Vorlesung
Statistische Methoden der Datenanalyse
Prof. Dr. Dr. Wolfgang Rhode

Wahrscheinlichkeitsverteilungen

Überblick

- Definitionen von Wahrscheinlichkeiten
  - Frequentistisch
  - Bayesisch
- Kombination von Wahrscheinlichkeiten
- Eindimensionale Verteilungen
  - Verteilungsfunktion und Wahrscheinlichkeitsdichte
  - Momente
  - Regeln über Mittelwerte und Varianzen
  - Gängige eindimensionale Verteilungen

Frequentistische Definition

- Wahrscheinlichkeit kann abhängig davon, ob a priori Wissen über den betrachteten Vorgang zur Definition benutzt werden kann, auf zwei Weisen eingeführt werden:

  1. Falls ein Ereignis auf *n* verschiedene und gleich wahrscheinliche Arten eintreten kann und *k* davon die Eigenschaft *A* haben, so ist die Wahrscheinlichkeit für das Auftreten von *A*

$$P(A) = \frac{k}{n} = \frac{\text{günstige}}{\text{mögliche}}\text{Fälle}$$

Frequentistische Definition

- Wahrscheinlichkeit kann abhängig davon, ob a priori Wissen über den betrachteten Vorgang zur Definition benutzt werden kann, auf zwei Weisen eingeführt werden:

  2. Ohne *a priori* Wissen: Die Eigenschaften A und nicht-A eines Experimentes werden n-fach unabhängig beobachtet. Dabei trete *k* mal die Eigenschaft *A* auf. Dann ist die Wahrscheinlichkeit *P(A)* gegeben durch

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

- Problem: Identische Wiederholungen von Experimenten sind schwer zu gewährleisten

Bayesische Definition

- Die Wahrscheinlichkeit *p(A|B)* ist ein quantitatives Maß der Plausibilität der Annahme *A* unter der Bedingung der bekannten Information gegeben durch die Annahme *B*
- *A* kann dabei eine beliebige logische Annahme sein
- *p(A|B)* lässt sich mit dem Satz von Bayes berechnen:

$$p(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

- Problem: Der vage Begriff „Plausibilität“ muss genau definiert werden
- Nutzung auch zur Parameterschätzung → Kapitel Schätzen

technische universität dortmund
Experimentelle Physik Vb
Astroteilchenphysik

Rechenregeln für bedingte Wahrscheinlichkeiten

- Summenregel: $p(A|B) + p(\bar{A}|B) = 1$
  - $\bar{A}$ ist das Komplement von A
- Produktregel:

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

- A, B beschreibt dabei die Annahme A und B seien wahr

technische universität dortmund
Experimentelle Physik Vb
Astroteilchenphysik

Kombination von Wahrscheinlichkeiten

- Gegeben seien die Ereignistypen *A* und *B* mit den Wahrscheinlichkeiten *P(A)* und *P(B)*, dann ist die Wahrscheinlichkeit für *A* <u>oder</u> *B*

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

- Wenn sich A und B ausschließen:

$$P(A \wedge B) = 0 \quad \text{und} \quad P(A \vee B) = P(A) + P(B)$$

- Als Spezialfall sei: B = Ā (nicht A), dann ist:

$$P(A \vee \overline{A}) = P(A) + P(\overline{A}) = 1$$

technische universität dortmund
Experimentelle Physik Vb
Astroteilchenphysik

Kombination von Wahrscheinlichkeiten

- Gegeben seien die Ereignistypen *A* und *B* mit den Wahrscheinlichkeiten *P(A)* und *P(B)*, dann ist die Wahrscheinlichkeit für *A* <u>und</u> *B*

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

- Sind *A* und *B* unabhängig

$$P(B|A) = P(B)$$

folgt

$$P(A \wedge B) = P(A) \cdot P(B)$$

Verteilungsfunktion und Wahrscheinlichkeitsdichte

- Ziel ist die Klassifizierung von möglichen Endzuständen eines statistischen Vorganges
  - Beispiel zur Klassifizierung: Bei einem Münzwurf werden z.B. die Zuordnungen Kopf → 0 und Zahl → 1 vorgenommen.
- Allgemein:
  - Wird dem Ereignis A_i die ganze Zahl i zugewiesen, liegt eine *diskrete* Zufallsvariable vor.
  - *Kontinuierliche* Zufallsvariablen werden genutzt, wenn es nicht möglich ist, die Ereignisse ganzen Zahlen zuzuordnen.

