Lösungen / Statistik 1/16

1.

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

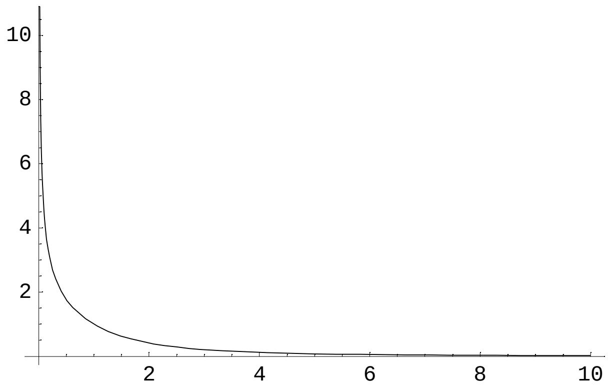
```
h = p * q; s = p + q;
solv = Solve[ s * h / 2 == 1, {q}]
```

$$\left\{\left\{q \to \frac{-p^2 - \sqrt{p}\ \sqrt{8 + p^3}}{2\ p}\right\},\ \left\{q \to \frac{-p^2 + \sqrt{p}\ \sqrt{8 + p^3}}{2\ p}\right\}\right\}$$

$$\frac{-p^{3/2} + \sqrt{8 + p^3}}{2\ \sqrt{p}}\ \Big/ .0 \to 0$$

```
ComplexInfinity → 0
```

$$\text{Plot}\left[\frac{-p^{3/2} + \sqrt{8 + p^3}}{2\ \sqrt{p}},\ \{p,\ 0,\ 10\}\right];$$




```
h * s /. solv[[2]] // Simplify
```

```
2
```

```
p * (p * q) + q * (p * q) /. solv[[2]] // Simplify
```

```
2
```

$$\text{f1}[x\_,\ p\_] := p\ x + \left(p * \frac{-p^{3/2} + \sqrt{8 + p^3}}{2\ \sqrt{p}}\right)\ //\ \text{Simplify};$$

$$\text{f2}[x\_,\ p\_] := -\frac{-p^{3/2} + \sqrt{8 + p^3}}{2\ \sqrt{p}}\ x + \left(p * \frac{-p^{3/2} + \sqrt{8 + p^3}}{2\ \sqrt{p}}\right)\ //\ \text{Simplify};$$

```
p x + (p * q) /. solv[[2]]
```

$$\frac{1}{2}\ \left(-p^2 + \sqrt{p}\ \sqrt{8 + p^3}\right) + p\ x$$

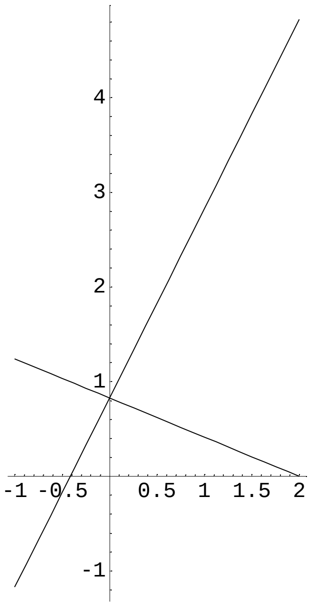
f1[x, 2]

$2\left(-1+\sqrt{2}+x\right)$

f2[x, 2]

$-\left(-1+\sqrt{2}\right)\,(-2+x)$

Plot[{f1[x, 2], f2[x, 2]}, {x, -1, 2}, AspectRatio → Automatic];



Remove[f]

$$f[x\_, p\_] := 0 \;/;\; x \le -\frac{-p^{3/2}+\sqrt{8+p^3}}{2\sqrt{p}};$$

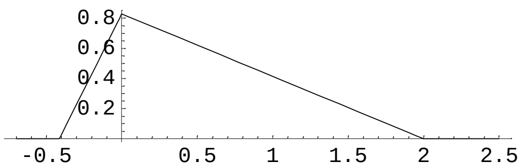
$$f[x\_, p\_] := f1[x, p] \;/;\; -\frac{-p^{3/2}+\sqrt{8+p^3}}{2\sqrt{p}} < x \;\&\&\; x \le 0;$$

f[x_, p_] := f2[x, p] /; 0 < x && x ≤ p;
f[x_, p_] := 0 /; p < x

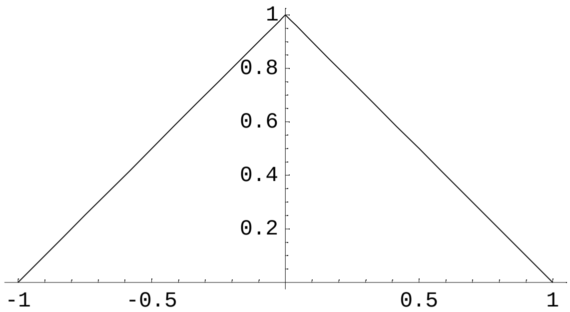
f[-1, 2]

