Stochastik Cheat Sheet

Begriffe

Deskriptive Statistik analysiert Messreihen (qualitiv, quantitativ) (Beschäftigt sich mit Stichproben)

Schliessende Statistik zieht Rückschlüsse von einer Stichprobe auf die Gesamtheit auf Grund der Verteilungsfunktion $P(E)$

Zufallsexperiment Ein Zufallsexperiment/-versuch ist ein (theoretisch) beliebig oft wiederholbarer Versuch, dessen Ausgang nicht vorhersagbar ist.

Hat ein Versuch n verschiedene Ausgänge, die alle gleich wahrscheinlich sind, so tritt jeder Ausgang ω_i mit einer Wahrscheinlichkeit von $P(\omega_i) = \frac{1}{n}$ ein

Elementarereignis Realistation eines Zufallsexperiments. Also ein einzelnes Ergebnis eines Zufallsexperiments (ω).

Ereignisraum Menge aller möglichen Ergebnisse eines Zufallversuchs (Ω)

Ereignis Beliebige Teilmenge A aus Ω. Alle möglichen Ereignisse sind $\mathcal{P}(\Omega)$.

Ein Ereignis, dass k verschiedene Ausgänge umfasst hat die Wahrscheinlichkeit $\frac{k}{n}$.

Komplementarereignis Menge aller Elementarereignisse $\overline{A}$, die zum Ereignisraum aber nicht zum Ereignis A gehören: $P(\overline{A}) = 1 - P(A)$

Absolute Häufigkeit Wie häufig ein Ereignis E in einer Versuchsreihe aus n Versuchen aufgetreten ist ($H_n(E)$)

Relative Häufigkeit Wie häufig ein Ereignis E im Verhältnis zu anderen Ereignissen aufgetreten ist ($R_n(E) = H_n(E)/n$).

Es gilt: $R_n(E) \approx P(E)$ für $n \to \infty$ Wobei $0 \leq R_n(E) \leq 1$ und $R_n(\Omega) = 1$

Zufallsvariable (X) lässt sich formal als Funktion beschreiben, die den Ergebnissen eines Zufallsexperiments Werte (Realisierungen) zuordnet: $X = X(\omega)$

Wahrscheinlichkeit (Laplace)

$$P(A) = \frac{\text{Zahl der günstigen Fälle}}{\text{Zahl der möglichen Fälle}} = \sum_{\omega_i \in A \subset \mathcal{P}} P(\{\omega_i\}) = \frac{|A|}{|\Omega|}$$

gilt nur für endliche Anzahl von Elementareignissen, die alle gleich wahrscheinlich sind!

Statistische Wahrscheinlichkeit (Mises)

$$P(A) = \lim_{n\to\infty} R_n(A) = \lim_{n\to\infty} \frac{H_n(A)}{n}$$

Sätze

Gleichwertigkeit	$(A \wedge B) \vee (\neg A \wedge \neg B) \Leftrightarrow A = B$
Gleichzeitigkeit	$A \wedge B \Leftrightarrow A \cap B$
Komplement	$(A \wedge \neg B) \vee (\neg A \wedge B) \Leftrightarrow A = \overline{B} \Leftrightarrow B = \overline{A}$
A und/oder B	$A \vee B \Leftrightarrow A \cup B$
Wenn A, dann auch B	$A \Rightarrow B \Leftrightarrow A \subset B$

Sind die Ereignisse B_i paarweise verschieden, dann lässt sich A in Teilergebnisse B_i zerlegen

Axiomatische Wahrscheinlichkeit (Kolmogoroff)

1. Axiom: Für jedes Ereignis A gilt: $0 \leq P(A) \leq 1$
2. Axiom (Normiertheit): Wenn Ω das Ereignis ist, dass alle Elementarergebnisse umfasst: $P(\Omega) = 1$
3. Axiom (Additionstheorem): Schliessen sich A und B gegenseitig aus: $P(A \cup B) = P(A) + P(B)$

Bernoulli Experiment

Voraussetzungen:

- Die Wahrscheinlichkeit für jedes Ereignis ist stets gleich gross
- Jeder Versuch ist unabhängig von allen anderen Versuchen

