Wahrscheinlichkeitsrechnung und Statistik für Biologen

2. Der Standardfehler

Dirk Metzler

24. März 2026

Inhaltsverzeichnis

1 Der Standardfehler

1.1 Ein Versuch

Versuchsaufbau: 14 Hirse-Pflanzen von einer Sorte wurden 7 Tage lang nicht mehr gegossen („trockengestresst“).

An den letzten drei Tagen wurde die Wasserabgabe der Pflanzen durch Wägung ermittelt und ein Mittelwert über drei Tage errechnet.

Zum Schluß des Versuchs wurden die Pflanzen abgeschnitten und die Blattfläche bestimmt.

$$\text{Transpirationsrate} = (\text{Wasserabgabe pro Tag})/\text{Blattfläche} \left[\frac{\text{ml}}{\text{cm}^2 \cdot \text{Tag}}\right]$$

Ein Ziel des Versuchs: die mittlere Transpirationsrate zu bestimmen. μ
(für diese Hirsesorte unter diesen Bedingungen)

In einem großen Versuch mit sehr vielen Pflanzen könnte man μ beliebig genau bestimmen.

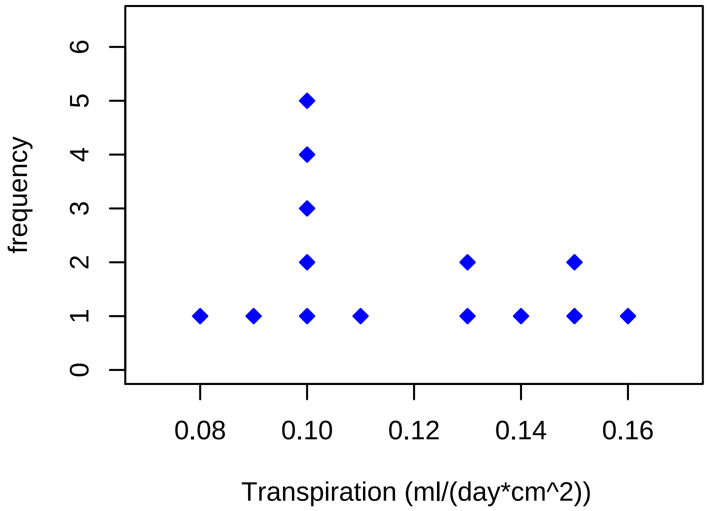
FRAGE: Wie genau ist die Schätzung von μ in diesem kleinen ($n = 14$) Versuch?

Beispiel inspiriert durch:

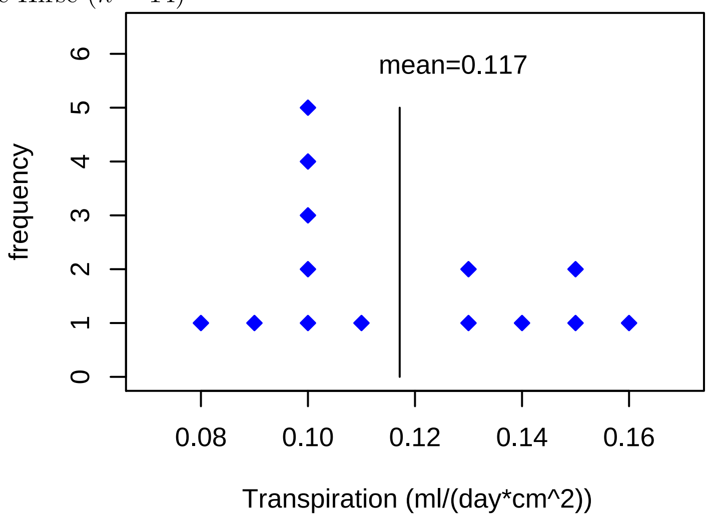
Literatur

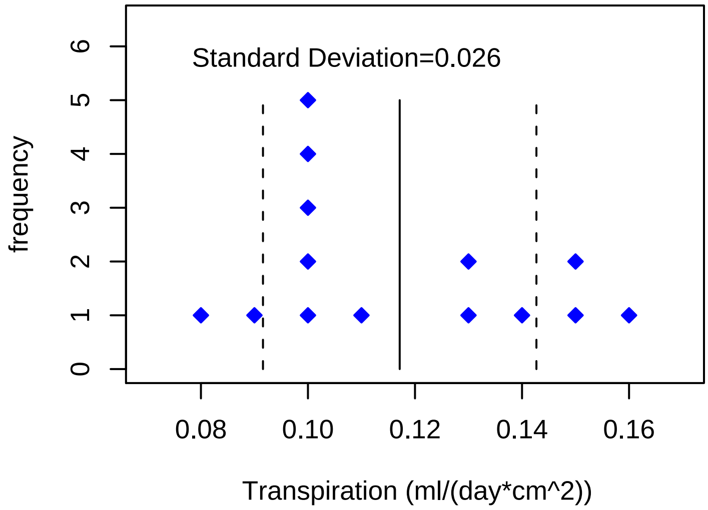
[BB05] V. Beyel and W. Brüggemann. Differential inhibition of photosynthesis during pre-flowering drought stress in Sorghum bicolor genotypes with different senescence traits. *Physiologia Plantarum*, 124:249–259, 2005.

