

Wahrscheinlichkeitsrechnung und
Statistik für Biologen
2. Der Standardfehler

Martin Hutzenthaler & Dirk Metzler

27. April 2010

Inhaltsverzeichnis

1	Eine kurze Wiederholung zur deskriptiven Statistik	2
1.1	Mittelwert und Standardabweichung	2
1.2	Ein Beispiel zur Warnung	3
2	Der Standardfehler	4
2.1	Ein allgemeiner Rahmen	4
2.2	Zur Verteilung von \bar{x}	12
2.3	Anwendungen	15
2.4	Zusammenfassung	19

1 Eine kurze Wiederholung zur deskriptiven Statistik

1.1 Mittelwert und Standardabweichung

Bachstelzen fressen Dungfliegen

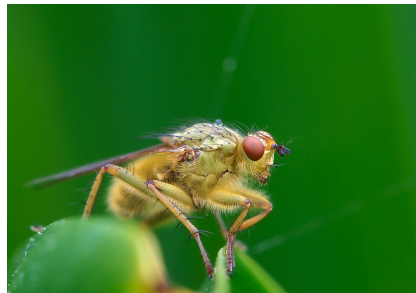
Räuber



Bachstelze (White Wagtail)
Motacilla alba alba

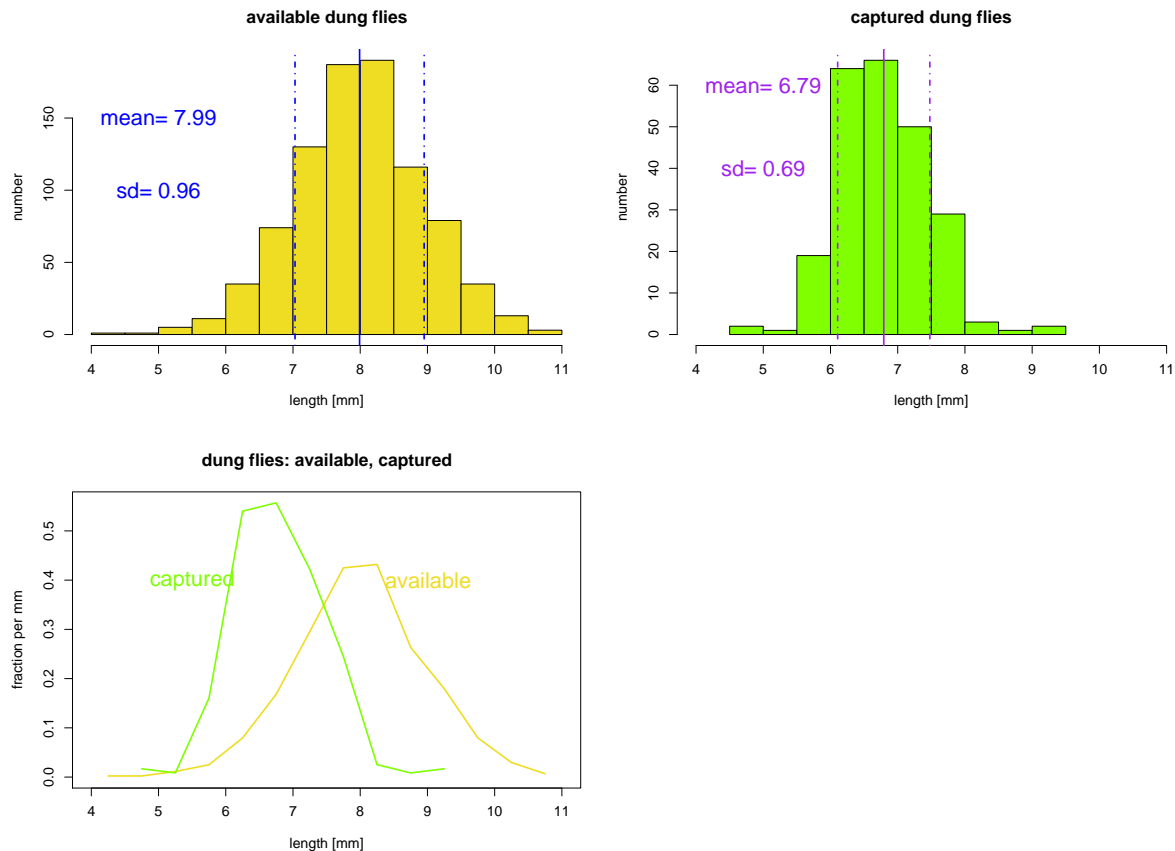
image (c) by Artur Mikolajewski

Beute



Gelbe Dungfliege
Scatophaga stercoraria

image (c) by Viatour Luc



Beobachtung

In der Biologie sind viele Datenverteilungen annähernd glockenförmig und können durch den **Mittelwert** und die **Standardabweichung** angemessen beschrieben werden.