Verteilungsfunktion und Wahrscheinlichkeitsdichte

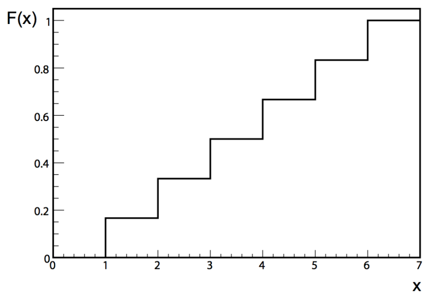
- Die Zufallsvariable *r* möge den möglichen Ausgang des Experimentes angeben. Sie wird mit der reellen Zahl *x* verglichen. Gesucht ist die Wahrscheinlichkeit dafür, dass ein Ereignis eintritt, bei dem die Zufallsvariable *r* kleiner ist als ein vorher gewähltes *x* ($r < x$). Dazu wird die Verteilungsfunktion gebildet:

$$F(x) = P(r \leq x)$$

- Die Verteilungsfunktion gibt die Summe aller Ereignisse unterhalb von *x* normiert auf die Gesamtzahl der Versuche an.

Verteilungsfunktion und Wahrscheinlichkeitsdichte

- Für einen Würfel mit sechs Seiten, ergibt sich für die Verteilungsfunktion F(x) eine sechsstufige Treppenfunktion, die monoton von 0 auf 1 ansteigt.



Verteilungsfunktion und Wahrscheinlichkeitsdichte

- Im Grenzfall einer kontinuierlichen Verteilung ist:

$$\lim_{x \to \infty} F(x) = \lim_{x \to \infty} P(r \leq x) = 1$$

- Da die Summe aus P(A) + P(Ā) = 1 ist, gilt:

$$P(r > x) = 1 - F(x) = 1 - P(r \leq x)$$

- Somit ist

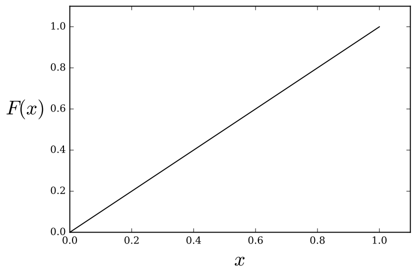
$$\lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} P(r \leq x) = 1 - \lim_{x \to -\infty} P(r > x) = 0$$

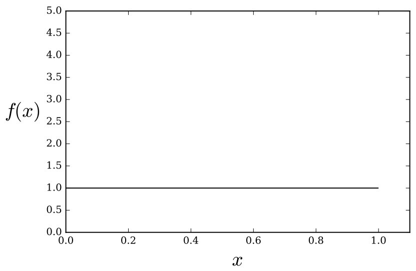
Verteilungsfunktion und Wahrscheinlichkeitsdichte

- Wenn die Verteilungsfunktion stetig differenzierbar ist, gilt:

$$\frac{dF(x)}{dx} = F'(x) = f(x)$$

- *f(x)* heißt dann Wahrscheinlichkeitsdichte von *r* und gibt ein Maß für die Wahrscheinlichkeit in dem Intervall $x \leq r \leq x + dx$ an





Verteilungsfunktion und Wahrscheinlichkeitsdichte

- Die Wahrscheinlichkeit, dass *r* kleiner ist als ein vorgewählter Wert *a*, ist gegeben durch:

$$P(r < a) = \int_{-\infty}^{a} f(x)\,\mathrm{d}x = F(a),$$

- die Wahrscheinlichkeit, dass *r* in einem Intervall zwischen *a* und *b* liegt, ist:

$$P(a \leq r \leq b) = \int_{a}^{b} f(x)\,\mathrm{d}x = F(b) - F(a).$$

- Insbesondere gilt bei Integration über den gesamten Bereich in *x*:

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1.$$

Allgemeine Eigenschaften einer Zufallsverteilung: Momente

- Der Mittelwert oder Erwartungswert *E(x)* bei einer diskreten Verteilung:

$$\bar{x} = E(x) = \sum_{i=1}^{n} (x_i \cdot P(x = x_i))$$

- Der Erwartungswert einer Funktion von diskreten *r* ist:

$$E[H(x)] = \sum_{i=1}^{n} (H(x_i) \cdot P(x = x_i))$$

Momente

- Erwartungswert für kontinuierlich verteilte x

$$E(x) = \bar{x} = \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x$$

- und für eine Funktion davon

$$E[H(x)] = \int_{-\infty}^{\infty} H(x) \cdot f(x)\,\mathrm{d}x$$

Momente

- Wichtige Charakteristika einer Verteilung sind ihre Breite und Symmetrie. Dazu betrachten wir als Spezialfall die Funktion:

$$H(x) = (x - c)^l$$

- Der Erwartungswert ergibt sich zu:

$$a_l = E[(x - c)^l]$$

Momente

- Berechnung von Momenten μ_l um den Mittelwert, heißen zentrale Momente:

$$\mu_l = E[(x - \bar{x})^l]$$

- Die Momente $\mu_0 = 1$ und $\mu_1 = 0$ sind trivial zu bestimmen.

