Optimierung von Stichprobenverfahren – Anwendungsbeispiele in der statistischen Qualitätssicherung

MANFRED BÖRGENS, WÖLFERSHEIM

Zusammenfassung: *In dieser Zeitschrift wurde ein stochastisches Optimierungsverfahren vorgestellt, das sich der Problemgruppe „Entscheidungen unter Unsicherheit“ zuordnen lässt (Börgens 2014). Es beruht auf zwei Verteilungen, die sich auf dieselbe Zufallsvariable (ZV) beziehen. Die eine Verteilung entsteht, wenn die ZV bei Vorliegen eines bestimmten Ereignisses E erhoben wird, die andere unter der Bedingung ¬ E. Das Optimalitätskriterium der Entscheidung auf E oder ¬ E beruht auf den Erträgen für richtige bzw. falsche Entscheidungen. Bei der Herleitung wurde ein Ansatz nach Bayes gewählt.*
Gut zu vermittelnde Anwendungsbeispiele finden sich bei der Auswertung von Stichproben im Rahmen der Qualitätssicherung. Diese sollen hier vorgestellt werden.

1 Einführung

In Börgens (2014) wurde ein Optimierungsverfahren beschrieben, das nur stochastische Grundkenntnisse voraussetzt. Die wichtigsten Ergebnisse aus diesem Beitrag werden in Abschnitt **2** kurz wiederholt.

Für den Stochastikunterricht ergeben sich durch die Behandlung der Optimierungsmethode und ihrer Anwendungen mehrere Perspektiven. Die Lernenden erhalten einen Einblick in Optimierungsaufgaben jenseits der Differentialrechnung, sie stärken ihre Kompetenz für die Modellierung statistischer Probleme und sie lernen mit der statistischen Qualitätssicherung ein industriell relevantes Arbeitsfeld kennen.

Alle Rechnungen lassen sich mit dem Taschenrechner oder mit Excel® durchführen.

Das Verfahren ermöglicht optimale Entscheidungen zwischen zwei Alternativen ¬ E = G und E = S. Das Optimalitätskriterium ist die Maximierung des erwarteten Ertrags. Als „Input“ werden zwei Wahrscheinlichkeitsverteilungen für dieselbe ZV benötigt, je eine unter der Bedingung G bzw. S. Für G und S müssen a-priori-Wahrscheinlichkeiten bekannt sein. Die Entscheidung kann A lauten („Annahme“ von G) oder Z („Zurückweisung“ von G). Die Erträge für die vier Kombinationen von G/S und A/Z gehören ebenfalls zum Input (4-Felder-Tafel). Damit erhält man eine Entscheidungsregel für A/Z.

2 Das Optimierungsverfahren

Die Realisation $X = x$ einer ZV X soll unter einer von zwei alternativen Bedingungen E oder ¬ E erfolgen. Es soll eine Entscheidung (unter Unsicherheit) herbeigeführt werden, ob $X = x$ eher E oder ¬ E nahelegt. Diese Entscheidung soll unter Ertragsgesichtspunkten optimiert werden.

Im Folgenden soll das Ereignispaar (¬ E, E) mit (G, S) identifiziert werden, was die Interpretation „Gut“ / „Schlecht“ nahelegt. Die Entscheidung wollen wir A nennen, wenn aufgrund der Realisation $X = x$ vom Vorliegen von G ausgegangen wird, ansonsten lautet die Entscheidung Z. Das Ereignispaar (A, Z) soll die Interpretation „Annahme“ / „Zurückweisung“ nahelegen.

Gemäß dem in Börgens (2014) angewendeten Bayes’schen Ansatz sind G / S konkurrierende Hypothesen mit den a-priori-Wahrscheinlichkeiten P(G) und $P(S) = 1 - P(G)$ sowie den Likelihoods $p_G(x) = P(X = x\,|G)$ und $p_S(x) = P(X = x\,|S)$. $p_G(x)$ und $p_S(x)$ sind somit Wahrscheinlichkeitsverteilungen für X. Die a-posteriori-Wahrscheinlichkeiten $P(G|\,X = x)$ und $P(S|\,X = x)$ lassen sich dann mit dem Satz von Bayes berechnen.

