Statistische Qualitätsregelung zur Einführung in die Beurteilende Statistik

Manfred Borovcnik, Klagenfurt

Kurzfassung: Statistisches Testen gilt als schwierig. Einige der Verständnisprobleme entstehen aber eigentlich dadurch, daß man die Verfahren von ihrem ursprünglichen Zusammenhang abgelöst und verallgemeinert hat. Der historische Umweg zum Testen über Fragen der Qualitätsregelung kann den Unterricht bereichern und die angesprochenen Schwierigkeiten vermeiden. Die grundlegenden Fragestellungen der Qualitätsregelung werden im folgenden dargestellt. Ein Grundtyp von Problemen wird dann ausführlich im Kontext der Qualitätsregelung behandelt. Schließlich werden Bezüge zum Testen von Hypothesen geknüpft; daraus wiederum ergeben sich Anregungen für den Unterricht zum Testen. Ein Ausblick auf Vertiefungen in der Qualitätskontrolle rundet die Ausführungen ab.

1. Die zwei Grundproblemstellungen in der Qualitätsregelung

Zunächst wird in den Fachjargon der Qualitätsregelung eingeführt. Dabei werden die üblichen Bezeichnungen und Begriffe erklärt. Dann werden die zwei Grundtypen von Problemstellungen erläutert. Das eine Verfahren ist die Gut-Schlecht-Prüfung, welches entscheiden hilft, ob eine Warensendung einen zu hohen Schlecht-Anteil aufweist oder den Lieferbedingungen genügt. Es wird üblicherweise beim Wareneingang beim Konsumenten angewendet, wobei die Bedingungen zwischen Konsumenten und Lieferanten ausgehandelt werden. Das andere Verfahren ist die laufende Qualitätsregelung während der Produktion. Dieses Verfahren hilft zu entscheiden, ob die produzierten Einheiten den geforderten Abmessungen hinsichtlich Sollwert und Streuung genügen, oder ob zu große Abweichungen davon aufgetreten sind, sodaß die Produktion zu unterbrechen und nach den Ursachen dafür zu suchen ist.

1.a Grundbegriffe

Zuerst werden die Fachbegriffe erklärt. Die Qualität kann nach Ausschuß oder Nicht-Ausschuß, nach der Zahl der Fehler pro Einheit oder nach den Abmessungen hinsichtlich eines bestimmten Qualitätsmerkmals erfolgen. Danach unterscheiden sich auch die Verfahren zur Prüfung der Qualität. Der Begriff statistisches Risiko wird erklärt. Wie immer die Entscheidung ausfällt, ob die Warensendung angenommen oder abgelehnt wird, der

Produktionsprozeß unterbrochen oder fortgesetzt wird, man kann bei der Entscheidung Fehler machen. Wenn man unterstellt, daß die zu überprüfende Teilmenge von Einheiten mittels Zufalls ausgewählt wird, kann man diese Risiken auch berechnen.

Die Prüfung von Produkten soll sicherstellen, daß alle Einheiten die vorgegebenen Forderungen an ihre Qualität erfüllen. Gegen die Überprüfung aller Einheiten, man nennt das im Jargon Totalkontrolle, spricht: Bei Totalkontrolle entstehen wirtschaftlich nicht vertretbare Kosten, bei manchen Prüfungen wird die Einheit zerstört. Die Prüfung von Teilmengen aller produzierten Einheiten, also von Stichproben, ermöglicht eine Balance zwischen der erforderlichen Genauigkeit der Kontrolle und deren Kosten. Stichprobenkontrollen werden etwa beim Wareneingang, während des laufenden Produktionsprozesses oder bei der Endprüfung eingesetzt. Mittlerweile setzt man Stichproben auch schon bei Inventuren von Lagern ein.

Folgende Benennungen sind im Fachjargon üblich:
Los: Das ist die Grundgesamtheit, eine Zusammenfassung von Einheiten ("Waren"), deren Qualität überprüft werden soll. Beispiele: Tagesproduktion an einem Fließband, Rohölförderung an einem Tag, Waggonladung Kohle.

Losumfang: Anzahl der Einheiten im Los; in Zeichen N.

Prüfeinheit: Einheit, die der Kontrolle unterzogen wird.

Probe: Stichprobe, das ist die Auswahl von Prüfeinheiten aus einem Los. Wenn die Gesetze der Wahrscheinlichkeitsrechnung anwendbar sein sollen, so muß die Probe nach dem reinen Zufallsprinzip erfolgen. Die Auswahl ist damit vergleichbar der Ziehung der Lotto-Kugeln. Von dieser idealen statistischen Vorstellung der Auswahl weicht man in der Praxis mehr oder weniger ab; man muß jedoch auf der Hut sein, daß die Auswahl der Einheiten für die Probe nicht mit den Eigenschaften bezüglich deren Qualität zusammenhängt.

Probenumfang: Umfang der Stichprobe, das ist die Anzahl der geprüften Einheiten; in Zeichen n.

Merkmal: Variable, eines oder mehrere Merkmale, welche die Qualität der Einheiten erfassen sollen; bezeichnet mit großen Buchstaben wie etwa X.

Zählprüfung: Die Qualität der Einheiten wird durch Zählen erfaßt. Bei der Gut-Schlecht-Prüfung teilt man die Prüfeinheiten nach Ausschuß oder in Ordnung ein; man zählt die Zahl a schlechter Einheiten in der Probe und beurteilt die Qualität des Loses anhand des Anteils a/n an Ausschuß in der Probe. Bei der Erfassung der Zahl der Fehler pro Einheit beurteilt man die Qualität aufgrund der mittleren Zahl der Fehler pro Einheit. In der Stochastik entsprechen dieser Situation qualitative bzw. quantitativ diskrete Merkmale. Die Zahl bzw. der Anteil schlechter Stücke ist dann eine hypergeometrisch verteilte (bzw.

näherungsweise binomial verteilte) Zufallsvariable, die Zahl der Fehler an einem Stück wird unter bestimmten Voraussetzungen mit der Poisson-Verteilung beschrieben.

Meßprüfung: Die Qualität der Einheiten wird durch Messen verschiedener Merkmale erfaßt; das Ergebnis der Prüfung ist eine physikalische Größe, etwa der Durchmesser einer Welle, gemessen in [cm]. Man beurteilt die Qualität des Loses aufgrund der mittleren Größe des Merkmals und dessen Streuung in der Probe. In der Stochastik entsprechen dieser Situation quantitativ stetige Variable. Die Variable, welche die Qualität beschreibt, wird i.a. durch die Normalverteilung beschrieben; diese Einschränkung ist jedoch nicht schwerwiegend, denn die statistische Beurteilung erfolgt aufgrund der Mittelwerte, die nach dem Zentralen Grenzverteilungssatz sowieso annähernd normalverteilt sind.

Statistische Risiken: Die grundsätzliche Vorgangsweise in der Qualitätskontrolle soll anhand der Überprüfung von Waren beim Wareneingang schon hier angedeutet werden. Der Produzent liefert die Ware, der Konsument prüft deren Qualität etwa hinsichtlich des Anteils schlechter Einheiten. Aus dem Los von gelieferten Einheiten wird nach dem Zufallsprinzip die Probe ermittelt. Danach prüft man die Qualität der Einheiten in der Probe und erhält damit die Zahl schlechter Stücke mit a bzw. den Schlecht-Anteil mit a/n. Im gesamten Los entspricht dieser Zahl a/n der (unbekannte) Schlecht-Anteil A/N. Weil man die Auswahl der Probe nur mittels Zufalls bewerkstelligte, kann man mit Hilfe der Wahrscheinlichkeitsrechnung zahlenmäßig eingrenzen, wie groß der Unterschied zwischen dem Schlecht-Anteil im Los und in der Stichprobe sein kann. Wenigstens kann man für ein solches Schwankungsintervall eine (große) Wahrscheinlichkeit angeben. Ein Rest-Risiko verbleibt allerdings für den Lieferanten wie für den Abnehmer: Trotz guter Qualitätslage im Los, d.h. kleinem A/N, kann es vorkommen, daß der Schlecht-Anteil a/n im Los sehr groß ist, d.h. über das berechnete Schwankungsintervall hinausgeht. Dies führt dann zu einer unberechtigten Reklamation der Lieferung. Oder: Trotz schlechter Qualitätslage im Los, d.h. A/N ist sehr groß, kann es vorkommen, daß die Stichprobe das nicht anzeigt. Das führt zur Annahme des schlechten Loses und den damit verbundenen Nachteilen für den Abnehmer. Die Wahrscheinlichkeitsrechnung allerdings sagt aus, daß es sich dabei um Ausnahmefälle handelt, die nicht allzu häufig auftreten.

1.b Gut-Schlecht-Prüfung

Das Problem der Prüfung einer Warenlieferung wird dargestellt. Die mathematischen Voraussetzungen werden erläutert, wobei klar wird, daß die zufällige Auswahl der Stichprobe die Verteilung der schlechten Stücke in der Stichprobe berechnen läßt. Damit lassen sich dann die Eigenschaften von Entscheidungsverfahren über die gelieferte Ware studieren.

Die Warenlieferung besteht aus einem endlichen Los vom Umfang N; es interessiert lediglich, ob die Einheiten gut oder schlecht sind. Eine Einheit, die schlecht ist, wird als

Ausschuß bezeichnet. Man kennt den Schlecht-Anteil A/N nicht, die Frage ist, wie erkennt man, ob dieser Schlecht-Anteil innerhalb akzeptierbarer Grenzen liegt oder ob dieser Anteil zu hoch ist.

