



Excel:

Zahlen · rechnen · Formeln

René Martin

Für alle Versionen – bis Office 365 (Version 1201)
einschließlich XVERWEIS und dynamische Arrayfunktionen

7.8 Statistik I

Im ersten Teil der Statistik-Funktionen soll es nicht um Statistik gehen, wie sie in der Schule in der Oberstufe oder in einer Reihe von Studiengängen gelehrt wird. Ähnlich wie in Mathematik I werden in diesem Kapitel einige wichtige Funktionen aus der Kategorie Statistik vorgestellt, die in die meisten Berechnungen Einzug gefunden haben.

Tabelle 7.33 Zusammenfassung: bereits erläuterte statistische Funktionen

Funktionsname mit Syntax	Beschreibung	Beispiel
MITTELWERT(Bereich 1; Bereich 2; Bereich 3;...)	berechnet den Durchschnitt	=MITTELWERT (C1:C117)
MIN(Bereich 1; Bereich 2; Bereich 3;...)	berechnet die kleinste Zahl	=MIN (C1:C117)
MAX(Bereich 1; Bereich 2; Bereich 3;...)	berechnet die größte Zahl	=MAX (C1:C117)
ANZAHL(Bereich 1; Bereich 2; Bereich 3;...)	zählt Zahlen in einem Bereich	=ANZAHL (C1:C117)
ANZAHL2(Bereich 1; Bereich 2; Bereich 3;...)	zählt Zahlen und Texte in einem Bereich	=ANZAHL2 (C1:C117)
ZÄHLENWENN ANZAHLLEEREZELLEN ZÄHLENWENN MITTELWERTWENN MITTELWERTWENN	werden ausführlich in Kapitel 7.1.6 beschrieben	

7.8.1 Beispiel: Überblick über eine Liste erhalten

In großen Listen kann man schnell den Überblick verlieren. Einige der statistischen Funktionen geben Aufschluss:

=MAX (G:G)

=MIN (G:G)

=MITTELWERT (G:G)

=ANZAHL (G:G)

=ANZAHLLEEREZELLEN (G:G)

ermitteln den größten und kleinsten Wert, den Durchschnitt, die Anzahl der vorhandenen Zahlen und die Anzahl der leeren Zellen.

KGRÖSSTE und KKLEINSTE

Die zweitgrößte Zahl wird mit

=KGRÖSSTE (G:G;2)

ermittelt. Analog findet

=KKLEINSTE (G:G;2)

die zweitkleinste Zahl. Man könnte eine fortlaufende Reihe bilden, um den zweitgrößten, drittgrößten, viergrößten ... Wert zu ermitteln oder auch die Funktion ZEILE zu Hilfe nehmen.

Hinweis

Auch die Funktion AGGREGAT liefert die kleinste, zweitkleinste, drittkleinste Zahl oder auch die größte, zweitgrößte, ...

MEDIAN

Schreibt man alle Zahlen nach ihrer Größe sortiert nebeneinander, dann liefert

=MEDIAN (G:G)

die Zahl, die sich in der Mitte dieser Liste befindet – normalerweise eine andere als der Mittelwert berechnet.

RANG.GLEICH

Will man nun die Position einer bestimmten Zahl innerhalb der geordneten Liste ermitteln, dann muss man RANG.GLEICH verwenden (die ältere Funktion RANG wird noch unterstützt):

=RANG.GLEICH (G2074;G:G;0)

=RANG.GLEICH (G2074;G:G;1)

Die 0 gibt die Position von der kleinsten Zahl an, 1 (oder eine beliebige andere Zahl) von der größten.

7.8.2 Beispiel: Mittelwert der zehn größten Werte

Aus einer Liste sollen nur die zehn höchsten Werte für einen Durchschnittswert hergenommen werden.

Dürfen bei den zehn größten auch Duplikate eingeschlossen werden, dann könnte die Lösung so aussehen:

```
= (KGRÖSSTE (B2:H9;1) +KGRÖSSTE (B2:H9;2) +KGRÖSSTE (B2:H9;3) +
KGRÖSSTE (B2:H9;4) +KGRÖSSTE (B2:H9;5) +KGRÖSSTE (B2:H9;6) +
KGRÖSSTE (B2:H9;7) +KGRÖSSTE (B2:H9;8) +KGRÖSSTE (B2:H9;9) +
KGRÖSSTE (B2:H9;10) ) /10
```

Mit der Funktion MITTELWERTWENN kann dies sehr viel leichter berechnet werden:

```
=MITTELWERTWENN (J3:J118; ">=" &KGRÖSSTE (J3:J118;10) )
```

Oder Sie schreiben es als Matrixfunktion:

```
{=SUMME (KGRÖSSTE (B2:H9; {1.2.3.4.5.6.7.8.9.10})) /10}
```

Oder Sie verwenden die Funktion AGGREGAT:

```
{=SUMME (AGGREGAT (14;6;B2:H9; {1.2.3.4.5.6.7.8.9.10})) /10}
```

Falls Duplikate nicht erwünscht sind, dann muss man mit zwei Hilfsspalten arbeiten. In der ersten Spalte werden die Werte sortiert:

```
=KGRÖSSTE ($B$2:$H$9; ZEILE () -1)
```

In der zweiten Spalte werden die Duplikate eliminiert:

```
=WENN (J2=MAX (J3:J118) ; "" ; J2)
```

Nun können wieder die zehn Spitzenwerte ermittelt werden und damit der Durchschnitt:

```
= (KGRÖSSTE (K2:K199;1) +KGRÖSSTE (K2:K199;2) +
KGRÖSSTE (K2:K199;3) +KGRÖSSTE (K2:K199;4) +
KGRÖSSTE (K2:K199;5) +KGRÖSSTE (K2:K199;6) +
KGRÖSSTE (K2:K199;7) +KGRÖSSTE (K2:K199;8) +
KGRÖSSTE (K2:K199;9) +KGRÖSSTE (K2:K199;10) ) /10
```

Oder Sie suchen in einer Zelle (beispielsweise in L2 den größten Wert: =MAX(B2:H9). Darunter wird dann der zweitgrößte, drittgrößte, ... Wert gesucht:

```
=KGRÖSSTE ($B$2:$H$9; ZÄHLENWENN ($B$2:$H$9; ">=" &L2) +1)
```

[=SUMME([AGGREGAT(14;6;B2:H9;[1.2.3.4.5.6.7.8.9.10])))]															
B	C	D	E	F	G	H	I	J	K	L	M	N	O	Ohne Duplikate	
4	83	79	11	90	15	90			Mit Duplikaten						
81	31	16	72	86	57	90			860	95	86			95	
50	37	27	9	79	19	42				90				90	
52	10	1	64	42	21	80			860	90				86	
17	81	20	48	21	64	23				90				83	
10	83	41	14	1	49	33				86				81	
63	56	13	69	65	60	95				83				80	
72	81	25	6	79	54	10				83				79	
										81				72	
										81				69	
										81				65	
										80				64	
										79				63	
										79				60	
										79				57	
										72				56	
										72				54	
										69				52	
										65				50	
										64				49	
										64				48	
										63				42	

Abbildung 7.117 Der Mittelwert der zehn größten Werte wird berechnet.

Tabelle 7.34 Zusammenfassung: bereits erläuterte statistische Funktionen

Funktionsname	Beschreibung	Beispiel
RANG.GLEICH(Zahl; Bezug; [Reihenfolge])	berechnet die Position eines Wertes innerhalb einer Liste (bei mehrmaligem Vorkommen – nimm die erste Zahl)	=RANG.GLEICH(A1;A:A;0)
RANG.MITTELW(Zahl; Bezug; [Reihenfolge])	berechnet die Position eines Wertes innerhalb einer Liste (bei mehrmaligem Vorkommen – nimm die mittlere Zahl)	=RANG.MITTELW(A1;A:A;0)
KKLEINSTE(Bereich;Position)	ermittelt den kleinsten, zweitkleinsten, drittkleinsten ... Wert einer Liste	=KKLEINSTE(A:A;2)
KGRÖSSTE(Bereich;Position)	ermittelt den kleinsten, zweitkleinsten, drittkleinsten ... Wert einer Liste	=KGRÖSSTE(A:A;2)
MEDIAN(Bereich)	ermittelt den „mittleren“ Wert eines Bereiches	=MEDIAN(A:A)

Ist allerdings bei der Potenzfunktion die Basis gesucht, so muss die Funktion LOG verwendet werden. Die folgende Funktion liefert 3:

`=LOG (125; 5)`

Der natürliche Logarithmus

Das Ganze funktioniert auch mit dem natürlichen Logarithmus zur Basis e (2,718281828).

