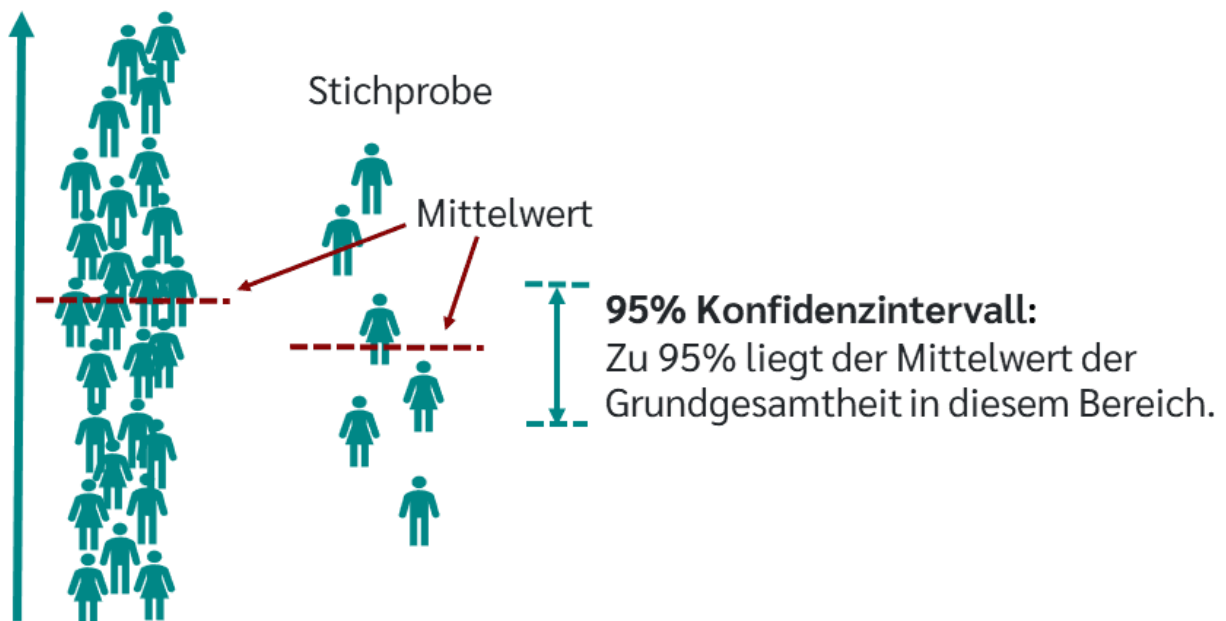


Was ist das Konfidenzintervall?

Das Konfidenzintervall KI ist der Bereich, in dem ein Parameter (z. B. der Mittelwert) mit einer bestimmten Wahrscheinlichkeit liegt.

Wenn mehrere Stichproben aus einer Grundgesamtheit gezogen werden, ist es sehr wahrscheinlich, dass jede Stichprobe einen anderen Mittelwert aufweist. Man möchte aber den Mittelwert der Grundgesamtheit kennen und nicht den der Stichprobe. Das Konfidenzintervall gibt nun den Bereich an, in dem der wahre Mittelwert der Grundgesamtheit mit einer bestimmten Wahrscheinlichkeit liegt.

Grundgesamtheit



Achtung! Die obige Erklärung ist weit verbreitet, weil sie einfach zu verstehen ist, aber nicht alle Experten halten sie für richtig. Richtig, aber komplizierter ist die folgende Definition:

Das 95 %-Konfidenzintervall (KI) ist ein Intervall, das anhand von Stichprobendaten aus einer unendlichen Reihe berechnet wird, von denen 95 % den Populationsparameter enthalten. Langfristig enthalten 95 % dieser Intervalle den wahren Mittelwert.

Wozu brauchen wir das Konfidenzintervall?

In der Statistik werden häufig Parameter der Grundgesamtheit anhand einer Stichprobe geschätzt, z. B. der [Mittelwert](#) oder die [Varianz](#). Dies sind jedoch nur Schätzungen und der wahre Wert in der Grundgesamtheit wird irgendwo um diese Schätzungen herum liegen. Es ist sehr nützlich, einen Bereich oder ein Intervall zu definieren, in dem der wahre Wert mit hoher Wahrscheinlichkeit liegt.

