

Grundlagen der Informatik

WS 20/21

Inhaltsverzeichnis

1	Lineare Optimierung	4
1.1	Simplexalgorithmus	4
2	Finanzmathematik	5
2.1	Folgen und Reihen	5
2.2	Einfache Verzinsung	5
2.2.1	Verzinsung mit Zinseszins	5
2.2.2	Stetige Verzinsung	5
2.2.3	Gemischte Verzinsung	6
2.2.4	Freigeld	6
2.3	Investitions- und Finanzrechnung	6
2.4	Rentenrechnung	6
2.4.1	Zinsperiode = Ratenperiode	6
2.4.2	Ratenperiode < Zinsperiode	7
2.5	Tilgungsrechnung	7
2.5.1	jährliche Ratentilgung	7
2.5.2	unterjährliche Ratentilgung	7
3	Deskriptive Statistik	8
3.1	Stichproben	8
3.2	Häufigkeiten, empirische Verteilungsfunktion	8
3.3	Klasseneinteilung	8
3.4	Statistische Parameter	9
3.5	Zweidimensionale Stichproben	9
3.6	Lineare Regression	10
4	Kombinatorik	11
4.1	Einführung	11
4.2	Permutation	11
4.3	Variation	11
4.4	Kombinationen	11
5	Wahrscheinlichkeitstheorie	12
5.1	Zufallsexperimente	12
5.2	Laplace Wahrscheinlichkeit	12
5.3	Axiome von Kolmogorow	12
5.4	Bedingte Wahrscheinlichkeit	13
5.5	Satz von Bayes	13
5.6	Unabhängigkeit von Ereignissen	13
6	Zufallsvariable und statistische Verteilung	14
6.1	Zufallsvariable	14

6.2	Spezielle diskret Verteilungen	14
6.2.1	Gleichverteilung	14
6.2.2	Bernoulli-Verteilung	15
6.2.3	Binomialverteilung	15
6.2.4	Geometrische Verteilung	16
6.2.5	Hypergeometrische Verteilung	16
6.2.6	Poisson-Verteilung	16
6.3	stetige Zufallsvariable	17
6.4	Spezielle stetige Verteilungen	17
6.4.1	Gleichverteilung	17
6.4.2	Normalverteilung	17
6.4.3	Exponentialverteilung	18
6.5	Notizen	18
7	Induktive Statistik und statistische Tests	18
7.1	Schätzfunktion	18
7.2	Maximum Likelihood Methode	18
7.3	Hypothesentest	19
7.4	Zentraler Grenzwertsatz	19
7.5	Varianztest	20

1 Lineare Optimierung

Unter Linearer Optimierung versteht man, dass eine bestimmte Funktion durch einstellen bestimmter 'Parameter' möglichst groß (oft für Gewinn benutzt) oder möglichst klein (für Verluste) wird. Dafür wird die lineare Zielfunktion:

$$g(\vec{x}) = g_0 + g_1x_1 \dots g_nx_n$$

hergenommen. Zum Beispiel ist g_i dann der Preis für die Herstellung eines Produkts und x_i dann die Anzahl wie oft das Produkt hergestellt werden soll und x_i sind eben die Parameter die eingestellt werden sollen. Dazu sind idR. Randbedingungen gegeben, wie ' x_1 darf höchstens doppelt so oft hergestellt werden wie x_2 '. Dann gibt es noch triviale Randbedingungen. Da man nicht negativ produzieren kann gilt meistens auch $x_i \geq 0$. Somit erhält man einige Ungleichungen und Gleichung. Zuerst schreibt man alle (Un)Gleichungen untereinander auf und dreht jedes Ungleichungszeichen (Multiplikation mit -1) um damit alle Ungleichungszeichen in die selbe Richtung zeigen. Nun versucht man Variablen zu eliminieren indem man die Gleichungen die man hat nach einer Variable umstellt und dann diese Variable in den Ungleichungen ersetzt. Beispiel:

$$x_1 + x_2 = 0 \Rightarrow x_1 = -x_2$$

x_1 kann in diesem Beispiel mit $-x_2$ in den Ungleichungen ersetzt werden. Wenn man die ganzen Gleichungen in ein Diagramm einzeichnet erhält man ein Vieleck mit Kanten die eben von den Ungleichungen dargestellt werden. Man kann jetzt jeden Schnittpunkt anschauen, die Werte berechnen die herauskommen wenn man die x-Koordinaten in die Gewinnungleichung einsetzt und vergleicht wie jeden Wert und sucht sich den höchsten/niedrigsten Wert als Lösung heraus. Bei vielen Parametern kann das schnell viele Berechnungen benötigen. Man kann auch das Diagramm wieder hernehmen. Hier sucht man sich einen geeigneten Startpunkt (eine Punkt an einer Kante) und schaut ob man ob die Kante ansteigt oder waagrecht verläuft und folgt dieser Kante bis man einen Schnittpunkt findet. Hier wird wiederum geschaut ob es eine Kante gibt die ansteigt oder waagrecht verläuft. Dies wird wiederholt bis das nicht mehr möglich ist. Der Letzt gefundene Schnittpunkt, welcher noch das Vieleck umspannt, erhält dann die Parameter für die Lösung.