Momente

- Für das zweite Moment gilt:

$$E[(r - \bar{x})^2] = \hat{\sigma}^2(x) = \mathrm{Var}(x) = \int\limits_{-\infty}^{\infty} (x - \bar{x})^2 f(x)\, \mathrm{d}x$$

- Die so definierte (empirische) Varianz ist ein Maß für die Breite der Verteilung.
- Die Wurzel aus der Varianz heißt Streuung oder Standardabweichung

$$\sigma = \sqrt{\sigma^2(x)}$$

Momente

- Das dritte Momente um den Mittelwert normiert auf die Standardabweichung heißt Schiefe (*skewness*). Es beschreibt die Symmetrie der Verteilung.

$$\gamma = \frac{\mu_3}{\sigma^3}$$

- $\gamma < 0$ linksschief
- $\gamma > 0$ rechtsschief

Momente

- Der Quotient des vierten zentralen Moments und dem Quadrat der Varianz wird als Wölbung (*kurtosis*) bezeichnet:

$$C = \mu_4/\sigma^4$$

- *C* ist groß, wenn die Verteilung über größere Ausläufer verfügt als die Gauß-Verteilung.
- Die Gauß-Verteilung selbst liefert *C=3*

Regeln über Mittelwerte und Varianzen

- Multiplikation jeder Zahl einer Verteilung mit (derselben) Konstanten:

$$H(x) = cx, \;\; c = \text{const}$$

- Es folgt, dass

$$E(c \cdot r) = c \cdot E(r), \;\text{und}\; \sigma^2(c \cdot r) = c^2 \cdot \sigma^2(r)$$

- Daher ist

$$\sigma^2(r) = E[(r-\bar{x})^2] = E[r^2 - 2r\bar{x} + \bar{x}^2] = E(r^2) - \bar{x}^2$$

Regeln über Mittelwerte und Varianzen

- Die <u>standardisierte Variable</u>

$$u = \frac{r-\bar{x}}{\sigma(r)}$$

- hat den Erwartungswert

$$E(u) = \frac{1}{\sigma(r)} E(r-\bar{x}) = \frac{1}{\sigma(x)}(\bar{x}-\bar{x}) = 0$$

- und die Varianz

$$\sigma^2(u) = \frac{1}{\sigma^2(x)} E[(r-\bar{x})^2] = \frac{\sigma^2(x)}{\sigma^2(x)} = 1$$

Regeln über Mittelwerte und Varianzen

- Der <u>wahrscheinlichste Wert</u>

$$P(x = x_m) = \text{maximal}$$

- Besitzt die Verteilung ein Maximum, heißt sie unimodal, sonst heißt sie multimodal

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = 0 \qquad\qquad \frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) < 0$$

Regeln über Mittelwerte und Varianzen

- Der Median ist derjenige Wert einer Verteilung, für den die Verteilungsfunktion $F = 0.5$ ist

$$F(x_{0.5}) = P(r \leq x_{0.5}) = 0.5$$

- Ist $f(x)$ stetig, gilt

$$\int_{-\infty}^{x_{0.5}} f(x)\,\mathrm{d}x = 0.5$$

- Ist die Verteilung unimodal, stetig und symmetrisch, dann ist Erwartungswert gleich dem wahrscheinlichsten Wert oder Median

Regeln über Mittelwerte und Varianzen

- Das Quartil einer Verteilung ist analog zu $x_{0.5}$ definiert als:

$$F(x_{1/4}) = 0.25, \quad F(x_{3/4}) = 0.75$$

unteres Quartil oberes Quartil

- Entsprechend sind Dezile (q=10%) und Quantile (q=beliebige Prozentsatze) definiert als

$$F(X_q) = \int_{-\infty}^{X_q} f(x)\,\mathrm{d}x = q$$

Regeln über Mittelwerte und Varianzen

- Der quadratische Mittelwert (*root mean square* = RMS) ist definiert als

$$x_{\mathrm{rms}} = \sqrt{E(x^2)} = \sqrt{\sigma^2(x) + \bar{x}^2}$$

- Ist der Erwartungswert gleich null gilt:

$$x_{\mathrm{rms}} = \sigma(x)$$

Momente vs. Lagemaße & Streuungsmaße

- Gerade wurden Momente betrachtet. Momente beschreiben eine Wahrscheinlichkeitsverteilung **nicht** die daraus gezogene Stichprobe.

- Lage- & Streuungsmaße werden zur Charakterisierung einer gezogenen Verteilung genutzt.

