[1 Streudiagramm und Korrelation 2](#_Toc74995731)

[1.1 Kovarianz 2](#_Toc74995732)

[1.2 Beispiel 2](#_Toc74995733)

[1.3 Eigenschaften Streudiagramm 3](#_Toc74995734)

[1.4 Korrelationsquoffizient 6](#_Toc74995735)

[1.4.1 Eigenschaften Korrelationsquoffizient 8](#_Toc74995736)

[1.5 Berechnung mit ti n-spire 9](#_Toc74995737)

[2 Die lineare Regression 12](#_Toc74995738)

[2.1 Diskrete Messwerte 12](#_Toc74995739)

[2.2 Beispiel 13](#_Toc74995740)

[2.3 Residuen 16](#_Toc74995741)

[2.4 Beispiel 16](#_Toc74995742)

[2.5 Beispiel aus dem Unterricht 17](#_Toc74995743)

# Beschreibende Statistik

## Ungewogenes arithmetisches Mittel (*Beobachtungswerte gegeben*)

Um das ungewogene arithmetische Mittel zu berechnen, addiert man alle gegebenen Beobachtungswerte bis und dividiert die so ermittelte Summe durch die Anzahl der Beobachtungswerten.

**Beispiel:**

Berechne das arithmetische Mittel.

|  |
| --- |
| 1. **Anzahl der Beobachtungswerte bestimmen** |
| Durch Abzählen stellen wir fest, dass es Beobachtungswerte gibt. |
| 1. **Formel aufschreiben** |
|  |
| 1. **Werte** **einsetzen** |
|  |
| 1. **Ergebnis berechnen** |
|  |

## Gewogenes artihmetisches Mittel *(Absolute Häufigkeiten gegeben)*

Um das gewogene arithmetische Mittel zu berechnen, addiert man zunächst die Produkte aller gegebenen Beobachtungswerte und ihrer absoluten Häufigkeiten von bis . Danach dividiert man die so ermittelte Summe durch die Anzahl der Beobachtungswerte .

**Beispiel**

Berechne das arithmetische Mittel.

|  |
| --- |
| 1. **Anzahl der Beobachtungswerte berechnen** |
|  |
| 1. **Formel aufschreiben** |
|  |
| 1. **Werte** **einsetzen** |
|  |
| 1. **Ergebnis berechnen** |
|  |

## Gewogenes artihmetisches Mittel (Relative Häufigkeiten gegeben)

Um das gewogene arithmetische Mittel zu berechnen, addiert man die Produkte aller gegebenen Beobachtungswerte und ihrer relativen Häufigkeiten von bis .

**Beispiel**

Berechne das arithmetische Mittel.

|  |
| --- |
| 1. **Formel aufschreiben** |
|  |
| 1. **Werte** **einsetzen** |
|  |
| 1. **Ergebnis berechnen** |
|  |

## Median

Der Median entspricht dem Wert, welcher größer oder gleich 50 % aller Werte ist.

**Beispiel 1 (gerade Anzahl)**

Wir fragen 12 Kinder nach ihrem Alter.

Die Antworten der Kinder lauten:

5, 3, 7, 4, 4, 3, 6, 4, 7, 8, 7, 6

Berechne den Median

|  |
| --- |
| 1. **Werte in aufsteigender Reihenfolge sortieren** |
|  |
| 1. **Median berechnen**   Da gerade ist, lautet die Formel zur Berechnung des Medians: |
| Mit erhält man demnach:  Aus der Tabelle lässt sich ablesen und .  Folglich gilt:  Der Median ist . |

**Beispiel 2 (ungerade Anzahl)**

Wir fragen 11 Kinder nach ihrem Alter.

Die Antworten der Kinder lauten:

5, 3, 7, 4, 4, 3, 6, 4, 7, 7, 6

Berechne den Median.

|  |
| --- |
| 1. **Werte in aufsteigender Reihenfolge sortieren** |
|  |
| 1. **Median berechnen**   Da ungerade ist, lautet die Formel zur Berechnung des Medians |
| Mit erhält man demnach:  Aus der Tabelle lässt sich ablesen .  Folglich gilt:  Der Median ist . |

## Erwartungswert

Der Erwartungswert ist eine Maßzahl zur Charakterisierung einer Wahrscheinlichkeitsverteilung.

Statt Maßzahl sagt man auch «**Kennzahl**» oder «**Kennwert**».

Der Erwartungswert ist ein «**Lageparameter**». Unter diesem Begriff werden alle Maßzahlen zusammengefasst, die eine Aussage über die Lage einer Verteilung machen.

Der Erwartungswert beschreibt die **zentrale Lage** einer Verteilung.

### Erwartungswert einer diskreten Verteilung

|  |
| --- |
| Ist eine diskrete Zufallsvariable, so heißt  der **Erwartungswert** von . |

**Beispiel 1**

Wir werfen einen Würfel. Berechne den Erwartungswert.

*Vorbereitung*

Die Zufallsvariable sei die Augenzahl beim Wurf eines symmetrischen Würfels.

Es gibt sechs mögliche Realisationen:

Alle sechs Realisationen haben die gleiche Wahrscheinlichkeit:

*Berechnung*

*Interpretation des Erwartungswerts*

Wenn man beispielsweise 1000-mal würfelt, die geworfenen Augenzahlen zusammenzählt und durch 1000 dividiert, ergibt sich mit hoher Wahrscheinlichkeit ein Wert in der Nähe von 3,5.

Das Beispiel zeigt, dass die Bezeichnung Erwartungswert irreführend sein kann: ist keineswegs der Wert, den man bei einem Wurf erwartet, denn 3,5 selbst kann nie als Augenzahl eintreten.

**Beispiel 1**

Wir spielen eine Runde Roulette. Berechne den Erwartungswert.

*Vorbereitung*

Die Zufallsvariable sei der Gewinn beim Roulette.

Wir setzen 1 € auf unsere Glückszahl. Falls wir gewinnen, erhalten wir 36 €. Unser Gewinn beträgt folglich 35 €, denn 1 € haben wir ja eingesetzt.

*Zur Erinnerung*: Beim Roulette kann man auf die Zahlen 0 bis 36 setzen.

Es gibt zwei Realisationen:

(falls wir verlieren)

(falls wir gewinnen)

Für die Wahrscheinlichkeiten gilt:

(in 36 von 37 Fällen verlieren wir)

(in 1 von 37 Fällen gewinnen wir)

*Berechnung*

*Interpretation des Erwartungswerts*

Wenn man beispielsweise 1000-mal auf seine Glückszahl setzt, die Gewinne und Verluste zusammenzählt und durch 1000 dividiert, ergibt sich mit hoher Wahrscheinlichkeit ein Wert in der Nähe von -3 Cent.

