Formelsammlung: Stochastik

1 Elementare Wahrscheinlichkeitsrechnung

1.1 Zufallsexperiment

1.1.1 Rechnen mit Ereignissen

A oder B tritt ein	$A \cup B$
A und B treten ein	$A \cap B$
A zieht B nach sich	$A \subseteq B$
A und B schliessen sich aus	$A \cap B = \emptyset$
A tritt nicht ein	$\overline{A}$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Distributivgesetz

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$
$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

Gesetz von Morgan

$$\overline{\overline{A}} = A$$

1.1.2 Elementare Rechenregeln für Wahrscheinlichkeiten

1) $0 \leq p(A) \leq 1$ für alle A
2) $p(\Omega) = 1$
3) sind A und B disjunkte Ereignisse, dann ist
$$p(A \cup B) = p(A) + p(B)$$
4) $p(\overline{A}) = 1 - p(A)$
5) aus $A \subseteq B$ folgt $p(A) \leq p(B)$
6) sind A_1, A_2,..., A_n paarweise disjunkten Ereignisse, dann gilt
$$p(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{k=1}^{n} p(A_k)$$
7) für zwei beliebige(sich nicht ausschliessende) Ereignisse A und B gilt
$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$
8) für drei beliebige(sich nicht ausschliessende) Ereignisse A, B und C gilt
$$p(A \cup B \cup C) = p(A) + p(B) + p(C) - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C)$$

1.1.3 Diskrete Gleichverteilung

$$p(E_i) = \frac{1}{n}$$ E_i: Elementarereignis

$$p(A) = \frac{Anzahl.günstige.Fälle}{Anzahl.mögliche.Fälle}$$ Laplace

1.1.4 Hauptsatz der Kombinatorik

Satz: Wir betrachten n aufeinanderfolgende Operationen, deren Reihenfolge vorgegeben ist. Wir nehmen an, die k-te Operation lasse sich auf genau m_k Arten durchführen. Die Anzahl der Möglichkeiten, alle n Operationen durchzuführen beträgt dann

$$m = m_1 * m_2 * m_3 * \ldots * m_n$$

1.1.5 Permutation ohne Wiederholung

Satz: Sei M eine Menge mit n Elementen. Die Anzahl der Permutationen von M beträgt

$$P(n) = n!$$

1.1.6 Permutation mit Wiederholung

Satz: Gegeben sei eine Menge M mit n Elementen und eine natürliche Zahl r. Aus den Elementen von M sollen r-gliedrige Folgen gebildet werden, wobei sich jedes Element beliebig oft wiederholen darf. Die Anzahl solcher Folgen beträgt

$$P(n,r) = n^r$$

1.1.7 Kombinatorik mit Wiederholung

Satz: Gegeben ist eine Menge M mit n Elementen sowie eine natürliche Zahl k, wobei $0 \le k \le n$. Die Anzahl Möglichkeiten, aus M eine Teilmenge mit genau k Elementen auszuwählen, ist gegeben durch den Binominalkoeffizient

$$C(n,k) = \binom{n}{k}$$

1.2 Bedingte Wahrscheinlichkeit

1.2.1 Allgemeine Definition

Definition: Sei Ω die Ergebnismenge eines Zufallsexperiments. Wir nehmen an, dass jedem Ereignis A eine Wahrscheinlichkeit p(A) zugeordnet ist. Sei B ein Ereignis mit $p(B) \neq 0$. Dann heisst

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

die bedingte Wahrscheinlichkeit von A unter der Bedingung B.

1.2.2 Satz der totalen Wahrscheinlichkeit

Satz: Sind A und B Ereignisse von positiver Wahrscheinlichkeit, dann gilt

$$p(A \cap B) = p(A) * p(B \mid A) = p(B) * p(A \mid B)$$

Sind A und B Ereignisse und ist $p(A) = 0$ oder $p(B) = 0$, dann gilt

$$p(A \cap B) = 0$$

Satz: Seien A_1, A_2,..., A_n Ereignisse von positiver Wahrscheinlichkeit. Wir nehmen an, dass sie paarweise disjunkt sind, das heisst

$$A_i \cap A_k = \emptyset$$

und das

$$A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$$

Für jedes Ereignis B gilt dann

$$p(B) = p(A_1) * p(B \mid A_1) + p(A_2) * p(B \mid A_2) + \ldots + p(A_n) * p(B \mid A_n)$$