Es gibt aber auch Ausnahmen.

1.2 Ein Beispiel zur Warnung

Kupfertolerantes Rotes Straußgras



Rotes Straußgras
Agrostis tenuis

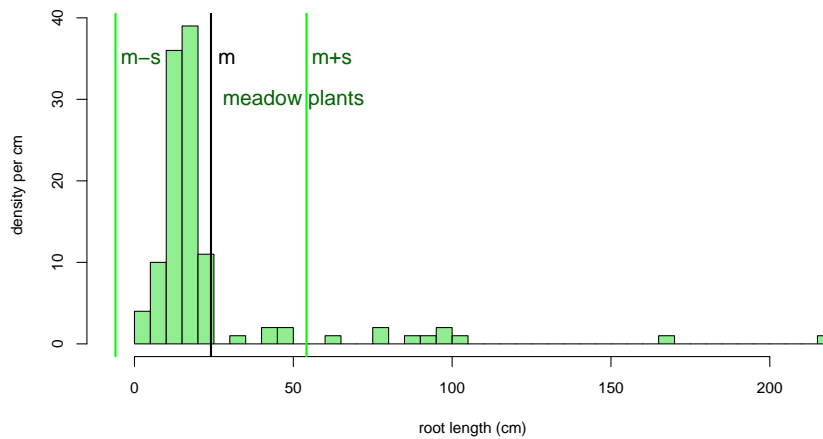
image (c) Kristian Peters



Kupfer
Cuprum

Hendrick met de Bles

Browntop Bent (n=50)

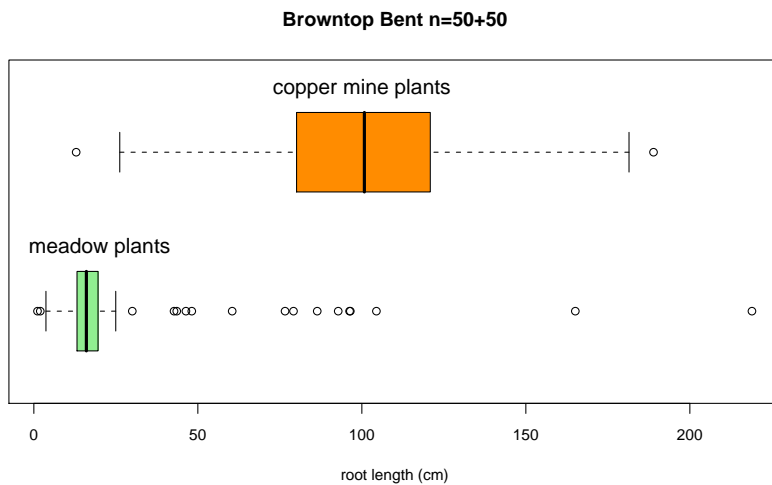


In etwa 2/3 der Wurzellängen innerhalb $[m-s, m+s]$???? **Nein!** Deutlich mehr.

Manche Verteilungen können nur mit mehr als zwei Variablen angemessen beschrieben werden.

z.B. mit den fünf Werten der Boxplots:

\min , Q_1 , median , Q_3 , \max



Schlussfolgerung

In der Biologie sind viele Datenverteilungen annähernd glockenförmig und können durch den [Mittelwert](#) und die [Standardabweichung](#) angemessen beschrieben werden.

Es gibt aber auch Ausnahmen. Also:
Immer die Daten erst mal graphisch untersuchen!

Verlassen sie sich **niemals** allein auf numerische Kenngrößen!