Bei einem fairen Spiel wäre der Erwartungswert gleich Null. Hier ist das Spiel unfair, da pro Runde im Schnitt ein Verlust von 3 Cent zu erwarten ist.

### Erwartungswert einer stetigen Verteilung

|  |
| --- |
| Ist eine stetige Zufallsvariable, so heißt  der Erwartungswert von . |

Dabei steht für die Dichtefunktion.

**Beispiel 1**

Ein Zufallsgenerator erzeugt zufällig eine Zahl zwischen -1 und 1.

Die Dichtefunktion des Zufallsgenerators ist:

Berechne den Erwartungswert.

*Interpretation des Erwartungswerts*

Wenn man bespielsweise 1000 Mal den Zufallsgenerator startet, die Zufallszahlen zusammenzählt und durch 1000 dividiert, ergibt sich mit hoher Wahrscheinlichkeit ein Wert in der Nähe von 0.

Da der Zufallsgenerator seine Werte symmetrisch im negativen und positiven Bereich streut, erwarten wir bei einer großen Anzahl an Zufallsexperimenten im Mittel den Wert 0.

**Beispiel 2**

Gegeben ist eine Zufallsvariable mit der Dichtefunktion

Berechne den Erwartungswert.

## Varianz

Die Varianz ist eine Maßzahl zur Charakterisierung einer Wahrscheinlichkeitsverteilung. Statt Maßzahl sagt man auch «**Kennzahl**» oder «**Kennwert**».

Die Varianz ist ein «Streuungsparameter». Unter diesem Begriff werden alle Maßzahlen zusammengefasst, die eine Aussage über die Streuung einer Verteilung machen.

Die Varianz beschreibt die **erwartete quadratische Abweichung** der Zufallsvariable von ihrem Erwartungswert.

**Nachteil** der Varianz ist, dass sie aufgrund der Quadrierung eine andere Einheit als die beobachteten Messwerte besitzt. Auf den ersten Blick können somit keine konkreten Aussagen über die Streuungsbreite abgeleitet werden. In der Praxis wird daher häufig die Standardabweichung, die sich aus Quadratwurzel der Varianz ergibt, herangezogen.

### Varianz einer diskreten Verteilung

In den folgenden beiden Abbildungen sind zwei Wahrscheinlichkeitsfunktionen dargestellt. Erkennst du den Unterschied zwischen einer kleinen und einer großen Varianz?

|  |  |
| --- | --- |
| Die Realisationen von sind **eng** um den Erwartungswert **gestreut**.  kleine Varianz | Chart, scatter chart  Description automatically generated |
| Die Realisationen von sind **breit** um den Erwartungswert **gestreut**.  grosse Varianz | Chart, scatter chart  Description automatically generated |

Im Folgenden schauen wir uns an, wie man die Varianz berechnet.

|  |
| --- |
| Ist eine diskrete Zufallsvariable, so heißt  die Varianz von . |

Dabei steht für den Erwartungswert.

|  |
| --- |
| **Verschiebungssatz für diskrete Zufallsvariablen** |

Der Verschiebungssatz erleichtert meist die Berechnung der Varianz.

**Beispiel 1**

Die Zufallsvariable sei die Augenzahl beim Wurf eines symmetrischen Würfels.

Es gibt sechs mögliche Realisationen:

Alle sechs Realisationen haben die gleiche Wahrscheinlichkeit:

|  |
| --- |
| 1. **Erwartungswert berechnen** |
| Der Erwartungswert ist |
| 1. **Varianz berechnen** |

## Standartabweichung

Die Standartabweichung beschreibt die erwartete Abweichung der Zufallsvariable von ihrem Erwartungswert.

### Standartabweichung einer diskreten Verteilung

In den folgenden beiden Abbildungen sind zwei Wahrscheinlichkeitsfunktionen dargestellt. Erkennst du den Unterschied zwischen einer kleinen und einer großen Standardabweichung?

|  |  |
| --- | --- |
| Die Realisationen von sind **eng** um den Erwartungswert **gestreut**.  kleine Standardabweichung |  |
| Die Realisationen von sind **breit** um den Erwartungswert **gestreut**.  grosse Standardabweichung |  |

Im Folgenden schauen wir uns an, wie man die Standardabweichung berechnet:

**Beispiel 1**

Die Zufallsvariable sei die Augenzahl beim Wurf eines symmetrischen Würfels.

Es gibt sechs mögliche Realisationen:

Alle sechs Realisationen haben die gleiche Wahrscheinlichkeit:

|  |
| --- |
| 1. **Varianz berechnen** |
| Die Varianz ist |
| 1. **Standardabweichung berechnen** |

**Beispiel 2**

Die Zufallsvariable sei der Gewinn beim Roulette.

Wir setzen 1 € auf unsere Glückszahl. Falls wir gewinnen, erhalten wir 36 €. Unser Gewinn beträgt folglich 35 €, denn 1 € haben wir ja eingesetzt.

*Zur Erinnerung*: Beim Roulette kann man auf die Zahlen 0 bis 36 setzen.

Es gibt zwei Realisationen:

(falls wir verlieren)

(falls wir gewinnen)

Für die Wahrscheinlichkeiten gilt:

(in 36 von 37 Fällen verlieren wir)

(in 1 von 37 Fällen gewinnen wir)

|  |
| --- |
| 1. **Varianz berechnen** |
| Die Varianz ist |
| 1. **Standardabweichung berechnen** |

### Standardabweichung einer stetigen Verteilung

In den folgenden beiden Abbildungen sind zwei Dichtefunktionen dargestellt. Erkennst du den Unterschied zwischen einer kleinen und einer großen Standardabweichung?

|  |  |
| --- | --- |
| Die Realisationen von sind **eng** um den Erwartungswert **gestreut**.  kleine Standardabweichung |  |
| Die Realisationen von sind **breit** um den Erwartungswert **gestreut**.  grosse Standardabweichung |  |

Im Folgenden schauen wir uns an, wie man die Standardabweichung berechnet:

**Beispiel 1**

Ein Zufallsgenerator erzeugt zufällig eine Zahl zwischen -1 und 1.

Die Dichtefunktion des Zufallsgenerators ist:

|  |
| --- |
| 1. **Varianz berechnen** |
| Die Varianz ist |
| 1. **Standardabweichung berechnen** |

**Beispiel 2**

Gegeben ist eine Zufallsvariable mit der Dichtefunktion:

|  |
| --- |
| 1. **Varianz berechnen** |
| Die Varianz ist |
| 1. **Standardabweichung berechnen** |

# Streuungsparameter

Unter dem Begriff Streuungsparameter werden alle statistischen Maßzahlen zusammengefasst, die eine Aussage über die Verteilung von einzelnen Werten um den Mittelwert machen.