In einer Vierfeldertafel (Tabelle 1) werden Erträge für die vier Kombinationen von G/S und A/Z eingetragen.

	G	S
A	h_{AG}	h_{AS}
Z	h_{ZG}	h_{ZS}

Tabelle 1: Vierfeldertafel

Beispielsweise steht h_{AG} für den zu erwartenden Ertrag, wenn richtigerweise auf G entschieden wird; h_{AS} und h_{ZG} stehen für die „Kosten“ der beiden möglichen Fehlentscheidungen.

Aus den a-priori-Wahrscheinlichkeiten, den Likelihoods und der Vierfeldertafel wurde die folgende Entscheidungsregel hergeleitet (Börgens 2014):

Man setzt $\beta := (h_{AG} - h_{ZG}) / (h_{ZS} - h_{AS})$.

(1) Entscheidung A $\Leftrightarrow$
$\boldsymbol{\beta \cdot P(\,G|X = x\,) \;\geq\; P(\,S|X = x\,)}$.

Wegen $P(S|X = x) = 1 – P(G|X = x)$ lässt sich dies umformulieren zu:

(2) Entscheidung A $\Leftrightarrow$
$$\mathbf{P(G|X = x) \geq (1 + \beta)^{-1}}$$

Mit $\alpha := \beta \cdot P(G) / P(S)$ ergibt sich aus (1):

(3) Entscheidung A $\Leftrightarrow$ $\boldsymbol{\alpha \cdot p_G(x) \geq p_S(x)}$

Für die Entscheidung Z wird das Ungleichheitszeichen umgekehrt.

3 Grundlagen der Stichprobenverfahren

Wir betrachten Zahlenpaare (n, c) mit $n \in \mathbf{N}$ und $c \in \mathbf{N}_0$, $c < n$. Dabei steht n für den *Stichprobenumfang,* c heißt *Annahmezahl;* diese Darstellung findet man auch in der nationalen und internationalen Normung, siehe z. B. DIN (2005). Diese Begriffe erfordern eine kurze Beschreibung des betrieblichen Umfelds:

Es wird bei solchen Stichprobenziehungen davon ausgegangen, dass bei der Produktion einer großen Anzahl gleichartiger Stücke aus jedem auszuliefernden Los eine zufällige Stichprobe des Umfangs n geprüft wird. Nur Lose mit maximal c fehlerhaften Stücken (Ausschuss) in der Stichprobe werden ausgeliefert, wobei es für das hier behandelte Optimierungsverfahren irrelevant ist, ob die Ausschussstücke in der Stichprobe ausgetauscht werden (was nicht immer möglich ist). Lose mit mehr als c fehlerhaften Stichprobenstücken werden zurückgewiesen; ihre Behandlung kann sehr unterschiedlich sein (Vollkontrolle oder Vernichtung oder anderes), spielt aber hier keine Rolle.

Das Problem, das sich dem Betrieb stellt, liegt auf der Hand: Welche Annahmezahl c soll gewählt werden? Sowohl die Auslieferung eines Loses, das später vom Kunden (berechtigt) reklamiert wird, als auch die Zurückweisung eines Loses, das die Qualitätsanforderungen erfüllt, sind Fehlentscheidungen, die möglichst selten vorkommen sollen; die Wahl von (n, c) soll u. a. die Kosten dieser Fehlentscheidungen minimieren.

Als Zufallsvariable X wählen wir die Anzahl fehlerhafter Stücke in der Stichprobe.

Das Optimierungsproblem ist natürlich nicht zu lösen ohne die Kenntnis, wann ein Los eine für den Kunden akzeptable Qualität aufweist.

