleichung	27
20	Kovarianz und Korrelation	28
20.1	Eigenschaften der Kovarianz	28
20.2	Die Varianz einer Indikatorsumme	29
20.3	Korrelationskoeffizient (Standardisierung der Kovarianz)	29
20.4	Eigenschaften	29
21	Diskrete Wahrscheinlichkeitsräume	30
21.1	Diskreter W -Raum	30
21.2	Ausblick	30

22	Wartezeitprobleme	32
22.1	geometrische Verteilung	32
22.2	negative Binomialverteilung	32
23	Die Poissonverteilung	33
23.1	Ausgangspunkt	33
23.2	Eigenschaften	33
24	Das Gesetz der großen Zahlen	34
24.1	Schwaches Gesetz der großen Zahlen	34
24.2	Stochastische Konvergenz	34
24.3	Schwaches Gesetz großer Zahlen von Jakob Bernoulli	34
25	Zentraler Grenzwertsatz	35
25.1	Standardisierte Partialsummen	35
25.2	ZGWS von de Moivre-Laplace	35
25.3	Approximation der Wahrscheinlichkeit	36
25.4	Stirling-Formel	36
26	Punktschätzer	36
26.1	Hypergeometrisches Modell und Binomialmodell	36
26.2	Maximum-Likelihood-Schätzer	37
27	Intervallschätzer	38
27.1	Idee	38
27.2	Konstruktion des Intervalls	38
27.3	Konfidenzgrenzen	38
27.4	Konstruktion der Vertrauensgrenzen	39
27.5	Approximative Konfidenzintervalle für große Stichprobenumfänge . . .	39
27.6	Planung des Stichprobenumfangs	39
28	Der statistische Test	40
28.1	Grundbegriffe	40
28.2	Gütefunktion	41
28.3	Philosophie des Testens	41
29	Der Binomialtest	41
29.1	Der einseitige Binomialtest	41
29.2	Der zweiseitige Binomialtest	42
29.3	Der approximative Binomialtest	43
29.4	Planung des Stichprobenumfangs	43
29.5	Zusammenhang zwischen Tests und Konfidenzintervallen	43

1 Zufallsexperiment

Zufallsexperiment

Modellierung zufallsabhängiger Vorgänge

Eigenschaften eines *idealen* ZE

- genau festgelegte Bedingungen (Versuchsbedingungen)
- mögliche Ergebnisse (Ausgänge)
- beliebig oft wiederholbar (unter gleichen Bedingungen)

Bei n-facher Versuchswiederholung: Ergebnisse als *n-Tupel*

Bei n-fach hintereinander ausgeführten Einzelexperimenten ergibt sich das kartesische Produkt $\Omega_1 \times \dots \times \Omega_n := \{(a_1, \dots, a_n) : a_j \in \Omega_j, j = 1, \dots, n\}$

Es gilt $\Omega_1 \times \dots \times \Omega_n = \Omega^n$, falls $\Omega_1 = \dots = \Omega_n$

2 Ereignis

- Teilmenge von Ω (Ergebnismenge, Grundraum)
- Elemente haben bestimmte Eigenschaft

sicheres Ereignis Ω

unmögliches Ereignis \emptyset

Vereinigung $A \cup B = \{\omega \in \Omega : \omega \in A \text{ oder } \omega \in B\}$

Durchschnitt $A \cap B = \{\omega \in \Omega : \omega \in A \text{ und } \omega \in B\}$

Diskunktheit $A \cap B = \emptyset$

Differenz $A \setminus B = \{\omega \in \Omega : \omega \in A \text{ und } \omega \notin B\}$

Komplementäreignis, Gegenereignis $A^c = \Omega \setminus A$

Partition: disjunkte Zerlegung von Ω

3 Zufallsvariablen

ZV sind ein natürliches und suggestives Darstellungsmittel von Ereignissen

ZV ist eine reellwertige Funktion mit Definitionsbereich Ω

Mittels ZV leichte Beschreibung der Ereignisse, weil sie bestimmtes Merkmal besitzen.
Zuordnung von Ω auf \mathbb{R}

$$X: \Omega \rightarrow \mathbb{R} \\ \omega \rightarrow X(\omega)$$

- Warum ZV?
Bsp: Gewinnausschüttung beim Münzwurf
- Allein aufgrund der Information von X kann i.A. nicht der Ausgang des ursprünglichen Experiments rekonstruiert werden
- Durch Bedingungen an eine ZV werden Ereignisse festgelegt

Mit ZV X, Y ist auch Summe, Differenz, Produkt, Maximum und Minimum definiert: mit ZV kann man wie mit Zahlen rechnen

Indikatorfunktion

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

⇒ Satz 3.3. (Regeln für Indikatorfunktionen)

Indikatorsumme:

Anzahl der Ereignisse die eintreten (Zählvariable)

Bsp: Trefferzahl als Zählvariable (Münzwurf)

$$A_j = \{(a_1, \dots, a_n) \in \Omega : a_j = 1\}$$

4 Relative Häufigkeiten

Zur Bewertung des "Gewissheitsgrades": *relative Häufigkeit*

$$r_{n,x}(A) := \frac{1}{n} \cdot |\{j : j = 1, \dots, n \text{ und } x_j \in A\}|$$

Wahrscheinlichkeit = Grenzwert relativer Häufigkeiten

ABER:

Die Stabilisierung relativer Häufigkeiten ist nur eine Erfahrungstatsache und kein mathematischer Sachverhalt!

5 Grundbegriffe der deskriptiven Statistik

Ziele

- Aufbereitung von Daten (Darstellung)
- Erkennen von Strukturen

Objekte nennt man *Untersuchungseinheiten, Beobachtungseinheiten oder Merkmalsträger*

Werte die von *Merkmal* bei einer Beobachtungseinheit angenommen werden heißen *Merkmalsausprägungen*

- Qualitative (artmäßig erfassbare): Geschlecht, Beruf, Nationalität
- Quantitative (zahlenmäßig erfassbare): Alter, Einwohnerzahl
 - diskret (isolierte Zahlenwerte): Zählvorgang
 - stetig: Messvorgang

Merkmale werden in stochastischen Modellen durch Zufallsvariablen beschrieben

- Totalerhebung: alle Merkmalsträger einer Grundgesamtheit werden in die Untersuchung einbezogen
- Stichprobe: zufällig gewonnene Teilmenge aus einer Grundgesamtheit

5.1 Empirische Häufigkeitsverteilung

a_1, \dots, a_s verschiedene Ausprägungen von Merkmal X

absolute Häufigkeiten

$$h_j := \sum_{i=1}^n 1_{\{x_i=a_j\}}$$

Anstelle von h_j Verwendung von relativen Häufigkeiten

$$r_j := \frac{h_j}{n}$$

5.2 Darstellungsformen

Darstellung	Darstellungsmittel
Säulendiagramm	Höhe der Säule
Kreisdiagramm	Kreissektor
Histogramm	Fläche

5.3 Relative Klassenhäufigkeiten

Sind unendlich viele Ausprägungen möglich, ist die Einteilung in *Klassen* sinnvoll.

Faustregel für Klassenzahl $s = \sqrt{n}$ (Klassenbreite konstant)

Dabei ist eine Klasse ein *halboffenes* Intervall $[a, b)$

Relative Klassenhäufigkeiten $k_j := \frac{1}{n} \sum_{i=1}^n 1_{\{a_j \leq x_i < a_{j+1}\}}$

Darstellungsform: *Histogramm*

Darstellungsmittel ist die *Fläche* des Rechtecks: $k_j = F_j = d_j \cdot (a_{j+1} - a_j)$

5.4 Kennzahlen für Lagemaße

Ein *Lagemaß* ordnet einer Stichprobe eine reelle Zahl zu, die das Zentrum der Stichprobe beschreibt

- arithmetisches Mittel
- Median
- getrimmtes Mittel

5.4.1 Arithmetisches Mittel

sensitiv auf Ausreißer

$$\bar{x} = \bar{x}_n = \frac{1}{n}(x_1 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i$$

5.4.2 Empirischer Median, Zentralwert

robust

Sortiere die Stichprobe der Größe nach

$$x_{med} = \begin{cases} x_{\frac{n+1}{2}:n} & \text{falls } n \text{ eine ungerade Zahl ist} \\ (x_{\frac{n}{2}:n} + x_{\frac{n}{2}+1:n})/2 & \text{falls } n \text{ eine gerade Zahl ist} \end{cases}$$

5.4.3 Getrimmtes Mittel

Robustifiziertes arithmetisches Mittel:

durch Weglassen von kleinen und großen Werten

$$\bar{x}_{n,\alpha} = \frac{1}{n-2 \cdot k} \sum_{j=k+1}^{n-k} x_{j:n}$$

ist das sogenannte α -getrimmte Mittel, wobei $k := [n \cdot \alpha], 0 < \alpha < 1/2$

5.5 Streuungsmaße

Eine vernünftige Forderung an ein Streuungsmaß ist, dass es sich bei Verschiebungen der Daten um einen Wert a nicht ändert

- empirische Varianz bzw. Standardabweichung
- Medianabweichung
- Quartilsabstand

5.5.1 empirische Varianz

mittlere quadratische Abweichung vom arithm. Mittel

$$s^2 = s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Die Wurzel aus der Varianz ergibt die Standardabweichung (beide werden von Ausreißern stark beeinflusst)

- Varianz von Beobachtungswerten, die in einer Stichprobe auftreten
- Erwartungswert der Stichprobenvarianz ist gleich der Varianz der GG

5.5.2 Medianabweichung

Median der absoluten Abweichungen vom Median (robust)

MAD = Median der Zahlen $|x_1 - x_{med}|, \dots, |x_n - x_{med}|$

5.5.3 Quartilsabstand

Merkmalswert unterhalb dessen ein vorgegebener Anteil α aller Fälle der Stichprobe liegen

Sei $x_{1:n} \leq \dots \leq x_{n:n}$ eine geordnete Stichprobe.

