Annahmekennlinien

Eine Anwendung der Poisson-Verteilung in der Praxis

J. Schärf, Wien

Eine gut funktionierende Qualitätskontrolle ist eine Voraussetzung für das Überleben eines Betriebes in einer freien Wirtschaft. Bei Massenfertigung ist die statistische Qualitätskontrolle, kurz SQK genannt, ein unentbehrliches Hilfsmittel der Qualitätssicherung. Man kann zwei Hauptzweige der SQK unterscheiden, die statistische Qualitätsregulierung (Kontrollkartentechnik) und die statistischen Stichprobenverfahren.

Im ersten Fall wird die Erzeugung mit Hilfe statistischer Methoden kontrolliert und gesteuert.

Im zweiten Fall wird mit Hilfe statistischer Methoden von der Qualität von Stichproben auf die Qualität der Grundgesamtheit geschlossen. Bei zerstörenden Prüfungen ist dies der einzig mögliche Weg (Photoblitze!).

Man kann ferner zwischen einer Prüfung quantitativer, also meßbarer Merkmale und einer Prüfung qualitativer Merkmale unterscheiden. Im ersten Fall spricht man von einer Variablenprüfung, im zweiten Fall von einer Attributprüfung. Bei der Attributprüfung wird nur zwischen "gut" und "schlecht" entschieden (z.B. Kaliberkontrolle).

Wir beschränken uns hier auf attributive Stichprobenprüfung.

Zunächst definieren wir die Begriffe Grundgesamtheit, Los, Stichprobe und Annahmezahl.

Die Menge der in Betracht gezogenen Einheiten heißt Gundgesamtheit.

Die Anzahl der Einheiten der Grundgesamtheit heißt Umfang der Grundgesamtheit.

Eine echte Teilmenge der Grundgesamtheit vom Umfang N heißt dann Los (Prüflos!), wenn sie als Gesamtheit beurteilt werden soll.

Eine Stichprobe vom Umfang n umfaßt n Einheiten, die einem Los zur Prüfung entnommen werden.

Die maximal zugelassene Anzahl fehlerhafter Einheiten in der Stichprobe, bei der das Los noch angenommen wird, heißt Annahmezahl c.

Modell: Eisenbahnwaggon mit 200 Kisten je 1000 Einheiten. Jeder Kiste soll eine Stichprobe vom Umfang 100 entnommen werden. Wenn in einer solchen Stichprobe 3 oder weniger Fehler auftreten, dann soll die Kiste angenommen werden. Treten 4 oder mehr Fehler in der Stichprobe auf, dann soll auf Kosten des Erzeugers jede in dieser Kiste enthaltene Einheit einzeln geprüft werden. Ergibt diese Prüfung, daß der vereinbarte Fehlerprozentsatz 3% überschritten ist, sich also mehr als 30 fehlerhafte Stücke in der Kiste befinden, dann ist der Inhalt dieser Kiste, selbstverständlich zu Lasten des Erzeugers, zu vernichten.

Erzeugung ...... Grundgesamtheit

200 Kisten ..... betrachtete Gesamtheit

Kiste .... Los vom Umfang N = 1000

n = 100 .... Stichprobenumfang

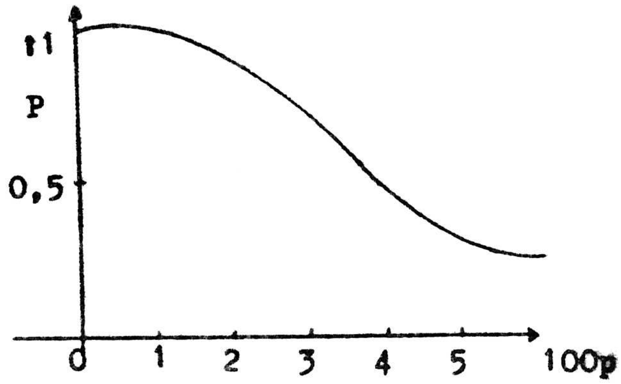
c = 3 .... Annahmezahl

Wenn die Grundgesamtheit einen Fehlerprozentsatz von z.B. 1% hat, dann gilt:

$$\text{Empirische Annahmewahrscheinlichkeit } P = \frac{\text{Anzahl der angenommenen Lose}}{\text{Anzahl der betrachteten Lose}}$$

(bei p = 1%)

Wenn man P für verschieden Fehlerprozentsätze ermittelt, so erhält man eine Funktion, deren graphische Darstellung <u>Annahmekennlinie</u> für n = 100 und c = 3 heißt.