0

Plot[{f[x, 2]}, {x, -0.7, 2.5}, AspectRatio → Automatic];



```
Plot[{f[x, 1]}, {x, -1, 1}, AspectRatio → Automatic];
```




```
e[n_, p_] := NIntegrate[x^n f[x, p], {x, -Infinity, Infinity}];
e[3, 2]
```

0.661504

$$\texttt{e[n\_, p\_] := Integrate}\left[\texttt{x\^{}n f[x, p]}, \left\{\texttt{x}, \frac{-p^{3/2}+\sqrt{8+p^3}}{2\sqrt{p}}, \texttt{p}\right\}\right];$$

```
e[3, 2]
```

$$\int_{\frac{4-2\sqrt{2}}{2\sqrt{2}}}^{2} x^3\, f[x, 2]\, dx$$

```
eNew[n_, p_] :=
```

$$\texttt{Integrate}\left[\texttt{x\^{}n f1[x, p]}, \left\{\texttt{x}, -\frac{-p^{3/2}+\sqrt{8+p^3}}{2\sqrt{p}}, 0\right\}\right] + \texttt{Integrate[x\^{}n f2[x, p], \{x, 0, p\}];}$$

```
eNew[3, 2]
```

$$\frac{5}{2} - \frac{29}{5\sqrt{2}} + \frac{8\sqrt{2}}{5}$$

```
{eNew[3, 1], eNew[2, 1], eNew[1, 1], eNew[0, 1]}
```

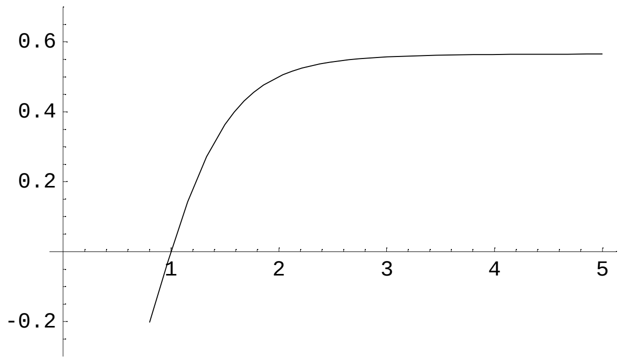
$$\left\{0, \frac{1}{6}, 0, 1\right\}$$

```
μ[p_] := eNew[1, p]; σ[p_] := Sqrt[eNew[2, p] - μ[p]^2];
γ[p_] := 1/σ[p]^3 * (eNew[3, p] - 3 μ[p] eNew[2, p] + 2 μ[p]^3)
```

```
{μ[1], σ[1], γ[1]}
```

$$\left\{0, \frac{1}{\sqrt{6}}, 0\right\}$$

```
Plot[γ[p], {p, 0.8, 5}, PlotRange → {-0.3, 0.7}];
```



2.

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

n_0 = 25
3 neutral
18 A Differenzen grösser 0 ==> A (1. Stichprobe)
4 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=18+4=22.
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=22.
P(X=18)= $\binom{22}{18} \left(\frac{1}{2}\right)^{18} \left(\frac{1}{2}\right)^{(22-18)} = \binom{22}{18} \left(\frac{1}{2}\right)^{22}$, 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[22, k] (1 / 2) ^22;  f[k]
```

$$\frac{\text{Binomial}[22,\ k]}{4194304}$$

```
f[18] // N
```

0.00174403

```
s[k_] := Sum[N[f[u]], {u, k, 22}];  s[k]
```

$$\frac{2.67983\times 10^{14}\ \text{Hypergeometric2F1}[1,\ -22+k,\ 1+k,\ -1]}{\text{Gamma}[23.-1.\ k]\ \text{Gamma}[1+k]}$$

s[k] ist die Wahrscheinlichkeit
P[K   18 (K grösser gleich 18)] = 1-P[K < 18 (K kleiner 18)]
P[K < 18 (K kleiner 18)] = 1-P[K   18 (K grösser gleich 18)] = 1-s[k]

