Statistik

1

```
Remove["Global`*"]
```

Daten

```
p[k_] := (45 - k) / (448 - k);
```

a

Jeder der 3, die kommen, kommt unabhängig. Die Wahrscheinlichkeiten multiplizieren sich. Das bedeutet, man hat: Günstige Fälle = 45^3, mögliche Fälle = 448^3. Man rechnet hier somit mit Variationen mit Wiederholung. Das führt dazu, dass z.B. eine Auswahl von Kunden AAB und eine Auswahl BAA als verschieden betrachtet werden. In der Realität liegen diese verschiedenen Fälle tatsächlich vor.

```
p[0]^3
```

$$\frac{91125}{89915392}$$

```
45^3 / 448^3
```

$$\frac{91125}{89915392}$$

```
N[%]
```

0.00101345

Rechnet man jedoch mit Kombinationen mit Wiederholung, so ergibt sich ein leicht anderes Resultat. Bei Kombinationen wird z.B. die Auswahl von Kunden AAB nicht von der BAA unterschieden. Permutiert man die Kombinationen wie wenn alle Elemente verschieden wären (Multiplikation mit k!), so erhält man (n-k+1)! / (n-1)!. Da aber nicht immer alle Elemente verschieden sind, kann man in der unten stehenden Formel nicht einfach im Zähler und im Nenner mit k¨multiplizieren. (n-k+1)! / (n-1)! ist ungleich den Variationen n^k. Erstaunich ist aber, dass die Abweichung des Resultates vom vorhergehenden nicht wesentlich ist.

```
Binomial[45 + 3 - 1, 3] / Binomial[448 + 3 - 1, 3]
```

$$\frac{1081}{1005760}$$

```
N[%]
```

```
0.00107481
```

(Dieses Resultat ist nicht richtig.)

b

Mit günstigen durch mögliche Fälle:

```
p[k_] := (45 - k) / (448 - k);
```

```
p[0] p[1] p[2]
```

$$\frac{7095}{7442848}$$

```
N[%]
```

```
0.000953264
```

Mit Kombinationen erhält man das gleiche Resultat:

```
Binomial[45, 3] / Binomial[448, 3]
```

$$\frac{7095}{7442848}$$

```
N[%]
```

```
0.000953264
```

c

```
Chance = 1 - Binomial[448 - 45, 3] / Binomial[448, 3]
```

$$\frac{4058295}{14885696}$$

```
N[%]
```

```
0.272631
```

2

```
Remove["Global`*"]
```

Daten

```
p = 12 / 100; q = 1 - p;
```

a

Methode 1

```
Chance = Binomial[(1 - 12 / 100) 1000, 10] / Binomial[1000, 10]
```

$$\frac{364539927967702599827}{1317047802309851064162}$$

```
N[%]
```

```
0.276786
```

Methode 2

12 Prozent von 1000 ist 880. Man greift 10 Stücke aus den 1000 heraus und will nur gute. Dan wiederholt man das Herausgreifen: Total 10 einzelne Stichproben. Das gibt die Produktwahrscheinlichkeit

```
Remove["Global`*"]
```

```
p[k_] := (880 - k) / (1000 - k);
```

```
p[0] p[1] p[2] p[3] p[4] p[5] p[6] p[7] p[8] p[9]
```

$$\frac{364539927967702599827}{1317047802309851064162}$$

```
N[%]
```

```
0.276786
```

