|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Begriffe** | | | | | | | | | |
| Statistik | | Entwicklung und Anwendung von Methoden zur Erhebung, Aufbereitung, Analyse und Interpretation von Daten | | | | | | | |
| Beschreibende Statistik | | Vollständige Kenntnis über das Untersuchungsobjekt | | | | | | | |
| Schliessende Statistik | | Für Untersuchung liegend die Daten des zu untersuchenden Objekts nur zum Teil vor. | | | | | | | |
| Hypothese | | Eine Hypothese ist eine Aussage deren Gültigkeit man für möglich hält, die aber nicht bewiesen oder verifiziert ist. | | | | | | | |
| Nullhypothese | | Die Nullhypothese H0 ist eine Aussage von der angenommen wird, dass sie stimmt. | | | | | | | |
| Alternativhypothese | | Die Alternativhypothese H1 beschreibt eine Annahme, sie ist also das Gegenteil der Nullhypothese. | | | | | | | |
| Fehler 1. Art (alpha) | | Fehlerhaftes Verwerfen einer Hypothese | | | | | | | |
| Fehler 2. Art (beta) | | Fehlerhaftes Annehmen einer Hypothese | | | | | | | |
| Zufällige Fehler | | Nicht reproduzierbar | | | | | | | |
| Systematische Fehler | | Reproduzierbar (können vermieden werden, unterliegt keinen grossen Schwankungen) | | | | | | | |
| Validierung | | Mache ich das Richtige (Überprüfung des Modells) | | | | | | | |
| Verifikation | | Mache ich es richtig (Verifiziertes Modell kann nicht valide sein) | | | | | | | |
| Merkmalsträger | | Der Gegenstand der statistischen Untersuchung | | | | | | | |
| Abgrenzungsmerkmal (sachlich, räumlich, zeitlich | | Sachlich: wer/was ist unter Merkmalsträger zu verstehen  z.B. Wer gilt als „Mitarbeiter“ eines Unternehmens | | | | | | | |
| Räumlich: Räumliche Grenzen, in denen der Merkmalsträger liegen muss z.B. ein Bürogebäude eines Konzerns | | | | | | | |
| Zeitlich: Zeitpunkt oder Zeitraum, an der ein Merkmalsträger „existieren“ muss, um Teil der Grundgesamtheit zu sein ≠ zum Zeitpunkt der Messung/Erhebung! | | | | | | | |
| Grundgesamtheit | | Die Menge aller Merkmalsträger die für eine Untersuchung in Frage kommen | | | | | | | |
| Merkmal | | Eigenschaften der Merkmalsträger die von Interesse sind | | | | | | | |
| Merkmalswert | | Der Wert der Beobachtung / Messung | | | | | | | |
| Primärstatistik | | Die Daten wurden genau für diesen Zweck erhoben (teuer) | | | | | | | |
| Sekundärstatistik | | Existierende Daten wobei es ungewiss ist, wie die Daten erhoben wurden. (günstig) | | | | | | | |
| Vollerhebung | | Befragung aller Merkmalsträger (Kosten und Umfang meist zu gross) | | | | | | | |
| Teilerhebung | | Befragung der essentiellen Merkmalsträger (wird meist gemacht) | | | | | | | |
| Diskrete Funktion | | Mit Lücken | | | | | | | |
| Stetige Funktion | | Ohne Lücken | | | | | | | |
| Formale Abhängigkeit | | Zahlenmässig begründete Abhängigkeit | | | | | | | |
| Sachliche Abhängigkeit | | Ist der Wert eines Merkmals kausal/ursachlich für den Wert eines zweiten Merkmals abhängig | | | | | | | |
| Menge | | Ungeordnet, ohne Redundanzen | | | | | | | |
| Tupel | | Geordnet, mit Redundanzen | | | | | | | |
| Zufallsexperiment | | Ein Experiment welches beliebig oft durchgeführt werden kann und das Ergebnis komplett vom Zufall abhängig ist | | | | | | | |
| Disjunkt | | Keine gemeinsame Teilmenge | | | | | | | |
| Zielgrösse | | Beschreiben die Grösse, die man optimieren möchte | | | | | | | |
| Einflussgrösse | | Sind Grössen welche die Zielgrösse beeinflussen. Es wird zwischen Streu,- und Störgrössen unterschieden. Man unterscheidet zwischen Steuergrössen und Störgrössen | | | | | | | |
| Steuergrössen | | Eine einstellbare Grösse (die man auch für eine gewisse Zeit halten kann) | | | | | | | |
| Störgrössen | | Eine Grösse deren Wert man nicht beeinflussen kann | | | | | | | |
| Faktoren | | Aus allen Einflussgrössen werden die wesentlichen/relevant Faktoren genannt. Es wird zwischen Quantitativen und Qualitativen Faktoren unterschieden: | | | | | | | |
| Quantitative Faktoren | | Quantitative Faktoren: Die Werte sind auf einer Ordinalskala beschreiben | | | | | | | |
| Qualitative Faktoren | | Qualitative Faktoren: Die Werte sind auf einer Nominalskala beschrieben | | | | | | | |
| Faktorstufen | | Die Werte die ein Faktor in einem Versuch annehmen soll, werden Faktorstufen genannt.  Kann ein Faktor nicht genau gemessen werden, so sollte der Abstand der Faktorstufen mindestens  6x die Varianz sein | | | | | | | |
| Komplexität | | Hohe Anzahl an Faktoren | | | | | | | |
| Kompliziertheit | | Unbekannte oder schwierig zu beschreibende Faktoren | | | | | | | |
| **Symbole** | | | | | | | | | |
|  | | Absolute Häufigkeit (Anzahl) | | | | | | | |
|  | | Relative Häufigkeit (Anteil) | | | | | | | |
|  | | Kumulierte absolute Häufigkeit | | | | | | | |
|  | | Kumulierte relative Häufigkeit | | | | | | | |
|  | | Mittelwert | | | | | | | |
|  | | Varianz | | | | | | | |
|  | | Standardabweichung | | | | | | | |
|  | | Arithmetisches Mittel (Durchschnitt) | | | | | | | |
|  | | Nullhypothese | | | | | | | |
|  | | Alternativhypothese | | | | | | | |
|  | | Elementarereignis 🡪 Teilmenge der Ergebnismenge | | | | | | | |