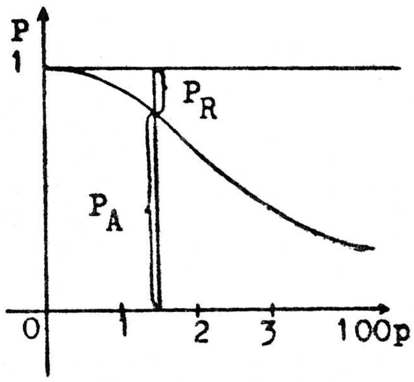
Wir berechnen für unser Modell die empirischen Annahmewahrscheinlichkeiten mittels gleich verteilter Zufallszahlen mit einem Tischcomputer.

Ein durchschnittlicher Schüler kann dieses BASIC-Programm in weniger als 15 Minuten schreiben.

```
100 RANDOM : FOR K=1 TO 200
110 FOR I=1 TO 8 : F(I)=0 : NEXT I
120 FOR I=1 TO 100 : Z=RND(100) : IF Z<9 THEN F(Z)=F(Z)+1
130 NEXT I
140 FOR I=2 TO 8 : F(I)=F(I)+F(I-1) : NEXT I
150 FOR I=1 TO 8 : IF F(I)<=3 THEN K(I)=K(I)+1
160 NEXT I,K
170 FOR I=1 TO 8 : PRINT I,K(I)/200 : NEXT I
```

Wir erhalten folgende Werte[1]:

100p	1	2	3	4	5	6
P_{emp}	0,99	0,87	0,66	0,41	0,22	0,11
P_{theor}	0,98	0,86	0,65	0,43	0,27	0,15

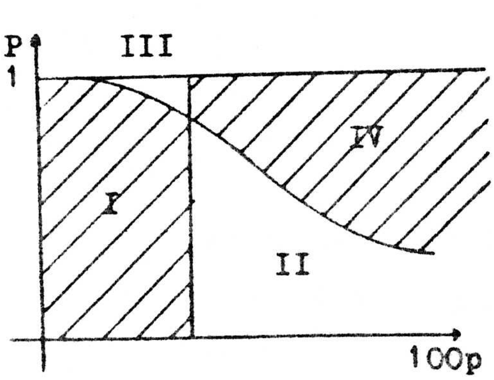


P_A ... Annahmewahrscheinlichkeit bei p = 1,5%

P_R ... Rückweisewahrscheinlichkeit bei p = 1,5%

Jede Stichprobenprüfung ist mit einem Risiko verbunden, das sich auf den Lieferanten und den Kunden verteilt.
Annahmekennlinien sind ein unentbehrliches Mittel für die Beurteilung von Lieferanten - und Kundenrisiko.

Wir unterscheiden bzgl. des Risikos vier Zonen:



Zone I Annahmewahrscheinlichkeit guter Lose.
Die Lieferung wird zu Recht angenommen.

[1] In der letzten Zeile befinden sich zum Vergleich die theoretischen Werte der Annahmewahrscheinlichkeit.

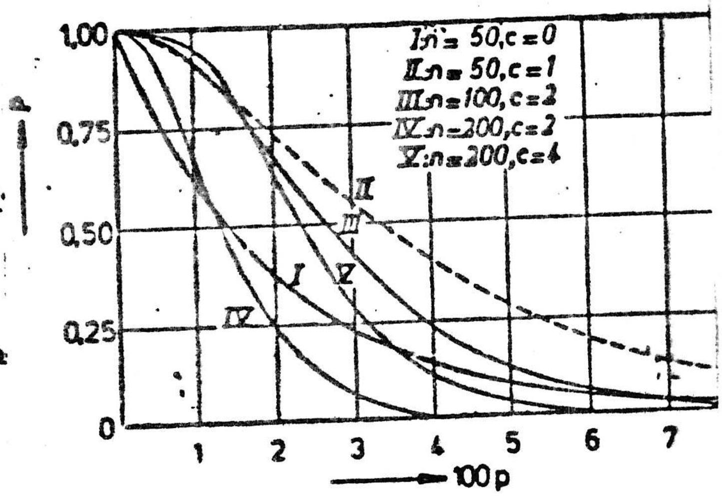
Zone II Annahmewahrscheinlichkeit schlechter Lose.
Die Lieferung wird zu Unrecht angenommen.