Die Lose sollen jeweils N gleichartige Stücke enthalten. Ein Los gilt als gut, wenn es höchstens M fehlerhafte Stücke enthält. Die Zahl M könnte beispielsweise mit dem Abnehmer der Ware vertraglich vereinbart worden sein; in diesem Fall hätte der Abnehmer ein Reklamationsrecht bei schlechten Losen, also bei mehr als M fehlerhaften Stücken. Man beachte, dass es manchmal, aber keineswegs immer, möglich ist, fehlerhafte Stichprobenstücke gegen fehlerfreie auszutauschen. Im Fall der Austauschmöglichkeit bezieht sich die Zahl M natürlich auf die verbleibenden fehlerhaften Stücke bei Auslieferung des Loses (dieser Fall wird uns im Fallbeispiel des Abschnitts **4** begegnen, der andere in Abschnitt **7**).

Als weitere Zufallsvariable Y wählen wir die Anzahl fehlerhafter Stücke im Los.

Nun kann man die Ereignispaare (A, Z) und (G, S) aus Abschnitt **2** angeben:

A	$X \leq c$	$\rightarrow$	Los ausliefern
Z	$X > c$	$\rightarrow$	Los zurückhalten
G	$Y \leq M$	$\rightarrow$	Los gut
S	$Y > M$	$\rightarrow$	Los schlecht

4 Empirische Verteilungen

Um die Anwendung der theoretischen Grundlagen aus den Abschnitten **2** und **3** vorzubereiten, soll zuerst ein Fallbeispiel vorgestellt werden.

Ein Betrieb stellt elektrische Klemmen her. In einer Produktionslinie umfassen die für den Verkauf bestimmten Lose jeweils $N = 1000$ Stück. Ein Los gilt als gut, wenn es höchstens 3% Ausschuss enthält; also ist $M = 30$. Aus jedem Los werden Stichproben mit $n = 50$ Stücken geprüft. Der Betrieb arbeitet mit der Vierfeldertafel in Tabelle 2 für die im Mittel erwarteten Erträge.

	G	S
A	250 €	-5250 €
Z	-110 €	-110 €

Tabelle 2: Vierfeldertafel für das Fallbeispiel

Wie kommen diese (fiktiv gewählten) Erträge zustande? Produktions- und Prüfkosten betragen zusammen pro Los 90 €. Der Verkaufspreis pro Los sei 340 €.

$h_{AG} = 250$ €: Verkaufspreis abzüglich Produktions- und Prüfkosten.

$h_{AS} = –5250$ €: h_{AG} abzüglich 5500 € Reklamations- und Folgekosten.

$h_{ZG} = h_{ZS} = -110$ €: Zusätzlich zu den Produktions- und Prüfkosten (–90 €) ist die Behandlung des zurückgewiesenen Loses zu bewerten; diese kann einen positiven oder negativen Ertrag erbringen (positiv z. B. bei Vollkontrolle und weiterer Verwertung, negativ z. B. bei Entsorgung) und wurde hier mit –20 € angesetzt.

Der Betrieb will die kostenminimale Annahmezahl c berechnen. Dafür ermittelt er empirisch, dass 96 % der Lose gut sind (a-priori-Wahrscheinlichkeit). Die empirisch ermittelte Verteilung der Anzahl der fehlerhaften Klemmen in den Stichproben ist in Abb. 1 sowie in den ersten drei Spalten von Tabelle 3 dargestellt. Auch hier wird (wie in Börgens (2014)) die relative Häufigkeitsverteilung als Schätzung für die Wahrscheinlichkeitsverteilung von X genommen. Damit sind die Likelihoods gegeben.

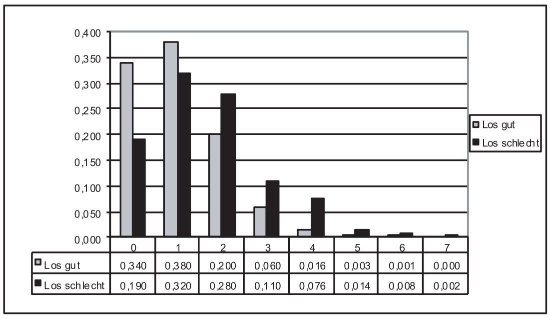

Abbildung 1: Empirische Häufigkeiten für Ausschusszahlen in der Stichprobe

Nun kann mit den Grundlagen aus den Abschnitten **2** und **3** die Entscheidungsregel für A, also für die Annahme eines Loses aufgrund der Stichprobe, berechnet werden. Diese Regel besteht lediglich aus der Bestimmung von c.