`=EXP (A1)`

potenziert die Zahl, die in A1 steht, mit der Basis e , und

`=LN (A1)`

liefert den natürlichen Logarithmus einer Zahl. Den Logarithmus zur Basis 10 können Sie auch mit

`=LOG10 (A1)`

ermitteln. Beispielsweise ergibt

`=LOG10 (1000000)`

die Zahl 6, da $10^6 = 1000000$.

9.2 Statistische Funktionen

9.2.1 Die Mitte

Die Aufgaben der statistischen Funktionen werden am besten an einem Beispiel deutlich. Lassen Sie uns ein solches konstruieren. Frau Lämpel ist Lehrerin in einer Klasse, in der sie Englisch unterrichtet. Sie hat 24 Schüler und Schülerinnen, die in der letzten Klassenarbeit folgende Noten erzielt haben. Die Noten stehen in einem Tabellenblatt in den Zellen A1:H7.

Tabelle 9.1 Bei der letzten Klassenarbeit hatten die Schüler folgende Noten.

Name	Note	Name	Note	Name	Note	Name	Note
Anton	1,00	Gerda	3,25	Martha	4,50	Traudl	4,25

Name	Note	Name	Note	Name	Note	Name	Note
Bert	4,00	Hildegard	2,75	Norbert	2,00	Ulla	5,25
Conny	2,50	Irmgard	3,50	Otto	5,50	Victoria	5,00
Det	5,00	Jürgen	2,75	Pedro	4,00	Walter	3,75
Ede	2,00	Karl	5,25	René	1,00	Xaver	3,50
François	1,75	Lotte	5,25	Stefan	2,25	Yasar	4,00

Das arithmetische Mittel

Um den Durchschnitt, das heißt das arithmetische Mittel, zu berechnen, kann die Funktion MITTELWERT verwendet werden:

=MITTELWERT (B2 : B7 ; D2 : D7 ; F2 : F7 ; H2 : H7)

GESTUTZTMITTEL

Sollen Ränder außer Acht gelassen werden, so ist die Funktion GESTUTZTMITTEL zu verwenden. Wird beispielsweise davon ausgegangen, dass von den 24 Schülern immer zwei Ausreißer sind (also $2/24 = 8,33333\%$), dann liefert

=GESTUTZTMITTEL (B2 : B25 ; 8,33333%)

den Wert 3,522727. Die beiden „Ausreißer“ sind Anton (oder René) mit der Note 1 und Otto mit 5,5. Somit verschiebt sich der Durchschnitt nach oben.

Der Durchschnitt liegt bei 3,5. Was aber, wenn Frau Lämpel die Noten sortiert? Das heißt, wenn sie von folgender Liste ausgeht.

Tabelle 9.2 Die Notenliste kumuliert

Note	Schüler		Note	Schüler
1	2		3,75	1
1,75	1		4	3
2	2		4,25	1
2,25	1		4,5	1
2,5	1		5	2
2,75	2		5,25	3
3,25	1		5,5	1
3,5	2			

Diese Liste kann übrigens mit der Funktion ZÄHLENWENN erzeugt werden. Nun muss der Durchschnitt mit der folgenden Funktion berechnet werden:

```
=SUMMENPRODUKT (H21:H35;I21:I35) /SUMME (I21:I35)
```

Modalwert

Das heißt, die Anzahl der Noten wird mit den Noten multipliziert und das Ergebnis addiert.

Was aber, wenn Frau Lämpel nun wissen möchte, welcher Wert am häufigsten auftaucht? Nun, sie könnte diese Tabelle nach der Spalte „Schüler“ sortieren lassen und erkennt sofort, dass sowohl die 4 als auch die 5,25 am häufigsten vorkommt. Die statistische Funktion, die dies ermittelt, lautet

```
=MODUS.EINF (B2:B7;D2:D7;F2:F7;H2:H7)
```

Da mehrere Werte in der Mitte liegen können existiert die Funktion MODUS.VIELF Sie ist eine Matrixfunktion und muss mit [Shift]+[Strg]+[Enter] beendet werden:

```
{=MODUS.VIELF (B2:B7;D2:D7;F2:F7;H2:H7) }
```

So erhält man 4; 5,25; #NV; #NV; ...

Diese beiden Funktionen ersetzen die „alte“ Funktion MODALWERT, die der Funktion MODUS.EINF entspricht:

```
=MODALWERT (B2:B7;D2:D7;F2:F7;H2:H7)
```

Sie liefert 4.

Median

Angenommen Frau Lämpel würde ihre Schüler nun der Notenwertung nach aufstellen und sich ansehen, welcher ihrer Schüler in der Mitte stünde, so könnte sie dies mit der Funktion MEDIAN berechnen. Eine Aufstellung ergibt, dass Irmgard als 12. Schülerin eine 3,5 erreicht und Walter eine 3,75. Der „mittlere“ Schüler liegt also dazwischen. Excel nimmt nun das arithmetische Mittel aus beiden Werten und berechnet den Median mit dem Wert 3,625:

```
=MEDIAN (H2:H7;F2:F7;D2:D7;B2:B7)
```

Quartil

Möchte Frau Lämpel nicht nur den „mittleren Schüler“ ermitteln, sondern auch den Schüler, der in der ersten Hälfte oder in der zweiten Hälfte in der Mitte liegt, so verwendet sie

die Funktion `QUARTILE.EXKL` oder `QUARTILE.INKL`. Sie liefern für die Quartile 1 (also 25%) die Werte 2,3125, beziehungsweise 2,4375 zurück. Der Unterschied: Stellt man die 24 sortiert der Größe nach auf zwischen der sechsten und der siebten Note, also zwischen 2,25 und 2,5. `QUARTILE.EXKL` berechnet: $2,25 + \frac{1}{4} \cdot (2,5 - 2,25)$, während `QUARTILE.INKL` berechnet: $2,25 + \frac{3}{4} \cdot (2,5 - 2,25)$. Oder wie die Hilfe es beschreibt: „basierend auf den Perzentilwerten 0 – 1 einschließlich oder ausschließlich. Die drei übrigen Parameter, also `QUARTILE.EXKL` (und auch `INKL`) entsprechen den Funktion `MIN`, `MAX` und `MEDIAN`. Die Funktion `QUARTILE.INKL` ersetzt die alte Funktion `QUARTILE`

Der Rang, der Drittgrösste und Drittkleinste

Möchte Frau Lämpel nun wissen, an welcher Stelle Ede mit seiner Note 2 steht, so kann sie die Funktion `RANG.GLEICH` oder `RANG.MITTELW` verwenden. Der Parameter „Reihenfolge“ liefert dabei die Position ausgehend vom Besten, das heißt kleinsten Wert (1) oder vom größten Wert (0). Diese beiden Funktionen ersetzen die alte Funktion `RANG`. Ede liegt in dieser Klassenarbeit auf

```
=RANG.GLEICH(2;A2:H7;1)
```

das heißt auf Rang 4 (nur Anton, Frieda und René sind besser) oder auf

```
=RANG.GLEICH(2;A2:H7;0)
```

das heißt auf Platz 20 von hinten.