1.1 Simplexalgorithmus

Da man schwer vier, fünf oder sechs Dimensionen zeichnen kann gibt es den Simplexalgorithmus welcher den Optimalwert rechnerisch ermitteln kann. Hier muss zuerst ein Tablett erstellt werden. Die Form des Tablets ist wie folgt:

$$\left(\begin{array}{cc} A & \vec{k} \\ \hline \vec{g} & G \end{array} \right) \text{ Wobei Die Matrix A}$$

Die Ungleichungen in Matrixform ist, $\vec{g} = G$ eine Lösung ist von der man weiß, dass sie die Gewinnfunktion erfüllt und sich $\vec{k} = A\vec{u} - \vec{b}$. Dabei ist \vec{u} auch der Vektor der die Gewinnfunktion erfüllt.

2 Finanzmathematik

2.1 Folgen und Reihen

Eine Zahlenfolge ist eine Abbildung, welche wie folgt definiert ist:

$$a = \mathbb{N} \rightarrow \mathbb{R}$$

Anstatt $a(i)$ schreibt man a_i für jedes einzelne Glied der Folge. Eine Reihe ist eine Aufsummierung aller Folgeglieder und wird unendliche Reihe genannt wenn unendlich viele Glieder aufsummiert werden bzw. endliche Reihe genannt wenn bis zu einem bestimmten index aufsummiert wird.

arithmetische Zahlenfolge: $a_n = a_1 + (n - 1)d, d \in \mathbb{R}$ fest

endliche arithmetische Reihe: $\frac{N(2a_1 + (N + 1)d)}{2}$

geometrische Zahlenfolge: $a_n = a_1 \cdot q^{n-1}$

geometrische Reihe: $a_1 \frac{q^N - 1}{q - 1}$

2.2 Einfache Verzinsung

Sollen Zinsen nicht miteinander verzinst werden (Zinsen werden auf ein anderes Konto ausgezahlt) spricht man von einfacher Verzinsung. Es gilt folgenden Formel:

$$K_n = K_0(1 + n \cdot i)$$

Wobei K_n das Kapitale nach n Zinsperioden darstellt. K_0 ist dabei das Startkapital, i die Zinsrate¹ und n eben die Laufzeit in Zinsperioden (meistens Monate oder Jahre).

2.2.1 Verzinsung mit Zinseszins

Hier werden die Zinsen mitverzinst (Zinsen werden aufs selbe Konto ausgezahlt). Die Formel lautet:

$$K_n = K_0(1 + i)^n$$

2.2.2 Stetige Verzinsung

Unter stetiger Verzinsung versteht man, dass die Zinsen nicht jeden Monat oder jeden Tag angerechnet werden sondern zu jedem Augenblick. Dafür wird folgende Formel verwendet:

$$K_n = K_0 e^{i \cdot n}$$

¹ $i \cdot 100 \equiv \text{Zinsrate in Prozent}$

2.2.3 Gemischte Verzinsung

Wenn man als Zinsperiode ein ganzes Jahr hat aber man die Verzinsung mitten in der Zinsperiode 'abbrechen' will wird folgende Formel benutzt:

$$K_n = K_0(1+i)^{\lfloor n \rfloor} \cdot (1+i(n-\lfloor n \rfloor)^2)$$

2.2.4 Freigeld

Die Idee hinter Freigeld ist dass Geld nicht wie Lebensmittel oä. verderben kann. Um das zu korrigieren kann man die Formel für einfache Verzinsung (wenn man Bargeld hat) bzw. die Formel für den Zinseszins hernehmen (Girokonto). Hier nimmt man als Zinsrate einfach den Negativwert.

2.3 Investitions- und Finanzrechnung

Die Investitionsrechnung wird verwendet um zu bestimmen ob sich eine Investition im Laufe der Zeit lohnen wird. Zuerst zieht man die Kosten der Investition vom Kapital ab. Gegenfalls schaut man ob es periodische Zahlungen gibt die man leisten muss (Wartungsarbeiten, Kreditszinsen etc.) und zieht diese auch ab. Man schätzt nun den Gewinn den die Investition im Monat/Jahr etc. macht, zieht noch eventuelle Zahlungen ab und erhält den Monats/Jahresgewinn. Dies macht man für alle Monate für die eine Schätzung machbar ist, verzinst diese dementsprechend und vergleicht den Gesamtgewinn mit dem Gewinn den man erhält wenn man nicht investiert mit dem über dem selben verzinsten Kapital, das man hat wenn man nicht investiert.

2.4 Rentenrechnung

Unter einer Rente versteht man im Allgemeinen eine regelmäßige Ein- oder Auszahlung. Es wird zwischen auch noch eine Unterscheidung getroffen wenn die Zinsperiode der Rentenperiode gleicht (man zahlt jeden Monat einen gewissen Betrag aufs Spargbuch und bekommt jeden Monat auch Zinsen auf das Geld) und wenn die Rentenperiode kleiner der Zinsperiode ist (man zahlt monatlich ein bekommt aber nur jährlich Zinsen).