3

```
Remove["Global`*"]
```

Daten

```
M1 = Table[12.4, {n, 1, 7}]
```

```
{12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4}
```

```
M2 = Table[12.5, {n, 1, 12}]
```

```
{12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5}
```

```
M3 = Table[12.6, {n, 1, 14}]
```

```
{12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6}
```

```
M4 = Table[12.7, {n, 1, 15}]
```

```
{12.7, 12.7, 12.7, 12.7, 12.7, 12.7,
 12.7, 12.7, 12.7, 12.7, 12.7, 12.7, 12.7, 12.7, 12.7}
```

```
M5 = Table[12.8, {n, 1, 11}]
```

```
{12.8, 12.8, 12.8, 12.8, 12.8, 12.8, 12.8, 12.8, 12.8, 12.8, 12.8}
```

```
M6 = Table[12.9, {n, 1, 5}]
```

```
{12.9, 12.9, 12.9, 12.9, 12.9}
```

```
M0 = Join[M1, M2, M3, M4, M5, M6]
```

```
{12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.5, 12.5, 12.5, 12.5, 12.5,
 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6,
 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.7, 12.7, 12.7, 12.7, 12.7,
 12.7, 12.7, 12.7, 12.7, 12.7, 12.7, 12.7, 12.7, 12.7, 12.7, 12.8, 12.8, 12.8,
 12.8, 12.8, 12.8, 12.8, 12.8, 12.8, 12.8, 12.8, 12.9, 12.9, 12.9, 12.9, 12.9}
```

```
N1 = Table[12.1, {n, 1, 5}]
```

```
{12.1, 12.1, 12.1, 12.1, 12.1}
```

```
N2 = Table[12.2, {n, 1, 11}]
```

```
{12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2}
```

```
N3 = Table[12.3, {n, 1, 15}]
```

```
{12.3, 12.3, 12.3, 12.3, 12.3, 12.3,
 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3}
```

```
N4 = Table[12.4, {n, 1, 14}]
```

```
{12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4}
```

```
N5 = Table[12.5, {n, 1, 12}]
```

```
{12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5}
```

```
N6 = Table[12.6, {n, 1, 7}]
```

```
{12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6}
```

```
N0 = Join[N1, N2, N3, N4, N5, N6]
```

```
{12.1, 12.1, 12.1, 12.1, 12.1, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2,
 12.2, 12.2, 12.2, 12.2, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3,
 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4,
 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5,
 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6}
```

a

```
Length[M0]
```

```
64
```

```
<< Statistics`DescriptiveStatistics`
```

```
LocationReport[M0]
```

{Mean → 12.6406, HarmonicMean → 12.639, Median → 12.6}

```
DispersionReport[M0]
```

{Variance → 0.0211806, StandardDeviation → 0.145535, SampleRange → 0.5,
MeanDeviation → 0.123145, MedianDeviation → 0.1, QuartileDeviation → 0.125}

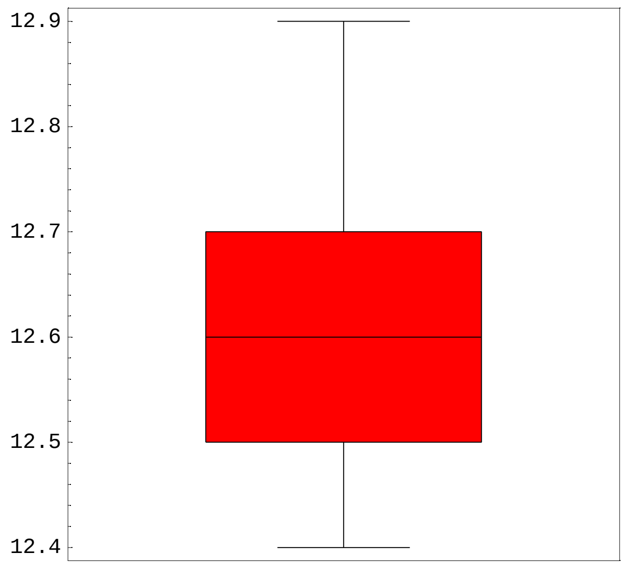
```
ShapeReport[M0]
```

{Skewness → 0.015996, QuartileSkewness → 0.2, KurtosisExcess → −0.911254}

```
<< Statistics`StatisticsPlots`
```

b

```
BoxWhiskerPlot[M0];
```




```
Length[N0]
```

64

```
<< Statistics`DescriptiveStatistics`
```

```
LocationReport[N0]
```

{Mean → 12.3594, HarmonicMean → 12.3577, Median → 12.4}

```
DispersionReport[N0]
```