Dem Los wird ein Teil, eine Stichprobe vom Umfang n, entnommen und auf gut - schlecht geprüft. Das Ergebnis der i-ten Prüfeinheit ist dann eine Zufallsvariable

$$X_i = \begin{cases} 0 & \text{gut} \\ 1 & \text{schlecht} \end{cases}$$

Die Anzahl aller schlechten Stücke in der Stichprobe ist dann die Zufallsvariable X mit

$$X = X_1 + X_2 + \ldots + X_n$$

Die Verteilung der schlechten Stücke X ist, bei zufälliger Auswahl, eine hypergeometrische mit den Parametern N, A und n, wobei das Augenmerk auf die unbekannte Zahl A gerichtet ist. Das Entscheidungsverfahren, ob das Los als gut eingestuft und daher angenommen oder als schlecht eingestuft und daher abgelehnt wird, kann durch folgende Bedingungen an die Zufallsvariable beschrieben werden:

Ereignis	Entscheidung
$X \leq c$	Los angenommen
$X > c$	Los abgelehnt

Die Zahl c heißt dabei Annahmekennzahl. Hat man die Zahl c bereits, so ist zu untersuchen, ob das damit verbundene Verfahren gut ist bzw. wie man dessen Eigenschaften überhaupt beurteilen soll. Klärt man den Zusammenhang zwischen den Eigenschaften eines Verfahren mit dem Wert der Annahmekennzahl, so erhält man eine Methode, den Wert von c zu bestimmen. Weitere Details dazu werden in Abschnitt 2 behandelt.

1.c Qualitätsregelung

Das Problem der Überprüfung der Qualität der produzierten Einheiten bei laufender Produktion wird dargestellt. Auch hier werden die mathematischen Voraussetzungen knapp referiert. Basis der Überlegungen ist, daß die Abmessungen normalverteilt sind und die laufend produzierten Einheiten einer unabhängigen Wiederholung derselben Normalverteilung entsprechen. Damit kann man die Verteilung für den Mittelwert von Stichproben berechnen - abhängig vom Sollwert natürlich. Mit dieser Referenzverteilung wird dann der konkrete Stichprobenmittelwert verglichen, was dann zur Entscheidung über Weiterführung oder Unterbrechung der Produktion führt.

Bei laufender Produktion wird überprüft, ob ein bestimmtes, quantitatives Qualitätsmerkmal X einen vorgegebenen Sollwert einhält oder nicht. Dabei wird unterstellt, daß das Merkmal X einer Normalverteilung $N(\mu, \sigma^2)$ folgt; μ entspricht dem Sollwert, σ der unvermeidbaren Variabilität des Prozesses. Im Laufe der Produktion kann sich μ (oder auch σ) ändern. Die Frage ist, wie erkennt man dies?

Man entnimmt der laufenden Produktion zu bestimmten Zeitpunkten einige Prüfeinheiten und bestimmt den Mittelwert dieser. Über die unvermeidliche Streuung des Prozesses hat man schon Vorinformationen aus einem sogenannten Probelauf oder aus Erfahrung. Im übrigen kann der Parameter σ auch Ziel einer Qualitätsbeurteilung sein, das aber soll im folgenden nicht behandelt werden. Die mittlere Abmessung des Merkmals in der Probe ist daher eine Zufallsvariable, die auch einer Normalverteilung gehorcht:

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Es gilt $\bar{X} \sim N(\mu, \sigma^2/n)$. Man hat aufgrund der Daten der Probe zu entscheiden, ob die Produktion unterbrochen und die Ursachen für eine Verschiebung des Sollwerts gesucht werden sollen, oder ob die Produktion weiterlaufen soll. Eine Entscheidungsregel über Eingreifen oder Nicht-Eingreifen kann man durch ein Intervall (UKG, OKG) beschreiben:

$\bar{X} < \text{UKG}$	**Produktion stoppen**
$\text{UKG} \leq \bar{X} \leq \text{OKG}$	**Produktion läuft weiter**
$\bar{X} > \text{OKG}$	**Produktion stoppen**

UKG steht dabei für untere Kontrollgrenze, OKG für obere Kontrollgrenze. Daneben gibt es noch sogenannte Warngrenzen WG, deren Überschreitung die Entnahme zusätzlicher Einheiten zur Prüfung zur Folge hat. Wieder stellt sich die Frage, welche Eigenschaften das Verfahren hat und wie man die Kontrollgrenzen bestimmt. Auf die entsprechenden Verfahren kann hier aus Platzgründen nicht eingegangen werden, es sei auf Rinne und Mittag (1989) oder auf Borovcnik (1987) verwiesen.

2. Gut-Schlecht-Prüfung

Die Prüfung einer Warenlieferung auf deren Qualität wird in diesem Abschnitt ausführlich behandelt. Meist wird diese Gut-Schlecht-Prüfung bei der Abnahmekontrolle von Waren bei Lieferung vom Produzenten an den Konsumenten durchgeführt. Es geht um die Beurteilung des Ausschußanteils der Lieferung. Stichprobenpläne werden zunächst als reine Handlungsanweisungen eingeführt, die angeben, wie viele Einheiten zu kontrollieren und wie viele fehlerhafte Einheiten maximal zu tolerieren sind, ehe man die Warensendung ablehnt. Sodann werden die Eigenschaften dieser Stichprobenpläne studiert. Als geeignetes Hilfsmittel wird dazu die sogenannte OC-Kurve eingeführt, die einen globalen Überblick über die Auswirkungen der Prüfung nach dem jeweiligen Plan angibt. Verfeinerte Interpretationen ergeben sich durch spezielle Parameter der OC-Kurven. Ferner werden Probenpläne mit Hilfe von Tabellen berechnet. Der Gebrauch solcher Tabellenwerke wird gezeigt. Bei der tiefergründigen Interpretation der eingeführten Parameter ergeben sich zunächst Berührpunkte mit der beurteilenden Statistik. Es zeigt sich jedoch bald, daß die gängigen Interpretationen nicht zielführend sind, es müssen daher weitere, für die Qualitätskontrolle wichtige Größen eingeführt werden. Den Abschluß dieses Abschnitts bilden Überlegungen zu Querverbindungen zwischen statistischer Qualitätskontrolle und beurteilender Statistik, wobei auch einige didaktische Anregungen für den Unterricht anfallen.

2.a Stichprobenpläne und ihre Eigenschaften

Stichprobenpläne sind Entscheidungsvorschriften, wie viele Einheiten zu prüfen sind und welche Zahl von fehlerhaften Einheiten maximal zu akzeptieren ist. Ferner wird die OC-Kurve dargestellt, welche die Wahrscheinlichkeit für die Annahme der Warensendung angibt, und zwar in Abhängigkeit vom tatsächlichen Ausschußprozentsatz. Es wird gezeigt, wie man diese OC-Kurven berechnet und welche Eigenschaften sie haben sollten, damit der Stichprobenplan gut wird. Weiters werden wichtige Parameter von solchen OC-Kurven eingeführt und erläutert. Diese Parameter erleichtern die Beurteilung von Prüfplänen wesentlich und führen zum Teil in die Nähe der statistischen Testtheorie.

Stichprobenplan als Handlungsanweisung: In Abschnitt 1.b wurde die Situation der Gut-Schlecht-Prüfung eingehend beschrieben. Ein Stichprobenplan ist eine Handlungsanweisung, wie man vorzugehen hat bei der Beurteilung eines Loses hinsichtlich dessen Ausschußanteils. Ein solcher Plan wird durch die drei Kennziffern N, n, c angegeben; N ist dabei die Losgröße, n die Zahl zu prüfender Einheiten und c ist die Annahmezahl. Die Entscheidung über das Los lautet so: Überschreitet die Zahl schlechter Einheiten X in der Probe die Annahmezahl c, so ist das Los als schlecht abzuweisen, ansonsten ist es anzunehmen. Zur Erinnerung, die Zufallsvariable X ist hypergeometrisch verteilt mit den Parametern N, A und n, in Zeichen

$$X \sim H(N, A, n)$$

Voraussetzung ist, daß die Stichprobe zufällig gezogen wird. Den Ausschußanteil $p=A/N$ geben Praktiker gerne als Prozentsatz an:

$$P = 100 \cdot \frac{A}{N} \%$$

Gerade dieser Ausschußprozentsatz P ist unbekannt. Einen solchen Plan kann man zunächst als reine Handlungsanweisung auffassen. Die Frage ist nur, wie wirkt sich ein solcher Probenplan aus? Dazu ein Beispiel.

Beispiel: Gegeben ist der Probenplan N=100, n=10, c=0. Mit L(P) werde die Wahrscheinlichkeit bezeichnet, das Los anzunehmen, falls der tatsächliche Ausschußprozentsatz im Los P beträgt. Die Funktion L(P) gibt einen ersten Eindruck von den Eigenschaften des Probenplans. Für L(P) muß man die Wahrscheinlichkeit des Ereignisses X≤0 mit den geeigneten Parameterwerten berechnen, und zwar mittels der hypergeometrischen Verteilung. Diese hypergeometrische Verteilung ist aber sehr unhandlich, man kann sie jedoch bei den üblichen Werten in der Qualitätskontrolle mit der Binomialverteilung, meist sogar mit der Poisson-Verteilung annähern. Das vereinfacht die Berechnung der Annahmewahrscheinlichkeiten sehr. Für die angesprochenen Approximationen gelten die folgenden Faustregeln:

$$\mathrm{H}(N, A, n) \quad \approx \quad \mathrm{B}(n, \tfrac{A}{N}) \quad \approx \quad \mathrm{P}(\lambda = n \cdot \tfrac{A}{N})$$

$$\text{falls } \tfrac{n}{N} \le 0{,}1 \qquad\qquad \text{falls } \tfrac{A}{N} \le 0{,}1$$

Bemerkung: Diese Verteilungen haben alle denselben Erwartungswert $n \cdot \frac{A}{N}$.