|  | | Ergebnismenge / Ergebnisraum (Menge aller möglicher Ausgänge eines Zufallsexperiments) | | | | | | | |
|  | | Siehe Wahrscheinlichkeiten | | | | | | | |
|  | | Besteht aus allen möglichen Ergebniskombinationen (Potenzmenge der Ergebnismenge) | | | | | | | |
|  | | Anzahl Messungen / Stichprobenumfang | | | | | | | |
| N | | Grösse der Grundgesamtheit | | | | | | | |
|  | | Mittlere absolute Abweichung | | | | | | | |
|  | | Spannweite | | | | | | | |
| **Ableitungsregeln** | | | | | | | | | |
|  |  | | |  |  | |  | |  |
|  |  | | |  |  | |  | |  |
|  |  | | |  |  | |  | |  |
|  |  | | |  |  | |  | |  |
|  |  | | |  |  | |  | |  |
|  |  | | |  |  | |  | |  |
| Linearitätsregel | | | |  | | | | | |
| Produktregel | | | |  | | | | | |
| Produktregel mit Konstante c | | | |  | | | | | |
| Kettenregel | | | |  | | | | | |
| Quotientenregel | | | |  | | | | | |
| **Experimente** | | | | | | | | | |
| Zyklischer Prozess des Experimentierens nach Shewhard (Plan 🡪 Do 🡪 Check 🡪 Act)   1. Hypothese aufstellen 2. Experiment durchführen 3. Hypothese überprüfen 4. Hypothese/Modell gegebenenfalls anpassen | | | | | | | | | |
| Hindernisse für den Erkenntnisgewinn:   1. Komplexität: Hohe Anzahl an Faktoren 2. Kompliziertheit: Unbekannte oder schwierig zu beschreibende Faktoren 3. Rauschen/Dynamik: Unterschiedliche Ergebnisse bei gleichen Faktoren | | | | | | | | | |
| Experimente werden immer nach einem bestimmen Schema durchgeführt:   1. Ausgangssituation beschreiben 2. Untersuchungsziele festlegen / Zielgrössen definieren 3. Faktoren auswählen und gewichten 4. Versuchsplanung erstellen 5. Versuche durchführen 6. Ergebnisse auswerten und Vertrauensintervalle bestimmen 7. Ergebnisse interpretieren und Massnahmen ableiten 8. Überprüfen der «Verbesserungen» | | | | | | | | | |
| Prozessmodell: | | | | | | | | | |
| **DoE: Design of Experiment** | | | | | | | | | |
| Wie sind Experimente zu planen, damit mit möglichst wenigen Einzelexperimenten der Zusammenhang zwischen Einflussfaktoren und Zielgrössen möglichst genau ermittelt werden können. | | | | | | | | | |
| **Vorgehen** | | | | | | | | | |
| 1. Ausgangssituation spezifizieren / Problem beschreiben / Ziel definieren    1. Kunde und dessen Bedürfnisse definieren    2. Liegen bereits Daten vor    3. Welche Probleme müssen gelöst werden    4. Welche Ressourcen (Zeit und Geld) stehen zur Verfügung (Kosten/Nutzen Analyse)    5. Betroffene Gruppen und deren Beziehung untereinander listen (Wiederstände, Supporter, Wissensträger) 2. Zielgrösse beschreiben: Dabei möglichst alle Grössen sammeln und diese dann auf die wichtigen Reduzieren | | | | | | | | | |
| **Einfluss-Zielgrössen-Matrix** | | | | | | | | | |
| 1. Für jede Zielgrösse eine Spalte anlegen 2. In der ersten Spalte alle Einflussgrössen sammeln und in Einflussgrössen und Steuergrössen unterteilen 3. Für jede Einflussgrösse das vorhandenen Wissen über Grösse und Einfluss auf jede Zielgrösse sammeln (z.B. stark, schwach, kein, linear, nicht linear) | | | | | | | | | |
| **Fehlerrechnung:** | | | | | | | | | |
| Die Fehlerrechnung wird benötigt um den Bereich abzuschätzen, in denen der tatsächliche Wert mit einer gewissen Wahrscheinlichkeit liegt. | | | | | | | | | |
| **Zufällige Fehler** | | | Nicht reproduzierbar | | | | | | |
| **Systematische Fehler** | | | Reproduzierbar | | | | | | |
| **Absoluter Fehler** | | | Bei Summen und Differenzen addieren sich die absoluten Fehler (gleiche Einheit wie Messwert) | | | | | | |
| Der absolute Fehler kann mithilfe der relativen Fehlers berechnet werden 🡪  = relativer Fehler \* Wert | | | | | | |
| **Relativer Fehler** | | | wobei = absolute Fehler und t = Messwert Bei Produkten und Quotienten addieren sich die relativen Fehler (einheitenlos 🡪 %) | | | | | | |
| Mit Hilfe des relativen Fehler lässt sich gut Abschätzen, welcher Faktor verbessert werden sollte. (der mit dem grösseren Fehleranteil) | | | | | | |
|  | | | Bei Potenzen kann der relative Fehler mit dem Exponenten multipliziert werden. Z.B \* 🡪 r da sich der relative Fehler bei Multiplikationen addiert (r\*r\* 🡪 fr +fr = 2 \* fr) | | | | | | |
| **Nennwert die Fehlerangabe ergänzen** | | | | | | | | | |
| Mindestens | | | Letzte Stelle des Messwertes + 1 Stelle (auf halbe gerundet) | | | | | | |
| Höchstens | | | Letzte Stelle des Messwertes (auf 0.3/0.4 gerundet) | | | | | | |
| Beispiel | | | *Gemessener Wert (****t****)* | | | ***Δ t*** *von* | | ***Δ t*** *bis* | |
| 15.32s | | | ± 0.005s | | ± 0.04s | |
| 15.3s | | | ± 0.05s | | ± 0.4s | |
| 15.320s | | | ± 0.0005s | | ± 0.004s | |
| **Beispiele (Masseinheiten beachten)** | | | | | | | | | |
| Schätzung eines Rechtecks  Länge wird abgelesen: 28.15 - 22.35 cm = 5.8 cm  Fehlerschätzung beim Ablesen: ± 0.05 cm  *//Subtraktion: absolute Fehler addieren sich (links und rechts)*  Länge des Rechtecks: 5.8 ± 0.1 cm | | | | | | | | | |
| Berechnung der Fläche  Breite des Rechtecks gegeben mit 0.9 ± 0.1 cm  Berechnung des relativen Fehlers ∆B = 0.1/0.9 = 11.1%  Relativer Fehler der Länge: ∆L = 0.1/5.8cm = 1.7%  *//Multiplikation: relative Fehler addieren sich*  ∆B + ∆L = 12.8%  Fläche A = L \* B A = 0.9 \* 5.8 = 5.2 ± 0.7 cm2  absoluter Fehler der Fläche 0.128 \* 5.22 cm2 = 0.668 cm2 | | | | | | | | | |