Zone III Rückweisewahrscheinlichkeit guter Lose.
Die Lieferung wird zu Unrecht zurückgewiesen.

Zone IV Rückweisewahrscheinlichkeit schlechter Lose.
Die Lieferung wird mit Recht zurückgewiesen.

Die Risikozonen III und II können durch die Wahl steilerer Annahmeknnlinien verkleinert werden.

Wir wollen nun anhand der nebenstehenden Abbildungen die wirtschaftliche Bedeutung der Annahmekennlinien untersuchen.



Wir betrachten zunächst die Stichprobenpläne II(n = 50; c = 1) und III (n = 100; c = 2).

"Prozentuelles Denken" führt auf Abwege; eine Stichprobe mit n,c verhält sich anders als eine doppelt so große Stichprobe mit der Annahmezahl 2c.

Bei etwa 1.5% Ausschuß der Grundgesamtheit haben beide Stichproben die Annahmewahrscheinlichkeit 0,82.

Bei p = 3% wird das Los II in 56 von 100 Fällen, das Los III aber nur in 42 von 100 Fällen angenommen.

Bei p = 4% und p = 5%, ... ist der Unterschied noch größer.

Ein Lieferant, der die schlechte Qualität seiner Erzeugnisse kennt, wird daher versuchen, daß das Schema II für die Abnahme gewählt wird. Als Vorwand für diese Wahl wird wohl die billigere Stichprobenentnahme herhalten müssen.
Der Kunde muß immer bedenken, daß $n = 100$, $c = 2$ keine Garantie für nur 2% im Los sind.
Er muß wissen, daß jedem Paar n, c genau eine Annahmekennlinie entspricht und umgekehrt.
Stichprobenpläne sollen sowohl die Interessen des Lieferanten als auch die des Kunden berücksichtigen.

<u>Berechnung von Annahmekennlinien</u>

Annahmekennlinien kann man mit der Hypergeometrischen Verteilung oder der Binomialverteilung oder mit der Poisson-Verteilung berechnen.
Hypergeometrische Verteilung:

$$P(x) = \frac{\binom{Np}{x} \cdot \binom{N-Np}{n-x}}{\binom{N}{n}}$$

ist die Wahrscheinlichkeit, daß sich in einer Stichprobe vom Umfang n, die einem Los vom Umfang N entnommen wurde, x defekte Einheiten befinden.
Für großes n und kleines p kann man statt der Hypergeometrischen Verteilung die Binomialverteilung verwenden.
Es gilt dann:

$$P(x) = \binom{n}{x} p^x q^{n-x}$$

Für $p \to 0 \wedge n \to \infty \wedge np \to \mu$ geht die Binomialverteilung in die Poisson-Verteilung über. Eine einfache Herleitung der Poisson-Verteilung stammt von Varangot:

$$P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$P(x-1) = \frac{n!}{(x-1)!(n-x+1)!} p^{x-1} q^{n-x+1}$$

$$\frac{P(x)}{P(x-1)} = \frac{n-x+1}{x} \cdot \frac{p}{q} = \frac{np(1 - \frac{x}{n} + \frac{1}{n})}{x} \cdot \frac{1}{q}$$

Für $n \to \infty \wedge p \to 0 \wedge np \to \mu \wedge q \to 1$ erhalten wir die Rekursionsformel

$$P(x) = \frac{\mu}{x} \cdot P(x-1)$$

$$P(1) = \frac{\mu}{1} \cdot P(0)$$

$$P(2) = \frac{\mu}{2} \cdot P(1) = \frac{\mu^2}{2!} P(0)$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$P(x) = \dots\dots\dots = \frac{\mu^x}{x!} P(0)$$

