|  |  |
| --- | --- |
| **Univariate Statistik Grundlagen** | |
| |  |  |  | | --- | --- | --- | | **Skala** | **Eigenschaft** |  | | Nominal | Codierung | Qualitativ, n. metrisch ohne Masseinheit | | Ordinal | Plus Reihenfolge | | Intervall | Plus Abstände | Quantitativ, metrisch mit Masseinheit | | Verhältnis | Plus absoluter Nullpunkt |  |  |  | | --- | --- | | Diskret | Merkmal auf einer Skala nur bestimmte Werte annehmen (Alter in ganzen Jahren) | | Stetig | Merkmale auf einer Skala jeden Wert annehmen | |  |
| Summenzeichen  Produktzeichen | **Absolute Häufigkeit**: Vorkommnisse eines Merkmales z.B. 1877 Mal (von n)  **Relative Häufigkeit**: Anzahl in % =  **Relative Häufigkeit kummuliert**: Summe der Relativen Häufigkeiten  **Dichte**: Relative Häufigkeit pro Merkmalseinheit = |
| **Mittelwerte** | |
| **Arithmetisches Mittel (Durchschnitt):**  Additiv/durchschnittlich stetige Rendite. Sinnvoll für metrische skalierte Merkmale TR: 1-Var-Stat: L1 🡪 x̅ | **Gewichteter Durchschnitt (Notendurchschnitt):**  Arithmetische Mittel der Werte xi mit den absoluten Häufigkeiten Wi (bzw. den relativen Häufigkeiten wi), k = Anzahl der verschiedenen Merkmalswerte oder Anzahl Klassen. TR: 1-Var-Stat: L1, L2 🡪 x̅ |
| **Median (Zentralwert):**  Links und rechts des Median liegen 50% der aufsteigend geordnete Werte. Interpolieren TR:  LinReg mit L2 (Relative Häufigkeit kummuliert; letzter Wert unter 50% und erster Wert über 50%), L1 (Klassenobergrenzen der beiden Werte), ONE, YES; dann f(0,5) oder f(50) | Perzentile, Dezile, Quartile oder allgemein Quantile  Median = Untergrenze der Medianklasse + ( Medianklasse)  J: Anzahl Werte in der Medianklasse bis zur Mitte  F: Anzahl Werte in der Medianklasse  Der Median ist das 0.5 – Quantil, das 2. Quartil (2/4) (bei Streuung), das 5. Dezil (5/10) oder das 50. Perzentil (50/100). |
| **Modus:**  Wert welcher am häufigsten vorkommt / Klasse mit der grössten Dichte (Modalklasse). Dichte ist nicht grösste relative Häufigkeit  **Dichte**: Relative Häufigkeit pro Merkmalseinheit = | **Verteilung:**  Rechtsschief: Modus < Median < Arithmetisches Mittel  Linksschief: Arithmetisches Mittel < Median < Modus  Symmetrisch: Modus = Median = Arithmetisches Mittel |
| **Geometrische Mittel (Zeitlicher Durchschnitt):**  Multiplikativ/durchschnittlich effektive Rendite. Bei zeitlich durchschnittlichen Wachstumsraten (zeitlich aufeinanderfolgen).  Umsatz, etc. immer 1: 20% Gewinn 🡪 1.2; 20% Verlust 🡪 0.8  Geometrisches Mittel: | Durchschnittliche Wachstumsrate: -1 |
|  | |
| **Streuungsmasse (Abweichung vom Mittelwert berechnen)** | |
| **Spannweite:**  Spannweite = Maximum – Minimum | Eliminierung von Ausreissern durch **Quartilsabstand** (Boxplot)  Quartilsabstand = 3. Quartil (75%) – 1. Quartil (25%)  Maximum – 3. Quartil – Median – 1. Quartil – Minimum |
| **Varianz:** Durchschnittliche quadratische Abweichung  TR: 1-Var: L1 (xi), L2 (wi) oder | **Standardabweichung:** Durchschnittliche Abweichung von jedem Messwert zum Mittelwert (Bsp. 0.29 Franken pro Stück). TR: 1-Var L1(Noten, Preis), L2(Anzahl Werte pro Wert L1) |