Ergebnisse des Versuchs



Trockengestresste Hirse ($n = 14$)





Transpirationsdaten: $x_1, x_2, \ldots, x_{14}$ $\overline{x} = \big(x_1 + x_2 + \cdots + x_{14}\big)/14 \quad = \frac{1}{14}\sum_{i=1}^{14} x_i$

$$\overline{x} = 0{,}117$$

Unsere Schätzung: $\mu \approx 0{,}117$

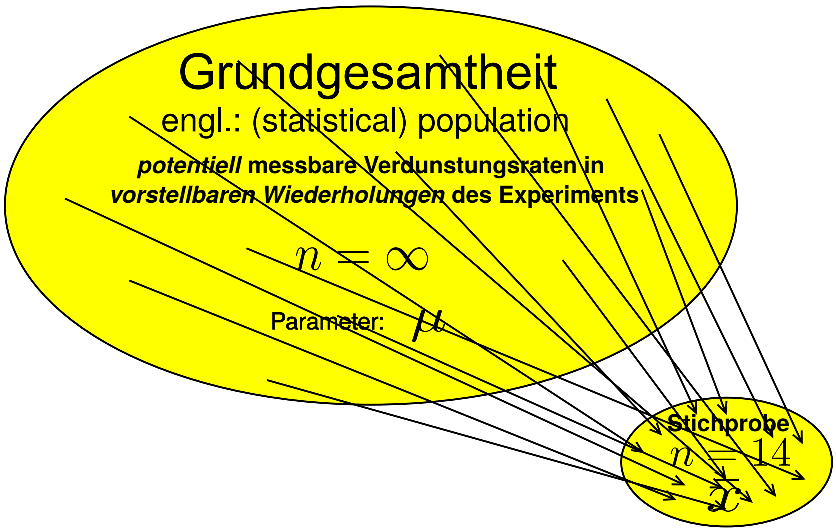
Wie genau ist diese Schätzung? Wie weit weicht der Schätzwert $\overline{x}$ von dem wahren Mittelwert μ ab?

1.2 Allgemeine Überlegung

Allgemeine Überlegung

Wir stellen uns vor, wir hätten den Versuch nicht 14 mal, sondern 100 mal, 1.000 mal, 1.000.000 mal wiederholt.

Unsere 14 Transpirationswerte betrachten wir als
zufällige Stichprobe
aus dieser großen Population von möglichen Werten.



Wir schätzen den Populationsmittelwert μ durch den Stichprobenmittelwert $\overline{x}$.

μ ist ein Parameter.

$\overline{x}$ ist eine Statistik.

Jede neue Stichprobe liefert einen neuen Wert von $\overline{x}$. $\overline{x}$ hängt vom Zufall ab: eine *Zufallsgröße* FRAGE: Wie variabel ist $\overline{x}$? Genauer: Wie weit weicht $\overline{x}$ typischerweise von μ ab?

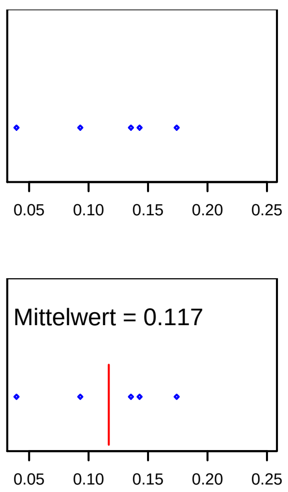
$$\overline{x} = \big(x_1 + x_2 + \cdots + x_n\big)/n$$

Wovon hängt die Variabilität von $\overline{x}$ ab?

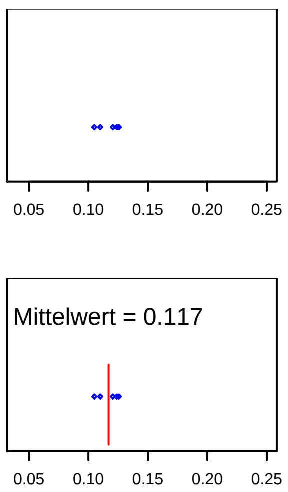
1.

von der Variabilität der einzelnen Beobachtungen

$x_1, x_2, \ldots, x_n$



x variiert viel
$\Rightarrow$ $\overline{x}$ *variiert viel*



x variiert wenig
$\Rightarrow$ $\overline{x}$ *variiert wenig*

2.

