Die t-Verteilung

Jonathan Harrington

Standard error of the mean (SE)

ist die Standardabweichung von Mittelwerten

Ich werfe 5 Würfel und berechne den Mittelwert der Zahlen

μ = 3.5 der wahrscheinlichste Wert

$$SE = \frac{\sigma}{\sqrt{5}}$$

Die Verteilung der Mittelwerte.
Bedeutung: ich werde nicht jedes Mal einen Mittelwert *m* = 3.5 bekommen, sondern davon abweichende Mittelwerte. Der SE ist eine numerische Verschlüsselung dieser Abweichung.

Standard error of the mean (SE)

$$SE = \frac{\sigma}{\sqrt{5}}$$

$$\sigma = \sqrt{\frac{\sum x^2}{n} - \mu^2}$$

```
sigma()/sqrt(5)
```

```
0.7637626
```

```
sigma <- function(unten=1, oben=6)
{
x = unten:oben
n = length(x)
m = mean(x)
sqrt((sum(x^2)/n - m^2))
}
```

Standard error of the mean (SE) und der Vertrauensintervall

95% Vertrauensintervall

$$\mu \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

qnorm(0.025)

3.5 - 1.96 * sigma()/sqrt(5)

oder

qnorm(0.025, 3.5, sigma()/sqrt(5))

2.003025

qnorm(0.975, 3.5, sigma()/sqrt(5))

4.996975

Bedeutung:

Wenn ich 5 Würfel werfe, dann liegt der Stichproben-Mittelwert, *m*, dieser 5 Zahlen zwischen 2.00 und 5.00 mit einer Wahrscheinlichkeit von 95% (0.95).

Probieren!

```
a = proben(1, 6, 5, 100)
sum(a < 2 | a > 5)
```

Standard error of the mean (SE) und der Vertrauensintervall

SE wird kleiner, umso größer n.

$$SE = \frac{\sigma}{\sqrt{n}}$$

umso größer n, umso weniger weicht m von μ ab.

Oder: Je mehr Würfel wir werfen, umso wahrscheinlicher ist es/sicherer wird es sein, dass *m* nah an μ ist. Im unendlichen Fall – wir werfen unendlich viele Würfel und berechnen deren Zahlenmittelwert – ist SE 0 (NULL) und m = μ = 3.5.

Standard error of the mean (SE) **wenn σ unbekannt ist**.

Lenneberg behauptet, dass wir im Durchschnitt mit einer Geschwindigkeit von 6 Silben pro Sekunde sprechen.

Hier sind 12 Werte (Silben/Sekunde) von einem Sprecher.

```
swerte
 [1]  6  5  6  9  6  5  6  8  5  6 10  9
```

Frage: sind die Werte überraschend?
(angenommen $\mu = 6$?).

Präzisere/bessere Frage: ist der Unterschied zwischen μ und *m* **signifikant**? (Oder: fällt *m* außerhalb des 95% Vertrauensintervalls von μ?).

Das Verfahren: **a one-sampled t-test**

Präzisere/bessere Frage: fällt *m* außerhalb des 95% Vertrauensintervalls von μ?

A. Um das Vertrauensintervall um μ zu berechnen, benötigen wir den SE.

B. Damit lässt sich ein Vertrauensintervall

m – *k SE* bis *m* + *k SE* setzen

(*k* ist eine gewisse Anzahl von SEs).

C. Wenn *m* (in diesem Fall 6.75) innerhalb dieses Intervalls fällt, ist das Ergebnis 'nicht signifikant' (konsistent mit der Hypothese, dass wir im Durchschnitt mit 6 Silben pro Sekunde sprechen).

A. Standard error of the mean (SE) berechnen

$$SE = \frac{\sigma}{\sqrt{n}} \qquad \hat{\sigma} = \sqrt{\frac{\sum x^2}{n-1} - m^2}$$

Aber das können wir nicht berechnen, weil wir σ nicht wissen! Wir können aber $\hat{\sigma}$ oder **unsere beste Einschätzung** von σ berechnen

In R kann $\hat{\sigma}$ ganz einfach mit sd() berechnet werden.