- Vorsicht: Als Mittelwert wird oft umgangssprachlich sowohl das 1. Moment (Erwartungswert) einer Wahrscheinlichkeitsverteilung, als auch der arithmetische Mittelwert einer Stichprobe bezeichnet. Letzterer ist jedoch ein Lagemaß.

Lagemaße

- Mittelwerte:

$$\underbrace{\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i}_{\text{arithmetisch}} \qquad \underbrace{\bar{x}_G = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}}_{\text{geometrisch}} \qquad \underbrace{\bar{x}_H = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}\right)^{-1}}_{\text{harmonisch}}$$

- Median:
  auf geordneten, metrischen Zufallsvariablen

$$\tilde{x}_{0.5} = \begin{cases} x_{\frac{n+1}{2}} & \text{falls } n \text{ ungerade,} \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right) & \text{falls } n \text{ gerade.} \end{cases}$$

Für p-Quantil 0.5 durch p ersetzen

Streuungsmaße

- Empirische Varianz:

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

- Empirische Stichprobenvarianz:

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

- Entropie (mittlere Information in x_i):

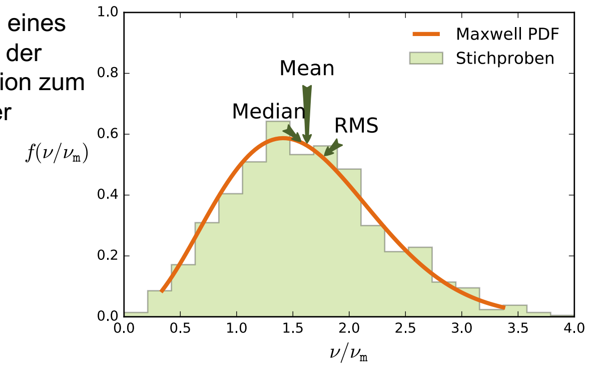
$$I = \sum_{i=1}^{n} f(x_i)\ \underbrace{\ln\frac{1}{f(x_i)}}_{\text{Information}}$$

Keine Streuung: $f(x_i) = 1 \rightarrow I = 0$
Maximale Streuung (alle rel. Häufigkeiten sind gleich):

$$f(x_i) = \frac{1}{n} \rightarrow I = \ln n$$

Beispiel aus Blobel - Lohrmann

- Maxwellsche Geschwindigkeitsverteilung eines idealen Gases als Funktion der Geschwindigkeit ν in Relation zum wahrscheinlichsten Wert der Geschwindigkeit ν_m



(diskrete) Gleichverteilung

- Einzelwahrscheinlichkeiten

$$P(X = a_m) = \frac{1}{n}$$

- Parameterbereich

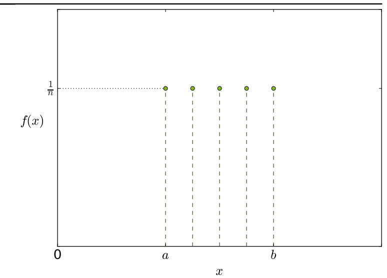
$$a, b \in \mathbb{R}, a < b$$

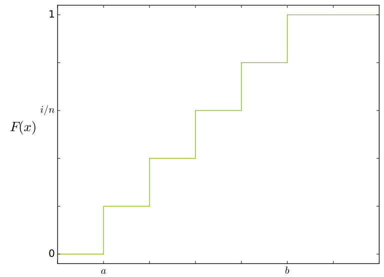
- Momente

$$m_k = \sum_{m=1}^{n} a_m^k \cdot P(X = a_m) = \sum_{m=1}^{n} a_m^k \cdot \frac{1}{n}$$

- Erwartungswert

$$\frac{1}{n}\sum_{m=1}^{n} a_m$$





(diskrete) Gleichverteilung

- Anwendungen:
  - Zufällige Versuche mit *n* gleichwahrscheinlichen Ausgängen:
  - Würfeln
  - Münzwurf
  - Thermodynamik: Räumliche Verteilung von Teilchen

(stetige) Gleichverteilung

- Dichtefunktion

  $$\frac{1}{b-a}\,,\; a \leqq x \leqq b$$

- Parameterbereich

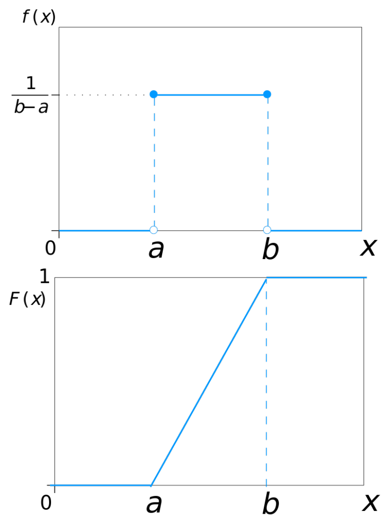
  $$a, b \in \mathbb{R}, a < b$$

- Momente

  $$\mu_{2k} = \frac{1}{2k+1}\left(\frac{b-a}{2}\right)^{2k}$$
  $$\mu_{2k-1} = 0$$