Für die Berechnung des Parameters λ der approximierenden Poisson-Verteilung muß man den Ausschußprozentsatz so umrechnen:

$$\lambda = n \cdot \frac{A}{N} = n \cdot \frac{P}{100}$$

Für P=1% ergibt sich λ=0,1; damit erhält man

$$\mathrm{L}(1\%) = \mathrm{W}(\mathrm{X} \le 0) = \frac{\lambda^0}{0!} \cdot e^{-\lambda} = \frac{0{,}1^0}{0!} \cdot e^{-0{,}1} = 0{,}90$$

Führt man die Berechnungen für einige Werte von P durch, so erhält man folgendes Bild der Annahmewahrscheinlichkeiten:

P%	1	2	3	5	10	20
λ	0,1	0,2	0,3	0,5	1	2
L(P)	0,90	0,82	0,73	0,60	0,37	0,13

Unterstellt man dem Produzenten einen sehr kleinen Ausschußprozentsatz, etwa 1%, so wird das Los mit einer Wahrscheinlichkeit von 0,90 angenommen, bei 5% ergibt sich eine Annahmewahrscheinlichkeit von 0,73; ist hingegen der Ausschuß sehr hoch, etwa 10%, so wird immer noch mit einer Wahrscheinlichkeit von 0,37 angenommen, ja selbst bei 20% Ausschuß im Los schlüpfen noch 13% der Lose durch die Kontrolle und werden angenommen. Hier zeigt sich auch schon, wie man die Werte zu deuten hat: Man kann nicht sagen, wie viele Lose zu Unrecht durchschlüpfen; man kann lediglich sagen, wenn dieser oder jener Ausschußanteil im Los zutrifft, dann ist die Wahrscheinlichkeit der Annahme so und so groß.

OC-Kurve eines Probenplans: Unter der sogenannten Operationscharakteristik oder kurz OC-Kurve eines Probenplans versteht man die Funktion, die jedem (fiktiven) Ausschußprozentsatz P die Annahmewahrscheinlichkeit $L(P)$ zuordnet. Die Kurve wird auch als Annahmekennlinie bezeichnet. Die OC-Kurve gibt also einen globalen Eindruck der Folgen des Probenplans. Der im vorhergehenden Beispiel behandelte Probenplan führt zu folgender OC-Kurve:

Man kann den Plan als streng aber wenig trennscharf bezeichnen. Streng ist er, weil die gute Qualitätslage von 1% nur mit 90% Wahrscheinlichkeit zur Annahme des Loses führt; wenig trennscharf ist er, weil eine vergleichsweise schlechte Qualitätslage von 10% noch immer mit Wahrscheinlichkeit von 0,37 zur Annahme führt.

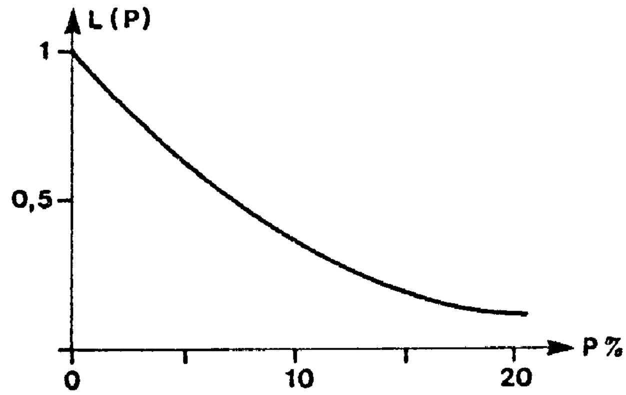


Abb. 1: OC-Kurve eines strengen, aber nicht trennscharfen Stichprobenplans

Ideal wäre eine OC-Kurve, die ganz sicher gute und schlechte Lose trennt; unterstellt man $P=1\%$ als Grenzpunkt guter Qualität, so sollten demnach Lose mit kleinerem P sicher angenommen, Lose mit größerem P sicher abgelehnt werden. Die ideale OC-Kurve sieht daher wie in nebenstehender Abbildung aus.

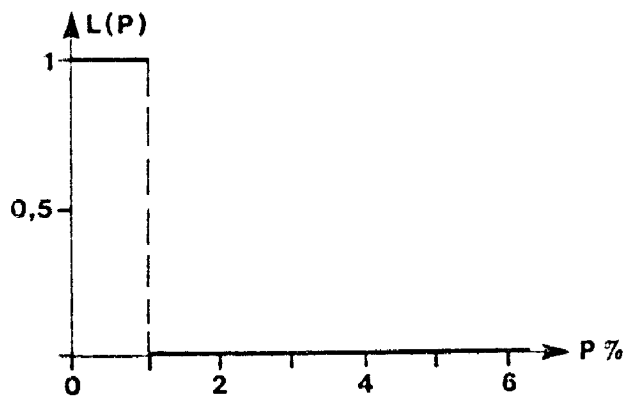


Abb. 2: OC-Kurve eines idealen Probenplans

Realistischerweise könnte man fordern, daß die Grenze zwischen guten und schlechten Qualitätslagen durch zwei Werte P_1 und P_2 bzw. das Intervall dazwischen beschrieben wird. Lose mit P_1 sollten noch mit sehr hoher Wahrscheinlichkeit angenommen, Lose mit P_2 dagegen sollten mit sehr kleiner Wahrscheinlichkeit noch angenommen werden; dazwischen sollte die OC-Kurve möglichst steil verlaufen. Das sieht dann etwa wie in nebenstehender Abbildung aus.

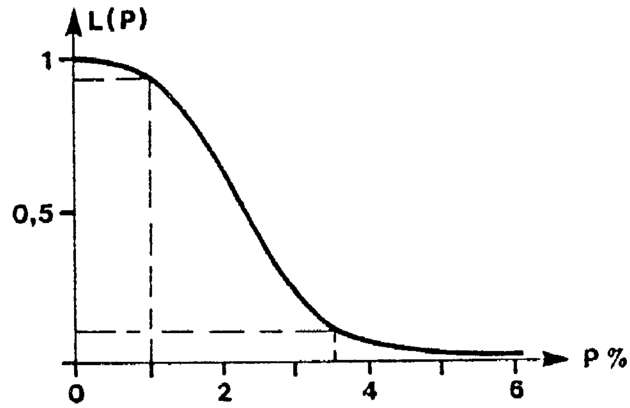


Abb. 3: OC-Kurve eines realistischen Probenplans

Wichtige Parameter eines Probenplans: Wie in der Wahrscheinlichkeitstheorie die Gestalt einer Verteilung über wenige Parameter wie den Erwartungswert und die Standardabweichung erfaßt und beurteilt wird, so gibt es in der Qualitätskontrolle einige wichtige lokale Anhaltspunkte, um die globale Gestalt der OC-Kurve zu beurteilen. Solche sind:

P_1 Gutgrenze AQL (acceptable quality level) - die Abkürzungen stammen allesamt aus der englischen Bezeichnung. Die Gutgrenze ist nicht eine Grenzlage von Qualität sondern bezeichnet die normale Fertigungsgüte, wie man sie etwa aus vergangenen Lieferungen kennt.

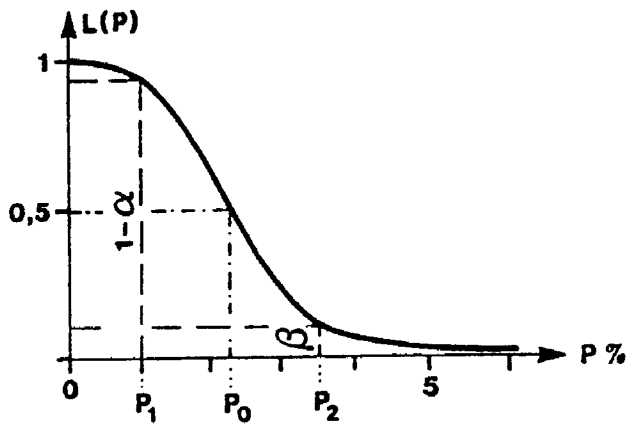


Abb. 4: Wichtige Punkte eines Probenplans: P_1 Gutgrenze, P_2 Schlechtgrenze, P_0 Indifferenzpunkt

α Produzentenrisiko, das ist die Wahrscheinlichkeit, das Los abzulehnen, obwohl der Produzent einen Ausschußanteil von P_1 oder besser einhält: $L(P_1) = 1 - \alpha$

P_2 Schlechtgrenze LQ (limiting quality), auch LTPD (lot tolerance percent defective) genannt. Sendungen, die diese Grenzlage überschreiten, sollen nur mit einer kleinen Wahrscheinlichkeit angenommen werden.

β Konsumentenrisiko, das ist die Wahrscheinlichkeit, das Los anzunehmen, obwohl der Produzent einen Schlechtanteil an der Schlechtgrenze oder gar darüber liefert: $L(P_2) = \beta$.

P_0 Indifferenzpunkt IQL (indifference quality level), das ist jene Qualitätslage, bei der Annahme und Ablehnung einem Münzwurf mit gleicher Wahrscheinlichkeit für beide Ausgänge gleich kommen. Man ist indifferent bezüglich der Entscheidung.

2.b Berechnung von Probenplänen und Interpretation der Eigenschaften

Im folgenden werden Stichprobenpläne, deren Parameter sowie deren OC-Kurven berechnet. Dazu werden Auszüge aus umfangreichen Tabellenwerken benützt. Der Umgang mit diesen Tabellen wird anhand von Beispielen erläutert. Aus Tabellen und Nomogrammen kann man also geeignete Stichprobenpläne mit vorgegebenen Eigenschaften finden. Im weiteren werden die Parameter von OC-Kurven und deren Interpretation noch genauer unter die Lupe genommen. Die naheliegende Deutung der auftretenden Wahrscheinlichkeiten für Fehlentscheidungen als Häufigkeiten auf lange Sicht erweisen sich dabei als nicht zielführend. Die Fehlerwahrscheinlichkeiten werden eher zu reinen Ordnungsgesichtspunkten von OC-Kurven von Stichprobenplänen, während für den Qualitätsingenieur andere Begriffe wie der des maximalen Durchschlupfs, also des maximalen Anteils von fehlerhaften Einheiten viel zentraler werden. Diese Durchschlupfgröße läßt tatsächlich eine Deutung auf längere Sicht zu, sie gibt nämlich den maximalen Fehleranteil bei häufiger Anwendung des Stichprobenplans an. Den Abschluß bilden einige Überlegungen zur Prüfdichte, das ist die Zahl der geprüften Einheiten bezogen auf die Zahl aller Einheiten in der Warensendung. Danach sind größere, homogene Warensendungen hinsichtlich der Überprüfung ihrer Qualität wirtschaftlicher.