2.4.1 Zinsperiode = Ratenperiode

Die allgemeine Formel lautet:

$$R_n = r \cdot q \cdot \frac{q^n - 1}{q - 1}$$

r ist die Ratenzahlung, n wieder die Anzahl der Ratenperioden, q berechnet sich durch $q = 1 + i$, wobei i wieder der Zins ist. Die Formel gilt für vorschüssige Ratenzahlungen und nachschüssige Zinszahlungen. Für die nachschüssige Ratenzahlung gilt:

$$R_n = r \cdot \frac{q^n - 1}{q - 1}$$

²Die Klammern sind Gaußsche Rundungsklammern. Sind sie wie hier nach unten gerichtet, wird abgerundet und wenn sie nach oben gerichtet sind, wird aufgerundet

2.4.2 Rentenperiode < Zinsperiode

Wenn die Rentenperiode kleiner der Zinsperiode ist muss man nur die Ratenzahlung r aus der Formel für die nachschüssige Ratenzahlung ändern. Es gilt für das neue r_e :

$$r_e = r \cdot \left(m + \frac{i}{2} \cdot (m + 1)\right) \Rightarrow R_n = r_e \frac{q^n - 1}{q - 1} = r \cdot \left(m + \frac{i}{2} \cdot (m + 1)\right) \cdot \frac{q^n - 1}{q - 1}$$

Wobei das m die Anzahl der Ratenzahlungen pro Zinsperiode ist.

2.5 Tilgungsrechnung

Bei der Tilgungsrechnung geht es idR. um die Rückzahlung eines Kredit (bzw. allgemein Schulden). Die Annuität bezeichnet damit den Betrag, der periodisch zurückgezahlt werden muss. Die Annuität setzt sich wiederum aus der Tilgungsrate (Betrag der zurückgezahlt werden soll) und dem Zinsbestandteil (Zinsen die pro Periode von den Restschulden dazukommen).

2.5.1 jährliche Ratentilgung

Soll die Tilgungsrate jedes Jahr (Periode) gleich bleiben ergibt sich folgende Formel:

$$T = \frac{S}{n}$$

Wobei T die Gesamtschuld ist und n die Anzahl der Jahre (perioden) der Kreditlaufzeit. Will man die Zinsrate für ein bestimmtes Jahr (hier r) ausrechnen gilt folgende Formel:

$$Z_r = T \cdot (n - r + 1) \cdot i$$

i ist dabei wieder der Zinssatz, n Die Kreditlaufzeit und T die Tilgungsrate.

Die Annuität ergibt sich dann wenn man die beiden speziellen Formeln nun in die allgemeine Formel von oben einsetzt und umformt:

$$A = T \cdot (1 + (n - r + 1)i)$$

Für die Annuitätentilgung gilt folgende Formel:

$$A = S \cdot q^n \frac{q - 1}{q^n - 1}$$

2.5.2 unterjährliche Ratentilgung

Soll mehrmals im Jahr getilgt werden muss der jährliche Zinssatz durch die Anzahl der Tilgungen im Jahr geteilt werden. Weiter beschreibt das n nicht mehr das Jahr sondern die Anzahl der Zinsperioden.

3 Deskriptive Statistik

3.1 Stichproben

Bei der Statistik versucht man allgemeine Aussagen über einen bestimmten Datensatz zu erheben. Bei Millionen von Daten ist es oft nicht sinnvoll alle Daten einzeln zu betrachten. Die Datenmenge wird auch als Beobachtungsmenge (B) bezeichnet. Jedem Element aus B wird dann ein bestimmtes Merkmal zugeordnet. Die Merkmale können diskret³ oder stetig⁴ sein. Weiter ist Urliste eine Teilmenge aus B mit den zugehörigen Merkmalswerten. Die Merkmalswerte aus der Urliste bilden dann die Stichprobe, welche geordnet oder ungeordnet sein kann. Der Stichprobenumfang (n) ist dann die Kardinalität der Stichprobenmenge.

3.2 Häufigkeiten, empirische Verteilungsfunktion

Bei Häufigkeiten wird zwischen der absoluten und relativen Häufigkeiten unterschieden. Dafür nimmt man die zuerst die Funktion $g(x)$ her, welche zählt wie oft ein Element x in der Stichprobe vorkommt. Dies ist dann die absolute Häufigkeit. Teilt man die absolute Häufigkeit durch n ⁵.

absolute Häufigkeit: $g(x) \neq 0 \Leftrightarrow x \in \{x_1, x_2, \dots, x_n\} \subset \mathbb{R} \ (\mathbb{N})$

relative Häufigkeit: $h(x) = \frac{1}{n}g(x)$

Die Summenhäufigkeit ($G(x)$) summiert alle Werte aus der Stichprobe bis einschließlich x . Die relative Häufigkeit nimmt die Summenhäufigkeit und teilt diese durch den Stichprobenumfang.

Summenhäufigkeit $G(x_i) = x_1 + x_2 + x_3 + \dots x_i \mid x_i < x_n$

relative Summenhäufigkeit: $H(x) = \frac{1}{n}G(x)$

Die relative Summenhäufigkeit wird auch die empirische Verteilfunktion genannt. Durch die Summenhäufigkeit werden Häufigkeiten zusammengefasst und man kann allgemeinere Aussagen für die zusammengefassten Häufigkeiten machen.