Bei der Binomialverteilung gilt

$$P(0) = q^n = (1-p)^n = (1 - \frac{np}{n})^n = (1 - \frac{\mu}{n})^n$$

Mit $\lim_{n \to \infty} (1 - \frac{\mu}{n})^n = e^{-\mu}$ erhalten wir somit:

$$\underline{P(x) = \frac{e^{-\mu} \cdot \mu^x}{x!}}$$

Querverbindungen zum EDV-Unterricht:

(1) Hypergeometrische Verteilung (2) Binomialverteilung	1. Version mit Direktformel
(3) Poissonverteilung (4) 1 - 3 in einem Programm	2. Version mit Rekursionsformel

In der Praxis hat man es oft nicht mit der Wahrscheinlichkeitsfunktion, sondern mit der Verteilungsfunktion zu tun.
Bei diskreten Verteilungen gilt:

$$F(x) = P(X \leq x) = \sum_{k=0}^{x} P(k)$$

Somit: $P(x) = F(x) - F(x-1)$

Querverbindung zur EDV: Tabellen der Verteilungsfunktionen der diskreten Verteilungen.
Nun zu unserer Annahmeknnlinie für $n = 100$ und $c = 3$.
Die Annahmewahrscheinlichkeit ist die Wahrscheinlichkeit dafür, daß sich in der Stichprobe c oder weniger Fehler befinden. Wir müssen daher mit den Verteilungsfunktionen arbeiten. Somit gilt:

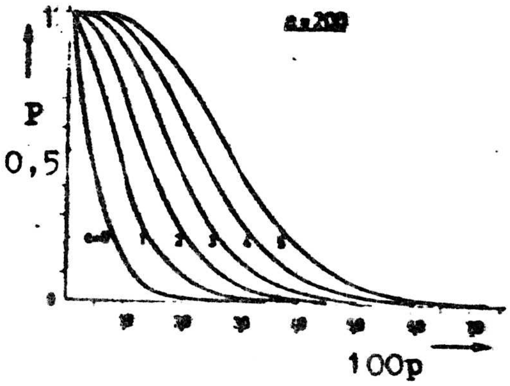
$$P_A(c) = F(c).$$

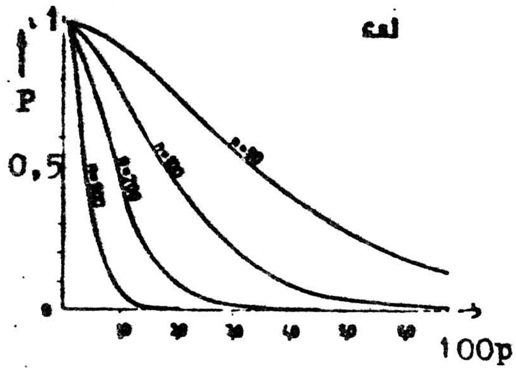
P	$\mu = np$	$F(\mu; c = 3)$
0	0	1
0,01	1	0,981
0,02	2	0,857
0,03	3	0,647
0,04	4	0,433
0,05	5	0,265
0,06	6	0,151
0,07	7	0,082

Die Werte der Poisson-Verteilungsfunktion entnehmen wir der Tabelle.

Querverbindungen zur EDV: Zeichnen einer Annahmekennlinie.

Es ergibt sich die Frage: Wie wirken sich Änderungen von n oder c auf die Form von Annahmekennlinien aus?

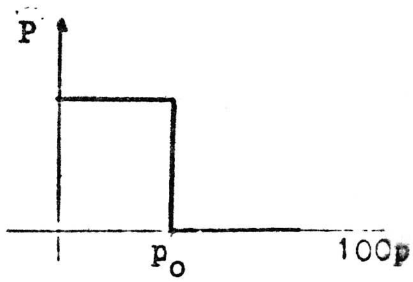




Wir erkennen:

Mit kleiner werdendem c und mit wachsendem n werden die Kurven steiler; die Trennschärfe wächst.

Grenzfall ist die nebenstehende Verteilung.



Es gibt viele Stichprobenpläne. Jeder dieser Pläne ist durch zwei Grundgrößen bestimmt, durch den Prüfungspunkt p_o und durch die Steilheit h_o.