Das α -Quantil mit $0 < \alpha < 1$ ist:

$$x_\alpha := \begin{cases} x_{[n\alpha+1]:n} & n \cdot \alpha \notin \mathbb{Z} \\ \frac{x_{n\alpha:n} + x_{n\alpha+1:n}}{2} & n \cdot \alpha \in \mathbb{Z} \end{cases}$$

Dabei ist $[y] = \max\{k \in \mathbb{Z} : k \leq y\}$

1. Quartil oder unteres Quartil

$$\alpha = 1/4 = 0.25 : x_{0.25}$$

2. Quartil oder Median

$$\alpha = 1/2 = 0.5 : x_{0.5} = x_{med}$$

3. Quartil oder oberes Quartil

$$\alpha = 3/4 = 0.75 : x_{0.75}$$

Quartilsabstand (interquartile range)

$$\text{IQR} = \text{oberes Quartil} - \text{unteres Quartil} = x_{0.75} - x_{0.25}$$

Die Quartilsdifferenz misst die Länge des Intervalls, das die Hälfte der mittleren Beobachtungen enthält

6 Endliche Wahrscheinlichkeitsräume

Eine Funktion P , die jedem Ereignis (also allen Teilmengen von Ω) eine reelle Zahl zuordnet $A \rightarrow P(A) \in \mathbb{R}$ heißt *Wahrscheinlichkeitsverteilung* oder *Wahrscheinlichkeitsmaß*, falls folgende drei Axiome erfüllt sind (*Kolmogorov*)

- (A1) $P(A) \geq 0$ **Nichtnegativität**
(A2) $P(\Omega) = 1$ **Normiertheit**
(A3) $P(A+B) = P(A)+P(B)$ **Additivität**

Das Paar (Ω, P) heißt endlicher Wahrscheinlichkeitsraum

- (i) $P(\emptyset) = 0$
- (ii) $P(A^c) = 1 - P(A)$
- (iii) $0 \leq P(A) \leq 1$
- (iv) $A \subset B \Rightarrow P(B \setminus A) = P(B) - P(A)$
- (v) $A \subset B \Rightarrow P(A) \leq P(B)$ ($A \subset B$: A impliziert B)
- (vi) $P(\sum A_i) = \sum P(A_i)$
- (vii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- (viii) $P(\cup A_i) \leq \sum P(A_i)$

Durch das System der Einzelwahrscheinlichkeiten ist ein Wahrscheinlichkeitsmaß eindeutig definiert
Das System der Einzelwahrscheinlichkeiten wird auch Verteilung genannt

6.1 Verteilung einer Zufallsvariablen

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

Die *Verteilung einer Zufallsvariablen* X ist das mit P^X bezeichnete W-Maß auf $X(\Omega)$

$$P^X(B) := P(X \in B) = \sum_{j: x_j \in B} P(X = x_j)$$

mit $B \subset X(\Omega)$.

Das Paar $(X(\Omega), P^X)$ heißt das durch X induzierte Zufallsexperiment.

- Beispiel: Maximum bei zweifachem Würfelwurf, $X(\omega) = \max\{i, j\}$

6.2 Subjektive Wahrscheinlichkeiten

Kolmogorovsches Axiomensystem auch bei subjektiven Einschätzungen anwendbar.

- sich paarweise ausschließende "Elementar-Hypothesen" $\omega_1, \dots, \omega_s$
- $p(\omega_j)$ sind die subjektiven Wahrscheinlichkeiten

7 Laplace-Modelle

Laplace- oder Gleichverteilung

$$\Omega = \{\omega_1, \dots, \omega_s\}$$

$$P(\{\omega_j\}) = p(\omega_j) = \frac{1}{s} = \frac{1}{|\Omega|}$$

Laplace-Experiment Gleichverteilung auf der Ergebnismenge (Homogenität)

- Zwei nicht unterscheidbare Würfel werden gleichzeitig geworfen (kein Laplace-Modell)
- Was versteht man unter Gleichverteilung?
alle Elementarereignisse haben gleiche Wahrscheinlichkeit
- Wie wird diese in der Regel verwendet?
$$WK = \frac{\text{Anzahl günstige}}{\text{Anzahl mögliche}}$$

8 Grundlagen der Kombinatorik

Kombinatorik: Lehre des Abzählens

8.1 Fundamentalprinzip des Zählens (Multiplikationsregel)

Für (a_1, \dots, a_k) gibt es $j_1 \cdot \dots \cdot j_k$ verschiedene k -Tupel, falls j_1 Möglichkeiten für die Wahl von a_1, \dots und j_k für die Wahl von a_k . (vollständige Induktion)

8.2 Permutationen und Kombinationen

M bel. n -elementige Menge

(i) k -Permutation aus M mit Wiederholung

k -Tupel in dem Elemente auch Mehrfach auftreten dürfen

$$\begin{aligned} M^k &= \{(a_1, \dots, a_k) : a_j \in M, j = 1, \dots, k\} \\ Per_k^n(mW) &:= \{(a_1, \dots, a_k) : a_j \in \{1, \dots, n\}, j = 1, \dots, k\} \\ |Per_k^n(mW)| &= n^k \end{aligned}$$

(ii) k -Permutation aus M ohne Wiederholung

k -Tupel in dem Elemente nicht Mehrfach auftreten können

$$\begin{aligned} Per_k^n(oW) &:= \{(a_1, \dots, a_k) \in Per_k^n(mW) : a_i \neq a_j, 1 \leq i \neq j \leq k\} \\ |Per_k^n(oW)| &= n^k := n \cdot (n-1) \cdot \dots \cdot (n-k+1) \end{aligned}$$

(iii) k -Kombination mit Wiederholung

k -Permutation mit Anordnungseigenschaft

$$\begin{aligned} Kom_k^n(mW) &:= \{(a_1, \dots, a_k) : 1 \leq a_1 \leq a_2 \leq \dots \leq a_k \leq n\} \\ |Kom_k^n(mW)| &= \binom{n+k-1}{k} \end{aligned}$$

(iv) k -Kombination ohne Wiederholung

k -Permutation mit Anordnungseigenschaft

$$\begin{aligned} Kom_k^n(oW) &:= \{(a_1, \dots, a_k) : 1 \leq a_1 < a_2 < \dots < a_k \leq n\} \\ |Kom_k^n(oW)| &= \binom{n}{k} \end{aligned}$$

Beweis:

- (i) und (ii) aus Multiplikationsregel
- (iv) Umformung k -Permutation in eine k -Kombination, es gibt $k!$ verschiedene k -Permutationen, welche gleiche k -Kombination bilden
- (iii) aus (iv) mit $b_j = a_j + j - 1$

9 Urnen und Teilchen/Fächer-Modelle

Abstrakte Beschreibung stochastischer Vorgänge

9.1 Urnenmodelle

Urne mit von 1 bis n nummerierten Kugeln; Ziehen von k Kugeln

- (1) **Ziehen unter Beachtung der Reihenfolge mit Zurücklegen**
 a_j ist Nummer der Kugel vom j-ten Zug
 $Per_k^n(mW) = \{1, \dots, n\}^k$
- (2) **Ziehen unter Beachtung der Reihenfolge ohne Zurücklegen**
 Notierung der Kugel nach jedem Zug ohne Zurücklegen
 $Per_k^n(oW)$
- (3) **Ziehen ohne Beachtung der Reihenfolge mit Zurücklegen**
 Notierung wie oft jede der n Kugeln gezogen wurde
 a_j gibt j-kleinste der Nummern der gezogenen Kugeln an
 $Kom_k^n(mW)$
- (4) **Ziehen ohne Beachtung der Reihenfolge ohne Zurücklegen**
 $Kom_k^n(oW)$

9.2 Teilchen/Fächer-Modelle

Teilchen/Fächer entsprechen Ziehungen/Kugeln; k Teilchen sollen auf n Fächer verteilt werden