| Bei Messgeräten ist der relative Fehler nicht auf den gemessenen Wert, sondern auf Messbereich bezogen. | | | | | | | | | |
| **Tipp für Rechnungen** mit Kombinationen von +/- und \*/:  1. Resultat berechnen ohne beachten der Fehlerangaben  2. Resultat berechnen unter Nutzen der Maximalwerte  3. Fehler ∆ ergibt sich durch die Differenz von 1. und 2. | | | | | | | | | |

|  |  |
| --- | --- |
| **Diagramme** | |
| Balkendiagramm | Y-Achse: Häufigkeit und X-Achse: Balken pro Klasse |
| Histogramm | Der Balken geht über die gesamte Klassenbreite |
| Polygonzug | Verbinden der Balken mit einer Linie, wobei jeweils der rechte Ecken verbunden wird. Beim Balkendiagramm wird die Mitte genommen. |

|  |  |  |
| --- | --- | --- |
| **Skalen** | | |
| Wir arbeiten hauptsächlich mit metrischen Skala (Intervall und Verhältnis) | | |
| Nominalskala (qualitativ) | Sind zwei Einheiten gleich oder ungleich?  Enthält Namen die gleichgewertet werden | |
| Geschlecht: {Feminin, Maskulin} Ortsname: {Berlin, Rom, Bern, Paris} Familienstand: {verheiratet, ledig, geschieden, verwitwet} | |
| Ordinalskala / Rangskala (qualitativ) | Es lässt sich zusätzlich eine Ordnung herstellen  Die Werte sind nicht mehr gleichgewichtet, sondern intensitätsmässig geordnet (in Klassen) | |
| Schulnote: {sehr gut, gut, genügend, schlecht} Umfragen: {Trifft zu, Trifft eher zu, Trifft eher nicht zu, Trifft nicht zu} Qualitätsstufe: {Standard, Business, First Class} | |
| Intervallskala (metrische Skala / Kardinalskala) (quantitativ) | Es lässt sich zusätzlich eine Aussage über die Abstände machen  Es kann der einfache Abstand (Intervall) gemessen werden. Hat keinen absoluten Nullpunkt | |
| Temperatur: {-12, .., 0, .., 42} Uhrzeit: {20:00, 0:00, 10:00} | |
| Verhältnisskala (metrische Skala / Kardinalskala) (quantitativ) | Es lässt sich zusätzlich eine Aussage über das Verhältnis machen  Hat einen absoluten Nullpunkt, deshalb Vergleich Aussagen möglich. Negative Werte sind nicht möglich. Besitzt das höchste Informationsniveau! | |
| Umsatz: {0M, 1M, 2M, 3M,.. } Alter: {0,1,..,40,.., gut, gut, genügend, schlecht} Gewicht: {0kg, 50kg, 60kg,..,80kg,120kg } | |
| **Häufigkeitsverteilung** | | |
|  | |  |

|  |  |  |
| --- | --- | --- |
| **Dichte:** | | |
| Wenn die **Klassenbreiten unterschiedlich gross** sind muss mit Dichte *di* gerechnet werden.  Ist die Klassenbreite gleich gross bzw. die Häufigkeit unklassifiziert, so ist *di = hi*. | | |
| **Häufigkeiten berechnen** | | |
| Häufigkeit für einen bestimmten Wert (z.B. 45) |  | |
| Arithmetisches Mittel  der Gesamtheit | **Unklassifiziert:** C:\Users\Michael Wieland\AppData\Local\Microsoft\Windows\INetCacheContent.Word\Screenshot from 2016-07-21 16-22-51.png  **Klassifiziert:** | |
| Modus (häufigster Wert) | Der Modus ist immer in der Klasse mit der höchsten Dichte. | |
| Median,  1. Quartil,  etc. | 1. Bestimmen von t    1. t = In wie viele Teile die Gesamtheit unterteilt ist (2=Median, 4=1.Quartil, etc.) 2. Klasse finden, in welcher der Median/Quartil liegt    1. (n/t) < H\_i 🡪 i = Klasse | |
| **Beispiel für klassifizierte Häufigkeit** | | |
| Welcher Anteil der Mitarbeiter ist < 45 Jahre alt? | | Anteil < 45 Jahre = 35%  Anteil > 45 Jahre = 100% - 35% = 75% |
| Rechnung mit Dichte: | | |
| d1 = 10/40 = 0.25 d2 = 15/10 = 1.5 d3 = 25/15 = 1.7  Modusklasse ist also Klasse 3, da sie die grösste Dichte hat. | | |

|  |  |  |
| --- | --- | --- |
| **Lagemasse / Lageparameter** | | |
| Arithmetisches Mittel / Mittelwert | Der Klassische Durchschnitt: Man addiert alle Messwerte und dividiert durch die Anzahl Messwerte | |
| Harmonisches Mittel | Ist zur Berechnung des Durchschnitts einzusetzen wenn das Merkmal aus einem Bruch hervorgeht. | |
| Bsp. Auf einer Strecke von 2 Kilomer benötigt ein Fahrzeug auf der Hinfahrt 10km/h und auf der Rückfahrt 30km/h | |
| Geometrisches Mittel | Ist die n-te Wurzel aus dem Produkt aller beobachteten Merkmalswerte  Verwendete man immer dann, wenn man Mittelwerte aus aufeinander aufbauenden Wachstumsfaktoren  berechnen will. Wichtig beim Geometrischen Mittel ist, dass man nicht den Prozentsatz selbst sondern die einzelnen Faktoren (Brüche)  einsetzt. | |
| Modus | Gibt den Wert an, der am häufigsten vorkommt | |
| Median | Die Mitte in einem geordneten Datensatz. Gibt es eine gerade Anzahl Elemente wird einfach der Schnitt der beiden in der Mitte liegenden Werte genommen. | |
| Quantil | Unterteilt die Gesamtheit in 2gleich grosse Teile | |
| Quartil | Unterteilt die Gesamtheit in 4 gleich grosse Teile | |
| Dezil | Unterteilt die Gesamtheit in 10 gleich grosse Teile | |
| Perzentil | Unterteilt die Gesamtheit in 100 gleich grosse Teile | |
| **Streumasse / Streuparameter** | | |
| Die Varianz und Standardabweichung werden in der Praxis für die Streuung eingesetzt. | | |
| Spannweite | Die Differenz zwischen dem grössten und kleinsten beobachteten Merkmal | |