Zwischen h_o, p_o, n und c bestehen die von Campell und Hamaker zum Teil theoretisch, zum Teil empirisch gefundenen Beziehungen

$$np_o = c + 0{,}67$$

$$\frac{\pi}{2} h_o^2 = np_o + 0{,}06 = c + 0{,}73$$

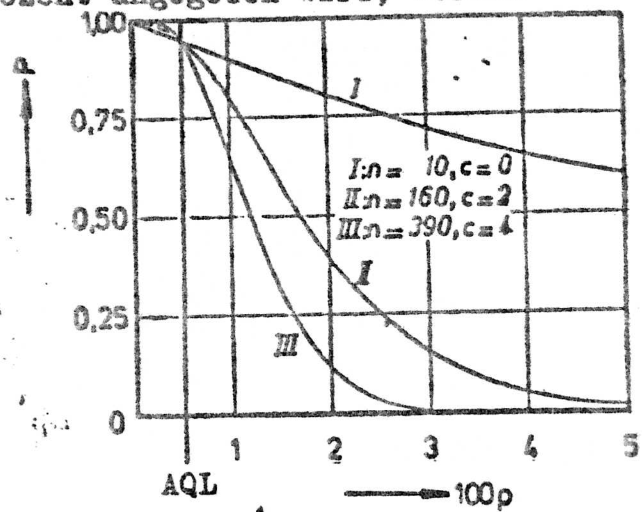
die auf der Poisson-Verteilung beruhen.

Sehr verbreitet ist ein Stichprobensystem, das als Prüfprozent AQL, die annehmbare Qualitätsgrenzlage (acceptable quality level) benützt.

Beim AQL-Wert, der üblicherweise in Prozent angegeben wird, werden die Lose in etwa 95 von 100 Fällen angenommen.

Wie berechnet man diese Annahmekennlinien?

Es sind c und P(AQL) = 0,95 AQL = 0,5% vorgegeben.



Für c = 0 und AQL = 0,5% erhalten wir aus der Tabelle[1] $\mu = 0,05$.

Somit $n \cdot 0,005 = 0,05 \Rightarrow$ <u>n = 10</u>.

Für c = 2 und AQL = 0,5% erhalten wir:

$F(\mu;2) = 0,95 \Rightarrow \mu = 0,80 \Rightarrow n \cdot 0,005 = 0,80 \Rightarrow$ <u>n = 160</u>.

Stichprobenpläne, bei denen die Entscheidung über das Los nach der Beurteilung einer einzigen Stichprobe erfolgt, heißen einfache Stichprobenpläne.

Bei doppelten Stichprobenplänen kann die Entscheidung bereits nach der Beurteilung einer Stichprobe fallen, sie muß aber nach der Beurteilung der zweiten Stichprobe fallen.

---

[1] $x = 0;\ F(\mu,0) = 0,95$

Vorgangsweise: Die Anzahl i_1 der fehlerhaften Einheiten der 1. Stichprobe wird mit c_1 verglichen.

Ist $i_1 \leq c_1$, dann wird das Los angenommen.

Ist $i_1 > r_1$, dann wird das Los gesperrt.

Gilt $c_1 < i_1 < r_1$, dann wird eine zweite Stichprobe mit dem Umfang n_2 durchgeführt.

Für die Auswertung dieser zweiten Prüfung sind die Summenzahlen $n_1 + n_2$, S_c, S_r ausschlaggebend.

Falls $i_1 + i_2 \leq S_c$ ist, so wird das Los angenommen.

Falls $i_1 + i_2 > S_r$ ist, so wird das Los gesperrt.

In der ÖNORM A 6650 ist ein vereinfachtes Verfahren für attributive Stichprobenprüfung angegeben, dem der AQL zugrunde liegt. Es sind dort für einfache und für doppelte Stichprobenpläne zu gegebenem Losumfang Stichprobenumfang nebtst zugehörigem c und r (Rückweisezahl) angegeben.

Auszüge aus den Tabellen befinden sich in der 13. Auflage meiner Funktionen- und Zahlentafeln.

