

## Bachelorarbeit

# **OC-Funktion bei diskreten Lebensdauerverteilungen anhand der Poisson-Verteilung und der geometrischen Verteilung**

Rabea Dagmar Aschenbruck

3. Juni 2014

Gutachterin: Prof. Dr. Christine Müller

Technische Universität Dortmund

Fakultät Statistik

Lehrstuhl Statistik mit Anwendungen im Bereich der Ingenieurwissenschaften



# Inhaltsverzeichnis

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Einleitung</b>  | <b>1</b>  |
| <b>2</b> | <b>Motivation und Einführung: Qualitätssicherung</b>           | <b>3</b>  |
| 2.1      | Eigenschaften von Qualität . . . . .                           | 3         |
| 2.2      | Qualitätsmanagement . . . . .                                  | 6         |
| 2.3      | Statistische Qualitätssicherung . . . . .                      | 7         |
| <b>3</b> | <b>Theorie</b>   | <b>9</b>  |
| 3.1      | Kurze Einführung . . . . .                                     | 9         |
| 3.1.1    | Zufallsvariable . . . . .                                      | 9         |
| 3.1.2    | Betrachtung diskreter Zufallsvariablen . . . . .               | 10        |
| 3.1.3    | Betrachtung stetiger Zufallsvariablen . . . . .                | 14        |
| 3.2      | Eindimensionale Wahrscheinlichkeitsverteilungen . . . . .      | 15        |
| 3.2.1    | Bernoulli-Verteilung . . . . .                                 | 16        |
| 3.2.2    | Binomialverteilung . . . . .                                   | 17        |
| 3.2.3    | Geometrische Verteilung . . . . .                              | 19        |
| 3.2.4    | Negative Binomialverteilung . . . . .                          | 20        |
| 3.2.5    | Poisson-Verteilung . . . . .                                   | 22        |
| 3.2.6    | Normalverteilung . . . . .                                     | 25        |
| 3.2.7    | $\chi^2$ -Verteilung . . . . .                                 | 25        |
| 3.3      | Zählende Abnahmeprüfung . . . . .                              | 27        |
| 3.3.1    | Einführung in die statistische Testtheorie . . . . .           | 27        |
| 3.3.2    | Allgemeine Durchführung der zählenden Abnahmeprüfung . . . . . | 30        |
| 3.3.3    | Test bei Poisson-verteilten Lebensdauern . . . . .             | 33        |
| 3.3.4    | Test bei geometrisch verteilten Lebensdauern . . . . .         | 34        |
| <b>4</b> | <b>Aufbau der Simulationsstudie</b>                            | <b>37</b> |
| 4.1      | Erläuterung der Funktionen . . . . .                           | 37        |
| 4.1.1    | Testdurchführung . . . . .                                     | 37        |

|          |   |           |
|----------|---|-----------|
| 4.1.2    | Wiederholung des Tests . . . . .  | 38        |
| 4.1.3    | Erstellen der OC-Funktion . . . . .   | 39        |
| 4.2      | Ablauf der Simulationsstudie . . . . .  | 40        |
| <b>5</b> | <b>Auswertung</b>   | <b>43</b> |
| 5.1      | Untersuchung der Simulationsstudie . . . . .  | 43        |
| 5.2      | Vergleich der simulierten mit der exakten OC-Funktion . . . . .                                 | 50        |
| 5.3      | Vergleich der Poisson- und der geometrischen Verteilung als Lebens-<br>dauervertelung . . . . . | 52        |
| <b>6</b> | <b>Zusammenfassung</b>  | <b>55</b> |
| <b>A</b> | <b>Programmiercode</b>  | <b>57</b> |
|          | <b>Literaturverzeichnis</b>   | <b>61</b> |
|          | <b>Erklärungen</b>  | <b>63</b> |

# Kapitel 1

## Einleitung

Um am Markt bestehen zu können, ist für Hersteller die Qualität ihrer Produkte von entscheidender Bedeutung. Daher wird auf Qualität in der Kommunikation nach außen regelmäßig Bezug genommen. Ein gutes Beispiel hierfür liefert das Qualitätsversprechen von Toyota, welches ein Zitat des Präsidenten von Toyota Deutschland untermauert: „Unser Ziel ist und bleibt höchste Kundenzufriedenheit. Deshalb streben wir täglich nach der besten Qualität unserer Autos und arbeiten kontinuierlich mit voller Kraft an ihrer Verbesserung.“ (vgl. Toyota Deutschland GmbH, 2014). Auch auf der Internetseite des Herstellers spielen Qualität und Kundenzufriedenheit eine große Rolle, indem unter anderen über aktuelle Rückrufaktionen informiert wird. Dabei steht eine geringe Anzahl von Rückrufen und der offene Umgang mit diesen ganz offensichtlich für nachhaltige Qualität. Im Jahr 2009 hatten vermehrte Rückrufaktionen das Image von Toyota schwer beschädigt. Imageschäden können für Unternehmen zu längerfristigen Nachteilen werden. So vermutete Ferdinand Dudenhöffer Anfang 2010, dass sich der Marktwert der Fahrzeuge senken werde und das Unternehmen die Folgen noch über einen Zeitraum von 5 Jahren spüren werde (vgl. TOMORROW FOCUS Media GmbH, 2010). Der Grund für die Rückrufaktionen war lediglich ein fragiles Gaspedal, welches aus Kostengründen, insbesondere in der Entwicklung, in über 8 Toyota-Modellreihen zum Einsatz kam. Das Gaspedal wurde nicht von Toyota selbst, sondern von einem Zulieferer gefertigt, wie dies etwa bei 75% der für die Herstellung des Autos benötigten Teile der Fall ist (vgl. Roman, 2010).

Mit dem soeben genannten Beispiel kann aufgezeigt werden, welche Ausmaße ein einziges, für sich genommen nur wenige Euro kostendes, falsch gefertigtes Teil haben kann. Daher sind Produzenten bei der Abnahme von Produktteilen der Zulieferer sehr auf Qualität bedacht. In der folgenden Bachelorarbeit wird eine Simulations-

studie durchgeführt, welche die Abnahmeprüfung untersucht. Es soll der Einfluss von verschiedenen Parametern analysiert werden, um so eine für Produzent und Zulieferer bestmögliche Abnahmeprüfung zu ermöglichen.

In Kapitel 2 wird in die Thematik der Qualitätssicherung eingeführt. Nachdem zu Beginn die Begriffe Qualität und Zuverlässigkeit definiert werden, wird anschließend die Motivation für Qualitätsmanagement kurz angesprochen und in die dafür nötigen Teilbereiche der statistischen Qualitätssicherung eingeführt. Im anschließenden Kapitel Theorie werden die, in der nachfolgenden Arbeit benötigten, mathematischen Grundlagen definiert und eingeführt. Insbesondere werden die zählenden Abnahmeprüfungen für Poisson-verteilte und geometrisch verteilte Lebensdauern hergeleitet sowie die zugehörigen exakten OC-Funktionen bestimmt, welche als Grundlage der Analyse der Bachelorarbeit dienen. Das vierte Kapitel erläutert die für die Durchführung der Simulationsstudie benötigten und programmierten R-Funktionen und beschreibt den Ablauf der Studie. Dies geschieht in Form der Festlegung der Variation des Signifikanzniveaus  $\alpha$ , des Stichprobenumfangs  $n$ , der Mindestlebensdauer  $t_0$  und des maximal akzeptierten Losschlechtanteils  $p_0$ . Auf diese Weise werden Informationen gewonnen, auf deren Grundlage Aussagen über den Einfluss der einzelnen Parameter getroffen werden können. Die Ergebnisse der im vierten Kapitel erläuterten Simulationsstudie werden im Kapitel 5 untersucht. Es werden zunächst die grafisch aufbereiteten Ergebnisse der Studie ausgewertet. Diese Untersuchung wird zeigen, dass sich die vier Parameter in ähnlicher Weise auf die OC-Funktionen auswirken, unabhängig davon, welche Verteilung die Lebensdauern modelliert. Anschließend werden die OC-Funktionen der Simulationsstudie mit den im Kapitel 3 hergeleiteten exakten OC-Funktionen verglichen, um auf diese Weise eine Verallgemeinerung der Ergebnisse der Studie auf die exakten OC-Funktionen zu rechtfertigen. Abschließend werden die Vor- und Nachteile der Darstellung der Lebensdauern über die Poisson-Verteilung und über die geometrische Verteilung diskutiert. Die Bachelorarbeit beendet eine Zusammenfassung des Vorgehens bei dieser Simulationsstudie und der zentralen Ergebnisse.

# Kapitel 2

## Motivation und Einführung: Qualitätssicherung

Zur Einführung in das Thema der Qualitätssicherung werden zunächst die Grundlagen definiert. Was versteht man unter "Qualität", ist bereits ein sehr geringer Unterschied zu einer Zielgröße zu beanstanden oder sind gewisse Qualitätsschwankungen - und wenn ja, in welchen Grenzen - akzeptabel? Nachdem dies erläutert und auch der Begriff der Zuverlässigkeit eingeführt wurde, folgt ein Überblick über Qualitätsmanagement. Dabei wird deutlich gemacht, worin aus Sicht eines Herstellers die Motivation für eine erfolgreiche Steuerung der Qualität, das Qualitätsmanagement, liegt. Hierzu werden die zur Verfügung stehenden Instrumente mit dem Schwerpunkt Qualitätssicherung betrachtet. Abschließend wird in die Teilbereiche der statistischen Qualitätssicherung eingeführt und der heutige Stand beschrieben.

### 2.1 Eigenschaften von Qualität

#### Qualität und Zuverlässigkeit

Was ist Qualität? Die Interpretation dieses Begriffs ist abhängig von der Betrachtungsweise und dem Aspekt, worüber eine Aussage getroffen werden soll. Eine allgemeine Definition lieferte 1992 die DIN ISO 8402 (nachzulesen in Rinne, Mittag, 1995, S. 9). Danach ist Qualität die

"Gesamtheit von Eigenschaften und Merkmalen eines Produkts oder einer Dienstleistung, die sich auf deren Eignung zur Erfüllung festgelegter oder vorausgesetzter Erfordernisse beziehen."

Diese Definition wurde im Dezember 2000 mit der EN ISO 9000:2000 aktualisiert. Diese Norm definiert Qualität als

**Definition 2.1 (*Qualität*)**

“Grad, in dem ein Satz inhärenter Merkmale Anforderungen erfüllt“ (DIN Deutsches Institut für Normung e.V., 2000, S. 18).

Qualität kann laut dieser Definition mit Adjektiven benannt und konkretisiert werden.

Bezogen auf verschiedene Produkte sind die in der Definition angesprochenen Anforderungen vielfältig. Dies soll anhand einiger Beispiele verdeutlicht werden. Den *Konsumenten* von Apple-Produkten ist häufig das Design wichtig. Für Käufer von Bio-Waren ist der Aspekt der Ressourcen-schonenden Herstellung entscheidend. Auch der Preis des Produktes muss sich im Rahmen der Kundenerwartung bewegen, ebenso wie der bereitgestellte Service des Produzenten bzw. Verkäufers. Diese genannten Beispiele betreffen die Erwartungen von Konsumenten an ein Produkt. Die Ziele der *Produzenten* können sich hiervon jedoch durchaus unterscheiden. Auf jeden Fall dürfte für sie ein langfristig guter Ertrag und der Markterfolg des Produkts wichtig sein. Des Weiteren sollte ein Produkt möglichst geringe Risiken für Haftungsansprüche verursachen und mit niedrigen Entwicklungs- und Produktdurchlaufzeiten gefertigt werden. Für die *Gesellschaft* sind Naturschutzaspekte von großer Bedeutung, wie zum Beispiel eine umweltschonende Produktion und Nutzung. Auch die Entsorgung ist von Interesse, wie das Thema Atommüll zeigt. Des Weiteren spricht für die Qualität eines Produkts, falls von ihm eine vergleichsweise geringere Gefährdung von Dritten ausgeht (vgl. Rinne, Mittag, 1995, S. 9).

Erweitert man den Betrachtungshorizont auf den gesamten Zeitraum einer möglichen Produktnutzung, sind die qualitätsbezogenen Eigenschaften um den Aspekt der Zuverlässigkeit zu ergänzen. Dieser bedeutet eine lange Lebensdauer des Produktes, also einen langen Zeitraum, in dem das Produkt funktionstüchtig ist. Allgemein wird der merkmalsbezogene Begriff Zuverlässigkeit in der EN ISO 9000:2000 definiert als

**Definition 2.2 (*Zuverlässigkeit*)**

“Zusammenfassender Ausdruck zur Beschreibung der Verfügbarkeit und ihrer Einflussfaktoren Funktionsfähigkeit, Instandhaltbarkeit und Instandhaltungsbereitschaft“ (DIN Deutsches Institut für Normung e.V., 2000, S. 26).



Diese Anforderung ist für alle drei Interessengemeinschaften von Bedeutung. Dabei steht für Konsumenten die lange Haltbarkeit und somit ein vermeintlich gutes Preis-Leistungs-Verhältnis im Vordergrund. Die Hersteller verbinden damit ein gutes Image und die Öffentlichkeit präferiert den Aspekt der Abfallvermeidung.

Es ist leicht zu erkennen, dass Qualität sehr unterschiedliche Ansprüche zu befriedigen hat. Demzufolge müssen zwischen der hohen Anzahl aller denkbaren qualitätsbeeinflussenden Anforderungen und den wirtschaftlich realisierbaren Eigenschaften eines Produktes Kompromisse gefunden werden. Diese werden in der Praxis durch Zielwerte und Toleranzgrenzen für jeden Zeitpunkt der Lebensdauer des Produkts umgesetzt. Toleranzgrenzen sind deswegen sinnvoll, da es bei Produktionen immer zu Schwankungen der Qualität kommt.

## Ursachen für Schwankungen

Was ist mit Schwankung der Qualität gemeint? Die Praxis zeigt, dass ein vorgegebener Zielwert bezüglich eines Merkmals des Produktes in der Realität bei den produzierten Einheiten nicht immer exakt erreicht wird. Dabei ist zu beachten, dass nach Rinne und Mittag drei verschiedene Arten von Schwankungen bei der Qualität eines Merkmals existieren. Es ist zum einen möglich, dass eine Abweichung der Qualität innerhalb einer Einheit auftritt. Dies kann zum Beispiel der Durchmesser einer Schraube sein, welcher nicht an jeder Stelle des Schraubenhalses dem Zielwert entspricht. Des Weiteren können Unterschiede zwischen den Einheiten auftreten. Das bedeutet, dass nicht jede Schraube als Durchmesser den Zielwert besitzt, sondern einige Einheiten einen anderen Wert aufweisen. Die dritte Art der Schwankung beschreibt die Abweichungen über die Zeit. Bei dem Schraubenbeispiel würde das bedeuten, dass die anfänglich produzierten Schrauben den Zielwert als Durchmesser aufweisen, die später produzierten Schrauben jedoch nicht. Die Gründe für diese Schwankungen zwischen Zielwert und tatsächlichem Wert werden im Folgenden erläutert (vgl. Rinne, Mittag, S. 10).