3.3 Klasseneinteilung

Oft ist es sinnvoll Stichprobenwerte in Klassen einzuteilen. Dafür wählt man zuerst die Anzahl an Klassen die man haben möchte (hie: l). Nun teilt man die den Stichprobenumfang durch l . Dadurch erhält man die Klassenbreite w (ggf. muss gerundet werden). Die Klassenbreite gibt an wie viele Einzelemenet zusammengefasst werden soll (durch Addition). Hat man alle Klassen zusammengefasst kann man nun allgemeinere Aussagen

³ \mathbb{N}_0 also alles was abzählbar ist wie Gehalt oder Anzahl von Besuchern etc.

⁴Intervalle in \mathbb{R} , beliebig genau ermittelbare Werte wie Zeit Länge

⁵Stichprobenumfang

über bestimmte Teile der Stichprobe machen (Aussagen über den Verdienst von Personen die weniger als 20h, zwischen 20h und 40h und mehr als 40h die Woche arbeiten als Beispiel).

Für die Klasseneinteilung sollte man sich an diese Regeln halten:

- Möglichst gleich große Klassen
- $5 \leq l^6 \leq 25$
- $l \approx \sqrt{n}$
- Klassengrenzen sollten ganze Zahlen sein
- Klassen sollten den Wertebereich der Stichprobe möglichst gut abdecken

3.4 Statistische Parameter

Desto allgemeiner man eine Stichprobe behandelt, desto übersichtlicher aber auch desto ungenauer werden die Informationen. Es gibt statistische Parameter welche die Stichprobe in eine einzelne Zahl zusammenfassen:

arithmetisches Mittel/Stichprobenmitte:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Median für n = ungerade:

$$\tilde{x} = x_{(\frac{n+1}{2})}$$

Median für n = gerade:

$$\tilde{x} = \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1})$$

Spannweite:

$$v := x_{(n)} - x_{(1)}$$

Varianz:

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

Streuung/Standardabweichung:

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

3.5 Zweidimensionale Stichproben

Eine Zweidimensionale Stichprobe ist wie eine normale Stichprobe nur dass da jetzt noch ein weiteres Merkmal auftritt, wodurch Zahlenpaare entstehen (zB. Anstatt nur Einkommen als Merkmal wird jetzt Einkommen und Steuern untersucht). Für zweidimensionale

⁶Anzahl der Klassen

Stichproben gelten (mit kleinen Änderungen) die selben Definitionen:

Stichenprobenumfang:	$n = \text{Anzahl der } (x,y) \text{ Zahlenpaare}$
absolute Häufigkeit:	$g(x, y) \neq 0 \Leftrightarrow x, y \in \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$
relative Häufigkeit:	$h(x, y) = \frac{1}{n}g(x, y)$
Summenhäufigkeit:	$G(x, y) = \sum_{\{(a,b): a \leq x, b \leq y\}} g(a, b)$
empirische Verteilungsfunktion:	$H(x, y) = \sum_{\{(a,b): a \leq x, b \leq y\}} h(a, b)$
Klassenhäufigkeiten:	$\tilde{h}_{ij} = \sum_{x \in KX_i, y \in KY_j} h(x, y)$
Randhäufigkeiten für ein festes x/y :	$h_{x/y}(x/y) = \sum_{x/y \in \mathbb{R}} h(x, y)^7$

3.6 Lineare Regression

Zeichnet man die Stichprobenwerte in ein Diagramm ein erhält man einige Punkt in einem Diagramm. Mit der linearen Regression sucht man eine Funktion, welche ungefähr der Richtung der Punkten entspricht. Für die Funktion $y = ax + b$ müssen a und b so gewählt werden, damit der Abstand von der Geraden zu den Punkten möglichst klein wird. Die einzelnen Abstände Δ werden dann quadriert und danach zusammenaddiert. a und b sollen so gewählt werden, dass Δ^2 möglichst klein wird. Diese Methode zur Ermittlung der Regressionsgeraden heißt Methode der kleinsten Quadrate". Die Kovarianz bei zweidimensionalen Stichproben wird wie folgt ermittelt:

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} \sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}$$

Teilt man die Kovarianz durch die Standardabweichung von jeweils x und y erhält man den empirischen Korrelationskoeffizienten:

$$r_{xy} = \frac{s_{xy}}{s_x \cdot s_y}$$

Weiter gilt für den Korrelationskoeffizienten r_{xy} :

- $-1 \leq r_{xy} \leq 1$
- Wenn $r_{xy} = 1$ oder $r_{xy} = -1$ liegen die Messpunkte direkt auf einer Geraden
- Wenn $r_{xy} = 0$ dann haben die Merkmale X und Y keinen Zusammenhang
- Gibt es mehr als 2 Merkmale kann man selber Paare bilden und sieht dadurch wie stark Merkmale direkt miteinander zusammenhängen

4 Kombinatorik

4.1 Einführung

Bei der Kombinatorik unterscheidet man zwischen der Permutation, Variation und Kombination und in diesen 3 Gruppen wird nochmals unterschieden ob Wiederholungen auftreten können oder nicht.

4.2 Permutation

Eine Anordnung von n unterscheidbaren Objekten (es können nicht zwei oder mehr gleiche Objekte auftreten) heißt Permutation ohne Wiederholung (P_n)

$$P_n = n!$$

Wobei n die Anzahl der ausgewählten Objekte ist.