Benutzung von Tabellen: Stichprobenpläne und deren OC-Kurven einschließlich der Parameter muß man nicht selbst berechnen. Man auf umfangreiche Tabellenwerke zurückgreifen, darunter der Military Standard ABC-STD-105, der im International Standard ISO 40080 übernommen wurde (siehe DGQ, 1981a). Diese Tabellen sind nach unterschiedlichen Gesichtspunkten geordnet

- AQL im ABC-STD-105
- LTPD (LQ) in Doge und Romig (1959)
- IQL in Philips (siehe Schaafsma und Willemze, 1961)

Der ABC-Standard 105 unterscheidet mehrere Prüfniveaus, die indirekt mit den Risiken α und β zusammenhängen. Der folgende Auszug daraus bezieht sich auf das Prüfniveau II, welches das gebräuchlichste ist. Zur verschärften Prüfung kommt es nach mehreren schlechten Lieferungen, zur reduzierten Prüfung kommt es in Ausnahmefällen bei besonders guter Qualität.

Auszug aus ABC-STD-105: Probenpläne für normale Prüfung - Prüfniveau II

N \ Stichprobenplan	AOL 0,040	AOL 0,065	AOL 0,10	AOL 0,15	AOL 0,25	AOL 0,40	AOL 0,65	AOL 1,0	AOL 1,5	AOL 2,5	AOL 4,0	AOL 6,5	AOL 10	AOL 15	AOL 25
2 ... 8	N	N	N	N	N	N	N	N	N	N bzw. 5-0	N bzw. 3-0	N bzw. 2-0	N bzw. 5-1	N bzw. 3-1	N bzw. 2-1
9 ... 15	N	N	N	N	N	N	N	N bzw. 13-0	8-0	5-0	3-0	2-0	5-1	3-1	3-2
16 ... 25	N	N	N	N	N	N	N bzw. 20-0	13-0	8-0	5-0	3-0	8-1	5-1	5-2	5-3
26 ... 50	N	N	N	N	N	N bzw. 32-0	20-0	13-0	8-0	5-0	13-1	8-1	8-2	8-3	8-5
51 ... 90	N	N	N	N bzw. 80-0	50-0	32-0	20-0	13-0	8-0	20-1	13-1	13-2	13-3	13-5	13-7
91 ... 150	N	N	N bzw. 125-0	80-0	50-0	32-0	20-0	13-0	32-1	20-1	20-2	20-3	20-5	20-7	20-10
151 ... 280	N	N bzw. 200-0	125-0	80-0	50-0	32-0	20-0	50-1	32-1	32-2	32-3	32-5	32-7	32-10	32-14
281 ... 500	N bzw. 315-0	200-0	125-0	80-0	50-0	32-0	80-1	50-1	50-2	50-3	50-5	50-7	50-10	50-14	50-21
501 ... 1 200	315-0	200-0	125-0	80-0	50-0	125-1	80-1	80-2	80-3	80-5	80-7	80-10	80-14	80-21	50-21
1 201 ... 3 200	315-0	200-0	125-0	80-0	200-1	125-1	125-2	125-3	125-5	125-7	125-10	125-14	125-21	80-21	50-21
3 201 ... 10 000	315-0	200-0	125-0	315-1	200-1	200-2	200-3	200-5	200-7	200-10	200-14	200-21	125-21	80-21	50-21
10 001 ... 35 000	315-0	200-0	500-1	315-1	315-2	315-3	315-5	315-7	315-10	315-14	315-21	200-21	125-21	80-21	50-21
35 001 ... 150 000	315-0	800-1	500-1	500-2	500-3	500-5	500-7	500-10	500-14	500-21	315-21	200-21	125-21	80-21	50-21
150 001 ... 500 000	1250-1	800-1	800-2	800-3	800-5	800-7	800-10	800-14	800-21	500-21	315-21	200-21	125-21	80-21	50-21
> 500 000	1250-1	1250-2	1250-3	1250-5	1250-7	1250-10	1250-14	1250-21	800-21	500-21	315-21	200-21	125-21	80-21	50-21

Auszug aus DIN 40080

Beispiele zur Benutzung der ABC-Tabelle:

- Das Los habe N=85 Einheiten, die Gutgrenze (AQL) sei mit P_1=0,65% festgelegt. Die Zeilen der ABC-Tabelle geben den Losumfang in Klassen an, man hat nach der Zeile 51-90 zu suchen; die Spalten geben die AQL in % an, zu suchen ist die Spalte mit 0,65, es ist die siebente. Im Kreuzpunkt von Zeile und Spalte liest man den Probenplan 20-0 ab. Das bedeutet, daß n=20 Einheiten zu prüfen sind und man das Los nur annimmt, wenn kein einziges Ausschußstück darunter zu finden ist, da c=0 ist.
- Für den Losumfang N=3500 und P_1=0,65% liest man den Probenplan 200-3 ab. Man hat daher n=200 Einheiten zu prüfen, c=3 ist die Annahmezahl.

Erst aus umfangreicheren Tabellen erhält man dann auch Auskunft über die Risiken α und β, letzteres natürlich nur für vorweg festgelegte Schlechtgrenzen P_2. Die folgende Tabelle aus Cameron (1952) ist nach dem Quotienten P_2/P_1 sowie nach α und β geordnet.

Beispiele zur Benutzung der Cameron-Tabelle:

- Für P_1=1% und P_2=4% sowie die Risiken α=0,05 und β=0,10 ist der Stichprobenplan (n,c) gesucht. Dabei ist vorausgesetzt, daß der Losumfang N genügend groß ist, damit die zur Erstellung der Tabelle verwendete Poisson-Approximation gewährleistet ist. Man beachte, daß in der Tabelle nicht die Prozentsätze P_i sondern die Anteile p_i = P_i/100

Tabelle zur Bestimmung von Probenplänen mit bestimmtem Konsumenten- und Produzentenrisiko, geordnet nach dem Quotienten p_1/p_2 (aus Cameron, 1952):

c	Values of p_2/p_1 for: $\alpha = .05$ $\beta = .10$	$\alpha = .05$ $\beta = .05$	$\alpha = .05$ $\beta = .01$	np_1	c	Values of p_2/p_1 for: $\alpha = .01$ $\beta = .10$	$\alpha = .01$ $\beta = .05$	$\alpha = .01$ $\beta = .01$	np_1
0	44.890	58.404	89.781	.052	0	229.105	298.073	458.210	.010
1	10.946	13.349	18.681	.355	1	26.184	31.933	44.686	.149
2	6.509	7.699	10.280	.818	2	12.206	14.439	19.278	.436
3	4.890	5.675	7.352	1.366	3	8.115	9.418	12.202	.823
4	4.057	4.646	5.890	1.970	4	6.249	7.156	9.072	1.279
5	3.549	4.023	5.017	2.613	5	5.195	5.889	7.343	1.785
6	3.206	3.604	4.435	3.286	6	4.520	5.082	6.253	2.330
7	2.957	3.303	4.019	3.981	7	4.050	4.524	5.506	2.906
8	2.768	3.074	3.707	4.695	8	3.705	4.115	4.962	3.507
9	2.618	2.895	3.462	5.426	9	3.440	3.803	4.548	4.130
10	2.497	2.750	3.265	6.169	10	3.229	3.555	4.222	4.771
11	2.397	2.630	3.104	6.924	11	3.058	3.354	3.959	5.428
12	2.312	2.528	2.968	7.690	12	2.915	3.188	3.742	6.099
13	2.240	2.442	2.852	8.464	13	2.795	3.047	3.559	6.782
14	2.177	2.367	2.752	9.246	14	2.692	2.927	3.403	7.477
15	2.122	2.302	2.665	10.035	15	2.603	2.823	3.269	8.181
16	2.073	2.244	2.588	10.831	16	2.524	2.732	3.151	8.895
17	2.029	2.192	2.520	11.633	17	2.455	2.652	3.048	9.616
18	1.990	2.145	2.458	12.442	18	2.393	2.580	2.956	10.346
19	1.954	2.103	2.403	13.254	19	2.337	2.516	2.874	11.082
20	1.922	2.065	2.352	14.072	20	2.287	2.458	2.799	11.825
21	1.892	2.030	2.307	14.894	21	2.241	2.405	2.733	12.574
22	1.865	1.999	2.265	15.719	22	2.200	2.357	2.671	13.329
23	1.840	1.969	2.226	16.548	23	2.162	2.313	2.615	14.088
24	1.817	1.942	2.191	17.382	24	2.126	2.272	2.564	14.853
25	1.795	1.917	2.158	18.218	25	2.094	2.235	2.516	15.623
26	1.775	1.893	2.127	19.058	26	2.064	2.200	2.472	16.397
27	1.757	1.871	2.098	19.900	27	2.035	2.168	2.431	17.175
28	1.739	1.850	2.071	20.746	28	2.009	2.138	2.393	17.957
29	1.723	1.831	2.046	21.594	29	1.985	2.110	2.358	18.742
30	1.707	1.813	2.023	22.444	30	1.962	2.083	2.324	19.532
31	1.692	1.796	2.001	23.298	31	1.940	2.059	2.293	20.324
32	1.679	1.780	1.980	24.152	32	1.920	2.035	2.264	21.120
33	1.665	1.764	1.960	25.010	33	1.900	2.013	2.236	21.919
34	1.653	1.750	1.941	25.870	34	1.882	1.992	2.210	22.721
35	1.641	1.736	1.923	26.731	35	1.865	1.973	2.185	23.525
36	1.630	1.723	1.906	27.594	36	1.848	1.954	2.162	24.333
37	1.619	1.710	1.890	28.460	37	1.833	1.936	2.139	25.143
38	1.609	1.698	1.875	29.327	38	1.818	1.920	2.118	25.955
39	1.599	1.687	1.860	30.186	39	1.804	1.903	2.098	26.770
40	1.590	1.676	1.846	31.066	40	1.790	1.887	2.079	27.587
41	1.581	1.666	1.833	31.938	41	1.777	1.873	2.060	28.406
42	1.572	1.656	1.820	32.812	42	1.765	1.859	2.043	29.228
43	1.564	1.646	1.807	33.686	43	1.753	1.845	2.026	30.051
44	1.556	1.637	1.796	34.563	44	1.742	1.832	2.010	30.877
45	1.548	1.628	1.784	35.441	45	1.731	1.820	1.994	31.704
46	1.541	1.619	1.773	36.320	46	1.720	1.808	1.980	32.534
47	1.534	1.611	1.763	37.200	47	1.710	1.796	1.965	33.365
48	1.527	1.603	1.752	38.082	48	1.701	1.785	1.952	34.198
49	1.521	1.596	1.743	38.965	49	1.691	1.775	1.938	35.032