- (1') **Unterscheidbare Teilchen, Mehrfachbelegung zugelassen**
 Unterscheidbarkeit entspricht Berücksichtigung der Reihenfolge
- (2') **Unterscheidbare Teilchen, keine Mehrfachbelegungen**
- (3') **Nichtunterscheidbare Teilchen, Mehrfachbelegungen zugelassen**
 Man kann nur feststellen wie viele Teilchen in einem Fach liegen
- (4') **Nichtunterscheidbare Teilchen, keine Mehrfachbelegungen**

9.3 Zusammenfassung

Urnenmodell	Teilchen/Fächer-Modell
Beachtung der Reihenfolge	Unterscheidbare Teilchen
ohne Beachtung der Reihenfolge	nicht unterscheidbare Teilchen
mit Zurücklegen	mit Mehrfachbelegung
ohne Zurücklegen	ohne Mehrfachbelegung

10 Das Paradoxon der ersten Kollision

Modellierung des Kollisions-Phänomens

X_n := Zeitpunkt der ersten Kollision beim sukzessiven rein zufälligen Besetzen von n Fächern

Wähle $n+1$ aus n mit Wiederholung

Kollision im k -ten "Versuch"

$$P(X_n \geq k+1) = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{n^k}$$

Komplementbildung liefert $P(X_n \leq k) = 1 - \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right)$ für $k > 2$

Beweis: Unterstellung Laplace-Modell, Verwendung der Multiplikationsregel
 $(a_1, \dots, a_{n+1}) \in A_{n,k} := \{X_n \geq k+1\}$

$$\begin{aligned} P(A_{n,k}) &= \frac{|A_{n,k}|}{|\Omega|} \\ &= \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1) \cdot n^{n+1-k}}{n^{n+1}} \end{aligned}$$

Für die ersten k Komponenten gilt $a_i \neq a_j$, also $Per_k^n(oW)$, danach mit Wiederholung für die restlichen Komponenten

$$\lim_{n \rightarrow \infty} P(X_n \leq \sqrt{n} \cdot t) = 1 - e^{-t^2/2} \quad (t > 0)$$

Beweis: Abschätzung von oben und unten mittels $P(X_n \leq k_n) \leq P(X_n \leq \sqrt{n} \cdot t) \leq P(X_n \leq k_n + 1)$ (genauer?) $t > 0$ gegeben, für jedes genügend große n existiert $k_n \in \mathbb{N}$ mit $k_n \leq \sqrt{n} \cdot t \leq k_n + 1$

11 Die Formel des Ein- und Ausschließens

Ziel: Berechnung der Wahrscheinlichkeit der Vereinigungen von Ereignissen

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

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) \\ - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ + P(A_1 \cap A_2 \cap A_3)$$

11.1 Formel des Ein- und Ausschließens, Siebformel, Formel von Poincaré-Sylvester

$$S_r := \sum_{1 \leq i_1 < \dots < i_r \leq n} P(A_{i_1} \cap \dots \cap A_{i_r}) \text{ (Summation aller } r\text{-elementigen Teilmengen)}$$

$$P\left(\bigcup_{j=1}^n A_j\right) = \sum_{r=1}^n (-1)^{r-1} \cdot S_r$$

Beweis: vollständige Induktion über n und Additionsgesetz

Aufspalten in $\bigcup_{i=1}^n A_i \cup A_{n+1}$ mittels Additionsgesetz, Anwendung der IV, geschickte Betrachtung der Indexmenge

11.2 Zählvariablen

$$X = 1_{A_1} + \dots + 1_{A_n}$$

$$P(X \geq 1) = P\left(\bigcup_{j=1}^n A_j\right) = \sum_{r=1}^n (-1)^{r-1} \cdot S_r$$

(min. ein A_i tritt ein: Summation der einzelnen A_i mittels Siebformel)

11.3 Austauschbarkeit

Gilt $\forall r \in \{1, \dots, n\} \forall i_1, \dots, i_r$ mit $1 \leq i_1 < \dots < i_r \leq n$:

$P(A_{i_1} \cap \dots \cap A_{i_r}) = p_r$, d.h. die Wahrscheinlichkeit des Durchschnitts hängt nur von der Anzahl r ab, nicht von der Wahl der Ereignisse, so heißen die Ereignisse A_1, \dots, A_n austauschbar.

$$P\left(\bigcup_{j=1}^n A_j\right) = \sum_{r=1}^n (-1)^{r-1} \binom{n}{r} P(A_1 \cap \dots \cap A_r)$$

11.4 Rencontre-Problem (min. ein Fixpunkt)

Die Zahlen $1, \dots, n$ werden einer rein zufälligen Permutation unterworfen.

Wie groß ist die Wahrscheinlichkeit, dass diese Permutation mindestens ein Element fest lässt?

$$\Omega = \text{Per}_n^{\text{OW}}$$

$$A_j = \{(a_1, \dots, a_n) \in \Omega : a_j = j\}$$

Ereignis mindestens ein Fixpunkt tritt auf $A_1 \cup \dots \cup A_n$

$$P(A_1 \cup \dots \cup A_n) = \sum_{r=1}^n (-1)^{r-1} \cdot \frac{1}{r!}$$

Beweis: Ereignisse sind austauschbar, da bestimmte Anzahl an Komponenten fest (r) und $n-r$ Komponenten des n -Tupels beliebig permutiert werden $\left(\frac{(n-r)!}{n!}\right) \cdot \binom{n}{r} = \frac{1}{r!}$
WK, dass min. 1 Element fest bleibt bei $n = 3$ schon über 50%

11.5 Rencontre-Zahl

Anzahl aller fixpunktfreien Permutationen (Komplementärereignis)

$$1 - \sum_{r=1}^n (-1)^{r-1} \cdot \frac{1}{r!} = - \sum_{r=2}^n (-1)^{r-1} \cdot \frac{1}{r!}$$

Anzahl $n! \sum_{r=2}^n (-1)^r \cdot \frac{1}{r!} = n! \sum_{r=0}^n (-1)^r \cdot \frac{1}{r!} =: \mathcal{R}_n$
Abschätzung mittels Exponentialreihe ergibt 63% als Grenzwert

- Wann wird die Formel besonders einfach?
Bei disjunkten Ereignissen
- In welchem Zusammenhang hatten wir in der Vorlesung mit austauschbaren Ereignissen zu tun?
Rencontre-Problem; Wahrscheinlichkeit, dass mindestens ein Fixpunkt in der Menge auftritt
- Es wird oft übersehen, dass ein vermeintlich "unwahrscheinliches" Ereignis in Wirklichkeit die Vereinigung vieler unwahrscheinlicher Ereignisse darstellt

12 Erwartungswert

$$E(X) := \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\})$$

Eigenschaften:

- (i) $E(X + Y) = E(X) + E(Y)$
 - (ii) $E(a \cdot X) = a \cdot E(X)$
 - (iii) $E(1_A) = P(A)$ (wichtig!)
 - (iv) $X \leq Y \Rightarrow E(X) \leq E(Y)$
- (i),(ii),(iv) folgen sofort aus Def.
 - (iii) $E(1_A) = \sum_{\omega \in \Omega} 1_A(\omega) \cdot P(\{\omega\}) = \sum_{\omega \in A} 1 \cdot P(\{\omega\}) = P(A)$

12.1 Erwartungswert einer Zählvariablen

$$X = \sum_{j=1}^n 1_{A_j} \Rightarrow E(X) = \sum_{j=1}^n P(A_j)$$

$$\text{Spezialfall } P(A_j) = p \Rightarrow E(X) = n \cdot p$$

12.2 Transformationsformel

$$X(\Omega) = \{x_1, \dots, x_k\}, k \text{ verschiedene Werte in } X(\Omega) \quad E(g(X)) = \sum_{j=1}^k g(x_j) \cdot P(X = x_j)$$

g ist Funktion auf Zufallsvariable X

Für $g(x) = x$ gilt

$$E(X) = \sum_{j=1}^k x_j \cdot P(X = x_j)$$

Der Erwartungswert hängt nur von Verteilung der ZV, nicht vom zugrundeliegenden W-Raum ab

Beweis: Summe Aufspalten in A_j die gleiches Ergebnis bei ZV, damit $X(\omega) = x_j, \dots$

- Was muss man sich unter dem Erwartungswert vorstellen?
Erwartungswert ist die auf lange Sicht erwartete Auszahlung pro Spiel

13 Stichprobenentnahme: Die hypergeometrische Verteilung

Die hypergeometrische Verteilung gibt dann Auskunft darüber, mit welcher Wahrscheinlichkeit in der Stichprobe eine bestimmte Anzahl von Elementen vorkommt, die die gewünschte Eigenschaft haben. Bedeutung kommt dieser Verteilung daher etwa bei Qualitätskontrollen zu.