Zur Strukturierung der Ursachen für die Streuung der wahren Werte um den Zielwert hat Timischl fünf Oberbegriffe gebildet: Mensch, Maschine, Methode, Material und Milieu (vgl. Timischl, 2002, S. 31). Diese Gruppen sind einzeln oder aber in Kombination untereinander für die Fehlerquellen verantwortlich. Unter dem Oberbegriff *Mensch* versteht Timischl, dass die Fehler aus menschlichem Versagen resultieren. Dies bedeutet bei einem Fertigungsprozess, dass ein Mensch eine Maschine falsch bedient oder regelt. Auf diese Weise wird der Fertigungsprozess verfälscht und es

können nicht mehr alle Einheiten mit dem vorgegebenen Zielwert produziert werden. Der Begriff *Maschine* fasst die Fehler zusammen, die ursächlich durch die im Fertigungsprozess eingesetzten Maschinen entstehen können. Zum Beispiel kann eine Materialermüdung dazu führen, dass die Maschine nach einiger Zeit nicht mehr die exakten Vorgaben der Produktionsnorm gewährleisten kann. Eine weitere Fehlerquelle kann durch falsche Ausrichtung oder Wartung der Maschine entstehen. Auch die Wahl der *Methode* ist laut Timischl eine Entstehungsquelle für die letztendliche Streuung. Für verschiedene Materialien sind verschiedene Bearbeitungsprozesse unterschiedlich effektiv. Auch Wartezeiten durch falsch bestimmte Prozessabläufe können die Einhaltung der Zielwerte verhindern. Der Oberbegriff *Material* fasst alle Fehler zusammen, welche aus der womöglich unterschiedlichen Materialbeschaffenheit resultieren. Schließlich umfasst der Begriff *Milieu* die Umwelteinflüsse, welche den Fertigungsprozess in seiner Präzision beeinträchtigen. Beispielhaft sind Luftdruck, Raumtemperatur oder Luftfeuchtigkeit zu nennen, aber auch Erschütterungen oder Staub zählt Timischl als mögliche Fehlerquellen auf (vgl. Timischl, 2002, S. 31).

## 2.2 Qualitätsmanagement

Für den Erhalt hoher Wettbewerbsfähigkeit wurde in der Vergangenheit für Unternehmen die Qualität immer wichtiger und stellt inzwischen für viele ein strategisches Unternehmensziel dar. Insoweit wollen sie ihren Kunden ein attraktives Preis-Leistungs-Verhältnis bieten. Gleichzeitig müssen Produzenten darauf achten, dass die Produkte möglichst kostengünstig hergestellt werden. Je später Fehler in dem gesamten Prozess der Produktentstehung aufgedeckt werden, desto teurer wird dies für den Produzenten. Die Ursache für einen Großteil der Fehler liegt in der unzulänglichen Planung und Entwicklung neuer Produkte, was häufig zu einer kostenintensiven Mängelbeseitigung nach Abschluss des Fertigungsprozesses führt. Vor diesem Hintergrund wird das Qualitätsmanagement immer bedeutender, da es Möglichkeiten der Kosteneinsparung bietet (vgl. Timischl, 2002, S. 5). Ein zeitgemäßes Qualitätsmanagement kann außerdem schnell auf Wünsche der Kunden reagieren und die Verbesserung der Qualität vorantreiben. Bei der Produktion von Gütern beinhaltet dies eine möglichst kurze und effektive Produktentwicklungszeit sowie der ständige Einsatz von qualitätssichernden Maßnahmen.

## 2.3 Statistische Qualitätssicherung

Die statistische Qualitätssicherung ist ein bedeutender Teilbereich des Qualitätsmanagements und beinhaltet nach Rinne und Mittag statistische Verfahren zur Auswertung von Prozessverläufen, zur Überwachung von Serienfertigungen und zur Endkontrolle (vgl. Rinne, Mittag, 1995, S. 15). Diese Verfahren werden angewendet, um die vorgegebenen Zielwerte und Toleranzgrenzen einzuhalten.

### Definition 2.3 (*Statistische Qualitätskontrolle*)

Die statistische Qualitätssicherung leitet anhand von Stichprobenprüfungen unter Verwendung mathematisch-statistischer Methoden Aussagen über den Grad der Erfüllung von Qualitätsforderungen ab (aus Rinne, Mittag, 1995, S. 17).

In der statistischen Qualitätssicherung werden in der Regel über Stichproben Aussagen getroffen und anschließend auf die Grundgesamtheit übertragen. Dabei kann es zu fehlerhaften Annahmen kommen. Diese werden jedoch aus Zeit- und Kostengründen in Kauf genommen, da es auch bei Vollkontrollen, beispielsweise auf Grund von möglichen Messfehlern, zu Fehlaussagen kommen kann.

### Aufgabenbereiche

Im Folgenden wird ein Überblick über einige Aufgabenbereiche der statistischen Qualitätssicherung gegeben. Die vermutlich älteste Form der Qualitätssicherung ist die *Abnahmeprüfung*. Hierbei wird eine endliche Gesamtheit von Produkteinheiten, in der Qualitätssicherung als Los bezeichnet, nach dem Fertigungsprozess in ihrer Qualität beurteilt. Dies erfolgt meistens im Auftrag des Konsumenten, also des potenziellen Käufers der Produkte, und mit dessen individuellen Qualitätsanforderungen. Es existieren zwei Arten der Abnahmeprüfung, die zählende und die messende Abnahmeprüfung. Diese basieren auf einer Stichprobe des Loses. Dabei muss das Los insoweit homogen sein, dass die enthaltenen Einheiten unter denselben Bedingungen gefertigt wurden und die Stichprobe zufällig aus dem Los gezogen wurde. Die Abnahmeprüfung betrachtet dann jeweils das Merkmal, welches den Qualitätsanforderungen des Konsumenten entsprechen soll. Bei der messenden Abnahmeprüfung wird die Ausprägung des interessierenden Merkmals der Einheiten der Stichprobe betrachtet. Bei der zählenden Abnahmeprüfung hingegen interessiert die Anzahl der Fehler oder die Anzahl der fehlerhaften Einheiten der Stichprobe aus dem Los. Bei beiden Vorgehen existieren Prüfpläne zur Durchführung der Abnahmeprüfungen.

Auf Grundlage dieser Prüfpläne wird entschieden, ob das Los den Qualitätsanforderungen genügt oder nicht (vgl. Rinne, Mittag, S. 141).

Eine weitere Prüfung im Bereich der statistischen Qualitätssicherung ist die *Zuverlässigkeitsprüfung*. Diese Abnahmeprüfung ist an dem Qualitätsmerkmal Zuverlässigkeit interessiert. Bei der Prüfung handelt es sich meistens um eine zerstörende Prüfung. Das bedeutet, dass die Einheiten der Stichprobe bis zum Funktionsausfall beobachtet werden. Zuverlässigkeitsprüfungen sind aufwändig, da zum Teil sehr viel Zeit bis zum Ausfall vergeht und die ausgefallenen Einheiten anschließend nicht mehr verwendet werden können. Daher ist der Stichprobenumfang bei Zuverlässigkeitsprüfungen meistens relativ klein. Als Qualitätsanforderung des Konsumenten ist eine Mindestlebensdauer gegeben, also eine Zeitspanne in der die Einheiten voll funktionsfähig sein sollen. Dies wird ebenfalls mittels eines Prüfplans getestet und anschließend entschieden, ob das Los verworfen werden sollte oder nicht. Dabei bedeutet Los verwerfen, dass es zu keinem Kauf des Loses von Seiten des Konsumenten kommt. Diese Bachelorarbeit untersucht im Folgenden das Verhalten von Zuverlässigkeitsprüfungen (vgl. Timischl, S. 281).

Der Vollständigkeit halber wird abschließend mit der Fertigungsüberwachung ein weiterer Aufgabenbereich der Qualitätssicherung betrachtet. Im Gegensatz zu der Abnahmeprüfung und der Zuverlässigkeitsprüfung ist das Ziel der *Fertigungsüberwachung*, dass bereits im laufenden Herstellungsprozess Qualitätsmängel erkannt werden. Diese Mängel sollen in einer Prüfstation bemerkt werden, welche einzeln von jeder produzierten Einheit passiert wird. Die Fertigungsüberwachung verfolgt zwei Ziele. Zum einen wird durch Stichprobenkontrollen geprüft, wie viele defekte Einheiten die Prüfstation unbehelligt passieren. Dies wird mittels kontinuierlicher Prüfpläne umgesetzt. Des Weiteren veranlasst die Fertigungsüberwachung bei nicht zufriedenstellender Qualität der Einheiten in den Stichproben eine Änderung des Prozesses. Mit Hilfe von sogenannten Regelkarten wird die zweite Zielsetzung der Fertigungsüberwachung umgesetzt. Weitere Informationen zu den Verfahren sind zu finden bei Timischl (vgl. Timischl, 2002, S. 177 ff.) sowie bei Rinne und Mittag (vgl. Rinne, Mittag, 1995, S. 3003 ff.).

# Kapitel 3

## Theorie

### 3.1 Kurze Einführung

Dieses Kapitel beginnt mit einem Überblick über die Grundlagen der Statistik und liefert die Definition für eine Zufallsvariable. Anschließend werden wichtige Funktionen und Parameter zunächst für diskrete und schließlich für stetige Zufallsvariablen vorgestellt.

#### 3.1.1 Zufallsvariable

Im weiteren Verlauf dieser Bachelorarbeit werden sich gegenseitig ausschließende Ausgänge von Zufallsexperimenten  $\omega_i$  betrachtet. Dabei bezeichnet  $\omega_i$ , mit  $i = 1, \dots, n$ , das  $i$ -te mögliche Ergebnis. Da der Ausgang eines solchen Experiments vom Zufall abhängt, kann das Ergebnis nicht mit Sicherheit vorhergesagt werden. Die insgesamt möglichen Ergebnisse werden in dem Ergebnisraum  $\Omega$  zusammengefasst. Teilmengen dieses Ergebnisraums heißen Zufallsereignisse. Es soll mit einer Wahrscheinlichkeit  $P$  bestimmt werden, wie hoch die Chance ist, dass ein bestimmtes Ereignis  $A$  eintritt. Diese Wahrscheinlichkeit muss folgende Axiome erfüllen:

**Axiom 3.1** (*Axiome von Kolmogorov*)

a)  $P(A) \geq 0$

b)  $P(\Omega) = 1$

c) falls  $A \cap B = \emptyset$ , so ist  $P(A \cup B) = P(A) + P(B)$

(vgl. Fahrmeir et al., 2010, S. 182)

Die Wahrscheinlichkeit  $P$  weist jedem Ereignis  $A \subset \Omega$  eine Zahl zwischen Null und Eins zu.

Eine Zufallsvariable  $X$  beschreibt ein Merkmal anhand der Ausprägung des Ergebnisses eines Zufallsvorgangs. Dabei ist  $x$  die Ausprägung, welche die Zufallsvariable  $X$  bei der Durchführung des Zufallsexperiments annimmt und wird Realisation genannt. Für die weitere Arbeit wird  $x \in \mathbb{R}$  angenommen.

**Definition 3.2 (Zufallsvariable)**

Bei einem Zufallsexperiment mit dem gegebenen Ergebnisraum  $\Omega$  wird die Abbildung

$$\begin{aligned} X : \Omega &\rightarrow \mathbb{R} \\ \omega &\mapsto X(\omega) = x \end{aligned}$$

eine (eindimensionale) Zufallsvariable genannt (vgl. Bamberg et al., 2008, S. 93).

Derart definierte Zufallsvariablen  $X$  sind entweder diskrete oder stetige Zufallsvariablen. Diese werden nun im Folgenden betrachtet und zugehörige Funktionen definiert. Des Weiteren werden jeweils Parameter für diskrete und stetige Zufallsvariablen eingeführt.

### 3.1.2 Betrachtung diskreter Zufallsvariablen

Eine Zufallsvariable  $X$  nennt man diskret, wenn  $X$  nur endlich viele Werte  $x_1, x_2, \dots, x_k$  oder abzählbar unendlich viele Werte  $x_1, x_2, \dots$  annehmen kann. Diese möglichen Ausprägungen werden in der Menge  $\mathcal{T}$  zusammengefasst. Dies führt unter Berücksichtigung der Axiome von Kolmogorov zu folgender Definition der Wahrscheinlichkeitsverteilung:

**Definition 3.3 (Wahrscheinlichkeitsfunktion)**

Die Wahrscheinlichkeitsfunktion  $f : \mathbb{R} \rightarrow [0, 1]$  einer diskreten Zufallsvariablen  $X$  ist für  $x \in \mathbb{R}$  definiert durch

$$f(x) = \begin{cases} P(X = x_i) = p_i, & \text{mit } x = x_i \in \mathcal{T} \\ 0, & \text{sonst} \end{cases}$$

(vgl. Fahrmeir et al., 2010, S. 229). Dabei gilt  $\sum_{x_i \in \mathcal{T}} f(x_i) = 1$ .

Mit Hilfe der soeben definierten Wahrscheinlichkeitsfunktion kann die Verteilungsfunktion  $F(x) = P(X \leq x)$  gebildet werden. Zu diesem Zweck wird das Ereignis  $A$  gesetzt als  $A = \{x_i \in \mathcal{T} \mid x_i \leq x\}$ .  $P(X \in A)$  beschreibt dann die Verteilungsfunktion einer diskreten Zufallsvariablen an der Stelle  $x$ .

**Definition 3.4 (Verteilungsfunktion für diskrete Zufallsvariablen)**

Die Verteilungsfunktion  $F : \mathbb{R} \rightarrow [0, 1]$  einer diskreten Zufallsvariablen  $X$  ist definiert durch

$$F(x) = P(X \leq x) = \sum_{i: x_i \leq x} f(x_i)$$

(vgl. Fahrmeir et al., 2010, S. 230). Dabei gilt für diese Funktion  $F$

$$\lim_{x \rightarrow -\infty} F(x) = 0 \text{ und } \lim_{x \rightarrow +\infty} F(x) = 1$$

(vgl. Rasch, 1995, S. 64).

Durch ein sogenanntes  $p$ -Quantil  $x^{(p)}$  kann der Definitionsbereich einer Verteilungsfunktion in zwei Mengen  $M_1 = \{x_i \in \mathcal{T} \mid x_i \leq x^{(p)}\}$  und  $M_2 = \{x_i \in \mathcal{T} \mid x_i > x^{(p)}\}$  geteilt werden. Es gilt

$$\frac{\#M_1}{\#\mathcal{T}} \leq p \text{ und } \frac{\#M_2}{\#\mathcal{T}} \geq 1 - p.$$

**Definition 3.5 ( $p$ -Quantil)**

Sei  $X$  eine diskrete Zufallsvariable mit Verteilungsfunktion  $F(x)$ . Der Wert  $x^{(p)}$  mit  $0 < p < 1$ , für den

$$F(x^{(p)}) \geq p$$

gilt und gleichzeitig für jedes  $x < x^{(p)}$

$$F(x) < p$$

zutritt, wird das  $p$ -Quantil der diskreten Zufallsvariable  $X$  mit Verteilungsfunktion  $F(x)$  genannt. Mit Hilfe der verallgemeinerten Inversen gilt nach Hartung

$$F^{-1}(p) = \inf\{x : F(x) \geq p\}$$

(vgl. Hartung, 2009, S. 114).

Das Quantil  $x^{(0.5)}$  wird der Median genannt. Dieses spezielle Quantil ist, genau wie jedes  $p$ -Quantil, ein Lageparameter, wovon im Folgenden weitere eingeführt werden. Anhand dieser numerischen Werte können Verteilungen untereinander verglichen werden. Der wohl bekannteste Lageparameter ist der Erwartungswert, welcher Angaben über das Zentrum der Verteilung der Zufallsvariablen enthält.

**Definition 3.6 (Erwartungswert einer diskreten Zufallsvariablen)**

Der Erwartungswert  $E(X)$  einer diskreten Zufallsvariablen  $X$  wird mit Hilfe der zugehörigen Wahrscheinlichkeitsfunktion  $f(x)$  wie folgt definiert:

$$E(X) = \sum_{i \geq 1} x_i \cdot f(x_i)$$

(vgl. Fahrmeir et al., 2010, S. 242).

Für den soeben definierten Erwartungswert gilt Folgendes:

**Satz 3.7 (Eigenschaften des Erwartungswerts)**

Seien  $X, Y$  Zufallsvariablen und  $a, b \in \mathbb{R}$ . Dann gilt

- a)  $E(aX) = aE(X)$
- b)  $E(X + Y) = E(X) + E(Y)$
- c)  $E(b) = b$

(vgl. Dehling, Haupt, 2004, S. 86).

Der bislang vorgestellte Erwartungswert gibt Auskunft über die Lage einer Verteilung, jedoch nicht über ihre Streuung. Dazu wird im Folgenden als Streuungsparameter die Varianz eingeführt.

**Definition 3.8 (Varianz einer diskreten Zufallsvariable)**

Die Varianz  $Var(X)$  einer diskreten Zufallsvariablen  $X$  wird bestimmt mit

$$Var(X) = \sum_{i \geq 1} (x_i - E(X))^2 \cdot f(x_i)$$

(vgl. Fahrmeir et al., 2010, S. 249).

Wesentliche Eigenschaften der Varianz beschreibt Rasch in seinem Werk (vgl. Rasch, 1995, S. 70) wie nachfolgend aufgeführt.



**Satz 3.9 (Eigenschaften der Varianz)**

a) Steinerscher Verschiebungssatz

Für die Zufallsvariable  $X$  gilt

$$\text{Var}(X) = E(X^2) - (E(X))^2.$$

b) Für die Zufallsvariable  $Y = aX + b$ , mit  $X$  Zufallsvariable und  $a, b \in \mathbb{R}$  gilt

$$\text{Var}(aX + b) = a^2 \cdot \text{Var}(X).$$

Abschließend wird die Faltung für stochastisch unabhängige und diskrete Zufallsvariablen eingeführt, da diese im weiteren Verlauf der Arbeit benötigt wird.

**Satz 3.10 (Faltung für Wahrscheinlichkeitsfunktionen)**

Seien  $X_1, X_2$  stochastisch unabhängige diskrete Zufallsvariablen mit zugehöriger Wahrscheinlichkeitsfunktion  $f_1$  bzw.  $f_2$  und  $\mathcal{T}_1$  sowie  $\mathcal{T}_2$  gegeben. Die Wahrscheinlichkeitsfunktion der Summe  $Z = X_1 + X_2$  ist gegeben durch

$$\tilde{f}(z) = \sum_{x_1 \in \mathcal{T}_1} f_1(x_1) \cdot f_2(z - x_1)$$

(Dehling, Haupt, 2004, S. 113).

**Beweis (Satz 3.10)**

Das Ereignis  $\{X_1 + X_2 = z\}$  kann ausgedrückt werden durch die disjunkte Vereinigung der Ereignisse  $\{X_1 = x_1, X_2 = x_2\}$ . Mit  $x_1 + x_2 = z \Leftrightarrow x_2 = z - x_1$  und unter Berücksichtigung der stochastischen Unabhängigkeit folgt

$$\begin{aligned} P(X_1 + X_2 = z) &= \sum_{\substack{x_1, x_2: \\ x_1 + x_2 = z}} P(X_1 = x_1, X_2 = x_2) \\ &= \sum_{x_1 \in \mathcal{T}_1} P(X_1 = x_1, X_2 = z - x_1) \\ &= \sum_{x_1 \in \mathcal{T}_1} P(X_1 = x_1) \cdot P(X_2 = z - x_1) \\ &= \sum_{x_1 \in \mathcal{T}_1} f_1(x_1) \cdot f_2(z - x_1) \end{aligned}$$

(aus: Dehling, Haupt, 2004, S. 114).

□

### 3.1.3 Betrachtung stetiger Zufallsvariablen

Eine Zufallsvariable  $X$  nennt man stetig, wenn zu zwei Ausprägungen eines Zufallsexperiments  $a$  und  $b$  auch jeder Wert des Intervalls  $[a, b]$  ein mögliches Ergebnis des Zufallsvorgangs ist.

#### Definition 3.11 (*Wahrscheinlichkeitsdichte*)

Für eine stetige Zufallsvariable  $X$  gilt

$$P(a \leq X \leq b) = \int_a^b f(z) dz.$$

Dabei wird die Funktion  $f(x)$  als (Wahrscheinlichkeits-)Dichte der stetigen Zufallsvariablen  $X$  bezeichnet (vgl. Fahrmeir et al., 2010, S. 230) und es gilt

$$\int_{-\infty}^{+\infty} f(z) dz = 1.$$

Die Verteilungsfunktion  $F(x) = P(X \leq x)$  kann direkt aus der Definition der Wahrscheinlichkeitsdichte gefolgert werden, da gilt:  $P(X \leq x) = P(-\infty < X \leq x)$ .

#### Definition 3.12 (*Verteilungsfunktion für stetige Zufallsvariablen*)

Die Verteilungsfunktion  $F(x)$  einer stetigen Zufallsvariablen  $X$  ist definiert durch

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(z) dz$$

(vgl. Fahrmeir et al., 2010, S. 273). Dabei gilt ebenfalls für diese Funktion

$$\lim_{x \rightarrow -\infty} F(x) = 0 \text{ und } \lim_{x \rightarrow +\infty} F(x) = 1$$

(vgl. Rasch, 1995, S. 64).

Bei stetigen Zufallsvariablen ist der Wert  $x^{(p)}$  wie folgt als das  $p$ -Quantil definiert:

#### Definition 3.13 ( *$p$ -Quantil*)

Für  $0 < p < 1$  nennt man die Zahl  $x^{(p)}$  das  $p$ -Quantil, falls

$$F(x^{(p)}) = p$$

gilt und  $F(x)$  die zur stetigen Zufallsvariablen  $X$  zugehörige Verteilungsfunktion ist (vgl. Fahrmeir et al., 2010, S. 287).

Weiterhin existiert die Umkehrfunktion  $F^{-1}$  der stetigen Verteilungsfunktion  $F$  mit

$$F^{-1}(p) = x^{(p)}.$$

Der zu den Lageparametern zählende Erwartungswert  $E(X)$  wird wie folgt bestimmt.

**Definition 3.14 (Erwartungswert einer stetigen Zufallsvariablen)**

Für eine stetige Zufallsvariable  $X$  mit zugehöriger Dichte  $f(x)$  ist der Erwartungswert  $E(X)$  definiert mit

$$E(X) = \int_{-\infty}^{+\infty} z \cdot f(z) dz$$

(vgl. Fahrmeir et al., 2010, S. 283).

Die Eigenschaften für den Erwartungswert aus Satz 3.7 gelten auch für den Erwartungswert einer stetigen Zufallsvariablen.

Die Varianz, mit deren Hilfe die Streuung zweier Verteilungen verglichen werden kann, ist für stetige Zufallsvariablen wie folgt definiert.

**Definition 3.15 (Varianz einer stetigen Zufallsvariablen)**

Für eine stetige Zufallsvariable  $X$  mit zugehöriger Dichte  $f(x)$  ist die Varianz  $Var(X)$  definiert mit

$$Var(X) = \int_{-\infty}^{+\infty} (z - E(X))^2 \cdot f(z) dz$$

(vgl. Fahrmeir et al., 2010, S. 283).

Auch bei der Varianz von stetigen Zufallsvariablen gelten die selben Eigenschaften wie bei diskreten Zufallsvariablen (vgl. Satz 3.9).

## 3.2 Eindimensionale Wahrscheinlichkeitsverteilungen

Die zuvor eingeführten Zufallsvariablen folgen Wahrscheinlichkeitsverteilungen. In diesem Abschnitt werden die für den weiteren Verlauf der Bachelorarbeit benötigten (Wahrscheinlichkeits-) Verteilungen eingeführt. Dabei beschreiben die Verteilungen

ein Modell zur Prognose der Wahrscheinlichkeit, mit der die möglichen Ereignisse der Zufallsvariablen eintreten.

### 3.2.1 Bernoulli-Verteilung

Gegeben sei ein Bernoulli-Experiment, also eine Zufallsvariable  $X$  mit nur zwei möglichen Ausprägungen (Erfolg und Misserfolg). Dabei wird das Ereignis, welches als Erfolg angesehen wird, mit der Zahl 1 kodiert und das Ereignis, welches für einen Misserfolg steht, mit 0. Die Eintretenswahrscheinlichkeit für den Erfolg ist  $\pi \in (0, 1)$ . Es soll gelten:

$$P(X = 1) = \pi \text{ und } P(X = 1^C = 0) = 1 - \pi,$$

wobei  $1^C$  das Gegenereignis zu 1 darstellt. Eine diskrete Zufallsvariable  $X$  folgt einer Bernoulli-Verteilung (kurz:  $X \sim \text{Ber}(\pi)$ ), wenn die zugehörige Wahrscheinlichkeitsfunktion die Form

$$f_{\text{Ber}(\pi)}(x) = P(X = x) = \begin{cases} 1 - \pi, & \text{für } x = 0, \\ \pi, & \text{für } x = 1, \\ 0, & \text{sonst} \end{cases}$$

besitzt. Die Verteilungsfunktion einer Bernoulli-Verteilung hat die Form

$$F_{\text{Ber}(\pi)}(x) = P(X \leq x) = \begin{cases} 0, & \text{falls } x < 0, \\ 1 - \pi, & \text{falls } 0 \leq x < 1, \\ 1, & \text{falls } 1 \leq x \end{cases}$$

(vgl. Vogel, 2000, S. 122 f.).

#### Satz 3.16 (*Erwartungswert und Varianz*)

Für  $X \sim \text{Ber}(\pi)$  gilt:

- a)  $E(X) = \pi$
- b)  $\text{Var}(X) = \pi(1 - \pi)$

#### Beweis (**Satz 3.16**)

Für  $X \sim \text{Ber}(\pi)$  gilt:

- a) Mit Definition 3.6 folgt direkt

$$E(X) = \sum_{x \in \{0,1\}} x \cdot f(x) = 0 \cdot (1 - \pi) + 1 \cdot \pi = \pi$$

b) Mit Definition 3.8 folgt direkt

$$\begin{aligned} \text{Var}(X) &= \sum_{x \in \{0,1\}} (x - \pi)^2 \cdot f(x) \\ &= (-\pi)^2 \cdot (1 - \pi) + (1 - \pi)^2 \cdot \pi \\ &= \pi^2 - \pi^3 + \pi - 2\pi^2 + \pi^3 = \pi \cdot (1 - \pi) \end{aligned} \quad \square$$

### 3.2.2 Binomialverteilung

Falls  $n$  identische Bernoulli-Experimente nacheinander durchgeführt werden, dann ist die Summe der Bernoulli-verteilten Zufallsvariablen  $X = \sum_{i=1}^n Y_i$  mit  $Y_1, \dots, Y_n$  uiv. wie  $Y_1 \sim \text{Ber}(\pi)$  binomialverteilt (kurz:  $X \sim \text{Bin}(n, \pi)$ ). Folglich repräsentiert diese Verteilung die Anzahl an Erfolgen von  $n$  unabhängigen Bernoulli-Experimenten. Die Binomialverteilung besitzt nach Rasch die Wahrscheinlichkeitsfunktion der Form

$$f_{\text{Bin}(n,\pi)}(x) = P(X = x) = \begin{cases} \binom{n}{x} \pi^x (1 - \pi)^{n-x}, & \text{für } x \in \{0, 1, 2, \dots, n\}, \\ 0, & \text{sonst} \end{cases}$$

(vgl. Rasch, 1995, S. 81). Die Verteilungsfunktion ist definiert als

$$F_{\text{Bin}(n,\pi)}(x) = P(X \leq x) = \sum_{z=0}^x \binom{n}{z} \pi^z (1 - \pi)^{n-z}.$$

Dabei wird

$$\binom{n}{k} = \frac{n!}{k! \cdot (n - k)!}$$

als Binomialkoeffizient bezeichnet. Für die folgende Arbeit wichtige Eigenschaften sind in Satz 3.17 zusammengetragen.

#### Satz 3.17

a) Nach Rasch gilt für  $n, k \in \mathbb{N}_0$  und  $k \leq n$

$$\binom{n}{k} = \binom{n}{n - k}$$

(vgl. Rasch, 1995, S. 14).

b) Des Weiteren gilt für alle  $n \in \mathbb{N}_0$  und ein beliebiges, aber festes  $m \in \mathbb{N}_0$

$$\sum_{i=0}^n \binom{i + m}{m} = \binom{n + m + 1}{m + 1}.$$

**Beweis (Satz 3.17)**

a) Man kann sich leicht überlegen, dass die Aussage aus der Symmetrie des Pascal'schen Dreiecks folgt.

Mathematischer Beweis für  $n, k \in \mathbb{N}_0$  und  $k \leq n$ :

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{(n-k)! \cdot k!} = \frac{n!}{(n-k)! \cdot (n-n+k)!} \\ &= \frac{n!}{(n-k)! \cdot (n-(n-k))!} = \binom{n}{n-k} \end{aligned}$$

b) Die Beweisführung erfolgt durch einen Induktionsbeweis.

IA:  $n = 0$  und  $m$  beliebig, aber fest.

$$\sum_{i=0}^0 \binom{i+m}{m} = \binom{m}{m} = 1 = \binom{m+1}{m+1} = \binom{n+m+1}{m+1}$$

IV: Für alle  $n \in \mathbb{N}_0$  und ein beliebiges, aber festes  $m \in \mathbb{N}_0$  folgt:

$$\sum_{i=0}^n \binom{i+m}{m} = \binom{n+m+1}{m+1}$$

IS: Es folgt für  $n+1 \in \mathbb{N}$  und  $m \in \mathbb{N}_0$  beliebig, aber fest:

$$\begin{aligned} \sum_{i=0}^{n+1} \binom{i+m}{m} &= \sum_{i=0}^n \binom{i+m}{m} + \binom{n+1+m}{m} \\ &= \binom{n+m+1}{m+1} + \binom{n+m+1}{m} \\ &= \frac{(n+m+1)!}{(m+1)! \cdot n!} + \frac{(n+m+1)!}{m! \cdot (n+1)!} \\ &= \frac{(n+m+1)! \cdot (n+1) + (n+m+1)! \cdot (m+1)}{(m+1)! \cdot (n+1)!} \\ &= \frac{(n+m+1)! \cdot (n+1+m+1)}{(m+1)! \cdot (n+1)!} \\ &= \frac{(n+1+m+1)!}{(m+1)! \cdot (n+1)!} \\ &= \binom{n+1+m+1}{m+1} \end{aligned}$$

□

**Satz 3.18 (Erwartungswert und Varianz)**

Für  $X \sim \text{Bin}(n, \pi)$  gilt:

a)  $E(X) = n \cdot \pi$

b)  $\text{Var}(X) = n \cdot \pi \cdot (1 - \pi)$

**Beweis (Satz 3.18)**

Für  $X \sim \text{Bin}(n, \pi)$  und mit  $Y_1, \dots, Y_n$  uiv. wie  $Y_1 \sim \text{Ber}(\pi)$  gilt:

a) Mit Definition 3.7 folgt direkt

$$E(X) = E\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n E(Y_i) = \sum_{i=1}^n \pi = n \cdot \pi$$

b) Beweisführung analog. □

**3.2.3 Geometrische Verteilung**

Gegeben ist ein Bernoulli-Experiment mit  $P(X = 1) = \pi$  und  $P(X = 1^C) = 1 - \pi$  mit  $\pi \in (0, 1)$ . Es interessiert nun, wie viele Fehlversuche nötig sind, bis zum ersten Mal Ereignis 1 eintritt. Dieses Ereignis 1 wird im Bereich der Qualitätssicherung als Ausfall einer untersuchten Einheit gesetzt und kann mit einer geometrisch verteilten Zufallsvariablen  $X$  (kurz:  $X \sim \text{Geo}(\pi)$ ) modelliert werden. Somit bildet die im Folgenden eingeführte geometrische Verteilung eine Lebensdauerverteilung, indem die abgeschlossenen Zeiteinheiten bis zum Eintritt des Ausfalls gezählt werden. Die Wahrscheinlichkeitsfunktion der Zufallsvariablen  $X$  ist dann definiert durch

$$f_{\text{Geo}(\pi)}(x) = P(X = x) = \begin{cases} (1 - \pi)^x \cdot \pi, & \text{für } x \in \mathbb{N}_0, \\ 0, & \text{sonst.} \end{cases}$$

Die Verteilungsfunktion der diskreten Zufallsvariablen besitzt für  $x \in \mathbb{N}_0$  die Form

$$F_{\text{Geo}(\pi)}(x) = P(X \leq x) = 1 - (1 - \pi)^{x+1}$$

(vgl. Vogel, 2000, S. 156 f.).