Bei der Permutation mit Wiederholung können verschiedene Sorten auftreten (Sorten sind Mengen mit gleichen Objekten). Für n Objekte und k Sorten gilt:

$$n = n_1 + n_2 + \dots + n_k \quad | \quad n_1, n_2, \text{ etc sind die Sorten}$$
$$P_n^{(n_1, n_2, \dots, n_k)} = \frac{n!}{n_1! \cdot n_2! \cdot \dots \cdot n_k!}$$

4.3 Variation

Die Variation ist eine Anordnung von k Objekten aus einer Menge aus n Objekten. Wenn ein Objekt mehrfach vorkommen kann spricht man von Variation mit Wiederholung ($V_n^{(k)}$) ansonsten spricht man von einer Variation ohne Wiederholung ($V_n^{(k)}$). Es gilt:

$$V_n^{(k)} = n(n-1)(n-2) \dots (n-k+1) = \frac{n!}{(n-k)!}$$
$$V_{w_n}^{(k)} = n^k$$

4.4 Kombinationen

Bei der Kombination ist im Vergleich zu Variation die Reihenfolge der Objekte nicht wichtig. Sind alle Objekte verschieden spricht man von einer Kombination ohne Wiederholung ($C_n^{(k)}$). Es gilt dann:

$$C_n^{(k)} = \binom{n}{k} = \frac{V_n^{(k)}}{k!}$$

Wenn aus einer Menge von Objekten n Sorten existieren und k Objekte ausgewählt werden und man dazu noch die Reihenfolge nicht berücksichtigt spricht man von einer Kombination mit Wiederholung ($C_{w_n}^{(k)}$). Es gilt:

$$C_{w_n}^{(k)} = \frac{(n-1+k)!}{k!(n-1)!} = \binom{n-1+k}{k}$$

5 Wahrscheinlichkeitstheorie

5.1 Zufallsexperimente

Ein Zufallsexperiment ist ein beliebig wiederholbarer Versuch, bei dem eine bestimmte Menge als Ausgang möglich ist, die auch bei jeder Wiederholung gleich bleiben muss. Der Ausgang kann dabei auch nicht vorhergesagt werden.

- Jeder einzelnen Ausgang heißt Ergebnis (oa. Elementarereignis) und wird mit ω bezeichnet.
- Die Menge aller Ereignisse heißt Ω
- Eine Teilmenge von Ω heißt Ereignis
- Ist A eine Teilmenge von Ω so ist \bar{A} das komplementäre Ereignis zu A . \bar{A} beinhaltet alle Ergebnisse die nicht in A vorkommen $\Rightarrow \bar{A} = \Omega \setminus A$

TODO: Ereignisalgebra

5.2 Laplace Wahrscheinlichkeit

Zufallsexperimente haben endlich viele Ereignisse. Um jetzt die Wahrscheinlichkeit eines bestimmten Ereignis mit anderen einfacher vergleichen zu können teilt man die Wahrscheinlichkeit A durch die Anzahl der Elemente aus Ω ⁸. Die Wahrscheinlichkeit ist jetzt in einem Intervall zwischen 1 und 0. Es gilt:

$$p(A) = \frac{|A|}{n} = \frac{\text{Anzahl der Elemente aus } A}{\text{Anzahl der Elemente aus } \Omega}$$

5.3 Axiome von Kolmogorow

Möchte man Zufallsexperimente durchführen, wobei die Ergebnisse verschiedene Wahrscheinlichkeiten haben, könnte man einen Versuch sehr oft durchführen um eine Näherung einer Wahrscheinlichkeit zu bekommen. Dies kann oft nicht durchgeführt werden, weshalb man Eigenschaften definiert hat, die Wahrscheinlichkeiten erfüllen müssen.

Ein Zufallsexperiment hat der Ereignismenge Ω , wobei \mathcal{A} eine Ereignisalgebra auf Ω ist. Nun heißt die Abbildung $p : \mathcal{A} \rightarrow \mathbb{R}$ Wahrscheinlichkeit oder auch Wahrscheinlichkeitsmaß wenn gilt:

1. $0 \leq p(A) \leq 1$
2. $p(\Omega) = 1$
3. Sind die Ereignisse A_1, A_2, \dots, A_n paarweise verschieden gilt $\bigcup_{i=1}^n A_i = \sum_{i=1}^n p(A_i)$

Für das Wahrscheinlichkeitsmaß gilt außerdem:

⁸Bei Laplace wird angenommen dass alle Ereignisse die selbe Wahrscheinlichkeit haben

- a) $p(\{\}) = 0$
- b) A und B sind Ereignisse: $(A \cap B) = \{\} \Rightarrow p(A \cup B) = p(A) + p(B)$
- c) $p(A) + p(\overline{A}) = 1$
- d) A und B sind Ereignisse: $p(A \cup B) = p(A) + p(B) - p(A \cap B)$
- e) A und B sind Ereignisse: $A \subset B \Rightarrow p(A) \leq p(B)$

5.4 Bedingte Wahrscheinlichkeit

Bedingte Wahrscheinlichkeiten setzen voraus dass eine bestimmtes Ereignis schon eingetroffen und man nun die Wahrscheinlichkeit der 'nächsten' Wahrscheinlichkeit berechnen möchte.