Für diesen Fall:

```
werte
 [1]  6  5  6  9  6  5  6  8  5  6 10  9

shut = sd(werte)
```

A. Standard error of the mean (SE) einschätzen

```
werte
 [1]  6  5  6  9  6  5  6  8  5  6 10  9
```

```
shut = sd(werte)
```

Einschätzung des Standard-Errors

$$\hat{SE} = \frac{\hat{\sigma}}{\sqrt{n}}$$

```
SEhut = shut/sqrt(12)
```

```
0.5093817
```

B. Vertrauensintervall: die t-Verteilung

Wenn die Bevölkerungs-Standardabweichung **eingeschätzt** werden muss, dann wird das Vertrauensintervall nicht mit der Normal- sondern der **t-Verteilung** mit einer gewissen Anzahl von **Freiheitsgraden** berechnet.

Die t-Verteilung ist der Normalverteilung recht ähnlich, aber die 'Glocke' und daher das Vertrauensintervall sind etwas breiter (dies berücksichtigt, die zusätzliche Unsicherheit die wegen $\hat{\sigma}$ entsteht).

Bei diesem one-sample t-test ist die Anzahl der Freiheitsgrade, df (degrees of freedom), von der **Anzahl der Werte in der Stichprobe** abhängig: **df = n – 1**

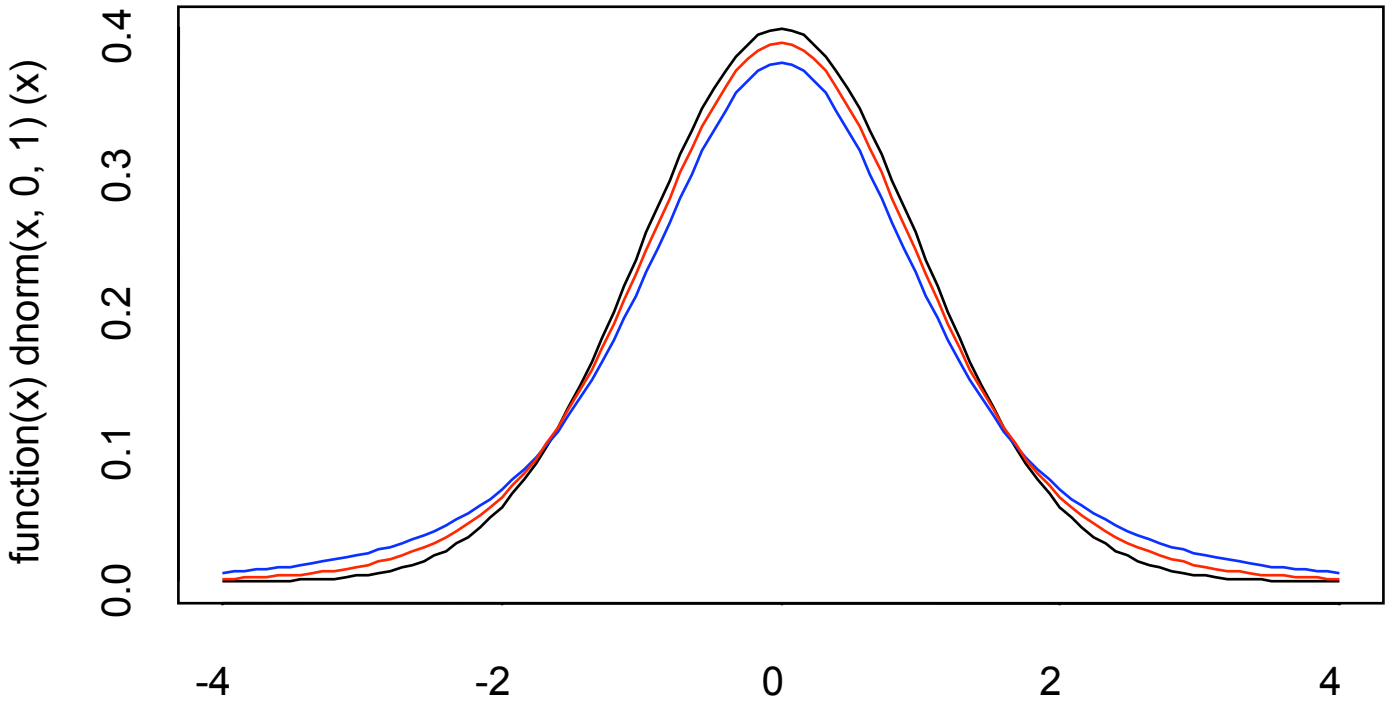
Je höher df, umso sicherer können wir sein, dass $\hat{\sigma} = \sigma$ und umso mehr nähert sich die t-Verteilung der Normalverteilung

Normalverteilung, $\mu = 0$, $\sigma = 1$.

```
curve(dnorm(x, 0, 1), -4, 4)
```

t-Verteilung, $\mu = 0$, $\sigma = 1$, df = 3

```
curve(dt(x, 3), -4, 4, add=T, col="blue")
```




```
curve(dt(x, 10), -4, 4, add=T, col="red")
```

