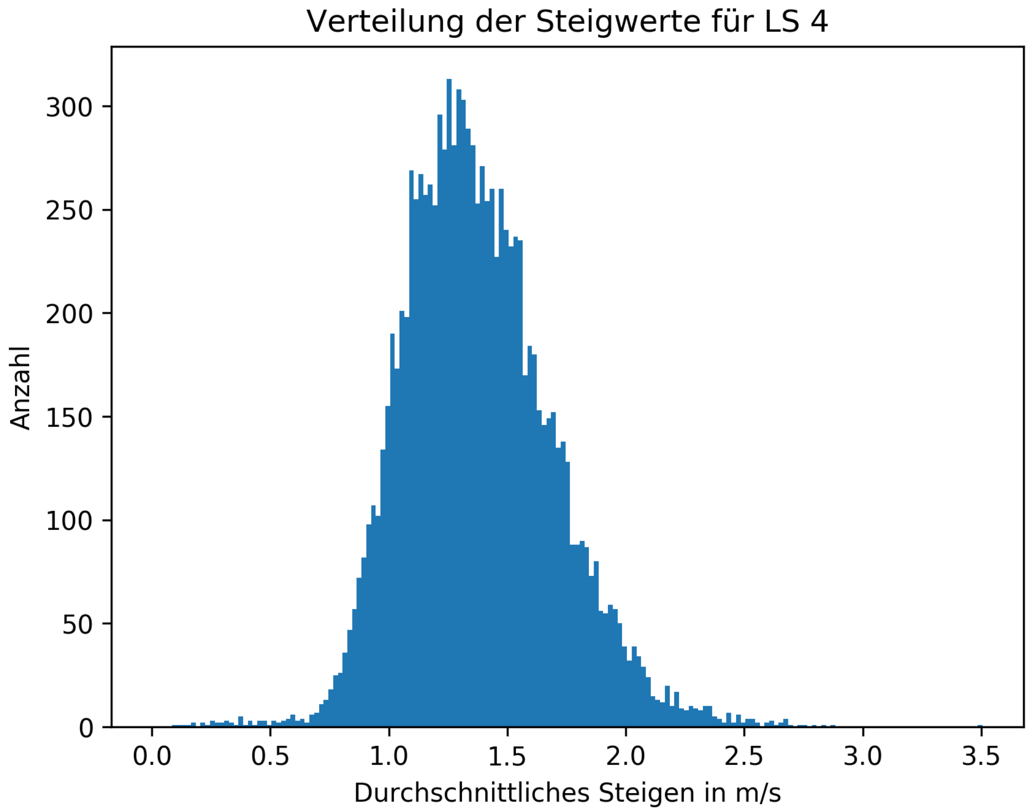
Verteilung des durchschnittlichen Steigens



Wie hoch statistische Fehler, die auf Grund der großen Schwankungen bei den Steigwerten des gleichen Flugzeugtyps entstehen, sind, lässt sich relativ genau beziffern.

In obigem Histogramm ist die Anzahl der Flüge mit Typ LS4, deren durchschnittliches Steigen in einem bestimmten Bereich liegt zu sehen. In der glockenähnlichen Form treten sehr hohe oder sehr niedrige durchschnittliche Steigwerte selten auf, es gibt aber viele Flüge mit durchschnittlichen Steigerten zwischen ein und zwei Metern pro Sekunde. Aus den durchschnittlichen Steigwerten der einzelnen Flüge eines Segelflugzeugs lassen sich für jeden Flugzeugtyp Mittelwert

$$\mu = \frac{1}{n}\sum_{i}^{n} x_i \quad \text{und Varianz} \quad \sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \mu\right)^2 \text{ berechnen,}$$

wobei x_i jeweils das durchschnittliche Steigen eines Fluges ist und n die Anzahl der Flüge. Die empirische Varianz ist also die durchschnittliche quadratische Abweichung vom Mittelwert. Die Standardabweichung σ ist die Wurzel aus der Varianz.

Eine hohe Varianz bei manchen Flugzeugtypen kann entweder auf eine große Heterogenität beim Können der Piloten oder in der Nutzungsart zurückzuführen sein.

