

Die t-Verteilung und die Prüfstatistik

Jonathan Harrington

Standard error of the mean (SE)

ist die Standardabweichung von Mittelwerten

Ich werfe 5 Würfel und berechne den Mittelwert der Zahlen

$\mu = 3.5$ 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.

$$SE = \text{sigma()}/\text{sqrt}(5)$$

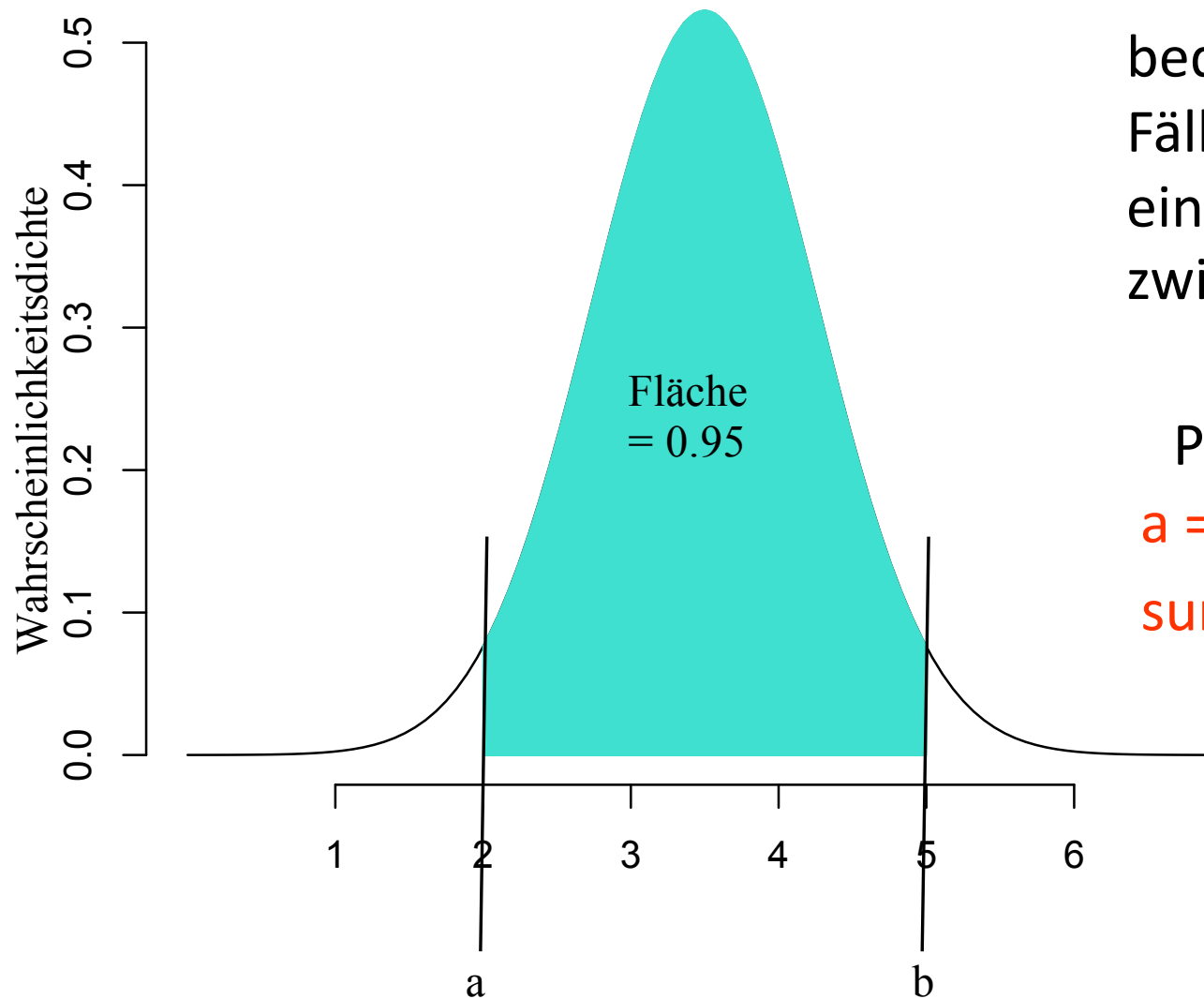
Standard error of the mean (SE)

SE wird kleiner, umso größer n .

$$SE = \frac{\sigma}{\sqrt{n}} \quad \text{umso größer } n, \text{ umso weniger weicht } m \text{ von } \mu \text{ ab.}$$

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 $m = \mu = 3.5$.