Dann ist die Wahrscheinlichkeit für eine Serie $E_1E_2E...E_n$ das Produkt der Einzelwahrscheinlichkeiten:

$$P(E_1E_2E...E_n) = P(E_1) \cdot P(E_2) \cdot \dots \cdot P(E_n)$$

Beispiele sind etwa würfeln oder Roulette (3×6 hintereinander)

Sätze

Additionssatz

Zwei beliebige Ereignisse $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Zwei sich ausschliessende Ereignisse $P(A \cup B) = P(A) + P(B)$

Drei beliebige Ereignisse

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) + P(A \cap B \cap C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \end{aligned}$$

Hilfskonstrukt $P(A) = P(A \cap B) + P(\overline{A} \cap \overline{B})$

Multiplikationssatz

Zwei beliebige Ereignisse
$P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$

Zwei unabhängige Ereignisse $P(A \cap B) = P(A) \cdot P(B)$

Bedingte Wahrscheinlichkeit

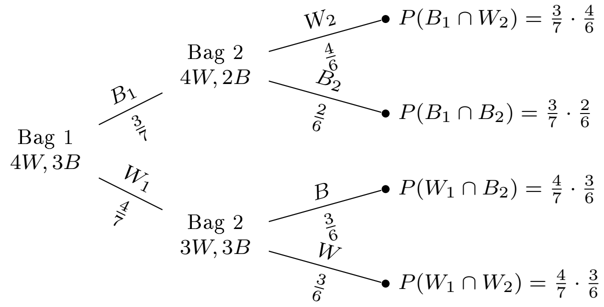
Bedingte Wahrscheinlichkeit (konditionale Wahrscheinlichkeit) $P(A|B)$ ist die Wahrscheinlichkeit des Eintretens eines Ereignisses A unter der Bedingung (auch Konditionalität), dass das Eintreten eines anderen Ereignisses B bereits bekannt ist.

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

wobei $P(A \cap B)$ die Wahrscheinlichkeit dafür ist, dass A und B gemeinsam auftreten.

Beipiel:

1. In einer Urne (Bag 1) befinden sich 4 weisse (W) und 3 schwarze (B) Kugeln.
2. Wir ziehen eine Kugel raus, dann befinden sich entweder noch 3 weisse und 3 schwarze Kugeln in der Urne (Bag 2)
3. Wir ziehen wieder eine Kugel raus ...



Totale Wahrscheinlichkeit

Sind nur bedingte Wahrscheinlichkeiten und die Wahrscheinlichkeiten des bedingenden Ereignisses bekannt, ergibt sich die totale Wahrscheinlichkeit von A aus:

$$P(A) = P(A \mid B) \cdot P(B) + P\left(A \mid \overline{B}\right) \cdot P\left(\overline{B}\right)$$

Anders ausgedrückt sind die Ereignisse A_i paarweise verschieden und der $\Omega = A_1 \cup A_2 \cup \dots \cup A_n$ dann gilt für das Ereignis E:

$$P(E) = \sum_{i=1}^{n} P(E|A_i) \cdot P(A_i)$$

Beispiel Gegeben seien zwei Anlagen (A_1 und A_2). A_1 produziert 20% Ausschuss E ($0.2 = P(E|A_1)$), A_2 produziert 10% Ausschuss ($0.1 = P(E|A_2)$). Zwei drittel der Produktion finden auf A_1 statt ($P(A_1) = 2/3$). Wie hoch ist die Wahrscheinlichkeit, dass ein Teil Ausschuss ist?
$P(E) = P(E|A_1) \cdot P(A_1) + P(E|A_2) \cdot P(A_2) = 0.2 \cdot 2/3 + 0.1 \cdot 1/3 \approx 0{,}167$

Unabhängigkeit von Ereignissen

Zwei Ereignisse sind stochastisch unabhängig wenn gilt:
$P(A|B) = P(A|\overline{B}) = P(A)$ d. h. es ist A egal, ob B eintritt.