Um den mittleren Steigwert möglichst exakt zu schätzen wurde der Mittelwert zusätzlich mit der Dauer des Fluges und dem Kurbelanteil gewichtet.

Ein Flug mit geringem Kurbelanteil und kurzer Dauer geht also weniger stark ein, als ein Flug mit langer Dauer und hohem Kurbelanteil. Dies verbessert die Aussagekraft der Steigwerte, da der Mehrwert an Information eines längeren Fluges auch mehr zählt.

Normalverteilung der Steigwerte

Einen zufällig aus den Daten gezogen durchschnittlichen Steigwert eines Fluges X nennen wir nun Zufallsvariable. In guter Näherung kann man davon ausgehen, dass das durchschnittliche Steigen normalverteilt ist. Wir schreiben X folgt der Normalverteilung

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

Würden wir den Wert der Zufallsvariable X beliebig oft ziehen, würde sich also in unsere Näherung eine Normalverteilung mit Mittelwert μ und Varianz σ^2 ergeben.

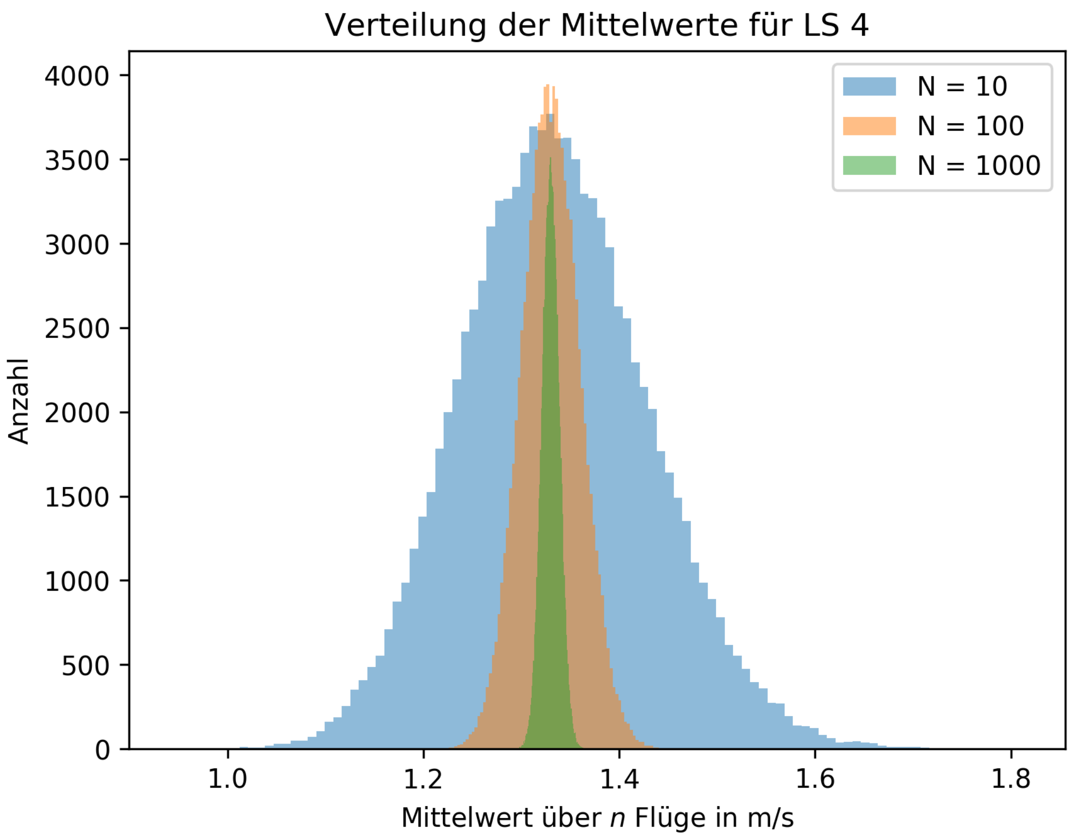
Normalverteilung des Mittelwerte

Die Grafik ‚Verteilung der Mittelwerte' zeigt die Verteilung des mittleren Steigens, wenn hunderttausend mal zufällig n Flüge aus allen Flügen der LS 4 ausgewählt werden und jeweils das mittlere Steigen $\overline{X}$ errechnet wird. Gut zu erkennen ist, dass der Mittelwert einer Normalverteilung folgt, deren Varianz mit zunehmender Stichproben-Größe immer geringer wird. Mit zehn Flüge ist unsere Schätzung sehr ungenau, mit tausend können wird den Bereich schon besser eingrenzen. Der geschätzte Mittelwert ist also nicht immer korrekt, aber je mehr Flüge vorhanden sind, desto sichererer können wir sein und desto weniger zählen individuelle Faktoren.

Die Standardabweichung dieser Verteilung und damit der Fehler unserer Schätzung für den Mittelwert ergibt sich dann als Quotient aus Standardabweichung der Steigwerte und der Wurzel der Anzahl an Flügen:

$$\sigma(\overline{X}) = \frac{\sigma}{\sqrt{n}} \text{ , es gilt also für den Mittelwert } \quad \overline{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$$

So lässt sich für den berechneten Mittelwert auch ein statistischer Fehler angeben. Um diesen möglichst gerung zu halten, wurden die verschiedenen Varianten der gleichen Spannweite eines Flugzeuges (beispielsweise JS 1 B, JS 1 C, JS 1 C TJ) immer zusammengelegt, sodass die Anzahl der Flüge höher war.



Wie hoch ist die Wahrscheinlichkeit, dass ein Flugzeug besser steigt als ein anderes?

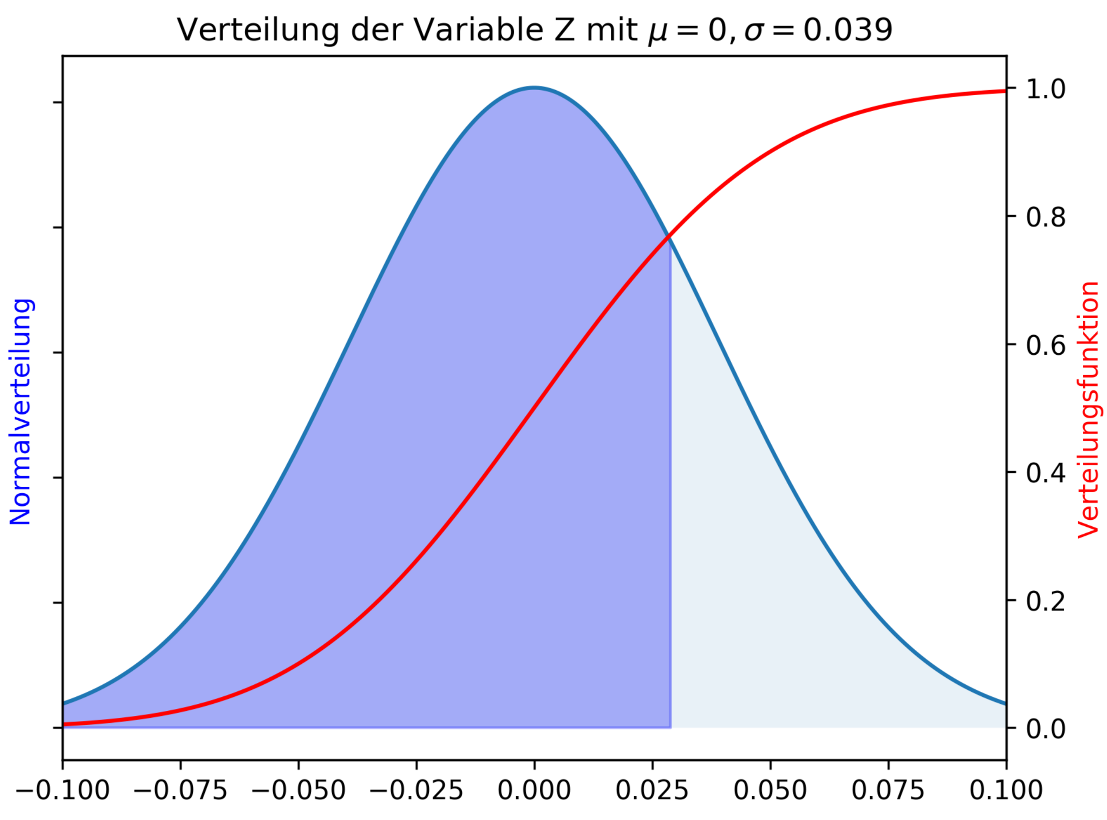
Aus statistischer Sicht geht es um einen Zweistichproben-Test zum Vergleich zweier arithmetischer Mittel aus zwei Grundgesamtheiten mit unterschiedlicher Varianz.
Seien $\overline{X}_1, \overline{X}_2$ unsere Schätzungen für die Mittelwerte der beiden Flugzeuge und $\sigma(\overline{X}_1)$, $\sigma(\overline{X}_2)$ deren Standardabweichungen. Dann ist auch die Differenz Z aus $\overline{X}_1$ und $\overline{X}_2$ normalverteilt mit

$$\mu_Z = \mu_1 - \mu_2 \quad \text{und Standardabweichung} \quad \sigma_Z = \sqrt{\sigma^2(\overline{X}_1) + \sigma^2(\overline{X}_2)}$$

Die Hypothese $\mu_1 > \mu_2$, also die Annahme dass ein Flugzeug besser steigt als das andere kann man prozentual beziffern, indem man betrachtet wie wahrscheinlich es ist, dass Z den Wert $\overline{X}_1 - \overline{X}_2$ annimmt.

Beispiel-Rechnung

Der neue Ventus mit 18 Metern hat ein mittleres Steigen von 1.667 mit einer Standardabweichung von 0.026, die JS 1 mit 18 Metern ein mittleres Steigen von 1.638 und eine Standardabweichung von 0.029. Nun wollen wir die Behauptung überprüfen, dass die JS 1 besser, beziehungsweise mindestens genauso gut steigt wie der Ventus. Dass heißt also, wir nehmen an, dass die Differenz der Mittelwerte μ_Z null beträgt, und nur der Zufall bei den eingereichten Flügen dazu geführt hat, dass der errechnete Mittelwert für den Ventus höher ist. Als erstes wird die Standardabweichung σ_Z ausgerechnet, ihr Wert beträgt 0.039. Wie wahrscheinlich ist es, dass die Zufallsvariable Z den Wert $\overline{X}_1 - \overline{X}_2 = 0.029$ oder einen größeren Wert annimmt?



In der obigen Grafik ist die Fläche unter der Normalverteilung mit Werten kleiner 0.029 farbig markiert. Die (rote) Verteilungsfunktion gibt für jeden Wert an, wieviel Prozent der gezogenen

Zufallszahlen kleiner als dieser Wert sind, sie setzt also das nicht markierte Gebiet in Relation zu dem markierten Gebiet.

Für 0.029 hat sie ungefähr den Wert 0.77. Das heißt, die Differenz aus dem Steigen des Ventus und dem Steigen der JS 1 erreicht in 77 Prozent der Fälle einen Wert kleiner als 0.029, falls diese tatsächlich genauso gut steigen. Nur in 23 Prozent der Fälle taucht höhere Wert auf.

In 77 Prozent der Fälle ist also unsere Annahme, dass die JS 1 im Steigen nicht schlechter als der Ventus ist, falsch. Oder anders ausgedrückt, die Wahrscheinlichkeit dass der Ventus besser steigt als die JS 1, beträgt 77 Prozent.

Für jedes Paar von Flugzeugen ist es möglich, mit den angegebenen Fehlern und Mittelwerten diese Wahrscheinlichkeit selber auszurechnen.

Ungenauigkeiten

Die Gewichtung mit Dauer des Fluges und Kurbelanteil muss auch in die Berechnung der Varianz und Standardabweichung eingehen. Dies wurde für diesen Bericht nicht getan, Varianz und Standardabweichung des Mittelwerts wurden nur mit den durchschnittlichen Steigwerten berechnet. Da die Standardabweichung aber bei kürzeren Flügen höher ist, würde eine solche Gewichtung den Fehlerbereich nochmals verkleinern, sodass unser Fehlbereich eine obere Abschätzung ist.