- Erwartungswert

  $$\frac{a+b}{2}$$



(stetige) Gleichverteilung

- Anwendungen:
  - Geometrische Wahrscheinlichkeit
  - Erzeugung von Zufallszahlen

Dreiecksverteilung

- Dichtefunktion

  $$\frac{2}{b-a}\left(1-\frac{2}{b-a}\left|x-\frac{a+b}{2}\right|\right)$$

- Parameterbereich

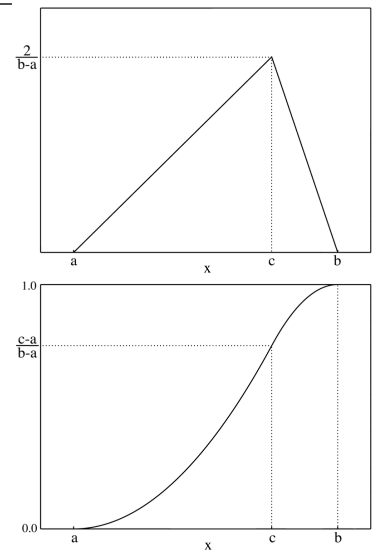
  $$a, b \in \mathbb{R}, a < b$$

- Momente

  $$\mu_{2k} = \frac{(a-b)^{(2k)}}{2^{(2k-1)}(2k+1)(2k+2)}\,,\; \mu_{2k-1} = 0$$

- Erwartungswert

  $$\frac{a+b}{2}$$

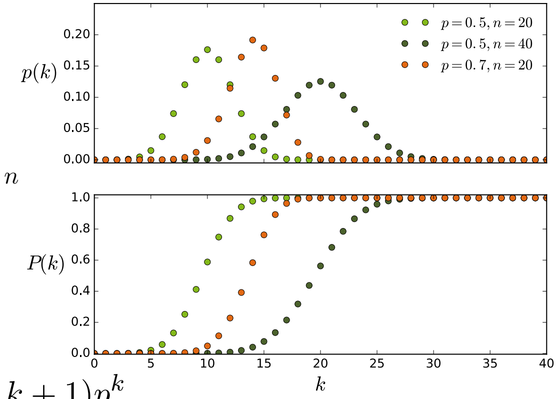


Dreiecksverteilung

- Anwendungen:
  - Verteilung der Summe zweier unabh. identisch (stetig) gleichmäßig verteilter Zufallszahlen
  - Zeitliche Planung der Durchführung eines Experiments:

a = optimistischster Wert, b = pessimistischster Wert, c = wahrscheinlichster Wert

Binominalverteilung



- Dichtefunktion

$$\binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \ldots, n$$

- Parameterbereich

$$0 < p < 1, n \in \mathbb{N}$$

- Momente

$$m_{(k)} = n(n-1)\ldots(n-k+1)p^k$$

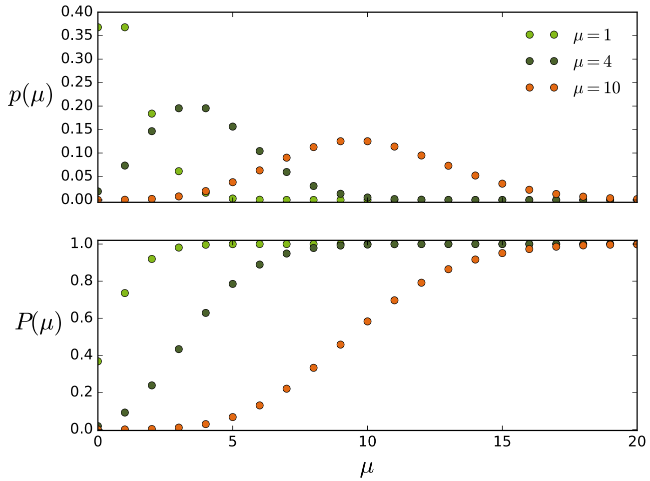
- Erwartungswert

$$np$$

Binominalverteilung

- Anwendungen:
  - Bernoulli-Schema
  - Statistische Qualitätskontrolle
  - Fehlerrechnung
  - Berechnung magnetischer Momente