Oder auch

```
=RANG.MITTELW(B11;A1:H7)
```

liefert 20,5, während

```
=RANG.MITTELW(B11;A1:H7;1)
```

4,5 ergibt. Der Unterschied erklärt sich damit, dass Ede mit seiner Note 2 diesen Platz auch mit Norbert teilt, der ebenfalls eine 2 geschrieben hat.

Umgekehrt liefern

```
=KKLEINSTE(A2:H7;4)
```

```
=KGRÖSSTE(A2:H7;21)
```

den viertkleinsten und den einundzwanziggrößten Platz, also 2. Und natürlich:

```
=KKLEINSTE(A2:H7;21)
```

=KGRÖSSTE (A2:H7; 4)

den viertschlechtesten Schüler, also denjenigen, der eine 5– (5,25) geschrieben hat: Karl, Lotte und Ulla sind nicht ganz so schlecht wie Otto.

Das geometrische Mittel

Eine weitere Funktion zur Berechnung der Mitte ist das geometrische Mittel, das in diesem Beispiel keinen Sinn macht, da es sich bei der geometrischen Mitte um Wachstumsangaben handelt und nicht um verteilte Werte. Würde man ihn dennoch berechnen, so ergibt

=GEOMITTEL (H2:H7; F2:F7; D2:D7; B2:B7)

den Wert 3,17276. Es ist die 24. Wurzel aus dem Produkt der Noten und entspricht damit folgender Funktion:

=POTENZ (PRODUKT (H2:H7; F2:F7; D2:D7; B2:B7) ; 1/24)

Das harmonische Mittel

Auch das harmonische Mittel ist hier fehl am Platz. Bei ihm werden die Kehrwerte der Zahlen addiert und durch die Anzahl geteilt. Das Ergebnis von

=HARMITTEL (B2:H7)

ist 2,77466. Soll nun ein Schwellenwert festgelegt werden, ab dem Beobachtungen akzeptiert werden, so kann die Funktion QUANTIL verwendet werden. Soll beispielsweise der Schwellenwert bei 10 % liegen, so ergibt das Quantil:

=QUANTIL (A2:H7; 0,1)

den Wert 1,825. Umgekehrt liefert

=QUANTILSRANG (B2:B25; 1,825)

den Wert 10 % (oder 0,1).

9.2.2 Abweichung

Auch für verschiedene Abweichungen stellt Excel eine Reihe von Funktionen zur Verfügung. Die größte und kleinste Zahl, das Maximum und Minimum wird mit den beiden Funktionen MAX und MIN ermittelt. Sie liefern 1 und 5,5.

Die Spannweite

Übrigens kann man statt dieser beiden Funktionen auch die Tabelle auf- oder absteigend sortieren. So erfährt man auch den kleinsten und den größten Wert. Nur hat das Sortieren zwei Nachteile: Es verändert die ursprüngliche Reihenfolge und bei der Eingabe neuer Datensätze muss erneut sortiert werden.

Die Differenz zwischen den beiden Zahlen beträgt 4,5; sie wird auch Spannweite genannt.

Die Mittelabweichung

Berechnet man nun die Differenz einer jeden Zahl vom Mittelwert, bildet daraus den Absolutwert, addiert diese und teilt sie durch die Anzahl, so erhält man die Mittelabweichung. Sie könnte auch mit der Funktion

=MITTELABW(B2:B7;D2:D7;F2:F7;H2:H7)

ermittelt werden. Das Ergebnis beträgt 1,1458333.

Tabelle 9.3 In der zweiten Klassenarbeit erhalten die Schüler andere Noten.

Name	Note	Name	Note	Name	Note	Name	Note
Anton	3,25	Gerda	4,00	Martha	3,75	Traudl	3,00
Bert	3,00	Hildegard	4,00	Norbert	3,25	Ulla	4,25
Conny	2,75	Irmgard	3,25	Otto	4,00	Victoria	4,00
Det	4,00	Jürgen	3,75	Pedro	4,00	Walter	4,50
Ede	3,75	Karl	4,00	René	1,75	Xaver	4,00
François	3,75	Lotte	2,75	Stefan	2,25	Yasar	3,00

Varianz und Standardabweichung

Der Durchschnitt liegt auch hier bei 3,5. Dagegen ist das Maximum 4,5, das Minimum 1,75 – die Spannweite beträgt lediglich 2,75, aber die mittlere Abweichung nur 0,5625, das heißt: Die Spannweite ist weniger deutlich ausgeprägt als bei der ersten Arbeit.

Ein besseres Maß für die Abweichung von der Mitte ist die Standardabweichung: Dabei wird die Differenz zum Mittelwert quadriert, diese wird addiert und durch die Anzahl geteilt. Dieses Maß wird Varianzen genannt. Zieht man aus ihm die Wurzel, so erhält man die Standardabweichung. Beide Funktionen finden sich in der Kategorie „Statistik“:

=VARIANZENA(F10:M15)

=STABWNA(F10:M15)

Sie liefern bei der ersten Klassenarbeit mit einer breiteren Streuung die Zahlen 1,828 und 1,352, bei der zweiten Klassenarbeit 0,4375 und 0,6614. In der zweiten Arbeit gruppieren sich die Werte also viel stärker um die arithmetische Mitte.

Achtung

Will man mit Standardabweichungen und Varianzen für Tendenzen arbeiten, so empfiehlt es sich, nicht durch die Anzahl, sondern durch die (Anzahl – 1) zu dividieren. Excel stellt für beide Aufgabenstellungen eine Lösung zur Verfügung:

```
=VARIANZA (G10:M15)
```

```
=STABWA (G10:M15) (
```

Sie liefern die Werte 1,9076 und 1,38116 beziehungsweise 0,4565 und 0,6757.

Diese Funktionen ersetzen die alten Funktionen VARIANZEN und VARIANZ, beziehungsweise STABWN und STABW. Außerdem stehen Ihnen noch die Funktionen STABW.N und STABW.S zur Verfügung. Sie ignorieren Texte und Wahrheitswerte.

Ein weiteres Maß für Abweichungen von Datenpunkten von deren Mittelpunkt ist die Möglichkeit, die Differenz zwischen den Datenpunkten und dem Mittelpunkt zu bilden und ihre Quadrate zu summieren. SUMQUADABW erledigt diese Aufgabe.

In diesem Zusammenhang sei noch eine weitere Funktion erwähnt:

Mit Mittelwert und Standardabweichung kann man nun die Verteilung standardisieren. Im ersten Notenbeispiel war der Durchschnitt 3,5, die Standardabweichung betrug 1,352. Daraus resultiert der zur Note 3 standardisierte Wert:

```
=STANDARDISIERUNG (3; 3,5; 1,352)
```

ergibt -0,3698.

9.2.3 Korrelationen

Frau Lämpel vermutet einen Zusammenhang zwischen den Noten, die ihre Schüler geschrieben haben, und deren Körpergröße. Sie gibt sie ein.

Tabelle 9.4 Noten und Körpergröße:

Name	Note	Körpergröße in cm	Name	Note	Körpergröße in cm
Anton	1,00	176,00	Martha	4,50	130,00
Bert	4,00	180,00	Norbert	2,00	181,00

Name	Note	Körpergröße in cm	Name	Note	Körpergröße in cm
Conny	2,50	155,00	Otto	5,50	155,00
Det	5,00	171,00	Pedro	4,00	165,00
Ede	2,00	154,00	René	1,00	180,00
François	1,75	170,00	Stefan	2,25	130,00
Gerda	3,25	184,00	Traudl	4,25	162,00
Hildegard	2,75	190,00	Ulla	5,25	162,00
Irmgard	3,50	185,00	Victoria	5,00	190,00
Jürgen	2,75	139,00	Walter	3,75	166,00
Karl	5,25	159,00	Xaver	3,50	143,00
Lotte	5,25	162,00	Yasar	4,00	120,00

Der Korrelationskoeffizient

Der Korrelationskoeffizient berechnet sich für die erste Spalte mit

=KORREL (B2 : B13 ; C2 : C13)

und liefert den Wert -0,074. Für die zweite Spalte ergibt sich -0,122. Der Korrelationskoeffizient bewegt sich zwischen den beiden Zahlen -1 und 1. Liegt der Wert bei -1, so stehen zwei Zahlenreihen in einem umgekehrten Verhältnis zueinander, bei 1 gibt es eine direkte Abhängigkeit. Da in Frau Lämpels Beispiel die Werte um 0 herum liegen, gibt es keinen Zusammenhang zwischen Körpergröße und Schulnote.