Wenn A und B Ereignisse sind, $p(B) \neq 0$ dann heißt

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

bedingte Wahrscheinlichkeit von A unter der Bedingung von B . Daraus folgt

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

Satz der totalen Wahrscheinlichkeit:

B sein ein Ereignis. A_1, A_2, \dots, A_n sind paarweise diskunkte Ereignisse die zusammen Ω ergeben $\Omega = \bigcup_{i=1}^n A_i$, dann gilt:

$$p(B) = \sum_{i=1}^n p(B|A_i) \cdot p(A_i)$$

5.5 Satz von Bayes

A_1, A_2, \dots, A_n sind paarweise disjunkte Ereignisse, $\Omega = \bigcup_{i=1}^n A_i$, B ist ein Ereignis und es gilt $p(B) > 0$ außerdem ist $i, j \in \mathbb{N}$, so gilt:

$$P(A_i|B) = \frac{p(B|A_i) \cdot p(A_i)}{\sum_{j=1}^n p(B|A_j) \cdot p(A_j)}$$

5.6 Unabhängigkeit von Ereignissen

A, B seien Ereignisse, A ist Unabhängig von B wenn gilt:

$$p(A) = p(A|B)$$

Daraus folgt, dass auch gelten muss:

$$p(B) = p(B|A)$$

6 Zufallsvariable und statistische Verteilung

6.1 Zufallsvariable

Eine Zufallsvariable ist eine Funktion die jedem Ereignis aus Ω eine Zahl in \mathbb{R} zuordnet. Die Zufallsvariable X ist somit eine Funktion und kein Wert.

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

X heißt diskrete Zufallsvariable, wenn deren Wertebereich endlich oder abzählbar unendlich ist.

Die Funktion $f(X) : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(X) = \begin{cases} p_i & \text{wenn } x = x_i \\ 0 & \text{für den Rest} \end{cases}$$

heißt Wahrscheinlichkeitsfunktion bzw. Dichtefunktion der diskreten Zufallsvariablen X . Die Verteilung von X ist definiert als

$$p_X(A) = \sum_{x_i \in A} f_X(x_i)$$

wobei $p_X = 2^{W_X} \rightarrow [0; 1]$, also die Potenzmenge des Wertebereichs. Weiter ist die Verteilungsfunktion definiert als

$$F_X : \mathbb{R} \rightarrow \mathbb{R} \text{ mit } F_X(x) = p(X \leq x) = \sum_{x_i < x} f_X(x_i)$$

Erwartungswert:

$$E(X) = \sum_{x_i \in W_X} x_i \cdot f_X(x_i)$$

Varianz:

$$\sum_{x_i \in W_X} (x_i - E(X))^2 \cdot f_X(x_i)$$

Varianz und Erwartungswert sind linear, dh. wenn eine neue Zufallsvariable gemacht wird, die (wie folgt) abhängig von Konstanten ist ($Z_n := a \cdot X + b$) dann gilt

$$E(Z_n) = a \cdot E(X) + b \mid \text{für den Erwartungswert}$$

$$Var(z) = a^2 \cdot Var(X) \mid \text{für die Varianz}$$

6.2 Spezielle diskretet Verteilungen

6.2.1 Gleichverteilung

Eine Verteilung mit der Dichte

$$f(x) = \begin{cases} \frac{1}{n} & \text{wenn } x \in W_X(|W_X| = n) \\ 0 & \text{für den Rest} \end{cases}$$

heißt Gleichverteilung. Für den Erwartungswert gilt

$$: E(X) = \frac{1}{n} \sum_{x \in W_X} x$$

Ist $W_x = 1, 2, 3, \dots, n$ gilt:

$$E(X) = \frac{n+1}{2}$$

6.2.2 Bernoulli-Verteilung

Wenn die Zufallsvariable nur 2 Werte liefern kann, spricht man von einer Bernoulli-Verteilung, es gilt:

- Dichte:

$$f(x) = \begin{cases} p & \text{für } x = 1 \\ 1 - p & \text{für } x = 0 \\ 0 & \text{für den Rest} \end{cases}$$

- Erwartungswert: $E(X) = p$
- Varianz: $V(X) = pq = p(1 - q)$ ⁹

6.2.3 Binomialverteilung

Die Binomialverteilung ist die Bernoulli-Verteilung n Mal durchgeführt, bzw. n Objekte aus denen x untersucht werden wenn nur eine Wahrscheinlichkeit p vorhanden ist.

- Dichte:

$$f(x) = \begin{cases} \binom{n}{x} p^x \cdot q^{n-x} & \text{für } x \in \{0, 1, 2, \dots, n\} \\ 0 & \text{für den Rest} \end{cases}$$

- Erwartungswert: $E(X) = \sum_{x=0}^n x \binom{n}{x} \cdot p^x \cdot q^{n-x} = n \cdot p$
- Varianz: $V(X) = \sum_{x=0}^n (x - n \cdot p)^2 \cdot \binom{n}{x} \cdot p^x \cdot q^{n-x} = n \cdot p \cdot q$

p ist dabei die Wahrscheinlichkeit für's Eintreten des Ereignis. q ist definiert durch $q = 1 - p$ und x ist die Anzahl wie oft das Ereignis eintreten soll.