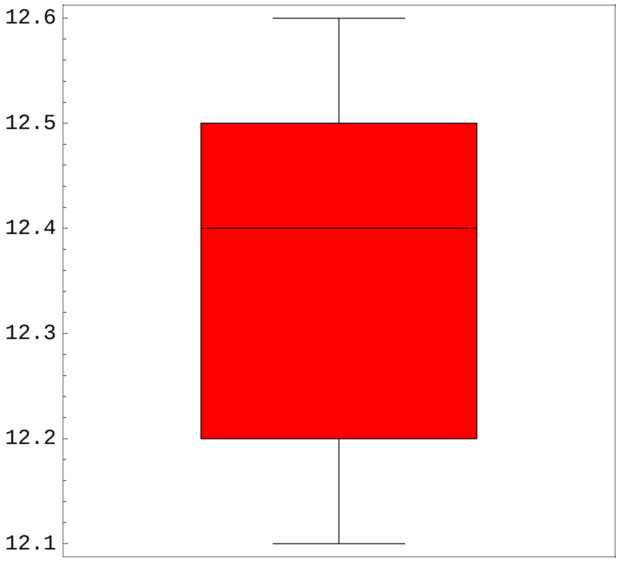
{Variance → 0.0211806, StandardDeviation → 0.145535, SampleRange → 0.5,
MeanDeviation → 0.123145, MedianDeviation → 0.1, QuartileDeviation → 0.125}

```
ShapeReport[N0]
```

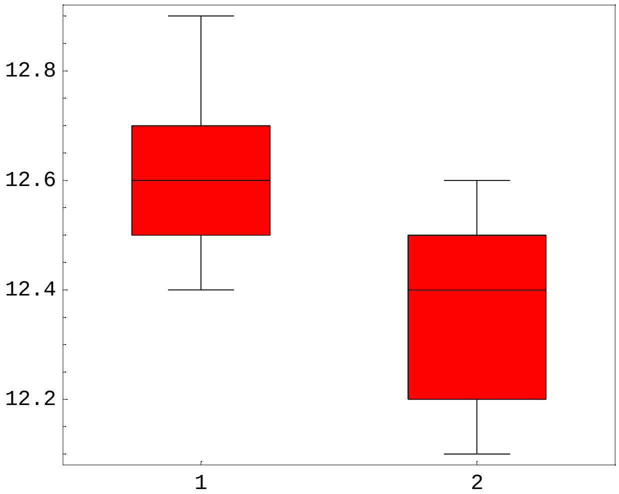
{Skewness → −0.015996, QuartileSkewness → −0.2, KurtosisExcess → −0.911254}

```
<< Statistics`StatisticsPlots`

BoxWhiskerPlot[N0];
```




```
BoxWhiskerPlot[M0, N0];
```



c

```
N7 = Table[12.1, {n, 1, 5}]

{12.1, 12.1, 12.1, 12.1, 12.1}

N8 = Table[12.1 + 0.6, {n, 1, 5}]

{12.7, 12.7, 12.7, 12.7, 12.7}
```

```
N0 = Join[N2, N3, N4, N5, N6, N8]
```

```
{12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.2, 12.3,
 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3, 12.3,
 12.3, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4, 12.4,
 12.4, 12.4, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5, 12.5,
 12.5, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.6, 12.7, 12.7, 12.7, 12.7, 12.7}
```

```
Length[N0]
```

```
64
```

```
<< Statistics`DescriptiveStatistics`
```

```
LocationReport[N0]
```

```
{Mean → 12.4063, HarmonicMean → 12.4044, Median → 12.4}
```

```
DispersionReport[N0]
```