1.2.3 Unabhängige Ereignisse

Satz: Zwei Ereignisse A und B heissen stochastisch unabhängig, wenn

$$p(B) = p(B \mid A)$$

oder, was logisch gleichwertig ist, wenn

$$p(A \cap B) = p(A) * p(B)$$

1.2.4 Satz von Bayes

Satz: Seien A_1, A_2, ..., A_n Ereignis von positiver Wahrscheinlichkeit. Wir nehmen an, dass sie sich paarweise ausschliessen (d.h. $A_i \cap A_k = \emptyset$ für $i \neq k$) und dass $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$. Dann gilt

$$p(A_k \mid B) = \frac{p(A_k) * p(B \mid A_k)}{p(A_1) * p(B \mid A_1) + p(A_2) * p(B \mid A_2) + \ldots + p(A_n) * p(B \mid A_n)}$$

1.2.5 Berechnung von Zuverlässigkeit

Serielles System

$$R = p_1 * p_2 * \ldots * p_n$$

Paralleles System

$$R = 1 - (1 - p_1) * (1 - p_2) * \ldots * (1 - p_n)$$

1.3 Zufallsvariablen

1.3.1 Diskrete Zufallsvariablen

Definition: Eine diskrete Zufallsvariable X wird durch die folgende Angabe festgelegt:

1. eine endliche Liste L der Zahlwerte, welche X annehmen kann. In der Liste L darf kein Wert zweimal vorkommen. Gewöhnlich werden die Zahlen aufsteigend(oder absteigend) geordnet. Also

   $L = \{x_1, x_2, \ldots, x_n\}$, wobei $x_1 < x_2 < \ldots < x_n$

2. eine Liste $P = \{p_1, p_2, \ldots, p_n\}$ der zugehörigen Wahrscheinlichkeiten p_k ist die relative Häufigkeit, mit welcher der Wert x_k angenommen wird.

1.3.2 Verteilfunktion

Definition: Sei X eine diskrete Zufallsvariable mit Wertemenge $L = \{x_1, x_2, \ldots, x_n\}$ und Wahrscheinlichkeiten $P = \{p_1, p_2, \ldots, p_n\}$. Für reelle Zahlen u setzt man $F(u) := P(X \le u)$, und man nennt die Zuordnung $u \mapsto F(u)$ die Verteilfunktion von X.

1.3.3 Erwartungswert

Definition: Gegeben sei eine diskrete Zufallsvariable X mit Wertemenge $L = \{x_1, x_2, \ldots, x_n\}$ und Wahrscheinlichkeiten $P = \{p_1, p_2, \ldots, p_n\}$. Der Erwartungswert von X ist

$$E(X) = \bar{x} = \sum_{k=1}^{n} p_k * x_k$$

1.3.4 Varianz und Standardabweichung

Definition: Gegeben sei eine Zufallsvariable X mit Wertemenge $L = \{x_1, x_2, \ldots, x_n\}$, mit Wahrscheinlichkeiten $P = \{p_1, p_2, \ldots, p_n\}$ und mit Erwartungswert $\bar{x} = E(X)$.

$$Var(X) = \sigma^2 := \sum_{k=1}^{n} p_k * (x_k - \bar{x})^2$$

heisst die Varianz von X. Die Wurzel daraus,

$$\sigma = \sqrt{Var(X)},$$

σ heisst die Standardabweichung der Zufallsvariblen X.

1.3.5 Die Binominalverteilung

Definition: Eine Zufallsvariable X heisst binominal verteilt (mit Parametern n und p), wenn gilt

- die Wertemenge ist {0, 1, 2, 3, ..., n}
- die Wahrscheinlichkeiten sind gegeben durch

$$P(X = k) = p_k = \binom{n}{k} * p_k * (1-p)^{n-k} \qquad \text{(für k = 0, 1, 2, ..., n)}$$

Satz: Die Zufallsvariable X sei binominal verteilt mit Parametern n und p. Dann gilt

1. $E(X) = n * p$
2. $Var(X) = n * p * (1-p)$

1.3.6 Verteilung nach Poisson

Definition: Eine Zufallsvariable X heisst poissonverteilt mit Parameter λ, wenn gilt

- die Wertemenge ist {0, 1, 2, ..., n, ...}
- die Wahrscheinlichkeit sind gegeben durch

$$P(X = k) = p_k = \frac{\lambda^k}{k!} e^{-\lambda} \qquad \text{(für k = 0, 1, 2, ...)}$$