vom Stichprobenumfang
n

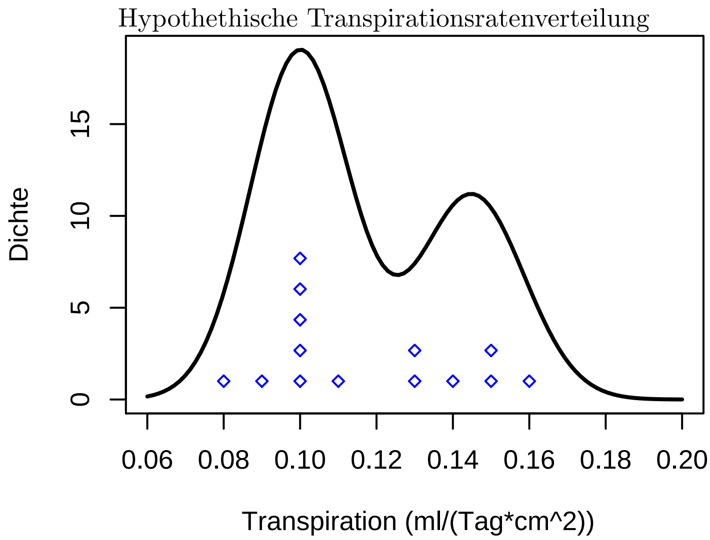
Je größer n, desto kleiner
die Variabilität von $\overline{x}$.

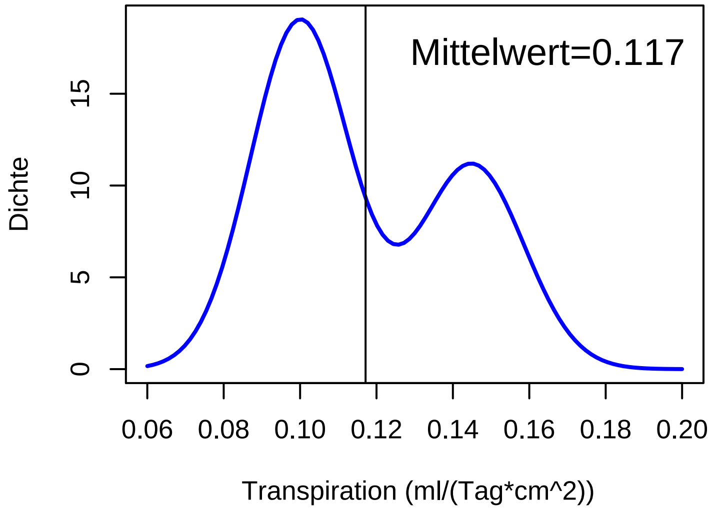
Um diese Abhängigkeit zu untersuchen, machen wir ein
(Computer-)Experiment.

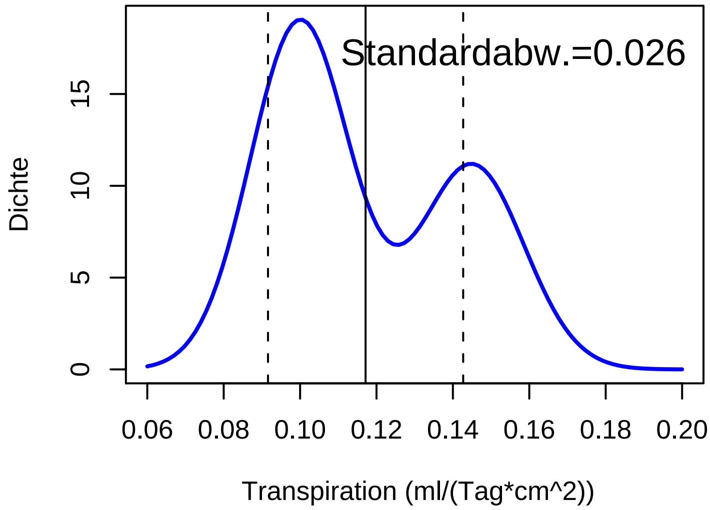
Experiment:

Wir nehmen eine Population, ziehen Stichproben, und schauen, wie $\overline{x}$ variiert.