Qualitätskontrolle: Ziehen von r roten (intakt) und s schwarzen (defekt) Kugeln ohne Zurücklegen

Modell der Binomialverteilung ohne Zurücklegen \Rightarrow Hypergeometrische Verteilung

Stochastisches Modell 1

a_j Nummer der j -ten gezogenen Kugel

$$\Omega = \text{Per}_n^{r+s}(oW) = \{(a_1, \dots, a_n) \in \{1, \dots, r+s\}^n : a_i \neq a_j, 1 \leq i \neq j \leq r+s\}$$

$$|\Omega| = (r+s)^n$$

Ziehen rein zufällig \Rightarrow Laplace-Modell

$$(i) E(X) = n \cdot \frac{r}{r+s}$$

$$(ii) P(X = k) = \frac{\binom{r}{k} \cdot \binom{s}{n-k}}{\binom{r+s}{n}}$$

Die durch die Einzelwahrscheinlichkeiten festgelegte Verteilung von X heißt *hypergeometrische Verteilung* mit den Parametern n, r und s

$$X \sim H(n, r, s)$$

Beweis?

Stochastisches Modell 2

$$\Omega = \text{Kom}_n^{r+s}(oW)$$

Dies lässt keinen Rückschluss mehr auf die Ziehungsreihenfolge zu (Anzahl!)

Diesselbe Verteilung trotz unterschiedlichen W-Raums

14 Mehrstufige Experimente

$\Omega := \Omega_1 \times \dots \times \Omega_n$, wobei Ω_j Ergebnismenge des Telexperiments ist

Zweistufige Experimente

$$\Omega = \Omega_1 \times \Omega_2$$

$$\text{Startverteilung } \sum_{a_1 \in \Omega_1} p_1(a_1) = 1$$

$$\text{Übergangswahrscheinlichkeiten } \sum_{a_2 \in \Omega_2} p_2(a_2|a_1) = 1$$

$$p(\omega) = p(a_1, a_2) = p_1(a_1) \cdot p_2(a_2|a_1)$$

$$P(\Omega) = 1$$

Mehrstufige Experimente

System von Übergangswahrscheinlichkeiten $p_j(a_j|a_1, \dots, a_{j-1})$

$$\text{wobei } \sum_{a_j \in \Omega_j} p_j(a_j|a_1, \dots, a_{j-1}) = 1$$

$$p(\omega) := p_1(a_1) \cdot p_2(a_2|a_1) \cdot p_3(a_3|a_1, a_2) \cdot \dots \cdot p_n(a_n|a_1, \dots, a_{n-1})$$

14.1 Erste Pfadregel

Deutet man ω als einen Pfad im Baumdiagramm, so besagt, dass die Wahrscheinlichkeit dieses Pfades gleich dem Produkt der an den Pfeilen des Pfades stehenden Übergangswahrscheinlichkeiten ist.

14.2 Zweite Pfadregel

$$P(A) := \sum_{\omega \in A} p(\omega)$$

Wahrscheinlichkeit von A berechnet sich als Summe der Wahrscheinlichkeiten aller zu A gehörenden Pfade

14.3 Produktexperimente

Das j -te Telexperiment kann ohne Kenntnis der Ergebnisse der früheren $j - 1$ Telexperimente durchgeführt werden, d.h. die Übergangswahrscheinlichkeiten hängen nicht von den Ergebnissen früherer Experimente ab.

$$p_j(a_j|a_1, \dots, a_{j-1}) =: p_j(a_j)$$

Somit geht die erste Pfadregel über in:

$$p(\omega) := p_1(a_1) \cdot \dots \cdot p_n(a_n)$$

15 Bedingte Wahrscheinlichkeiten

Teilinformationen über das Ergebnis eines *bereits abgeschlossenen* Vorgangs führen zu Neubewertung der Eintrittswahrscheinlichkeit des uns interessierenden Ereignisses

$$P(A|B) := \frac{P(A \cap B)}{P(B)}$$

Bedingte Wahrscheinlichkeit von A unter der Bedingung B festgehaltenes B definiert mit $A \rightarrow P(A|B)$ ein W-Maß

15.1 Zusammenhang zwischen Übergangswahrscheinlichkeiten

$$P(A \cap B) = P(B) \cdot P(A|B)$$

allgemeine Multiplikationsregel

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \dots \cdot P(A_n|A_1 \cap \dots \cap A_{n-1})$$

Übergangswahrscheinlichkeiten in gekoppelten Experimenten lassen sich als bedingte Wahrscheinlichkeiten auffassen

15.2 Formeln

(i) **Formel von der totalen Wahrscheinlichkeit**

$$P(B) = \sum_{j=1}^s P(A_j) \cdot P(B|A_j) = \sum_{j=1}^s P(A_j) \cdot \frac{P(A_j \cap B)}{P(A_j)} = \sum_{j=1}^s P(A_j \cap B)$$

(ii) **Bayes-Formel**

$$P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{j=1}^s P(A_j) \cdot P(B|A_j)}$$

Beweis

- (i) Idee $P(B) = P(\Omega \cap B)$
- (ii) Nenner ist $P(B)$

Für $n = 2$ nimmt Formel der totalen Wahrscheinlichkeit folgende Form an (Pfadregel)

$$P(B) = \sum_{j=1}^s p_1(e_j) \cdot p_2(b|e_j)$$

15.3 Subjektive Wahrscheinlichkeiten

Wenn bei der Bayes-Formal A_j als "Ursache" oder "Hypothese" für das Eintreten von B aufgefasst wird:

- vor Beobachtung eines stochastischen Vorgangs $P(A_j)$ (a-priori-Wahrscheinlichkeit)
- Wenn B eintritt so ist $P(A_j|B)$ die a-posteriori-Wahrscheinlichkeit, dafür dass A_j Ursache von B ist

15.4 Bewertung medizinischer Tests

- **Sensitivität** eines Tests: Wahrscheinlichkeit, dass eine kranke Person als krank erkannt wird
- **Spezifität** eines Tests: Wahrscheinlichkeit, dass eine gesunde Person als gesund erkannt wird

K: krank, N: negativer Test

$$P(K|N^c) = \frac{P(K) \cdot P(N^c|K)}{P(K) \cdot P(N^c|K) + P(K^c) \cdot P(N^c|K^c)}$$

16 Stochastische Unabhängigkeit

$P(A|B) = P(A)$, d.h. das Eintreten von B hat wahrscheinlichkeitstheoretisch keinen Einfluss auf das Eintreten von A

$$\Rightarrow P(A \cap B) = P(A) \cdot P(B)$$

Wenn diese Gleichung erfüllt ist sind A und B stochastisch unabhängig bzgl. P

Disjunkte Ereignisse sind nur bei $P(A) = 0$ oder $P(B) = 0$ stochastisch unabhängig

- paarweise Unabhängigkeit impliziert nicht die (vollst.) Unabhängigkeit

A_1, \dots, A_n heißen stochastisch unabhängig, falls

$$P\left(\bigcap_{j \in T} A_j\right) = \prod_{j \in T} P(A_j)$$

für $T \subset \{1, \dots, n\}$ mit $2 \leq |T| \leq n$

folgende Aussagen sind äquivalent

(i) A_1, \dots, A_n sind stochastisch unabhängig

(ii) Es ist
$$P\left(\left(\bigcap_{i \in I} A_i\right) \cap \left(\bigcap_{j \in J} A_j^c\right)\right) = \prod_{i \in I} P(A_i) \cdot \prod_{j \in J} P(A_j^c)$$

für jede Wahl disjunkter Teilmengen $I, J \subset \{1, \dots, n\}$

- (ii) \Rightarrow (i) klar, $J = \emptyset$
- (i) \Rightarrow (ii) Induktion $k = |J|$

16.1 Produktexperimente

$$\Omega = \Omega_1 \times \dots \times \Omega_n$$

$$p(\omega) = p_1(a_1) \cdot \dots \cdot p_n(a_n), \omega = (a_1, \dots, a_n)$$

ist P, also das Produkt W-Maß $P := \prod P_j$

Ereignisse, die sich auf verschiedene Komponenten des Produktexperiments beziehen sind stochastisch unabhängig

16.2 Blöcke-Lemma

Aufteilung von n unabhängigen Ereignisse in zwei disjunkte Blöcke B und C

\Rightarrow Ereignisse B und C sind unabhängig

17 Gemeinsame Verteilung von Zufallsvariablen

$X, Y : \Omega \rightarrow \mathbb{R}$ ZV mit Wertebereich x_1, \dots, x_r bzw. y_1, \dots, y_s

$P(X = x_i, Y = y_j) = P(\{X = x_i\} \cap P(\{Y = y_j\})$ heißt **gemeinsame Verteilung** von X und Y