## Wichtige Streuungsparameter

|  |  |
| --- | --- |
| **Spannweite** |  |
| **Interquartilsabstand** |  |
| **Mittlere absolute Abweichung** |  |

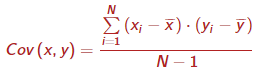
# Streudiagramm und Korrelation

Das Streudiagramm und die Berechnung der Korrelation werden bei der Analyse von Zusammenhang metrischer Merkmale angewendet.

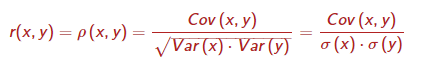
|  |  |
| --- | --- |
|  |  |
| **Streuungsdiagramm** | **Korrelation** |

Das Streudiagramm und der metrische Korrelationskoeffizient gehören nicht zwangsläufig zusammen, lassen sich gemeinsam aber gut erklären. Voraussetzung für die Anwendung ist, dass die zu untersuchenden Merkmale mindestens intervall skalliert sind.

## Kovarianz

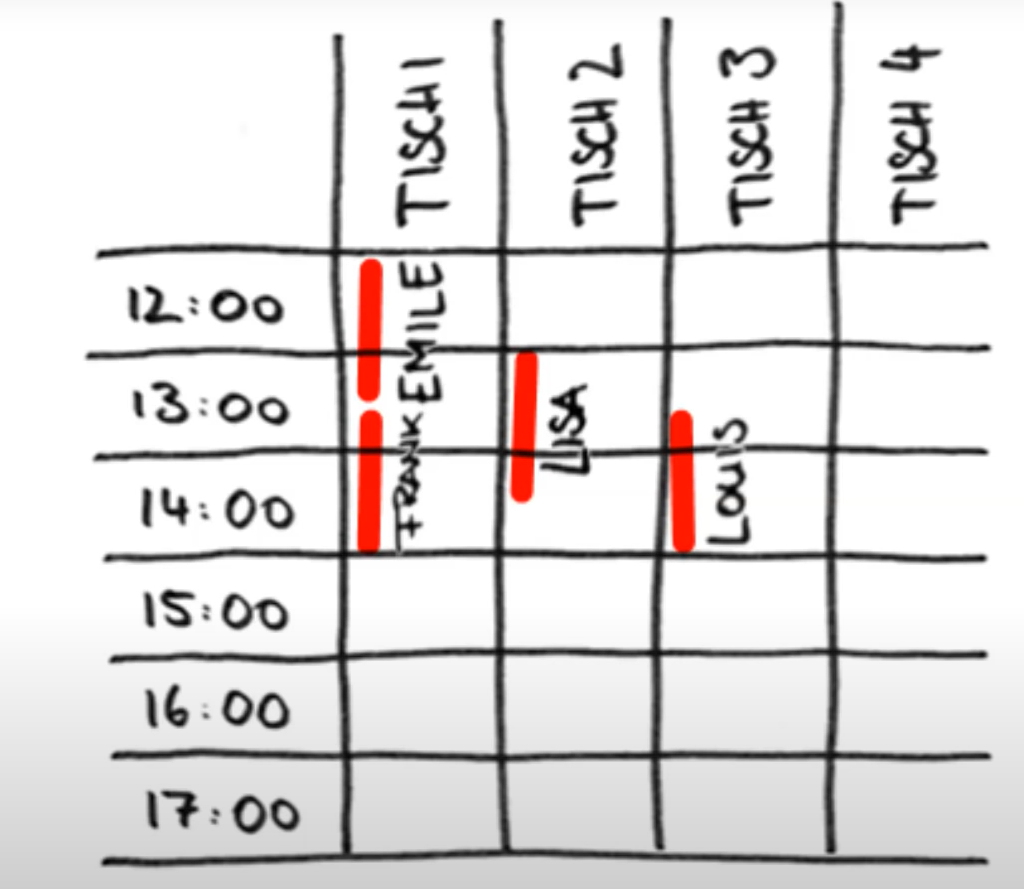


Für die Berechnung der Korrelationsquoffizienten, muss die Kovarianz berechnet werden:



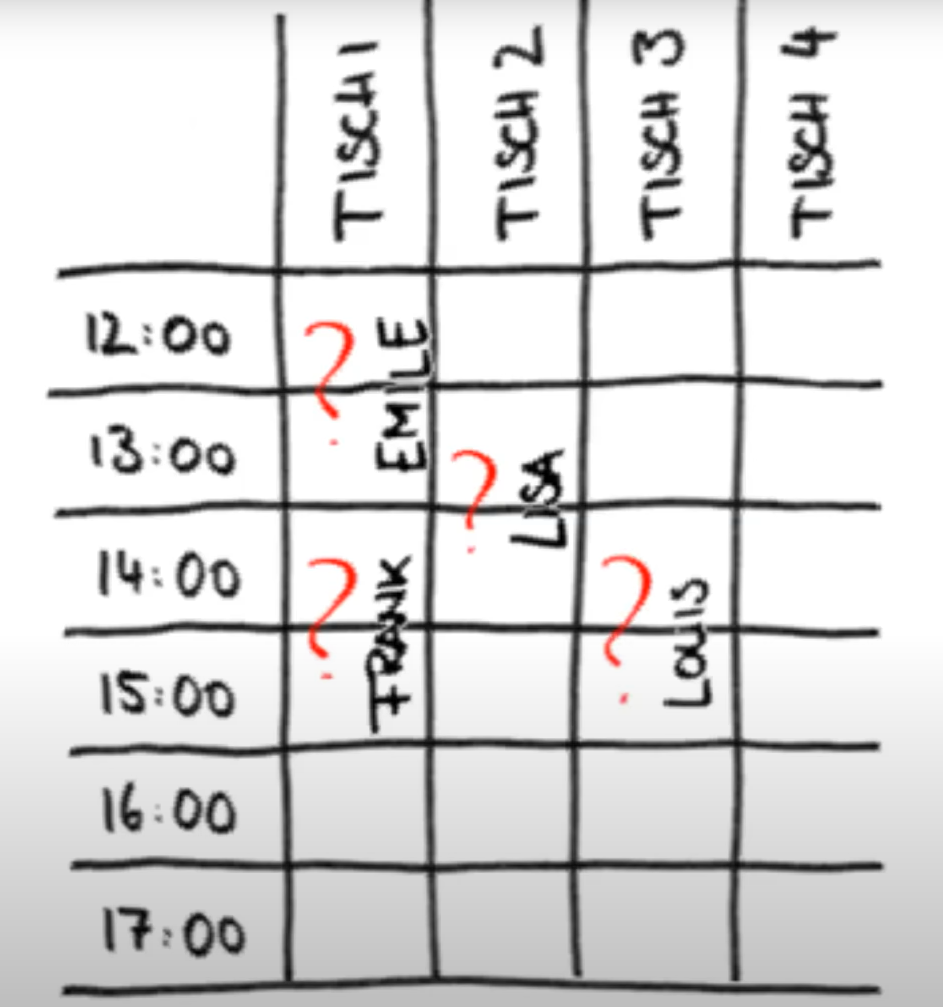
## Beispiel

Aureli besitzt ein gut laufendes Restaurant. Es läuft sogar so gut, dass Gäste immer zuerst ihren Tisch reservieren müssen. Aureli plant pro Reservierung und Tisch 90 Mintuten ein:



Ihre Gäste sind aber sehr unterschiedlich. Manche essen schnell und verlassen das Restaurant bereits nach kurzer Zeit wieder. Andere hingegen dehnen ihr Essen deutlich länger aus. Für Aureli ist das ein Problem. Essen die Gäste schnell hätte Sie für den Tisch ohne Weiteres ein 2. Reservierung entgegennehmen können. Oder noch schlimmer. Die Gäste essen länger, obschon der Tisch eigentlich weiter reserviert wurde. Dann erscheinen die neuen Gäste, obschon der Tisch noch nicht wieder frei ist.

Aureli möchte herausfinden, wie lange die Gäste den Tisch belegen werden.



Damit Aureli einige Informationen zum Verhalten der Gäste haben möchte, notiert sie immer, wenn Sie Zeit hat, wie lange ihre geblieben sind. Sie notiert sich **wieviele** Gäste einen Tische besetzt haben und **wie lange** die Gäste geblieben sind:

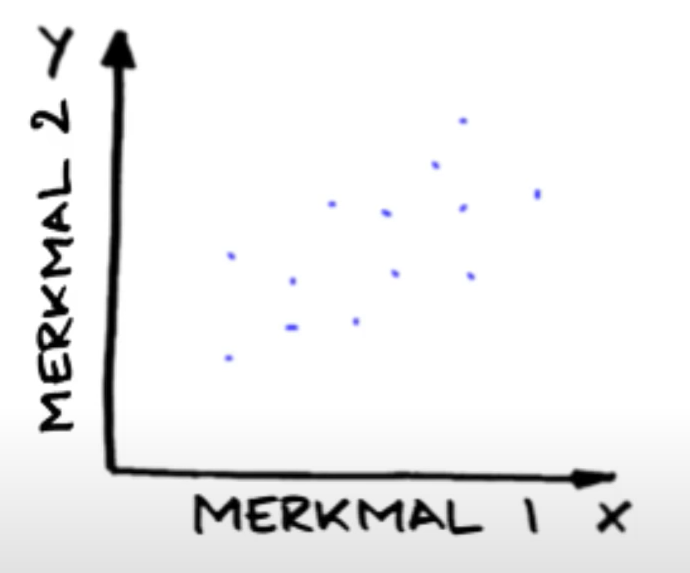


Ihre notierten Werte sehen wie folgt aus:

|  |  |  |  |
| --- | --- | --- | --- |
| **Nr.** | **Uhrzeit** | **Gäste** | **Dauer (min.)** |
| 1 | 12:30 | 5 | 105 |
| 2 | 17:00 | 2 | 55 |
| 3 | 15:30 | 1 | 50 |
| 4 | 12:00 | 2 | 60 |
| 5 | 17:00 | 4 | 80 |
| 6 | 16:00 | 4 | 70 |

## Eigenschaften Streudiagramm

Das Streudiagramm hilt bei der grafischen Darstellung von metrischen Merkmalen. Hier schreibt man die Merkmale an die waagrechte und das andere an die senkrechte Achse. Für jede statistische Wert wird ein Punkt an der Stelle, wo die Werte die Ausprägung der beiden Merkmale entspricht.



Um Ihre Werte bildlich darzustellen, trägt sie die verschiedenen Merkmale in Diagramme ein:

|  |  |
| --- | --- |
|  |  |
| Hier untersucht Sie den Zusammenhand zwischen der **Uhrzeit** und der **Dauer** des Ausenthalts ihrer Gäste. | Das Gleiche macht Sie hier mit der **Dauer** und **Anzahl** **Gäste**. |

|  |  |
| --- | --- |
|  | Diese Diagramme werden Streudiagramme genannt. |

Mit diesen Diagrammen haben wir die Zusammenhänge der Dauer und Anzahl Gäste sowie Uhrzeit und Dauer veranschaulicht. Nun analysieren wir, welche Merkmale zusammenhängen:

|  |  |
| --- | --- |
|  |  |
| In diesem Diagramm sehen wir, dass kleine Gruppen mindestens tendenziell die Tische kurz belegen. Wir sehen auch das grössere Personen gruppen länger bleiben. Wenn dieses Verhalten längerfristig eintrifft, so hätte Aureli eine gute Möglichkeit gefunden, die die Dauer der Tischbelegungen besser abzuschätzen. ***Dies vermittelt den Eindruck, dass je mehr Gäste einen Tischbesetzen, desto lang der Aufenthalt dauert.*** | Bei diesem Diagramm sehen wir, dass es bei frühen und genau so bei späten Tageszeiten Reservierungen gibt, die mal kurz und mal lange einzlene Tische belegen. ***Die Uhrzeit der Reservierung ist also kein guter Indikator für Dauer der Tischbelegung.*** |

Wir sehen also, dass das linke Diagramm uns hilft zu analysieren, wie lange ein Tisch besetzt sein wird. Der Unterschied ist:

|  |  |
| --- | --- |
|  |  |
| In diesem Diagramm verlaufen die Punkte ungefähr auf einer Linie. | In diesem Diagramm hingegen sind die Punkte im gesamten Diagramm verteilt. |

## Korrelationsquoffizient

Eine Kennzahl, die angibt, wie gut die Punkte in einem Streudiagramm eine Linie formen, ist der metrische Korrelationsquoffizient. Berechnen wir nun den Korrelationsquoffizient zur Uhrzeit und Dauer:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  |  | | --- | --- | --- | --- | | **Nr.** | **Uhrzeit** | **Gäste** | **Dauer (min.)** | | 1 | 12:30 | 5 | 105 | | 2 | 17:00 | 2 | 55 | | 3 | 15:30 | 1 | 50 | | 4 | 12:00 | 2 | 60 | | 5 | 17:00 | 4 | 80 | | 6 | 16:00 | 4 | 70 |   **x** = Uhrzeit  **y** = Dauer |
| 1. Als erstes berechnen wir den arithmetischen Mittelwert für für die Uhrzeit |  |
| 1. Nun berechnen wir den arithmetischen Mittelwert für die Dauer der Tischebesetzung |  |
| 1. Jetzt setzen wir die Werte in die Formel |  |
| 1. Wenn wir nun die Formel ausrechnen, bekommen wir für folgenden Wert |  |