Übrigens benötigt man zur Berechnung des Korrelationskoeffizienten die Kovarianz (KOVARIANZ). Sie ist der Mittelwert der für alle Datenpunkte gebildeten Produkte der Abweichungen.

9.2.4 Trends

Frau Lämpel sieht ein, dass es keinen Zusammenhang zwischen diesen beiden Größen gibt, aber sie stellt fest, dass die Noten der Klassenarbeiten immer besser werden. Sie errechnet folgende (Mittel-)Werte.

Tabelle 9.5 Die Mittelwerte der Klassenarbeiten

Termin der Arbeit	Nummer	Klassendurchschnitt
01.10.2019	1	3,50

Termin der Arbeit	Nummer	Klassendurchschnitt
02.11.2019	2	3,50
04.12.2019	3	3,20
05.01.2020	4	3,20
06.02.2020	5	3,10
10.03.2020	6	3,00
11.04.2020	7	3,10
13.05.2020	8	2,80
14.06.2020	9	2,70
16.07.2020	10	

Sie trägt diese in ein Diagramm ein:

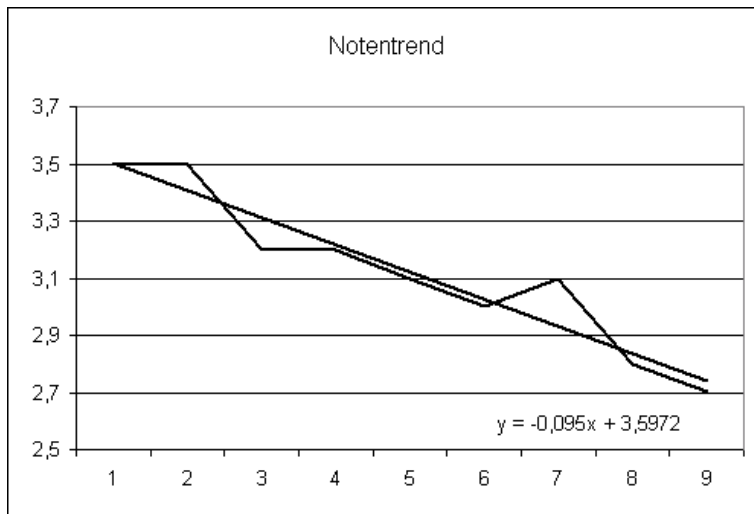


Abbildung 9.10 Die Linie wurde mit der Option „Trend“ eingeschaltet.

Wenn Frau Lämpel nun wissen möchte, wie die nächste Arbeit voraussichtlich ausfällt, so markiert sie die Spalte neben den Noten bis zur zehnten Arbeit. Sie wählt aus dem Funktionsassistenten die Funktion TREND aus und markiert bei y-Werte die vorhandenen neun Noten, bei NEUE_X_WERTE die Reihe der Zahlen 1 bis 10. Da es sich hier um eine Matrixfunktion handelt, muss die Tastenkombination [Shift] + [Strg] + [Enter] gedrückt werden.

Tabelle 9.6 Das Ergebnis der Trendberechnung

Termin der Arbeit	Nummer	Klassendurchschnitt	Trend
01.10.2019	1	3,50	3,50
02.11.2019	2	3,50	3,41
04.12.2019	3	3,20	3,31
05.01.2020	4	3,20	3,22
06.02.2020	5	3,10	3,12
10.03.2020	6	3,00	3,03
11.04.2020	7	3,10	2,93
13.05.2020	8	2,80	2,84
14.06.2020	9	2,70	2,74
16.07.2020	10		2,65

Zu diesem Ergebnis gelangen Sie auch mit der Funktion `PROGNOSE.LINEAR`.



Achtung

Beachten Sie, dass `PROGNOSE.LINEAR` die Parameter in der Reihenfolge $x / y_Werte / x_Werte$ haben möchte; dagegen die Funktion `TREND` in der Reihenfolge: $y_Werte / x_Werte / x$.

Die Trendachse schneidet die y-Achse und hat eine bestimmte Steigung. Die beiden Werte können mit den Funktionen

`=ACHSENABSCHNITT(D2:D11;B2:B11)` und

`=STEIGUNG(D2:D11;B2:B11)`

berechnet werden. Sie liefern 3,5972 und -0,095 (also eine sehr flache, negative Steigung). Dies kann auch in einem Diagramm dargestellt werden. Man markiert die (vorhandenen) x- und y-Werte (hier die Zahlen 1 bis 9 und 3,5 bis 2,7) und lässt sich dazu ein Diagramm erstellen. Über das Kontextmenü „Trendlinie Hinzufügen“ kann zusätzlich eine Trendlinie und deren Funktion eingefügt werden.

Ist nur ein Trendwert gesucht, so kann auch die Funktion `SCHÄTZER` verwendet werden. Vermutet man zwischen Werten, denen kein Trend zu Grunde liegt, einen linearen Zusammenhang, dann kann die Funktion `RGP` verwendet werden, um eine Matrix zu berechnen, in der die Elemente liegen, die eine Gerade bilden.

Variation

Geht Frau Lämpel allerdings von einer exponentiellen Steigerung aus, so müsste sie statt TREND die Funktion VARIATION verwenden. Damit sieht das Bild etwas anders aus.

Tabelle 9.7 Das Ergebnis der exponentiellen Trendberechnung mit VARIATION

Termin der Arbeit	Nummer	Klassendurchschnitt	Trend
01.10.2019	1	3,50	3,52
02.11.2019	2	3,50	3,41
04.12.2019	3	3,20	3,31
05.01.2020	4	3,20	3,21
06.02.2020	5	3,10	3,11
10.03.2020	6	3,00	3,02
11.04.2020	7	3,10	2,93
13.05.2020	8	2,80	2,84
14.06.2020	9	2,70	2,75
16.07.2020	10		2,67

Vermutet man zwischen Werten, denen kein Trend zu Grunde liegt, einen exponentiellen Zusammenhang, dann kann die Funktion RKP verwendet werden, um eine Matrix zu berechnen, in der die Elemente liegen, die eine solche Kurve bilden.

9.2.5 Wahrscheinlichkeiten

Binominalverteilte Zufallsvariable

Frau Lämpel hat während der vielen Klassenarbeiten festgestellt, dass ihre Schüler mit einer Wahrscheinlichkeit von 20 % spicken, schummeln, mogeln und abschreiben. Wie groß, so fragt sie sich, ist nun die Wahrscheinlichkeit, dass sie zwei Schüler beim Betrügen erwischen wird, wenn sie zehn Schüler überprüft? Hierfür kann die Funktion der Binomialverteilung (oder Bernouilli-Verteilung) helfen. Die Zahl der Erfolge sind 2, die Versuche 10 bei einer Wahrscheinlichkeit von 0,2. Da die Daten nicht kumuliert werden, wird dort die Zahl 0 eingegeben.

=BINOM.VERT(2;10;0,2;0)

So erhält sie das Ergebnis 0,3019898 oder 30 %, das heißt, mit einer Wahrscheinlichkeit von 1 zu 2 spicken zwei von zehn Schülern bei einer Arbeit. Daneben finden Sie noch die alte Funktion BINOMVERT.