B. Vertrauensintervall um $\mu = 6$

```
mu = 6
n = length(swerte)
SEhut = sd(swerte)/sqrt(n) # eingeschätzter SE
frei = n - 1 # Freiheitsgrade
mu  + SEhut * qt(0.025, frei) # untere Grenze
```

4.878858

```
mu  + SEhut * qt(0.975, frei) # obere Grenze
```

7.121142

C. Signifikant?

Auf der Basis dieser Stichprobe liegt μ zwischen 4.878858 und 7.121142 mit einer Wahrscheinlichkeit von 95%.

Frage: angenommen $\mu = 6$ sind die Werte überraschend?

```
mean(swerte)
[1] 6.75
```

Nein.

The two-sampled t-test

Meistens werden wir **2 Stichprobenmittelwerte** miteinander vergleichen wollen (und wesentlich seltener wie im vorigen Fall einen Stichprobenmittelwert, *m*, mit einem Bevölkerungsmittelwert, μ).

Zwei Händler, X und Y, verkaufen Äpfel am Markt.

Die Äpfel von Y sind teuerer, weil seine Äpfel mehr wiegen (behauptet Y).

Ich kaufe 20 Äpfel von X, 35 von Y. Ich wiege jeden Apfel und berechne:

	X	Y
Gewicht-Mittelwert	mx = 200	my = 220
Gewicht S-abweichung	sx = 20	sy = 30
Anzahl	nx = 20	ny = 35

Ist dieser Unterschied mx – my = 200 – 220 = – 20 g **signifkant**?

Hypothesen

H0: Es gibt keinen signifikanten Unterschied zwischen den Mittelwerten.

= die Wahrscheinlichkeit, dass der Unterschied zwischen diesen Mittelwerten 0 sein könnte ist mehr als 0.05 (kommt öfter als 5 Mal pro Hundert vor).

H1: Es gibt einen signifikanten Unterschied zwischen den Mittelwerten

= die Wahrscheinlichkeit, dass der Unterschied zwischen diesen Mittelwerten 0 sein könnte ist weniger als 0.05 (kommt seltener als 5 Mal pro Hundert vor).

Vorgang

Wir nehmen an, dass mx – my = -20 g **eine Stichprobe aus einer Normalverteilung ist**.

1. Wir müssen die Parameter μ, σ (und dann SE) dieser Normalverteilung **einschätzen.**

2. Wir erstellen ein 95% Vertrauensintervall fuer die t-Verteilung.

3. Wenn dieses Vertrauenintervall 0 einschließt, ist H0 akzeptiert (kein signifikanter Unterschied zwischen mx und my) sonst H1 (der Unterschied ist signifikant).

1. μ, SE einschätzen

Die beste Einschätzung von μ ist der Mittelwertunterschied unserer Stichprobe

Fuer diesen Fall mu = mx – my = – 20

1. SE einschätzen

Die beste Einschätzung von SE

$$\sqrt{\frac{(n_x-1)s_x^2+(n_y-1)s_y^2}{n_x+n_y-2}} \quad \text{x} \quad \sqrt{\frac{1}{n_x}+\frac{1}{n_y}}$$

X Y

Gewicht-Mittelwert	mx = 200	my = 220
Gewicht S-abweichung	sx = 20	sy = 30
Anzahl	nx = 20	ny = 35

Für diesen Fall, SEhut = 7.525339

Bitte in R-Befehle umsetzen und bestätigen.

```
nx = 20
ny = 35
sx = 20
sy = 30
z = ((nx - 1) * sx^2) + ((ny - 1) * sy^2)
nenn = nx + ny - 2
SEhut = sqrt(z/nenn) * sqrt(1/nx + 1/ny)
```

```
[1] 7.525339
```

95% Vertrauensintervall

df = nx + ny - 2

-20 - qt(0.025, df) * SEhut -4.906081

-20 + qt(0.025, df) * SEhut -35.09392

$\mu = -20$

SEhut =7.525339

Der Unterschied zwischen den Mittelwerten liegt zwischen -35.09392g und -4.906081g mit einer Wahrscheinlichkeit von 95%

Der Unterschied zwischen den Mittelwerten liegt zwischen -35.09392g und -4.906081g mit einer Wahrscheinlichkeit von 95%

Die Wahrscheinlichkeit, dass der Unterschied zwischen den Mittelwerten 0 sein könnte ist daher weniger als 5% (kommt weniger als 5 Mal pro 100 Stichproben vor).