Satz: Die Zufallsvariable X sei nach Poisson verteilt, mit Parameter λ. Dann gilt

3. $E(X) = \lambda$
4. $Var(X) = \lambda$

1.4 Stetige Zufallsvariablen

1.4.1 Allgemeine Definition

Definition: Eine stetige Zufallsvariable X ist gegeben durch eine so genannte Dichtefunktion

f: IR → IR

welche die folgende Bedingungen erfüllt:

- $f(x) \geq 0$ für alle x
- $\int_{-\infty}^{\infty} f(x)dx = 1$

Die Wahrscheinlichkeit, dass der Wert von X in ein Intervall $a \leq x \leq b$ fällt, beträgt

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$$

Definition: Gegeben sei eine stetige Zufallsvariable X mit Dichtefunktion f: IR → IR

1. Der Erwartungswert von X ist
$$E(X) = \bar{x} = \int_{-\infty}^{\infty} x * f(x)dx$$
2. Die Varianz von X ist gegeben durch
$$Var(X) = \int_{-\infty}^{\infty} (x - \bar{x})^2 * f(x)dx$$
3. Die Standardabweichung von X ist die Quadratwurzel der Varianz
$$\sigma(X) = \sqrt{Var(X)}$$

1.4.2 Die Normalverteilung

Definition: Eine stetige Zufallsvariable X mit der Dichtefunktion

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

heisst normalverteilt mit den Parametern μ und σ^2. Man schreibt dafür auch

$$X \sim N(\mu, \sigma^2)$$

- Der Erwartungswert von X ist $E(X) = \mu$
- Die Varianz von X ist $Var(X) = \sigma^2$

1.4.3 Die Exponentialverteilung

Definition: Eine stetige Zufallsvariable X heisst exponential verteilt, wenn ihre Dichtefunktion die Form

$$y = f(x) = \lambda * e^{-\lambda x} \text{ falls } x \geq 0$$

hat. Man rechnet leicht nach, dass

$$E(X) = \frac{1}{\lambda}$$

$$\sigma(X) = \frac{1}{\lambda}$$

1.4.4 Berechnung der Wahrscheinlichkeiten

Um die Wahrscheinlichkeit

$$P(a \leq X \leq b) = \frac{1}{\sigma\sqrt{2\pi}} * \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

zu bestimmen, geht man über zur standardisierten Zufallsvariable $U := \frac{X-\mu}{\sigma}$, insbesondere berechnet man die standardisierten Grenzen

$$u_1 = \frac{a-\mu}{\sigma}, \quad u_2 = \frac{b-\mu}{\sigma}$$

Dann ist

$$\frac{1}{\sigma\sqrt{2\pi}} * \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} * \int_{u_1}^{u_2} e^{-\frac{u^2}{2}} du$$

Das letzte Integral kann mit eine Tabelle näherungweise ermittelt werden:

$$\frac{1}{\sqrt{2\pi}} * \int_{u_1}^{u_2} e^{-\frac{u^2}{2}} du = \Phi(u_2) - \Phi(u_1)$$

1.4.5 Approximation der Binominalverteilung durch die Normalverteilung

Definition: Eine Zufallsvariable X ist näherungsweise normal verteilt, sofern

1) sie sich darstellen lässt als Summe von vielen Zufallsvariablen(von viele „zufälligen Einflüssen“)
2) diese einzelnen Zufallsvariablen voneinander unabhängig sind
3) der Einfluss dieser einzelnen Summanden (also dieser zufälligen Einflüsse) geringfügig ist gegenüber der Wirkung der gesamten Summe.

Satz: Ist die Zufallsvariable X binomisch verteilt mit Erwartungswert $E(X) = n * p$ und Varianz $\sigma^2 = n * p * (1 - p)$, und ist ferner

$$n * p * (1 - p) > 9$$

dann ist X näherungsweise normal verteilt, und es gilt

$$P(a \le X \le b) = \frac{1}{\sigma\sqrt{2\pi}} * \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

1.5 Mehrere Zufallsvariablen

1.5.1 Unabhängigkeit

Definition: Zwei Zufallsvariablen X und Y sind stochastisch, wenn für jeden Teilbereich I der Wertemenge von X und jeden Teilbereich J der Wertemenge von Y gilt
$$P(X \in I \cap Y \in J) = P(X \in I) * P(Y \in J)$$

