

α - und β - Fehler bzw. Fehler 1. und 2. Art

Zur Verdeutlichung der möglichen Fehler bei statistischen Entscheidungen beschränke ich mich im folgenden auf *einseitige* Fragestellungen. Es wird also von der Frage ausgegangen, ob sich durch die Einnahme eines Medikaments das Reaktionsvermögen verlangsamt, ob $\mu > \mu_0$ ist.

Gehen wir wiederum von einem Mittelwert in unserer Stichprobe von 105 ms aus. Die Wahrscheinlichkeit dafür, eine derartige Abweichung oder eine noch größere unter Geltung der Nullhypothese zu bekommen, wird als Überschreitungswahrscheinlichkeit bezeichnet. Wir erhalten sie, indem wir den Mittelwert von 105 z-transformieren.

In unserem konkreten Fall ergibt das:

$$z = (105 - 100) / 2,5 = 2$$

Die Fläche von -unendlich bis 2 ergibt die Wahrscheinlichkeit, bei einem Mittelwert der Stichprobenkennwerteverteilung von 100 und einem Standardfehler des Mittelwertes von 2,5 einen Mittelwert kleiner als 105 zu bekommen. Diese Wahrscheinlichkeit beträgt 0,9772

Die Überschreitungswahrscheinlichkeit ist nun nichts anderes als $1 - 0,9772 = 0,0228$

Wie wir bereits wissen, ist ein derartiges Ergebnis unter Voraussetzung der Nullhypothese nicht unmöglich. Die Wahrscheinlichkeit für ein derartiges Ergebnis - 105 ms oder mehr - beträgt 0,0228

Aufgrund eines Signifikanzniveaus von $p \leq 0,05$ verwerfen wir die Nullhypothese und entscheiden uns für die Alternativhypothese. Eine Überschreitungswahrscheinlichkeit von 0,0228 ist zwar unwahrscheinlich, sie ist aber nicht unmöglich. Verwerfen wir die Nullhypothese aufgrund dieser Überschreitungswahrscheinlichkeit, so besteht daher immer noch ein Restrisiko, dass die Nullhypothese dennoch richtig gewesen wäre. Die Wahrscheinlichkeit für dieses Restrisiko entspricht exakt der Überschreitungswahrscheinlichkeit. Diese ist nämlich die Wahrscheinlichkeit *unter Geltung der Nullhypothese* eine Abweichung von 105 oder mehr ms zu bekommen. Sie wird daher auch als *Irrtumswahrscheinlichkeit* oder als Wahrscheinlichkeit für einen α -Fehler (Fehler 1.Art) bezeichnet. Es ist dies die Wahrscheinlichkeit fälschlicherweise die Nullhypothese zu verwerfen.

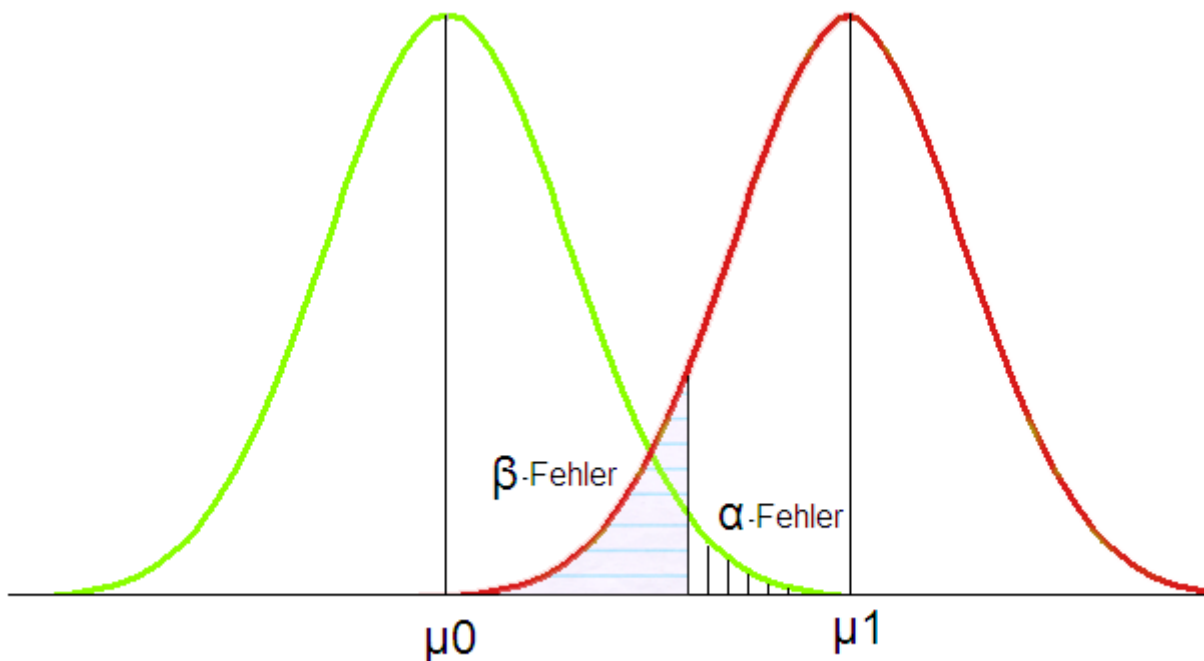
Nun stehen wir bei der Frage, ob wir die Nullhypothese verwerfen oder akzeptieren sollen, vor *zwei* Entscheidungen: Entscheidung zugunsten der Nullhypothese oder Entscheidung gegen die Nullhypothese. Beide Entscheidungen können falsch sein. Den Fehler, sich fälschlicherweise für die Nullhypothese zu entscheiden, obwohl die Alternativhypothese richtig wäre, wird als β -Fehler bzw. Fehler 2.Art bezeichnet.