Poissonverteilung



- Dichtefunktion

$$e^{-\mu}\frac{\mu^k}{k!}, k = 0, 1, \ldots$$

- Parameterbereich

$$\mu > 0$$

- Momente

$$m_{(k)} = \mu^k$$

- Erwartungswert

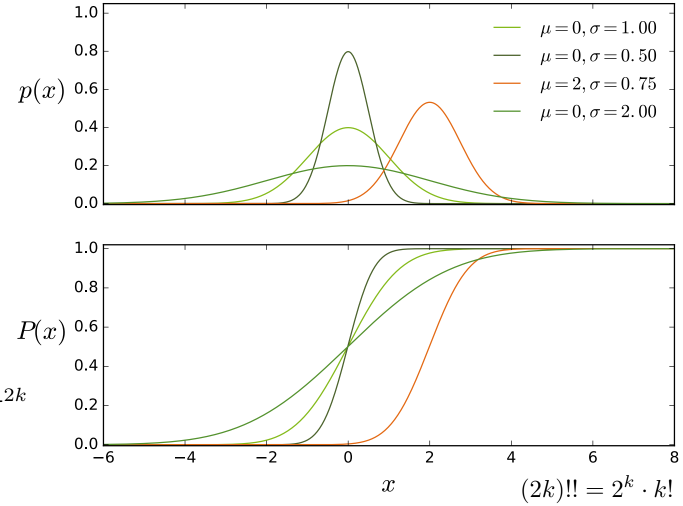
$$\mu$$

Poissonverteilung

- Anwendungen:
  - Zählraten:
  - Ereignisse in Teilchendetektoren
  - Radioaktive Zerfälle

Normalverteilung

- Dichtefunktion
  $$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
- Parameterbereich
  $$\mu \text{ reell}, \sigma > 0$$
- Momente
  $$\mu_{2k-1} = 0$$
  $$\mu_{2k} = (2k-1)!!\sigma^{2k}$$
- Erwartungswert
  $$\mu$$



$$(2k)!! = 2^k \cdot k!$$

Normalverteilung

- Das unbestimmte Integral über die Normalverteilung kann nicht analytisch berechnet werden
  → Nachschauen in Tabellen, Berechnung mit dem Computer

$$\int_{\mu-n\sigma}^{\mu+n\sigma} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x$$

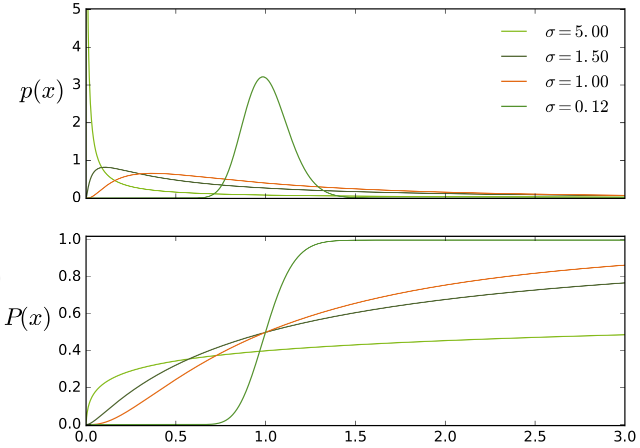
- 68,27% der Fläche liegt innerhalb des 1σ-Bereichs um den Mittelwert,
- 95,45% liegen innerhalb des 2σ-Bereichs,
- 99,73% liegen innerhalb des 3σ-Bereichs

Normalverteilung

- Anwendungen:
  - Grundlegende Verteilung in der Wahrscheinlichkeitsrechnung und zur statistischen Auswertung von Versuchs-, Beobachtungs- und Messergebnissen:
  - Fehlerrechnung
  - Intensitätsverteilung ausgedehnter kosmischer Quellen
  - ...

Logarithmische Normalverteilung

- Dichtefunktion
  $$\frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}}, \; x > 0$$
- Parameterbereich
  $$\mu \text{ reell}, \sigma > 0$$
- Momente
  $$m_k = e^{\left(k_m u + \frac{k^2\sigma^2}{2}\right)}$$
- Erwartungswert
  $$e^{\left(\mu + \frac{\sigma^2}{2}\right)}$$

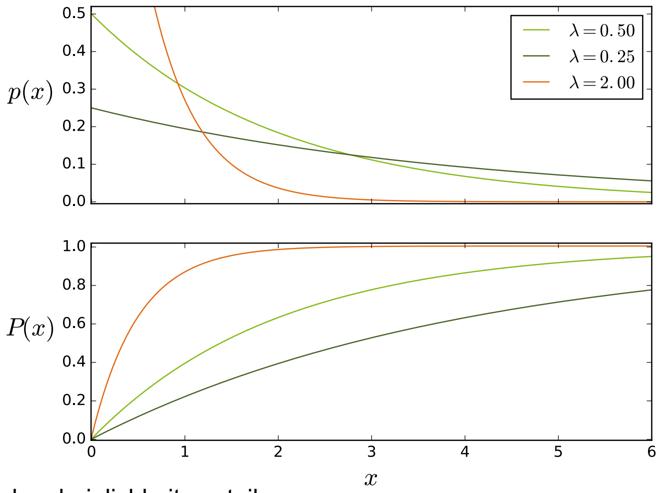


Logarithmische Normalverteilung

- Anwendungen:
  - Lebensdauer- sowie Festigkeitsprobleme
  - Konzentrationsuntersuchungen
  - Bestimmung natürlicher Größen:
  - Partikelgröße in Wolken
  - Verteilung der Materie im Universum (Elementarteilchen, Galaxien)