- (X, Y) ist zweidimensionaler Zufallsvektor
- Es gilt: $\{X = x_i\} = \sum_{j=1}^s \{X = x_i, Y = y_j\}$
- Die Verteilungen der Komponenten eines Zufallsvektors nennt man **Randverteilungen** oder **Marginalverteilungen**
- Beispiel: Zweifacher Würfelwurf, ZV: 1. Wurf, Maximum
- Die gemeinsame Verteilung ist i.A nicht durch die beiden Marginalverteilungen festgelegt
- Kontingenztafeln (Vierfeldertafel) zur Darstellung gemeinsamer empirischer Häufigkeitsverteilungen
- Ausprägung zweier Merkmale (ZV): Daten von der Form (x_i, y_l) in l -ter U-Einheit

17.1 Funktionen von ZV

$g(X, Y)(\omega) := g(X(\omega), Y(\omega))$ ist ZV $g(X, Y)$

$$E(g(X, Y)) = \sum_{i=1}^r \sum_{j=1}^s g(x_i, y_j) \cdot P(X = x_i, Y = y_j)$$

Für die Berechnung des Erwartungswerts muss nicht erst die Verteilung berechnet werden

17.2 Unabhängigkeit von ZV

X, Y stochastisch unabhängig falls $P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$ für alle $x, y \in \mathbb{R}$

17.3 Summen unabhängiger ZV

X, Y unabhängig

Faltungsformel $P(X + Y = u) = \sum_{x_i + y_j = u} P(X = x_i) \cdot P(Y = y_j)$

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

Standardmodell für unabhängige ZV

Projektion auf die Komponenten $X_j(a_1, \dots, a_n) = a_j$

18 Die Binomialverteilung und die Multinomialverteilung

Ausgangspunkt: n unabhängige Treffer/Niete-Experimente

$$p_j(1) = 1 - p_j(0) = p$$

Bernoulli-Kette der Länge n $p(\omega) = \prod_{j=1}^n p_j(a_j) = p^i \cdot (1-p)^{n-i}$

Zählvariable $X := \sum_{j=1}^n 1_{A_j}$ (für Anzahl Treffer)

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

Eine ZV X besitzt **Binomialverteilung mit Parametern n und p** : $X \sim B(n,p)$, falls $P(X=k)$ mit obiger Gleichung für alle k erfüllt

$$E(X) = n \cdot E(1_{A_i}) = n \cdot p \text{ (da } A_i \text{ unabhängig und Additivität des EW)}$$

- Wie kommt man auf den Binomialkoeffizient?
 - $n!$ Möglichkeiten der Permutation dividiert durch die $k!$ Einsen, die beliebig vertauscht sein können und die $(n-k)!$ Nullen
- stochastisch unabhängige Ereignisse mit gleicher Wahrscheinlichkeit sind binomialverteilt
- Additionsgesetz für Binomialverteilung
 - $X \sim B(m,p), Y \sim B(n,p) \Rightarrow X+Y \sim B(m+n,p)$

18.1 Verallgemeinerung der Binomialverteilung

Experiment mit Ausgängen $1, \dots, s$ Ausgang k mit Wahrscheinlichkeit p_k , Modellierung als Produktexperiment

Anzahl der Treffer k -ter Art in n Versuchen: $X_k := \sum_{j=1}^n 1_{A_j(k)}$.

Ziel: Bestimmung der Verteilung der Zufallsvektoren X_k

$$|\{X_1 = i_1, \dots, X_s = i_s\}| = \binom{n}{i_1} \cdot \binom{n-i_1}{i_2} \cdot \dots \cdot \binom{n-i_1-\dots-i_{s-1}}{i_s} = \frac{n!}{i_1! \cdot i_2! \cdot \dots \cdot i_s!} := \binom{n}{i_1, \dots, i_s}$$

(Multinomialkoeffizient)

Ein Zufallsvektor (X_1, \dots, X_s) besitzt eine **Multinomialverteilung** falls

$$P(X_1 = i_1, \dots, X_s = i_s) = \frac{n!}{i_1! \cdot i_2! \cdot \dots \cdot i_s!} p_1^{(i_1)} \cdot p_2^{(i_2)} \cdot \dots \cdot p_s^{(i_s)}$$

$(X_1, \dots, X_s) \sim Mult(n; p_1, \dots, p_s)$

Eigenschaften

Sei $(X_1, \dots, X_s) \sim \text{Mult}(n; p_1, \dots, p_s)$

(i) $X_k \sim B(n, p_k)$

(ii) Sei T_1, \dots, T_l Partition der Menge $\{1, \dots, s\}$

$$Y_r := \sum_{k \in T_r} X_k$$

$$q_r := \sum_{k \in T_r} p_k, r = 1, \dots, l$$

$$\Rightarrow (Y_1, \dots, Y_l) \sim \text{Mult}(n; q_1, \dots, q_l)$$

Beweis: folgt aus Darstellung: Treffer k-ter Art

19 Varianz

Abweichung einer ZV X von ihrem Erwartungswert

$$\text{Varianz } V(X) := E(X - E(X))^2 = \sum_{j=1}^n (x_j - E(X))^2 \cdot P(X = x_j)$$

$$\text{Standardabweichung } \sigma(X) := \sqrt{V(X)}$$

- Varianz (wie auch $E(X)$) hängt nur von Verteilung, nicht von der speziellen Gestalt des W -Raumes ab
- bei Gleichverteilung

$$E(X) = \frac{1}{n} \sum_{j=1}^n x_j = \bar{x}$$

$$V(X) = \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2 = \frac{n-1}{n} s^2$$

19.1 Eigenschaften der Varianz

- (i) $V(X) = E(X - a)^2 - (E(X) - a)^2, a \in \mathbb{R}$
- (ii) $V(X) = E(X^2) - (E(X))^2$
- (iii) $V(X) = \min_{a \in \mathbb{R}} E(X - a)^2$
- (iv) $V(a \cdot X + b) = a^2 \cdot V(X)$
- (v) $V(X) \geq 0$ und $V(X) = 0 \Leftrightarrow P(X = a) = 1$ für ein $a \in \mathbb{R}$

19.2 Standardisierung einer Zufallsvariablen

Ist X ZV so heißt die *lineare Transformation*

$$X \rightarrow Z := \frac{X - E(X)}{\sqrt{V(X)}}$$

die *Standardisierung von X*

Es gilt $E(Z) = 0$ und $V(Z) = 1$

19.3 Elementare stochastische Ungleichung

- **Markov-Ungleichung**
Für alle $\epsilon > 0$ $P(|X| \geq \epsilon) \leq \frac{1}{\epsilon} \cdot E(|X|)$
- **Tschebyschow-Ungleichung**
Für alle $\epsilon > 0$ $P(|X - E(X)| \geq \epsilon) \leq \frac{1}{\epsilon^2} \cdot V(X)$

20 Kovarianz und Korrelation

Maßzahl für den Zusammenhang zweier statistischer Merkmale

- gleichsinnig: > 0
- ungleichsinnig: < 0
- unkorreliert: $= 0$

Kovarianz von X, Y ZV $C(X, Y) = E((X - E(X)) \cdot (Y - E(Y)))$
 $V(X + Y) = V(X) + V(Y) + 2 \cdot C(X, Y)$

- Varianzbildung ist nicht additiv
- Kovarianz hängt von der gemeinsamen Verteilung von X und Y ab

20.1 Eigenschaften der Kovarianz

- (i) $C(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y)$
- (ii) $C(X, Y) = C(Y, X), C(X, X) = V(X)$
- (iii) $C(X + a, Y + b) = C(X, Y)$
- (iv) $C(X, Y) = 0$, falls X und Y unabhängig
- (v) $C\left(\sum_{i=1}^m a_i X_i, \sum_{j=1}^n b_j Y_j\right) = \sum_{i=1}^m \sum_{j=1}^n a_i \cdot b_j \cdot C(X_i, Y_j)$
- (vi) $V(X_1 + \dots + X_n) = \sum_{j=1}^n V(X_j) + 2 \cdot \sum_{1 \leq i < j \leq n} C(X_i, X_j)$

Sind die ZV X_1, \dots, X_n stochastisch unabhängig so gilt

$$V(X_1 + \dots + X_n) = \sum_{j=1}^n V(X_j)$$

Diese Formel bleibt unter der schwächeren Bedingung der **paarweisen Unkorreliertheit** $Cov(X_i, X_j) = 0$ gültig

Zwei ZV heißen **unkorreliert**, falls $C(X, Y) = 0$

Aus Unkorreliertheit folgt nicht Unabhängigkeit

20.2 Die Varianz einer Indikatorsumme

Varianz der Indikatorsumme $X = \sum_{j=1}^n 1_{A_j}$

Kovarianz $C(1_A, 1_B) = E(1_A \cdot 1_B) - E(1_A) \cdot E(1_B) = E(1_{A \cap B}) - E(1_A) \cdot E(1_B) = P(A \cap B) - P(A) \cdot P(B)$

Varianz $V(1_A) = P(A)(1 - P(A)) = P(A) - P(A)^2$

$$V(X) = \sum_{j=1}^n P(A_j)(1 - P(A_j)) + 2 \cdot \sum_{1 \leq i < j \leq n} (P(A_i \cap A_j) - P(A_i)P(A_j))$$