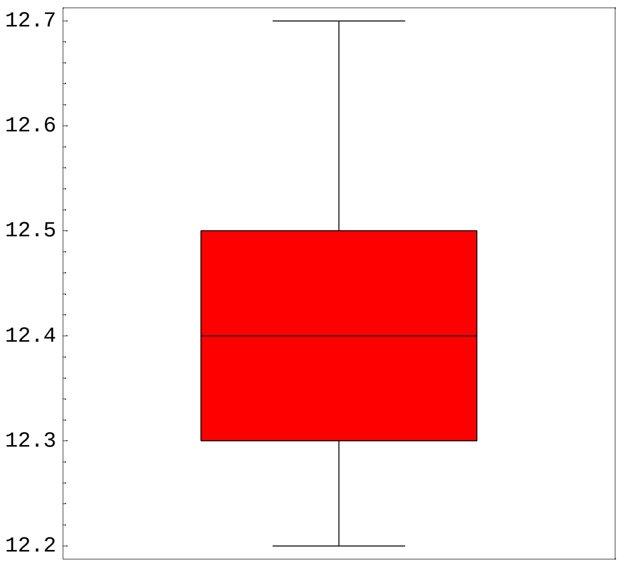
```
{Variance → 0.0228175, StandardDeviation → 0.151054, SampleRange → 0.5,
 MeanDeviation → 0.123438, MedianDeviation → 0.1, QuartileDeviation → 0.1}
```

```
ShapeReport[N0]
```

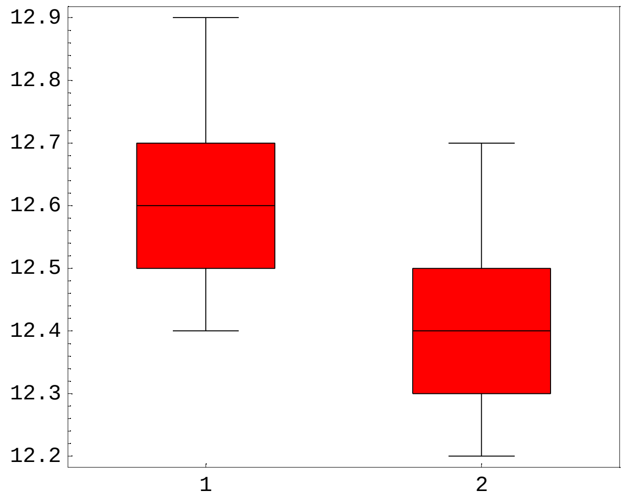
```
{Skewness → 0.33899, QuartileSkewness → 0., KurtosisExcess → -0.837027}
```

```
<< Statistics`StatisticsPlots`
```

```
BoxWhiskerPlot[N0];
```



```
BoxWhiskerPlot[M0, N0];
```



4

```
Remove["Global`*"]
```

a

```
<< Statistics`ContinuousDistributions`
```

```
μ = 5.295;   σ = 0.005;
```

```
PDF[NormalDistribution[μ, σ], x]
```

$79.7885\, e^{-20000.\,(-5.295+x)^2}$

```
CDF[NormalDistribution[μ, σ], x]
```

$\frac{1}{2}\,(1 + \mathrm{Erf}[141.421\,(-5.295 + x)])$

```
2 CDF[NormalDistribution[μ, σ], μ - 0.007]
```

0.161513

b

```
1 - CDF[NormalDistribution[μ, σ], 5.3]
```

0.158655

5.

```
Remove["Global`*"]
```

```
n = 150;
xQuer = 315.0;
σQuadrat = 900;
σ = Sqrt[σQuadrat];
α = 0.01;
```