Die gleiche Berechnung machen wir nun für Dauer und Anzahl Gäste:

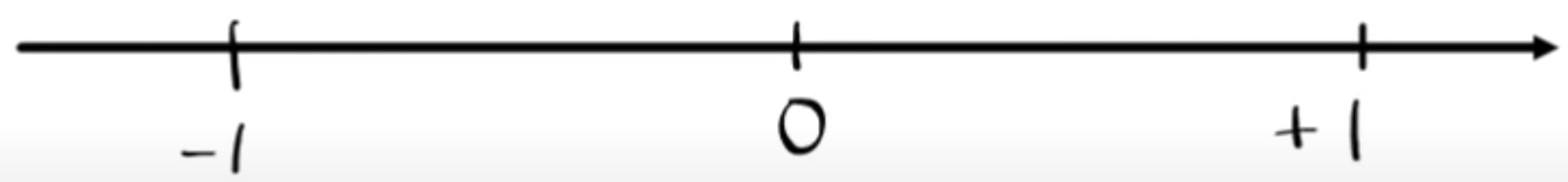
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  |  | | --- | --- | --- | --- | | **Nr.** | **Uhrzeit** | **Gäste** | **Dauer (min.)** | | 1 | 12:30 | 5 | 105 | | 2 | 17:00 | 2 | 55 | | 3 | 15:30 | 1 | 50 | | 4 | 12:00 | 2 | 60 | | 5 | 17:00 | 4 | 80 | | 6 | 16:00 | 4 | 70 |   **x** = Anzahl Gäste  **y** = Dauer |
| 1. Als erstes berechnen wir den arithmetischen Mittelwert für für die Anzahl der Gäste |  |
| 1. Nun berechnen wir den arithmetischen Mittelwert für die Dauer der Tischebesetzung |  |
| 1. Jetzt setzen wir die Werte in die Formel |  |
| 1. Wenn wir nun die Formel ausrechnen, bekommen wir für folgenden Wert |  |

So ergibt sich aus den Berechnungen für die Diagramme folgende Korrelationsquoffiziente:

|  |  |
| --- | --- |
|  |  |
|  |  |

### Eigenschaften Korrelationsquoffizient

Der Korrelationsquoffizient ist eine statistische Masszahl, welcher immer einen Wert zwischen und annimmt:



|  |  |  |
| --- | --- | --- |
|  |  |  |
| Je näher der Wert bei liegt, um so deutlicher formen die Punkte eine Gerade mit einer negativen Steigung. | Je näher der Wert bei liegt, um so weniger ähneln die Punkte einer Geraden. | Je näher der Wert bei liegt, um so deutlicher formen die Punkte eine Gerade mit einer positiven Steigung.  Und wenn wert genau lieght, bedeutet das, dass die Punkte genau auf einer Geraden liegen. |

Nun analysieren wir diese Eigenschaften bei den Werten von Aurelis Diagrammen:

|  |  |
| --- | --- |
|  |  |
| Hier sehen wir, dass die Punkte fast eine Gerade bilden. Auch Korrelationsquoffizent liegt nahe . | Bei diesem Diagramm sehen wir aber, dass die Punkte ziemlich verstreut sind. Auch der Korrelationsquofizient ist nahe . |

Was bringt es also den metrischen Korrelationsquoffizenten zu berechnen? Die Berechnung des metrischen Korrelationsquoffizienten bringt jeden etwas, der etwas über eine Sache wissen muss, die schwer zu erheben ist oder erst in der Zukunft bekannt wird und er stattdessen auf eine andere Grösse ausweichen will.

|  |  |
| --- | --- |
| Red Alert Icon Stock Illustration - Download Image Now - iStock | Korrelation Kausalität (Ursache Wirkung) |

## Berechnung mit ti n-spire

|  |  |
| --- | --- |
| Auf ***ctrl*** und ***doc*** |  |
| Im ***Menü Liste & Spreadsheet hinzufügen*** wählen |  |
| Nun erscheint ein Sheet für die Füllung der Werte |  |
| Nun haben wir die Werte eingetragen |  |
| Auf ***menu*** dann auf ***Statistik***, dann auf ***Statistische Berechnung*** und dann auf ***Statistik mit 2 Variablen***. Und nun folgende Werte eintragen: |  |
| Das Resultat |  |

# Die lineare Regression

Die Realität liefert uns oft Messwerte zu Einzelbeobachtungen. Wenn wir bspw. über die finanzielle Situation eines Unternehmens wissen wollen, dann schauen wir uns dessen Bilanz an. Die Bilanz jedoch gibt es jedoch genau einmal in Jahr.



Wenn wir auf Diät gehen, wiegen wir uns vielleicht jede Woche oder jeden zweiten Tag und unterstellen, dass unser Gewicht zwischendurch auch irgendwo dazwischen gelegen hat.

## Diskrete Messwerte



Solche Einzelwerte geben uns also sogenannte **diskrete** Messwerte. Solche diskreten Messwerte zu besitzen ist schon mal sehr wertvoll. Wenn wir uns aber beim Diätplan fragen zu welchem Zeitpunkt unser Gewicht am gerinsten oder der Gewichstverlust am Höchsten war, stossen wir schnell an unsere Grenzen.

|  |  |
| --- | --- |
|  |  |
| Diskret heisst mit Lücken. Veranschaulicht haben wir solche Messwerte mit Streudiagramme. | Stetige Messwerte bedeuten Werte ohne Lücken und Sprünge. Das ist also das, was wir aus der Realität bekommen. |