Nehmen wir an, die Verteilung aller möglichen Transpirationswerte sieht folgendermaßen aus:



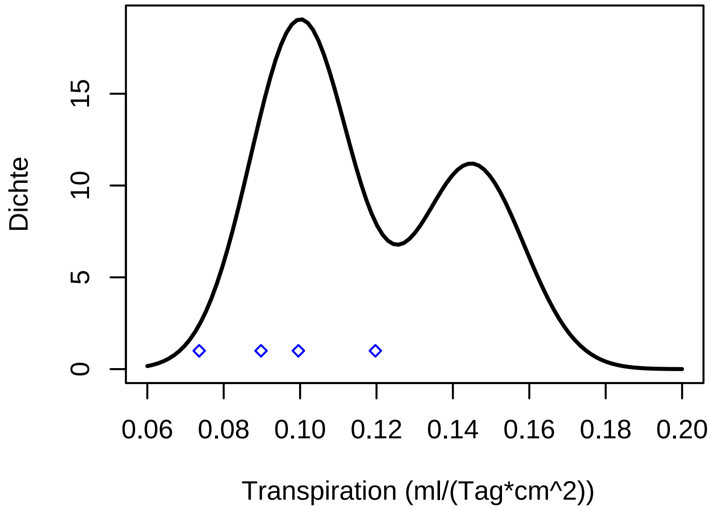


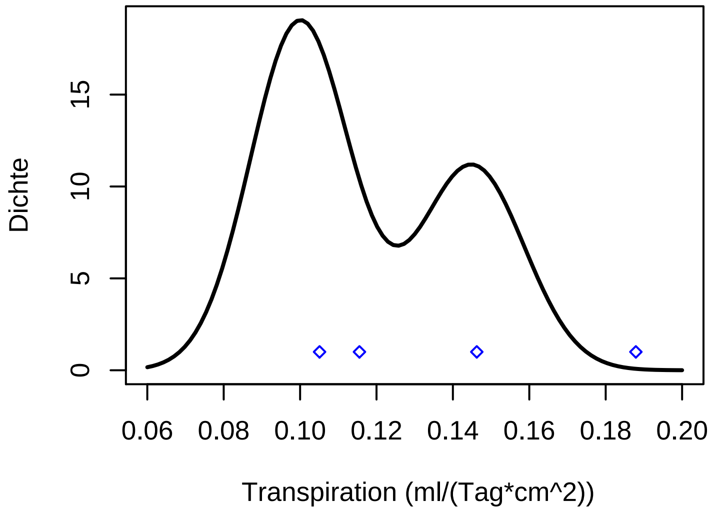


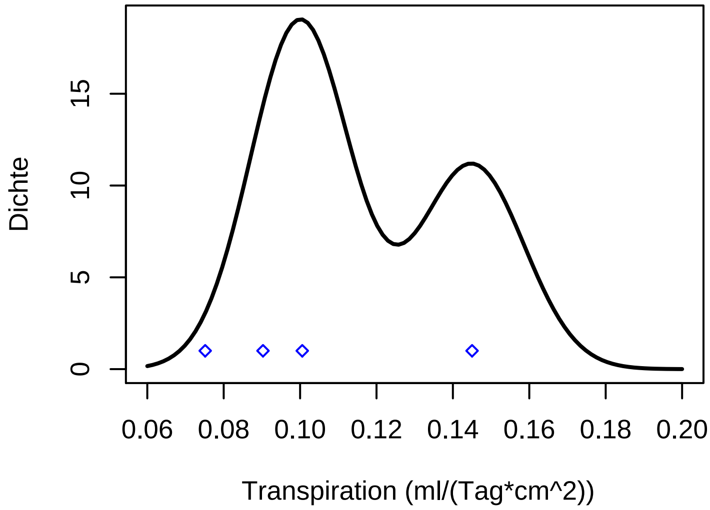
Wir beginnen mit kleinen Stichproben:

$$n = 4$$

Eine Stichprobe vom Umfang 4 Eine zweite Stichprobe vom Umfang 4 Eine dritte Stichprobe vom Umfang 4

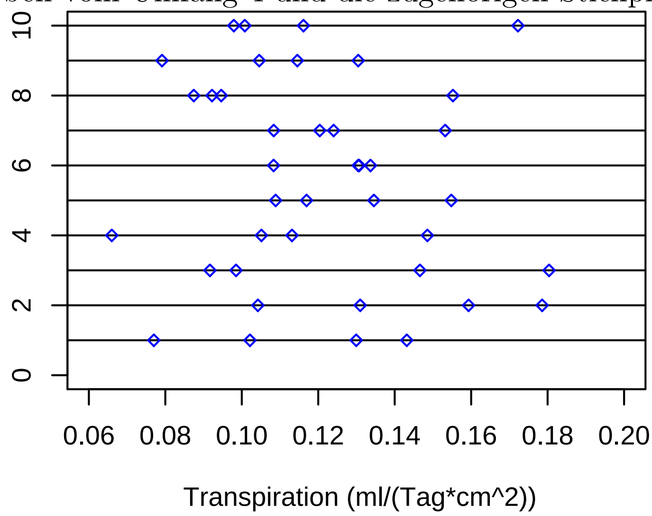

Dichte
0
5
10
15
0.06
0.08
0.10
0.12
0.14
0.16
0.18
0.20
Transpiration (ml/(Tag*cm^2))