Theorem von Bayes

Ausgangslage: $P(\text{Wirkung}|\text{Ursache})$ ist meist recht einfach. Interessant ist aber $P(\text{Ursache}|\text{Ereignis})$.

$$P(A_j|E) = \frac{P(A_j) \cdot P(E|A_j)}{\sum_{i=1}^{n} P(A_i) \cdot P(E|A_i)} \qquad P(A|E) = \frac{P(A) - P(E|A)}{P(E)}$$

Beispiel: Wie hoch ist die Wahrscheinlichkeit, dass A_j die Ursache für E ist? (Wir haben ein defektes Teil in der Hand):
$P(A_1|E) = \frac{2/3 \cdot 0.2}{2/3 \cdot 0.2 + 1/3 \cdot 0.1} = \frac{4}{5}$: 80% für Anlage A_1.

Unterscheidung der Ansätze

Frequentistischer Ansatz	Bayescher Ansatz
• A-posterio Ansatz • beliebige Wiederholbarkeit • relativen Häufigkeiten	• A-priori Ansatz • Kolmogoroff-Axiome • bedingte Wahrscheinlichkeit • Aussagen über unwiederholbare Ereignisse
Exakte Anworten auf uninteressante Frage	Unexakte Antworten auf interessante Fragen

Kombinatorik

Voraussetzung: Alle betrachteten Fälle sind gleich wahrscheinlich.

Urnenmodell In einer Gefäss befinden sich n unterscheidbare Kugeln. Nach und nach werden k davon zufällig ausgewählt

Permutationen Es gibt $n!$ Möglichkeiten, um n Elemente einer Menge anzuordnen

Binominalkoeffizient $\binom{n}{k} = \frac{n!}{k!\cdot(n-k)!}$

		Zurücklegen	
		mit	**ohne**
Reihenfolge	**mit**	$\lvert\Omega\rvert = n^k$ z. B. Zahlenschloss mit $k = 3$ Ziffern von 0-9 ($n = 10$) $\to \lvert\Omega\rvert = 10^3$	$\lvert\Omega\rvert = \frac{n!}{(n-k)!} = \mathrm{npr}(n,k)$ z. B. $k = 5$ Plätze bei einem Rennen mit $n = 12$ Läufern $\to \lvert\Omega\rvert = \frac{12!}{7!}$
	ohne	$\lvert\Omega\rvert = \binom{n+k-1}{k} = \mathrm{ncr}(n,k)$ z. B. $k = 2$ Würfel mit $n = 6$ Seiten $\to \lvert\Omega\rvert = \binom{6+2-1}{2}$	$\lvert\Omega\rvert = \binom{n}{k}$ z. B. Zahlenlotto: 6 (k) aus 49 (n) Kugeln: $\lvert\Omega\rvert = \binom{49}{6}$

Verteilungen

Wahrscheinlichkeitsfunktion

Beschreibt die Wahrscheinlichkeit dafür, dass eine Zufallsvariable X einen bestimmten Wert x annimmt.

- Diskrete Verteilungen: $p(x) = P(X = x)$
- Stetigen Verteilungen:

$$P(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x \text{ bzw. } \mu([a,b]) = \int_a^b f(x)\,\mathrm{d}x$$

Binominalverteilung

Beschreibt die Anzahl der Erfolge in einer Serie von gleichartigen und unabhängigen Versuchen, die jeweils genau zwei mögliche Ergebnisse haben („Erfolg" oder „Misserfolg") (Bernoulli-Prozesse). Für n Versuche mit k Erfolgen (E) gilt ($p = P(E)$):

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}$$

Es werden bei n Versuchen genau k Erfolge erzielt. Für $p \neq 1/2$ ist die Binominalverteilung unsymmetrisch.

