|  |  |  |  |
| --- | --- | --- | --- |
| **DISKRETE VERTEILUNGEN**  Bernoulliverteilung  X~{0,1}, E[X]=p, Var[X]=p(1-p), P(X=1)=p, P(X=0)=1-p  Binomialverteilung (Anz der Erfolge beim n-maligen Ziehen mit zurücklegen)  E[X]=np, Var[X]=Var[∑Xi]=∑Var[Xi]=nVar[Xi]=np(1-p)  dbinom(k,n,p)=P(X=k), pbinom(k,n,p)=F(k),  qbinom(q,n,p)=, rbinom(k,n,p) k bin-vertilte Zahl  Hypergeometrische Verteilung  n mal(ohneZurückleg), M-El-te mit Erfolg, N-El-te ohne  dhyper(k,M,N,n)=P(X=k), phyper(k,M,N,n)=F(k)  Poisson-Verteilung  P(X) in Zeit ∆t ist bekannt  E[X]=Var[X]=  dpois(k,)=P(X=k), ppois(k,)=F(k)  Gleichverteilung  Alle Werte sind gleichwahrscheinlich  sample(1:N,n) n-Anz ZV-len zwischen 1 und N  **STETIGE VERTEILUNGEN**  Normalverteilung  , E[x]=µ, Var[x]=,  max(f(x))=, Wendepunkte:  Fall: Standardnormalverteilung      dnorm(x,)=f(x), pnorm(x,)=F(x),  qnorm(  Exponentialverteilung  Modellierung von Lebensdauern, Wartezeiten.  Sei im Intervall [0,t] von t Zeiteinheiten, dann beschreibt die Exponentialverteilung die Wartezeit X bis Eintreten eines Ereignisses  : Durchschnit Eintreten eines Ereignisses pro Zeiteinheit **x**: Zeitabstand (höchstens) zw. zwei Ereignissen  dexp(x,)=f(x), pexp(x,)=F(x)  Gedächtnislos:  Gleichverteilung  dunif(x,a,b)=f(x), punif(x,a,b)=F(x),  runif(n) n Z-Zahlen zw (0;1)  Chi-Quadrat-Verteilung  dchisq(x,n)=f(x), pchisq(x,n)=F(x)  t-Verteilung  ist t-verteilt mit n Freiheitsgraden  dt(y,n)=f(y), pt(y,n)=F(y) | **BESCHREIBENDE STATISTIK**  Modalwert(e)  - am häufigsten auftretende Ausprägungen(insbes bei qualitativen Merkmahlen)  Mittelwert mean(x)    Schwerpunkt der Daten (empfindlich gegenüber Ausreißern)  Median median(x)  Nicht empfindlich (robust) gegenüber Ausreißern  Spannweite  Stichprobenvarianz var(x)  Stichprobenstandardabweichung s sd(x)  p-Quantil xp  quantile(x,p)  Boxplot    Korrelation – grafische Darstellung zw multivariaten Daten x und y durch ein Streudiagramm (Punkte)  Empirische Kovarianz sxy cov(x,y)  Für hat Punktewolke steigende , fallende Tendenz  Empirischer Korrelationskoeffizient r cor(x,y)  Näherungsweise lineare Zusammenhang zw x und y , falls |r|~1  Regressionsgerade    Ungleichung von Chebyschev  Sei eine Stichprobe mit s>0. Es sei ) mit N( der Anzahl der Elemente in . Dann gilt:  Z.B: k=2: Mehr als (1-1/4)\*100% der Daten liegen in der 2s Bereich usw.  **WAHRSCHEINLICHKEITSTEORIE**  De Morgan‘schen Regel  Axiome von Kolmogorov:  Laplace-Exmeriment  Zufallsexperiment mit n gleichwahrscheinlichen Elementarereignissen  Kombinatorik  **Allgemeines Zählprinzip**  Anz der Mögl-ten, für ein k-stufiges Zufallsexperiment mit Varianten im i-ten Schritt:  Anzahl der **Permutationen** einer n-elementigen Menge  n-maliges Ziehen ohne Zurücklegen mit Beachtung der Reihenfolge  \* für n unterscheidbare Elemente: n!  \* für k Klassen mit je nicht unterscheidbaren Elementen(n=∑)  Anz der k-elementigen Teilmengen einer n-el-gen Menge k-maliges Ziehen aus einer n-elementigen Menge | Bedingte Wahrscheinlichkeit  Die Wahrscheinlichkeit für das Eintreten von E, wenn F eingetreten ist.  Multiplikationsregel (E,F )  Satz von der totalen Wahrscheinlichkeit  Ereignisse bilden eine disjunkte Zerlegung bzw eine Partition von :    Formel von Bayes  Hilfreich, wenn man kennt, nicht aber  P  Unabhängigkeit  Die Ereignisse heißen **stochastisch** **unabhängig**, wenn die Info über das Eintreten des Ereignis die Wahr-keit für das Eintreten des anderen nicht ändert. Bedeutet **nicht unbedingt kausale** **Unabhängigkeit**!  Falls E, F unabhängig, gilt auch,F bzw. unabhäng  Disjunkheit zweier Ereignissen:  Aus Disjunkheit zweier Ereignissen folgt nicht dass sie unabhängig sind da, jedes kann sein  **ZUFALLSVARIABLEN**  Verteilungsfunktion F(x)  ,  Eigenschaften der Verteilungsfunktion:  rechtsseitig stetig, d.h:  monoton wachsend    ABER  Diskrete Zufallsvariablen  Eigenschaften der Wahrscheinlichkeitsverteilung:  Es gilt:  F(x) ist eine rechtsseitig stetige Treppenfunktion mit Sprüngen bei den Realisationen xi  Stetige Zufallsvariablen P(X=xi) = 0  Eigenschaften Wahrscheinlichkeitsdichte f(x):  Es gilt:  ist stetig und  Mehrdimensionale ZV  Erwartungswert  Erwartungswert E[X]=μ  Sei g(X) eine Funktion der ZV X, dann gilt: | **Eigenschaften des Erwartungswerts**:  Varianz und Kovarianz  *Standardabweichung*  **Eigenschaften der Varianz**  Kovarianz Cov[X,Y]  Wenn X, Y stochastisch unabhängig ⇒Cov[X,Y]=0  Verschiebungssatz:  **Eigenschaften der Kovarianz**  Quantile  Das p-Quantil ist der kleinste Wert ∈R für den gilt  p-Quantil einer stetigen ZV mit streng monoton wachsendem F(x):  Chebyshev Ungleichung  Es gilt für jedes beliebige :  Schwaches Gesetz der großen Zahlen  Mit Chebyshev-Ungleichung und  folgt:  ist eine Folge von unabhängigen, identisch verteilten ZV mit EW und Varianz .  Dann gilt für ein beliebig kleines :  d.h. der konvergiert stochastisch gegen den EW . **ZGWS** Wahrscheinlichkeitsaussagen über  n, und bekannt für n unabhängige identisch verteilte ZV  Verteilung nicht bekannt.  Für **hinreichend große n** gilt dann näherungsweise:  Faustregel für Größe von n:  n > 30: Verteilung ist schief aber ohne markante примечательный Ausreißer (Exponentialverteilung)  n > 15: Verteilung annähernd примерно symmetrisch (Binomialverteilung)  n <= 15: Verteilung annähernd normalverteilt  Stichprobenverteilungen für normalverteilte Grundgesamtheiten  Ja nachdem, was gegeben ist, wählen: |