Normal verteilte Zufallsvariable

Der Durchschnitt ihrer ersten Arbeit liegt bei 3,5 und einer Standardabweichung von $\sigma = 1,352$. Wie groß ist nun die Wahrscheinlichkeit, dass einer ihrer Schüler, der zufällig ausgewählt wurde, eine 4 oder schlechter hat? Die Normalverteilung (oder Gauß'sche Verteilung) beträgt für „x“ = 4, „Mittelwert“ = 3,5, „Standabwn“ = 1,352 und „Kumuliert“ = 1 die Formel:

=NORM.VERT (4; 3,5; 1,352; 1)

oder den Wert 0,64424, das heißt mehr als 64 %. Noch enthalten ist die alte Funktion NORMVERT.

Negativbinomialverteilten Zufallsvariablen

Übrigens könnte mit der Funktion NEGBINOM.VERT (früher: NEGBINOMVERT) die Wahrscheinlichkeit einer negativbinomialverteilten Zufallsvariablen berechnet werden, das heißt, man kann berechnen, wie wahrscheinlich es ist, dass es eine gewisse Anzahl von Misserfolgen gibt, bis ein positives Ereignis n-mal eintritt.

Hypergeometrisch-verteilte Zufallsvariable

Noch ein Beispiel: Die Noten von 1 bis 6 lassen sich auf einer Skala abtragen, so dass sich 21 verschiedene Notenwerte ergeben (1, 1-, 1-2, ... 5-6, 6+ und 6). Acht dieser 21 Noten sind schlechter oder gleich schlecht als 4. Wie hoch, so fragt sich Frau Lämpel, ist nun die Wahrscheinlichkeit, wenn sie zufällig drei Schüler auswählt, dass sie zwei schlechte und einen guten erwischt? Hier hilft die hypergeometrische Verteilung. Die Erfolge der Stichprobe betragen 2 (zwei schlechte Schüler), die Größe der Stichprobe ergibt 3 (drei ausgewählte Schüler). Die Erfolge der Grundgesamtheit betragen 8 (es gibt acht Noten {4-, 4-5, 5+, 5-, 5-6, 6+ und 6} und der Umfang der Grundgesamtheit ergibt 21 (mögliche Noten). Somit erhält

=HYPGEOM.VERT (2; 3; 8; 21)

den Wert 0,27368, das entspricht mehr als 27 %. In älteren Excel-Versionen hieß die Funktion HYPGEOMVERT.

Die Poisson-Verteilung

Eine Annäherung an die Binomialverteilung stellt die Poisson-Verteilung dar. Sie liefert die Grenzverteilung zur Binomialverteilung.

Beispiel: Die Wahrscheinlichkeit, dass einer ihrer Schüler am Tage vor einer Klassenarbeit vom Elternhaus wegläuft, beträgt 2 %. Wie groß, will Frau Lämpel wissen, ist die Wahrscheinlichkeit, dass von 1000 Schülern, die pro Jahr eine Arbeit schreiben, vier von zu Hause weglaufen? Dazu verwendet sie die Poisson-Verteilung, wobei für x die Zahl vier einzugeben ist. Der Mittelwert beträgt $1000 \cdot 0,002$, und da nicht kumuliert wird, ist in dieses Feld der Wert 0 einzugeben. Die Formel sieht wie folgt aus:

`=POISSON.VERT(4;1000*0,002;0)`

Das Ergebnis lautet: 0,09, also fast 10 %. Sie ersetzt die alte Funktion POISSON.

In diesem Zusammenhang sollen noch drei weitere Funktionen erwähnt werden:

Soll der Standardfehler der geschätzten y-Werte für alle x-Werte einer Regression berechnet werden, dann kann die Funktion STFEHLERYX herangezogen werden.

Die Wahrscheinlichkeit für einen Wert, der in einem Intervall von zwei Werten liegt, wird mit der Funktion WAHRSCHBEREICH berechnet.

Mit der Funktion WEIBULL.VERT (früher: WEIHBULL) werden die Wahrscheinlichkeiten einer weibullverteilten Zufallsvariablen berechnet.

Binomialverteilung

Wenn Sie die Kombinationen kumulieren (mit der Funktion HÄUFIGKEIT oder ZÄHLENWENN), erhalten Sie eine binomialverteilte Datenreihe, die man leicht mit einem Diagramm visualisieren kann:

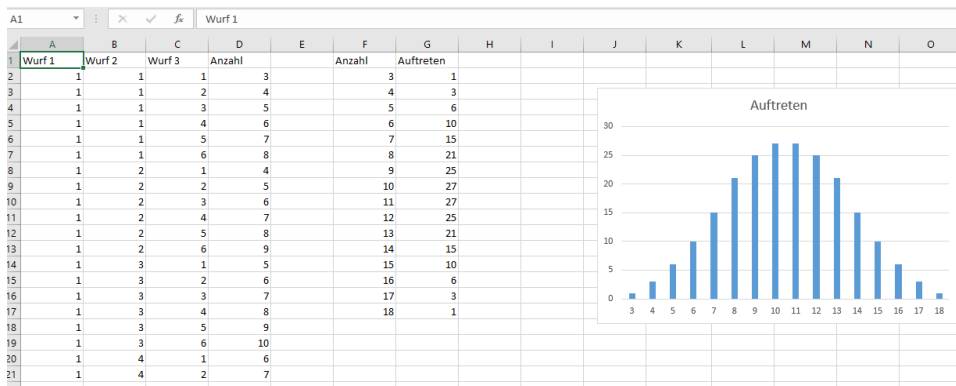


Abbildung 9.11 Binomialverteilung

Carl Friedrich Gauss hat – ausgehend von der Binomialverteilung – für die Normalverteilung eine Formel entwickelt:

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

Sie wurde auf dem alten 10-DM-Schein dargestellt und so seine Leistung gewürdigt:



Abbildung 9.12 Die Normalverteilung nach Gauss

Man kann sie in Excel leicht nachbauen:

$$=1/(\$E\$2*WURZEL(2*PI()))*EXP(-0,5*((A2-\$G\$2)/\$E\$2)^2)$$

oder natürlich die Funktion

$$=NORM.VERT(A2;\$G\$2;\$E\$2;FALSCH)$$

verwenden. Mit einem Graphen kann man diese Glockenform darstellen:

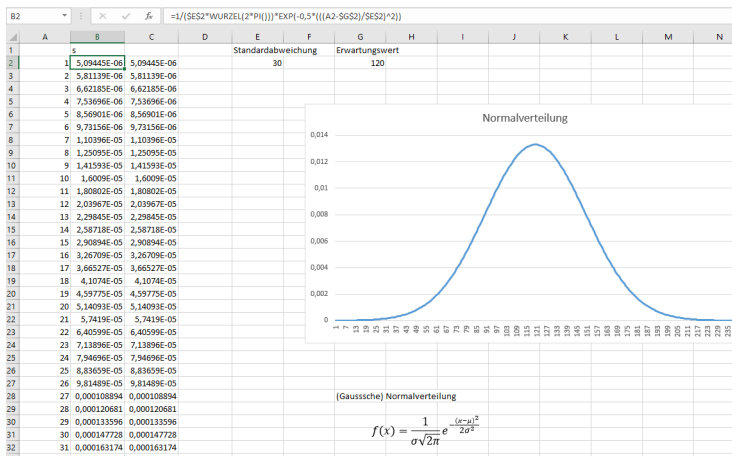


Abbildung 9.13 Die Normalverteilung

Ändert man nun Standardabweichung und Erwartungswert, verändert die Kurve ihre Form:

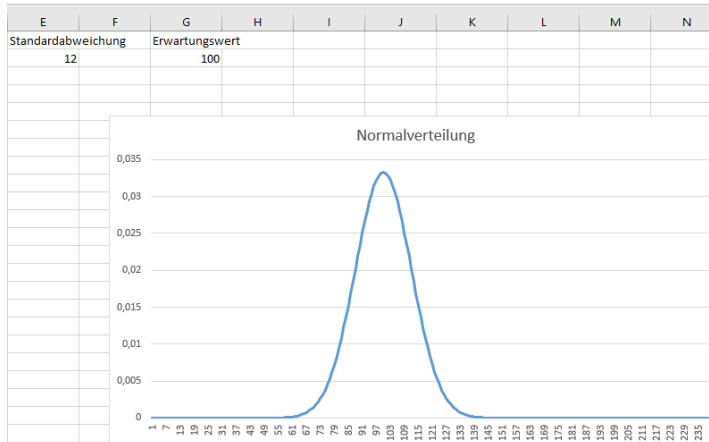


Abbildung 9.14 Die Normalverteilung mit anderen Werten für σ und μ

9.2.6 Konfidenzintervall

In Frau Lämpels erster Klassenarbeit schreiben 24 Schüler mit. Der Klassendurchschnitt liegt bei 3,5, die Standardabweichung bei 1,352. Bei 1000 Schülern ihrer Schule, so schätzt sie, liegt die Wahrscheinlichkeitsangabe (oder das Konfidenzintervall) bei 90 %, das heißt, das Vertrauensniveau wird in 10 % der Fälle nicht erreicht. Mit diesen Werten kann sie in Excel mit der KONFIDENZ das Vertrauensintervall berechnen. Alpha entspricht 10 %, die Standardabweichung liegt bei 1,352 und der Umfang der Stichprobe bei 24. Das ergibt:

`=KONFIDENZ.NORM(10%;1,352;24)`

Es liefert den Wert 0,4539, also etwas weniger als die Hälfte. Früher lautet diese Funktion KONFIDENZ. Bei einer Student t-verteilten Zufallsvariablen verwenden Sie die Funktion KONFIDENZ.T[^].

9.2.7 Tests und Verteilungen

An dieser Stelle alle Tests und Verteilungen aufzuführen und zu erläutern, würde den Rahmen dieses Buches sprengen. Deshalb soll an einem Beispiel erläutert werden, was ein solches Testverfahren leistet, wie es anzuwenden ist und welche Zahlen in die Excel-Funktionen einzugeben sind.

Frau Lämpels Klasse besteht aus Jungen und Mädchen, deren Klassenarbeiten benotet werden. Da sie weniger an den Viertelnoten (also 2+ oder 3-) interessiert ist, stellt sie eine Liste auf, wie viele Jungen und wie viele Mädchen welche Noten (in der ersten Arbeit) geschrieben haben.

Tabelle 9.8 Die Schüler nach Geschlecht und Noten aufgelistet

Name	Mädchen	Jungen
Anton	1	
Bert		4
Conny	2,5	
Det		5
Ede		2
François	1,75	
Gerda	3,25	
Hildegard	2,75	
Irmgard	3,5	
Jürgen		2,75
Karl		5,25
Lotte	5,25	
Martha	4,5	
Norbert		2
Otto		5,5
Pedro		4
René	1	
Stefan		2,25
Traudl	4,25	
Ulla	5,25	
Victoria	5	
Walter		3,75
Xaver		3,5
Yasar		4

Mit Hilfe der Funktion

=WENN (ODER (B11-GANZZAHL (B11)=0,25;B11-GANZZAHL (B11)=0,75) ;
RUNDEN (B11;0) ;WENN (B11="";"";B11))

werden die Noten gerundet.

Tabelle 9.9 Die gerundeten Noten sehen wie folgt aus:

Name	Mädchen	Jungen
Anton		1
Bert		4
Conny	2,5	
Det		5
Ede		2
François		2
Gerda	3	
Hildegard	3	
Irmgard		3,5
Jürgen		3
Karl		5
Lotte	5	
Martha	4,5	
Norbert		2
Otto		5,5
Pedro		4
René	1	
Stefan		2
Traudl		4
Ulla	5	
Victoria	5	
Walter		4
Xaver		3,5
Yasar		4

Mit der Funktion HÄUFIGKEIT lässt Frau Lämpel die Notenverteilung anzeigen. Dazu ist im Bereich „Daten“ die Matrix einzugeben, die gesucht ist, also in unserem Fall die

Notenwerte von 1 bis 6. Mit Klasse ist die Notenliste gemeint (einmal für die Mädchen, einmal für die Jungs). Da es sich um eine Matrixformel handelt, ist sie mit [Shift] + [Strg] + [Enter] zu beenden. Sie hat beispielsweise folgende Gestalt:

```
{=HÄUFIGKEIT (F2:F25;$I2:$I12) }
```

Tabelle 9.10 Die Noten werden nach ihrer Häufigkeit aufgelistet.

Noten:	Mädchen	Jungen
1	0	2
1,5	0	0
2	0	4
2,5	1	0
3	2	1
3,5	1	1
4	1	4
4,5	1	0
5	4	1
5,5	0	1
6	0	0

Dies sind die realen Daten der ersten Klassenarbeit. Auch wenn der Notendurchschnitt der Mädchen etwas besser ist als der Schnitt der Jungen, so soll gezeigt werden, dass es keinen Zusammenhang zwischen Geschlecht und Noten gibt. Dazu werden die Summen der Jungs und der Mädchen einerseits und der Notenkategorien andererseits gebildet.

Tabelle 9.11 Die Werte werden addiert:

Noten	Mädchen	Jungen	Summe
1	0	2	2
1,5	0	0	0
2	0	4	4
2,5	1	0	1
3	2	1	3
3,5	1	1	2
4	1	4	5
4,5	1	0	1

Noten	Mädchen	Jungen	Summe
5	4	1	5
5,5	0	1	1
6	0	0	0
Summe	10	14	24

Das Verhältnis der Mädchen zur Gesamtzahl der Schüler der Klasse beträgt 10/24. Die theoretisch zu erwartenden Häufigkeiten liegen bei den Mädchen also bei $2 \cdot 10/24$, bei den Jungen bei $2 \cdot 14/24$, also bei den Werten 0,8333 und 1,1666.

Um diese Formel schnell von links nach rechts (von den Mädchen zu den Jungen) und von oben nach unten (über die gesamten Notenwerte) zu ziehen, kann mit relativen, gemischten und absoluten Bezügen gearbeitet werden:

=10/24*2 entspricht:

$$=J\$13/\$L\$13*\$L2$$

Tabelle 9.12 Damit ergeben sich folgende Häufigkeitswerte:

Noten	Mädchen	Jungen	Summe der Häufigkeiten
1	0,83333333	1,16666667	2
1,5	0	0	0
2	1,66666667	2,33333333	4
2,5	0,41666667	0,58333333	1
3	1,25	1,75	3
3,5	0,83333333	1,16666667	2
4	2,08333333	2,91666667	5
4,5	0,41666667	0,58333333	1
5	2,08333333	2,91666667	5
5,5	0,41666667	0,58333333	1
6	0	0	0
	10	14	

Wenn also die bei Unabhängigkeit zu erwartenden Werte mit den beobachteten Häufigkeiten übereinstimmen, kann die Hypothese von der Unabhängigkeit als bestätigt angesehen werden. Je weiter sie aber voneinander abweichen, desto größer ist die Abweichung

von der Unabhängigkeit, das heißt, die beiden Untersuchungsvariablen sind als voneinander abhängig anzusehen.

Um vernünftig weiterrechnen zu können, muss die Zeile mit der Note 6 und die mit der Note 1,5 gelöscht werden, da diese Noten nicht auftauchen.

Tabelle 9.13 Werte, die nicht verwendet werden, werden gelöscht.

Noten	Mädchen	Jungen	Häufigkeit
1	0,83333333	1,16666667	2
2	1,66666667	2,33333333	4
2,5	0,41666667	0,58333333	1
3	1,25	1,75	3
3,5	0,83333333	1,16666667	2
4	2,08333333	2,91666667	5
4,5	0,41666667	0,58333333	1
5	2,08333333	2,91666667	5
5,5	0,41666667	0,58333333	1
	10	14	24

Bildet man die Differenz zwischen den erwarteten Werten und den gemessenen Werten und quadriert sie (um aus negativen Werten positive zu erzeugen und um eine bekannte Verteilung zu erhalten), teilt sie anschließend durch die erwarteten Werte und addiert diese Werte zum Schluss auf, so erhält man in obigem Beispiel die Zahl 7,88. Je größer diese Zahl von 0 abweicht, umso eher kann die ursprüngliche Hypothese verworfen werden, je kleiner, umso eher stimmt sie.