1.5.2 Summen von Zufallsvariablen

Satz: Seien X_1 und X_2 zwei Zufallsvariablen, und sei

$$Y = a_1 * X_1 + a_2 * X_2 \text{ wo } a_1, a_2 \in \text{IR}$$

Dann gilt

$$E(Y) = a_1 * E(X_1) + a_2 * E(X_1)$$

Sind X_1 und X_2 unabhängig, dann gilt weiter

$$Var(Y) = a_1^2 * Var(X_1) + a_2^2 * Var(X_2)$$

2 Statistik

2.1 Stichproben

2.1.1 Das arithmetisches Mittel

Gegeben sei eine Folge von n reellen Zahlen a_1, a_2, ..., a_n. Das arithmetische Mittel dieser Zahl ist

$$\overline{a} = \frac{a_1 + a_2 + \ldots + a_n}{n}$$

In einer Folge von Dateien a_1, a_2, ..., a_n kann es natürlich vorkommen, dass sich einige Werte wiederholen. Wir nehmen an, a_1 komme m_1-mal vor, a_2 komme m_2-mal vor, ..., a_k komme m_k-mal vor. Das arithmetische Mittel dieser Zahl ist

$$\overline{a} = \frac{a_1 * m_1 + a_2 * m_2 + \ldots + a_n * m_n}{m_1 + m_2 + \ldots + m_n}$$

2.1.2 Das geometrische Mittel

Gegeben sei eine Folge von n reellen Zahlen a_1, a_2, ..., a_n. Wir setzen voraus, dass die Werte alle positiv sind. $a_k > 0$. Das geometrische Mittel der Zahl ist

$$\overline{a}_{geo} = \sqrt{a_1 + a_2 + \ldots + a_n}$$

2.1.3 Das harmonische Mittel

Gegeben sei eine Folge von n reellen, strikt positiven Zahlen a_1, a_2, ..., a_n. Wir nehmen an a_1 komme m_1-mal vor, a_2 komme m_2-mal vor, ..., a_k komme m_k-mal vor. Das harmonische Mittel dieser Zahl ist

$$\overline{a}_{har} = \frac{m_1 + m_2 + \ldots + m_n}{\frac{m_1}{a_1} + \frac{m_2}{a_2} + \cdots + \frac{m_n}{a_n}}$$

2.1.4 Der Median

Definition: Gegeben sei eine aufsteigende(oder absteigende) Folge von n reellen Zahlen:

$$a_1 \le a_2 \le a_3 \le \ldots \le a_n$$

Der Median ist

$Md = a_{r+1}$ falls n = 2r+1 ungerade

$Md = \frac{1}{2}(a_r + a_{r+1})$ falls n = 2r gerade ist

2.1.5 Quantile

Regel:

(1) Ist n*α keine ganze Zahl, dann nimmt man die kleinste Zahl k, welche grösser als n*α ist, und setzt $Q_\alpha = a_k$

(2) Ist n*α eine ganze Zahl, also n*α = k, dann setzt man $Q_\alpha = \frac{1}{2}(a_k + a_{k+1})$

Beispiele:

(a) Zahlenfolge: 1; 2; 2; 4; 5; 7; 9; 12; 15; 20; 21; 24; 27; 28
unteres Quartil: $Q_{1/4}$= 4
oberes Quartil: $Q_{3/4}$= 21
Media $Md = \frac{9+12}{2} = 10,5$

(b) Zahlenfolge: -5,1; -4,9; -4,7; -3,9; -3,0; -2,0; -1,9; -0,8; 1.2; 2,2; 3,0; 3,1
10%-Quantil: $Q_{1/10}$= -4,9
oberes Quartil: $Q_{3/4} = \frac{1,2+2,2}{2} = 1,7$
Median $Md = \frac{-2,0-1,9}{2} = 1,95$

2.1.6 Masse für die Streuung

Definition: Die empirische Varianz der Folge V ist

$$S_V^2 = \frac{\sum_{i=1}^{a}(v_i - \bar{v})^2}{n-1}$$

Definition: Die empirische Standardabweichung ist

$$S_V = \sqrt{\frac{\sum_{i=1}^{a}(v_i - \bar{v})^2}{n-1}}$$

2.2 Korrelation

2.2.1 Kovarianz und Regressionsgerade

Gegeben seien n Punktpaare ($x_1|y_1$), ...,($x_n|y_n$). Es seien $\bar{x}$ bzw. $\bar{y}$ der Mittelwert und S_x^2 bzw. S_y^2 die empirische Varianz

1) Die Grösse
$$C_{xy} = \frac{1}{n-1}\sum_{k=1}^{a}(x_k - \bar{x})(y_k - \bar{y})$$
heisst die empirische Kovarianz der Stichprobe. Diese Grösse kann positiv oder negativ sein.