Exponentialverteilung

- Dichtefunktion
  $$\lambda e^{-\lambda x}$$
- Parameterbereich
  $$\lambda > 0$$
- Momente
  $$m_k = \frac{\lambda}{\lambda - k}$$
- Erwartungswert
  $$\frac{1}{\lambda}$$

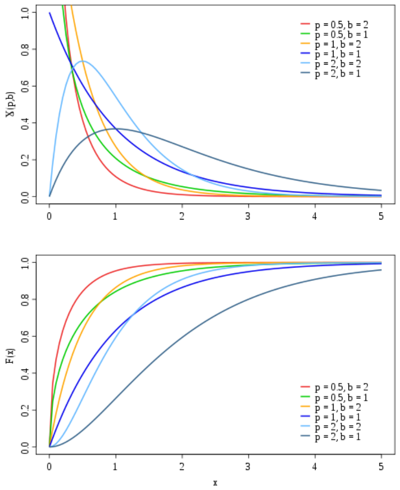


Exponentialverteilung

- Anwendungen:
  - Zeit zwischen zwei Anrufen
  - Lebensdauern bei radioaktiven Zerfällen

Gamma-Verteilung

- Einzelwahrscheinlichkeiten
  $$\frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$
- Parameterbereich
  $$\lambda > 0$$
- Momente
  $$m_k = \left(\frac{\lambda}{\lambda - k}\right)^\alpha$$
- Erwartungswert
  $$\frac{\alpha}{\lambda}$$



technische universität dortmund
Experimentelle Physik Vb
Astroteilchenphysik

Gamma-Verteilung

- Anwendungen:
  - Verallgemeinerung der Exponentialverteilung
  - Wahscheinlichkeitstheorie
  - Versicherungsmathematik

Prof. Dr. Dr. W. Rhode
Wahrscheinlichkeitsverteilungen
Statistische Methoden der Datenanalyse

technische universität dortmund
Experimentelle Physik Vb
Astroteilchenphysik

t-Test - Einstichproben Test (Vorgriff)

Fragestellung:
Teste, ob die gegebenen Daten zu einer Normalverteilung mit vorgegeben Erwartungswert $\mu=\mu_0$ passen

Test-Statistik t, mit arithmetischem Mittel x ...

$$t = \sqrt{n}\,\frac{x - \mu_0}{\sqrt{S^2}}$$

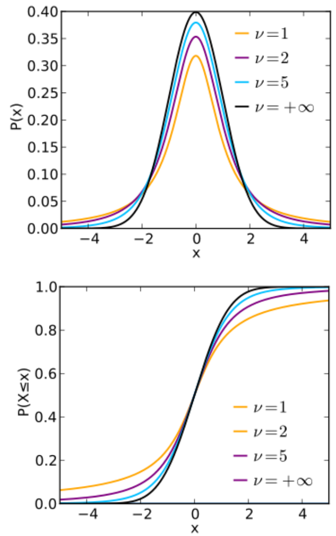
Die Test-Statistik ist unter der Nullhypothese t-verteilt mit n-1 Freiheitsgraden.

Prof. Dr. Dr. W. Rhode
Wahrscheinlichkeitsverteilungen
Statistische Methoden der Datenanalyse

technische universität dortmund
Experimentelle Physik Vb
Astroteilchenphysik

t-Verteilung

- Dichtefunktion
  $$\frac{\Gamma(\frac{\nu+1}{2})(1+\frac{x^2}{\nu})^{-\frac{n\nu+1}{2}}}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}$$
- Parameterbereich
  $$\nu \in \mathbb{N}$$
- Momente
  $$m_{2k-1} = 0, 2k \leq \nu$$
  $$m_{2k} = \nu^k \frac{(2k-1)!!(\nu-2k-2)!!)}{(\nu-2)!!},$$
  $$2k \leq \nu - 1$$
- Erwartungswert
  $$0, \nu \geq 2$$



Prof. Dr. Dr. W. Rhode
Wahrscheinlichkeitsverteilungen
Statistische Methoden der Datenanalyse

t-Verteilung

- Anwendungen:
  - Prüfen von Erwartungswerten
  - Regressions- und Korrelationsanalysen
  - T-Test