| Zentraler Quartilsabstand / Interquartilsabstand | Die Differenz zwischen dem ersten und dritten Quartil    80% Dezilabstand = D9 – D1 | |
| Mittlere absolute Abweichung | Der Durchschnitt der Summe aller Differenzen zum Mittelwert  Die Betragsstriche der mittleren absoluten Abweichung ist unvorteilhaft (Fallunterscheidung). Deshalb arbeitet man viel öfter mit der Varianz | |
| Varianz | Durch das Quadrieren wird der Varianzwert sehr gross, weshalb man eher mit der Standardabweichung rechnet.    Rsp. vereinfacht:  C:\Users\Administrator\Dropbox\share\ExEv\varianz_umformung.png  **oder**  **../../../Desktop/Screen%20Shot%202016-07-21%20at%2015.47.24.p** | |
| Standardabweichung | Arithmetisches Mittel der Abweichung vom Mittelwert der Gesamtheit | |
|  | |
| Variationskoeffizient (%) | Die Standardabweichung im Verhältnis zum arithmetischen Mittel | |
| Der Langstreckenläufer und der Weitspringer erbringen, relativ gesehen eine nahezu gleichmässige Leistung | |
| **Boxplot** | | |
| Der Boxplot vermittelt einen schnellen Eindruck, in welchem Bereich die Daten liegen und wie sie sich in diesem Bereich aufteilen.   1. Der geordnete Datensatz wird in 4 Abschnitte aufgeteilt, die etwa gleich viele Werte umfassen    1. Minimum    2. Maximum    3. **Median**    4. Beim unteren Quartil (Min 25% aller Messwerte kleiner/gleich und Max 75% aller Messwerte (grösser/gleich)    5. Beim oberen Quartil Max 25% aller Messwerte kleiner/gleich und Min 75% aller Messwerte (grösser / gleich) 2. Die oberen und unteren Enden der Quartile mit Strichen verbinden = B 3. Verbindungslinie zwischen Min und unterem Quartil sowie eine Verbindungslinie zwischen Max und oberem Quartil = Whisker | | |
|  | | |
| **Zeitreihen** | | |
| X-Achse = Zeit / Y-Achse = Merkmalswerte 🡪 Punktdiagramm | | |
| **Gleitender Mittelwert** | | |
| Ziel: Glättung der Zeitreihe/Kurve, in dem die hohen und niedrigen Werte gegeneinander Abgeglichen werden. | | |
| Man berechnet immer das arithmetische Mittel über eine Auswahl aller Messwerte und verschiebt diese Auswahl kontinuierlich nach vorne. Aus den neuen Messwerten (arithemtische Mittel) wird anschliessend eine neue Zeitreieh erstellt. | | |
| ../../../Desktop/Screen%20Shot%202016-07-21%20at%2008.50.38.p  ../../../Desktop/Screen%20Shot%202016-07-21%20at%2008.50.43.p  ../../../Desktop/Screen%20Shot%202016-07-21%20at%2008.52.02.p  ../../../Desktop/Screen%20Shot%202016-07-21%20at%2008.50.56.p | | |
| **Regressionsanalyse** | | |
|  | | Die Regressionsanalyse ist das Bindeglied zwischen diskreten Messungen (mit Lücken) und für die Analysen wichtige Stetigkeit (ohne Lücken).   Das Ziel dabei ist die Abweichung zwischen den beiden Funktion so klein wie möglich zu halten  Man ermittelt damit den Zusammenhang zwischen X und Y Achse (linear, exponentiell, etc.) |
| **Regressionsfunktion für lineare Zusammenhänge** | | |
| Der gesuchte Wert sollte auf der Y-Achse, der gegeben auf der X-Achse liegen. | | |
| Regressionsgerade | Beschreibt den Zusammenhang zwischen dem unabhängigen Merkmal X und dem abhängigen Merkmal Y. | |
| Regressionsparameter | Gibt den tendenziellen Wert des Merkmals Y an, wenn der Wert des Merkmalswert x gleich Null ist. | |
| Regressionsparameter | Gibt als Steigungsmass an, um wie viele Einheiten sich der Wert des Merkmals Y tendenziell änder, wenn der Wert des Merkmals X um eine Einheit erhöht wird. | |
|  |  | |
|  |  | |
| **Beispiel** | 12 Studenten gingen im letzten Semester neben dem Studium einer Erwerbstätigkeit nach. In der nachfolgenden Tabelle sind der zeitliche Aufwand (Std./Woche) für die Erwerbstätigkeit X und der zeitliche Aufwand (Std./Woche) für das Studium Y angegeben.    Ein Student der 6 Stunden pro Woche erwerbstätig ist, will anhand der vorliegenden Daten ermitteln, wieviel Zeit er für sein Studium aufbringen kann. Es besteht ein linearer Zusammenhang zwischen Erwerbstätigkeit und Studium (Siehe Regressionsanalyse) | |
| Arbeitstabelle für die Regressionsgeraden und    = = 7; = = 35  = = -0.49  = 35 – (-0.49)\*7 = 38.43 | |
| **Resultat für** | : -0.49x + 38.43  Auf den Fall 6 Studen Erwerbstätigkeit angewendet folgt:  -0.49\*6 + 38.43 = 35.49 🡪 der Student kann tendenziell von einem Aufwand von 35.49 Stunden pro Woche für sein Studium ausgehen. Der tatsächliche Aufwand wird jedoch davon abweichen, da noch weitere Faktoren als die Erwerbstätigkeit einen Einfluss auf die Höhe der Studiendauer haben. | |
|  | = = -1.73  = 7 – (-1.73)\*35 = 67.55 | |
| **Resultat für** | -1.73y + 67.55  Die Regressionsgerade beschreibt die Tendenz des Zusammenhangs zwischen dem Zeitaufwand für da Studium und dem Zeitaufwand für die Erwerbstätigkeit.  Es kann somit der tendenziell anfallende Zeitaufwand für die Erwerbstätigkeit bestimmt werden. | |
| **Beispiel für den Zusammenhang mit Häufigkeitsverteilung** | Ein Unternehmen bezahlt seine 200 Beschäftigen (n) im Produktionsbereich nach Tarifgruppen aus. Aus der nachstehenden Tabelle kann die Verteilung der 124 weiblichen und der 76 männlichen Beschäftigten auf die Tarifgruppen ersehen werden.    Beschreiben sie den Zusammenhang zwischen den Merkmalen Geschlecht (X) und Tarifgruppenzugehörigkeit: | |