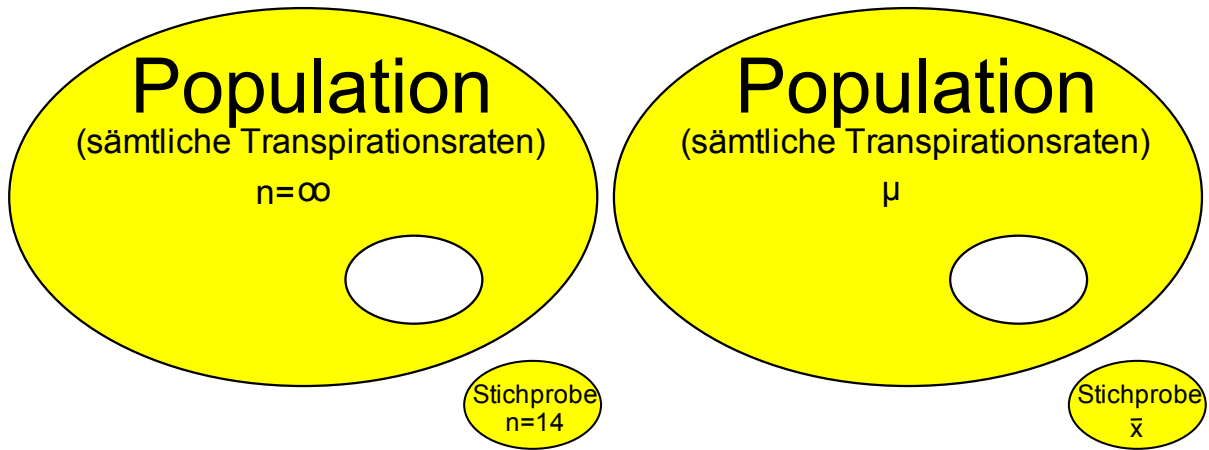
2 Der Standardfehler

2.1 Ein allgemeiner Rahmen

Ein allgemeiner Rahmen

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
 \bar{x} .

Jede neue Stichprobe liefert einen neuen Wert von \bar{x} .
 \bar{x} hängt vom Zufall ab: eine *Zufallsgröße*

FRAGE: Wie variabel ist \bar{x} ?

Genauer: Wie weit weicht \bar{x} typischerweise von μ ab?

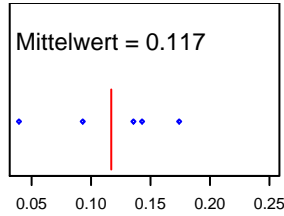
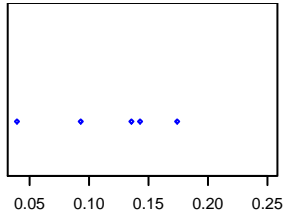
$$\bar{x} = (x_1 + x_2 + \dots + x_n)/n$$

Wovon hängt die Variabilität von \bar{x} ab?

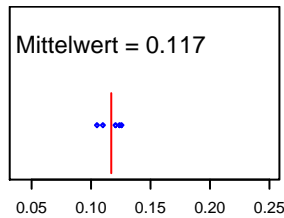
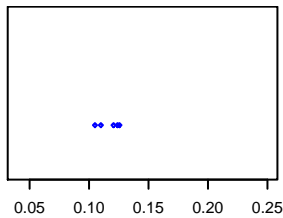
1.

von der Variabilität der einzelnen Beobachtungen

$$x_1, x_2, \dots, x_n$$



x variiert viel
 $\Rightarrow \bar{x}$ variiert viel



x variiert wenig
 $\Rightarrow \bar{x}$ variiert wenig

2.

vom Stichprobenumfang

n

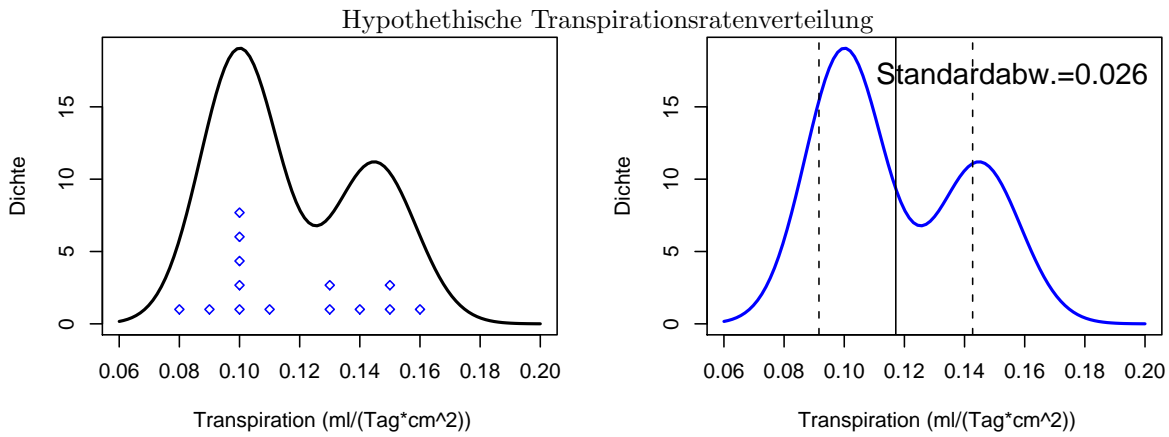
Je größer n , desto kleiner
 die Variabilität von \bar{x} .