Tabelle 9.14 Die Erwartungswerte

Noten	Mädchen	Jungen
1	0,83333333	0,5952381
2	1,66666667	1,1904762
2,5	0,81666667	0,5833333
3	0,45	0,3214286
3,5	0,03333333	0,0238095
4	0,56333333	0,402381

Noten	Mädchen	Jungen
4,5	0,81666667	0,58333333
5	1,76333333	1,2595238
5,5	0,41666667	0,297619
		12,6171

Ein Wort zur Funktion HÄUFIGKEIT. Die Syntax dieser Funktion lautet:

=HÄUFIGKEIT(Daten;Klassen)

Daten sind die Zellen, in denen die Werte stehen, die gezählt werden. Mit Klassen sind die Gruppen gemeint. Sie können aufsteigen oder absteigend sortiert werden.

Da HÄUFIGKEIT als Matrixfunktion arbeitet, müssen sämtliche Zielwerte markiert werden und die Funktion anschließend mit [Shift] + [Strg] + [Enter] abgeschlossen werden.

{=HÄUFIGKEIT(A\$1:D\$14;F\$1:F\$14)}

{=HÄUFIGKEIT(F2:F25;\$I2:\$I12)}						
E	F	G	H	I	J	K
	Mädchen	Jungen		Häufigkeit:	Mädchen	Jungen
Anton			1	1	0	2
Bert			4	1,5	0	0
Conny	2,5			2	0	4
Det		5		2,5	1	0
Ede		2		3	2	1
François		2		3,5	1	1
Gerd	3			4	1	4
Hildegard	3			4,5	1	0
Irmgard	3,5			5	4	1
Jürgen		3		5,5	0	1
Karl	5			6	0	0
Lotte	5					
Martha	4,5					
Norbert		2				
Otto		5,5				
Petra		4				
René		1				
Stefan		2				
Traudl	4					
Ulla	5					
Victoria	5					
Walter		4				
Xaver		3,5				
Yaşar		4				

Abbildung 9.15 HÄUFIGKEIT

Das gleiche Ergebnis würde man auch mit der Funktion ZÄHLENWENN erreichen:

=ZÄHLENWENN(\$A\$1:\$D\$14;"<="&F1;\$A\$1:\$D\$14;">"&F2)

Jedoch ist HÄUFIGKEIT recht flexibel wenn es um die Anzahl unterschiedlicher Elemente geht – wenn Sie eine Zahlenreihe haben, liefert

=SUMME(WENN(HÄUFIGKEIT(A1:A1000;A1:A1000)>=1;1;0))

die Anzahl der unterschiedlichen Zahlen. Zum gleichen Ergebnis würde man auch mit der Funktion SUMMENPRODUKT gelangen:

=SUMMENPRODUKT(1/ZÄHLENWENN(\$A\$1:\$A\$1000;\$A\$:\$A\$1000))

Da SUMMENPRODUKT eine Matrixfunktion ist, muss sie nicht mit [Shift] + [Strg] + [Enter] beendet werden.

Chierteilung

Frau Lämpel geht nun von einem Signifikanzniveau von 15 % aus. Um die Wahrscheinlichkeit zu bestimmen, ob ein Zusammenhang besteht, bestimmt sie die Freiheitsgrade. Das ist das Produkt aus den um 1 verminderten Zeilen und Spalten, also $(9-1) \cdot (2-1) = 8$ Freiheitsgrade. Nun liefert die Funktion

=CHIU.VERT.RE(12,6171;8)

den Wert 0,1257, also unter der 15%-Grenze. Früher hieß diese Funktion CHIVERT. Den gleichen Wert erhält sie, wenn sie die Funktion CHITEST einsetzt:

=CHITEST(W11:X19;R11:S19)

Übrigens könnte Frau Lämpel die Zahl 12,6171 rechnerisch aus dem Ergebnis des Chitests mit der folgenden Funktion ermitteln:

=CHIU.INV.RE(0,0969;8)

wobei die Zahl 8 die Anzahl der Freiheitsgrade ist. Früher hieß die Funktion CHIINV.

Das Bestimmtheitsmaß

Teilt man den Wert 12,6171 durch die Summe aus 12,6171 und der Schüleranzahl 24 und zieht daraus die Wurzel, so erhält man den Pearson'schen Kontingenzkoeffizienten, der in einem Wertebereich zwischen 0 und 1 liegt. In Frau Lämpels Fall liegt er bei 0,587. Mit ihm kann man nun das BESTIMMTHEITSMASSE berechnen.

In diesem Zusammenhang seien noch einige weitere Funktionen erwähnt:

Ein Maß, wie asymmetrisch eine eingipflige Häufigkeitsverteilung um den Mittelwert liegt, liefert die Funktion SCHIEFE. Sie ergibt bei den vorliegenden Noten der ersten Klassenarbeit die Zahl -0,2616, bei der zweiten -0,937. Diese negative Zahl zeigt eine Verteilung an, deren Gipfel sich mehr hin zu Werten kleiner dem Mittelwert erstreckt.

Die Funktion KRITBINOM liefert den kleinsten Wert, für den die kumulierten Wahrscheinlichkeiten der Binomialverteilung größer oder gleich einer Grenzwahrscheinlichkeit sind.

Die Wölbung einer Verteilung im Vergleich zu der Normalverteilung wird mit der Kurtosis (KURT) gemessen: Eine positive Kurtosis weist auf eine schmale und spitze Verteilung hin, eine negative auf eine flache Verteilung.

Weitere Verteilungsfunktionen, die Excel zur Verfügung stellt, sollen hier nicht ausführlich beschrieben werden – dies würde den Rahmen des Buches sprengen.

Standardabweichung in der Pharmazie

Ein einfaches Beispiel: Eine pharmazeutische Firma prüft regelmäßig Produkte. In zwei Zellen werden der Produktname und die Spezifikation hineingeschrieben, in einer anderen Zelle wird daraus der Überschriftstext ermittelt:

```
= "Calculation of Standardcurve of "&A1&" in "&B1
```

In Spalte B befinden sich die Volumina, die gemessen werden, in Spalte C und D zwei Messergebnisse. In Spalte E wird das Ergebnis berechnet:

```
=C5-D5
```

Mit einem Doppelklick auf das Kästchen kann die Berechnung nach unten gezogen werden (Abbildung 10.9).

In einer zweiten Tabelle, die sich weiter unten befindet, werden die Volumina erneut aufgelistet. Nun soll der Mittelwert der Zahlen berechnet werden, deren Werte in Spalte E zu den entsprechenden aus B gehören. Man könnte die Mittelwertfunktion durch Anklicken der einzelnen Zellen berechnen:

```
=MITTELWERT(E5;E10;E15)
```

E5					
	A	B	C	D	E
2	Calculation of Standardcurve of Hexachloral in Hexapidin#77				
3					
4	Sample	Inj. Vol (µl)	Area (mV x m)	Std. blank	standard-blank
5	Std. dil. to 13 µg/ml	10	32.78	1.73	31.05
6	Std. dil. to 13 µg/ml	20	68.50	3.65	64.85
7	Std. dil. to 13 µg/ml	30	102.58	5.56	97.02
8	Std. dil. to 13 µg/ml	40	138.42	7.43	130.99
9	Std. dil. to 13 µg/ml	50	170.73	9.47	161.26
10	Std. dil. to 13 µg/ml	10	33.14	1.86	31.28
11	Std. dil. to 13 µg/ml	20	68.49	3.93	64.56
12	Std. dil. to 13 µg/ml	30	105.26	5.66	99.60
13	Std. dil. to 13 µg/ml	40	137.13	7.37	129.76
14	Std. dil. to 13 µg/ml	50	168.64	9.66	158.98
15	Std. dil. to 13 µg/ml	10	33.29	1.85	31.44
16	Std. dil. to 13 µg/ml	20	68.93	3.76	65.17
17	Std. dil. to 13 µg/ml	30	103.96	5.75	98.21
18	Std. dil. to 13 µg/ml	40	137.47	7.48	129.99
19	Std. dil. to 13 µg/ml	50	168.69	9.40	159.29

Abbildung 9.16 Die erfassten Daten

Stehen die Volumina in Spalte B regelmäßig untereinander (also beispielsweise 10, 20, 30, 40, 50, 10, 20, ...), so kann die Berechnung problemlos heruntergezogen werden. Was tut man aber, wenn dies nicht der Fall ist?