Man berechnet aus der Vierfeldertafel $\beta = 18/257$. Wegen $p(G) = 96/100$ folgt: $\alpha = 432/257$.

x	$p_G(x)$	$p_S(x)$	$\alpha^* p_G(x) - p_S(x)$		Ertrag für $c = x$
0	0,340	0,190	0,382	A	-31,56
1	0,380	0,320	0,319	A	33,98
2	**0,200**	**0,280**	**0,056**	**A**	**45,53**
3	0,060	0,110	-0,009	Z	43,65
4	0,016	0,076	-0,049	Z	33,55
5	0,003	0,014	-0,009	Z	31,71
6	0,001	0,008	-0,006	Z	30,41
7	0,000	0,002	-0,002	Z	30,00

Tabelle 3: Alternative Verteilungen und Entscheidung für das Fallbeispiel

Nach (3) wird das Los angenommen, falls $\alpha \cdot p_G(x) - p_S(x) \geq 0$. Die vorletzte Spalte von Tabelle 3 weist aus, dass dies für $x \leq 2$ der Fall ist, also ist c = 2. Die letzte Spalte wird erst in Abschnitt **8** verwendet.

Lose werden also ausgeliefert, falls in der Stichprobe maximal zwei fehlerhafte Stücke gefunden werden.

5 Anwendung der Binomialverteilung

Die Anzahl fehlerhafter Stücke ist binomialverteilt (in jeder Menge produzierter Teile, z. B. im Los oder in der Stichprobe), falls der Herstellungsprozess „beherrscht" ist (dies wird oft auch mit „statistisch kontrolliert" bezeichnet). Prozessbeherrschung ist ein zentrales Thema der modernen Qualitätssicherung, siehe z. B. Timischl (2012). In einem beherrschten Produktionsprozess treten fehlerhafte Stücke im Los zufällig auf und der Erwartungswert ihrer Anzahl ist bekannt und mit den Qualitätserwartungen vereinbar. Diese Zufälligkeit bedeutet mathematisch, dass jedes produzierte Teil dieselbe Wahrscheinlichkeit p für einen Defekt aufweist. Unter diesen Voraussetzungen lässt sich von den Kriterien am Ende von Abschnitt **2** am leichtesten (2) anwenden. Man benötigt dann den Satz von Bayes nicht, der für (3) verwendet wurde, sondern setzt nur die Eigenschaften der Binomialverteilung ein. Bei einer Binomialverteilung mit den Parametern m und q soll $p_{m,q}$ für die Verteilung und $f_{m,q}$ für die Verteilungsfunktion stehen

Wie in Abschnitt **3** sei $X = x$ eine Realisation der ZV „Anzahl fehlerhafter Stücke in der Stichprobe". Ein gutes Los darf maximal $M - x$ fehlerhafte Stücke im „Rest", also unter den $N - n$ nicht geprüften Stücken haben. Die Wahrscheinlichkeit dafür ist

$$P\left(G \mid X = x\right) = f_{N-n,p}\left(M - x\right) = \sum_{k=0}^{M-x} p_{N-n,p}\left(k\right) =$$

$$= \sum_{k=0}^{M-x} \binom{N-n}{k} \cdot p^{k} \cdot (1-p)^{N-n-k} .$$

(Die leere Summe wird = 0 gesetzt; dies tritt hier bei $x > M$ auf.)

$P(G \mid X = x)$ fällt also streng monoton mit x, somit lässt sich das maximale x mit $P(G \mid X = x) \geq (1 + \beta)^{-1}$ eindeutig bestimmen; nach (2) ist dieses x das gesuchte c für den Stichprobenplan. Wir erhalten damit:

(4) $\mathbf{c = \max \{ x < n \mid f_{N-n,p}(M - x) \geq (1 + \beta)^{-1} \}}$.