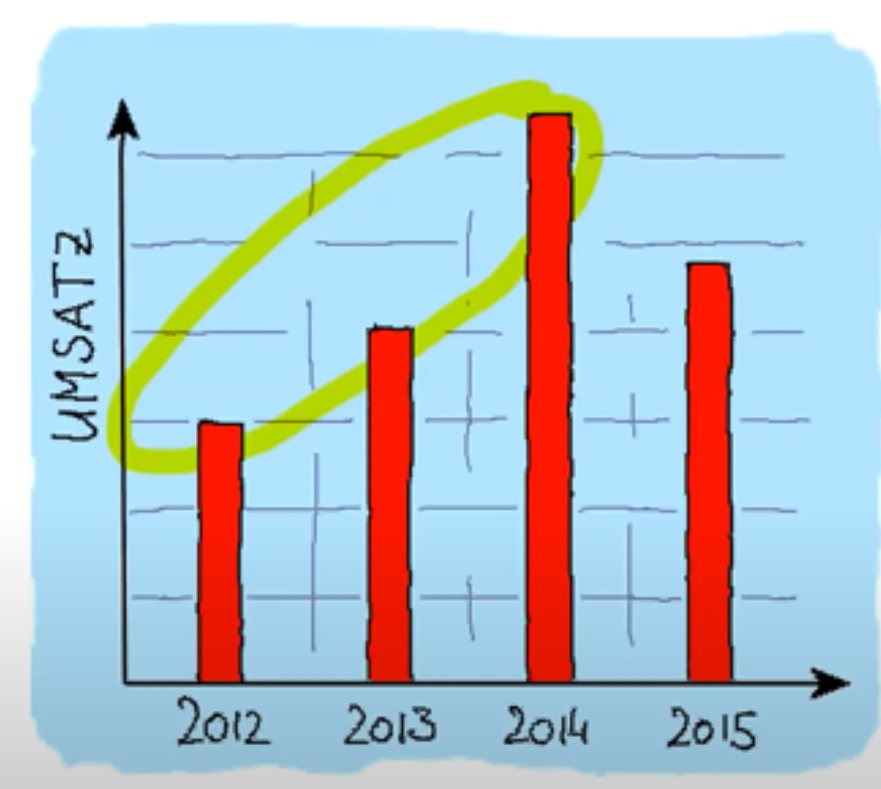
Die regressive Analyse ist das Bindeglied zwischen der diskreten Messungen und der stetigen.



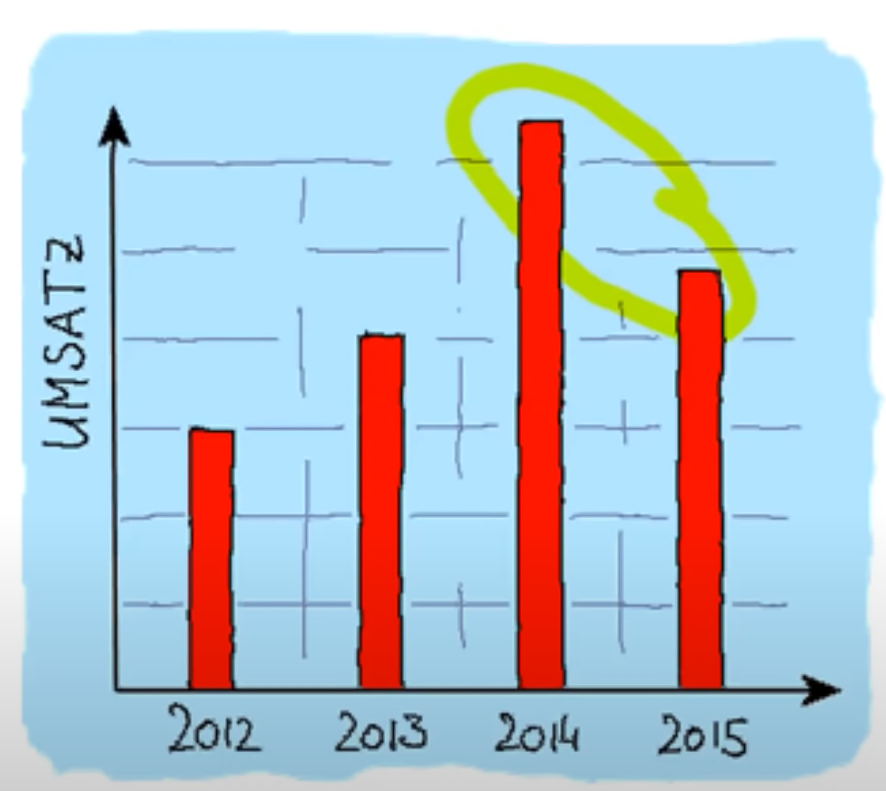
## Beispiel

Mark arbeitet nach seinem Studium im Controlling bei einer Firma, welche Bio- Limonaden verkauft. Der Mark für Bio- Limonanden wächst und daran hat die Firma viel Geld verdient. Das Wachstum hat den Chef der Firma dazu verleitet die Preis für die Flasche Bio- Limonade immer weiter zu erhöhen.

Zunächst hat es auch gut funktioniert und zu steigenden Umsätze geführt.