| **Resultat** | Im Unternehmen arbeiten deutlich mehr Frauen als Männer. Um eine Aussage über die Häufigkeitsverteilung der Geschlechter in den verschiedenen Tarifgruppen zu machen, müssen die relativen Häufigkeiten mit folgender Formel berechnet werden.  Bsp. = 38.43 🡪 relative Häufigkeit, weiblich in G1    Mit der relativen Häufigkeitswerten kann man nun feststellen, dass es in den Tarifgruppen 1 und 4 zu einer Verschiebung kommt, jedoch die Gruppen 2 und 3 geschlechtsunabhängig sind. | |

|  |  |
| --- | --- |
| **Wahrscheinlichkeitsrechnung / Stochastik** | |
| Menge | Ungeordnet, ohne Redundanzen |
| Tupel | Geordnet, mit Redundanzen |
| Zufallsexperiment | Ein Experiment welches beliebig oft durchgeführt werden kann und das Ergebnis komplett vom Zufall abhängig ist (z.B Werfen eines Würfels) |
| Elementarereignis | Ist ein möglicher Ausgang des Zufallsexperiments, wobei zwei Elementarereignisse sich immer gegenseitig ausschliessen. |
| Ergebnismenge | Umfasst alle möglichen Elementarereignisse eines Zufallsexperiments. Z.B {1,2,3,4,5,6} |
| Ereignis | Eine Teilmenge der Ergebnismenge. Z.B {2,4,6} |
| System der Ereignisse | Bei einem Zufallsvorgang gemessene Ereignisse, bilden zusammen ein System von Ereignissen. Dieses weist Eigenschaften auf, welche es ermöglichen Relation (Durchschnitt, Vereinigung, etc.) zu bilden. |
| Unmögliches Ereignis | Die leere Menge |
| Disjunkte Ereignisse | A und B sind disjunkt, wenn sie keine gemeinsame Teilmenge besitzen. |
| Laplace Experiment  (gut Fälle / alle Fälle) | Ein Experiment bei dem jedes Ergebnis dieselbe Wahrscheinlichkeit hat und die Ergebnismenge endlich/abzählbar ist. |
| Wie gross ist die Wahrscheinlichkeit im Lotto (6 aus 49) genau drei Richtige anzukreuzen? |
| Unabhängige Ereignisse | Zwei Ereignisse A und B sind voneinander unabhängig wenn gilt: |
| Beispiel: Es wurden folgende Wahrscheinlichkeiten berechnet: P(A) = 0,65; P(A|B) = 0,75;   * Da P(A) ≠ P(A|B) sind die beiden Ereignisse A und B abhängig. |

|  |  |
| --- | --- |
| Additionssatz | Die Wahrscheinlichkeit das A **oder** B eintritt   1. A und B sind vereinbar      1. A und B sind unvereinbar   C:\Users\Michael Wieland\AppData\Local\Microsoft\Windows\INetCacheContent.Word\additionssatz.png |
| Multiplikationssatz |  |
| Es gibt 5 Kugel, 4 rote und 1 weise. Wie gross ist die Wahrscheinlichkeit, dass man ohne zurücklegen zwei rote zieht?  🡪 Da man nicht zurücklegt, ist die Wahrscheinlichkeit der weiteren Kugeln abhängig von der von den vorherig gezogenen.  C:\Users\Michael Wieland\AppData\Local\Microsoft\Windows\INetCacheContent.Word\Multiplikationssatzbaum.png |
| Bedingte Wahrscheinlichkeit | Die Wahrscheinlichkeit von A unter der Voraussetzung von B |
| ../../../../Desktop/Screen%20Shot%202016-08-05%20at%2012.59.4   |  |  |  |  | | --- | --- | --- | --- | |  | **S** | **s** | **Total** | | **l** | 0.32 | 0.08 | *0.4* | | **l** | 0.18 | 0.42 | *0.6* | | **Total** | *0.5* | *0.5* | *1* | |
| Komplementäre Wahrscheinlichkeit | Die Wahrscheinlichkeit, dass A nicht eintritt |

|  |  |  |  |
| --- | --- | --- | --- |
| **Korrelation** | | | |
| Die Korrelation ist eine Kennzahl für den Zusammenhang zwischen mehreren Streudiagrammen | | | |
| Formale Abhängigkeit | | Zahlenmässig begründete Abhängigkeit | |
| Sachliche Abhängigkeit | | Ist der Wert eines Merkmals kausal/ursachlich für den Wert eines zweiten Merkmals abhängig | |
| Kovarianz | | Merkmalswertkombinationen (x\_i, y\_i) | |
| Kovarianz bei Stichproben | | Bei Stichproben verwendet man eine korrigierte Varianz, wobei man nicht nur n sondern durch n – 1 teilt | |
| Korrelationskoeffizient  Lineare Abhängigkeit | | Es resultiert immer ein Wert r zwischen -1 und 1:   1. Je mehr der Wert bei -1 liegt, desto mehr ähneln die Punkte im Streudiagramm einer Gerade mit negativer Steigung (stark linear abhängig) 2. Je mehr der Wert bei 0 liegt, desto grösser ist die Streuung der Punkte (linear unabhängig) 3. Je mehr der Wert bei +1 liegt, desto mehr ähneln die Punkte im Streudiagramm einer Gerade mit positiver Steigung (stark linear abhängig) | |
| **Permutationen und Kombinatorik** | | | |
|  | | | |
| **Permutationen** | | | |
| Permutationen  ohne Wiederholung/Zurücklegen (Merke: 0!=1) | | Anzahl Möglichkeiten n Objekte anzuordnen = Fakultät | |
| Eine Maschine muss vier Aufträge A, B, C, D nacheinander abarbeiten. Wie viel Anordnungen sind möglich: 4!=24  🡪 für den ersten Platz gibt es 4 Möglichkeiten, für den zweiten 3, usw. 🡪 n! | |
| Permutationen  mit Wiederholung/Zurücklegen | | Bei **identischen** Elementen werden diese in Klassen zusammengefasst. Es gibt dabei k Klassen mit jeweils nk identischen Elementen | |
| Beispiel: | | Von einem 6-stelligen Zahlenschloss weiss man, dass es sich mit einer bestimmten Folge der Ziffern 1, 1, 4, 4, 4 und 8 öffnen lässt. Wie viele Versuche sind maximal notwendig um das Zahlenschloss zu öffnen?  Gegeben sind n=6 Ziffern, die in k=3 Klassen von untereinander gleich Ziffern zerfallen. Die Klasse «1» enthält n1 = 2 Elemente, die Klasse «4» n2 = 3 und die Klasse «8» n3 = 1 Element. | |