Falls alle Ereignisse gleichwahrscheinlich und alle Schnitte gleiche Wahrscheinlichkeit haben:

$$V(X) = nP(A_1)(1 - P(A_1)) + n(n-1)(P(A_1 \cap A_2) - P(A_1)^2)$$

$$V(X) = n \cdot p(1 - p)$$

Hier ist $Y \sim \text{Bin}(n, p)$, also die Indikatorfunktion binomialverteilt

20.3 Korrelationskoeffizient (Standardisierung der Kovarianz)

Korrelationskoeffizient von X und Y

$$r := r(X, Y) = \frac{C(X, Y)}{\sqrt{V(X) \cdot V(Y)}}$$

Der Korrelationskoeffizient ist ein Maß für die Güte der linearen Vorhersagen

20.4 Eigenschaften

- (i) $C(X, Y)^2 \leq V(X) \cdot V(Y)$ (Cauchy-Schwarz-Ungleichung)
 - (ii) $|r(X, Y)| \leq 1$
 - (iii) Es gilt $|r(X, Y)| = 1$ genau dann wenn $\exists a, b \in \mathbb{R}$ mit $P(Y = a + b \cdot X) = 1$
 - $r(X, Y) = +1$ folgt $b > 0$, d.h. Y wächst mit wachsendem X
 - $r(X, Y) = -1$ folgt $b < 0$, d.h. Y fällt mit wachsendem X
- Situation in der Datenanalyse: Untersuchung eines statistischen Zusammenhangs (Körpergröße, Gewicht) mittels Punktwolken
 - empirischer Korrelationskoeffizient macht Aussage über Stärke und Richtung eines linearen Zusammenhangs

21 Diskrete Wahrscheinlichkeitsräume

Modellierung durch endlichen W-Raum nicht möglich

$$\begin{aligned} &\text{Erweiterung: } \Omega = \{\omega_1, \omega_2, \dots\} \text{ abzählbar unendlich} \\ &p(\omega_j) \geq 0 \\ &\sum_{j=1}^{\infty} p(\omega_j) = 1 \end{aligned}$$

21.1 Diskreter W-Raum

(A1) $P(A) \geq 0$ für alle $A \subset \Omega$ (Nichtnegativität)

(A2) $P(\Omega) = 1$ (Normiertheit)

(A3) $P\left(\sum_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$ (σ -Additivität)

Es bleiben gültig

- Folgerungen (i) bis (viii) (Kapitel 6)
- Siebformel (Kapitel 11)
- Formel der totalen WK und Formel von Bayes (Kapitel 15)
- Erwartungswert $E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\})$ falls $< \infty$

21.2 Ausblick

Bisher ist Ω diskret, also *endlich* oder *abzählbar unendlich*.
W-Maß ist dann eine Mengenfunktion

$\Omega = \mathbb{R}$

\Rightarrow man kann nicht mit Potenzmengen arbeiten

$\Rightarrow \sigma$ -Algebren

(Ω, \mathcal{A}, P) heißt W-Raum, falls

(i) $\Omega \neq \emptyset$

(ii) \mathcal{A} σ -Algebra:

(a) $\Omega \in \mathcal{A}$

(b) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

(c) $A_n \in \mathcal{A}, n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

(iii) P W-Maß auf \mathcal{A}

(A1) $P(A) \geq 0, A \in \mathcal{A}$

(A2) $P(\Omega) = 1$

(A3) $A_n \in \mathcal{A}, n \in \mathbb{N}$, paarweise disjunkt

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

- σ -Additivität \Rightarrow Additivität (Umkehrung gilt nicht)
- im Endlichen fällt beides zusammen

22 Wartezeitprobleme

22.1 geometrische Verteilung

Warten auf den ersten Treffer: die geometrische Verteilung

Sei X die **Anzahl der Nieten vor dem ersten Treffer**, also

$$\{X = k\} = \{\omega_{k+1}\}$$

$$P(X = k) = (1 - p)^k \cdot p \quad X \sim G(p) \quad (X \text{ ist geometrisch verteilt})$$

$$E(X) = \frac{1-p}{p} = \frac{1}{p} - 1$$

$$V(X) = \frac{1-p}{p^2}$$

22.2 negative Binomialverteilung

Warten auf den r -ten Treffer

Tupel mit r Einsen und $j - r$ Nieten; jedes Tupel hat Wahrscheinlichkeit $(1 - p)^{j-r} \cdot p^r$ und es gibt $\binom{j-1}{r-1}$ Möglichkeiten $r - 1$ Treffer unter den ersten $j - 1$ Versuchen auszuwählen

$$p_{r,j} := \binom{j-1}{r-1} \cdot (1-p)^{j-r} \cdot p^r \quad (\text{Summe aller } W_k \text{ liefert } 1)$$

Modellierung als Produktexperiment mit dem Grundraum Ω^j

Sei X die **Anzahl der Nieten vor dem r -ten-Treffer**, also benötigt man r X_i , welche geometrisch verteilt sind

$$\text{negative Binomialverteilung } P(X = k) = \binom{k+r-1}{k} \cdot p^r \cdot (1-p)^k \quad (X \sim Nb(r, p))$$

$$E(X) = r \cdot \frac{1-p}{p}$$

$$V(X) = r \cdot \frac{1-p}{p^2}$$

Namensgebung der negative Binomialverteilung kommt von

$$\text{negative Binomialverteilung } (X = k) = \binom{-r}{k} \cdot p^r \cdot (-(1-p))^k$$

23 Die Poissonverteilung

23.1 Ausgangspunkt

Binomialverteilung $B(n, p_n)$ mit $p_n := \lambda/n$

Untersuchung des Verhaltens für $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} B(n, p_n)(\{k\}) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

Wegen $\sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$ ist dies ein W-Maß auf $\mathbb{N} \cup \{0\}$

Poisson-Verteilung

$$P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, k \in \mathbb{N} \cup \{0\}$$

$$X \sim Po(\lambda)$$

- Bin: Verteilung der Summe von n Indikatoren unabhängiger Ereignisse mit gleicher Wahrscheinlichkeit
- Poisson: Jedes Ereignis besitzt kleine Wahrscheinlichkeit $p_n = \lambda/n$ und tritt selten ein; die Wahrscheinlichkeit des Eintretens von k dieser Ereignisse konvergiert gegen einen festen, nur von λ und k abhängenden Wert

23.2 Eigenschaften

- Falls $X \sim Po(\lambda)$, so gilt $E(X) = V(X) = \lambda$
- $X \sim Po(\lambda), Y \sim Po(\mu) \Rightarrow X + Y \sim Po(\lambda + \mu)$
- Rutherford-Geiger-Experiment: radioaktiver Zerfall Poisson-verteilt

24 Das Gesetz der großen Zahlen

- Stabilisierung relativer Häufigkeiten ist Erfahrungstatsache

24.1 Schwaches Gesetz der großen Zahlen

X_1, \dots, X_n stochastisch unabhängige ZV auf diskretem W-Raum mit
 $E(X_1) = \dots = E(X_n) =: \mu$ und
 $V(X_1) = \dots = V(X_n) =: \sigma^2$
Dann gilt für jedes $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \cdot \sum_{j=1}^n X_j - \mu \right| \geq \epsilon \right) = 0$$

- Beweis: Anwendung der Tschebyschow-Ungleichung
- WK, dass bei großem n der Abstand zwischen arithmetischen Mittel und EW $\geq \epsilon$ ist, ist Null

24.2 Stochastische Konvergenz

Y_1, Y_2, \dots ZV und $a \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} P(|Y_n - a| \geq \epsilon) = 0 \text{ für jedes } \epsilon > 0$$

so ist $(Y_n)_n$ gegen a konvergent

Das schwache Gesetz der großen Zahlen besagt also, dass die Folge der arithmetischen Mittel von unabhängigen ZV mit gleichem Erwartungswert μ und gleicher Varianz, stochastisch gegen μ konvergiert

24.3 Schwaches Gesetz großer Zahlen von Jakob Bernoulli

A_1, \dots, A_n stochastisch unabhängige Ereignisse mit $P(A_1) = \dots = P(A_n) = p$

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \cdot \sum_{j=1}^n 1_{A_j} - p \right| \geq \epsilon \right) = 0$$

Die Wahrscheinlichkeit, dass sich relative Trefferhäufigkeit R_n in einer Bernoulli-Kette vom Umfang n von der Trefferwahrscheinlichkeit p um weniger als einen beliebig kleinen vorgegebenen Wert ϵ unterscheidet konvergiert für $n \rightarrow \infty$ gegen 1

- relative Häufigkeiten stabilisieren sich (mathematisch quantifizierbar)
- Starkes Gesetz der großen Zahlen (stärkere Konvergenz als stochastische) direkte Konvergenz zu μ