95% Konfidenzintervall



bedeutet: in 95/100
Fälle erwarte ich
einen Mittelwert
zwischen 2 und 5.

Probieren

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

```
a = qnorm(0.025, mu, SE)
```

2.003053

```
b = qnorm(0.975, mu, SE)
```

4.996947

Berechnungen wenn σ unbekannt ist

1. SE muss eingeschätzt werden
2. Verwendung der t-Verteilung statt der Normalverteilung

σ ist unbekannt

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

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

werte

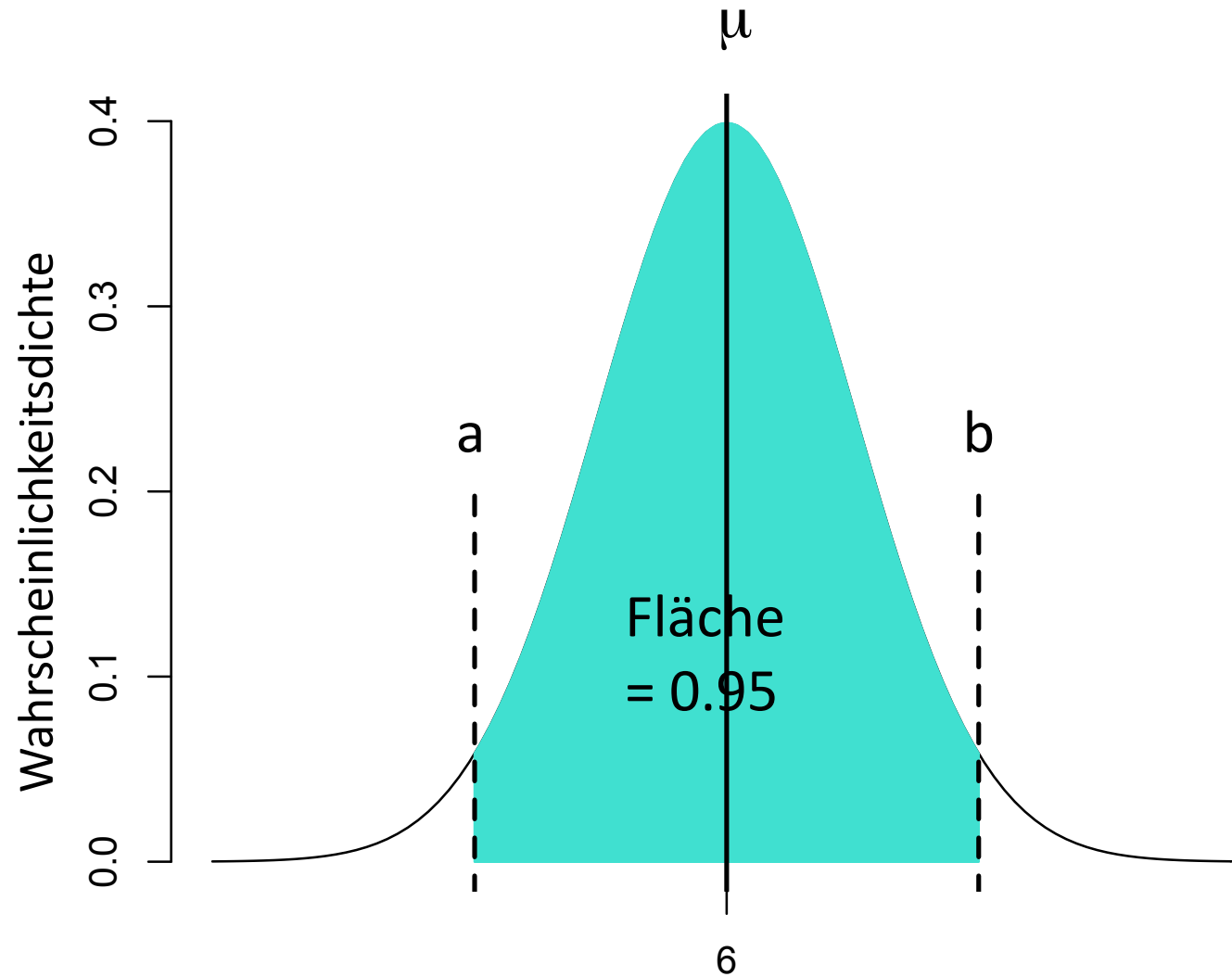
[1] 6 5 6 9 6 5 6 8 5 6 10 9

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% Konfidenzintervalls von μ ?**).