angegeben sind. Die erste Spalte entspricht den Vorgaben für α und β, der Wert 4,057 kommt dem Quotienten p_1/p_2 = 4 am nächsten. In der entsprechenden Zeile liest man in der Vorspalte c=4 und in der letzten Spalte den Wert np_1 =1,970 ab. Für den Probenumfang errechnet man daraus n = 1,970/0,01 = 197. Der Stichprobenplan lautet also n=197 und c=4.

- Man bestimme den Probenplan für P_1 = 1% und P_2 = 3% sowie die Risiken α = β = 0,05. Nun hat man in der zweiten Spalte diejenige Zahl zu suchen, die dem Quotienten 3 am nächsten kommt, das ist 3,074. In derselben Zeile liest man c=8 und np_1 = 4,095 ab. Der Probenplan lautet daher 410 - 8.

Interpretation der Gutgrenze P_1 und mittlerer Schlechtanteil: Was bedeutet eigentlich die Gutgrenze AQL? Stellt sie eine Art zugelassene Qualitätslage dar, einen sozusagen tolerierten Ausschußanteil? Was passiert, wenn der Produzent diesen Hinweis P_1 für seine Qualitätslenkung als Mittelwert versteht? Wenn er also seine Produktion so abstimmt, daß im Mittel der Ausschußanteil P_1 eingehalten wird? Als Folge würde sich ergeben, daß viele der Lose einen etwas höheren Anteil an schlechten Einheiten aufweisen. Wenn aber die OC-Kurve sehr steil von P_1 nach P_2 fällt, dann würden diese häufigen und von daher normalen Lose eine viel zu niedrige Annahmewahrscheinlichkeit aufweisen. Es würde daher viel zu oft passieren, daß Lose abgelehnt werden, die Folgekosten und der Vertrauensverlust wären unzumutbar hoch.

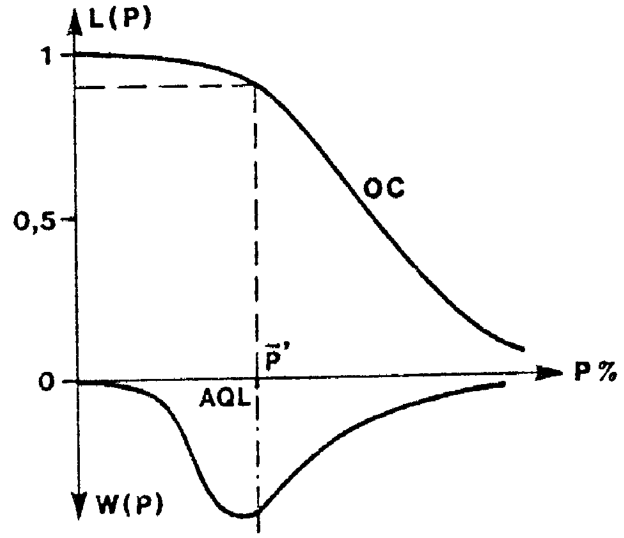


Abb. 5: OC-Kurve und Verteilung des Schlechtanteils - Mittlerer Schlecht-Anteil ist gleich AQL

Der Produzent muß daher den mittleren Ausschußanteil weit unter dem Niveau P_1 halten, die Gutgrenze P_1 ist eigentlich nur eine Rechengröße, die einen sehr indirekten Hinweis auf die Qualitätslenkung gibt, jedenfalls gibt sie keinen Aufschluß über den mittleren Ausschußanteil über längere Zeit bzw. über die normale Qualitätslage. Man vergleiche dazu die nebenstehende Abbildung, die eine brauchbare Verteilung der Fehler in der Produktion wiedergibt.

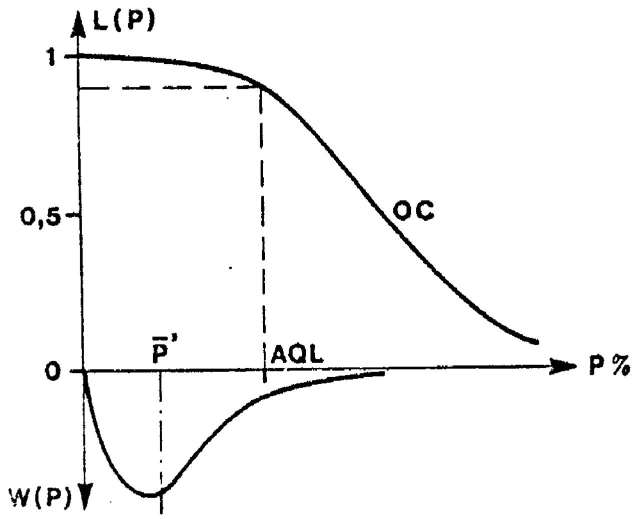


Abb. 6: OC-Kurve und Verteilung des Schlechtanteils - Mittlerer Schlecht-Anteil weit unter AQL

Die Interpretation der Schlechtgrenze P_2 und maximaler Durchschlupf: Auch die Schlechtgrenze (limiting quality) ist nur eine Rechengröße, wovon man sich leicht durch folgende Überlegung überzeugen kann. Auf dem Niveau P_2 beträgt die Annahmewahrscheinlichkeit lediglich β, etwa $\beta=0{,}1$. Schon Annahmewahrscheinlichkeiten, die unter 0,9 fallen, sind für den Hersteller in der Regel unzumutbar, weshalb er sie zu vermeiden sucht. Schon gar nicht wird der Hersteller riskieren, einen Ausschußanteil P in der Nähe von der Schlechtgrenze P_2 zu produzieren. Die Punkte (P_1, $1-\alpha$) und (P_2, β) auf der OC-Kurve sind also nur Anhaltspunkte, Rechengrößen sozusagen, damit man geeignete Stichprobenpläne auswählen kann. P_1 gibt dabei noch einen indirekten Hinweis für den Produzenten zur Steuerung der Qualität. Die Interpretation von P_2 ist da viel schwieriger und kann dem Konsumenten eigentlich kein Bild von der tatsächlich gelieferten Qualität verschaffen. Ihn interessiert nicht dieser völlig unrealistische Punkt sondern vielmehr, wie sich der Ausschußanteil entwickelt, falls er den Probenplan längerfristig anwendet.

Nehmen wir an, es besteht zusätzlich die Vereinbarung, daß ein abgelehntes Los auf Kosten des Lieferanten total kontrolliert wird und daß dabei auftretende Ausschußeinheiten durch gute ersetzt werden. Der Durchschlupf D an schlechten Einheiten nach Prüfung hängt dann vom Ausschußprozentsatz so ab:

$$D = D(P) = P \cdot L(P) + 0 \cdot [1 - L(P)]$$

Bei P% Ausschuß wird nämlich mit Wahrscheinlichkeit $L(P)$ angenommen, mit Wahrscheinlichkeit $1 - L(P)$ aber wird total kontrolliert, sodaß dann keine schlechte Einheit mehr in der Lieferung ist. $D(P)$ ist ein mittlerer Durchschlupf AOQ (average outgoing quality), die wiederum vom unbekannten Ausschußprozentsatz abhängt. Der maximale Wert des mittleren Durchschlupfs in Abhängigkeit von P AOQL (average outgoing quality limit) sagt jedoch viel über das Verhalten eines Probenplans aus:

$$D_{max} := \max_P D(P)$$

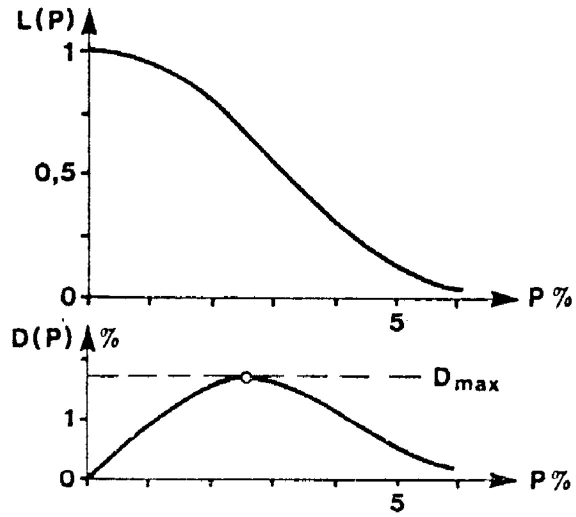


Abb. 7: OC-Kurve, darunter mittlerer und maximaler Durchschlupf

Zwar ist D_{max} kein Garant, daß nicht auch Lose mit noch höherem Anteil an Ausschuß angenommen werden, denn es ist ja nur das Maximum von *mittleren* Anteilen. Wendet man aber den Probenplan häufig an, so ist es wirklich im Durchschnitt ein *maximaler* Ausschußprozentsatz. Der Ausschußanteil P_{Dmax}, der den maximalen Durchschlupf verursacht, liegt in der Regel in der Zone mit zu geringen Annahmewahrscheinlichkeiten (mit $L(P) \leq 0{,}9$). Solche Lose werden i.a. nicht