Einfacher Stichprobenplan zur attributiven Kontrolle [1]

Losumfang N AQL		0,4%			0,65%			1,0%			1,5%		
von	bis	n	c	r	n	c	r	n	c	r	n	c	r
1	25	*32	0	1	*20	0	1	*13	0	1	*8	0	1
26	50	*32	0	1	20	0	1	13	0	1	8	0	1
51	90	32	0	1	20	0	1	13	0	1	8	0	1
91	150	32	0	1	20	0	1	13	0	1	32	1	2
151	280	32	0	1	20	0	1	50	1	2	32	1	2
281	500	32	0	1	80	1	2	50	1	2	50	2	3
501	1 200	125	1	2	80	1	2	80	2	3	80	3	4
1 201	3 200	125	1	2	125	2	3	125	3	4	125	5	6
3 201	10 000	200	2	3	200	3	4	200	5	6	200	7	8
10 001	35 000	315	3	4	315	5	6	315	7	8	315	10	11
35 001	150 000	500	5	6	500	7	8	500	10	11	500	14	15

*Wenn der Stichprobenumfang n laut Tabelle gleich oder größer als der tatsächliche Losumfang N ist, dann 100%-Prüfung durchführen.

Beispiel: N = 490, AQL = 1,5% n = 50, c = 2, r = 3

Doppelter Stichprobenplan zur attributiven Kontrolle [1]

Losumfang N AQL		0,4%			0,65%			1,0%			1,5%		
von	bis	n_1 n_1+n_2	c_1 S_c	r_1 S_r	n_1 n_1+n_2	c_1 S_c	r_1 S_r	n_1 n_1+n_2	c_1 S_c	r_1 S_r	n_1 n_1+n_2	c_1 S_c	r_1 S_r
1	90	*80 160 Δ	0 1	2 2	*50 100 Δ	0 1	2 2	*32 64 Δ	0 1	2 2	*20 40 Δ	0 1	2 2
91	150	*80 160 Δ	0 1	2 2	*50 100 Δ	0 1	2 2	32 64 Δ	0 1	2 2	20 40	0 1	2 2
151	280	*80 160 Δ	0 1	2 2	50 100 Δ	0 1	2 2	32 64	0 1	2 2	20 40	0 1	2 2
281	500	80 160 Δ	0 1	2 2	50 100	0 1	2 2	32 64	0 1	2 2	32 64	0 3	3 4
501	1 200	80 160	0 1	2 2	50 100	0 1	2 2	50 100	0 3	3 4	50 100	1 4	4 5
1 201	3 200	80 160	0 1	2 2	80 160	0 3	3 4	80 160	1 4	4 5	80 160	2 6	5 7
3 201	10 000	125 250	0 3	3 4	125 250	1 4	4 5	125 250	2 6	5 7	125 250	3 8	7 9
10 001	35 000	200 400	1 4	4 5	200 400	2 6	5 7	200 400	3 8	7 9	200 400	5 12	9 13
35 001	150 000	315 630	2 6	5 7	315 630	3 8	7 9	315 630	5 12	9 13	315 630	7 18	11 19

Δ Wahlweise entsprechenden Einfachplan verwenden
* Wenn der Stichprobenumfang n laut Tabelle gleich oder größer als der tatsächliche Losumfang N ist, dann 100%-Prüfung durchführen.

Beispiel: N = 9989, AQL = 1,5%
$n_1 = 125$, $c_1 = 3$, $r_1 = 7$; $n_1 + n_2 = 250$, $S_c = 8$, $S_r = 9$

[1] Auszug aus ÖNORM A6650, Attributive Stichprobenprüfung (mit Genehmigung des Österreichischen Normungsinstituts).

Beispiele:

Einfacher Stichprobenplan: Zu $N = 810$ und $AQL = 1{,}5\%$ finden wir $n = 80$, $c = 3$ und $r = 4$.

Doppelter Stichprobenplan: Zu $N = 810$ und $AQL = 1{,}5\%$ finden wir $n_1 = 50$, $c_1 = 1$, $r_1 = 4$, $n_1 + n_2 = 100$, $S_c = 4$, $S_r = 5$.

Literatur:

ÖNORM 6650, DIN 40080

Willenze, Schaafsma, Moderne Qualitätskontrolle, Eindhoven: Philips Technische Bibliothek.

Qualitätssicherung von eingekauften Güter, Philips.

Sadowy, Industrielle Statistik, Vogel-Verlag.

Istvan Vincze, Mathematische Statistik mit industriellen Anwendungen.

Prof.Dr.J. Schärf
TGM
Wexstraße 17
1200 Wien