⁹ $p = p(X = 1), q = p(X = 0)$

6.2.4 Geometrische Verteilung

Anzahl der Versuche die gebraucht werden bis der 1. Erfolg bei einem Bernoulli-Experiment eintritt. n ist dabei die Anzahl der Versuche (Wie hoch ist die Wahrscheinlichkeit, dass das Ereignis bis zum n -ten Versuch eintritt).

- Dichte:

$$f(x) = \begin{cases} q^n p & \text{für } n \in \mathbb{N}_0 \\ 0 & \text{für den Rest} \end{cases}$$

- Erwartungswert: $E(X) = \frac{q}{p}$

- Varianz: $V(X) = \frac{q}{p^2}$

6.2.5 Hypergeometrische Verteilung

Aus einer Grundmenge (Gesamtmenge) N werden M Elemente eine bestimmte Eigenschaft (Aus 10 (N) Lampen sind 2 (M) defekt). Aus N werden nun n Elemente betrachtet. f gibt dann die Wahrscheinlichkeit für x auftauchende Elemente an (Aus 10 Lampen sind 2 defekt, es werden 5 (n) Lampen untersucht. Wie Wahrscheinlich ist es dass eine (x) aus den 5 Lampen kaputt ist).

- Dichte: $\frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}$

- Erwartungswert: $E(X) = n \cdot \frac{M}{N}$

- Varianz: $V(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}$

6.2.6 Poisson-Verteilung

Bei der Poisson-Verteilung ist der Erwartungswert λ oft schon gegeben, da man gerade ihn zur Berechnung für die Dichte benötigt. Poisson eignet sich für Stichproben mit hohem Umfang ($n = 100, 1000, \text{etc}$) aber niedrigem Erwartungswert ($\lambda = 2$) Es gilt:

- Dichte: $f(x) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$

- Erwartungswert: $E(X) = \lambda$

- Varianz: $V(X) = \lambda$

Dabei ist k zum Beispiel Anzahl der defekten Geräte.

6.3 stetige Zufallsvariable

Eine Zufallsvariable heißt stetig wenn W_X ein Intervall in \mathbb{R} , die Dichtefunktion f_X gilt, dass sie nie negativ ist und und die Verteilfunktion $F_X(x)$ das Integral aus f_X , also

$$F_X(x) = \int_{-\infty}^x f_X(u) du.$$

- Erwartungswert: $E(X) = \int_{-\infty}^{\infty} x \cdot (f_X(x)) dx$
- Varianz: $V(X) = \int_{-\infty}^{\infty} (x - E(X))^2 \cdot f_x(x) dx = E(X^2) - E(X)^2$

6.4 Spezielle stetige Verteilungen

6.4.1 Gleichverteilung

- Dichte:

$$f(x) = f(x, a, b) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{für den Rest} \end{cases}$$

- Erwartungswert: $E(X) = \frac{a+b}{2}$
- Varianz: $V(X) = \frac{(b-a)^2}{12}$

6.4.2 Normalverteilung

Die Normalverteilung ist die häufigste Verteilung in der Praxis.

Dabei ist $W_X = \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, somit gilt:

- Dichte: $f(x, \mu, \sigma) = \varphi(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Erwartungswert: $E(X) = \mu$
- Varianz: $V(X) = \sigma^2$
- Standardnormalverteilung ($\mu = 0, \sigma = 1$): $\Phi(x) = \Phi(x, 0, 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$

Umrechnung von X ins Standardnormalverteilung Z :

$$Z(X) = \frac{X - \mu}{\sigma} = \frac{X - E(X)}{\sqrt{Var(x)}}$$

6.4.3 Exponentialverteilung

Die Exponentialverteilung ist ein Model für die Lebensdauer für Produkte.

- Dichte: $f(x) = f(x, \lambda) = \lambda e^{-\lambda x}$ für $x \geq 0$; ansonsten 0 für $x < 0$
- Erwartungswert: $E(X) = \frac{1}{\lambda}$
- Varianz: $Var(X) = \frac{1}{\lambda^2}$

6.5 Notizen

Das Intervall wird durch die Integration der Dichtefunktion ermittelt.

Varianz und Erwartungswert haben den Wert 0 wenn das die Intervallgrenzen gleich sind (bzw. wenn nur ein Punkt gegeben ist und nicht zwei)

Auswahltabelle

	diskret?	stetig?
anzahl?	Binomial/Hypergeometrische	Poisson
erstes Mal?	geometrisch	Exponential

7 Induktive Statistik und statistische Tests

7.1 Schätzfunktion

Die Schätzfunktion \bar{X} gibt dann einen Wert für bestimmte Parameter (zB. Erwartungswert) Nun Werden verschiedene Stichproben erhoben und (unter Beachtung der vermuteten Verteilung) werden dann die Parameter (zB. Erwartungswert) ausgerechnet. Ist zB. $\bar{X} = E(X)$ ist \bar{X} erwartungstreu für $E(X)$. Allgemein ist jede Funktion $U = g(X_1, X_2, \dots, X_n)$ eine Schätzfunktion für einen Parameter a einer Verteilung X (a darf aber nicht in X vorkommen).