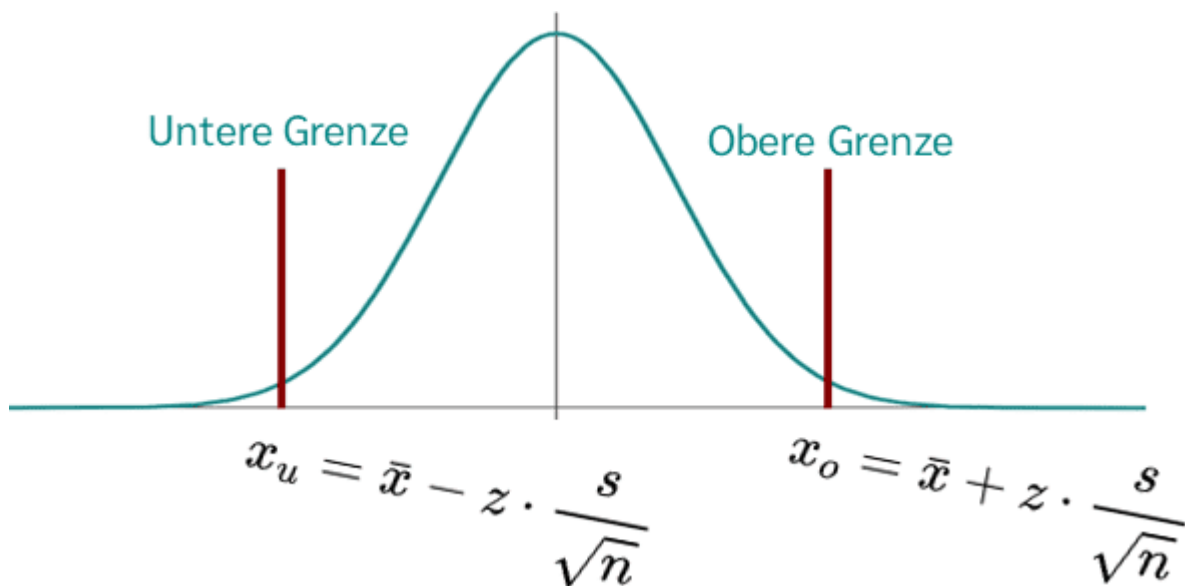
Konfidenzintervall berechnen

Zur Berechnung des Konfidenzintervalls wird die Verteilungsfunktion des betreffenden Parameters (z. B. des Mittelwerts) in der Grundgesamtheit benötigt. Nimmt man an, dass diese Verteilung normalverteilt ist, so ergibt sich das Konfidenzintervall für den Mittelwert mit:

$$KI = \bar{x} \pm z \cdot \frac{s}{\sqrt{n}}$$

Standardabweichung s
 Stichprobengröße n
 Mittelwert \bar{x}
 z-Wert für das Konfidenzniveau z
 Untere/obere Grenze

Dabei ist \bar{x} der Mittelwert der Stichprobe, n der Stichprobenumfang und s die Standardabweichung der Stichprobe. Plus und Minus geben die untere bzw. obere Grenze des Konfidenzintervalls an.



Wenn die Stichprobe klein ist, wird anstelle der Normalverteilung die t-Verteilung verwendet. Dann wird der z-Wert durch t ersetzt und die Formel lautet