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

Experiment:

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

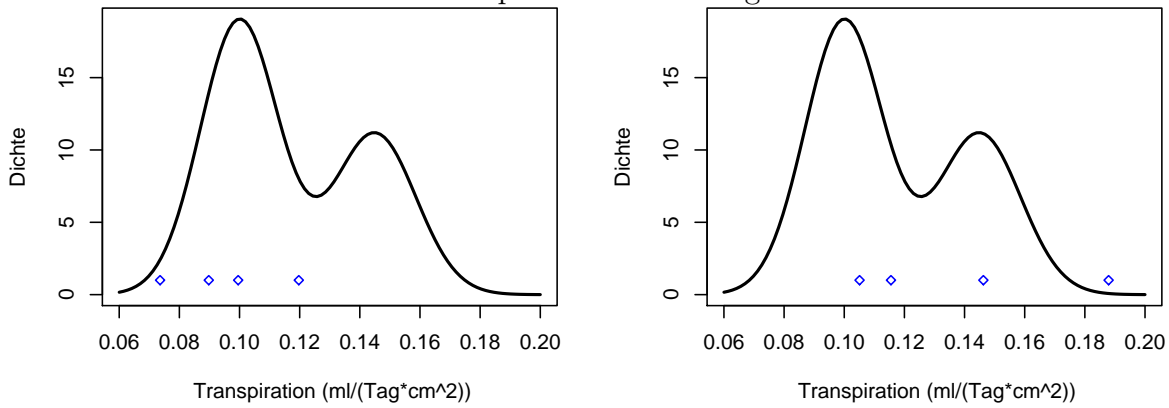
Nehmen wir an, die Verteilung aller möglichen Transpirationswerte sieht folgendermaßen
 aus:



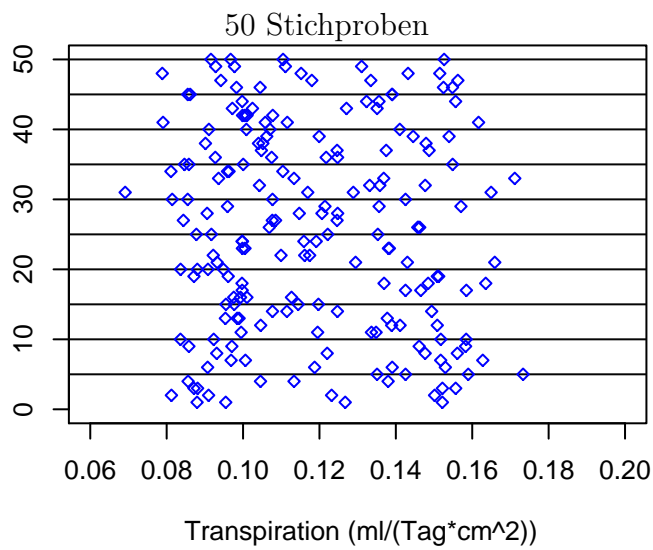
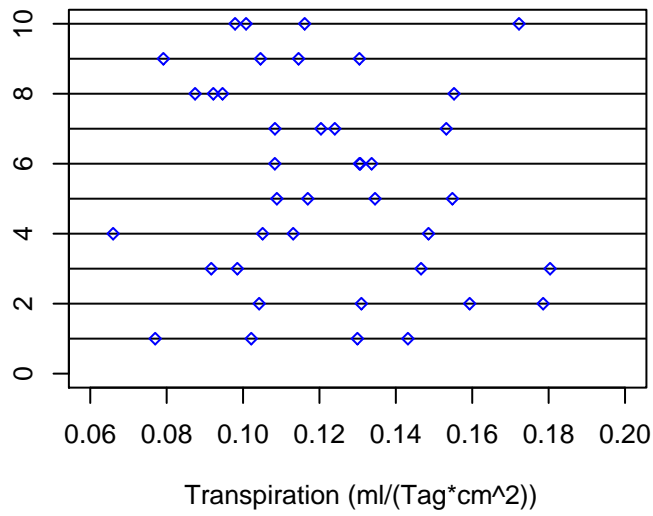
Wir beginnen mit kleinen Stichproben:

$$n = 4$$

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

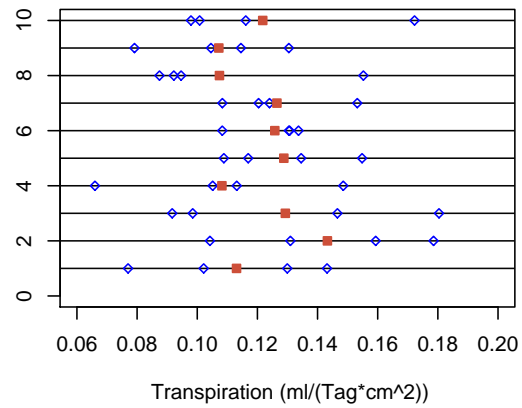
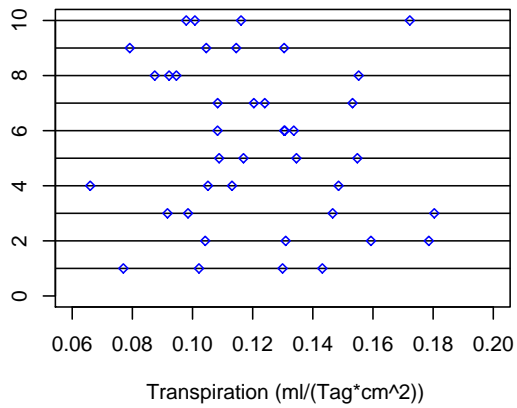


10 Stichproben

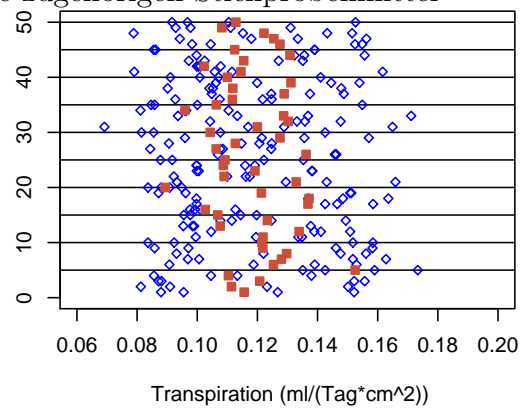
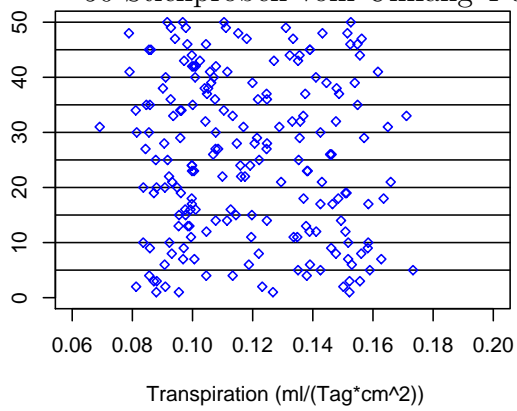


Wie variabel sind
die Stichprobenmittelwerte?

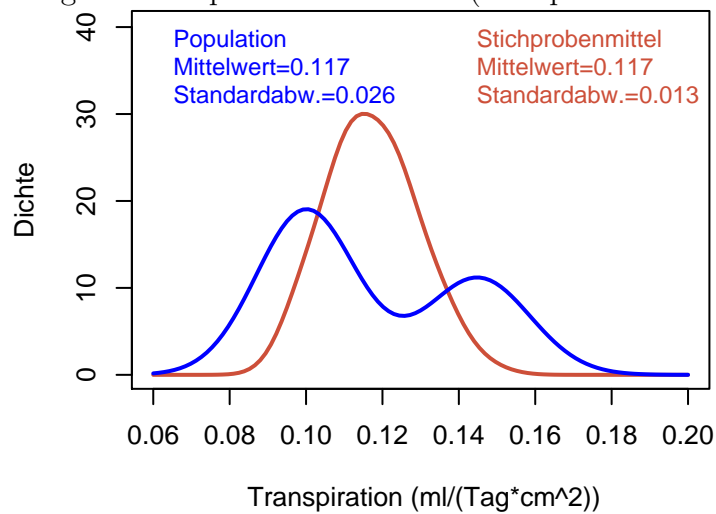
10 Stichproben vom Umfang 4 und die zugehörigen Stichprobenmittel



50 Stichproben vom Umfang 4 und die zugehörigen Stichprobenmittel



Verteilung der Stichprobenmittelwerte (Stichprobenumfang $n = 4$)