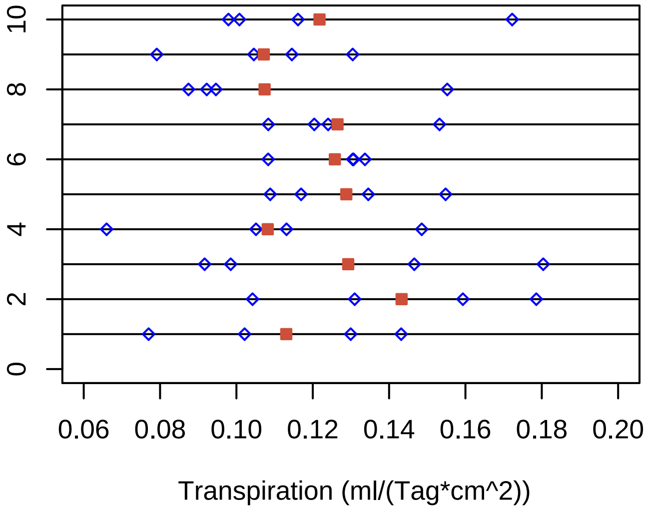

Dichte
0
5
10
15
0.06
0.08
0.10
0.12
0.14
0.16
0.18
0.20
Transpiration (ml/(Tag*cm^2))



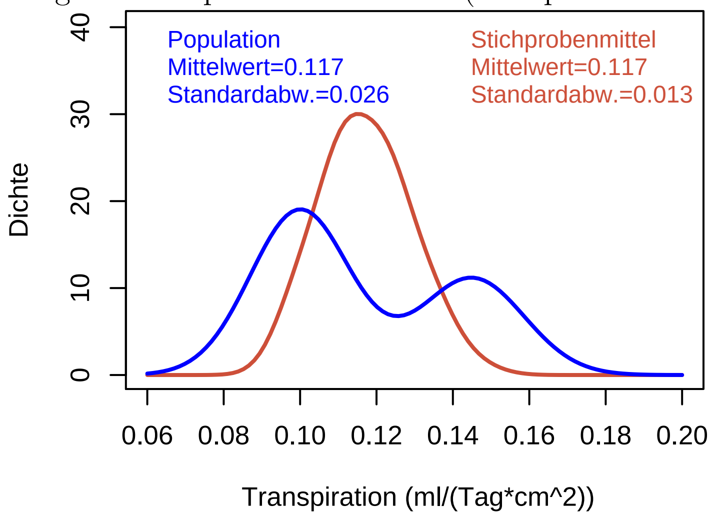
Wie variabel sind die Stichprobenmittelwerte?

10 Stichproben vom Umfang 4 und die zugehörigen Stichprobenmittel





Verteilung der Stichprobenmittelwerte (Stichprobenumfang $n = 4$)

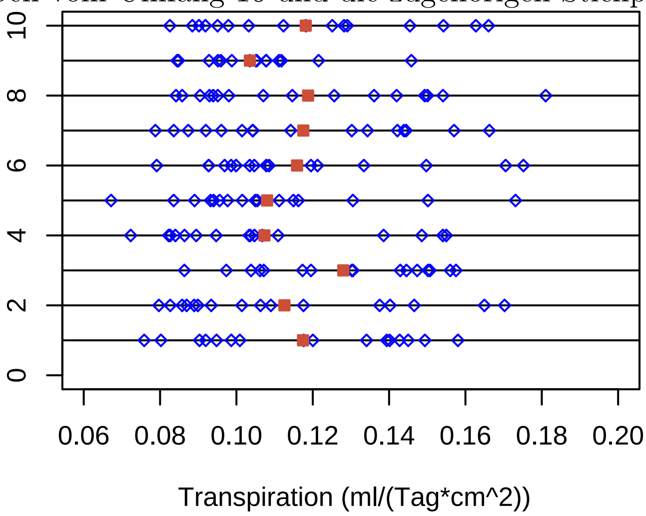


Population: Standardabweichung $= 0{,}026$

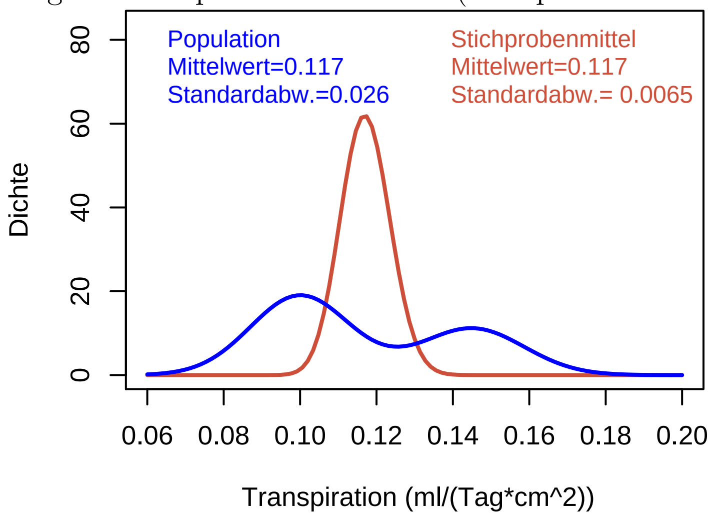
Stichprobenmittelwerte ($n = 4$): $\begin{aligned} Standardabweichung &= 0{,}013 \\ &= 0{,}026/\sqrt{4} \end{aligned}$