25 Zentraler Grenzwertsatz

Bernoulli-Kette der Länge n

A_j Treffer im j -ten Versuch

$$X_j = 1_{A_j}$$

$$S_n = X_1 + \dots + X_n \sim \text{Bin}(n, p)$$

ZV muss standardisiert werden

25.1 Standardisierte Partialsummen

$$S_n^* = \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - np}{\sqrt{npq}}$$

(Anzahl der Treffer mit $X_j = 1_{A_j}$, wobei A_j unabhängige Ereignisse sind mit $P(A_j) = p$)

$$E(S_n^*) = 0, V(S_n^*) = 1$$

Für großes n nähert sich das Histogramm mit Klassenbreite $\frac{1}{\sqrt{npq}} = x_{n,j+1} - x_{n,j}$ und

Höhe $P(S_n^* = x_{n,j}) \cdot \sqrt{npq} = h_{n,j}$ der Gaußschen Glockenkurve an
(Dichte der standardisierten Normalverteilung)

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

25.2 ZGWS von de Moivre-Laplace

Normalapproximation der Binomialverteilung

$$S_n \sim B(n, p)$$

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right) = \int_a^b \varphi(x) dx$$

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - np}{\sqrt{npq}} \leq b\right) = \int_{-\infty}^b \varphi(x) dx$$

$$\Phi(t) = \int_{-\infty}^t \varphi(x) dx$$

(Verteilungsfunktion der standardisierten Normalverteilung)

Eigenschaften

- $\varphi(x) = \varphi(-x)$
- $\lim_{t \rightarrow \infty} \Phi(t) = \int_{-\infty}^{\infty} \varphi(x) dx = 1$
- für $n \rightarrow \infty$ Binomialverteilung konv. gg. Normalverteilung

25.3 Approximation der Wahrscheinlichkeit

$$P(k \leq S_n \leq l) = \sum_{j=k}^l B(n, p)(\{j\}) = \sum_{j=k}^l \binom{n}{j} p^j (1-p)^{n-j}$$

Bessere Approximation

$$\begin{aligned} P(k \leq S_n \leq l) &= P\left(\frac{k-np}{\sqrt{npq}} \leq \frac{S_n-np}{\sqrt{npq}} \leq \frac{l-np}{\sqrt{npq}}\right) \\ &\equiv \Phi\left(\frac{l-np+\frac{1}{2}}{\sqrt{npq}}\right) - \Phi\left(\frac{k-np-\frac{1}{2}}{\sqrt{npq}}\right) \\ &= \Phi(X_n, l) - \Phi(X_n, k) \\ &\quad \text{(Stetigkeitskorrektur)} \end{aligned}$$

Das Erstaunliche am ZGWS ist, dass das asymptotische stochastische Verhalten einer Summe von unabhängigen und identisch verteilten ZV nur vom Erwartungswert und der Varianz, nicht jedoch von der speziellen Gestalt der Verteilung bestimmt wird

25.4 Stirling-Formel

$$n! \sim n^n e^{-n} \cdot \sqrt{2\pi n}$$

26 Punktschätzer

26.1 Hypergeometrisches Modell und Binomialmodell

Grundgesamtheit GG zerfällt in zwei Teilmengen (*Dichotome GG*)

$$p = \frac{\text{Anzahl Elemente mit bestimmter Eigenschaft E}}{\text{Anzahl aller Elemente der GG}}$$

⇒ Schätzen von p aufgrund der Stichprobe

Einkleidung in Urnenmodell

- G hat N Kugeln
- r rote Kugeln haben Eigenschaft E

Modellierung: n-maliges Ziehen **ohne** Zurücklegen

$$H(n, r, N-r)(\{K\}) = \frac{\binom{r}{k} \binom{N-r}{n-k}}{\binom{N}{n}}$$

In der Praxis ist N sehr groß und man betrachtet Ziehen **mit** Zurücklegen

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

wobei $r/N = p$

Das einfachere Binomialmodell ist eine gute Approximation des hypergeometrischen Modells wenn N groß ist im Verhältnis zu n

- Wahrscheinlichkeitsrechnung: Binomialmodell ist bekannt: man studiert Verteilungseigenschaften und Wahrscheinlichkeiten $\{S_n = k\}$
- Statistik: konkrete Realisierung k von S_n : man schließt auf Parameter p

Punktschätzer: relative Trefferhäufigkeit

$$\hat{p} = \frac{k}{n}$$

konkrete Realisierung $R_n = \frac{S_n}{n}$ aus Stichprobe

Wahrscheinlichkeit, dass Schätzung richtig

$$P_p(R_n = p) \begin{cases} 0, & \text{falls } np \notin \{0, 1, \dots, n\} \\ \binom{n}{np} \cdot p^{np} \cdot (1 - p)^{n-np}, & \text{falls } np \in \{0, 1, \dots, n\} \end{cases}$$

- \hat{p} ist **erwartungstreu**, d.h. $E_p(R_n) = \frac{1}{n} \cdot E_p(S_n) = \frac{1}{n} \cdot n \cdot p = p$
- \hat{p} ist **konsistent**, d.h. $V(R_n) = \frac{1}{n^2} \cdot E_p(S_n) = \frac{p \cdot (1 - p)}{n}$ und die Schätzung ist damit umso genauer je größer das n ist, da Varianz abnimmt

26.2 Maximum-Likelihood-Schätzer

Das Modell unter welchem die beobachteten Daten die größte WK des Eintretens besitzen, wird gewählt

Likelihood-Funktion

$$L_k(p) := P_p(S_n = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

Das Ergebnis k ist fest und man unter sucht **die Wahrscheinlichkeit des Eintretens von k** unter verschiedenen durch einen Parameter p gekennzeichneten Modellen

$$L_k(p^*) = \max_{0 \leq p \leq 1} L_k(p) \text{ heißt Maximum-Likelihood-Schätzung}$$

Bilden der Ableitung und Nullsetzen: bestes $p = \frac{k}{n}$

Die relative Trefferhäufigkeit ist eine Maximum-Likelihood-Schätzung für p (i.A. allerdings nicht erwartungstreu)

27 Intervallschätzer

27.1 Idee

Angabe von Intervall mit plausiblen Parameterwerten für p

$$I_n = [p_u(R_n), p_o(R_n)] \subset [0, 1] \text{ mit } R_n = S_n/n$$

I_n heißt **Vertrauensintervall** für p zur **Vertrauenswahrscheinlichkeit** $1 - \alpha$, falls

$$\forall p \in (0, 1) : P_p(p \in I_n) = P_p[p_u(R_n) \leq p \leq p_o(R_n)] \geq 1 - \alpha$$

D.h., das zufällige Intervall I_n enthält mit Mindestwahrscheinlichkeit $1 - \alpha$ den Parameter p

- Vertrauensintervall = Konfidenzintervall
- Vertrauenswahrscheinlichkeit = Konfidenzniveau

27.2 Konstruktion des Intervalls

- Anwendung der Tschebyschow-Ungleichung für R_n
- Anstelle R_n , nutze Trefferzahl $S_n = R_n \cdot n$

27.3 Konfidenzgrenzen

- (i) Funktion $p_o(S_n)$ heißt **obere Konfidenzgrenze** für p zur **Konfidenzwahrscheinlichkeit** $1 - \beta$, falls

$$\forall p \in (0, 1) : P_p(p \leq p_o(S_n)) \geq 1 - \beta$$

- (ii) Funktion $p_u(S_n)$ heißt **untere Konfidenzgrenze** für p zur **Konfidenzwahrscheinlichkeit** $1 - \beta$

$$\forall p \in (0, 1) : P_p(p_u(S_n) \leq p) \geq 1 - \beta$$

Aus zwei einseitigen Konfidenzintervallen mit Konfidenzwahrscheinlichkeit $1 - \alpha/2$ lässt sich ein zweiseitiges Konfidenzintervall konstruieren

$$\{p_u(S_n) \leq p \leq p_o(S_n)\} = \{p_u(S_n) \leq p\} \cap \{p \leq p_o(S_n)\}$$

$$P(A \cap B) \geq P(A) + P(B) - 1$$

27.4 Konstruktion der Vertrauensgrenzen

Unterschreitungsschranke

$$\forall p \in [p_o(k), 1] : P_p(S_n \leq k) = \sum_{j=0}^k \binom{n}{j} \cdot p^j \cdot (1-p)^{n-j} \leq \beta$$

Überschreitungsschranke

$$\forall p \in [0, p_u(k)] : P_p(S_n \geq k) = \sum_{j=k}^n \binom{n}{j} \cdot p^j \cdot (1-p)^{n-j} \leq \beta$$

Beweis: mittels Fallunterscheidung

27.5 Approximative Konfidenzintervalle für große Stichprobenumfänge

mit Hilfe von ZGWS von Moivre-Laplace (Kapitel 25)

$$\begin{aligned} P_p(S_n \leq k) &= P_p \left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq \frac{k - np}{\sqrt{np(1-p)}} \right) \\ &\equiv \Phi \left(\frac{k - np + 0.5}{\sqrt{np(1-p)}} \right) = \beta \end{aligned}$$