Um sich diesen zweiten Fehler besser veranschaulichen zu können, gehen wir zunächst davon aus, der Mittelwert μ_1 unter Geltung der Alternativhypothese sei uns bekannt. Nehmen wir also probenhalber an, nach Einnahme des Medikaments würde eine durchschnittliche Reaktionszeit von 110 ms auftreten. μ_1 sei also 110. Auch für diese Population können wir die Stichprobenverteilung der Mittelwerte berechnen. Der Standardfehler dieser Verteilung

sei ebenfalls 2,5. Beide Verteilungen - die Kennwertverteilung der Mittelwerte unter Geltung der H_0 und die Kennwertverteilung unter Geltung der H_1 - überschneiden sich in einem Bereich.

Die Fläche *links* vom Mittelwert von 105 stellt die Wahrscheinlichkeit dar, irrtümlicherweise sich *gegen* die H_1 zu entscheiden. Dies ist die Wahrscheinlichkeit des β -Fehlers.

Stellen wir dies graphisch dar:



Die verschiedenen Entscheidungsalternativen und die damit verbundenen Wahrscheinlichkeiten lassen sich schematisch wie folgt zusammenfassen:

Entscheidung zugunsten von:	Tatsächlich richtig ist	
	H_0	H_1
H_0	$1 - \alpha$	β
H_1	α	$1 - \beta$

Die richtige Entscheidung *zugunsten* der H_1 wird auch als *Teststärke* bezeichnet. Es ist dies die Wahrscheinlichkeit, uns richtigerweise für die Alternativhypothese zu entscheiden.

Da aber μ_1 in den Sozialwissenschaften in der Regel nicht bekannt ist - wir vermuten nur, daß das Medikament die Reaktionszeit verlangsamt, wissen aber nicht, um wie viel - kann der β -Fehler nicht genau angegeben werden. Wir können daher nicht die Wahrscheinlichkeit dafür angeben, uns fälschlicherweise für die H_0 zu entscheiden.

Aus diesem Grunde sagt man bei einer Entscheidung zugunsten der H_0 , *die Nullhypothese konnte aufgrund der empirischen Stichprobe nicht verworfen werden*, nicht aber: sie wurde angenommen.

Bei der Hypothesenbildung wird der α -Fehler deshalb klein gehalten, um die mit der Alternativhypothese verbundenen *Kosten* zu minimieren. Eine Minimierung des α -Fehlers bedeutet ein *konservatives* Testen. Man stelle sich als Beispiel einen Produzenten von Reifen vor, der überlegt, ein neues und kostspieliges Verfahren einzuführen. Ein statistischer Test soll zeigen, inwiefern die neue Produktion tatsächlich zu einer besseren Reifenqualität führt. Wegen der hohen Umstellungskosten wird sich der Produzent nur im Falle einer hohen Signifikanz - also bei einem möglichst kleinen α -Fehler für die neue Produktion entscheiden.