Erhöhen wir den Stichprobenumfang von 4auf 16

10 Stichproben vom Umfang 16 und die zugehörigen Stichprobenmittel



Verteilung der Stichprobenmittelwerte (Stichprobenumfang $n = 16$)



Population: Standardabweichung = 0,026

Stichprobenmittelwerte ($n = 16$): $$\begin{aligned} Standardabweichung &= 0{,}0065 \\ &= 0{,}026/\sqrt{16} \end{aligned}$$

Die allgemeine Regel

Die Standardabweichung des Mittelwerts einer Stichprobe vom Umfang n ist $1/\sqrt{n}$ mal der Standardabweichung der Population.

Die Standardabweichung der Population bezeichnet man mit σ *(sigma)*.

Die Regel schreibt man häufig so:

$$\sigma(\overline{x}) = \frac{1}{\sqrt{n}}\sigma(X)$$

In der Praxis ist der Wert des Parameters σ unbekannt.

Es wird durch eine Statistik, nämlich die korrigierte Stichproben-Standardabweichung

$$s = \sqrt{\tfrac{1}{n-1}\textstyle\sum_i (x_i - \overline{x})^2},$$

geschätzt:

$$\sigma = ??$$

$$\sigma \approx s$$

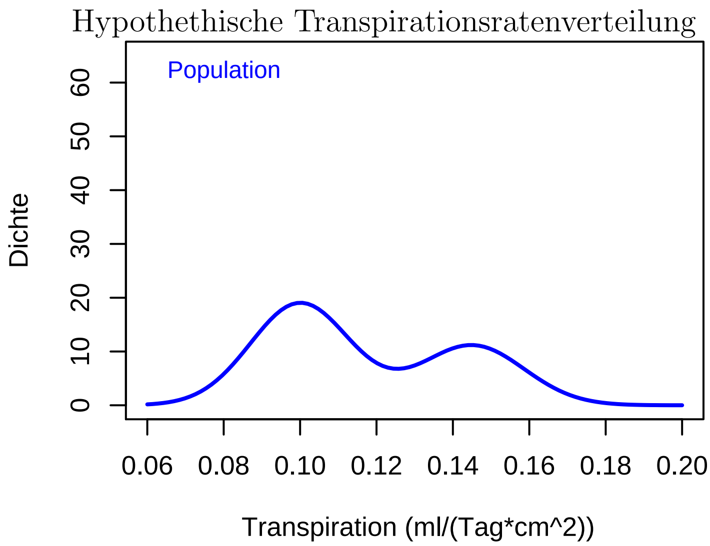
Die geschätzte Standardabweichung von $\overline{x}$ $s/\sqrt{n}$ nennt man den Standardfehler. (Englisch: *standard error of the mean*, *standard error*, *SEM*)

$$\text{SEM} = \frac{s}{\sqrt{n}} = \frac{\sqrt{\frac{1}{n-1}\sum_i (x_i-\overline{x})^2}}{\sqrt{n}} = \sqrt{\frac{\sum_i (x_i-\overline{x})^2}{n\cdot(n-1)}}$$

1.3 Zur Verteilung von $\overline{x}$

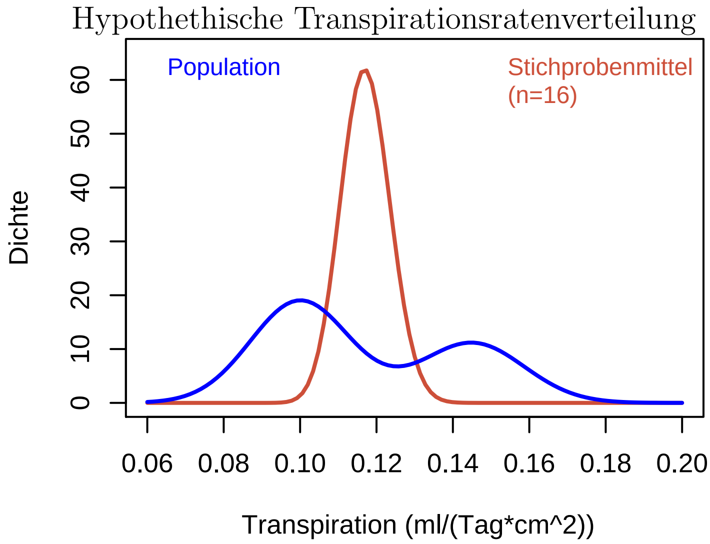
Wir haben gesehen:

Auch wenn die Verteilung von x mehrgipflig & asymmetrisch ist



ist die Verteilung von $\overline{x}$ trotzdem (annähernd) eingipflig & symmetrisch

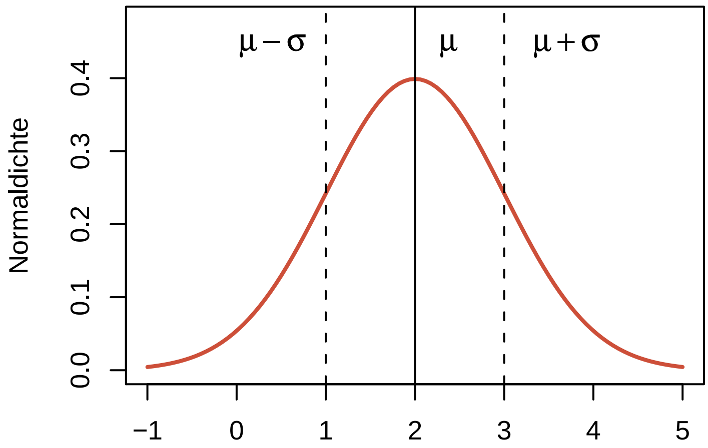
(wenn der Stichprobenumfang n nur groß genug ist)



Die Verteilung von $\overline{x}$ hat annähernd eine ganz bestimmte Form:

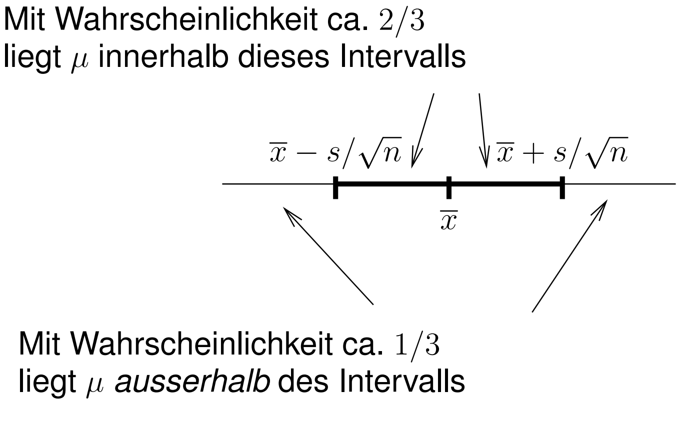
die Normalverteilung.

Dichte der Normalverteilung



Die Normalverteilungsdichte heisst auch *Gauß'sche Glockenkurve* (nach Carl Friedrich Gauß, 1777-1855)

Wichtige Folgerung



Demnach:

Es kommt durchaus vor, dass $\overline{x}$ von μ um mehr als $s/\sqrt{n}$ abweicht.

1.4 Anwendungen

ANWENDUNG 1: Welche Werte von μ sind plausibel?

$$\overline{x} = 0{,}12$$
$$s/\sqrt{n} = 0{,}007$$

Frage: Könnte es sein, dass $\mu = 0{,}115$?

Antwort: Es ist gut möglich.

Abweichung $\overline{x} - \mu = 0{,}120 - 0{,}115 = 0{,}005$.

Standardfehler $s/\sqrt{n} = 0{,}007$

Abweichungen dieser Größe kommen häufig vor.

(Die Frage, welche Abweichungen *nicht* mehr plausibel sind, untersuchen wir später.)

ANWENDUNG 2: Vergleich von Mittelwerten Beispiel: Springkrebs

Galathea: Carapaxlänge in einer Stichprobe

Männchen: $\overline{x}_1 = 3{,}04$ mm $s_1 = 0{,}9$ mm $n_1 = 25$

Weibchen: $\overline{x}_2 = 3{,}23$ mm $s_2 = 0{,}9$ mm $n_2 = 29$

Die Weibchen scheinen größer zu sein.

Ist das ernst zu nehmen?
Oder könnte es nur *Zufall* sein?

Wie genau sind die beiden Mittelwerte?

Männchen: $\overline{x}_1 = 3{,}04$ mm $s_1 = 0{,}9$ mm $n_1 = 25$

$$s_1/\sqrt{n_1} = 0{,}18 \text{ [mm]}$$

Mit Schwankungen von $\pm 0{,}18$ (mm) in $\overline{x}_1$ müssen wir rechnen.

Wie genau sind die beiden Mittelwerte?

Weibchen: $\overline{x}_2 = 3{,}23$ mm $s_2 = 0{,}9$ mm $n_2 = 29$

$$s_2/\sqrt{n_2} = 0{,}17 \text{ [mm]}$$

Es ist nicht unwahrscheinlich, dass $\overline{x}_2$ um mehr als $\pm 0{,}17$ (mm) vom wahren Mittelwert abweicht.