Da sich der Mittelwert als SUMME/ANZAHL berechnen lässt, können statt diesen beiden Funktionen die entsprechenden Bedingungsfunktionen verwendet werden:

```
=SUMMEWENN($B$5:$B$19;A25;$E$5:$E$19) /  
ZÄHLENWENN($B$5:$B$19;A25)
```

Neben dem Durchschnitt soll allerdings auch die Standardabweichung berechnet werden. Es gibt keine Funktion STABWNWENN oder Ähnliches. Da die Berechnung der Standardabweichung sehr viel komplexer ist als der Mittelwert, erscheint es wenig sinnvoll, sie in ihre Bestandteile zu zerlegen und über Wenn-Funktionen zu berechnen. Es geht allerdings auch anders.

```
=STABW.N($E$5:$E$19)
```

berechnet die Standardabweichung aller Zahlen von E5 bis E19.

```
{=STABW.N(WENN(A25=$B$5:$B$19;$E$5:$E$19;""))}
```

Da die Zelle A25 nicht gleichzeitig mit allen Zellen von B5 bis B19 verglichen werden kann, muss mit einer Matrixfunktion gerechnet werden. Das heißt, Sie beenden den Funktionsassistenten oder die Eingabe mit [Shift] + [Strg] + [Enter]. In der Eingabezeile sind danach die beiden geschweiften Klammern sichtbar.

Weiter unten werden über die Dosierung und die berechneten Mittelwerte die Steigung, der y-Achsenabschnitt und die Korrelation der beiden Datenreihen ermittelt:

Statistische Funktionen

=STEIGUNG (B25:B29;C25:C29)

=ACHSENABSCHNITT (B25:B29;C25:C29)

=KORREL (B25:B29;C25:C29)

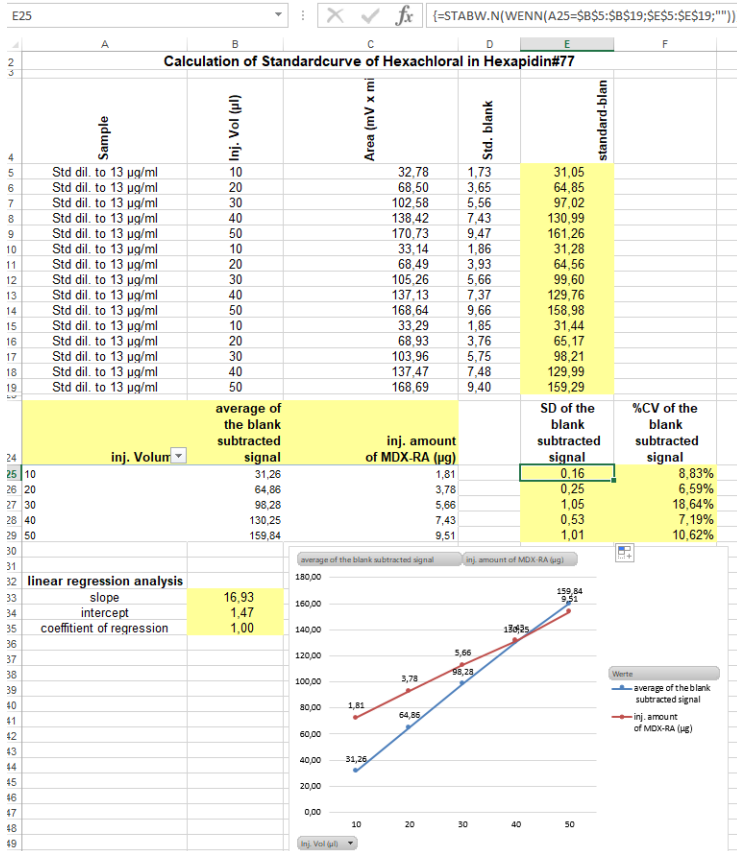


Abbildung 9.17 Die ermittelten Werte

Damit wird auf einem zweiten Blatt weitergerechnet. Über einen Bezug kann man sich den Wert holen. Auf diesem Blatt steht jeweils dreimal untereinander:

Hexachloral active sample 1
 Hexachloral active sample 1
 Hexachloral active sample 1
 Hexachloral active sample 2

Hexachloral active sample 2
 Hexachloral active sample 2
 Hexachloral active sample 3 ...

Dies kann über eine Zeilenfunktion gelöst werden. Die erste Formel in Zeile 8 wird berechnet über:

=standard!\$A\$1&" active sample "&GANZZAHL((ZEILE()+1)/3)-2

Dabei ist „standard“ der Name des Blattes, in dem das pharmazeutische Produkt in Zelle A1 steht. In Zeile 8 liefert Zeile() die Zahl 8. Zu ihr wird 1 addiert, das Ergebnis durch 3 geteilt und abgerundet. Dies liefert 3. Um auf die Zahl 1 zu kommen, muss nur noch 2 abgezogen werden. Diese Funktion wird nun für alle Samples nach unten gezogen. Um unter jeder Dreiergruppe einen Strich zu erhalten, könnte man die bedingte Formatierung verwenden. Über das Register Start | Bedingte Formatierung wird eingestellt: „Formel ist“

=GANZZAHL((ZEILE()-1)/3)=(ZEILE()-1)/3

D8

Abbildung 9.18 Das Auswertungsblatt

Damit wird unter der Zeile 10, 13, 16, ... die Formatierung verwendet – nämlich ein Strich. Und mit einer ähnlichen Funktion kann nun der Mittelwert der jeweils drei zusammengehörigen Zellen berechnet werden:

```
=WENN (GANZZAHL ( (ZEILE () +1) / 3) = (ZEILE () +1) / 3 ;  
MITTELWERT (C8:C10) ; "" )
```

9.3 Weitere statistische Hilfsmittel

Sind in den Excel-Optionen die Analysefunktionen aktiviert, so weist Excel in diesem Register Daten | Analyse den Punkt Datenanalyse auf. Über diesen kann eine Reihe von Funktionen, die mit dem Funktionsassistenten zur Verfügung stehen, auf Datenbereiche angewendet werden.

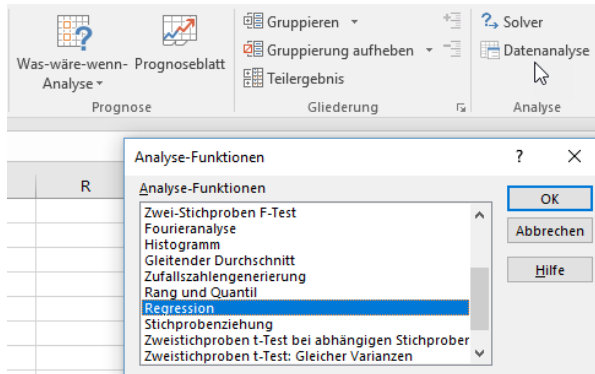


Abbildung 9.19 Die Analysefunktionen

Diese Analysefunktionen helfen dabei, dass Sie von großen Wertebereichen nicht jede Funktion einzeln berechnen müssen, sondern mit einem Mausklick alle wichtigen Funktionsergebnisse zur Verfügung haben.

Dabei werden zum Teil nicht nur die Rechenergebnisse angezeigt, sondern auch Diagramme über bestimmte Verteilungen erzeugt.

Beispielsweise liefert der Punkt Regression eine Reihe von Informationen über einen Wertebereich: