

Geschäfte machen mit Statistik.

Teil 1: Die magischen fünf Prozent

Der Laden

Zu Beginn meiner Karriere als Statistiker wurde ich (Abb. 1) während einer Dienstreise in den USA von meinem Gastgeber als Mitglied des „Ladens“ („shop“) von Prof. Schumacher in Freiburg vorgestellt. Reagierte ich zunächst einigermaßen verwundert – wir betreiben doch keinen Kiosk –, so finde ich heute die Vorstellung eines „Ladens“ für Teile unserer statistischen Arbeit recht zutreffend. Das Angebot in dem Laden umfasst gewisse Entscheidungen, z. B., ob ein Ergebnis von statistischer Signifikanz (übersetzt: Bedeutung) ist. Die Kundschaft besteht in meinem Fall zumeist aus Doktoranden der Zahnmedizin und deren Betreuern. Der Leser sollte zudem noch wissen, dass das Material, also Daten und Hintergrundinformationen, von den Fragestellern für gewöhnlich selbstverantwortlich eingebracht wird.



Dr. rer. nat. Thomas Gerds

Die Geschäftsphilosophie

Einer der ersten „Kunden“, die mein damals noch neues „Geschäft“ betreten, brachte Bruchfestigkeitswerte von verschraubten Implantat-Abutment-Systemen. Es waren für verschiedene Materialien jeweils acht Werte erhoben worden (Tab. 1). Eine Frage bestand darin, ob sich zwei der Materialien hinsichtlich der Bruchfestigkeit unterscheiden. Mein erster Gedanke war: „Wie bitte? Mit nur acht Werten soll ich eine allgemein für diese Materialien geltende Aussage machen? Also mit sehr sparsamen Informationen eine eventuell weitreichende Entscheidung treffen? Unmöglich! Allein die möglichen Konsequenzen: Ich sage, das eine Material sei besser, das Ergebnis wird veröffentlicht, die Industrie stellt Ihre Produktion um, die Patienten werden nur noch mit diesem Material versorgt, und so weiter!“ Ein solches Projekt erschien mir höchst riskant und wenig vertrauenswürdig.“

**„Die Statistik ist wie eine Laterne im Hafen.
Sie dient dem betrunkenen Seemann
mehr zum Halt als zur Erleuchtung.“**

Hermann Joseph Abs, deutscher Bankier (1901-1994)

Schließlich warf ich einen genaueren Blick auf die 16 Daten. Und plötzlich dämmerte es mir. Ich wusste doch Bescheid über die Ware, die ich hier zu verkaufen hatte (zur Erinnerung: es handelt sich nur um die Entscheidung über die sta-

tistische Bedeutung des Unterschieds zwischen den beiden Materialien). Entscheidungstheorie ist nämlich ein fester Bestandteil der mathematischen Statistik, und die habe ich studiert. So wusste ich, dass die *Neymann-Pearson*-Theorie [für einen Überblick siehe etwa 2] eine Strategie liefert, die gewissermaßen eine Qualitätskontrolle der Ware ermöglicht: Wenn ich mich streng an diese Strategie halten würde, dann wären in meiner *gesamten zukünftigen Karriere* höchstens 5% der Entscheidungen falsch zugunsten des einen oder des anderen Materials. Die Qualität der Ware wäre zudem noch einigermaßen unabhängig von der Art der Entscheidungen, der Anzahl der Kunden, und auch von der eventuell kleinen Datenmengen. Wer würde solch eine Strategie nicht verfolgen wollen?

Hintergrundinformation

Der kalifornische Statistiker *E. L. Lehmann* [2] zitiert zwei Größen seines Fachs, *J. Neyman* (1894-1981) und *E. S. Pearson* (1895-1980), wie folgt: „Ohne die Hoffnung zu haben, dass man weiß, ob eine bestimmte Hypothese wahr oder falsch ist, können wir jedoch nach Regeln suchen, die unsere Handlungen bestimmen, so dass wir versichern können, auf lange Sicht nicht zu oft falsch zu liegen.“ Einen statistischen Test, das heißt eine Regel, die eine Hypothese annimmt oder verwirft, beschreiben die beiden Autoren als „eine Verhaltensregel“.

Der Preis ist heiß

Was bedeutet diese Strategie nun für den Kunden? Der Laden ist zwar vertrauenswürdig, denn er liefert selten falsche Entscheidungen, dafür bürgen schließlich die mathematischen Beweise der *Neymann-Pearson*-Theorie. Aber was ist der Preis dafür?

Ganz einfach lässt er sich anhand des obigen Beispiels erklären. Bei acht Messwerten pro Gruppe werde ich (natürlich nicht ohne vorher die angemessene statistische Analyse durchgeführt zu haben) einfach keine Entscheidung fällen! Weder das eine noch das andere Material hat signifikant bessere Werte. Die daraus leider viel zu oft gezogene Schlussfolgerung, es gäbe tatsächlich keinen signifikanten Unterschied, ist aber falsch [1].

Wenn der Auftraggeber meines Kunden bereit wäre, mehr zu bezahlen, etwa in Form einer größeren Studie, dann würde er damit die Entscheidungsfindung wahrscheinlicher machen. In der Tat wird eine größere Studie auch mehr Geld kosten. Damit nicht zuviel investiert wird, ist ein Vorgespräch angesagt, worin die Investitionsmenge so abgeschätzt wird, dass ein zufriedenstellendes Ergebnis erwartet werden kann.

Implantat-Abutment Nr. 1	Implantat-Abutment Nr.2
667 ,66	794 ,08
888 ,43	721 ,95
1048 ,39	579 ,43
845 ,68	720 ,92
975 ,93	688 ,78
770 ,19	902 ,31
788 ,07	866 ,38
637 ,61	763 ,24

Tabelle 1 Die acht Bruchfestigkeitswerte (gemessen in Newton) können mit dem bloßen Auge verglichen werden. Es handelt sich um einen Auszug der in [3] veröffentlichten Daten.

4 Kundenzufriedenheit

Warum hat sich die magische 5%-Signifikanzgrenze etabliert? Aufgrund der Vorstellung eines Ladens lässt sich dieses Phänomen durch das Preis-Leistungs-Verhältnis erklären: Warum bevorzugt ein Kunde einen Laden mit 5% falschen Entscheidungen gegenüber einem Laden mit nur 1% oder sogar 20%? Die Antwort ist durch eine Mehrheitsentscheidung zu begründen: Der erste Laden (1%) ist den meisten Kunden zu teuer, der zweite (20%) produziert Billigware wie eine 10-Euro-Stichsäge.

Literatur

1. Gerdts T, Türp JC, Antes G: Von der Trennschärfe einer statistischen Entscheidung und dem Testen auf Gleichheit. Dtsch Zahnärztl Z 60, 547-548 (2005)
2. Lehmann EL: The Fisher, Neyman-Pearson theories of testing hypotheses: One theory or two? J Am Stat Assoc 88, 1242-1249 (1993) [im Weltnetz frei unter < <http://www.jstor.org/view/01621459/di985998/98p0012j/0>>]
3. Strub JR, Gerdts TA: Fracture strength and failure mode of five different single-tooth implant-abutment combinations. Int J Prosthodont 16, 167-171 (2003)

T. Gerdts, Freiburg