Mittels Umkehrfunktion: Bestimmung von β

27.6 Planung des Stichprobenumfangs

z.B. Tschebyschow-Ungleichung

28 Der statistische Test

Signifikanter Wert: Unterschied zum reinen Raten bedeutungsvoll

28.1 Grundbegriffe

Beziehungen zwischen Schätz- und Testproblemen

Gemeinsamkeiten

- Daten als Realisierung einer ZV
- die für möglich erachteten Verteilungen bilden den Modell-Rahmen

Unterschiede

- Schätzproblem: Entscheidung für unbekannte Parameter
- Testproblem: Aufteilung der Menge in zwei disjunkte Teilmengen
 $\theta = \theta_0 + \theta_1$
Hypothese $H_0 : \vartheta \in \theta_0$
Hypothese $H_1 : \vartheta \in \theta_1$
 \Rightarrow Aufteilung des Stichprobenraums

Testtheorie

- Hypothesen H_0 und H_1 werden nicht symmetrisch behandelt
- H_0 ist die **Nullhypothese** (die man widerlegen möchte)
- H_1 ist die Alternativhypothese (die man bestätigt haben möchte)
- Entscheidung für H_1 : die Nullhypothese H_0 wird verworfen
- Entscheidung für H_0 : die Nullhypothese H_0 wird nicht verworfen

Fehler

Entscheidung	Wirklichkeit $\vartheta \in \theta_0$	Wirklichkeit $\vartheta \in \theta_1$
für H_0	richtige Entscheidung	Fehler 2. Art
für H_1	Fehler 1. Art	richtige Entscheidung

Fehler sind unvermeidbar, aber Wahrscheinlichkeiten für Fehlentscheidungen können durch vernünftige Test klein gehalten werden

- Fehler 1. Art
 - Nullhypothese wird fälschlicherweise abgelehnt
 - durch α kontrolliert
- Fehler 2. Art
 - Nullhypothese wird fälschlicherweise nicht abgelehnt
 - wird kleiner mit größerem Stichprobenumfang

28.2 Gütefunktion

Gütefunktion ordnet jedem ϑ die **Verwerfungswahrscheinlichkeit der Hypothese H_0**

\Rightarrow Vorgabe einer Schranke $\alpha \in (0, 1)$ für Fehler 1. Art

$\Rightarrow g(\vartheta) \leq \alpha \forall \vartheta \in \theta_0$

Ein solcher Test heißt **Signifikanz-Test** zum **Niveau α**

- Je kleiner α ist, desto bedeutungsvoller (signifikanter) stellt sich im Fall einer Ablehnung von H_0 der erhaltene Widerspruch zu H_0 dar.
- Kleines α dient der Sicherung der Alternative
- α berücksichtigt nicht Fehler 2. Art (Signifikanztest)
- die **Power** der Gütefunktion für $\vartheta \in \theta_1$ wird hier nicht eingeschränkt, sondern mit α maximiert

28.3 Philosophie des Testens

Test ist eine Art Widerspruchsbeweis: Man geht von der Annahme der Richtigkeit von H_0 aus. Wenn der beobachtete Wert der Teststatistik im Bereich der unplausiblen Werte liegt, dann betrachtet man dies als Widerspruch

- Ein Nichtverwerfen der Nullhypothese bedeutet nicht unbedingt, dass sie zutrifft, sondern nur, dass sie nicht genügend unplausibel ist, um verworfen zu werden
- Wer erst aufgrund eines Datensatzes zu einer Hypothese kommt, braucht neue Daten um diese Hypothese zu bestätigen
- α charakterisiert nur, dass bei Unterstellung der Gültigkeit von H_0 die Wahrscheinlichkeit für eine Ablehnung von H_0 höchstens α beträgt

29 Der Binomialtest

Modellannahme: $Bin(n, p), p \in (0, 1)$

$S_n \sim Bin(n, p)$ Anzahl der Therapie-Erfolge

p Erfolgswahrscheinlichkeit einer neuen Therapie

p_0 Erfolgswahrscheinlichkeit einer herkömmlichen Therapie

29.1 Der einseitige Binomialtest

$H_0 : p \leq p_0, H_1 : p > p_0$

Fehler 1. Art

Die Behauptung der Überlegenheit einer neuen Therapie

Fehler 2. Art

Das Nichtvertreten einer wahren Forschungshypothese

- Verwerfen von H_0 falls zu viele Treffer aufgetreten sind
- Gütefunktion

$$g_{n,k}(p_0) = \sum_{j=k}^n \binom{n}{j} \cdot p_0^j \cdot (1-p_0)^{n-j} \leq \alpha$$

- Fehlerniveau für α ausschöpfen und Fehler 2. Art möglichst klein zu halten

$$k = \min \left\{ l \in \{0, 1, \dots, n\} : \sum_{j=l}^n \binom{n}{j} \cdot p_0^j \cdot (1-p_0)^{n-j} \leq \alpha \right\}$$

Beispiel $p_0 = 1/2, n = 20, \alpha = 0.1$

$$g_{20,14}(0.5) = \left(\frac{1}{2}\right)^{20} \cdot \sum_{j=14}^{20} \binom{20}{j} \equiv 0.0577 \leq 0.1$$

- Die Nullhypothese $H_0 : p \leq 1/2$ wird zum Niveau $\alpha = 0.1$ abgelehnt, falls mindestens 14 Treffer beobachtet werden
- Die Nullhypothese $H_0 : p \leq 1/2$ wird zum Niveau $\alpha = 0.1$ nicht abgelehnt, falls weniger als 14 Treffer beobachtet werden

29.2 Der zweiseitige Binomialtest

zweiseitiges Testproblem

$$H_0 : p = p_0, H_1 : p \neq p_0$$

- zweiseitiger kritischer Bereich $\kappa_1 = \{0, 1, \dots, k_u\} \cup \{k_o, \dots, n\}$
- **Gütefunktion**

$$g_{n,k_u,k_o}(p) = \sum_{j=0}^{k_u} \binom{n}{j} \cdot p^j \cdot (1-p)^{n-j} + \sum_{j=k_o}^n \binom{n}{j} \cdot p^j \cdot (1-p)^{n-j}$$

- $k_u = k_u(n, \alpha/2), k_o = k_o(n, \alpha/2)$
- Spezialfall $p_0 = 1/2 \Rightarrow k_u(n, \alpha/2) = n - k_o(n, \alpha/2)$
kritischer Bereich ist symmetrisch zum Erwartungswert
- Man wird immer einseitig testen, wenn man sicher ist, dass der (wahre) Parameter nur in einer Richtung von dem Parameter der Nullhypothese abweichen kann

29.3 Der approximative Binomialtest

Im Fall $np(1-p) \geq 9$ kann man zu Testen der Nullhypothese die Normalapproximation der Binomialverteilung verwenden. Anstatt S_n nimmt man $\frac{S_n - np_0}{\sqrt{np_0(1-p_0)}}$

Es ist $P_{p_0} \left(\frac{S_n - np_0}{\sqrt{np_0(1-p_0)}} \leq x \right) \equiv \Phi(x)$

Mittels kritischem Wert $\Phi^{-1}(1-\alpha)$ ergibt sich Wahrscheinlichkeit α

29.4 Planung des Stichprobenumfangs

Bernoulli-Kette mit unbekanntem p

$H_0 : p \leq p_0, H_1 : p > p_0$

Unterschied $p > p_0$ ist **relevant** falls es p_1 gibt mit $p \geq p_1$ und $p_1 > p_0$

29.5 Zusammenhang zwischen Tests und Konfidenzintervallen

- Ein Konfidenzintervall enthält alle plausiblen Parameter, also alle Parameter, die mit den Beobachtungen vereinbar sind

Testproblem $H_0 : p = p_0$

Konfidenzintervall $[p_u(k), p_o(k)]$

Testproblem $H_0 : p = p'$

kritischer Bereich $\{0, 1, \dots, k_u(p')\} \cup \{k_o(p'), \dots, n\}$

Man definiert Menge I_n derjenigen Parameter p' bei denen der zweiseitige Test bei Vorliegen der konkreten Trefferhäufigkeit k die Nullhypothese $H_0 : p = p'$ nicht ablehnt

$$I_n(k) := \{p' \in (0, 1), k \in \kappa_0(p')\} = \{p' \in (0, 1) : k \notin \kappa_1(p')\}$$

$I_n(k)$ ist Konfidenzintervall zum Niveau $1 - \alpha$

- Konfidenzintervall ist informativer als Test
- Konstruktion von Konfidenzintervallen mit Tests
- Gute Tests zum Niveau α , also Tests mit großer Power, liefern auch gute (kurze) Konfidenzintervalle zur Vertrauenswahrscheinlichkeit $1 - \alpha$