| **Variationskoeffizient**  Streuung von verschiedenen Merkmalen mit unterschiedlichen Grössenordnung im Verhältnis zum Arithmetischen Mittelwert (in %) Gibt an wie gross die Standardabweichung im Vergleich zum AM ist. | **Volatilität (Finanzmathematik)**  Auf ein Jahr hochgerechnete Standardabweichung der stetigen (also logarithmierten) Renditen.   1. Index oder Kurs in L1 🡪 L2 stor data L1 (Erster L1 / Letzter L2 löschen) 2. L3 stor data L1/L2 (Effektive Rendite) 3. L1 stor data ln(L3) (Stetige Rendite)   AM der stetigen Rendite (Ø stetige Rendite pro Monat) =  Standardabweichung =  n = 4,12,52,250 (Quartal, Monate, Wochen, Börsentage) |
| **Indexzahlen** | |
| **Relative (prozentuale) Veränderung**  L2 = L1 (Daten) kopieren und Formel löschen  Ersten Eintrag in L1 und letzten Eintrag in L2 löschen  L3 = L1/L2 relative prozentuale Veränderung zum Vorjahr | **Zusammengesetzer Index (Preisindex von Laspeyres)**    Indexreihen lassen sich nur vergleichen, wenn beide den gleichen Basiszeitpunk haben. |
| . I= (  103,4 / 1.042 = 99.2  100,7 \* 1.042 = 104.9 | |  |  |  | | --- | --- | --- | | **Monat** | **LIK 2005** | **LIK 2010** | | August 10 | 103.4 | 99.2 | | Dezember 10 | 104.2 | 100 | | März 11 | 104.9 | 100.7 | |
| **Korrelation (Vermutete oder gestützte Zusammenhänge zwischen zwei Merkmalen)** | |
| **Kovarianz (Daten zentrieren)**  Je grösser X, desto grösser Y (Je kleiner X, desto kleiner Y) 🡪 Positiv  Je grösser X, desto kleiner Y (Je kleiner X, desto grösser Y) 🡪 Negativ | **Korrelationskoeffizient (Daten standardisieren)**  TR: LinReg[ax+b] L1, L2, ONE, YES   |  |  |  | | --- | --- | --- | | r (Linearer Zusammenhang) | Stärke des funkt. Zusammenhang | R^2 | | 0 | Kein nachweisbarer funkt. Zusammenhang | 0% | | ±0.25 | 5% | | ±0.5 | Schwach | 25% | | ±0.75 | Mittel | 50% | | ±1.0 | Stark | 100% | |
| **Regression (Funktionaler Zusammenhang zwischen zwei Merkmalen, durch Bestimmung der Parameter)** | |
| **Methode der kleinsten Quadrate (Trendrechnung t statt x)**  Ziel: Gerade, die möglichst nahe bei allen Punkten in einem Streupunktdiagramm liegt. Nähe wird in diesem Zusammenhang definiert, als senkrechter Abstand im Quadrat zwischen dem Datenpunkt und dem Punkt auf der Geraden. Idee: Quadratische Abweichung minimieren: Für welche Parameter ist die quadratische Abweichung minimal?   |  |  | | --- | --- | | Gleichung 1 | Gleichung 2 | | Kleinste Quadrate = (y – (Funktion (Tabelle oben)))^2  Gleichung nach a und b ableiten (=0)  X und Y mittels TR ausrechnen |
| **Bestimmtheitsmass**  Totale Varianz = erklä. Varianz (R^2) + nicht erklä. Varianz (1-R^2)  1. LinReg: L1 (x – Werte), L2 (y – Werte), YES 🡪 R^2 (Anteil erkl. Va. %)  2. L3 = funktion(L1) 🡪 Schätzer  3. L1 = L2 – L3 Differenz yi - Schätzer  1 – Var- Stas L2: (Sx)^2 🡪 Total  1 – Var- Stas L3: (Sx)^2 🡪 Erklärbare Varianz  1 – Var- Stas L1: (Sx)^2 🡪 Nicht erklärbare Varianz | 61,5% der Varianz der Investitionstätigkeit ist durch die Zinsentwicklung erklärbar, 38.5% sind nicht erklärbar. Es gibt noch andere Gründe für die Schwankung der Investition.  Je grösser R^2 desto besser passt sich die Regressionskurve der Punktewole an. |
| **Zeitreihenanalyse (Regelmässigkeiten in der Entwicklung zeitlich geordneter Beobachtungen)** | |
|  | |