Tabelle von Kennziffern zur Beurteilung von Operationscharakteristiken nach DGQ (1981a):

n - c	p'_{90} (L = 90 %)	p'_{50} (L = 50 %)	p'_{10} (L = 10 %)	D_{max} = AOQL	p'_{AOQL}
2- 0	5,27	34,7	115	18	50,0
2- 1	26,60	83,9	195	42	81,0
3- 0	3,50	23,1	76,8	12	33,3
3- 1	17,70	55,9	130	28	54,0
3- 2	36,70	89,1	177	46	75,7
5- 0	2,10	13,9	46,1	7,4	20,0
5- 1	10,6	33,6	77,8	17	32,4
5- 2	22,0	53,5	106	27	45,4
5- 3	34,9	73,4	134	39	59,0
8- 0	1,31	8,66	28,8	4,6	12,5
8- 1	6,65	21,0	48,6	11	20,3
8- 2	13,8	33,4	66,5	17	28,4
8- 3	21,8	45,9	83,5	24	36,9
8- 5	39,4	70,9	116	40	54,4
13- 0	0,808	5,33	17,7	2,8	7,7
13- 1	4,09	12,9	29,9	6,5	12,5
13- 2	8,48	20,6	40,9	11	17,5
13- 3	13,4	28,2	51,4	15	22,7
13- 5	24,2	43,6	71,3	24	33,5
13- 7	35,8	59,0	90,5	34	44,6
20- 0	0,527	3,47	11,5	1,8	5,0
20- 1	2,66	8,39	19,5	4,2	8,1
20- 2	5,51	13,4	26,6	6,9	11,4
20- 3	8,73	18,4	33,4	9,7	14,8
20- 5	15,8	28,4	46,4	16	21,8
20- 7	23,3	38,3	58,9	22	29,0
20- 8	27,2	43,3	65,0	26	32,8
20-10	35,1	53,3	77,0	33	40,3
32- 0	0,328	2,16	7,19	1,2	3,1
32- 1	1,66	5,24	12,2	2,6	5,1
32- 2	3,44	8,35	16,6	4,3	7,1
32- 3	5,45	11,5	20,9	6,1	9,2
32- 5	9,85	17,7	29,0	9,9	13,6

zur Prüfung vorgelegt. D_{max} ist daher eine Art Garantie, eine Obergrenze für den Ausschußanteil, das mag insbesondere bei neuen Lieferpartien wichtig sein. Wenn man über den Durchschlupf auf lange Sicht Aussagen macht, müßte man die Verteilung des Ausschußanteils in der Produktion kennen oder wenigstens über die mittlere Fertigungsgüte $\overline{P}$ Bescheid wissen, was aber eher die Ausnahme darstellt. Aus der folgenden DGQ-Tabelle erhält man die wichtigsten Daten zu verschiedenen Stichprobenplänen (*n*,*c*).

Beispiel zum Lesen der DGQ-Tabelle: Für den Stichprobenplan n=8 und c=0 erhält man aus der obigen Tabelle folgende Punkte auf der OC-Kurve:

Fortsetzung der Tabelle von Kennziffern zur Beurteilung von Operationscharakteristiken nach DGQ (1981a):

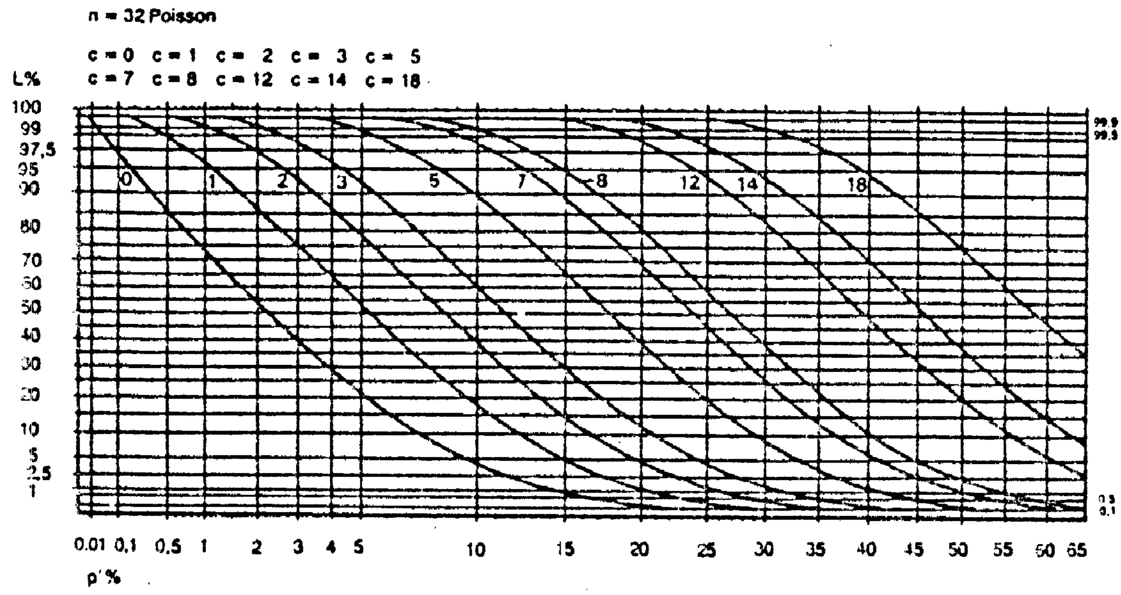
n - c	p'_{90} (L = 90%)	p'_{50} (L = 50%)	p'_{10} (L = 10%)	D_{max} = AOQL	p'_{AOQL}
32- 7	14,6	24,0	36,8	14	18,1
32- 8	17,0	27,1	40,6	16	20,5
32-10	21,9	33,3	48,1	21	25,2
32-12	27,0	39,6	55,6	25	30,0
32-14	32,2	45,8	62,9	29	34,8
32-18	42,7	58,3	77,4	39	44,7
50- 0	0,21	1,39	4,61	0,74	2,0
50- 1	1,06	3,36	7,78	1,7	3,2
50- 2	2,20	5,35	10,6	2,7	4,5
50- 3	3,49	7,34	13,4	3,9	5,9
50- 5	6,30	11,3	18,6	6,3	8,7
50- 7	9,31	15,3	23,5	9,0	11,6
50- 8	10,9	17,3	26,0	10	13,1
50-10	14,0	21,3	30,8	13	16,1
50-12	17,3	25,3	35,6	16	19,2
50-14	20,6	29,3	40,3	19	22,3
50-18	27,3	37,3	49,5	25	28,6
50-21	32,5	43,3	56,4	29	33,5
80- 0	0,131	0,866	2,88	0,46	1,2
80- 1	0,665	2,10	4,86	1,1	2,0
80- 2	1,38	3,34	6,65	1,7	2,8
80- 3	2,18	4,59	8,35	2,4	3,7
80- 5	3,94	7,09	11,6	4,0	5,4
80- 7	5,82	9,59	14,7	5,6	7,3
80- 8	6,79	10,8	16,2	6,4	8,2
80-10	8,78	13,3	19,3	8,2	10,1
80-12	10,8	15,8	22,2	9,9	12,0
80-14	12,9	18,3	25,2	12	13,9
80-18	17,1	23,3	30,9	16	17,9
80-21	20,3	27,1	35,2	18	20,9

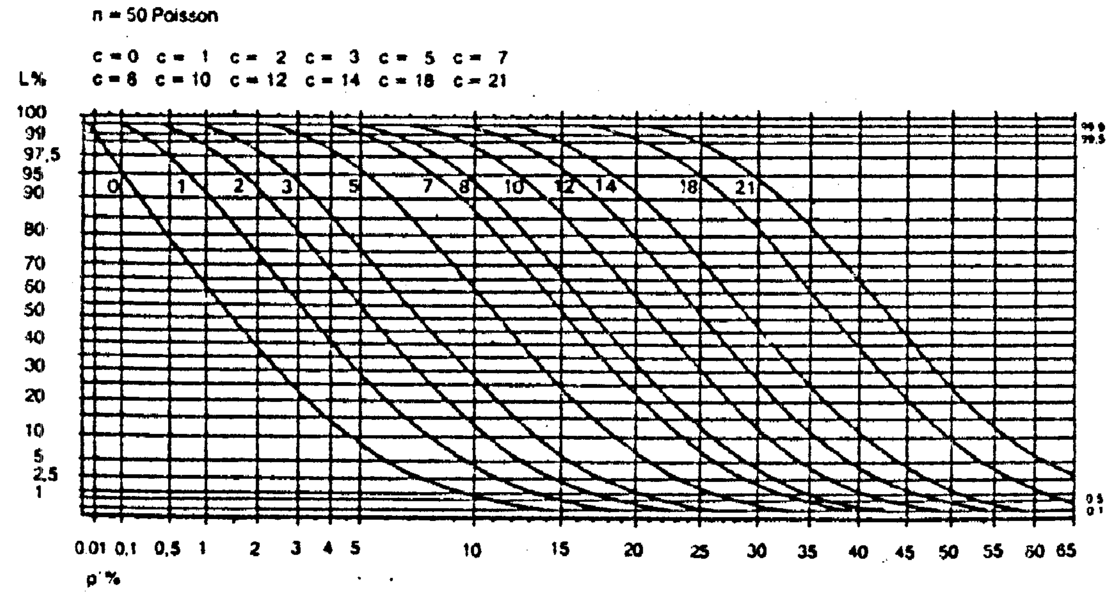
P_1	P_0	P_2
(1,31, 0,9)	(8,66, 0,5)	(28,8, 0,1)

Ferner liest man 12,5% aus der letzten Spalte für jenen Prozentsatz an Ausschuß, der den maximalen Durchschlupf D_{max} = 4,6 (vorletzte Spalte) nach sich zieht.