Die Differenz der Mittelwerte

$$3{,}23 - 3{,}04 = 0{,}19 \text{ [mm]}$$

ist kaum größer als die zu erwartenden Schwankungen.

Es könnte also einfach *Zufall* sein, dass $\overline{x}_2 > \overline{x}_1$

GENAUER FORMULIERT:

Wenn in Wirklichkeit die Populationsmittelwerte gleich sind, $\mu_{Weibchen} = \mu_{Männchen}$ kann es trotzdem leicht passieren, dass die Stichprobenmittelwerte $\overline{x}_2$ und $\overline{x}_1$ so weit auseinander liegen.

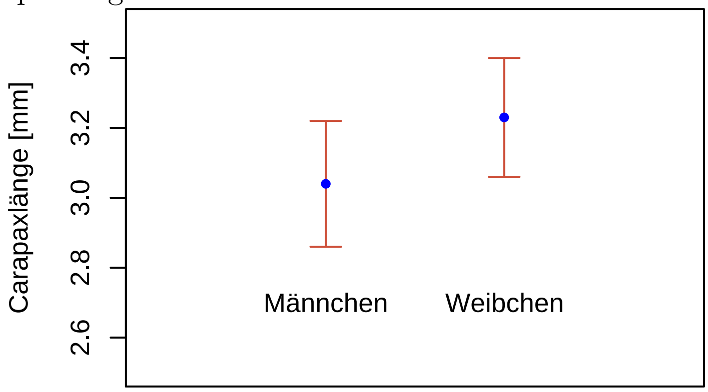
Der Statistiker sagt: Die Differenz der Mittelwerte ist (statistisch) *nicht signifikant.*

nicht signifikant = könnte gut Zufall sein

ANWENDUNG 3:

Wenn man Mittelwerte graphisch darstellt, sollte man auch ihre Genauigkeit
$(\pm s/\sqrt{n})$
anzeigen.

Carapaxlängen: Mittelwerte ± Standardfehler nach Geschlecht



ANWENDUNG 4: Bei der Versuchsplanung: Wieviele Beobachtungen brauche ich? (Wie groß sollte die Stichprobenlänge n sein?)

Wenn man weiß, welche Genauigkeit (Standardfehler $s/\sqrt{n}$) für $\overline{x}$ man erreichen will,

und wenn man (aus Erfahrung oder aus einem Vorversuch) s ungefähr kennt, dann kann man das notwendige n ungefähr abschätzen: $s/\sqrt{n} = g$ (g = gewünschter Standardfehler)

Beispiel: Gestresste Transpirationswerte bei einer anderen Hirse-Sorte: $\overline{x} = 0{,}18$ $s = 0{,}06$ $n = 13$

Nehmen wir an, der Versuch soll wiederholt werden und man will Standardfehler $\approx 0{,}01$ erreichen.

Wie groß sollte n sein?

Lösung: gewünscht: $s/\sqrt{n} \approx 0{,}01$ aus dem Vorversuch wissen wir:

$$\begin{aligned} s &\approx 0{,}06 \\ \sqrt{n} &\approx \frac{0,06}{0,01} = 6 \\ n &\approx 36 \end{aligned}$$

1.5 Zusammenfassung

ZUSAMMENFASSUNG

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ.
- Aus dieser Population ziehen wir eine Zufallsstichprobe vom Umfang n, mit Stichprobenmittelwert $\overline{x}$.
- $\overline{x}$ ist eine Zufallsgröße mit Mittelwert μ und Standardabweichung $\sigma/\sqrt{n}$.
- Man schätzt die Standardabweichung von $\overline{x}$ mit $s/\sqrt{n}$, wobei s die *korrigierte* Stichprobenabweichung ist.
- $s/\sqrt{n}$ nennt man den *Standardfehler*.
- Schwankungen in $\overline{x}$ von der Größe $s/\sqrt{n}$ kommen häufig vor. Solche Schwankungen sind *„nicht signifikant“:* sie könnten gut Zufall sein.

Was Sie u.a. erklären können sollten

- Wieso ist der (Stichproben-)Mittelwert $\overline{x}$ eine Zufallsvariable?
- Verteilungseigenschaften von $\overline{x}$
- Was genau ist mit “Standardfehler” gemeint?
- Was ist der Unterschied zwischen dem Standardfehler und...
  - ...der (Stichproben-)Standardabweichung?
  - ...der Standardabweichung des Mittelwerts?
- Wieso muss man beim Berechnen des Standardfehlers aus Daten durch $n-1$ (oder $\sqrt{n-1}$) und zusätzlich durch n (oder $\sqrt{n}$) teilen?
- Anwendungen des Standardfehlers bei der deskriptiven Datenanalyse und der Versuchsplanung.