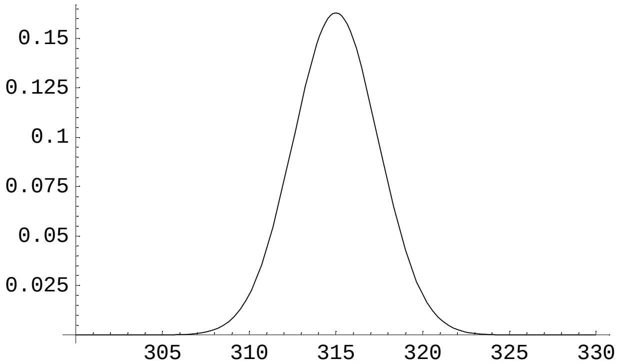
```
σ // N
```

30.

```
NV[x_, μ_, σ_, n_] := 1 / (σ / Sqrt[n] Sqrt[2 Pi]) E^(-1 / 2 ((x - μ) Sqrt[n] / σ) ^2);
NV[μ, xQuer, σ, n]
```

$$\frac{e^{-\frac{1}{12}(-315.+\mu)^2}}{2\sqrt{3\pi}}$$

```
Plot[NV[μ, xQuer, σ, n], {μ, 300, 330}];
```




```
Integrate[NV[μ, xQuer, σ, n], {μ, 280, 340}]
```

1.0000000000

Integriere nur über die halbe Normalverteilung:

```
Integrate[NV[μ, xQuer, σ, n], {μ, xQuer, x}] == (1 - α) / 2
```

$4.38463\times10^{-15} + 0.5\,\mathrm{Erf}[-90.9327 + 0.288675\,x] == 0.495$

```
rootOben =
 FindRoot[Evaluate[Integrate[NV[μ, xQuer, σ, n], {μ, xQuer, x}]] == (1 - α) / 2, {x, 320}]
```

{x → 321.309}

```
cOben = x /. rootOben
```

321.309

```
cUnten = xQuer - (cOben - xQuer)
```

308.691

```
Integrate[NV[μ, xQuer, σ, n], {μ, cUnten, cOben}]
```

0.98999999985

```
1 - Integrate[NV[μ, xQuer, σ, n], {μ, cUnten, cOben}]
```

```
0.01000000015
```

Das ist etwa . Lösung: [cUnten, cOben] ist das Konfidenzintervall.

6.

```
Remove["Global`*"]
```

n_0 = 20
0 neutral
17 A Differenzen grösser 0 ==> A (1. Stichprobe)
3 Differenzen kleiner 0 ==> B (2. Stichprobe)

H0 ==> beide Verfahren sind gleich
H1 ==> beide Verfahren sind nicht gleich, Verfahren verschieden

Frage: Wahrscheinlichkeit, dass unter Annahme von H0 trotzdem eine Abweichung zwischen A und B auftritt,dass also H0 falsch sein muss?
Sei x_i der i-te Wert der aus der 1. Stichprobe mit Verfahren A, x_i' der i-te Wert der aus der 2. Stichprobe mit Verfahren B.
Sei $d_i = x_i - x_i$ '. Dann ist das Auftreten einer negativen Differenz gleich wahrscheinlich wie das Auftreten einer positiven Differenz.
Wir nehmen an,dass die Verteilungen A und B angesichts der Messwerte x_i und x_i' stetig sind Messungen sind hier (keine Anzahlen). Daher kann man angesichts der daraus resultierenden Wahrscheinlichkeitsdichten $P(X_i=X_i') = 0$ setzen.
Weiter sind die Zufallsgrössen für i=1,2,...,n unabhängig.
Damit wird neu n=17+3=20.
So ergibt sich z.B.für A neu eine Binomialverteilung mit $p=\frac{1}{2}$, $q = 1-\frac{1}{2}=\frac{1}{2}$ und n=20.
$P(X=17)=\binom{20}{17}\left(\frac{1}{2}\right)^{17}\left(\frac{1}{2}\right)^{(20-17)} = \binom{20}{17}\left(\frac{1}{2}\right)^{20}$, P(X=k) := f(k)

