FORMEN UND GRÖßEN: VERTEILUNGSTHEORIE IN DER ANWENDUNG

A.F. Bissell

Übersetzt von G. Fillbrunn

In diesem Artikel über eine Anwendung der Wahrscheinlichkeitsrechnung und Statistik in der Wirtschaft befassen wir uns mit einem Problem, das jedem Hersteller von Gütern, deren Größe in einer gewissen Bandbreite variiert, begegnet. Die Frage lautet, wieviel Stücke in den einzelnen Größen hergestellt werden sollen.

In einigen Fällen kann diese Frage nur mit Hilfe einer detaillierten Marktuntersuchung beantwortet werden. Zum Beispiel dann, wenn man etwa wissen möchte, welcher Prozentsatz von Hausfrauen jeweils die verschiedenen vorgesehenen Packungsgrößen eines neuen Reinigungsmittels oder neuer Frühstücksflocken kaufen wird. Entsprechend wird der Hersteller von technischen Bauteilen (Scharnieren, Schrauben, Rohrverbindungen) Daten über den Bedarf seiner Kunden hinsichtlich der von ihm geplanten Größen benötigen.

Es gibt noch andere Fälle, bei denen ein einfaches statistisches Modell zusammen mit einigen Daten, die zur Schätzung der zugehörigen Parameter dienen, gut weiterhelfen kann. Typische Beispiele begegnen einem, wenn man die Anforderungen an verschiedene Kleidungsstücke wie Hüte, Schuhe, Hemden usw. herausfinden möchte. Die Größe dieser Artikel muß offensichtlich mit den Maßen derer, die sie tragen sollen, in Zusammenhang gebracht werden - Hutgrößen mit den Kopfmaßen, Schuhe mit der Fußlänge usw. Wenn statistische Daten über die Verteilung dieser Maße in den anzusprechenden Bevölkerungsteilen verfügbar sind oder ohne weiteres beschafft werden können, besteht die Möglichkeit, die benötigten Mengen dadurch zu schätzen, daß man die Eigenschaften eines geeigneten Verteilungstyps benutzt.

Bei bestimmten Kleidungsstücken können Daten für zwei (oder mehr) Maße erforderlich sein, wie zum Beispiel die Taillenweite und Beinlänge für Hosen oder die Fußlänge und Fußbreite für Schuhe, die sowohl in 'Paßformen' als auch in Größen angeboten werden. Somit werden manchmal bivariable (oder sogar multivariable) Daten benötigt, die mit geeigneten statistischen Verfahren verarbeitet

werden. Dies werden wir hier nicht weiter verfolgen, sondern wir wollen uns ein einfaches Beispiel mit einer Variablen vornehmen. Die vorkommenden Daten sind angenommen, typisch ist dagegen das Problem.

Wir nehmen einmal an, daß ein Hemdenhersteller mit einer jährlichen Stückzahl von ungefähr einer halben Million (sagen wir 10000 Stück pro Woche) seine Erzeugnisse auf das metrische System umstellen und in Größen mit 1cm-Intervallen produzieren möchte. Durch eine geeignete Stichprobe, die aus einer entsprechenden Grundgesamtheit (diese besteht aus den erwachsenen männlichen Briten) stammt, erhält er Realisierungen der Variablen für die Hemdengröße, die von dem einzelnen Merkmalsträger bevorzugt wird. Die Halsweite mag für hochgeschlossene Hemden die geeignetste Variable sein, während für Freizeithemden wohl der Brustumfang zu nehmen ist (hier dürften auch größere Abstände zwischen den einzelnen Größen möglich sein). Wir wollen den ersten Fall betrachten. Der Produzent findet wahrscheinlich heraus, daß die ermittelten Halsweiten zu einer ungefähren Normalverteilung gehören mit dem Erwartungswert von etwa 38 cm und der Standardabweichung von ungefähr 1,5 cm. Wir wollen annehmen, daß er 37,8 cm als Schätzwert für μ und 1,4 cm als solchen für σ erhält.