Poisson-Verteilung

Spezialfall der Binominalverteilung, der die Verteilung von seltenen Ereignissen beschreibt. n wird weggelassen (d. h. schlägt ein Blitz ein? ist entscheidbar, wie oft er einschlägt kann hingegen nicht berechnet werden)

$$P_\lambda(k) = \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda}$$

Mit $\lambda = g \cdot w$ und $g =$Anzahl Ereignisse pro Einheitsintervall während des Intervalls w. Beispiel: Blitzhäufigkeit $= 10$ Einschläge pro km^2. Wie gross ist die Wahrscheinlichkeit für Blitzeinschläge pro ha (10,000 km^2)? $\lambda = 10/\mathrm{km}^2 \cdot 0.01\mathrm{km}^2 = 0.1$

Geometrische Verteilung

Beschreibt, wie oft ein Versuch wiederholt werden muss, bis E Eintritt.

$$p_k = (1-p)^{k-1}p$$

Normal- oder Gauss-Verteilung

Dichtefunktion:

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

Erwartungswert $E(x) = \mu$ entspricht dem Mittelwert

Varianz $\mathrm{Var}(x) = \sigma^2$ (Standardabweichung^2)

Dichtefunktion der Standartnormalverteilung

$$f_x(X) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

Verteilungsfunktion

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}\,\mathrm{d}t$$

Maximum

$$x_{\mathrm{max}} = \mu \qquad f(x_{\mathrm{max}}) = \frac{1}{\sigma\sqrt{2\pi}}$$

Wendepunkte

$$x_{\mathrm{wende}} = \mu \pm \sigma \qquad f(x_{\mathrm{wende}}) = \frac{1}{\sigma\sqrt{2\pi\mathrm{e}}}$$

Transformation

Eine Normalverteilung mit beliebigen μ und σ und der Verteilungsfunktion F hat folgende Beziehung zur Standardnormalverteilung ($\mathcal{N}(0,1)$):

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Für $X \sim \mathcal{N}(\mu,\sigma^2)$ gilt die Transformation $Z = \frac{X-\mu}{\sigma}$. Z ist dann standardnormalverteilt ($\mathcal{N}(\mu,\sigma^2) \to \mathcal{N}(0,1)$)

Die Wahrscheinlichkeit, dass ein Ereignis X, bei denen die Wahrscheinlichkeiten normalverteilt sind, im Intervall $[a,b]$ liegt lässt sich wie folgt berechnen:

$$P(a \le X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Die Werte für $\Phi(x)$ lassen sich einer Tabelle entnehmen

Zufallsvariablen

Diskret	**Stetig**
Nur abzählbare Werte Verteilungsfunktion: $X \le x_i$	Nur Werte in einem Interval Verteilungsfunktion
$F_X(x) = P(X \le x) = \sum_{x_i \le x} f_X(x_i)$	$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\,\mathrm{d}u$
Die Grenzwerte streben gegen 1 bzw. gegen 0 für $x \to \pm\infty$	
Wahrscheinlichkeitsfunktion (Zähldichte)	Dichtefunktion (1te Ableitung der Verteilungsfkt.)
$P(X = x_i) = f_X(x_i)$ $f_X(x_i) \ge 0$ und $\sum f_X(x_i) = 1$ $P(a \le X \le b) = \sum_{a \le x_i \le b} f_X(x_i)$	$f_X(x) = \frac{\mathrm{d}F_x(x)}{\mathrm{d}x}$
Erwartungswert (artithm. Mw.) $\mu \equiv E(x) = \sum_i f_X(x_i) \cdot x_i$	Erwartungswert $E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\,\mathrm{d}x$
Varianz $\sigma^2 = E(X - E(X))^2 = \sum_i (x_i - E(X))^2 f_X(x_i) = E(X^2) - (E(X)^2)$	Varianz $\sigma^2 = \int_{-\infty}^{\infty} (x - E(X)^2) \cdot f_X(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)\,\mathrm{d}x - (E(x))^2$

Beschreibende Statistik

Histogramm Säulendiagramm. Höhe der Säulen widerspiegelt die relative oder absolute Häufigkeit, die x-Achse zeigt die einzelnen Ereignisse

Streuungsdiagrammm Jedem Datenpaar (x_i, y_i) wird ein Punkt der x,y-Ebene zugeordnet

Lineare Regression geht von einem linearen Zusammenhang zwischen den Merkmalen X und Y aus: $Y = \alpha + \beta \cdot X + \varepsilon$. Wobei ε ein Störterm ist. X heisst Regressor, Y heisst Regressand (die zu erklärende Variable)

Revision: 1.0, Datum: 29. Mai 2013