Um weiter zu kommen betrachten wir die neue Alternativhypothese H1, welche postuliert, dass die x_i-Werte im Durchschnitt wesentlich grösser sind als die x_i'-Werte.
H0 wird dann abgelehnt, wenn die Anzahl der positiven Differenzen k^+ einen kritischen Wert $k_{(1-\alpha)}$ überschreitet oder gleich ist.
Vorderung: $P(K^+ \;\; k_{(1-\alpha)})$, = Irrtumswahrscheinlichkeit.
Es gilt: $P(K^+ \;\; k_{(1-\alpha)}) = \sum_{i=k\_(1-\alpha)}^{n} P(K^+ = i)$ ==> $k_{(1-\alpha)}$ = ?

```
f[k_] := Binomial[20, k] (1/2)^20; f[k]
```

$$\frac{\text{Binomial}[20,k]}{1048576}$$

```
f[20] // N
```

9.53674×10^{-7}

```
s[k_] := Sum[N[f[u]], {u, k, 20}];  s[k]
```

$$\frac{2.3202\times10^{12}\ \text{Hypergeometric2F1}[1,\ -20+k,\ 1+k,\ -1]}{\text{Gamma}[21.-1.\ k]\ \text{Gamma}[1+k]}$$

s[k] ist die Wahrscheinlichkeit
P[K 17 (K grösser gleich 17)] = 1-P[K < 17 (K kleiner 17)]
P[K < 17 (K kleiner 17)] = 1-P[K 17 (K grösser gleich 17)] = 1-s[k]

```
Prepend[Table[{k, s[k], 1 - s[k]}, {k, 0, 20}],
  {"n", "s[k]=α", "1-s[k]=1-α"}] // TableForm
```

n	s[k]=α	1-s[k]=1-α
0	1.	0.
1	0.999999	9.53674×10^{-7}
2	0.99998	0.0000200272
3	0.999799	0.000201225
4	0.998712	0.00128841
5	0.994091	0.00590897
6	0.979305	0.0206947
7	0.942341	0.0576591
8	0.868412	0.131588
9	0.748278	0.251722
10	0.588099	0.411901
11	0.411901	0.588099
12	0.251722	0.748278
13	0.131588	0.868412
14	0.0576591	0.942341
15	0.0206947	0.979305
16	0.00590897	0.994091
17	0.00128841	0.998712
18	0.000201225	0.999799
19	0.0000200272	0.99998
20	9.53674×10^{-7}	0.999999

Ein Wert k 17 (k grösser gleich 17) kommt daher mit der Wahrscheinlichkeit 0.00128841... kleiner vor, je nach k. Für grössere -Werte (alpha-Werte) müsste man k kleiner haben, damit P[K k, grösser gleich] (kleiner gleich alpha) eintrifft. Für k 18 (k grösser gleich 18) hat daher die Alternative eine sehr kleine Wahrscheinlichkeit.
Da die eingetroffene reale Situation, die Alternative also mit k = 17 mit p= 0.00128841... so unwahrscheinlich aber dennoch real ist, müsste man H0 z.B. bei einem Signifikanzniveau von = 0.05 oder bei = 0.01ablehnen, denn es ist 0.00128841... < 0.01. Die Abweichung von einem erwarteten Resultat, dass die negativen und die positiven Differenzen etwa gleich oft vorkommen müsten, ist hier zu gross.

7.

```
Remove["Global`*"]
```

```
f[k_] := Binomial[40, k] 0.5^k 0.5^(40 - k)
```

```
Sum[f[k], {k, 0, 5}]
```

6.91306×10^{-7}

8.

Daten

```
Remove["Global`*"]
```

```
a1 = 50 - 8;
a2 = 50 - 2;
b1 = 4 * 2;
b2 = 2;
```

a

```
matrix = {{"Frau", a1, b1, a1 + b1},
  {"Mann", a2, b2, a2 + b2}, {"Total", a1 + a2, b1 + b2, a1 + a2 + b1 + b2},
  {" ", "Gewöhnlich", "Gepanzert", "Total"}}; matrix // MatrixForm
```

Frau	42	8	50
Mann	48	2	50
Total	90	10	100
	Gewöhnlich	Gepanzert	Total

b

```
10 / 100 // N
```

0.1

c

```
2 / 100 // N
```

0.02

d

```
2 / 50 // N
```

0.04