Falls es kein x mit (4) gibt, müsste man alle Lose zurückweisen; dies wäre zu interpretieren als schwerwiegendes Qualitätsproblem in der Produktion.

p_G und p_S werden in Abschnitt **8** benötigt. Für die Lernenden ist das eine weitere Gelegenheit, den Satz von Bayes anzuwenden:

$$p_G(x) = P(X = x \mid G) = \frac{P(G \mid X = x) \cdot P(X = x)}{P(G)} = $$
$$= \frac{f_{N-n,\,p}(M - x) \cdot p_{n,\,p}(x)}{f_{N,\,p}(M)}$$
$$p_S(x) = P(X = x \mid S) = \frac{P(S \mid X = x) \cdot P(X = x)}{P(S)} = $$
$$= \frac{\left(1 - f_{N-n,\,p}(M - x)\right) \cdot p_{n,\,p}(x)}{1 - f_{N,\,p}(M)}.$$

Was kann man aus diesem Abschnitt lernen? Beherrschte Prozesse erlauben die Verwendung der Binomialverteilung; die Erhebung einer empirischen Verteilung wie in Abschnitt **4** ist nicht erforderlich. Die wesentliche zusätzliche Information steckt in der Ausschusswahrscheinlichkeit p, deren Kenntnis das Kriterium (4) ermöglicht. Ein Los wird angenommen, wenn sich maximal c fehlerhafte Stücke in der Stichprobe finden, und c ist mit (4) leicht zu berechnen. Ein ausgearbeitetes Beispiel wird in Abschnitt **7** gegeben.

6 Anwendung der Poisson-Verteilung

Bei manchen produzierten Gütern lassen sich die Lose nicht in „Stücke" einteilen. Solche Lose sind z. B. festgelegte Mengen einer Flüssigkeit, eines Baustoffs oder eines Gewebes (etwa 1 Liter einer Chemikalie, 10 m^2 einer beschichteten Platte oder eines Kleiderstoffs). In diesem Fall sprechen wir von „kontinuierlichen" Losen (im Gegensatz zu den bisher betrachteten „diskreten" Losen); dann werden nicht „fehlerhafte Stücke", sondern „Fehler" im Los gezählt. Fehler können mikroskopisch kleine Fremdkörper in einer Flüssigkeit sein, Kratzer oder Einschlüsse auf einer beschichteten Platte, Web- oder Farbfehler im Stoff. Ein Los gilt als gut, wenn es höchstens M Fehler enthält. Der Hersteller prüft aus jedem Los einen zufällig ausgewählten Anteil r, also z. B. 100 cm^2 der Plattenoberfläche (dann wäre $r = 0{,}001$). Das Los wird ausgeliefert, wenn die Stichprobe höchstens c Fehler enthält. M und c sind also wie in Abschnitt **5** gewählt; r tritt an die Stelle von n; eine Losgröße N gibt es hier nicht, sondern ein Los wird als eine Einheit betrachtet.

Schaut man sich unter diesen Voraussetzungen noch einmal das Fallbeispiel in Abschnitt **4** an, so könnte es ebenso als Beispiel für kontinuierliche Lose genommen werden. Denn bei der Anwendung des Kriteriums (3) wurden keine Verteilungsparameter benötigt, sondern nur die empirische Verteilung aus Abb. 1 und Tabelle 3 (dort wäre dann lediglich „Fehler" statt „Ausschuss" bzw. „fehlerhafte Stücke" einzusetzen). Das gilt natürlich über das Fallbeispiel hinaus:

Liegen die Likelihoods als empirische Verteilungen vor, spielt es keine Rolle, ob es sich um diskrete oder kontinuierliche Lose handelt. In beiden Fällen wird die Entscheidung A mit (3) getroffen.

