Achtung, Mathe!

Die Regel der Drei

Björn und Sören Christensen

Stellen Sie sich folgendes vor: Sie werden bei der Arbeit beauftragt, die eingehenden Bestellungen auf Schäden zu überprüfen. Sie schauen sich, sagen wir, dreißig Teile an. Wenn davon nun kein einziges Teil zu beanstanden wäre, würden Sie vermutlich als Ergebnis nicht zu der Einschätzung kommen, dass in der Gesamtlieferung gar keine defekten Teile sind. Es dürfte stattdessen sinnvoller sein, eine obere Schranke anzugeben: „Ich vermute, dass der Anteil der defekten Teile in der Gesamtlieferung nicht mehr als x beträgt.“

Aber was ist x? – In der Praxis nutzt man dafür oft eine einfach zu merkende Daumenregel, die „Regel der Drei“: Wenn Sie N Teile untersucht haben, die alle ok waren, dann können Sie ziemlich sicher davon ausgehen, dass in der Gesamtlieferung der Anteil der defekten Teile nicht höher als 3/N ist. In unserem Beispiel ist N=30. Sie können nach Ihrer Stichprobe also statistisch abgesichert sagen, dass in der Gesamtlieferung nicht mehr als 3/30 = 10 % der Teile defekt sein werden.

Wenn Sie eine größere Stichprobe von, sagen wir, N=100 Teilen angesehen haben und dabei alle ok waren, dann sagt die Regel, dass der Gesamtanteil nicht mehr als 3/100 = 3 % betragen wird. Die Regel ist natürlich nicht auf defekte Teile beschränkt. Wenn Sie zum Beispiel die ersten N Wörter im Diktat Ihres Kindes kontrolliert haben und alle fehlerfrei waren, dann können Sie – sofern die Konzentration Ihres Kindes nicht nachgelassen hat – davon ausgehen, dass nicht mehr als ein Anteil von 3/N der Wörter falsch sind.

Regel hilft dabei, den Fehleranteil abzuschätzen

Wenn Sie jetzt keine Lust auf mathematische Herleitungen haben, können Sie das Lesen an dieser Stelle einfach beenden. Für alle anderen folgt hier die Erklärung der „Regel der Drei“: Wenn wir die unbekannte wirkliche Fehlerwahrscheinlichkeit p – in unserem Fall die Wahrscheinlichkeit für ein defektes Teil – nennen, dann ist die Wahrscheinlichkeit dafür, dass man bei N Beobachtungen keine Fehler findet $(1\text{-}p)^N$. Nun suchen wir ein 95 %-Konfidenzintervall, das heißt, wir müssen das größte p so finden, dass die Wahrscheinlichkeit, dass keine defekten Teile beobachtet werden, 5 % beträgt.

Das heißt, man muss die Gleichung $(1\text{-}p)^N = 5\ \%$ nach p auflösen. Durch Logarithmieren und Umstellen erhält man $\ln(1\text{-}p)=\ln(0{,}05)/N$. Da nun $\ln(0{,}05)= -2{,}995\ldots$ und $\ln(1\text{-}p)$ für kleine p ungefähr -p ist, ergibt sich ungefähr p=3/N. Egal, ob mit Herleitung oder ohne, die Regel der Drei hilft im Alltag sehr einfach, um grob abzuschätzen, wie groß ein Fehleranteil in einem Bereich maximal sein dürfte, wenn in einer Stichprobe keine Fehler aufgetreten sind.

Björn Christensen ist Professor für Statistik und Mathematik an der FH Kiel. **Sören Christensen** ist Professor für Stochastik an der Christian-Albrechts-Universität Kiel.

SH Journal, Wochenendbeilage im shz, 22.04.2023