2) Die Gerade
$$g: y - \bar{y} = a\cdot(x - \bar{x}), \quad \text{wo } a = \frac{C_{xy}}{S_x^2}$$
heisst die erste Regressionsgerade(oder Ausgleichsgerade) der Stichprobe. Bei dieser Geraden nimmt die Summe der Quadrate der Abweichung der y-Werre den kleinstmöglichen Wert an.

3) Die Gerade

$$g^* = x - \bar{x} = a^* \cdot (y - \bar{y}), \qquad \text{wo } a^* = \frac{C_{xy}}{S_y^2}$$

heisst die zweite Regressionsgerade. Sie minimalisiert die Abweichungen der x-Werte.

Beide Regressionsgeraden gehen durch den Schwerpunkt $(\bar{x} \mid \bar{y})$ der Stichprobe.

2.3 Testen von Hypothesen

2.3.1 Vorgehen

Unsere Testregeln können aus zweierlei Gründen zu falschen Entscheiden führen:

(1) Es kann vorkommen, dass die Hypothese H_0 richtig ist, und wir sie trotzdem verwerfen. Der Bereich V ist aber so gewählt, dass dies nur mit dem geringen Risiko von α = 4,1% geschieht. Man nennt das einen Fehler der ersten Art.

(2) Es kann auch vorkommen, dass wir die Hypothese H_0 als richtig anschauen, obwohl sie falsch ist. Man nennt das einen Fehler der zweiten Art. Das Risiko solcher Fehler ist im Allgemeinen schwer abzuschätzen.

Vorgehen zum Testen

1. Zuerst muss man sich darüber klar sein, welche Vermutung man bestätigen oder widerlegen will. Das geschieht indem man die so genannte Nullhypothese H_0 aufstellt.
2. Man plant ein Zufallsexperiment: Ziehen einer geeigneten Stichprobe
3. Nun wählt man das so genannte Signifikanzniveau α (in der Praxis meistens 5%, manchmal auch 1%). α ist die Wahrscheinlichkeit dafür, einen Fehler der ersten Art zu begehen (also die Hypothese H_0 fälschlicherweise zu verwerfen).
4. Nun muss man den Verwerfungsbereich V so festlegen, dass das Signifikanzniveau eingehalten wird.
5. Man führt das Experiment durch und entscheidet.

2.4 Parameterschätzung bei der Normalverteilung

2.4.1 *Fall 1: σ^2 ist bekannt, μ ist unbekannt*

Die beste Schätzfunktion (im Sinne des Maximum-Likelihood-Prinzips) ist der Mittelwert

$$\overline{X} = T_\mu(X_1, X_2, \ldots, X_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Als Schätzung für μ nehmen wir also den Mittelwert

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

2.4.2 *Fall 2: σ^2 ist unbekannt, μ ist bekannt*

Das Maximum-Likelihood-Prinzip führt hier auf die Schätzfunktion

$$T = T_{\sigma^2}(X_1, X_2, \ldots, X_n) = \frac{1}{n}\sum_{k=1}^{n}(X_k - \mu)^2$$

Als Schätzung für σ^2 nehmen wir die Grundgesamtheit Varianz also

$$\frac{1}{n}\sum_{k=1}^{n}(x_k - \mu)^2$$

2.4.3 *Fall 3: σ^2 und μ sind unbekannt*

In der Praxis ist das der geläufige Fall. Mit dem Maximum-Likelihood-Prinzip ergeben sich für die unbekannten μ und σ^2 die Schätzfunktionen

$$\overline{X} = T_\mu(X_1, X_2, \ldots, X_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

$$T = T_{\sigma^2}(X_1, X_2, \ldots, X_n) = \frac{1}{n}\sum_{k=1}^{n}(X_k - \mu)^2$$

Zum Schätzen μ und σ^2 nehmen wir

- Den Mittelwert $\overline{x} = \dfrac{x_1 + x_2 + \ldots + x_n}{n}$
- die empirische Varianz $S_X^2 \; \dfrac{1}{n-1}\sum_{k=1}^{n}(x_k - \overline{x})^2$