Wenn $E(U) = a$ gilt, so ist die Schätzfunktion U erwartungstreu.

7.2 Maximum Likelihood Methode

Für den Erwartungswert eignete sich das arithmetische Mittel über alle Kopien von X also aus allen X_i . Da aber es aber viele Parameter und diese auch verschieden in anderen Verteilungen sein können, nutzt man die Maximum-Likelihood-Methode um bestimmte Parameter zu bestimmen. Es gilt folgende Formel:

$$L(\vec{x}, a) = \prod_{i=1}^n f(x_i, a)$$

Das Ergebnis wird nach a abgeleitet und auf Nullstellen untersucht, da sich dort die Extremstellen befinden. Ggfs. muss noch einmal abgeleitet werden um den Höhepunkt zu bekommen. Zu beachten ist, dass das a ein beliebiger Parameter einer Verteilung sein kann. Die verschiedenen x_i Werte sind Realisierungen von a .

7.3 Hypothesentest

Bei einem Hypothesentest wird eine Aussage (Hypothese) auf ihre Wahrscheinlichkeit überprüft. Zum Beispiel Erwartungswert (Erwartungswert der Größe von Männern ist 1.72m).

Dabei unterscheidet man zwischen einseitigen und symmetrischen Tests. Erwartungswert von Größen ist symmetrisch da wenn die Hypothese 'links oder rechts' abweicht verworfen wird. Einseitige Test schauen ob ein bestimmter Wert nicht überschritten worden ist ('Anzahl' defekter Smartphones bei eine Shop, '90%' von Akkus gehen nach 2 Jahren kaputt).

Ferner muss noch die Irrtumswahrscheinlichkeit α festgelegt werden (oft 0.1, 0.05, 0.01). Eine Hypothese kann zu 70% zutreffen, aber die restlichen 30% bieten ein hohes Risiko (Smartphones werden zurückgeschickt, Stichprobe vom Hersteller ist aber in Ordnung). α ist daher meist klein. $1 - \alpha$ ist die Sicherheitswahrscheinlichkeit.

Symmetrisch Test haben zwei Grenzen (g_u, g_o). Die zwei Grenzbereiche werden links und rechts gleich (durch α) aufgeteilt.

Berechnung

1. Eine Hypothese für X aufstellen
2. Irrtumswahrscheinlichkeit α festlegen, wenn nicht schon gegeben
3. Grenzen berechnen. Für einseitige Test entweder Obergrenze oder Untergrenze, für symmetrische Tests Ober- und Untergrenze festlegen. Allgemein muss gelten: $p(X < g_u) + p(X > g_o) = 1$. Die Grenzen berechnet man aus der Umkehrung der Verteilfunktion:

- Einseitig: $g_u = F_X^{-1}(\alpha)$ oder $g_o = F_X^{-1}(1 - \alpha)$
- Symmetrisch: $g_u = F_X^{-1}\left(\frac{\alpha}{2}\right)$ und $g_o = F_X^{-1}\left(1 - \frac{\alpha}{2}\right)$

4. Stichprobe wird durchgeführt und mit den Grenzen verglichen. Ist das beobachtete Merkmal in der Stichprobe wird die Hypothese angenommen.

Bei einer Normalverteilung hilft es die Normalverteilung in die Standardnormalverteilung tzu bringe, da man dadurch die Grenzwerte oft in einer Tabelle gegeben sind.

7.4 Zentraler Grenzwertsatz

Der zentrale Grenzwertsatz sagt aus, dass wenn n Zufallsvariablen (alle n Zufallsvariable sind Kopien voneinander) mit beliebiger Verteilung miteinander addiert werden ($Y_n = X_1 + X_2 + \dots + X_n$), sich aus der Summe annähernd eine Normalverteilung bildet. Außerdem ist die Wahl des n wichtig (somit auch die Anzahl an Stichproben). Die

Faustregel lautet:

$$n = \frac{9}{p \cdot (1 - p)}$$

Für Varianz und Erwartungswert gilt noch folgendes:

- Erwartungswert: $E(Y_n) = n \cdot E(X)$
- Varianz: $Var(Y_n) = n \cdot Var(X)$

7.5 Varianztest

Um bei der Normalverteilung die Varianz überprüfen zu können, nutzt man die χ^2 -Verteilung. Diese setzt sich aus der Summe von m unabhängigen, standardnormalverteilten Zufallsvariablen, welche alle jeweils einzeln quadriert werden:

$$X := Z_1^2 + Z_2^2 + \dots + Z_m^2 \quad (1)$$

Man spricht von einer χ^2 -Verteilung mit m Freiheitsgraden. Für Varianz und Erwartungswert gilt noch folgendes:

- Erwartungswert: $E(X) = m$
- Varianz: $Var(X) = 2m$

Eine Zufallsvariable der Form

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

heißt χ^2 -verteilt mit $m = n - 1$ Freiheitsgraden. s^2 ist dabei die empirische Varianz und \bar{x} der Mittelwert der Stichprobe.

Ablauf: Skript extra Seiten zu χ^2 ist schon gut zusammengefasst.