Dodge und Romig (1959) haben Stichprobenpläne auch nach dem maximalen Durchschlupf AOQL angeordnet. Für einige Stichprobenpläne ist im folgenden der gesamte Verlauf der OC-Kurven wiedergegeben (aus DGQ, 1981a).

Prüfdichte: Den Anteil der geprüften Stücke *n/N* nennt man auch Prüfdichte. Ist das Los einigermaßen homogen hinsichtlich der zu überprüfenden Qualitätsvariablen, so hängt die





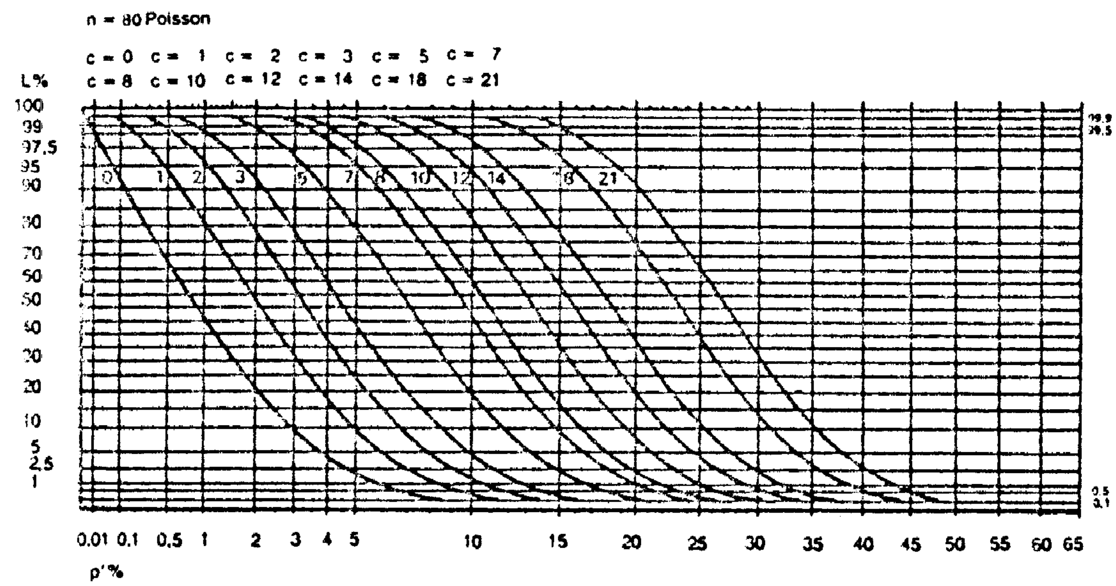


Abb. 8: Vollständige OC-Kurven für ausgewählte Probenpläne

Information, die man aus einer Stichprobe erhält, in erster Linie vom Umfang der Stichprobe ab und nicht von der Losgröße. Haben wir es einmal mit einer Sendung von N=1000 und zum anderen mit N=10000 Einheiten zu tun und entnehmen wir in beiden Fällen eine Probe mit n=100 Einheiten, so ist die daraus resultierende Information gleichwertig, d.h. die Risiken α und β bzw. der maximale Durchschlupf sind gleich, obwohl die Prüfdichte einmal 10%, das andere Mal 1% beträgt. Die Prüfung größerer Lose ist somit wirtschaftlicher als die von kleineren. Vorsicht jedoch, denn es ist nicht angebracht, verschiedene Lose mit unterschiedlicher Qualität zu vermischen und das inhomogene "Mischlos" auf einmal zu prüfen. Damit würde die Wirtschaftlichkeit großer Lose wieder zunichte gemacht. Die folgende Tabelle soll belegen, daß die Prüfdichte mit zunehmendem Losumfang abnehmen kann, obwohl die Risiken α und β gleichzeitig kleiner werden. Der Tabelle liegen die Gutgrenze P_1=2% und die Schlechtgrenze P_2=6% zugrunde, die Werte für α und β wurden aus umfangreicheren Tabellen entnommen.

N	n	100 n/N	c	α	β
400	50	12,5	2	0,08	0,42
750	75	10,0	3	0,07	0,34
1375	110	8,0	4	0,07	0,21
2665	160	6,0	6	0,04	0,16
5625	225	4,0	8	0,04	0,08
15000	300	2,0	10	0,04	0,03
56250	450	0,8	14	0,04	0,005
375000	750	0,2	22	0,03	0,000
1500000	1500	0,1	44	0,006	0,000

2.c *Stichprobenpläne und statistisches Testen*

Im folgenden wird die Entscheidung über Annahme oder Ablehnung eines Loses als statistischer Test formuliert. Daran knüpfen sich einige didaktische Überlegungen, die zeigen sollen, wo man die Qualitätskontrolle mit Vorteil zur Einführung in die Beurteilende Statistik verwenden kann. Die Entscheidungssituation ist ganz natürlich, die OC-Kurve zur Beurteilung des Entscheidungsverfahrens ist naheliegend. Die OC-Kurve stellt nämlich ein Szenario nach dem Motto "Was wäre, wenn" dar, man spielt also alle möglichen Fälle durch. Klar wird auch die Rolle der zufälligen Auswahl, damit man diese Kurve überhaupt erst berechnen kann. Es entfällt auch die beim Testen so schwierige Wahl der Nullhypothese, weil diese eher als einer von mehreren Orientierungspunkten auf der OC-Kurve erscheint. Die Interpretation von Fehlerwahrscheinlichkeiten ist ein leidiges Problem in der beurteilenden Statistik, das üblicherweise mit der Auswirkung bei häufiger Testanwendung angepackt wird. In der Qualitätskontrolle verblaßt diese Deutung auf längere Sicht angesichts viel wichtigerer Gesichtspunkte des Verfahrens, etwa des maximal einzukalkulierenden Durchschlupfs an schlechten Einheiten.

Probenplan als statistischer Test: Formuliert man die Gut-Schlecht-Prüfung als statistischen Test, so hat man erst die Hypothesen über die Grundgesamtheit aufzustellen. Als Nullhypothese wählt man $P=P_1$, der Anteil an schlechten Einheiten im Los ist gleich der Gutgrenze P_1; als Gegenhypothese wählt man $P=P_2$. Die Basis der Entscheidung über das Los ist die Testgröße X, Zahl der schlechten Einheiten in der Stichprobe vom Umfang n. Unter der Voraussetzung der zufälligen Entnahme der Probe ist X hypergeometrisch verteilt mit den Parametern N, A, und n, wobei A mit P über $A = P/N$ zusammenhängt. Das heißt genauer:

- Unter H_0 gilt: $X \sim H(N, A_1 = P_1/N, n)$
- Unter H_1 gilt: $X \sim H(N, A_2 = P_2/N, n)$

In der Gut-Schlecht-Prüfung zeichnet man neben der Nullhypothese noch einen weiteren Punkt aus, nämlich die Schlechtgrenze P_2 und man testet eigentlich nur den 'Punkt' P_1 gegen den 'Punkt' P_2. Wenn Null- und Gegenhypothese jeweils nur eine Verteilung festlegen, so nennt man das einen Alternativtest.

Der Probenplan ist eine Vorschrift N, n, c mit der Folge, daß von den geprüften Einheiten nicht mehr als c schlecht sein dürfen, wenn H_0 nicht verworfen werden soll.

Ereignis	Entscheidung	
	bei Gut-Schlecht-Prüfung	beim Testen
$X \le c$	*für* H_0	T_0 H_0 nicht abgelehnt
$X > c$	gegen H_0	T_1 H_0 abgelehnt

Bemerkung: Die Entscheidung für H_i wird hier und im folgenden mit T_i bezeichnet. Die möglichen Fehlentscheidungen sind der Fehler 1. Art, wobei H_0 zutrifft, aber dennoch abgelehnt wird, sowie Fehler 2. Art, wobei H_0 nicht zutrifft, aber nicht abgelehnt wird. Für die Wahrscheinlichkeiten dieser Fehler gilt dann:

$W(T_1 \mid H_0) = \alpha$	W(Ablehnung | Los ok) = Produzentenrisiko
$W(T_0 \mid H_1) = \beta$	W(Nicht-Ablehnung | Los zu schlecht) = Konsumentenrisiko

Die Fehler beim Alternativtest sind aber nicht die einzigen Anhaltspunkte zur Beurteilung des Testverfahrens. In der Statistik erweitert man Null- und Gegenhypothese zu

$$H_0\colon P \le P_1 \quad H_1\colon P > P_1$$

und berechnet die Wahrscheinlichkeit, H_0 nicht abzulehnen, in Abhängigkeit von dem 'wahren' Wert des Parameters, das ist die sogenannte Gütefunktion $W(T_1 | P)$. Für P aus H_0, d.h. $P \leq P_1$ ist ein möglichst kleiner Wert der Gütefunktion erwünscht, die Gütefunktion gibt in diesem Bereich den Fehler 1. Art an, das Maximum wird bei $P=P_1$ erreicht. Für P aus H_1, d.h. $P>P_1$ ist ein möglichst großer Wert erwünscht, weil die Gütefunktion in diesem Bereich die Wahrscheinlichkeit der richtigen Ablehnung angibt, das ist die Gegenwahrscheinlichkeit zum Fehler 2. Art.

Ablehnen von H_0

Gütefunktion	kleine	große Werte
Wahrer Zustand	H_0	H_1
OC-Kurve	große	kleine Werte

Annehmen von H_0

In der Qualitätskontrolle bevorzugt man die Annahmewahrscheinlichkeit des Loses, das ist die Nullhypothese, in Abhängigkeit vom wahrem Ausschußprozentsatz; das ist die Gegenwahrscheinlichkeit zur Gütefunktion, jetzt einmal *ohne* die Bezugspunkte P_1 und P_2:

$$L(P) = W(T_0 | P) = 1 - W(T_1 | P)$$

Später führt man die Bezugspunkte P_1 und P_2 wiederum ein und bettet das Testproblem P_1 gegen P_2 in das allgemeine Problem $P \leq P_1$ gegen $P \geq P_2$ (2!) ein.