Mit dieser Verteilung muß er für seinen wöchentlichen Ausstoß (10000 Hemden) die Kontingente der verschiedenen Größen festlegen. Somit ist das Problem darauf zurückgeführt, für die verschiedenen 1cm-Halsweitenklassen die jeweiligen Bevölkerungsanteile zu schätzen (er mag auch zu erwägen haben, welche Kragenweite für jemanden mit einer gegebenen Halsweite benötigt wird, so daß ihm der Kragen weder zu eng noch zu weit ist).

Für jeden beliebigen Halsweitenbereich $(x_1; x_2]$ kann der zugehörige Anteil aus der Tabelle für die Standardnormalverteilung entnommen werden. Ausgehend von der Standardisierung

$$U = \frac{X-\mu}{\sigma}$$

findet man mit $u_1 = \frac{x_1-\mu}{\sigma}$ den Anteil der Werte von U, die höchstens so groß wie u_1 sind. Analog erhält man mit $u_2 = \frac{x_2-\mu}{\sigma}$ den entsprechenden Prozentsatz hinsichtlich der oberen Grenze x_2 des Halsweitenintervalls. Der Anteil der Werte, die in $(x_1; x_2]$ liegen, ist dann

$$P(U \leq u_2) - P(U \leq u_1)$$

Tabelle 1. Anteile der erwachsenen Männer bei verschiedenen Halsweitenintervallen

Halsweite [cm] $(x_1;x_2]$	Entsprechende Standardisierungen u_1	u_2	$P(U \leq u_2)$	Anteil für das Halsweitenintervall
höchstens 33	$(-\infty)$	-3,43	0,0003	0,0003
(33;34]	-3,43	-2,71	0,0034	0,0031
(34;35]	-2,71	-2,00	0,0228	0,0194
(35;36]	-2,00	-1,29	0,0985	0,0757
(36;37]	-1,29	-0,57	0,2843	0,1858
(37;38]	-0,57	0,14	0,5557	0,2714
(38;39]	0,14	0,86	0,8051	0,2494
(39;40]	0,86	1,57	0,9418	0,1367
(40;41]	1,57	2,29	0,9890	0,0472
(41;42]	2,29	3,00	0,9987	0,0093
über 42	3,00	(∞)	(1)	0,0013

(Die Werte für u wurden auf zwei Dezimalstellen gerundet, da die Tabellen der Standardnormalverteilung im allgemeinen nur diesen Genauigkeitsgrad berücksichtigen.)

Tabelle 1 enthält die Zwischen- und Endergebnisse für die Halsweiten im (ungefähren) 3σ-Bereich um μ.

Nun müssen noch aus den Prozentsätzen die wöchentlichen Produktionskontingente gewonnen werden. Durch Multiplikation der Werte der letzten Spalte von Tabelle 1 mit 10000 ergeben sich die 'exakten' Werte der Tabelle 2. Es ist sinnvoll, diese auf Werte zu runden, die für die Praxis günstig sind, z.B. auf Fünfziger, Hunderter oder Gros. Außerdem dürfte es unwirtschaftlich sein, für die sehr kleinen und sehr großen Weiten, von denen geringe Stückzahlen benötigt werden, wöchentliche Kontingente festzulegen. Wenn der Fabrikant Hemden mit diesen Extremmaßen herstellen möchte, wird er diese wohl in einer Stückzahl von einigen Hundert auf einmal und in mehrwöchigen Abständen produzieren.

Tabelle 2. Kontingente bei 10000 Hemden pro Woche

Halsweite [cm] $(x_1;x_2]$	'Exaktes' Kontingent	Geeignet gerundete Kontingente für einen reduzierten Größenbereich
höchstens 33	3	Keine Angabe des Kontingents
(33;34]	31	
(34;35]	194	200
(35;36]	757	750
(36;37]	1858	1900
(37;38]	2714	2750
(38;39]	2494	2550
(39;40]	1367	1400
(40;41]	472	450 *
(41;42]	93	Keine Angabe des Kontingents
über 42	13	

* Es wurde derart gerundet, daß sich die Summe 10000 ergibt.

Tabelle 2 enthält daher einige geeignet gerundete Werte für einen reduzierten Größenbereich. Sie mahnt uns, daß die Ergebnisse einer statistischen Analyse interpretiert und in einer Weise dargestellt werden müssen, die mit den praktischen Erfordernissen der Anwendung in Einklang steht.