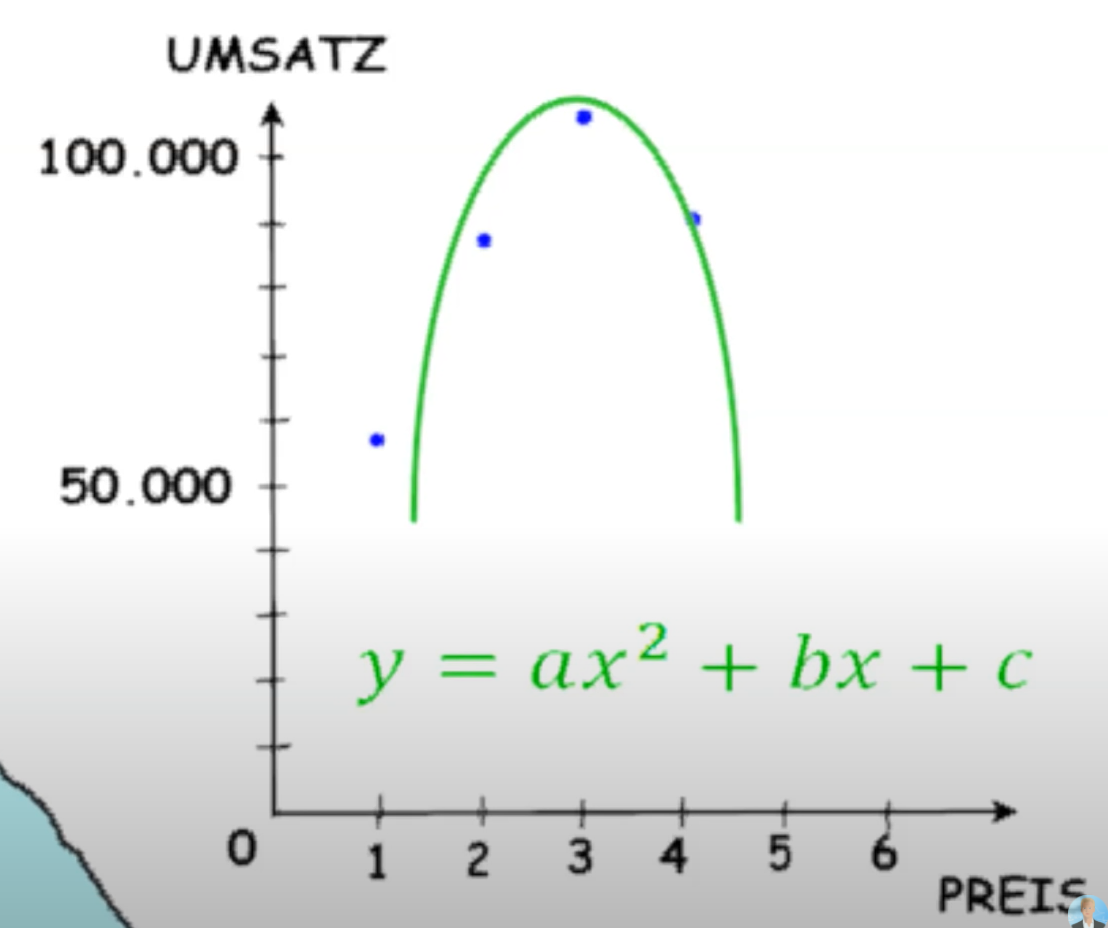


Nach der letzten Preiserhöhung jedoch fielen die Verkauszahlen.



Rene weiss vom Studium, dass die Nachfrage oft mit steigenden Preisen sinkt. Der Chef von Rene gibt ihm den Auftrag, auszurechnen bei welchem Preis der Umsatz wohl maximal würde.

Rene rechnet mit der linearen Preisabsatzfunktionen:



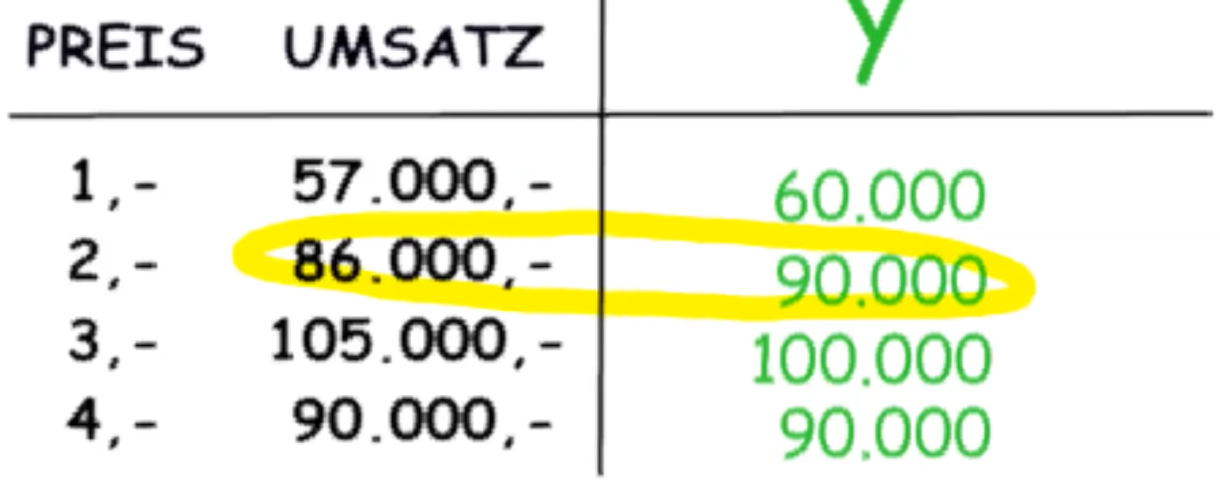
Wir nehmen für die Koeffizienten (**a,** **b** , **c**) der Funktion folgende Werte an:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  | | --- | --- | --- | | **Preis** | **Umsatz** | **y** | | 1.- | 57’000 | 54’000 | | 2.- | 86’000 | 86’000 | | 3.- | 105.000 | 96’000 | | 4.- | 90’000 | 84’000 |   **a** = -11’000  **b** = 65’000  **c** = 0  x = Preis |
| 1. Wir rechnen für den Verkaufspreis von 1 Euro. 2. Das machen wir dann auch für die anderen Verkaufspreisen |  |

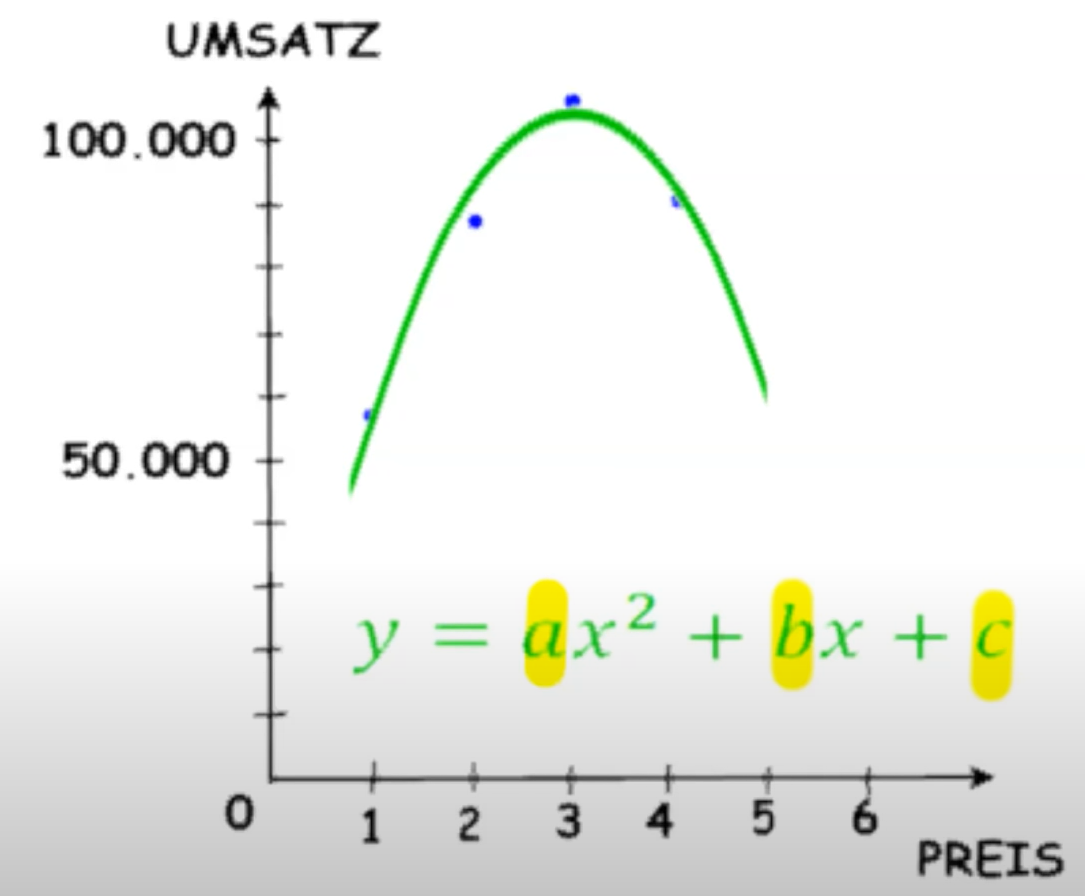
Wir nehmen für die Koeffizienten (**a,** **b**, **c**) nun neue Werte an:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  | | --- | --- | --- | | **Preis** | **Umsatz** | **y** | | 1.- | 57’000 | 60’000 | | 2.- | 86’000 | 90’000 | | 3.- | 105.000 | 100’000 | | 4.- | 90’000 | 90’000 |   **a** = -10’000  **b** = 60’000  **c** = 10’000  x = Preis |
| 1. Wir rechnen für den Verkaufspreis von 1 Euro. 2. Das machen wir dann auch für die anderen Verkaufspreisen |  |