Das Verfahren: a one-sampled t-test

Fällt m außerhalb des 95% Konfidenzintervalls von μ ?
= kommt 6.75 zwischen a und b vor?



1. Einschätzung von SE

Die beste Einschätzung von SE ist **die Standardabweichung der Stichprobe, s** :

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

In R:

`SE = sd(werte)/sqrt(length(werte))`

2. die t-Verteilung

Wenn die Bevölkerungs-Standardabweichung **eingeschätzt** werden muss, dann wird das Konfidenzintervall 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 Konfidenzintervall sind etwas breiter (dies berücksichtigt, die zusätzliche Unsicherheit wegen der Einschätzung von SE).

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

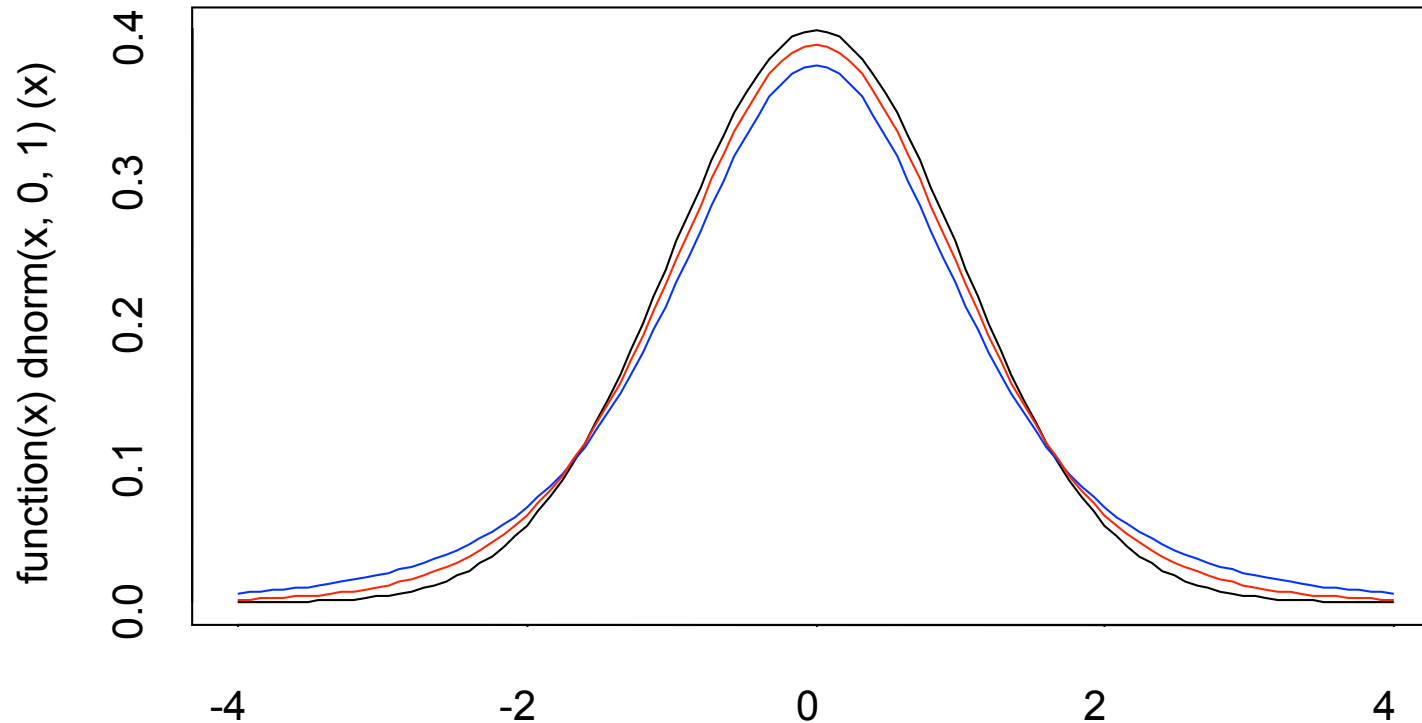
Je höher df, umso sicherer können wir sein, dass $\hat{SE} = SE$ und umso mehr nähert sich die t-Verteilung der Normalverteilung

