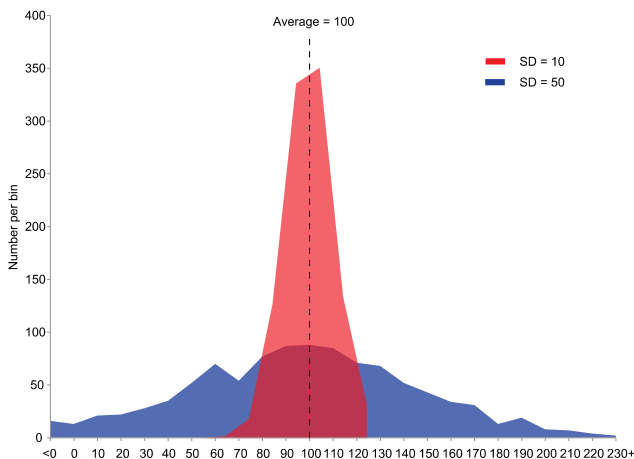


Der **t-Test** bezeichnet eine Gruppe von statistischen Test. Zur Durchführbarkeit müssen folgende Voraussetzungen erfüllt sein:

- Sofern mehrere Stichproben untersucht werden: Die Varianzen (s^2 , eine Kennzahl für die Streuung!), der beiden Stichproben unterscheiden sich nicht *signifikant*. *Gegenbeispiel*: Beispiel für zwei Stichprobenstreuungen, die sich in der Varianz *signifikant* („wesentlich“) unterscheiden.



Ob sich die beiden Stichprobenvarianzen signifikant unterscheiden, lässt sich mit dem **F-Test** prüfen. Eine Abwandlung des t-Tests, der weniger bekannte **WELCH-Test**, setzt keine gleiche Varianz voraus.

- Die genutzten Stichprobenumfänge sind genügend groß, um Grundgesamtheiten zu repräsentieren.
- Die Stichprobenwerte entstammen allesamt aus normalverteilten Grundgesamtheiten.

Es werden folgende t-Tests unterschieden:

- Der **Einstichproben-t-Test (= Einfacher t-Test)** prüft ob sich der Mittelwert einer Stichprobe von einem Sollwert unterscheidet. *Beispiel einer Fragestellung*: Die Packungen eines Paracetamol-Präparates besitzen laut Aufdruck 400 mg Paracetamol pro Tablette. Die stichprobenartige chemische Analyse ergab im Mittel jedoch 395 mg pro Tablette. Ist das nur Zufall und auf eine ungenügend hoher Zahl an Stichproben zurückzuführen oder ist tatsächlich weniger Paracetamol enthalten, als angegeben?
- Der **Zweistichproben-t-Test (auch Doppelter t-Test)** prüft anhand der Mittelwerte zweier Stichproben, wie sich die Mittelwerte zweier Grundgesamtheiten zueinander verhalten. *Beispiel einer Fragestellung*: 2

Packungen Paracetamol werden jeweils für sich stichprobenartig auf den Paracetamolgehalt der Tabletten geprüft und die Mittelwerte verglichen:

Packung A: 395 mg Packung B: 399 mg

Ist das nur Zufall und auf eine ungenügend hoher Zahl an Stichproben zurückzuführen, oder unterscheiden sich tatsächlich die wahren Paracetamolgehalte?

- Der **Abhängige t-Test (auch Paardifferenzentest)** prüft für zwei verbundene (abhängige) Stichproben, ob sich die mittlere Differenz der Messwerte sich von Wert $\Delta = 0$ unterscheidet. Dabei wird vorausgesetzt, dass die Differenzen normalverteilt sind. *Beispiel einer Fragestellung*: 10 Personen mit Fieber wurde ein Paracetamolpräparat gegeben. Dabei sank das Fieber im Mittel um 1,3 °C. Ist das nur Zufall und auf eine ungenügend hoher Zahl an Stichproben zurückzuführen, oder besitzt das Paracetamol fiebersenkende Wirkung?

Die t-Tests laufen darauf hinaus, dass zuerst eine Behauptung (Hypothese) formuliert wird. Dies ist in der Regel die **Nullhypothese (H_0)**: „Wir geht davon aus, die sich die beiden Vergleichsgrößen nur zufällig unterscheiden und auf die Stichprobenbildung zurückzuführen sind“.

Anschließend wird dann mit einem Rechenverfahren eine **Prüfgröße (= Prüfstatistik = Teststatistik)** berechnet und mit einem tabellierten Wert verglichen. Überschreitet die berechnete *Teststatistik (Prüfgröße)* den tabellierten Wert, wird die Nullhypothese verworfen, d.h. man geht von einem tatsächlichen Unterschied aus. Diese Verwerfung unterliegt aber einer gewissen Irrtumswahrscheinlichkeit (α -Fehler). Endgültige Gewissheit, ob sich die tatsächlichen Werte unterscheiden, bekommt man also auch mit dem t-Test nicht. Je nachdem welche Irrtumswahrscheinlichkeit (α) man sich einräumt, muss man die *Teststatistik* mit einem anderen Tabellenwert vergleichen.

Räumt man sich eine Irrtumswahrscheinlichkeit von unter $\alpha = 5\%$ bei der Ablehnung der Nullhypothese, so sagt man auch: „Die Werte unterscheiden sich statistisch *signifikant*“. Ist man mit sich strenger und räumt sich nur eine Irrtumswahrscheinlichkeit von unter $\alpha = 1\%$ ein, so sagt man, die Werte unterscheiden sich „statistisch *hoch signifikant*“

Tabellenkalkulationsprogramme können mit der Funktion **TTEST** (bzw. **TTEST**) direkt die Irrtumswahrscheinlichkeit α berechnen.