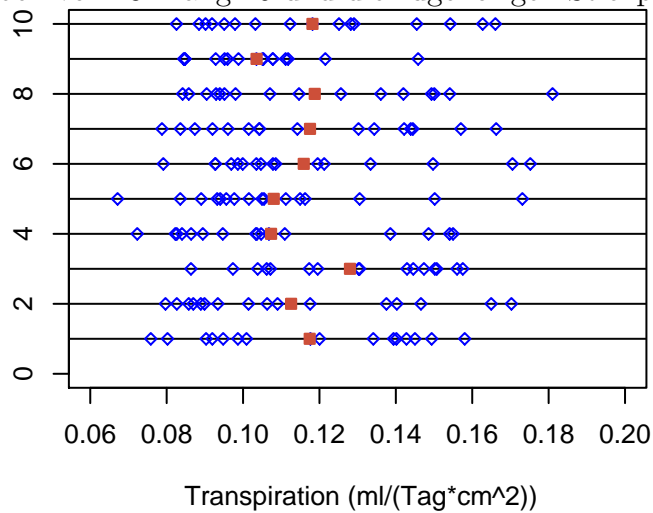


Population: Standardabweichung = 0,026

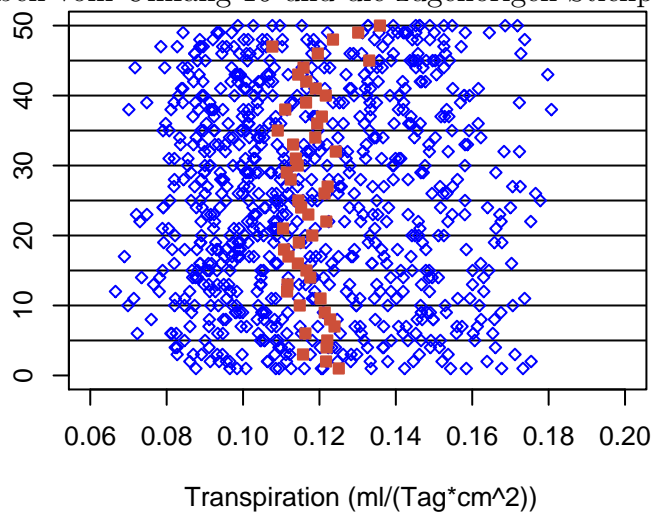
Stichprobenmittelwerte ($n = 4$): $Standardabweichung = 0,013$
 $= 0,026/\sqrt{4}$

Erhöhen wir den Stichprobenumfang von 4 auf 16

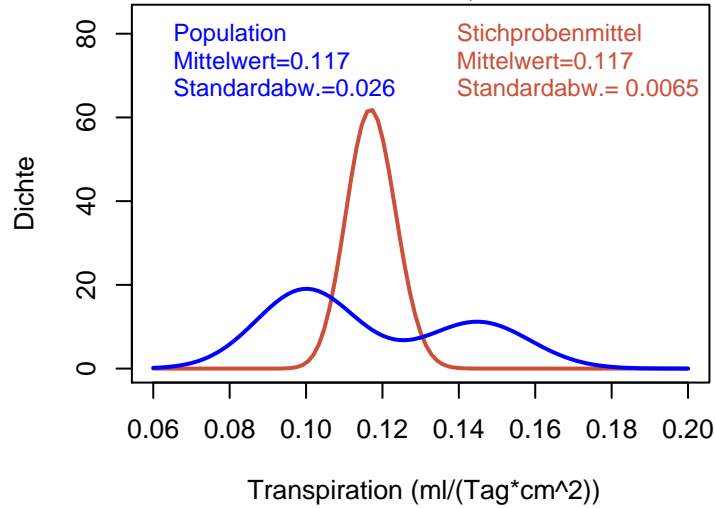
10 Stichproben vom Umfang 16 und die zugehörigen Stichprobenmittel



50 Stichproben vom Umfang 16 und die zugehörigen Stichprobenmittel



Verteilung der Stichprobenmittelwerte (Stichprobenumfang $n = 16$)



Population: Standardabweichung = 0,026

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

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:

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

In der Praxis ist σ unbekannt.

Es wird durch die Stichproben-Standardabweichung s geschätzt:

$$\sigma = ??$$

$$\sigma \approx s$$

Die geschätzte Standardabweichung von \bar{x}

$$s/\sqrt{n}$$

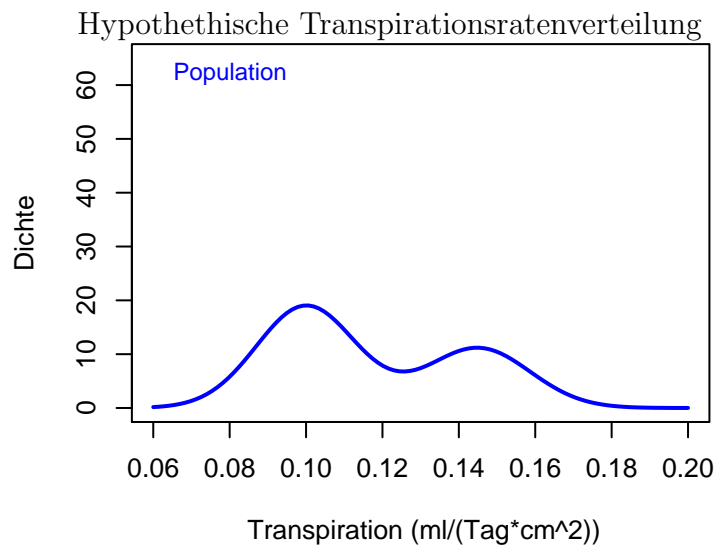
nennt man den
Standardfehler.

(Englisch: *standard error of the mean, standard error, SEM*)

2.2 Zur Verteilung von \bar{x}

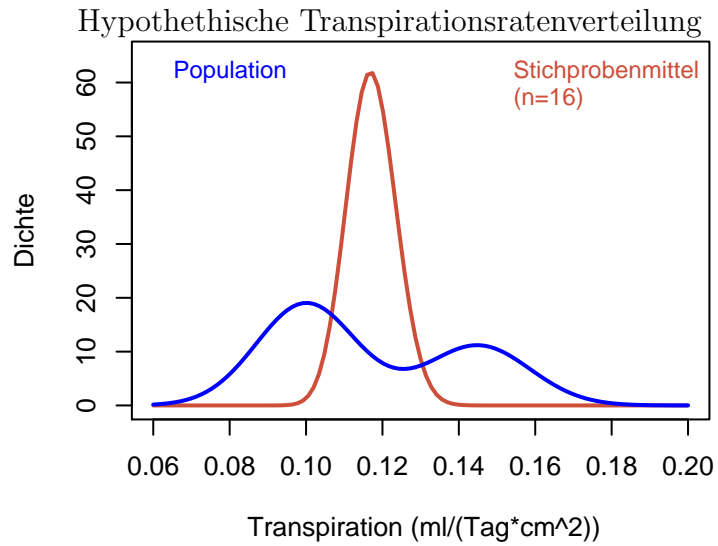
Wir haben gesehen:

Auch wenn die Verteilung von x mehrgipfelig & asymmetrisch ist



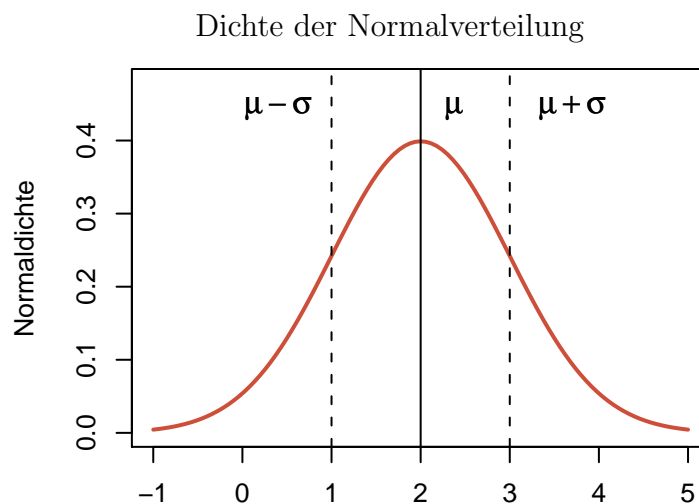
ist die Verteilung von \bar{x} trotzdem (annähernd) eingipflig & symmetrisch

(wenn der Stichprobenumfang n nur groß genug ist)



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

die Normalverteilung.

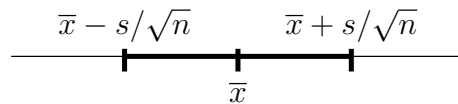




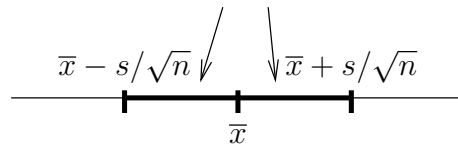
Die Normalverteilungsdichte heisst auch *Gauß'sche Glockenkurve* (nach Carl Friedrich Gauß, 1777-1855)

Wichtige Folgerung

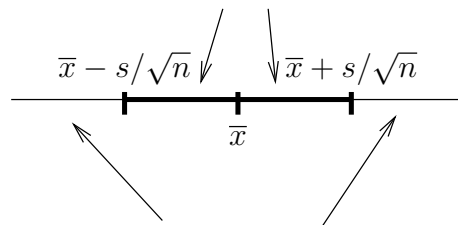
Wir betrachten das Intervall



Mit Wahrscheinlichkeit ca. $2/3$
liegt μ innerhalb dieses Intervalls



Mit Wahrscheinlichkeit ca. $2/3$
liegt μ innerhalb dieses Intervalls



Mit Wahrscheinlichkeit ca. $1/3$
liegt μ *ausserhalb* des Intervalls

Demnach:

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

2.3 Anwendungen

ANWENDUNG 1: Welche Werte von μ sind plausibel?

$$\begin{aligned}\bar{x} &= 0,12 \\ s/\sqrt{n} &= 0,007\end{aligned}$$

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

Antwort: Es ist gut möglich.