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

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

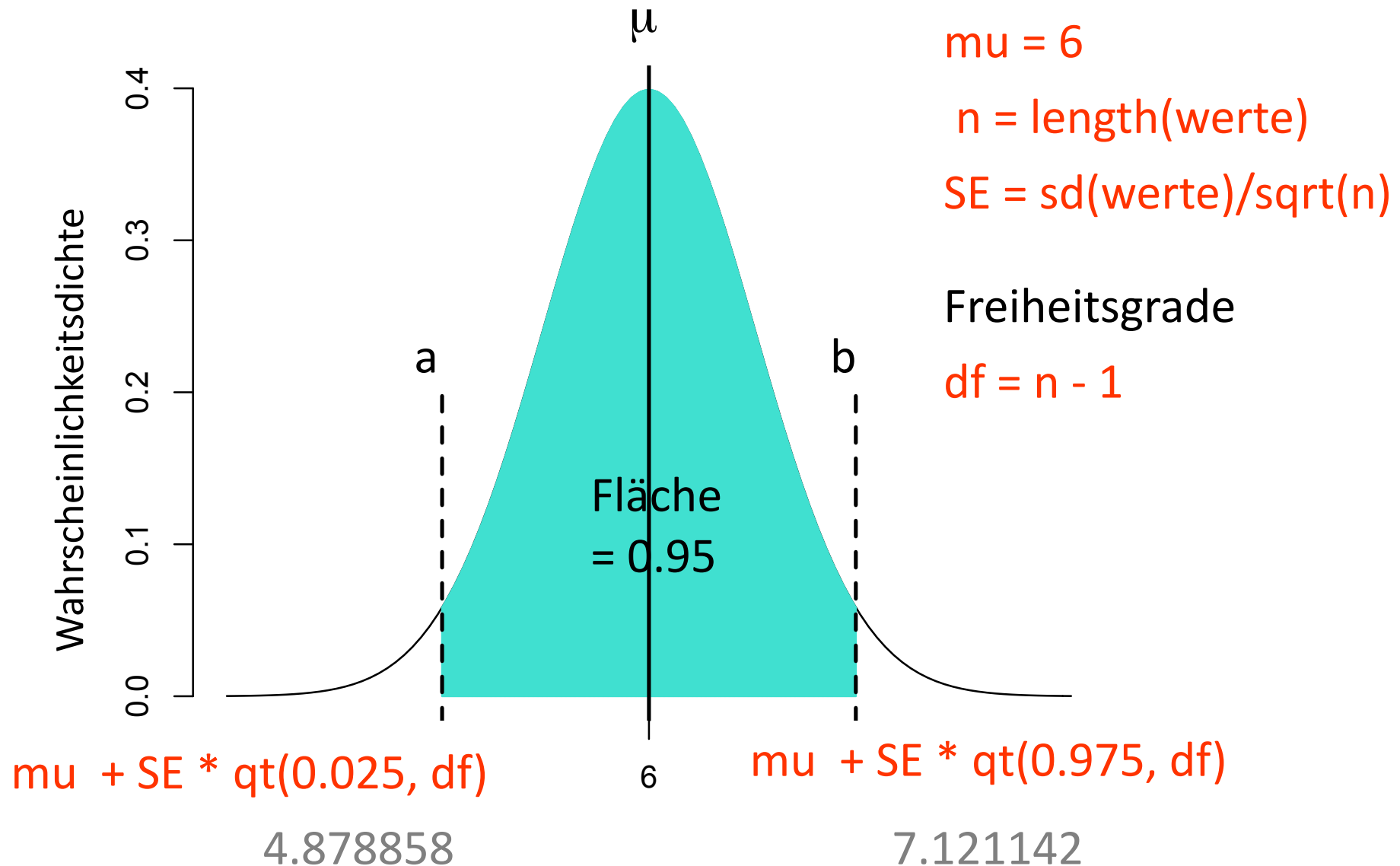
t-Verteilung, $\mu = 0$, $SE = 1$, $df = 3$

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



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

Fällt m außerhalb des 95% Konfidenzintervalls von μ ?
= kommt 6.75 zwischen a und b vor?



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

Frage: angenommen $\mu = 6$ ist der Stichprobenmittelwert $m = 6.75$ überraschend?

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).

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

Was ist die Wahrscheinlichkeit, dass die Jahreszeit einen Einfluss auf die Dauern hat?

$x = c(20, 15, 19, 22, 17, 16, 23, 18, 20)$

$y = c(18, 15, 17, 24, 15, 12, 14, 11, 13, 17, 18)$

Was ist die Wahrscheinlichkeit, dass die Jahreszeit einen Einfluss auf die Dauern hat?

= Was ist die Wahrscheinlichkeit, dass der Unterschied zwischen den Mittelwerten der beiden Stichproben von 0 (Null) abweicht?

95% Konfidenzintervall um die **Mittelwertunterschiede** berechnen.

$$\mu = \text{mean}(x) - \text{mean}(y)$$

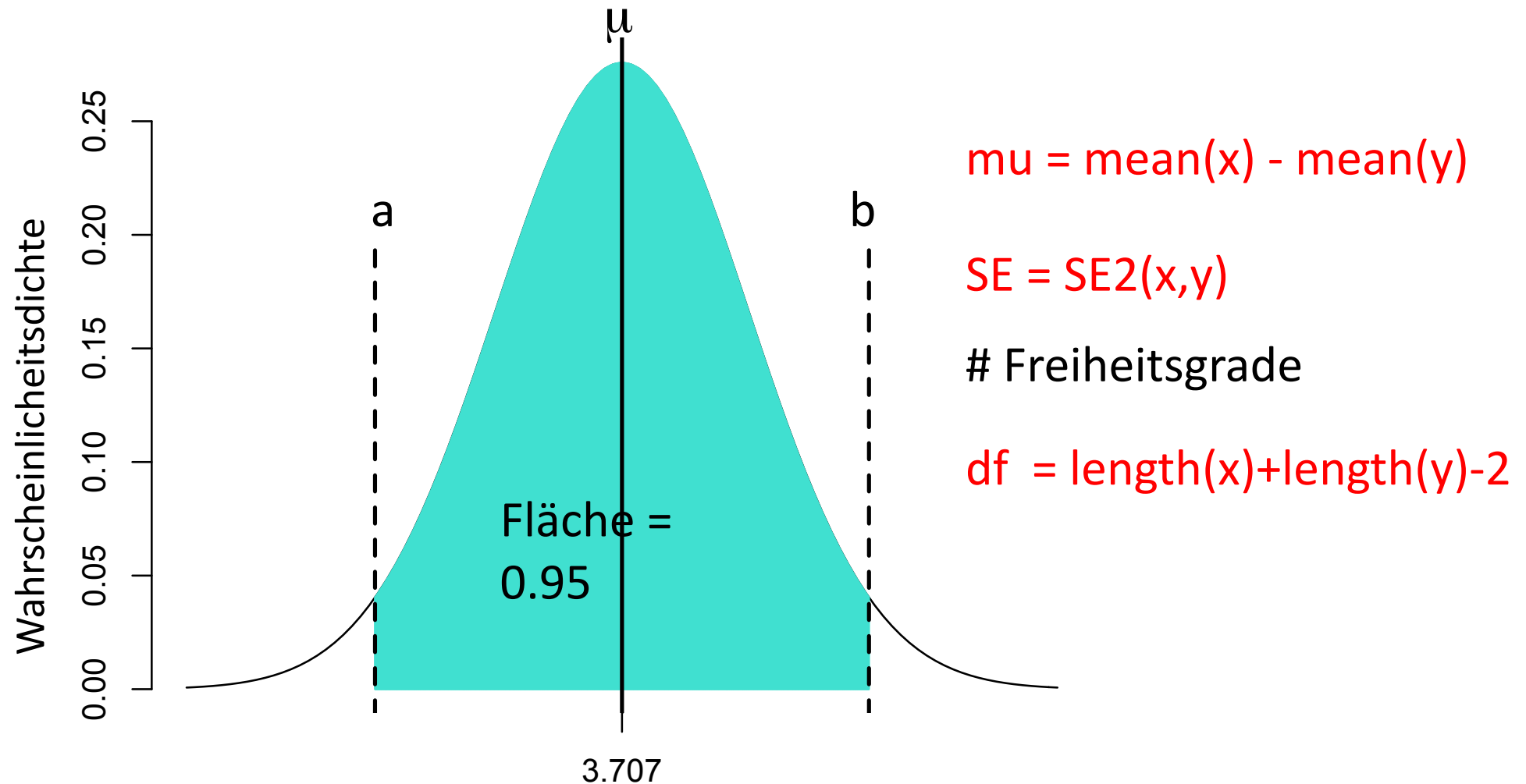
Kommt 0 (Null) innerhalb dieses Konfidenzintervalls vor?

$\hat{SE}$ der Mittelwertunterschiede

$$\hat{SE} = \sqrt{\frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}} \times \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$$

$$SE = SE2(x, y)$$

Was ist die Wahrscheinlichkeit, dass sich die Mittelwerte der Winter- und Sommerzeiten unterscheiden?



Kommt 0 (Null) zwischen a und b vor?

$$a = \mu + qt(0.025, df) * SE$$

[1] 0.03094282

$$b = \mu + qt(0.975, df) * SE$$

[1] 6.110471

Der Unterschied zwischen den Mittelwerten
liegt zwischen 0.03 und 6.11 mit einer
Wahrscheinlichkeit von 0.95

Die Wahrscheinlichkeit, dass sich die Mittelwerte
nicht unterscheiden ist weniger als 5%.

Der t-test() Funktion

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

prüft die Wahrscheinlichkeit, dass $\mu = 0$

mu/SE bedeutet: 0 und μ sind 2.12 SEs voneinander entfernt

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

Freiheitsgrade

Die Wahrscheinlichkeit, dass $\mu = 0$, ist 0.04794

95% Konfidenzintervall für μ

Die Jahreszeit hat einen signifikanten Einfluss auf die Dauer ($t_{18} = 2.1$, $p < 0.05$). Oder: $t_{18} = 2.1$, $p < 0.05$

Die t-test() Funktion: Formel-Methode

```
xlab = rep("winter", length(x))  
ylab = rep("sommer", length(y))  
jahreszeit = factor(c(xlab, ylab))  
d = c(x, y)
```

```
d.df = data.frame(dauer = d, J = jahreszeit)
```

```
t.test(dauer ~ J, var.equal=T, data=d.df)
```

Kriterien für eine t-test Durchführung

```
mfdur = read.table(file.path(pfad, "mfdur.txt"))
```

```
head (mfdur)
```

	duration	Gender
1	115.250	F
2	74.687	F
3	124.813	F

Hat Gender einen Einfluss auf die Dauer? (was ist die Wahrscheinlichkeit, dass der Unterschied zwischen den Dauermittelwerten von M und F = 0?)

Kriterien für eine t-test Durchführung

Sind die Verteilungen pro Stufe normalverteilt?

`shapiro.test()`

ja

nein

`wilcox.test()`

Unterschiede in der Varianz?

`var.test()`

nein (Default)

`t.test()`

ja

`t.test(..., var.equal=T)`

```
with(mfdur, tapply(duration, Gender, shapiro.test))
```

Stufe F des Faktors Gender

\$F

Shapiro-Wilk normality test

data: X[[1L]]

W = 0.9866, p-value = 0.9037

Die Wahrscheinlichkeit,
dass die Werte
normalverteilt sind.

\$M

Shapiro-Wilk normality test

data: X[[2L]]

W = 0.9528, p-value = 0.08804

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

```
var.test()
```

prüft ob das Verhältnis zwischen Varianzen signifikant von 1 abweicht.

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.test(duration ~ Gender, data = mfdur)
```

```
      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 $F(40, 40) = 0.8, p > 0.05$

(Das Verhältnis zwischen den Varianzen weicht nicht signifikant ab von 1.)

Wenn keine Normalverteilung

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

```
wilcox.test(duration ~ Gender, data = mfdur)
```

```
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
```

Gender hat einen signifikanten Einfluss auf die Dauer
(Wilcoxon rank sum test, $p < 0.001$)

Normalverteilung, Varianzen sind unterschiedlich

```
t.test(duration ~ Gender, data = mfdur)
```

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

Gender hat einen signifikanten Einfluss auf die Dauer ($t_{79.3} = 3.7$, $p < 0.001$). Oder $t_{79.3} = 3.7$, $p < 0.001$

...sonst

```
t.test(duration ~ Gender, var.equal=T, data = mfdur)
```

Beispiel. t-test Fragen, Frage 1(a)

```
tv = read.table(file.path(pfad, "tv.txt"))  
head(tv)  
with(tv, table(V))
```

```
# boxplot  
boxplot(d ~ V, data=tv)
```

```
# Prüfen, ob sie einer Normalverteilung folgen  
with(tv, tapply(d, V, shapiro.test))  
# alles OK
```

```
# Prüfen, ob sich die Varianzen unterscheiden  
var.test(d ~ V, data=tv)  
# Die Varianzen unterscheiden sich signifikant. Daher:
```

```
t.test(d ~ V, data = tv)
```

Der Vokal hatte einen signifikanten Einfluss auf die Dauer ($t[12.5] = 4.3$, $p < 0.001$)

Data-Frame **dr**

(a) Hat Position einen Einfluss auf F1?

(b) Hat Dialekt einen Einfluss auf F1?

(c) Hat Position einen Einfluss auf die Dauer?