| Bei der **Kombinatorik** geht es darum, aus n Elementen, k auszuwählen und anschliessend in eine Ordnung zu bringen. | | | |
| Binomialkoeffizient | | | Aus n Optionen, k auswählen  Im Rechner ist das die Funktion nCr(n,k) |
| **Kombinationen** ohne Wiederholung/Zurücklegen  (Mit Beachtung der Anordnung) | | | Im Rechner ist das die nPr(n,k) Funktion |
| Aus 5 Bewerber soll eine Rangliste der ersten 3 Plätze gemacht werden. Wie viele verschiedene Listen sind möglich? V3(5) = |
| **Kombinationen** ohne Wiederholung/Zurücklegen  (ohne Beachtung der Anordnung) | | | Ist gleich dem Binomialkoeffizienten    Im Rechner ist das die Funktion nCr(n,k) |
| Bsp. Lotto  Beim Lotto müssen aus 49 Zahlen 6 Zahlen ausgewählt werden. Wie viel Tipps sind möglich? |
| **Kombinationen** mit Wiederholung/Zurücklegen  (Mit Beachtung der Anordnung) | | |  |
| In einem Einkaufsladen gibt es unterschiedlich bemalte Vasen zu kaufen. Der Kunde möchte 3 Vasen für seinen Garten kaufen. Wie viele Möglichkeiten hat er, die Vasen in seinem Garten auf 4 Plätzen anzuordnent. |
| Bei einem Ziffernschloss muss man eine 5-stellige Zahl einstellen, die aus den Ziffern 0-9 gebildet wird. Wie viele Kombinationen gibt es? |
| **Kombinationen** mit Wiederholung/Zurücklegen  (ohne Beachtung der Anordnung) | | |  |
| In einem Rat werden 3 Sitze neu vergeben, es bewerben sich 6 Verbände darauf. Die wiederholte Auswahl eines Verbandes ist möglich. Wie viele mögliche Sitzverteilungen gibt es?  k= 3 n = 6 |
| **Zufallsvariablen** | | | |
| Zusammenhang Zufallsvariable und Merkmal  ../../../Desktop/Screen%20Shot%202016-07-21%20at%2010.25.58.p | | | |
| Realisation | Wert der Zufallsvariable für ein Ereignis z.B. Im Monopoly ist die Summe der Augenzahlen zweier Würfel entscheidend, wie weit ein Spieler vorrücken darf:  Zufallsvariable = Augensumme Realisationen = {2,3,4, … , 12} | | |
| Eine Zufallsvariable hat für jedes Ereignis eine bestimmte Wahrscheinlichkeit, dass das Ereignis eintreffen kann. | | | |
| **Diskrete Massenfunktion** | | | * Kann mit einem Stabdiagramm veranschaulicht werden (Ordinate (Y) = Wahrscheinlichkeit, Abszisse (X) = Ereigniswerte) * Wahrscheinlichkeit kann direkt abgelesen werden * Hat Lücken und nur positive Werte * Die Summe aller Einzelwahrscheinlichkeiten entspricht 1 = Fläche unter dem Graphen |
| **Stetige** **Dichtefunktion** / **Kontinuierliche** **Verteilungsfunktion** | | | * Auf der X-Achse sind unendliche viele Werte * Die Summe aller Einzelwahrscheinlichkeiten entspricht 1 = Fläche unter dem Graphen |
| Erwartungswert | | |  |
| Varianz | | |  |
| **Beispiele:** | | | |
| Ein Zufallsvorgang besteht im dreimaligen Werfen einer Münze. Entscheidend ist die Anzahl an Wappen.  Geben Sie die Wahrscheinlichkeits- und Verteilungsfunktion an.  Zufallsvariable: Anzahl Wappen  Realisation: 0,1,2,3   |  |  |  | | --- | --- | --- | | xi | f(xi) | F(xi) | | 0 | 0.125 | 0.125 | | 1 | 0.375 | 0.500 | | 2 | 0.375 | 0.875 | | 3 | 0.125 | 1.000 |  1. Berechnen Sie den Erwartungswert. 2. Berechnen Sie die Varianz und die Standardabweichung | | | |
| **Stichproben** | | | |
| Zufallsstichprobe | | | Aus der Grundgesamtheit werden Elemente zufällig ausgewählt |
| Einfache Stichprobe | | | Die Elemente der Stichprobe haben alle die gleiche Wahrscheinlichkeit |

|  |  |
| --- | --- |
| Geschichtete Stichprobe | Ist es möglich, Elemente mit gleichen Eigenschaften in Gruppen einzuteilen ist es sinnvoller, Teilstichproben pro Gruppe/Schicht zu nehmen, um genauer Aussagen über die Gesamtheit zu machen |
| **Schätzverfahren** | |
| Ist der Mittelwert, Standardabweichung und die Verteilungsfunktion nicht bekannt müssen diese mit Hilfe von Schätzfunktionen geschätzt werden. Ziel dabei ist es, von einer Stichprobe auf die Grundgesamtheit zu schliessen und dabei den Fehler einer falschen Schätzung zu minimieren. | |
| **Punktschätzung** | |
| Schätzfunktion für den Mittelwert |  |
| Schätzfunktion für Varianz und Standardabweichung |  |
| **ACHTUNG:** Evtl. Wurzel ziehen! Wir arbeiten meist mit der Standardabweichung | |
|  | |
|  | |
| P = Wahrscheinlichkeit (Anteilswerte) | |

|  |  |
| --- | --- |
| **Intervallschätzung** | |
| **Konfidenzintervall für den Mittelwert** | |
| Beispiel: Bekanntheitsgrad (unbekannte Varianz)  Ein Chemieunternehmen möchte den Bekanntheitsgrad eines von ihm hergestellten Waschmittels in Erfahrung bringen. Dazu werden 400 Personen zufällig ausgewählt und befragt. Das Waschmittel war 30 % der Befragten zumindest namentlich bekannt. Erstellung des zentralen 95%-Konfidenzintervalls für Θ.  **Schritt 1:** Festlegung der Verteilungsform von P  Wie: n \* P \* (1 - P) > 9 = 400 \* 0.3 \* 0.7 = 84 > 9 🡪 wahr, also approximativ normalverteilt  **Schritt 2:** Festlegung der Varianz / Standartabweichung von P  Resultat: = 0.02  **Schritt 3:** Ermittlung des Quantilswertes z 🡪 Gemäss geg. Konfidenzintervall (Unterscheidung einseitig/Beidseitig(zentral))  Resultat aus Tabelle  **Schritt 4:** Berechnung des maximalen Schätzfehlers  Resultat: z \* = 1.96 \* 0.02 = 0.04  **Schritt 5:** Ermittlung der Konfidenzgrenze  Resultat: W(0.30 – 0.04 Θ 0.30 + 0.04) = 0,95  W(0.26 Θ 0.34) = 0.95  Der Bekanntheitsgrad in der Grundgesamtheit wird mit einer Wahrscheinlichkeit von 95% vom Intervall [26%; 34%] überdeckt. | |