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

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

Abweichungen dieser Größe kommen häufig vor.

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

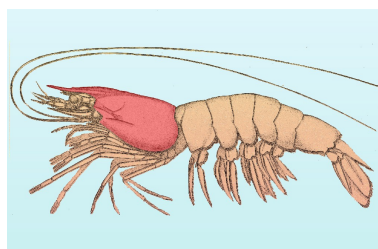
ANWENDUNG 2: Vergleich von Mittelwerten

Beispiel: Springkrebs



Galathea squamifera
image (c) by Matthias Buschmann

Vergleich der Carapaxlänge:



(c): public domain

Galathea: Carapaxlänge in einer Stichprobe

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

Weibchen: $\bar{x}_2 = 3,23$ 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: $\bar{x}_1 = 3,04$ 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 \bar{x}_1 müssen wir rechnen.

Wie genau sind die beiden Mittelwerte?

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

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

Es ist nicht unwahrscheinlich, dass \bar{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 $\bar{x}_2 > \bar{x}_1$

GENAUER FORMULIERT:

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

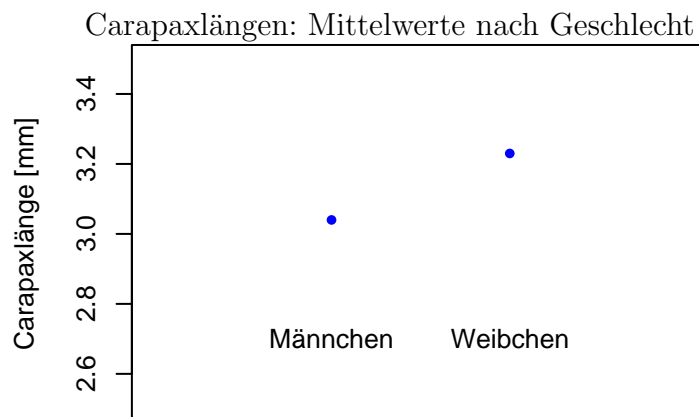
Der Statistiker sagt:
Die Differenz der Mittelwerte ist (statistisch) *nicht signifikant*.

nicht signifikant = könnte Zufall sein

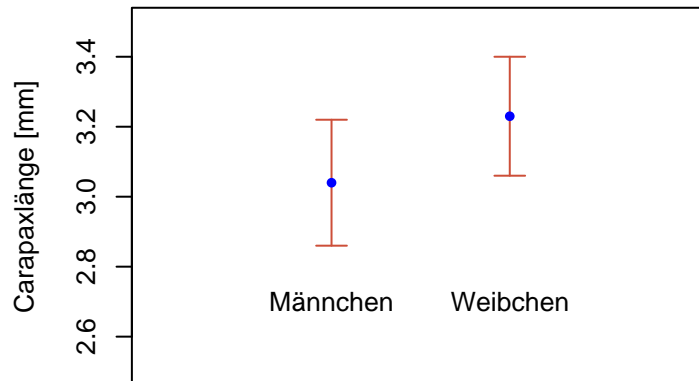
ANWENDUNG 3:

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

Also nicht so: sondern so:



Carapaxlängen: Mittelwerte \pm 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 \bar{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: $\bar{x} = 0,18$ $s = 0,06$
 $n = 13$

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: $s \approx 0,06$, $\sqrt{n} \approx \frac{0,06}{0,01} = 6$ $n \approx 36$

2.4 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 \bar{x} .
- \bar{x} ist eine Zufallsgröße mit Mittelwert μ und Standardabweichung σ/\sqrt{n} .
- Man schätzt die Standardabweichung von \bar{x} mit s/\sqrt{n} .
- s/\sqrt{n} nennt man den *Standardfehler*.
- Schwankungen in \bar{x} von der Größe s/\sqrt{n} kommen häufig vor. Solche Schwankungen sind „*nicht signifikant*“: sie könnten Zufall sein.