```
Prepend[Table[{k, s[k], 1 - s[k]}, {k, 0, 22}],
  {"n", "s[k]=α", "1-s[k]=1-α"}] // TableForm
```

n	s[k]=α	1-s[k]=1-α
0	1.	0.
1	1.	2.38419×10^{-7}
2	0.999995	5.48363×10^{-6}
3	0.999939	0.0000605583
4	0.999572	0.000427723
5	0.997828	0.00217175
6	0.99155	0.00845027
7	0.973761	0.0262394
8	0.9331	0.0669003
9	0.856861	0.143139
10	0.738266	0.261734
11	0.584094	0.415906
12	0.415906	0.584094
13	0.261734	0.738266
14	0.143139	0.856861
15	0.0669003	0.9331
16	0.0262394	0.973761
17	0.00845027	0.99155
18	0.00217175	0.997828
19	0.000427723	0.999572
20	0.0000605583	0.999939
21	5.48363×10^{-6}	0.999995
22	2.38419×10^{-7}	1.

Ein Wert k   18 (k grösser gleich 18) kommt daher mit der Wahrscheinlichkeit 0.00217... 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 = 18 mit p= 0.00217... so unwahrscheinlich aber dennoch real ist, müsste man H0 z.B. bei einem Signifikanzniveau von   = 0.05 oder bei   = 0.01ablehnen, denn es ist 0.00217... < 0.01. Die Abweichung von einem erwarteten Resultat, dass die negativen und die positiven Differenzen etwa gleich oft vorkommen müsten, ist hier zu gross.

3.

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

Gegeben ist von einem Datensatz mit n = 4984 Messungen der gerundete Mittelwert $\bar{x}n$ = 652 und die gerundete Standardabweichung StD = 184. Weiter ist bekannt, dass die Messungen in Klassen eingeteilt worden waren. Es handelt sich also hier approximativ um Mittelwert und Standardabweichung von mittleren Klassenwerten.
Nun ist bekannt geworden, dass eine Messvorrichtung, mit der j = 196 Werte gemessen worden sind, die Klassenwerte 650 statt richtig 670 geliefert hat. Ebenfalls sind bei dieser Messvorrichtung 212 Werte der Klassengrösse 750 als unsinnig abgetan und unterdrückt worden. n ist also zu klein eingerechnet. Berechne den korrigierten Mittelwert und die korrigierte Standardabweichung in 2 Stufen.
Bemerkung zur Lösung dieser Aufgabe: Diese soll mit einem Rechner ausgeführt werden. Eine Musterlösung wird bei Gelegenheit unter dem folgenden Link bereitgestellt:

```
<< Statistics`DescriptiveStatistics`;
n0 = 4984;
xM0 = 652;
StD0 = 184;
j0 = 196;
m0 = 212;
xj0 = 650;
xj1 = 670;
xm0 = 750;
QS[xM_, n_, StD_] := (n - 1) StD^2 + n xM^2;
```

a Korrigierter Mittelwert und korrigierte Standardabweichung Stufe 1

```
xM1 = (n0 xM0 - j0 xj0 + j0 xj1) / n0
```

$$\frac{58098}{89}$$

```
N[%]
```

652.787

```
QS0 = QS[xM0, n0, StD0]
```

2287422784

```
StD1[n_, QS_, xM_] := Sqrt[QS / (n - 1) - n / (n - 1) xM^2];
```

```
StD1[n0, QS0, xM1]
```

$$4\sqrt{\frac{909967622}{443487}}$$

```
N[%]
```

```
181.189
```

In der erster Stufe ist der Mittelwert von 652 auf etwa 652.79 gestiegen und die Standardabweichung von 184 auf etwa 181.19 gesunken.

b Korrigierter Mittelwert und korrigierte Standardabweichung Stufe 2

```
xM2 = (n0 xM1 + m0 xm0) / (n0 + m0)
```

$$\frac{284374}{433}$$

```
N[%]
```

```
656.753
```

```
QS1 = QS[xM1, n0, StD0]
```

$$\frac{204035841696}{89}$$

```
QS0 - QS1
```

$$-\frac{455213920}{89}$$

```
N[%]
```

$$-5.11476\times 10^{6}$$

```
QS2[QS1_, m_, xm_] := QS1 + m xm^2;
```

```
QS2[QS1, m0, xm0]
```

$$\frac{214649091696}{89}$$

```
N[%]
```

$$2.41179\times 10^{9}$$

```
StD2[n_, m_, QS_, xM_] := Sqrt[QS / (n + m - 1) - (n + m) / (n + m - 1) xM^2];
```

```
StD2[n0, m0, QS2[QS1, m0, xm0], xM2]
```

$$8\sqrt{\frac{20548193565}{40039943}}$$

```
N[%]
```

```
181.23
```

In der ersteun und zweiten Stufe ist der Mittelwert von 652 über etwa 652.79 auf etwa 656.8 gestiegen und die Standardabweichung von 184 über etwa 181.19 auf etwa 181.23 gesunken.