| ../../../Desktop/Screen%20Shot%202016-07-24%20at%2014.23.45.p  Genauigkeit erhöhen:   * Konfidenzgrenze behalten, Umfang n erhöhen * Umfang n behalten, Konzidenzniveau senken | |
| Konfidenzintervall für beidseitig begrenzt:  ../../../Desktop/Screen%20Shot%202016-07-21%20at%2015.33.53.p  1- α gibt die Wahrscheinlichkeit an, dass sich die Zufallsvariable/Stichprobenfunktion innerhalb des Intervalls befindet. (Konfidenzintervall) | |
| Der tägliche Kaffeekonsum in einem Büro:   |  |  | | --- | --- | | xi | f(xi) | | 1 | 20 | | 2 | 30 | | 3 | 40 | | 4 | 10 |   Wie hoch ist die Wahrscheinlichkeit, dass der Stichprobenmittelwert bei n=100 im Intervall (2.3;2.5) liegt?  Varianz-Berechnung siehe "Streuparameter" 🡪 0.84  ../../../Desktop/Screen%20Shot%202016-07-21%20at%2016.15.52.p  ../../../Desktop/Screen%20Shot%202016-07-21%20at%2016.12.18.p  Intervall von z=-1,09 bis +1,09 => 0.8621-0.1379 = 0.7242 🡪 72.42% | |
| **notwendiger Stichprobenumfang n beim Konfidenzintervall für das arithmetische Mittel** | |
| In diesem Fall wird gefordert, dass die Schätzung ein vorgegebenes Mindestmass an Genauigkeit e besitzt und dass diese Mindestgenauigkeit mit einer vorgegebenen Konfidenz bzw. Sicherheit erzielt wird.  gegeben: Konfidenz, Genauigkeit (e) 🡪 e = x̅ -μ  gesucht: Stichprobenumfang n | |
| Beispiel: Zuckerabfüllung  gegeben: Konfidenz z = 1.96 Genauigkeit e = 0.2 g Standartabweichung σ = 1.2  = 138.3  Es müssen 139 Packungen entnommen werden, um die gewünschte Genauigkeit zu erzielen (wegen n 30 ist im Falle einer beliebig verteilten Grundgesamtheit die Approximation durch die Normalverteilung zulässig). | |
| Beispiel: Eine Molkerei liefert an eine Lebensmittelkette 40'000 Flaschen Milch mit 1000ml Soll-Füllmenge. Die Stichproben haben eine durchschnittliche Füllmenge von 1000.25ml. Aufgrund von zahlreichen Kontrollen weiss man, dass die Ist-Füllmenge normalverteilt mit einer Streuung von = 1.2ml ist. Wie viele Flaschen Milch müssen der Lieferung entnommen werden, wenn folgende Dinge gegeben sind. gegeben: z= 1.96 (=95% Konf. Int beidseitig) e=0.25ml ( z = 🡪 nach n auflösen = 88.51 🡪 89 Stichproben | |
| **Konfidenzintervall für die Varianz** | |
| Es wird die folgende Schätzfunktion verwendet: | |
| Voraussetzungen für eine erwartungstreue Schätzung: das Merkmal X ist in der Grundgesamtheit normalverteilt und die Entnahme erfolgt mit Zurücklegen. | |
| **-Zweiseitiges Konfidenzintervall** | |
| Formel:  (Konfidenzintervall)  n =Stichprobenumfang/Freiheitsgrade  y = Mit k = n-1 und z.B. 95% Konf. Int 🡪 0.975 (und 0.025 (, in der Chi2 Tabelle 🡪 y herausfinden | |
| **-Einseitiges Konfidenzintervall (nach oben begrenzt)** | |
| Formel:  (Konfidenzintervall)  n = Stichprobenumfang/Freiheitsgrade  y = Mit | |
| **Testverfahren** | |
| Signifikanzniveau | Das Signifikanzniveau wird meist bei 5% angesetzt. Ist der Wert kleiner wie 5% wird angenommen, dass ein Ergebnis signifikant ist. |
| Hypothese / Nullhypothese H0 | Die Nullhypothese H0 ist eine Aussage von der angenommen wird, dass sie stimmt. |
| Alternativhypothese H1 | Die Alternativhypothese H1 beschreibt eine Annahme, sie ist also das Gegenteil der Nullhypothese. |
| Fehler 1. Art (alpha) | Fehlerhaftes Verwerfen einer Hypothese |
| Fehler 2. Art (beta) | Fehlerhaftes Annehmen einer Hypothese |
| Parametertest | Man möchte wissen, ob der angegebene Benzinverbrauch eines Autos eingehalten wird. μ=10l/100km, σ=1l/100km  Es werden 25 Autos getestet. Dabei kommt der Mittelwert 10.2L/100km zustande. Liegt das Ergebnis im Bereich statistischer Schwankungen, wenn 1-α=0.95  ../../../../Google%20Drive/HSR/Module/2FS16/ExEv/Screen%20Shot%202016-08-05%20at%2009.00Annahme: Normalverteilter Strichprobenmittelwert  z für zweiseitige Tests: 1.96  Intervall:  ≤ x̅ ≤  9.61≤ x̅ ≤ 10.39 🡪 Nullhypothese annehmen! |
| Anteilswert  (unbekannte Wahrscheinlichkeit) | ../../../../Google%20Drive/HSR/Module/2FS16/ExEv/Screen%20Shot%202016-08-05%20at%2010.19 |

|  |  |  |
| --- | --- | --- |
| **Verteilungen** | | |
| **Zentraler Grenzwertsatz** | Der zentrale Grenzwertsatz besagt, dass sich mit grösserem Stichprobenumfang n, alle Verteilungen der Normalverteilung approximieren. Als Faustregel gilt, dass bei einem n über 30 Stichproben die Normalverteilung genommen werden kann. | |
| **Stetige Verteilungen** | | |
| Stetige Verteilungen sind überabzählbar. Das heisst, sie beinhalten so viele Werte, dass diese nicht einfach gezählt werden können. | | |
| **Normalverteilung** |  | |
| **Standardnormalverteilung** | Die Standardnormalverteilung ist der einfachste Fall der Normalverteilung, wenn der Mittelwert = 0 und die Varianz = 1 ist. | |
| **Z-Wert berechnen** |  | |
| Kleiner: Z Wert direkt ablesen | |
| Grösser: 1 – Z Wert aus der Tabelle | |
| Beidseitig: Min/Max Z-Werte herauslesen und Differenz bilden | |