Daher akzeptieren wir H1:

H1: Es gibt einen signifikanten Unterschied zwischen den Mittelwerten

Die benötigten Dauern (Minuten) an 9 Tagen im Winter in die Arbeit zu fahren sind:

20 15 19 22 17 16 23 18 20

Die entsprechenden Dauern an 11 Tagen im Sommer sind:

18 15 17 24 15 12 14 11 13 17 18

Ist der Unterschied zwischen den durchschnittlichen Sommer- und Winterzeiten signifikant ($p < 0.05$)?

Eine R-Funktion schreiben, **SE2(x,y)**, um

$$\sqrt{\frac{(n_x-1)s_x^2+(n_y-1)s_y^2}{n_x+n_y-2}} \quad \text{x} \quad \sqrt{\frac{1}{n_x}+\frac{1}{n_y}}$$

zu berechnen.

```
x = c(10, 15, 19, 9, 12, 8)
y = c(14, 11, 9, 10, 4, 4, 19, 10)
SE2(x, y)
```

```
[1]  2.502747
```

```
SE2 <- function(x, y)
{
nx = length(x)
ny = length(y)
sx = sd(x)
sy = sd(y)
num = ((nx - 1) * sx^2) + ((ny - 1) * sy^2)
den = nx + ny - 2
sqrt(num/den) * sqrt(1/nx + 1/ny)
}
```

```
x = c(20, 15, 19, 22, 17, 16, 23, 18, 20)
y = c(18, 15, 17, 24, 15, 12, 14, 11, 13, 17, 18)

# SE
SEhut =        = SE2(x,y)

  # μ
  d =          mean(x) - mean(y)

# Anzahl der Freiheitsgrade

  df =         length(x) + length(y) - 2

# Vertrauensintervall

  d - qt(0.025, df) * SEhut        d + qt(0.025, df) * SEhut

      [1] 6.110471                 [1] 0.03094282
```

Die t-test() Funktion

```
> t.test(x, y, var.equal=T)
```

Die Wahrscheinlichkeit, dass der Unterschied zwischen dem Mittelwert von x und dem Mittelwert von y gleicht 0 (Null)

95% Vertrauensintervall

```
data:  x and y
t = 2.1223, df = 18, p-value = 0.04794
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.03094282 6.11047132
sample estimates:
mean of x mean of y
 18.88889  15.81818
```

t=2.1233 bedeutet: die Werte 18.88889 und 15.81818 sind 2.1233 SEs voneinander entfernt

t-test () fortgesetzt

vowlax, vowlax.l, vowlax.spkr

Segmentliste, Etikettierungen, Sprecherlabels
(ungespannte deutsche Vokale)

Unterschieden sich die beiden Sprecher (67 und 68) in der Dauer der “E” Vokale?

```
temp = vowlax.l=="E"
dE = dur(vowlax[temp,])
lab = vowlax.spkr[temp]
```

Entweder konventionell

```
temp = lab == "67"
   x = dE[temp]
   y = dE[!temp]
   t.test(x, y, var.equal=T)
```

Oder, die Formel-Methode

```
t.test(dE ~ lab, var.equal=T)
```

Kriteria für eine t-test Durchführung

zwei Stichproben, x und y. Sind die Mittelwerte von x und y voneinander signifikant unterschiedlich?

```
pfad = "Das Verzeichnis, wo die Daten gespeichert ist"
mfdat = read.table(paste(pfad, "mfdur.txt", sep="/"))

x = mfdat[,1]

y = mfdat[,2]
```

Kriteria für eine t-test Durchführung

x und y

Sind x und y normalverteilt?

shapiro.test(x) shapiro.test(y)

ja nein

Sind die Varianzen von x und y voneinander signifikant unterschiedlich?

wilcox.test(x, y)

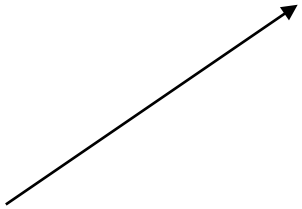
var.test(x, y)

ja nein

t.test(x,y) t.test(x,y, var.equal=T)

```
shapiro.test(y)

Shapiro-Wilk normality test
data:  y
W = 0.9866, p-value = 0.9037
```



Die Wahrscheinlichkeit, dass die Werte normalverteilt sind.

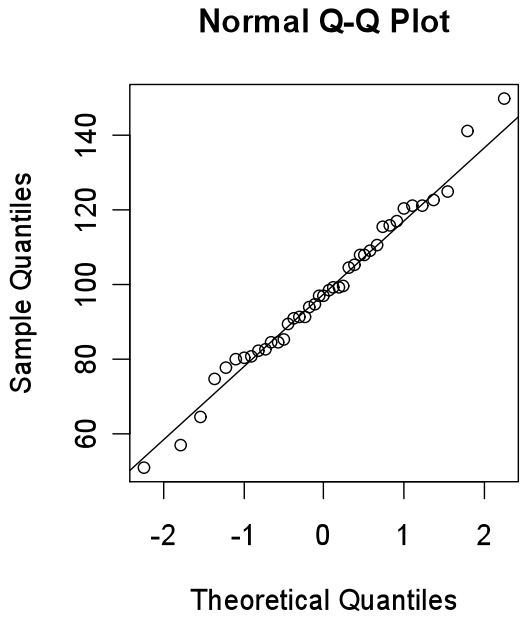
Wenn $p < 0.05$ dann weicht die Stichprobe signifikant von einer Normalverteilung ab, und der t-test soll nicht eingesetzt werden.

qqnorm()

Je mehr die Werte von der geraden Linie abweichen, umso unwahrscheinlicher ist es, dass die Werte einer Normalverteilung folgen.

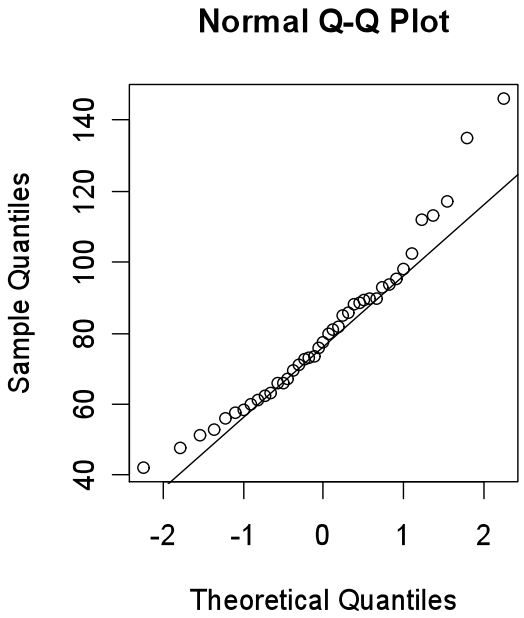
qqnorm(y)

qqline(y)



qqnorm(x)

qqline(x)



shapiro.test(y)

p-value = 0.9037

shapiro.test(x)

p-value = 0.08804

var.test()

prüft ob die Varianzen der beiden Stichproben voneinander signifikant abweichen.

Um signifikante Unterschiede zwischen Varianzen festzustellen, wird ein **F-test** und die **F-Verteilung** verwendet – diese Verteilung ist das gleiche wie die t-Verteilung hoch 2.

```
var(y)
[1] 428.9193
var(x)
[1] 516.3584
var(y)/var(x)
[1] 0.830662
var.test(y,x)

        F test to compare two variances

data:  x and y
F = 0.8307, num df = 40, denom df = 40, p-value = 0.5601
alternative hypothesis: true ratio of variances is not equal to
1
95 percent confidence interval:
 0.4429731 1.5576553
sample estimates:
ratio of variances
          0.830662
```

Der Unterschied zwischen den Varianzen ist nicht signifikant (genauer: das Verhältnis zwischen den Varianzen weicht nicht signifikant ab von 1).

F(40, 40) = **0.83**, p > 0.05

Wenn keine Normalverteilung

Wilcoxon Rank Sum and Signed Rank Tests (Mann-Whitney test)

```
wilcox.test(y, x)
```

```
        Wilcoxon rank sum test with continuity correction
data:  x and y
W = 1246, p-value = 0.0001727
alternative hypothesis: true location shift is not equal to 0
```

Der Unterschied zwischen x und z ist signifikant.
(Wilcoxon rank sum test, W = 1246, $p < 0.001$)

Normalverteilung, Varianzen sind unterschiedlich

t.test(y, x)

Welch Two Sample t-test
data: x and y
t = 3.6947, **df = 79.321**, **p-value = 0.0004031**
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  8.183973 27.297539
sample estimates:
mean of x mean of y
 97.95751 80.21676

Der Unterschied zwischen x und y ist signifikant (t = 3.69, df = 79.3, p < 0.001)

…sonst t.test(y,x, var.equal=T)