Im Falle eines *beherrschten* Produktionsprozesses treten analog zum Abschnitt **5** Fehler im Los zufällig auf und der Erwartungswert ihrer Anzahl ist bekannt und mit den Qualitätserwartungen vereinbar. Was ändert sich gegenüber diskreten Losen? Da man von einem zufälligen Auftreten der Fehler ausgeht, ist die Fehlerzahl in jeder produzierten Menge Poisson-verteilt. Der Parameter λ der Poisson-Verteilung für das Gesamtlos ist auch der Erwartungswert. Daraus ergibt sich der Parameter $\lambda \cdot r$ für die Stichprobe und $\lambda \cdot (1-r)$ für den Rest des Loses. Bei einer Poisson-Verteilung mit Parameter λ soll p_λ für die Verteilung und f_λ für die Verteilungsfunktion stehen.

$X = x$ ist hier eine Realisation der ZV „Anzahl Fehler in der Stichprobe". Ein gutes Los darf maximal $M - x$ fehlerhafte Stücke im „Rest", also außerhalb der Stichprobe haben. Die Wahrscheinlichkeit dafür ist

$$P(G \mid X = x) = f_{\lambda \cdot (1-r)}(M - x) = \sum_{k=0}^{M-x} p_{\lambda \cdot (1-r)}(k) =$$
$$= e^{-\lambda(1-r)} \cdot \sum_{k=0}^{M-x} \frac{(\lambda(1-r))^k}{k!}.$$

$P(G \mid X = x)$ fällt streng monoton mit x, also lässt sich das maximale x mit $P(G \mid X = x) \geq (1 + \beta)^{-1}$ eindeutig bestimmen; nach (2) ist dieses x das gesuchte c für den Stichprobenplan. Somit gilt:

(5) $\mathbf{c = \max\{\, x \mid f_{\lambda \cdot (1-r)}(M - x) \geq (1 + \beta)^{-1} \,\}}$.

Falls es kein x mit (5) gibt, muss man alle Lose zurückweisen.

p_G und p_S (für Abschnitt **8**) werden wieder mit dem Satz von Bayes bestimmt:

$$p_G(x) = P(X = x | G) = \frac{P(G | X = x) \cdot P(X = x)}{P(G)} =$$

$$= \frac{f_{\lambda\cdot(1-r)}(M - x) \cdot p_{\lambda\cdot r}(x)}{f_\lambda(M)}$$

$$p_S(x) = P(X = x | S) = \frac{P(S | X = x) \cdot P(X = x)}{P(S)} =$$

$$= \frac{\left(1 - f_{\lambda\cdot(1-r)}(M - x)\right) \cdot p_{\lambda\cdot r}(x)}{1 - f_\lambda(M)}.$$

Ein ausgearbeitetes Beispiel wird in Abschnitt **7** gegeben.

7 Fallbeispiel

Das folgende Anwendungsbeispiel führt die beiden Varianten (diskrete Lose, siehe Abschnitt **5**; kontinuierliche Lose, siehe Abschnitt **6**) parallel durch. Dies verfolgt den Zweck, die weitgehende methodische Übereinstimmung deutlich zu machen, aber auch die Abweichungen.

Ein Betrieb ermittelt empirisch durch Langzeitbeobachtung, dass ein produziertes Los einer bestimmten Ware im Durchschnitt 1,8 fehlerhafte Stücke bzw. Fehler aufweist. Gute Lose dürfen gemäß einer Liefervereinbarung maximal 3 fehlerhafte Stücke / Fehler enthalten. Jedes Los soll in Zukunft vor Warenausgang einer zufälligen Stichprobenprüfung unterzogen werden; der Stichprobenumfang soll immer 5% eines Loses ausmachen. Tabelle 4 zeigt die Vierfeldertafel (fiktive Werte; Interpretation wie bei Tabelle 2).

	G	S
A	60 €	-510 €
Z	-10 €	-10 €

Tabelle 4: Vierfeldertafel für das Fallbeispiel

Im Falle der diskreten Lose sei die Losgröße N = 60. Wir erhalten damit die Parameter in Tabelle 5. Die Stichprobenumfänge entsprechen dabei 5% des Loses. $\lambda = 1{,}8$ wurde als Parameter der Poisson-Verteilung vom Schätzwert (empirischer Mittelwert) übernommen. Der Erwartungswert der Binomialverteilung für das Gesamtlos ist N·p; auch hier übernehmen wird den empirischen Mittelwert 1,8, so dass sich aus N = 60 ergibt: p = 0,03. β wurde nach Abschnitt 2 berechnet.