| **In einer Frabrik wird Zucker abgefüllt. Der Mindestinhalt jeder Tüte soll 1000g beinhalten. Gegeben: μ=1002g, σ=1,5g**  **Wie hoch ist die Wahrscheinlichkeit, dass eine Tüte das Sollgewicht unterschreitet?  🡪Tabelle🡪0.0918🡪9.18%** | |
| **Chi-Quadrat-Verteilung**  **(Varianz!)** | **Voraussetzungen** | |
| * Zufallsvariablen sind unabhängig und normalverteilt | |
| **Wird aus der Normalverteilung abgeleitet** | |
| **Anwendung:**   * **Schätzung von Verteilungsparametern (z.B. Varianz)** | |
| **../../../Desktop/Screen%20Shot%202016-07-21%20at%2015.14.13.p** | |
| **Erwartungswert**  **../../../Desktop/Screen%20Shot%202016-07-21%20at%2015.15.47.p** | **Varianz**  **../../../Desktop/Screen%20Shot%202016-07-21%20at%2015.15.53.p** |
| **T-Verteilung** | **Voraussetzungen** | |
| Freiheitsgrade: k=n-1 | |
| Das Füllgewicht von Leberwürsten ist normalverteilt. Das Soll Mindestgewicht ist 125g.Aus den täglich produzierten 600 Würsten werden 26 gewogen:  ../../../Desktop/Screen%20Shot%202016-07-24%20at%2014.33.09.p  Erstellen Sie das zentrale 95% Konfidenzintervall.   1. berechnen der Stichprobenparameter: 2. Festlegen der Verteilungsform von x̅  X normalverteilt und σ2 unbekannt 🡪t-verteilt mit k=n-1 Freiheitsgraden 3. Festlegen der Standardabweichung von x̅ Varianz unbekannt, ohne Zurücklegen n/N<0.05 🡪 4. Festlegen von t 1-α =0.95 (zweiseitiges Intervall!) und k=25 🡪 Tabelle: t=2.060 5. Berechnen der Maximalen Schätzfehlers: t\*σx̅ = 2.060 \* 0.34 = 0.70 g 6. Berechnen der Konfidenzgrenzen W(124.58 -0.70 ≤ μ ≤ 124.58 + 0.70)=0.95 W(123.88 ≤ μ ≤ 125.28)=0.95 | |
| **Exponentialverteilung** | Anwendungen:   * Zeitspanne zwischen zwei Anrufen in einer Telefonzentrale * Dauer eines Telefongesprächs. * Lebensdauer eines Geräts, wenn Defekte durch äussere Einflüsse und nicht durch Verschleiß verursacht werden. | |
| **Wahrscheinlichkeitsdichte: λ = Durchschn. Eintritt**  **../../../Desktop/Screen%20Shot%202016-07-21%20at%2013.23.11.p** | |
| **Verteilungsfunktion: (entspricht der aufsummierten Wahrsch.)**  **../../../Desktop/Screen%20Shot%202016-07-21%20at%2013.30.24.p** | |
| **Erwartungswert**  **../../../Desktop/Screen%20Shot%202016-07-21%20at%2013.31.56.p** | **Varianz**  **../../../Desktop/Screen%20Shot%202016-07-21%20at%2013.32.07.p** |
| Ein Geschäft wird täglich zwischen 10.00 und 11.00 Uhr von durchschn. 3,5 Kunden besucht. Wie gross ist die Wahrscheinlichkeit, dass der Abstand zwischen dem Eintreffen zweier Kunden höchstens 0.2 Stunden beträgt?  **../../../Desktop/Screen%20Shot%202016-07-21%20at%2013.52.34.p** | |
| **Weibull Verteilung** | Beschreibt die Lebensdauer von Geräten oder Materialen mit Abnutzungserscheinung | |
| **Diskrete Verteilungen** | | |
| **Bernoulli Experiment** | Ist ein Zufallsexperiment mit genau zwei möglichen Ergebnissen (Treffer oder Niete). | |
| **Bernoulli Kette (binomial Verteilt)** | Bsp: Wie gross ist die Wahrscheinlichkeit mit 5 Würfel, 2x eine Sechs zu würfeln n = 5 k = 2 p = 1/6 q = 5/6 | |
|  | 20% der von einer Maschine produzierten Bolzen sind unbrauchbar. Wie hoch ist die Wahrscheinlichkeit, dass von 4 zufällig ausgewählten Bolzen höchstens zwei unbrauchbar sind?   * Genau zwei Möglichiche Ausgänge: defekt / i.O   = Summe aus P(0 unbrauchbar) + P(1 unbr.) + P(2 unbr.) | |
| **Binomialverteilung** | **Voraussetzungen** | |
| * Die Experimente sind voneinander unabhängig * Es gibt nur zwei Ausgangsmöglichkeiten * Anzahl der Versuche ist fix * Das Experiment wird immer identisch durchgeführt * Mit grösserem n, nähert sich die Binomialverteilung, ähnlich der Dichtefunktion einer Normalverteilung an | |
| **Wahrscheinlichkeitsfunktion (Bernoulli)** | |
| **Erwartungswert µ (relativ zum Nullpunkt)** | |
| **Varianz** | |
| Gegen eine Krankheit wurde ein neues Medikament entwickelt. Die Heilungschance liegt bei 90%. Wie gross ist die Wahrscheinlichkeit, dass bei 5 zufällig gewählten Patienten mindestens 4 geheilt werden? | |
| **Poissonverteilung** | **Voraussetzungen** | |
| * Wie hoch ist die Wahrscheinlichkeit, dass das Ereignis in einem Intervall genau oder höchstens x-Mal eintritt, wenn bekannt ist, dass in diesem Intervall das Ereignis im Mittel μ -Mal auftritt. * Hängt stark vom Mittelwert (μ) ab * typische Beispiele: Druckfehler/Seite, Arbeitsunfälle/Tag * Die Ereignisse treten unabhängig voneinander auf. (z.B Telefonanrufe) | |
| Seltene Ereignisse häufen sich (z.B. Bitfehler bei Flugzeugabstürzen) 🡪 Verteilung der seltenen Ereignisse | |
| Ankunftsrate: Eintreffende Ereignisse / Zeit  z.B. 24 Kunden in 8 Stunden 🡪 24/8=3 | |
| BEACHTE: Die Werte in der Verteilung Tabelle sind auf kumuliert 🡪 Ist ein einzelner Wert gesucht, muss die Differenz zum vorherigen Wert berechnet werden vorherige Tabellenwert abgezogen oder einfach die Formel verwendet werden. | |
| **Wahrscheinlichkeitsfunktion** | |
| **Erwartungswert:** | |
| Durchschnittlich 1 Telefonanruf/Minute. Wie gross ist die Wahrscheinlichkeit, dass 2 Anrufe pro Minute eingehen?    Achtung: Falls bis zu 2 Anrufe gefragt sind, müssen diese auf kumuliert werden.  Bei mehr als 2: | |
| **Rechteckverteilung** | Alle Realisationen in einem bestimmten Intervall [a, b] sind gleich wahrscheinlich | |