2.5 Prüfverteilungen

2.5.1 *Definition*

Sei X eine stetige Zufallsvariable mit Dichtefunktion f(x), und sei η eine Zahl mit $0 \le \eta \le 1$. Das η-Quantil von X ist die Zahl u, für die gilt

$$P(X \le u) = \int_{-\infty}^{u} f(x)dx = \eta$$

2.5.2 *Die Chi-Quadrat-Verteilung*

Gegeben seien n Zufallsvariablen X_1, X_2, ..., X_n von denen wir annehmen,

- dass sie unabhängig sind und
- dass sie alle standard normal verteilt sind (also μ=0 und σ=1)

Nun bildet man die neue Zufallsvariable

$$Y := X_1^2 + X_2^2 + \ldots + X_n^2$$

Definition: Die Verteilung von Y heisst die χ^2-Verteilung mit n Freiheitsgraden.

Satz: Die Zufallsvariablen Y hat den Erwartungswert E(Y) = n und die Varianz Var(Y) = 2n.

2.5.3 *Die t-Verteilung (Student-Verteilung)*

Gegeben seien zwei Zufallsvariablen X und Y. Wir setzen voraus

- dass X normal verteilt ist mit Erwartungswert 0 und Varianz und
- dass Y eine χ^2-Verteilung mit n Freiheitsgraden hat.

Nun bildet man die neue Zufallsvariable

$$T := \frac{X}{\sqrt{\frac{Y}{n}}}$$

Definition: Die Verteilung von T heisst die t-Verteilung mit n Freiheitsgraden

Satz: Die Zufallsvariable T hat den Erwartungswert E(T)=0 und die Varianz $Var(T) = \frac{n}{n-2}$

2.6 Konfidenz-Schätzungen

2.6.1 Vertrauensintervalle für den Mittelwert bei bekannter Varianz

Bestimmung eines Konfidenzintervalls für den Mittelwert μ einer Normalverteilung bei bekannt Varianz σ^2

Schritt 1: Wahl eines Konfidenzniveaus γ (95%, 90%, 99% oder dergleichen)

Schritt 2: Man bestimmt u so, dass $\Phi(u) = \frac{1+\gamma}{2}$

Schritt 3: Man berechnet den Mittelwert $\bar{x}$ der Stichprobe.

Schritt 4: Man berechnet $a := \frac{u * \sigma}{\sqrt{n}}$

Das Konfidenzintervall für den Mittelwert der Grundgesamtheit ist

$$\bar{x} - a \leq \mu \leq \bar{x} + a$$

2.6.2 Vertrauensintervalle für die Varianz

Bestimmung eines Konfidenzintervalls für die Varianz σ^2 einer Normalverteilung

Schritt 1: Wahl eines Konfidenzniveaus γ (95%, 90%, 99% oder dergleichen)

Schritt 2: Mit Hilfe der Tabelle über χ^2-Verteilung mit (n-1) Freiheitsgraden bestimmt man

$$v := \frac{1-\gamma}{2} - Quantil \text{ und } w := \frac{1+\gamma}{2} - Quantil$$

Schritt 3: Man berechnet die empirische Varianz s^2 der Stichprobe. Das Konfidenzintervall für die Varianz der Grundgesamtheit ist

$$\frac{(n-1)s^2}{w} \leq \sigma^2 \leq \frac{(n-1)s^2}{v}$$

2.6.3 *Vertrauensinterballe für den Mittelwert bei unbekannter Varianz*

Bestimmung eines Konfidenz-Intervalls für den Mittelwert μ einer Normalverteilung unbekannter Varianz σ^2.

Schritt 1: Wahl eines Konfidenzniveaus γ (95%, 90%, 99% oder dergleichen)

Schritt 2: In der Tabelle der t-Verteilung mit n-1 Freiheitsgraden (wobei n der Umfang der Stichprobe ist) sucht man

$$w := \frac{1+\gamma}{2} - Quantil$$

Schritt 3: Man berechnet den Mittelwert $\bar{x}$ und die empirische Varianz s^2 der Stichprobe.

Schritt 4: Man berechnet $a := \frac{w * s}{\sqrt{n}}$

Das Konfidenzintervall für den Mittelwert μ ist

$$\bar{x} - a \leq \mu \leq \bar{x} + a$$