$$KI = \bar{x} \pm t \cdot \frac{s}{\sqrt{n}}$$

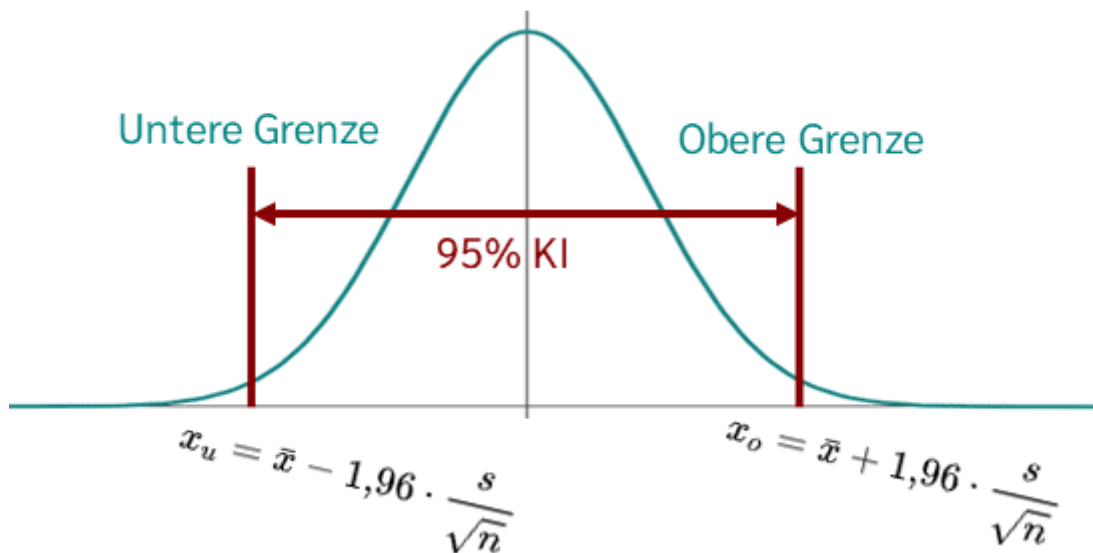
Konfidenzintervall 95%

Für die Berechnung des Konfidenzintervalls muss die Wahrscheinlichkeit festgelegt werden, mit der der Mittelwert der Grundgesamtheit innerhalb des Intervalls liegt. Als Wahrscheinlichkeit wird sehr häufig das Konfidenzniveau von 95% oder 99% verwendet. Diese Wahrscheinlichkeit wird auch als Konfidenzoeffizient bezeichnet.

Für das **95% Konfidenzintervall** und das **99% Konfidenzintervall** ergeben sich die folgende z-Werte:

Konfidenzniveau	95%	99%
z-Wert	1,96	2,58

Wenn ein Konfidenzintervall von 95 % angegeben wird, kann man zu 95 % sicher sein, dass der wahre Wert des Parameters innerhalb dieses Intervalls liegt.



p-Wert

Ein P-Wert ist ein statistisches Maß dafür, wie wahrscheinlich es ist, dass ein beobachteter Unterschied zwischen Gruppen zufällig auftritt. Es gibt an, wie stark die Evidenz gegen die Nullhypothese ist, dass es keinen Unterschied zwischen den Gruppen gibt. Ein kleiner P-Wert (z. B. $<0,05$) deutet darauf hin, dass die Wahrscheinlichkeit, dass der Unterschied zufällig auftritt, sehr gering ist.

Der P-Wert wird **berechnet**, indem die Wahrscheinlichkeit bestimmt wird, den beobachteten oder einen noch stärkeren Unterschied zwischen den Gruppen zu erhalten, wenn die Nullhypothese wahr ist. Dies wird durch die Verwendung statistischer Tests wie des t-Tests oder der ANOVA erreicht, die den P-Wert berechnen, basierend auf der Verteilung der Stichproben und der Stärke des Unterschieds.

p-hacking

p-Hacking, auch als specification searching bekannt, bezeichnet die Verzerrung und Manipulation von Forschungsergebnissen, deren Ziel es ist, den p-Wert (d. h. die statistische Signifikanz) durch nachträgliche Anpassung der Testparameter unter eine bestimmte Grenze zu reduzieren.

Der p-Wert wird „gehackt“, also künstlich unter die 5%-Grenze gedrückt. Auf diese Weise wird eine vermeintliche statistische Signifikanz der Ergebnisse erzeugt. Die häufige Fehlinterpretation von p-Werten sowie die Anwendung von p-Hacking haben zu unzähligen falschen Forschungsergebnissen geführt, die der Wissenschaft geschadet haben.

Hacking kann als Reaktion wissenschaftlicher Autoren auf die Tatsache angesehen werden, dass Studien mit signifikanten Ergebnissen für die Publikation präferiert werden und Analysen, die nicht signifikante Resultate aufweisen, unveröffentlicht bleiben (File-Drawer-Problem). Mithilfe von Metaanalysen ist es möglich, p-Hacking aufzudecken.

Quelle:

<https://de.wikipedia.org/wiki/P-Hacking>