Vor- und Nachteile des Qualitätskontrolle-Kontexts: Ausgangspunkt ist ein Probenplan und das Studium von dessen Eigenschaften. Das Entscheidungsproblem über Annahme und Ablehnung des Loses ist natürlich. Die Testgröße 'Anzahl schlechter Einheiten' ist naheliegend, ihre Verteilung ist aufgrund der zufälligen Auswahl auch klar. Überhaupt wird die Rolle der zufälligen Auswahl besonders einsichtig und ermöglicht so die Herleitung der hypergeometrischen Verteilung für die Testgröße. Das ist bei anderen Testproblemen längst nicht so einfach. Dort wird nämlich die Stichprobe durch n-fache unabhängige Wiederholung desselben Experiments beschrieben. Wenn es dann um Tests um den Mittelwert geht, muß man die Verteilung des Mittelwerts erst bestimmen, was im Falle der Normalverteilung keine leichte Aufgabe ist. Am schwerwiegendsten ist aber, daß hinter komplizierten Begriffen die Wichtigkeit der zufälligen Auswahl verhüllt bleibt.

Bei der Beurteilung eines Probenplans wird die Annahmewahrscheinlichkeit des Loses global in Abhängigkeit von allen möglichen Ausschußanteilen im gesamten Los studiert, das ist die OC-Kurve. Diese OC-Kurve ergibt sich somit als natürliches Instrument, man untersucht ja praktisch alle möglichen Fälle für das Los. Klar, daß man nicht einen der Werte von P für wirklich zutreffend hält, man studiert die Folgen nach dem Motto "was wäre, wenn?" Bei der OC-Kurve handelt es sich praktisch um bedingte Wahrscheinlichkeiten der Annahme des Loses, bedingt auf den jeweiligen Wert von P.

Es entfällt zunächst auch die schwierige Wahl der Nullhypothese und der Gegenhypothese. Es sei nur auf die indirekte Art des statistischen Testens verwiesen, wonach man eigentlich immer das, was man statistisch nachweisen möchte, in die Gegenhypothese zu setzen hat. Der Grund liegt grob darin, daß man nur den Fehler 1. Art kontrolliert, der Fehler 2. Art sich praktisch so nebenbei einstellt und von Fall zu Fall auch recht groß sein kann. Man spezifiziert ja bei der OC-Kurve eines gegebenen Probenplans ja überhaupt keine besonderen Punkte. Wenn man allerdings Probenpläne klassifiziert oder auswählt, braucht man schon einige Anhaltspunkte. Erst jetzt kommen die Gut- und Schlechtgrenzen P_1 und P_2 ins Spiel, aber auch andere Punkte wie der Indifferenzpunkt P_0.

Naheliegende Interpretationen der Risiken α und β jedoch scheitern. Eine frequentistische Deutung von α ist nicht möglich, denn dann müßten immer Lose von der Qualitätslage α vorliegen. Weiters scheitert die Interpretation von P_1 als durchschnittlicher Qualitätslage, vielmehr muß P_1 weit über dem üblichen Durchschnitt von Schlecht-Anteil über längere Zeit hinweg liegen. Die Grenzen sowie die statistischen Risiken α und β erweisen sich als reine Rechengrößen zur Bewertung von Probenplänen. Viel bedeutender wird der mittlere bzw. der maximale mittlere Durchschlupf. Der letztere garantiert etwa eine Obergrenze für den mittleren Durchschlupf bei längerer Anwendung des Probenplanes.

Nachteilig erscheint, daß die Berechnung der OC-Kurve umständlich ist. Erst einmal taucht auch die hypergeometrische Verteilung auf, die man im Unterricht gerne aus Zeitgründen oder wegen der damit verbundenen Kombinatorik vermeiden möchte. Die verwendeten Approximationen durch die Binomial- und Poisson-Verteilung sind zwar nicht ganz so schwierig, aber doch aufwendig zu erklären. Tabellen und Nomogramme zum Ablesen der Wahrscheinlichkeiten können den Rechenaufwand verringern, erfordern aber einige Zeit zum Erklären und einige Gewöhnung von seiten der Lernenden. Der Computer kann jedoch in Zukunft einige dieser Nachteile wieder aufheben.

3. Ausblicke

In diesem Abschnitt werden einige Vertiefungen in die Qualitätskontrolle angeschnitten. Die didaktischen Konsequenzen der Einbettung einer Einführung in das statistische Testen in den Kontext der Qualitätskontrolle im Unterricht werden nocheinmal auf den Punkt gebracht.

Vertiefung der Gut-Schlecht-Prüfung: Ein Probenplan, der die Überprüfung von z.B. 100 Einheiten vorschreibt, könnte in zwei Abschnitte zu 50 Einheiten zerlegt werden. Ist das Ergebnis nach der ersten Probe außergewöhnlich gut, so könnte man gleich das Los annehmen, ist hingegen das Ergebnis sehr schlecht, so könnte man gleich ablehnen - ohne die zweite Probe zu entnehmen. Eine solche Prüfung heißt *mehrfacher* Probenplan. Die obige Überlegung deutet an, daß mehrfache Probenpläne bei gleichen Eigenschaften im Durchschnitt eine geringere Zahl an zu prüfenden Einheiten erfordern.

Eine weitere Verfeinerung des Probenplans besteht darin, daß man schrittweise kleine Proben entnimmt und je nach bis dahin gezählter schlechter Einheiten eine Entscheidung für oder gegen das Los trifft oder eine weitere Probe zieht. Diese Vorgangsweise nennt man sequentielle Probenpläne, sie führt zu einer weiteren Reduktion des Prüfaufwands. Für Details sei auf Bücher zur Qualitätsregelung verwiesen, z.B. auf Rinne-Mittag (1989), oder auf ein Manuskript des Autors (Borovcnik, 1987). OC-Kurven zur Auswahl eines geeigneten Planes findet man in DGQ (1981b).

Qualitätsregelung: In Abschnitt 1.c wurde das Problem der laufenden Qualitätsregelung angesprochen. Es kann hier aus Platzgründen nicht ausführlich behandelt werden. Die Qualitätsregelung kann für den Mittelwert, die Streuung oder andere Parameter der Stichprobe (und damit der Grundgesamtheit) erfolgen. Man hat es jetzt mit einem Prozeß zur Erzeugung von Daten zu tun. Das Verfahren 'Prüfen auf Sollwert' läuft im wesentlichen auf einen Signifikanztest der H_0: aktueller 'Mittelwert = Sollwert' hinaus. In der Praxis werden die Daten der Stichprobe auf eine sogenannte Qualitätsregelkarte eingetragen. Fällt der Mittelwert der Probe aus einem bestimmten Intervall hinaus, so muß die laufende Produktion unterbrochen und nach den Ursachen dafür gesucht werden. Für Details sei wiederum auf Rinne-Mittag (1989) oder Borovcnik (1987) verwiesen.

Statistisches Testen im Kontext der Qualitätsprüfung: Es wurde im Zusammenhang mit der Gut-Schlecht-Prüfung einiges dazu gesagt, wie der Kontext die mathematischen Begriffe sinnvoll erscheinen läßt. Ganz besonders wichtig daran ist, daß die Fragestellungen ganz natürlich sind und daß die Deutung der Begriffe einfacher wird. Manche der Begriffe, die in der vom Kontext abgeschoteten mathematischen Form gar so wichtig werden, sind im Kontext entweder anders zu verstehen oder eben nicht so wichtig, weil andere Kenngrößen für den Anwender entscheidend sind. Dazu wurde das Beispiel des α- und β-Fehlers gebracht. Diese sind reine Rechengrößen und deswegen so wichtig, weil man damit die Probenpläne besser ordnen kann. Für die Praxis zählen aber Kennzahlen für den Durchschlupf viel mehr. Im Wechselspiel zwischen Anwendungskontext und Mathematik ergibt sich eine Hilfe für den Aufbau geeigneter Vorstellungen von den Begriffen, siehe dazu auch Borovcnik (1992). Für weitere Anregungen, wie man Qualitätskontrolle in den Unterricht einbauen kann, sei auf DIFF verwiesen, wo auch eine Fülle von Beispielen enthalten ist.

Literatur:

Borovcnik, M.: *Statistische Qualitätskontrolle*, Manuskript zu einer Vorlesung, 1987.

Borovcnik, M.: *Stochastik im Wechselspiel von Intuitionen und Mathematik*, Bibliographisches Institut, Mannheim, 1992.

Cameron, J.M.: Tables for Constructing and for Computing the Operating Characteristics of Single Sampling Plans, *Industrial Quality Control* IX (1952/53) 1, 37-39.

DGQ: *Stichprobenprüfung anhand qualititativer Merkmale*, Deutsche Gesellschaft für Qualität, Frankfurt, 1981a.

DGQ: *Stichprobenprüfung für kontinuierliche Fertigung anhand qualitativer Merkmale*, Deutsche Gesellschaft für Qualität, Frankfurt, 1981b.

DGQ: *Methoden zur Ermittlung geeigneter AQL-Werte*, Deutsche Gesellschaft für Qualität, Frankfurt, 1981c.

DIFF: Das Aufgabenfeld Qualitätskontrolle, *Mathematik - Aufgabenstellen im Stochastikunterricht*, Deutsches Institut für Fernstudien, Tübingen, 1989.

Dodge, H.-F. und H.G. Romig: *Sampling Inspection Tables, Single and Double Sampling*, 2nd ed., Wiley, New York, 1959

Rinne, H. und H.-J. Mittag: *Statistische Methoden der Qualitätssicherung*, Hanser, München, 1989.

Schaafsma, A.H. und F.G. Willemze: *Moderne Qualitätskontrolle*, Philips Technische Bibliothek, Eindhoven, 1961.