Immer gibt es aber Abweigungen zwischen theoretischen und tatsächlichen Umsätzen:



Rene muss also die Werte für **a,** **b**, **c** so bestimmen, dass die Abweichungen (auch Fehler genannt) möglichst klein werden



Das Verfahren, welches hier zur Anwendung kommt, heisst **Regressionsrechnung**.

Nach der Ableitung der Werte ist die Berechnung des maximum Preiswertes gar kein Problem mehr:

|  |  |
| --- | --- |
| Hier die Werte der Quoffienten berechnet und in der Formel eingetragen |  |
| Da wir jedoch den Maximalwert haben möchten, drehen wir die berechnung um: |  |
| Hier der Maximalwert für den Verkaufspreis für einen maximalen Umsatz |  |

## Residuen

Ein Residuum, ganz grob gesagt, ist für eine bestimmte Beobachtung der Fehler (Abweichung), den die Vorhersage des gerechneten Regressionsmodells (Regressionsgerade) für diese Beobachtung gemacht hat. Sie sind eine wichtige Kennzahl bei der Regression. Man schaut sie sich typischerweise während der Modelldiagnose genauer an, um zu überprüfen ob die **Annahmen dieses Modells** plausibel sind.

Man berechnet das Residuum für die -te Beobachtung durch:

Grafisch kann man sich das schön veranschaulichen. Das Residuum für eine Beobachtung ist der farbig markierte Abstand:

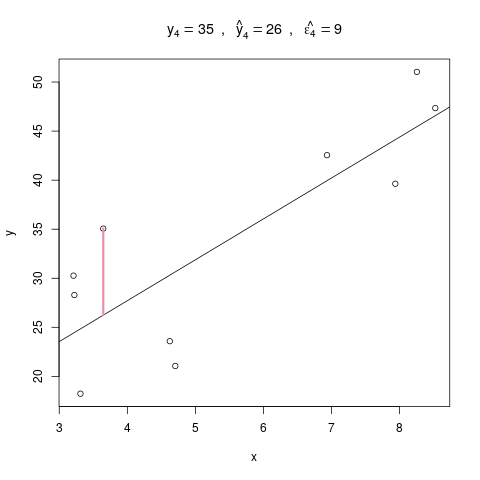


Abbildung 1: Residuen

## Beispiel

Für die vierte Beobachtung in dieser Grafik liegt der wahre -Wert bei 35, und der vorhergesagte (also der auf der Regressionsgeraden) bei 26. Das Residuum, der Fehler der Schätzung, liegt also bei .

## Beispiel aus dem Unterricht

In einem Lebensmittelbetrieb werden in einem klimatisierten Raum täglich 1000 Pakete Hüppen mit Schokoladefüllung hergestellt.

Da wegen einer kurzfristig angesetzten anstehenden Renovation die Klimaanlage einige Zeit mit reduzierter Leistung arbeiten wird, misst die Produktionsleiterin an zwölf Tagen die Temperatur T (in Grad Celsius) im Produktionsraum und ermittelt gleichzeitig die Anzahl A hergestellter Pakete, die nicht einwandfrei sind:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **T** | 16.8 | 17.2 | 17.4 | 17.6 | 17.7 | 17.2 | 18.4 | 18.1 | 17.2 | 17.5 | 18.9 | 17.0 |
| **A** | 61 | 58 | 58 | 76 | 82 | 64 | 95 | 95 | 70 | 73 | 101 | 58 |

Die Produktionsleiterin rechnet damit, dass während der Renovation die Temperatur auf 20 Grad Celsius steigen wird. Sie überlegt sich nun, mit wie vielen nicht einwandfreien Paketen sie in dieser Zeit rechnen muss.

Sie vermutet einen linearen Zusammenhang zwischen der Temperatur und der Anzahl nicht einwandfreier Pakete.

**Frage 1:**

Welche der folgenden Grössen ist zur Beurteilung, ob tatsächlich ein linearer Zusammenhang vorliegt, am besten geeignet?

**Antwort:**

Korrelationskoeffizient nach Bravais-Pearson

**Frage 2:**

Berechnen Sie den Korrelationskoeffizienten nach Bravais-Pearson für die Temperatur und die Anzahl! Sie können als Hilfsmittel Excel verwenden, wenn Sie möchten. Runden Sie auf zwei Stellen nach dem Komma.

**Antwort:**

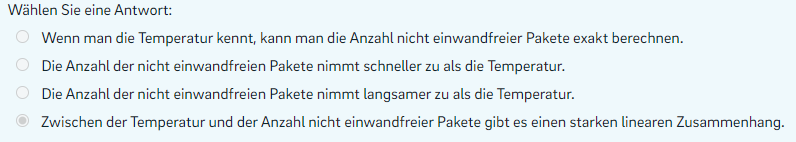
Der Korrelationskoeffizient beträgt



**Frage 3:**

Der Korrelationskoeffizient ist jetzt zu interpretieren. Geben Sie an, welche Interpretation korrekt ist:

**Antwort:**



**Frage 4:**

Die Produktionsleiterin möchte den linearen Zusammenhang jetzt auch bestimmen. Sie berechnet daher die Gleichung der Regressionsgerade A = a + bT; dabei ist A die Anzahl der nicht einwandfreien Pakete, T die Temperatur.

Bestimmen Sie die Grössen a und b. Runden Sie sie auf ganze Zahlen.

**Antwort:**

Die Antwort lautet:

**Frage 5:**

Mit wie vielen nicht einwandfreien Paketen muss die Produktionsleiterin also bei 20 Grad Celsius rechnen.

**Antwort:**

Variationsquoeffizienz