Es sollen nun die Regeln (4) und (5) zur Anwendung kommen. Dafür werden in den Tabellen 6 und 7 $f_{N\text{-}n\,,\,p}(M - x)$ und $f_{\lambda\,(1-r)}(M - x)$ berechnet. Das ist z. B. mit Excel® leicht möglich, da dort die Verteilungsfunktionen verfügbar sind.

	diskret	kontinuierlich
Losgröße	N = 60	-
Stichprobenumfang	n = 3	r = 0,05
Reklamationsgrenze	M = 3	
Verteilung	Binomialverteilung	Poisson-Verteilung
Parameter in Verteilung für das Los	N = 60 p = 0,03	λ = 1,8
Parameter in Verteilung für die Stichprobe	n = 3 p = 0,03	λ·r = 0,09
β	0,14	
$(1+\beta)^{-1}$	$^{50}/_{57} \approx$ **0,877**	

Tabelle 5: Verteilungsparameter für das Fallbeispiel

Die Tabellen 6 und 7 zeigen, dass als optimale Annahmezahl $c = 0$ zu wählen ist, da in der zweiten Spalte $(1+\beta)^{-1}$ nur für $x = 0$ übertroffen wird. Die letzten drei Spalten wurden schon für Abschnitt **8** berechnet.

Lose werden also nur dann ausgeliefert, wenn die Stichprobe völlig fehlerlos ist.

x	$f_{N\text{-}n,\,p}(M - x) = f_{57,\,0{,}03}(3 - x)$		p_G	p_S	Ertrag für c = x
0	**0,908**	**A**	**0,92700**	**0,79150**	**6,18**
1	0,756	Z	0,07157	0,19558	0,32
2	0,487	Z	0,00143	0,01271	-0,26

Tabelle 6: Werte der Verteilungsfunktion (Binomialverteilung)

x	$f_{\lambda\,(1\text{-}r)}(M - x) = f_{1{,}71}(3 - x)$		p_G	p_S	Ertrag für c = x
0	**0,905**	**A**	**0,92830**	**0,79608**	**4,65**
1	0,755	Z	0,06964	0,18570	-1,10
2	0,490	Z	0,00204	0,01736	-1,92
3	0,181	Z	0,00002	0,00084	-1,96

Tabelle 7: Werte der Verteilungsfunktion (Poisson-Verteilung)

8 Erwarteter Ertrag

Aus Börgens (2014) übernehmen wir den erwarteten Ertrag

$$H = h_{AG}\cdot P(A \mid G)\cdot P(G) + h_{AS}\cdot P(A \mid S)\cdot p(S)$$
$$+ h_{ZG}\cdot P(Z \mid G)\cdot p(G) + h_{ZS}\cdot P(Z \mid S)\cdot p(S)$$
$$P(A|G) = \Sigma_{x \leq c}\, p_G(x) \qquad P(Z|G) = 1 - P(A|G)$$
$$P(A|S) = \Sigma_{x \leq c}\, p_S(x) \qquad P(Z|S) = 1 - P(A|S)$$

H soll zuerst für das Fallbeispiel aus Abschnitt **4** berechnet werden. c, P(G), $p_G(x)$ und $p_S(x)$ sind dort angegeben. Damit erhält man Tabelle 8.

$P(A \mid G) = \Sigma_{x \leq 2}\, p_G(x) = 0{,}92$	$h_{AG} = 250$
$P(A \mid S) = \Sigma_{x \leq 2}\, p_S(x) = 0{,}79$	$h_{AS} = -5250$
$P(Z \mid G) = 1 - P(A \mid G) = 0{,}08$	$h_{ZG} = -110$
$P(Z \mid S) = 1 – P(A \mid S) = 0{,}21$	$h_{ZS} = -110$
P(G) = 0,96 P(S) = 0,04	
H = 45,528 (Als durchschnittlicher **Gewinn** pro Los ist 45,528 € zu erwarten.)	