F-Test (Vorgriff)

Fragestellung:

Test, ob die Varianzen zweier Stichproben X, Y mit Umfang n_X, n_Y aus unterschiedlichen, normalverteilten Grundgesamtheiten gleich sind

H_0: $\sigma_x^2 = \sigma_y^2$

H_1: $\sigma_x^2 < \sigma_y^2$

Test-Statistik ist der Quotient der geschätzten Varianzen

$$F = \frac{S_Y^2}{S_X^2} = \frac{\frac{1}{n_Y-1}\sum_{i=1}^{n_Y}(y_i - y)^2}{\frac{1}{n_X-1}\sum_{i=1}^{n_X}(x_i - x)^2}$$

Unter H_0 ist die Test-Statistik F-verteilt mit n_Y-1 Freiheitsgraden im Zähler und n_X-1 Freiheitsgraden im Nenner.

F-Verteilung

- Dichtefunktion

$$\frac{\frac{m}{n}^{m/2} x^{m/2-1}(1 - \frac{mx}{n})^{-\frac{m+n}{2}}}{B(m/2, n/2)}$$

- Parameterbereich

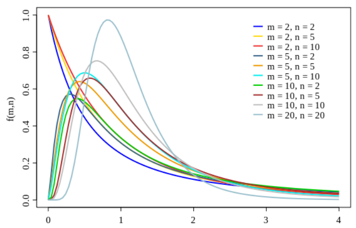
$$m, n \in \mathbb{N}$$

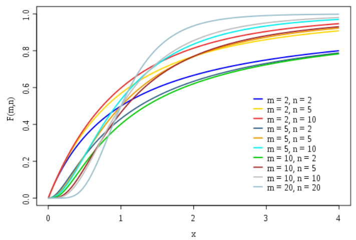
- Momente

$$m_k = \frac{\Gamma(\frac{m}{2}+k)\Gamma(\frac{n}{2}-k)n^k}{\Gamma(m/2)\Gamma(n/2)m^k}$$

- Erwartungswert

$$n/(n-2), n \geq 3$$





F-Verteilung

- Anwendungen:
  - Vergleich von Streuungen
  - Varianz- und Kovarianzanalyse
  - F-Test

χ^2-Verteilung

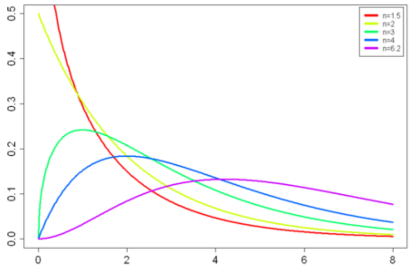
- **Fragestellung:**
  Teste, ob n gemessene Daten y_i einem angenommenen Modell $f(x_i)$ folgen
- Dichtefunktion

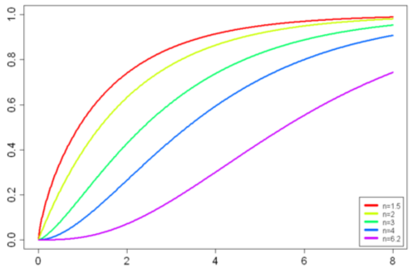
$$\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{(\frac{n}{2}-1)}e^{-\frac{x}{2}}$$

- Parameterbereich: $n \in \mathbb{N}$
- Momente

$$m_k = 2^k \frac{\Gamma\left(k+\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}$$

- Erwartungswert: n





χ^2-Verteilung

- Anwendungen:
  - Prüfen von Streuungen
  - χ^2-Test

Beta-Verteilung

- Dichtefunktion

$$\frac{1}{B(p,q)}x^{p-1}(1-x)^{q-1}\,,\ 0<x<1$$

- Parameterbereich

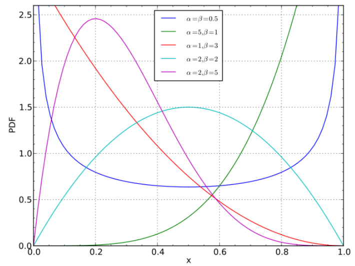
$$p>0\,,\ q>0$$

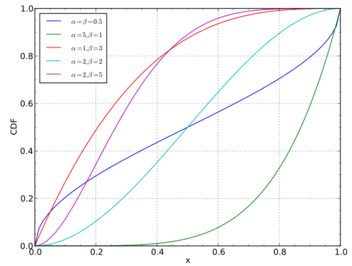
- Momente

$$m_k = \frac{B(k+p,q)}{B(p,q)}$$

- Erwartungswert

$$\frac{p}{p+q}$$





Beta-Verteilung

- Anwendungen:
  - Korrelationsanalyse