Tabelle 8: Erwarteter Ertrag für empirische Verteilung

Dieses H taucht auch in Tabelle 3 auf (dort auf volle Cent gerundet). Um zu verdeutlichen, dass es sich tatsächlich um den maximal möglichen Ertrag handelt, wurde in Tabelle 3 in der letzten Spalte der jeweilige Ertrag angegeben, der sich bei der Berechnung von H (wie in Tabelle 8) durch Variation der Annahmezahlen c ergeben hätte.

Nun soll H für das Fallbeispiel in Abschnitt **7** berechnet werden. Tabelle 9 bezieht sich auf diskrete Lose, verwendet also die Binomialverteilung. P(G) wurde bisher nicht angegeben. Das ist auch nicht nötig, da es sich berechnen lässt:
$P(G) = f_{N,\,p}(\,M\,) = f_{60,\,0{,}03}(\,3\,) \approx 0{,}894$.

$P(A \mid G) = p_G(0) \approx 0{,}927$	$h_{AG} = 60$
$P(A \mid S) = p_S(0) \approx 0{,}7915$	$h_{AS} = -510$
$P(Z \mid G) = 1 - p_G(0) \approx 0{,}073$	$h_{ZG} = -10$
$P(Z \mid S) = 1 – p_S(0) \approx 0{,}2085$	$h_{ZS} = -10$
P(G) ≈ 0,894 P(S) ≈ 0,106	
H ≈ 6,18 (Als durchschnittlicher **Gewinn** pro Los ist 6,18 € zu erwarten.)	

Tabelle 9: Erwarteter Ertrag für Binomialverteilung

$f_G(0) = p_G(0)$ und $f_S(0) = p_S(0)$ sind Tabelle 6 entnommen. H wurde mit Hilfe der ungerundeten Zwischenergebnisse berechnet.

Dieses maximale H finden wir (gerundet) auch in der letzten Spalte von Tabelle 6, zusammen mit den anderen Erträgen, die sich durch Variation der Annahmezahlen c ergeben hätten.

Analog geht man bei kontinuierlichen Losen vor. P(G) berechnet man dann mit der Poisson-Verteilung: $P(G) = f_{\lambda}(\,M\,) = f_{1{,}8}(\,3\,) \approx 0{,}891$. Damit erhält man Tabelle 10.

$P(A \mid G) = P_G(0) \approx 0{,}928$	$h_{AG} = 60$
$P(A \mid S) = P_S(0) \approx 0{,}796$	$h_{AS} = -510$
$P(Z \mid G) = 1 - P_G(0) \approx 0{,}072$	$h_{ZG} = -10$
$P(Z \mid S) = 1 – P_S(0) \approx 0{,}204$	$h_{ZS} = -10$
P(G) ≈ 0,891 P(S) ≈ 0,109	
H ≈ 4,647 (Als durchschnittlicher **Gewinn** pro Los ist 4,647 € zu erwarten.)	

Tabelle 10: Erwarteter Ertrag für Poisson-Verteilung

Dieses maximale H finden wir (gerundet) wieder in der letzten Spalte von Tabelle 7, zusammen mit den anderen Erträgen, die sich durch Variation der Annahmezahlen c ergeben hätten.

Literatur

Börgens, M. (2014): Optimierungsverfahren für alternative statistische Verteilungen. In: *Stochastik in der Schule* 34 (2), S. 2–8.

DIN (Deutsches Institut für Normung, Hrsg.) (2005): Qualitätssicherung und angewandte Statistik, Verfahren II: Probenahme und Annahmestichprobenprüfung, Normen. DIN-Taschenbuch 225, 3. Aufl. Berlin: Beuth-Verlag.

Timischl, W. (2012): Qualitätssicherung – Statistische Methoden, 4. Aufl. München/Wien: Hanser-Verlag

Anschrift des Verfassers

Manfred Börgens
Technische Hochschule Mittelhessen
Wilhelm-Leuschner-Straße 13
61169 Friedberg
manfred.boergens@mnd.thm.de