**Satz 3.19 (Erwartungswert und Varianz)**

Für  $X \sim \text{Geo}(\pi)$  gilt:

a)  $E(X) = \frac{1-\pi}{\pi}$

b)  $\text{Var}(X) = \frac{1-\pi}{\pi^2}$

**Beweis (Satz 3.19)**

Für  $X \sim \text{Geo}(\pi)$  gilt:

a) Mit Definition 3.6 und  $\sum_{z=0}^{\infty} z \cdot s^z = \frac{s}{(1-s)^2}$  für  $1 < s < 1$  folgt direkt:

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \cdot (1 - \pi)^x \pi = \pi \sum_{x=0}^{\infty} x \cdot (1 - \pi)^x \\ &= \pi \cdot \frac{1 - \pi}{(1 - (1 - \pi))^2} = \frac{\pi(1 - \pi)}{\pi^2} = \frac{1 - \pi}{\pi} \end{aligned}$$

b) Mit  $\sum_{z=0}^{\infty} z^2 \cdot s^z = \frac{s(1+s)}{(1-s)^3}$  für  $1 < s < 1$  gilt

$$\begin{aligned} E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot (1-\pi)^x \pi = \pi \sum_{x=0}^{\infty} x^2 \cdot (1-\pi)^x \\ &= \pi \cdot \frac{(1-\pi)(1+1-\pi)}{(1-1+\pi)^3} = \frac{\pi(1-\pi)(2-\pi)}{\pi^3} = \frac{(1-\pi)(2-\pi)}{\pi^2}. \end{aligned}$$

Mit Satz 3.9 a) folgt

$$\begin{aligned} \text{Var}(X) &= E(X^2) - (E(X))^2 = \frac{(1-\pi)(2-\pi)}{\pi^2} - \frac{(1-\pi)^2}{\pi^2} \\ &= \frac{(1-\pi)(2-\pi) - (1-\pi)^2}{\pi^2} = \frac{(1-\pi)(2-\pi-1+\pi)}{\pi^2} = \frac{1-\pi}{\pi^2}. \end{aligned}$$

□

### 3.2.4 Negative Binomialverteilung

Gegeben ist erneut ein Bernoulli-Experiment mit  $P(X = 1) = \pi$  und  $P(X = 1^c) = 1 - \pi$  mit  $\pi \in (0, 1)$ . Es interessiert nun, wie viele Fehlversuche vor dem  $n$ -ten Eintreten des Ereignisses 1 auftreten. Dabei soll gelten:  $n \in \mathbb{N}$ . Dieses Szenario kann mit einer negativ binomialverteilten Zufallsvariable  $X$  (kurz:  $X \sim \text{negBin}(\pi, n)$ ) modelliert werden. Die zugehörige Wahrscheinlichkeitsfunktion besitzt dann die Form

$$f_{\text{negBin}(\pi, n)}(x) = P(X = x) = \begin{cases} \binom{x+n-1}{x} \cdot \pi^n \cdot (1-\pi)^x, & \text{für } x \in \mathbb{N}_0, \\ 0, & \text{sonst} \end{cases}$$

(vgl. Vogel, 2000, S. 158).

Die Verteilungsfunktion hat die Form

$$F_{\text{negBin}(\pi, n)}(x) = P(X \leq x) = \sum_{z=0}^x \binom{z+n-1}{z} \cdot \pi^n \cdot (1-\pi)^z.$$

Es ist anzumerken, dass diverse Zusammenhänge zwischen der geometrischen Verteilung und der negativen Binomialverteilung existieren, auf die im weiteren Verlauf der Arbeit zurückgegriffen wird.

#### Satz 3.20

- a) Es gilt  $f_{\text{negBin}(\pi, 1)} = f_{\text{Geo}(\pi)}$  (vgl. Vogel, 2000, S. 158).  
 b) Außerdem gilt für unabhängig identisch verteilte Zufallsvariablen  $X_1, \dots, X_n \sim \text{Geo}(\pi)$

$$\sum_{i=1}^n X_i \sim \text{negBin}(\pi, n).$$



**Beweis (Satz 3.20)**

- a) Wählt man den Parameter  $n$  der negativen Binomialverteilung  $n = 1$ , dann ist leicht zu erkennen, dass die Wahrscheinlichkeitsfunktion der negativen Binomialverteilung bei jeweils gleichem Parameter  $\pi$  der Wahrscheinlichkeitsfunktion der geometrischen Verteilung entspricht

$$f_{negBin(\pi,1)} = \left\{ \begin{array}{ll} \binom{x}{x} \cdot \pi \cdot (1 - \pi)^x, & \text{für } x \in \mathbb{N}_0, \\ 0, & \text{sonst} \end{array} \right\} = f_{Geo(\pi)}.$$

- b) Die Beweisführung erfolgt durch einen Induktionsbeweis.

IA:  $n = 1$

Sei  $X_1 \sim Geo(\pi)$  mit Wahrscheinlichkeitsfunktionen  $f_{Geo(\pi)}$  sowie mit  $\mathcal{T}_1 = \mathbb{N}_0$ . Mit Satz 3.20 a) folgt direkt:

$$\sum_{i=1}^1 X_i = X_1 \sim negBin(\pi, 1).$$

IV:  $\sum_{i=1}^n X_i \sim negBin(\pi, n)$  für alle  $n \in \mathbb{N}$ .

IS:  $Z = \sum_{i=1}^{n+1} X_i = \sum_{i=1}^n X_i + X_{n+1}$

Nach Definition 3.10 folgt:

$$\begin{aligned} \tilde{f}(z) &= \sum_{x \in \mathbb{N}_0} f_{negBin(\pi,n)}(x) \cdot f_{Geo(\pi)}(z - x) \\ &= \sum_{x=0}^z f_{negBin(\pi,n)}(x) \cdot f_{Geo(\pi)}(z - x) \\ &= \sum_{x=0}^z \binom{x+n-1}{x} (1 - \pi)^x \cdot \pi^n \cdot (1 - \pi)^{z-x} \cdot \pi \\ &= \pi^{n+1} \cdot (1 - \pi)^z \sum_{x=0}^z \binom{x+n-1}{x} \end{aligned}$$

Für die Summe über den Binomialkoeffizienten gilt mit Satz 3.17:

$$\begin{aligned} \sum_{x=0}^z \binom{x+n-1}{x} &\stackrel{3.17a)}{=} \sum_{x=0}^z \binom{x+n-1}{n-1} \\ &\stackrel{3.17b)}{=} \binom{z+(n-1)+1}{(n-1)+1} = \binom{z+n}{n} \\ &\stackrel{3.17a)}{=} \binom{z+n}{z} \end{aligned}$$

Eingesetzt in die obige Gleichung folgt:

$$\begin{aligned}\tilde{f}(z) &= \pi^{n+1} \cdot (1 - \pi)^z \sum_{x=0}^z \binom{x+n-1}{x} \\ &= \binom{z+n}{z} \cdot \pi^{n+1} \cdot (1 - \pi)^z \\ &= \binom{z+(n+1)-1}{z} \cdot \pi^{n+1} \cdot (1 - \pi)^z \\ &= f_{negBin(\pi, n+1)}(z)\end{aligned}$$

Insgesamt folgt also  $Z = \sum_{i=1}^{n+1} X_i \sim negBin(\pi, n+1) \forall n \in \mathbb{N}$ .  $\square$

### Satz 3.21 (Erwartungswert und Varianz)

Für  $X \sim negBin(\pi, n)$  gilt:

- a)  $E(X) = \frac{n(1-\pi)}{\pi}$
- b)  $Var(X) = \frac{n(1-\pi)}{\pi^2}$

#### Beweis (Satz 3.21)

Für  $X \sim negBin(\pi, n)$  gilt:

- a) Mit Satz 3.20 folgt  $X = \sum_{i=1}^n Y_i \sim negBin(\pi, n)$  mit  $Y_1, \dots, Y_n$  uiv. wie  $Y_1 \sim Geo(\pi)$ .

Es gilt also mit Satz 3.19 und Satz 3.7:

$$\begin{aligned}E(X) &= E\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n E(Y_i) \\ &= n \cdot E(Y_1) = n \cdot \frac{1-\pi}{\pi} = \frac{n(1-\pi)}{\pi}\end{aligned}$$

- b) Beweisführung analog.  $\square$

### 3.2.5 Poisson-Verteilung

Die im Folgenden dargestellte Poisson-Verteilung ist in unterschiedlichen Bereichen einsetzbar. Mit ihr lassen sich unter anderem Eintretenswahrscheinlichkeiten für unabhängige Ereignisse modellieren, die in einer bestimmten Zeitspanne eintreten. Andererseits ist sie eine gängige Möglichkeit, Lebensdauern zu modellieren. Falls die diskrete Zufallsvariable  $X$  Poisson-verteilt ist (kurz:  $X \sim Poi(\lambda)$ ), dann hat die Wahrscheinlichkeitsfunktion die Form

$$f_{Poi(\lambda)}(x) = P(X = x) = \begin{cases} \frac{\lambda^x \cdot e^{-\lambda}}{x!}, & \text{für } x \in \mathbb{N}_0 \text{ und } \lambda > 0, \\ 0, & \text{sonst.} \end{cases}$$

Die zugehörige Verteilungsfunktion hat die Form

$$F_{Poi(\lambda)}(x) = P(X \leq x) = \begin{cases} 0, & \text{falls } x < 0, \\ \sum_{z \leq x} \frac{\lambda^z \cdot e^{-\lambda}}{z!}, & \text{falls } x \geq 0 \end{cases}$$

(vgl. Vogel, 2000, S. 159 f.).

**Satz 3.22 (Erwartungswert und Varianz)**

Für  $X \sim Poi(\lambda)$  gilt:

- a)  $E(X) = \lambda$
- b)  $Var(X) = \lambda$

**Beweis (Satz 3.22)**

Für  $X \sim Poi(\lambda)$  gilt:

- a) Es gilt also mit Definition 3.6 und mit  $\sum_{z=0}^{+\infty} \frac{\lambda^z}{z!}$ :

$$\begin{aligned} E(X) &= \sum_{x=0}^{+\infty} x \cdot \frac{\lambda^x \cdot e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=1}^{+\infty} \frac{\lambda^x}{(x-1)!} = \lambda \cdot e^{-\lambda} \sum_{x=1}^{+\infty} \frac{\lambda^{x-1}}{(x-1)!} \\ &= \lambda \cdot e^{-\lambda} \sum_{y=0}^{+\infty} \frac{\lambda^y}{y!} = \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda \end{aligned}$$

- b) Mit Satz 3.9 und mit  $\sum_{z=0}^{+\infty} \frac{\lambda^z}{z!}$  folgt:

$$Var(X) = E(X^2) - (E(X))^2$$

Es wird nun  $E(X^2)$  bestimmt:

$$\begin{aligned} E(X^2) &= \sum_{x=0}^{+\infty} x^2 \cdot \frac{\lambda^x \cdot e^{-\lambda}}{x!} \\ &= \sum_{x=0}^{+\infty} x(x-1) \cdot \frac{\lambda^x \cdot e^{-\lambda}}{x!} + \sum_{x=0}^{+\infty} x \cdot \frac{\lambda^x \cdot e^{-\lambda}}{x!} \\ &= \lambda^2 \cdot e^{-\lambda} \sum_{x=2}^{+\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda \\ &= \lambda^2 \cdot e^{-\lambda} \cdot e^{\lambda} + \lambda = \lambda^2 + \lambda \end{aligned}$$

Mit Einsetzen von  $E(X^2) = \lambda^2 + \lambda$  und  $E(X) = \lambda$  folgt schließlich:

$$Var(X) = E(X^2) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \quad \square$$

Wenn unabhängig und identisch Poisson-verteilte Zufallsvariablen aufsummiert werden, dann folgt diese Summe ebenfalls einer Poisson-Verteilung (vgl. Vogel, 2000, S. 162).

**Satz 3.23 (Additionssatz für Poisson-verteilte Zufallsvariablen)**

Für  $X_1, \dots, X_n$  unabhängig identisch verteilt wie  $X_1 \sim Poi(\lambda)$  gilt

$$\sum_{i=1}^n X_i \sim Poi(n\lambda)$$

für alle  $n \in \mathbb{N}$ .

**Beweis (Satz 3.23)**

Die Beweisführung erfolgt durch einen Induktionsbeweis.

IA:  $n = 1$

Der Nachweis ist trivial. Sei  $X_1 \sim Poi(\lambda)$  mit Wahrscheinlichkeitsfunktion  $f_{Poi(\lambda)}$  sowie mit  $\mathcal{T} = \mathbb{N}_0$ . Es gilt

$$\sum_{i=1}^1 X_i = X_1 \sim Poi(1\lambda).$$

IV:  $\sum_{i=1}^n X_i \sim Poi(n\lambda)$  für alle  $n \in \mathbb{N}$ .

IS:  $Z = \sum_{i=1}^{n+1} X_i = \sum_{i=1}^n X_i + X_{n+1}$ .

Nach Definition 3.10 folgt:

$$\begin{aligned} \tilde{f}(z) &= \sum_{x \in \mathbb{N}_0} f_{Poi(n\lambda)}(x) \cdot f_{Poi(\lambda)}(z-x) \\ &= \sum_{x=0}^z \frac{(n\lambda)^x \cdot e^{-n\lambda}}{x!} \cdot \frac{\lambda^{z-x} \cdot e^{-\lambda}}{(z-x)!} \\ &= \sum_{x=0}^z \frac{(n\lambda)^x \cdot \lambda^{z-x} \cdot e^{-(n+1)\lambda}}{x! \cdot (z-x)!} \cdot \frac{z!}{z!} \\ &= \frac{e^{-(n+1)\lambda}}{z!} \sum_{x=0}^z \binom{z}{x} \cdot (n\lambda)^x \cdot \lambda^{z-x} \cdot \frac{((n+1) \cdot \lambda)^z}{((n+1) \cdot \lambda)^z} \\ &= \frac{((n+1) \cdot \lambda)^z \cdot e^{-(n+1)\lambda}}{z!} \sum_{x=0}^z \binom{z}{x} \frac{(n\lambda)^x}{((n+1) \cdot \lambda)^x} \cdot \frac{(n\lambda)^{z-x}}{((n+1) \cdot \lambda)^{z-x}} \\ &= f_{Poi((n+1)\lambda)}(z) \sum_{x=0}^z f_{Bin(z, \frac{n}{n+1})}(x) \\ &= f_{Poi((n+1)\lambda)}(z) \end{aligned}$$

Insgesamt folgt also  $Z = \sum_{i=1}^{n+1} X_i \sim Poi((n+1) \cdot \lambda)$  für alle  $n \in \mathbb{N}$ .  $\square$

### 3.2.6 Normalverteilung

Die vermutlich am häufigsten verwendete Verteilung ist die Normalverteilung. Diese Verteilung modelliert eingipflige und symmetrische Häufigkeitsverteilungen metrischer Merkmale. In der statistischen Schätz- und Testtheorie findet sie häufig Anwendung, da viele Verteilungen approximativ normalverteilt sind. Eine normalverteilte Zufallsvariable  $X$  (kurz:  $X \sim \mathcal{N}(\mu, \sigma^2)$ ) besitzt die Dichtefunktion

$$f_{\mathcal{N}(\mu, \sigma^2)}(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \text{ für alle } x \in \mathbb{R}.$$

Dabei ist  $\mu \in \mathbb{R}$  und es soll gelten  $\sigma^2 > 0$ . Die Verteilungsfunktion besitzt die Form

$$F_{\mathcal{N}(\mu, \sigma^2)}(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz$$

(vgl. Vogel, 2000, S. 126 f.).

Die Parameter geben Auskunft über die Art der Verteilung. Wird  $\mu = 0$  und  $\sigma^2 = 1$  gewählt, dann erhält man einen Spezialfall der Normalverteilung. Eine Zufallsvariable heißt standardnormalverteilt, wenn sie der Normalverteilung mit  $\mu = 0$  und  $\sigma^2 = 1$  folgt (kurz:  $X \sim \mathcal{N}(0, 1)$ ) (aus Vogel, 2000, S. 127 f.).

#### Satz 3.24 (Erwartungswert und Varianz)

Für  $X \sim \mathcal{N}(\mu, \sigma^2)$  gilt:

- a)  $E(X) = \mu$
- b)  $Var(X) = \sigma^2$

(vgl. Vogel, 2000, S. 126).

### 3.2.7 $\chi^2$ -Verteilung

Sind die unabhängig und identisch verteilten Zufallsvariablen  $Y_1, \dots, Y_n$  gegeben und folgen diese einer Standardnormalverteilung, dann ist die Summe von  $n$  quadrierten standardnormalverteilten Zufallsvariablen  $\chi^2$ -verteilt mit  $n$  Freiheitsgraden, das heißt

$$X = \sum_{i=1}^n Y_i^2 \sim \chi_n^2.$$

Die zugehörige Dichtefunktion hat die Form

$$f_{\chi_n^2}(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \cdot \Gamma(\frac{n}{2})} \cdot x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}}, & \text{für } x > 0, \\ 0, & \text{sonst.} \end{cases}$$

Dabei steht  $\Gamma(x)$  für den Wert der Gamma-Funktion an der Stelle  $x$  (vgl. Vogel, 2000, S. 140).

Die Verteilungsfunktion hat die Form

$$F_{\chi_n^2}(x) = P(X \leq x) = \int_{-\infty}^x f_{\chi_n^2}(z) dz.$$

### Satz 3.25 (Erwartungswert und Varianz)

Für  $X \sim \chi_n^2$  gilt:

a)  $E(X) = n$

b)  $Var(X) = 2n$

(vgl. Vogel, 2000, S. 140).

Die  $\chi^2$ -Verteilung ist eine weitere Verteilung, die im Zusammenhang mit einer anderen Verteilung steht: Die Werte der Verteilungsfunktion der Poisson-Verteilung können mit Hilfe der Verteilungsfunktion der  $\chi^2$ -Verteilung dargestellt werden (vgl. Rasch, 1995, S. 86).

### Satz 3.26

Für  $X \sim Poi(\lambda)$  und  $Y \sim \chi_{2(n+1)}^2$  gilt  $\forall n \in \mathbb{N}_0$

$$F_{\chi_{2(n+1)}^2}(2\lambda) = \int_{-\infty}^{2\lambda} \frac{1}{2^{n+1} \Gamma(n+1)} z^n e^{-\frac{z}{2}} dz = 1 - \sum_{z=0}^n \frac{\lambda^z e^{-\lambda}}{z!} = 1 - F_{Poi(\lambda)}(n).$$

### Beweis (Satz 3.26)

Die Beweisführung erfolgt durch einen Induktionsbeweis.

IA: Mit  $n = 0$  folgt

$$\begin{aligned} F_{\chi_2^2}(2\lambda) &= \int_{-\infty}^{2\lambda} \frac{1}{2 \cdot \Gamma(1)} e^{-\frac{z}{2}} dz = \int_0^{2\lambda} \frac{1}{2} e^{-\frac{z}{2}} dz = [-e^{-\frac{z}{2}}]_0^{2\lambda} \\ &= -e^{-\lambda} + e^0 = 1 - \frac{\lambda^0 e^{-\lambda}}{0!} = 1 - \sum_{z=0}^0 \frac{\lambda^z e^{-\lambda}}{z!} \\ &= 1 - F_{Poi(\lambda)}(0) \end{aligned}$$

IV: Es gilt für alle  $n \in \mathbb{N}_0$

$$\begin{aligned} F_{\chi_{2(n+1)}^2}(2\lambda) &= \int_{-\infty}^{2\lambda} \frac{1}{2^{n+1} \cdot \Gamma(n+1)} z^n e^{-\frac{z}{2}} dz \\ &= 1 - \sum_{z=0}^n \frac{\lambda^z e^{-\lambda}}{z!} = 1 - F_{Poi(\lambda)}(n) \end{aligned}$$

IS: Mit  $n + 1$  folgt:

$$\begin{aligned} F_{\chi^2_{2(n+2)}}(2\lambda) &= \int_{-\infty}^{2\lambda} \frac{1}{2^{n+2} \cdot \Gamma(n+2)} z^{n+1} e^{-\frac{z}{2}} dz \\ &= \int_0^{2\lambda} \frac{1}{2^{n+1}(n+1)!} z^{n+1} \cdot \frac{1}{2} e^{-\frac{z}{2}} dz \end{aligned}$$

mit partieller Integration folgt:

$$\begin{aligned} &= \left[ \frac{z^{n+1}}{2^{n+1}(n+1)!} (-e^{-\frac{z}{2}}) \right]_0^{2\lambda} - \int_0^{2\lambda} \frac{(n+1)z^n}{2^{n+1}(n+1)!} (-e^{-\frac{z}{2}}) dz \\ &= \frac{(2\lambda)^{n+1}}{2^{n+1}(n+1)!} (-e^{-\lambda}) + \int_0^{2\lambda} \frac{z^n}{2^{n+1} \cdot \Gamma(n+1)} e^{-\frac{z}{2}} dz \end{aligned}$$

mit IV folgt:

$$\begin{aligned} &= \frac{\lambda^{n+1} (-e^{-\lambda})}{(n+1)!} + 1 - \sum_{z=0}^n \frac{\lambda^z e^{-\lambda}}{z!} \\ &= 1 - \sum_{z=0}^{n+1} \frac{\lambda^z e^{-\lambda}}{z!} = 1 - F_{Poi(\lambda)}(n+1) \end{aligned}$$

□

### 3.3 Zählende Abnahmeprüfung

Die theoretische Grundlage für die zählende Abnahmeprüfung liefert die statistische Testtheorie. Daher wird zunächst eine Einführung in die besonderen Begrifflichkeiten dieses Gebietes gegeben und anschließend das Vorgehen bei der Durchführung einer zählenden Abnahmeprüfung vorgestellt. Schließlich werden die für den weiteren Verlauf dieser Arbeit benötigten zählenden Abnahmeprüfungen konkret für ein Poisson-verteiltes Los und für ein geometrisch verteiltes Los hergeleitet.

#### 3.3.1 Einführung in die statistische Testtheorie

Die statistische Testtheorie ist ein Teilbereich der Statistik, in dem Aussagen als sogenannte Hypothesen formuliert und anschließend überprüft werden. Dabei wird konkret unter einer Hypothese eine Behauptung über einen Aspekt von einer oder mehreren Zufallsvariablen verstanden. Ein Verfahren zur Überprüfung dieser Hypothesen ist der Hypothesentest. Dieser greift auf Realisationen der betrachteten Zufallsvariablen zurück und ist somit zufallsabhängig. Aus diesem Grund kann ein Test niemals verifizieren oder falsifizieren, da in jedem Fall eine Fehlentscheidung möglich ist (vgl. Rinne, Mittag, 1995, S. 112).

Bei den Hypothesentests wird nach Rinne und Mittag zwischen Parametertests und Anpassungstests unterschieden (vgl. Rinne, Mittag, 1995, S. 112). Bei einem Parametertest beinhaltet die Hypothese Informationen über den Parameter der Verteilung der Zufallsvariablen der Zufallsstichprobe. In Anpassungstests wird dagegen geprüft, ob die betrachteten Zufallsvariablen mit unbekannter identischer Verteilung einer bestimmten Verteilung folgen. Im Folgenden wird die Betrachtung von Anpassungstests vernachlässigt, da diese für die Lösung der in dieser Bachelorarbeit beschriebenen Probleme nicht von Interesse sind.

Ein statistisches Testproblem eines Parametertests besteht aus einer Nullhypothese  $H_0$  und einer Alternativhypothese  $H_1$ , welche sich gegenseitig ausschließen. Die Hypothesen beinhalten den unbekannt Parameter  $\theta$  der Verteilung der zu untersuchenden Zufallsstichprobe (vgl. Mosler, Schmid, 2008, S. 238).

**Definition 3.27 (Testproblem)**

Es gelte  $\Theta_0, \Theta_1 \subset \Theta$  und  $\Theta_1 = \Theta_0^c$ . Die Testhypothesen lauten

$$H_0 : \theta \in \Theta_0 \text{ gegen } H_1 : \theta \in \Theta_1.$$

Sei weiterhin  $\Theta \subseteq \mathbb{R}$  und  $\theta_0 \in \Theta$ . Falls die Hypothesen formuliert sind als

$$H_0 : \theta = \theta_0 \text{ gegen } H_1 : \theta \neq \theta_0,$$

nennt man dies einen zweiseitigen Test. Falls man

$$H_0 : \theta \leq \theta_0 \text{ gegen } H_1 : \theta > \theta_0 \text{ oder } H_0 : \theta \geq \theta_0 \text{ gegen } H_1 : \theta < \theta_0$$

testet, bezeichnet man dies als ein einseitiges Testproblem.

Ein statistischer Test bewertet die Gültigkeit der Hypothesen durch die Anwendung einer Entscheidungsregel. Mit Hilfe einer auf der Stichprobe basierenden Prüfgröße kann entschieden werden, ob eher  $H_0$  oder  $H_1$  auf die Grundgesamtheit zutrifft. Diese Prüfgröße nennt man auch Teststatistik (Fahrmeir et al., 2010, S. 415).

**Definition 3.28 (Test)**

Gegeben seien die Zufallsvariablen  $X_1, \dots, X_n$ . Um einen Hypothesentest bezüglich eines Testproblems laut Definition 3.27 durchzuführen, ist eine Teststatistik definiert durch  $S(x_1, \dots, x_n)$ . Dabei sind  $x_1, \dots, x_n$  die Realisationen der oben genannten Zufallsvariablen. Des Weiteren ist ein kri-



tischer Bereich  $K \subset \mathbb{R}$  gegeben (vgl. Mosler, Schmid, 2008, S. 239). Der statistische Test hat die Form

$$\varphi(x_1, \dots, x_n) := \begin{cases} 1, & \text{falls } S(x_1, \dots, x_n) \in K, \\ 0, & \text{falls } S(x_1, \dots, x_n) \notin K. \end{cases}$$

Falls  $\varphi(x_1, \dots, x_n) = 1$  kann die Nullhypothese abgelehnt (verworfen) werden, bei  $\varphi(x_1, \dots, x_n) = 0$  kann die Nullhypothese nicht verworfen werden.

Auf Grund der bereits angesprochenen Zufallsabhängigkeit kann ein Parametertest auch falsche Schlüsse über die Parameter der Verteilung der Zufallsvariable ziehen.

|  | $\theta \in \Theta_0$                | $\theta \in \Theta_1$               |
|--|--------------------------------------|-------------------------------------|
| $\varphi(x_1, \dots, x_n) = 0$<br>$H_0$ wird nicht verworfen | richtige<br>Entscheidung             | Fehler 2. Art<br>( $\beta$ -Fehler) |
| $\varphi(x_1, \dots, x_n) = 1$<br>$H_0$ wird verworfen       | Fehler 1. Art<br>( $\alpha$ -Fehler) | richtige<br>Entscheidung            |

**Tabelle 3.1:** Fehlerarten bei der Durchführung von Hypothesentests

In Tabelle 3.1 sind die möglichen Ausgänge eines statistischen Tests dargestellt, die bei der Durchführung von Hypothesentests auftreten können. Es wird von einem Fehler erster Art gesprochen, falls die Hypothese  $H_0$  verworfen wird, obwohl sie stimmt. Der Fehler zweiter Art bezeichnet die Situation, dass die Testentscheidung auf  $H_0$  fällt, obwohl die Alternativhypothese zutrifft. Im Allgemeinen ist es laut Mosler und Schmid nicht möglich, die beiden Fehlerwahrscheinlichkeiten  $\alpha$  und  $\beta$  simultan zu minimieren (vgl. Mosler, Schmid, 2008, S. 242). Die Fehlerwahrscheinlichkeit  $\alpha$  wird kontrolliert, indem der Fehler erster Art des Hypothesentests vorgegeben wird. Dies erreicht man, indem die Entscheidungsregel höchstens mit einer Wahrscheinlichkeit von  $\alpha$  einen Fehler erster Art begeht. Ein derart konstruierter Test wird Test zum Signifikanzniveau  $\alpha$  genannt, wobei in der Praxis meistens  $\alpha = 0.05$  gewählt wird (Fahrmeir et al., 2010, S. 416 f.). Durch das Signifikanzniveau  $\alpha$  und durch die Verteilung der Teststatistik wird der kritische Bereich  $K$  eines Tests aus Definition 3.28 bestimmt.

Zu jedem Test kann eine Gütefunktion erstellt werden. Diese bewertet die Güte des Tests anhand der Testentscheidung, indem die Wahrscheinlichkeit angegeben wird, bei dem Parameter  $\theta \in \Theta$  die betrachtende Nullhypothese abzulehnen (vgl. Mosler,

Schmid, 2008, S. 240). In diesem Zusammenhang bedeutet Güte einen kleinen Fehler zweiter Art.

**Definition 3.29 (Gütefunktion)**

Sei ein Test laut Definition 3.28 gegeben zu den Zufallsvariablen  $X_1, \dots, X_n$ . Des Weiteren sei  $P(\varphi(X_1, \dots, X_n) = 1 \mid \theta)$  die Wahrscheinlichkeit, dass die Teststatistik im kritischen Bereich liegt, wenn  $\theta \in \Theta$  der wahre Parameter ist. Die Funktion  $G : \Theta \rightarrow [0, 1]$  mit

$$G(\theta) := P(\varphi(X_1, \dots, X_n) = 1 \mid \theta)$$

heißt Gütefunktion des Tests.

In engem Zusammenhang mit der Gütefunktion steht die Operations-Charakteristik-Funktion (kurz: OC-Funktion). Diese gibt die Wahrscheinlichkeit an, dass die Nullhypothese bei dem Parameter  $\theta \in \Theta$  nicht abgelehnt wird und stellt somit die Gegenwahrscheinlichkeit zur Gütefunktion dar (vgl. Mosler, Schmid, 2008, S. 240).

**Definition 3.30 (OC-Funktion)**

Es gelten dieselben Voraussetzungen wie in Definition 3.29. Die OC-Funktion  $L : \Theta \rightarrow [0, 1]$  zum Test aus Definition 3.28 ist definiert durch

$$L(\theta) := 1 - G(\theta), \text{ für } \theta \in \Theta.$$

Um einen Test zum Signifikanzniveau  $\alpha$  zu erhalten, soll, wie bereits angesprochen, die Fehlerwahrscheinlichkeit erster Art auf  $\alpha$  minimiert werden. Daher muss

$$\alpha = \sup_{\theta \in \Theta_0} G(\theta) = 1 - \inf_{\theta \in \Theta_0} L(\theta)$$

gelten. Insofern ist  $\alpha$  die maximale Wahrscheinlichkeit für den Fehler erster Art. Des Weiteren kann mittels

$$\beta = 1 - \inf_{\theta \in \Theta_1} G(\theta) = \sup_{\theta \in \Theta_1} L(\theta)$$

die maximale Wahrscheinlichkeit, einen Fehler zweiter Art zu begehen, bestimmt werden (vgl. Mosler, Schmid, 2008, S. 240 f.).

### 3.3.2 Allgemeine Durchführung der zählenden Abnahmeprüfung

Eine Abnahmeprüfung dient zur Entscheidungsfindung, ob das Prüfling des Produzenten vom Konsumenten angenommen (gekauft) werden sollte oder nicht. Sie wird

vom Konsumenten durchgeführt. Aus dem Los mit einem Losumfang von  $N$  wird eine Zufallsstichprobe von  $n$  Einheiten gezogen. Dabei wird im Weiteren angenommen, dass diese Loseinheiten aus derselben Produktionslinie stammen und zufällig zusammengestellt werden. Auf Grund dieser Annahmen können die Einheiten im Los als unabhängig und identisch verteilt angesehen werden. Mit Hilfe der Stichprobe wird entschieden, ob das Los angenommen wird oder nicht. Dies geschieht, indem die einzelnen Objekte auf eigene Qualitätsansprüche geprüft werden.

Für die weitere Arbeit wird der Aspekt der Zuverlässigkeit als Qualitätsanspruch betrachtet, also eine Zuverlässigkeitsprüfung durchgeführt. Das bedeutet, dass die Einheiten der Stichprobe auf ihre Mindestlebensdauer hin überprüft werden. Diesen Aspekt betrachtet die Zuverlässigkeitsfunktion  $R(t)$  (reliability function), indem die Wahrscheinlichkeit angegeben wird, dass eine Einheit länger als  $t$  abgeschlossene Zeiteinheiten funktionsfähig ist:  $R(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t)$  (vgl. Meyna, Pauli, 2010, S. 33). Dabei ist die Verteilungsfunktion als Ausfallwahrscheinlichkeit der Einheiten nach  $t$  abgeschlossenen Zeiteinheiten zu verstehen.

**Definition 3.31 (Zuverlässigkeitsfunktion)**

Die Zuverlässigkeitsfunktion  $R : [0, \infty) \rightarrow [0, 1]$  einer Zufallsvariablen  $T \sim F(t)$  ist definiert durch

$$R(t) := 1 - F(t).$$

Oftmals enthalten produzierte Mengen einige Einheiten, die nicht funktionsfähig sind. Diese Einheiten haben  $t = 0$  abgeschlossene Zeiteinheiten funktioniert und folglich gilt für diese Einheiten  $R(0) \leq 1$ . Weiterhin gilt  $\lim_{t \rightarrow +\infty} R(t) = 1 - \lim_{t \rightarrow +\infty} F(t) = 1 - 1 = 0$ . Die Zuverlässigkeitsfunktion  $R(t)$  ist eine monoton fallende Funktion (vgl. Meyna, Pauli, 2010, S. 33).

Die Durchführung einer Abnahmeprüfung erfolgt durch die Anwendung eines statistischen Tests aus Definition 3.28. Es soll geprüft werden, ob der Ausschussanteil des Loses  $p$  höchstens genauso groß ist, wie der maximal akzeptierte Ausschussanteil  $p_0$ . Die zugehörige Nullhypothese lautet  $H_0 : p \leq p_0$ . Dabei kann  $p$  als der Anteil der Einheiten aufgefasst werden, welcher bis zum Zeitpunkt  $t_0$  ausgefallen ist ( $p = P(T \leq t_0)$ ). Dieser Überlegung folgt die Umformung der Nullhypothese:

$$\begin{aligned} p \leq p_0 &\Leftrightarrow P(T \leq t_0) \leq p_0 \Leftrightarrow F(t_0) \leq p_0 \\ &\Leftrightarrow 1 - F(t_0) \geq 1 - p_0 \\ &\Leftrightarrow R(t_0) \geq 1 - p_0. \end{aligned}$$

Nach der Umformung beinhaltet die Nullhypothese eine Aussage über die Zuverlässigkeitsfunktion und ist somit dem in dieser Arbeit betrachteten Qualitätsmerkmal angepasst. Der Anteil  $p_0$  und der Zeitpunkt  $t_0$  stehen für die Qualitätsansprüche des Konsumenten und werden dementsprechend von eben diesem vorgegeben. Insgesamt lauten die Hypothesen

$$H_0 : R(t_0) \geq 1 - p_0 \text{ gegen } H_1 : R(t_0) < 1 - p_0.$$

Der statistische Test der Abnahmeprüfung beruht auf den Zuverlässigkeitsdaten der Stichprobe. Diese werden unter anderem in sogenannten “End-of-Life-Tests“ (kurz: EOL-Tests) erhoben (vgl. Meyna, Pauli, 2010, S. 579 f.). Ein EOL-Test beobachtet die  $n$  Einheiten der Stichprobe bis zum Ausfall und ermittelt so die Lebensdauern  $t_1, \dots, t_n$ , was unter Umständen zeitaufwändig sein kann. Die in dem Test ausgefallenen Einheiten werden nicht durch neue Einheiten ersetzt. Mit Hilfe der so erhobenen Lebensdauern wird eine Teststatistik  $S(t_1, \dots, t_n)$  gebildet und diese mit dem kritischen Bereich  $K$  zur Testentscheidung verglichen. Dabei ist der kritische Bereich über die Verteilung der Teststatistik und das Signifikanzniveau  $\alpha$  definiert.

Die Testentscheidung  $\varphi(t_1, \dots, t_n) = 1$  bedeutet für den Konsumenten, dass das Los den gewünschten Qualitätsansprüchen zum Signifikanzniveau  $\alpha$  nicht genügt. Das Los sollte in diesem Fall verworfen werden. Falls die Testentscheidung  $\varphi(t_1, \dots, t_n) = 0$  ist, spricht dies für einen Kauf des Loses.

Es bestehen noch weitere Möglichkeiten der zählenden Abnahmeprüfung. Neben dem EOL-Test kann beispielsweise ein Lebensdauertest durchgeführt werden (vgl. Meyna, Pauli, 2010, S. 576 ff.). Dabei werden die Stichprobeneinheiten über die Mindestlebensdauerzeit  $t_0$  beobachtet und die Anzahl der Ausfälle registriert. Auch bei diesem Verfahren werden die ausgefallenen Einheiten nicht ersetzt. Anhand der Ausfallzeitpunkte im Zeitraum 0 bis  $t_0$  wird über die Qualität des Loses entschieden. Ein weiteres Verfahren der Abnahmeprüfung ist die Sequentialprüfung (vgl. Meyna, Pauli, 2010, S. 559 ff.). Dieses oft auch Folgeprüfung genannte Vorgehen betrachtet die Elemente des Loses einzeln nacheinander. Nach jeder Betrachtung wird entschieden, wie im Weiteren verfahren wird. Das Los wird entweder abgelehnt oder angenommen oder die Prüfung wird fortgesetzt, indem die nächste Einheit kontrolliert wird. In der Praxis jedoch wird in bestimmten Zeitintervallen die Zahl der ausgefallenen Einheiten vermerkt. Weitere Abnahmeprüfungen thematisieren Meyna und Pauli in Kapitel 18.1 ihrer Veröffentlichung zur Zuverlässigkeitstechnik (Meyna, Pauli, 2010, S. 556 ff.).

### 3.3.3 Test bei Poisson-verteiltern Lebensdauern

Seien die Zufallsvariablen  $T_1, \dots, T_n$  gegeben und seien diese unabhängig identisch verteilt wie  $T_1 \sim Poi(\lambda)$ . Dabei geben die Realisationen der Zufallsvariablen die Anzahl an abgeschlossenen Zeiteinheiten an, bevor der Ausfall eintritt.

Das zu untersuchende Merkmal ist die Lebensdauer der Einheiten des Loses. Die Ansprüche an die Zuverlässigkeit in Form der Mindestlebensdauer  $t_0$  sollen mindestens  $(1 - p_0)$  der Einheiten erfüllen. Die interessierenden Testhypothesen lauten dementsprechend

$$H_0 : R_{Poi(\lambda)}(t_0) \geq 1 - p_0 \text{ gegen } H_1 : R_{Poi(\lambda)}(t_0) < 1 - p_0.$$

Dabei ist  $R_{Poi(\lambda)}(t_0)$  die Zuverlässigkeitsfunktion auf Grundlage der Verteilungsfunktion  $F_{Poi(\lambda)}$  und  $p_0$  der maximale Ausschussanteil. Des Weiteren kann  $R_{Poi(\lambda)}(t_0)$  mittels der Definition 3.31 und des Satzes 3.26 wie folgt umgeformt werden:

$$R_{Poi(\lambda)}(t_0) = 1 - F_{Poi(\lambda)}(t_0) = F_{\chi^2_{2(t_0+1)}}(2\lambda).$$

Dies eingesetzt in  $H_0$  führt zu:

$$\begin{aligned} F_{\chi^2_{2(t_0+1)}}(2\lambda) &\geq 1 - p_0 \\ \Leftrightarrow 2\lambda &\geq F_{\chi^2_{2(t_0+1)}}^{-1}(1 - p_0) \\ \Leftrightarrow \lambda &\geq \frac{1}{2} F_{\chi^2_{2(t_0+1)}}^{-1}(1 - p_0) =: \lambda_0. \end{aligned}$$

Die Testhypothesen können folglich geschrieben werden als

$$H_0 : \lambda \geq \lambda_0 \text{ gegen } H_1 : \lambda < \lambda_0.$$

Die Teststatistik  $S(T_1, \dots, T_n)$  des Parametertests bei Poisson-verteiltern Zufallsvariablen ist  $S(T_1, \dots, T_n) = \sum_{i=1}^n T_i$ . Es gilt für die Teststatistik unter  $H_0$  und nach Satz 3.23  $S(T_1, \dots, T_n) \sim Poi(n\lambda_0)$ . Der Test lehnt  $H_0$  ab, wenn  $S(t_1, \dots, t_n) \in K = [0, k)$  mit  $k = F_{Poi(n\lambda_0)}^{-1}(\alpha)$  gilt. Zusammenfassend hat der Test nach Definition 3.28 die Form

$$\varphi(t_1, \dots, t_n) = \begin{cases} 1, & \text{falls } S(t_1, \dots, t_n) = \sum_{i=1}^n t_i < k, \\ 0, & \text{falls } S(t_1, \dots, t_n) = \sum_{i=1}^n t_i \geq k. \end{cases}$$

Die OC-Funktion für Poisson-verteilte Lebensdauern in Abhängigkeit von  $\lambda$   $L_{Poi}^\lambda : (0, \infty) \rightarrow [0, 1]$  wird nun im Folgenden bestimmt. Mit  $c = k - 1$  folgt für die OC-Funktion

$$\begin{aligned} L_{Poi}^\lambda(\lambda) &= 1 - P(\varphi(t_1, \dots, t_n) = 1 \mid \lambda) \\ &= 1 - P(TS < k \mid \lambda) = 1 - P(TS \leq c \mid \lambda) \\ &= 1 - F_{Poi(n\lambda)}(c) \\ &= 1 - \sum_{z=0}^c \frac{(n\lambda)^z \cdot e^{-(n\lambda)}}{z!}. \end{aligned}$$

Des Weiteren wird die OC-Funktion bezüglich des Ausschussanteils  $p$  bestimmt. Dafür wird zunächst  $p$  unter Beachtung des Satzes 3.26 in Abhängigkeit von  $\lambda$  gesetzt

$$\begin{aligned} p &= F_{Poi(\lambda)}(t_0) \\ &= \int_{2\lambda}^{+\infty} f_{\chi_{2(t_0+1)}^2}(z) dz \\ &= [F_{\chi_{2(t_0+1)}^2}(z)]_{2\lambda}^{+\infty} \\ &= F_{\chi_{2(t_0+1)}^2}(+\infty) - F_{\chi_{2(t_0+1)}^2}(2\lambda) \\ &= 1 - F_{\chi_{2(t_0+1)}^2}(2\lambda). \end{aligned}$$

Anschließend wird so umgeformt, dass  $\lambda$  in Abhängigkeit von  $p$  gilt

$$\begin{aligned} p &= 1 - F_{\chi_{2(t_0+1)}^2}(2\lambda) \\ \Leftrightarrow F_{\chi_{2(t_0+1)}^2}(2\lambda) &= 1 - p \\ \Leftrightarrow \lambda &= \frac{1}{2} F_{\chi_{2(t_0+1)}^2}^{-1}(1 - p). \end{aligned}$$

Mit  $\lambda = \frac{1}{2} F_{\chi_{2(t_0+1)}^2}^{-1}(1 - p)$  kann nun die OC-Funktion  $L_{Poi}^\lambda$  zu  $L_{Poi}^p$  umgeformt und die OC-Funktion somit in Abhängigkeit von  $p$  dargestellt werden. Es gilt für die OC-Funktion  $L_{Poi}^p : (0, 1) \rightarrow [0, 1]$

$$L_{Poi}^p(p) = 1 - \sum_{z=0}^c \frac{(n \cdot \frac{1}{2} \cdot F_{\chi_{2(t_0+1)}^2}^{-1}(1 - p))^z \cdot e^{-\frac{1}{2} \cdot F_{\chi_{2(t_0+1)}^2}^{-1}(1 - p)}}{z!}.$$

### 3.3.4 Test bei geometrisch verteilten Lebensdauern

Seien die unabhängig und identisch wie  $T_1 \sim Geo(\pi)$  verteilten Zufallsvariablen  $T_1, \dots, T_n$  gegeben mit  $f_{Geo(\pi)}(T_1 = t_1) = \pi \cdot (1 - \pi)^{t_1}$ . Die Zufallsvariablen modellieren die Anzahl an abgeschlossenen Zeiteinheiten, bevor der Ausfall eintritt.

Erneut ist die Lebensdauer das interessierende Merkmal und die Qualitätsanforderung eine Mindestlebensdauer  $t_0$ . Der maximal akzeptierte Ausschussanteil ist  $p_0$ . Die interessierenden Testhypothesen lauten

$$H_0 : R_{Geo(\pi)}(t_0) \geq 1 - p_0 \text{ gegen } H_1 : R_{Geo(\pi)}(t_0) < 1 - p_0,$$

wobei  $R_{Geo(\pi)}(t_0)$  die Zuverlässigkeitsfunktion auf Basis von  $F_{Geo(\pi)}$  darstellt. Mit Hilfe der Definition 3.31 und der Verteilungsfunktion der geometrischen Verteilung kann  $R_{Geo(\pi)}(t_0)$  wie folgt umgeformt werden:

$$\begin{aligned} R_{Geo(\pi)}(t_0) &= 1 - F_{Geo(\pi)}(t_0) \\ &= 1 - 1 + (1 - \pi)^{t_0+1} \\ &= (1 - \pi)^{t_0+1}. \end{aligned}$$

Durch einsetzen in die Nullhypothese  $H_0$  ergibt sich:

$$\begin{aligned} (1 - \pi)^{t_0+1} &\geq 1 - p_0 \\ \Leftrightarrow 1 - \pi &\geq \sqrt[t_0+1]{1 - p_0} \\ \Leftrightarrow \pi_0 := 1 - \sqrt[t_0+1]{1 - p_0} &\geq \pi. \end{aligned}$$

Die Testhypothesen können folglich geschrieben werden als

$$H_0 : \pi \leq \pi_0 \text{ gegen } H_1 : \pi > \pi_0.$$

Die Teststatistik  $S(T_1, \dots, T_n)$  des Parametertests bei geometrisch verteilten Zufallsvariablen ist  $S(T_1, \dots, T_n) = \sum_{i=1}^n T_i$ . Dabei ist die Teststatistik unter  $H_0$  und nach Satz 3.20  $S(T_1, \dots, T_n) \sim \text{negBin}(\pi_0, n)$ . Die Testentscheidung lehnt  $H_0$  ab, wenn  $S(t_1, \dots, t_n) \in K = [0, k)$  mit  $k = F_{\text{negBin}(\pi_0, n)}^{-1}(\alpha)$  gilt. Zusammenfassend hat der Test nach Definition 3.28 die Form

$$\varphi(t_1, \dots, t_n) = \begin{cases} 1, & \text{falls } S(t_1, \dots, t_n) = \sum_{i=1}^n t_i < k, \\ 0, & \text{falls } S(t_1, \dots, t_n) = \sum_{i=1}^n t_i \geq k. \end{cases}$$

Die OC-Funktion für geometrisch verteilte Lebensdauern in Abhängigkeit von  $\pi$   $L_{Geo}^\pi : (0, 1) \rightarrow [0, 1]$  wird nun im Folgenden bestimmt. Mit  $c = k - 1$  folgt für die OC-Funktion

$$\begin{aligned} L_{Geo}^\pi(\pi) &= 1 - P(\varphi(t_1, \dots, t_n) = 1 \mid \pi) \\ &= 1 - P(TS < k \mid \pi) = 1 - P(TS \leq c \mid \pi) \\ &= 1 - F_{\text{negBin}(\pi, n)}(c) \\ &= 1 - \sum_{z=0}^c \binom{z+n-1}{z} \cdot \pi^n \cdot (1 - \pi)^z. \end{aligned}$$

Ähnlich wie bei der OC-Funktion für Poisson-verteilte Lebensdauern wird im Weiteren ebenfalls die OC-Funktion in Abhängigkeit von  $p$  bestimmt. Dafür wird  $\pi$  in Abhängigkeit von  $p$  gesetzt:

$$\begin{aligned}
 p &= F_{Geo(\pi)}(t_0) \\
 \Leftrightarrow p &= 1 - (1 - \pi)^{t_0+1} \\
 \Leftrightarrow (1 - \pi)^{t_0+1} &= 1 - p \\
 \Leftrightarrow 1 - \pi &= |{}^{t_0+1}\sqrt{1-p}| \\
 \Leftrightarrow \pi &= 1 - |{}^{t_0+1}\sqrt{1-p}|.
 \end{aligned}$$

Mit  $\pi = 1 - |{}^{t_0+1}\sqrt{1-p}|$  folgt nun für die exakte OC-Funktion für geometrisch verteilte Lebensdauern  $L_{Geo}^p : (0, 1) \rightarrow [0, 1]$

$$L_{Geo}^p(p) = 1 - \sum_{z=0}^c \binom{z+n-1}{z} \cdot (1 - |{}^{t_0+1}\sqrt{1-p}|)^n \cdot (|{}^{t_0+1}\sqrt{1-p}|)^z.$$



# Kapitel 4

## Aufbau der Simulationsstudie

In diesem Kapitel wird der Aufbau der Simulationsstudie zur Untersuchung der OC-Funktion erläutert. Diese Studie wurde beispielhaft für diskrete Lebensdauerverteilungen anhand von Poisson-verteilten und von geometrisch verteilten Lebensdauern durchgeführt. Die Vorgehensweise beruht auf den in Kapitel 3 hergeleiteten Tests für die jeweilige Verteilung. Die Durchführung wurde mit der Statistiksoftware R 3.0.2 (R Core Team, 2013) realisiert. Die dazu programmierten Funktionen (vgl. Anhang A) werden im Folgenden dargestellt und stehen auf der beiliegenden CD als Datei zur Verfügung.

### 4.1 Erläuterung der Funktionen

#### 4.1.1 Testdurchführung

Die mit *Test* bezeichnete Funktion führt den eigentlichen Test zur Bewertung des Loses aus. Die Eingabeparameter dieser Funktion sind neben der Angabe der Verteilung der Lebensdauern ( $V = \text{“Poi“}$  oder  $V = \text{“Geo“}$ ) die Stichprobengröße  $n \in \mathbb{N}$ , der Parameter  $\theta$  der Lebensdauerverteilung, den für die Hypothesen benötigten Parameter  $\theta_0$  und das Signifikanzniveau  $\alpha \in (0, 1)$ . Falls die Abnahmeprüfung für Poisson-verteilte Lebensdauern durchgeführt werden soll, dann gilt  $\theta, \theta_0 \in \mathbb{R}^+$ . Im Fall einer Abnahmeprüfung für geometrisch verteilte Lebensdauern muss hingegen gelten  $\theta, \theta_0 \in (0, 1)$ .

Zunächst werden  $n$  Zufallszahlen mit dem Parameter  $\theta$  aus der Verteilung, welche  $V$  benennt, gezogen. Bei dem Test für die Poisson-verteilten Lebensdauern gilt  $T_1, \dots, T_n$  sind unabhängig identisch verteilt wie  $T_1 \sim \text{Poi}(\theta)$ ; bei dem anderen Test stammen die unabhängig identisch verteilten Zufallszahlen  $T_1, \dots, T_n$  aus der Ver-

**Algorithm 1** Testdurchführung

---

```

1: procedure TEST( $V, n, \theta, \theta_0, \alpha$ )
2:   if  $V = \text{Poi}$  then
3:      $t_1, \dots, t_n \leftarrow n$  Zufallszahlen werden gezogen aus  $\text{Poi}(\theta)$ 
4:      $TS \leftarrow$  Summe der Zufallszahlen  $t_1, \dots, t_n$ 
5:      $k \leftarrow F_{\text{Poi}(n\theta_0)}^{-1}(\alpha)$ 
6:   else
7:      $t_1, \dots, t_n \leftarrow n$  Zufallszahlen werden gezogen aus  $\text{Geo}(\theta)$ 
8:      $TS \leftarrow$  Summe der Zufallszahlen  $t_1, \dots, t_n$ 
9:      $k \leftarrow F_{\text{negBin}(\theta_0, n)}^{-1}(\alpha)$ 
10:  end if
11:  if  $TS < k$  then
12:     $TE \leftarrow 1$  ▷ Los verwerfen
13:  else
14:     $TE \leftarrow 0$  ▷ Los nicht verwerfen
15:  end if
16:  return  $TE$ 
17: end procedure

```

---

teilung  $\text{Geo}(\theta)$ . Diese Zufallszahlen repräsentieren jeweils von einer Einheit aus der Stichprobe die Anzahl der abgeschlossenen Zeiteinheiten, bevor der Ausfall in einem EOL-Test eingetreten ist. Anschließend wird auf Grundlage dieser Zufallszahlen die Teststatistik  $S(t_1, \dots, t_n)$  gebildet. Konkret bedeutet dies, dass in beiden Fällen die gezogenen Zufallszahlen aufsummiert werden. Anschließend wird der kritische Wert  $k$  bestimmt. Falls es sich bei den gezogenen Zufallszahlen um Zufallszahlen aus der Poisson-Verteilung handelt, gilt  $k = F_{\text{Poi}(n\theta_0)}^{-1}(\alpha)$ . Falls hingegen der Fall  $V = \text{Geo}$  eintritt, dann wird  $k$  der Wert  $F_{\text{negBin}(\theta_0, n)}^{-1}(\alpha)$  zugewiesen. Schließlich wird die Testentscheidung bestimmt, indem geprüft wird, ob die Teststatistik  $S(t_1, \dots, t_n)$  kleiner als der kritische Wert  $k$  ist. Ist dies der Fall, kann das Los verworfen werden und die Funktion  $\text{Test}$  gibt eine 1 aus. Gilt hingegen  $TS \geq k$ , dann gibt die Funktion eine 0 aus.

### 4.1.2 Wiederholung des Tests

Die zweite programmierte Funktion heißt *Wiederholung* und bildet die relative Häufigkeit, dass der zugrunde liegende Test nicht ablehnt. Diese Funktion besitzt neben den Eingabeparametern der Funktion  $\text{Test}$  zusätzlich noch den Parameter  $W \in \mathbb{N}$ ,

welcher die Anzahl an Wiederholungen angibt.

---

**Algorithm 2** Wiederholung des Tests
 

---

```

1: procedure WIEDERHOLUNG( $V, n, \theta, \theta_0, \alpha, W$ )
2:    $x \leftarrow$  Vektor mit  $W$  leeren Einträgen
3:   for  $i = 1$  to  $W$  do
4:      $x[i] \leftarrow$  Test( $V, n, \theta, \theta_0, \alpha$ )
5:   end for
6:   return  $1 - \text{mean}(x)$ 
7: end procedure

```

---

Die Funktion *Wiederholung* führt zunächst die Funktion *Test*  $W$ -mal aus und speichert jeweils die Testentscheidung. Anschließend wird die relative Häufigkeit gebildet, dass der Test nicht abgelehnt hat. Dies geschieht auf Basis der  $W$  Testentscheidungen. Diese relative Häufigkeit wird schließlich ausgegeben.

### 4.1.3 Erstellen der OC-Funktion

---

**Algorithm 3** Erstellen der OC-Funktion
 

---

```

1: procedure OCFUNKTION( $V, n, p_0, t_0, \alpha, W, \theta_a, \theta_b, \Delta_\theta$ )
2:   if  $V = \text{Poi}$  then
3:      $\theta_0 \leftarrow \frac{1}{2} F_{\chi^2_{2(t_0+1)}}^{-1}(1 - p_0)$ 
4:   else
5:     if  $V = \text{Geo}$  then
6:        $\theta_0 \leftarrow 1 - |(1 - p_0)^{\frac{1}{t_0+1}}|$ 
7:     end if
8:   end if
9:    $\Theta \leftarrow \text{seq}(\theta_a, \theta_b, \Delta_\theta)$ 
10:   $OC \leftarrow$  Vektor mit  $\text{length}(\Theta)$  leeren Einträgen
11:  for  $j = 1$  to  $\text{length}(\Theta)$  do
12:     $OC[j] \leftarrow$  Wiederholung( $V, n, \Theta[j], \theta_0, \alpha, W$ )
13:  end for
14:  return  $OC$ 
15: end procedure

```

---

Für die Durchführung der Simulation ist schließlich noch die Funktion *OCFunktion* erstellt worden. Diese Funktion besitzt insgesamt neun Eingabeparameter. Es

muss die Verteilung  $V$  angegeben werden (“Poi“ für Poisson-verteilte Lebensdauern; “Geo“ für geometrisch verteilte Lebensdauern), aus welcher die Lebensdauern gezogen werden sollen, ebenso wie der Stichprobenumfang  $n \in \mathbb{N}$ . Des Weiteren werden Eingaben für  $p_0 \in (0, 1)$  sowie  $t_0 > 0$  benötigt, um das  $\theta_0$  zu bestimmen. Auch bei dieser Funktion muss das Signifikanzniveau  $\alpha \in (0, 1)$  und die Anzahl der Wiederholungen  $W \in \mathbb{N}$  eingegeben werden. Die Parameterwerte  $\theta_a, \theta_b$  und  $\Delta_\theta$  werden benötigt, um den Parameterbereich zu bestimmen, welcher simuliert wird. Dabei muss  $\Delta_\theta \leq \theta_b - \theta_a$  gelten, da der Parameterwert  $\Delta_\theta$  die Sequenzlänge zwischen den interessierenden Parameterwerten bestimmt. Für den Fall, dass  $\theta$  für den Parameter der Poisson-Verteilung steht, gilt  $\theta_a, \theta_b, \theta_\Delta \in \mathbb{R}^+$ . Falls jedoch die OC-Funktion für geometrisch verteilte Lebensdauern simuliert werden soll, dann gilt  $\theta_a, \theta_b, \theta_\Delta \in (0, 1)$ .

Zunächst wird abhängig von  $V$   $\theta_0$  in Abhängigkeit von  $p_0$  und  $t_0$  gebildet. Für Poisson-verteilte Zufallszahlen ist  $\theta_0 = \frac{1}{2} F_{\chi^2_{2(t_0+1)}}^{-1}(1 - p_0)$ . Falls  $V = \text{“Geo“}$ , dann wird  $\theta_0$  mittels  $1 - |(1 - p_0)^{\frac{1}{t_0+1}}|$  gebildet. Nachdem  $\theta_0$  bestimmt wurde, wird ein Vektor  $\Theta$  erstellt. Dieser beinhaltet eine Sequenz von  $\theta_a$  bis  $\theta_b$  in den Schritten  $\theta_\Delta$ . Schließlich wird mit jedem Eintrag  $\theta \in \Theta$  die Funktion *Wiederholung* ausgeführt und der mit dieser Funktion bestimmte Anteil, dass der Test nicht abgelehnt hat, gespeichert und ausgegeben.

## 4.2 Ablauf der Simulationsstudie

Die Simulationsstudie soll den Einfluss von verschiedenen Parametern auf die OC-Funktion einer zählenden Abnahmeprüfung bei Poisson-verteilten oder geometrisch verteilten Lebensdauern untersuchen. Da diese Funktion von mehreren verschiedenen Parametern beeinflusst wird, muss jeder Parameter einzeln betrachtet werden. Dies ist erforderlich, um die Auswirkung auf genau einen Parameter zurückführen zu können. Daher bekommen die Parameter der OC-Funktion Standardausprägungen zugewiesen, die in Tabelle 4.1 dargestellt sind.

|          | Variation 1 | Standardausprägung | Variation 2 |
|----------|-------------|--------------------|-------------|
| $\alpha$ | 0.01        | 0.05               | 0.1         |
| $n$      | 50          | 100                | 150         |
| $t_0$    | 1           | 5                  | 10          |
| $p_0$    | 0.05        | 0.1                | 0.15        |

**Tabelle 4.1:** Standardausprägung und Variation der Parameter in der Studie

Die Simulationsstudie wird für jede der beiden Verteilungen der Lebensdauern durchgeführt, zunächst für Poisson-verteilte und anschließend für geometrisch verteilte Lebensdauern. Zu Beginn wird der Parameter  $\alpha$  variiert, dem die Ausprägungen 0.01, 0.05 und 0.1 zugewiesen werden (siehe auch Tabelle 4.1). Folglich wird die Funktion *OCFunktion* dreimal aufgerufen, jeweils mit einer der drei Ausprägungen von  $\alpha$ . In der gesamten Simulationsstudie wurden jeweils 200000 Wiederholungen durchgeführt ( $W = 200000$ ). Zur Visualisierung stellt eine Grafik die drei auf diesem Wege bestimmten Annahmewahrscheinlichkeiten in Abhängigkeit von  $\theta$  dar. Eine weitere Grafik zeigt diese Annahmewahrscheinlichkeiten bezüglich des Losschlechtanteils  $p$ . Dies ist auf Grund der in Kapitel 3 dargestellten Abhängigkeit zwischen  $\theta$  und  $p$  ohne weitere Simulationen mittels Umrechnungen möglich. Die beiden Grafiken werden im Kapitel 5 ausgewertet und analysiert. Anschließend folgt dasselbe Vorgehen für die drei anderen Parameter  $n$ ,  $t_0$  und  $p_0$  mit den jeweils der Tabelle 4.1 zu entnehmenden Variationen. Es ist noch anzumerken, dass logischerweise jeweils die Kurven mit den Standardausprägungen identisch sind.

Die Wahl der Variationsausprägungen ist relativ willkürlich geschehen. Bei dem Parameter  $\alpha$  sind die drei Ausprägungen die in der Praxis am häufigsten verwendeten Werte. Für die Variation des Parameters  $n$  wurden die Größen 50, 100, 150 gewählt, da diese sehr realistische und vermutlich in der Praxis häufig gewählte Größenbereiche darstellen. Gleiches gilt für die Parameter  $t_0$  und  $p_0$ . Obwohl unterschiedliche Produkte unterschiedliche Lebenserwartungen haben, sind die Zeiträume 1 Jahr, 5 Jahre und 10 Jahre in der Industrie/Wirtschaft gebräuchliche Lebensdauern. Auch der Ausschussanteil ist mit 0.05, 0.1 und 0.1 realitätsnah gewählt.

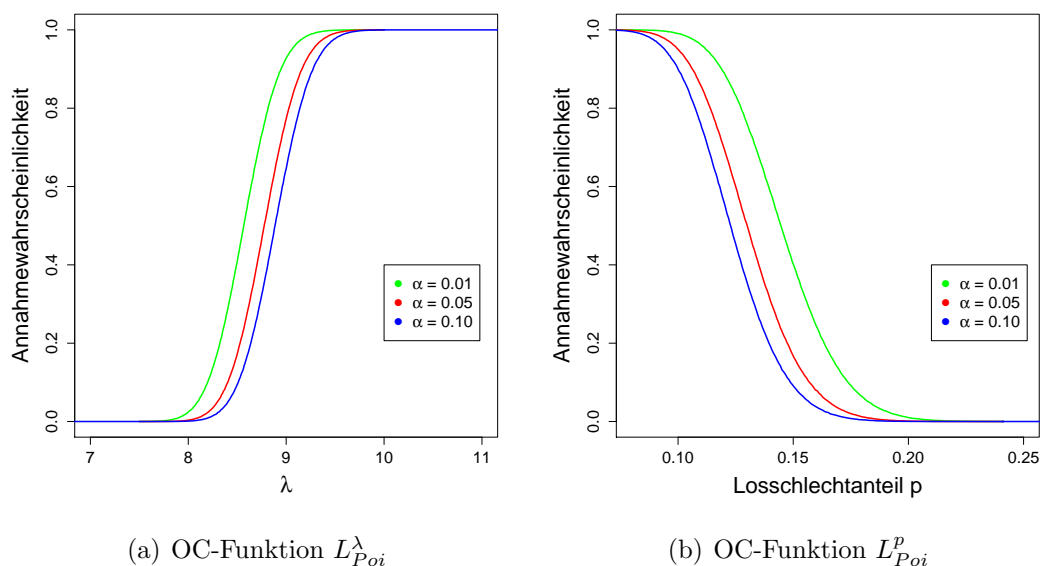


# Kapitel 5

## Auswertung

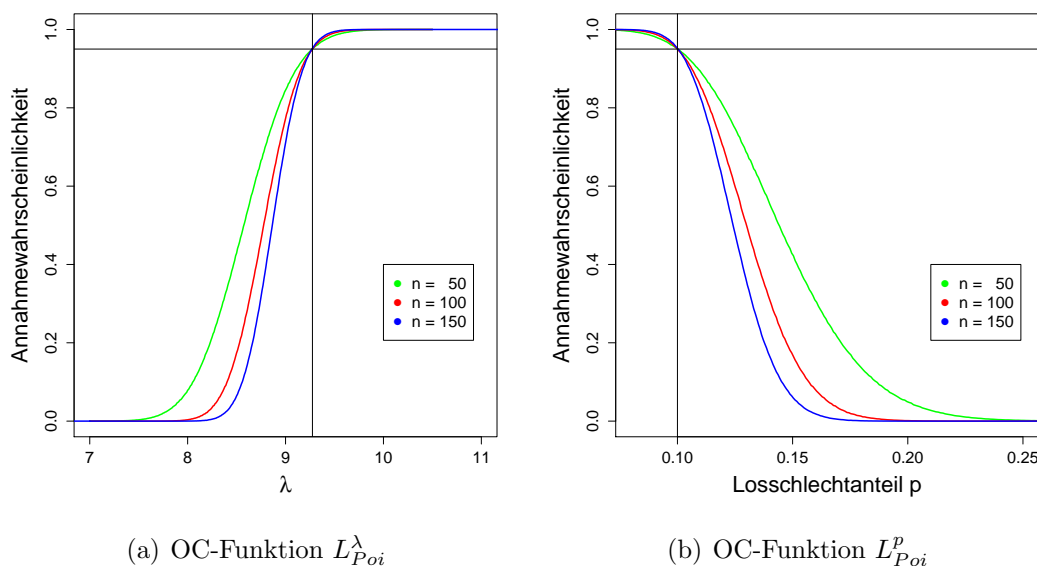
Die im Folgenden betrachteten Daten zeigen das Ergebnis der in Kapitel 4 beschriebenen Simulationsstudie. Sie werden in diesem Kapitel grafisch dargestellt und analysiert. Anschließend werden beispielhaft die simulierten OC-Funktionen mit den exakten OC-Funktionen verglichen. Danach erfolgt ein Vergleich der OC-Funktionen auf Basis von Poisson-verteilten und geometrisch verteilten Lebensdauern, um die Anwendbarkeit zu diskutieren. Die in diesem Kapitel gezeigten Grafiken befinden sich in Form von PDF-Dateien auf der beiliegenden CD.

### 5.1 Untersuchung der Simulationsstudie



**Abbildung 5.1:** OC-Funktion  $L_{Poi}$  abhängig von  $\lambda$  bzw.  $p$  mit den festen Parametern  $n = 100$ ,  $t_0 = 5$  und  $p_0 = 0.1$ . Das Signifikanzniveau  $\alpha$  wird variiert.

Zunächst werden die Simulationsergebnisse für Poisson-verteilte Lebensdauern betrachtet. Abbildung 5.1 zeigt die OC-Funktionen auf Grundlage der Daten, welche mit variierendem Signifikanzniveau  $\alpha$  simuliert wurden. In Abbildung 5.1(a) ist bei den OC-Funktionen  $L_{Poi}^\lambda$  zu erkennen, dass die monoton steigenden Kurven nahezu identisch verlaufen. Je größer  $\lambda$  wird, desto unwahrscheinlicher wird es, dass das Los nicht verworfen wird. Zusätzlich ist zu erkennen: Je größer das Signifikanzniveau  $\alpha$  gewählt wurde, desto weiter sind die Kurven nach rechts verschoben. In Abbildung 5.1(b) erkennt man drei monoton fallende Kurven mit nahezu identischem Kurvenverlauf. Je geringer das Signifikanzniveau  $\alpha$  gewählt wurde, desto weiter ist die zugehörige OC-Funktion  $L_{Poi}^p$  nach rechts verschoben.

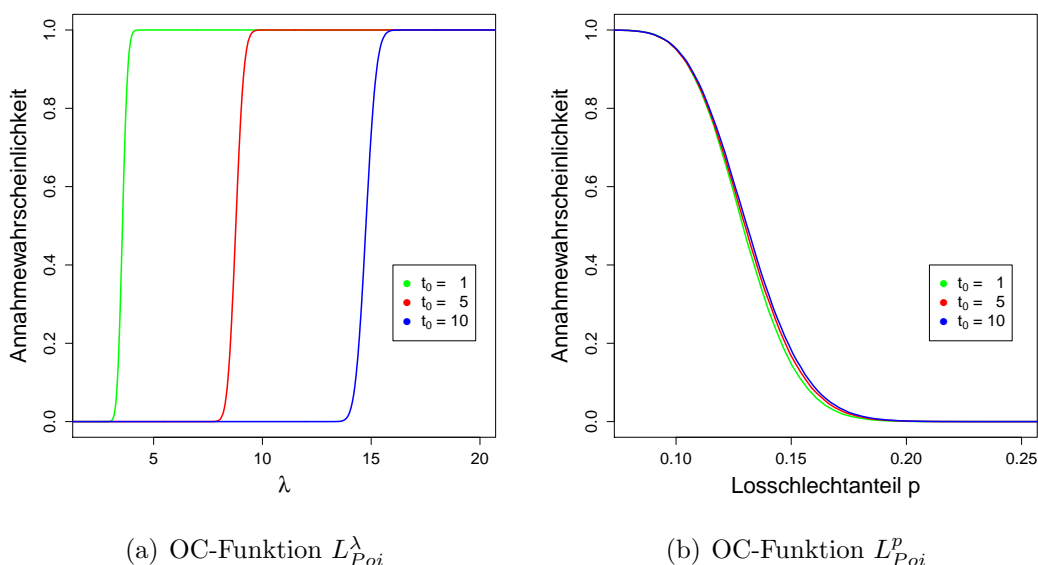
(a) OC-Funktion  $L_{Poi}^\lambda$ (b) OC-Funktion  $L_{Poi}^p$ 

**Abbildung 5.2:** OC-Funktion  $L_{Poi}$  abhängig von  $\lambda$  bzw.  $p$  mit den festen Parametern  $\alpha = 0.05$ ,  $t_0 = 5$  und  $p_0 = 0.1$ . Die Stichprobengröße  $n$  wird variiert.

In Abbildung 5.2 sind die Daten der Simulationsstudie grafisch dargestellt, bei welchen die Stichprobengröße  $n$  variiert wurde. Es ist zu erwarten, dass eine Abnahmeprüfung auf Basis eines größeren Stichprobenumfangs eine höhere Güte aufweist. Dies hängt mit der Überlegung zusammen, dass ein Test basierend auf mehr Daten eine geringere Wahrscheinlichkeit besitzt, eine Fehlentscheidung zu treffen. Die von  $\lambda$  abhängigen OC-Funktionen in Abbildung 5.2(a) weisen nicht das gleiche Steigungsverhalten auf. Es ist deutlich zu erkennen, dass eine auf einem weniger großen Stichprobenumfang basierende Abnahmeprüfung eine geringere Güte aufweist. Dies drückt der flachere Kurvenverlauf bei niedriger Stichprobengröße aus. Folglich besitzt, wie erwartet, eine Abnahmeprüfung mit größerem Stichproben-



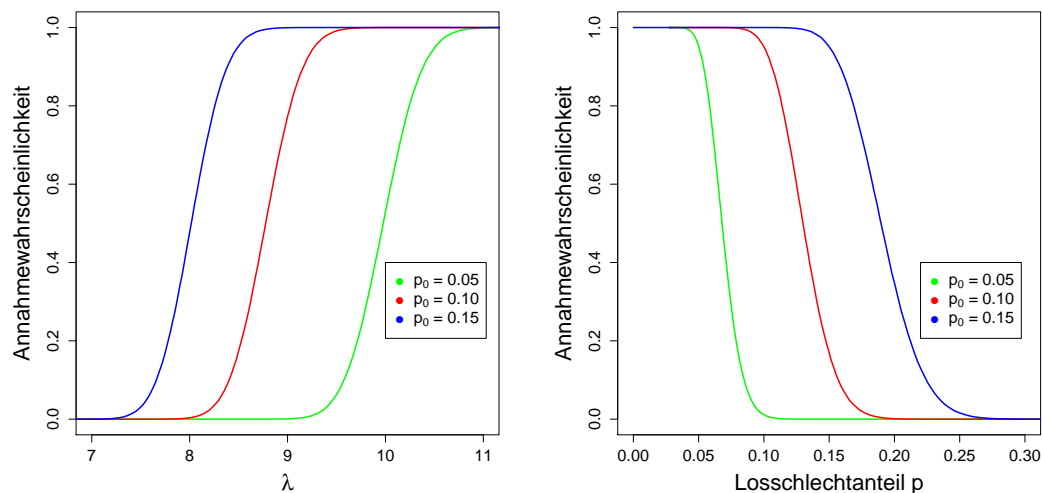
umfang eine höhere Güte. Die OC-Funktionen schneiden sich alle in dem Punkt  $(\lambda_0 = \frac{1}{2}F_{\chi^2_{2(t_0+1)}}^{-1}(1-p_0), 1-\alpha) = (\frac{1}{2}F_{\chi^2_{12}}^{-1}(0.9), 0.95) \approx (9.27, 0.95)$ , da dieser durch die Nullhypothese mittels  $\lambda_0$  und das Signifikanzniveau  $\alpha$  vorgegeben ist. In Abbildung 5.2(b) zeigt sich ebenfalls die unterschiedliche Güte bei verschiedenen Stichprobenumfängen. Die Schnittstelle der drei Kurven bei  $p = 0.1 = p_0$  und der Annahmewahrscheinlichkeit  $0.95 = 1 - \alpha$  ist durch die festgelegten Parameter wiederum gegeben. Des Weiteren ist zu erkennen, dass mit wachsendem Stichprobenumfang der Fehler zweiter Art für jedes  $p > p_0$  geringer wird. Dieser Zusammenhang ist jedoch nicht linear.



**Abbildung 5.3:** OC-Funktion  $L_{P_{oi}}$  abhängig von  $\lambda$  bzw.  $p$  mit den festen Parametern  $\alpha = 0.05$ ,  $n = 100$  und  $p_0 = 0.1$ . Die Mindestlebensdauer  $t_0$  wird variiert.

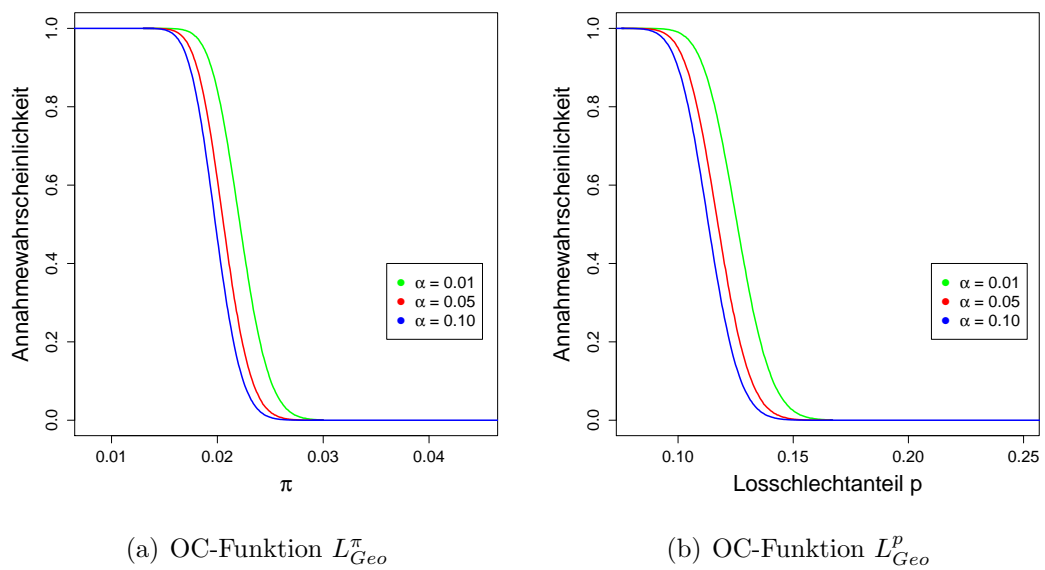
Abbildung 5.3 stellt die Daten dar, die durch die Variation der Mindestlebensdauer  $t_0$  in der Simulationsstudie entstanden sind. Bei Betrachtung der Grafik 5.3(a) muss mit Hinblick auf die anderen Grafiken beachtet werden, dass hier eine andere Skalierung verwendet wurde. Es ist zu erkennen, dass die Variation der Mindestlebensdauer keinen deutlichen Einfluss auf die Güte der Abnahmeprüfung hat, da die Verlaufsmuster der Kurven weitgehend identisch sind: Je höher die Mindestlebensdauer ist, desto weiter ist die Kurve nach rechts verschoben. Dies ist auf die von  $t_0$  abhängige Nullhypothese zurückzuführen. In Abbildung 5.3(b) kann ein minimaler Einfluss von  $t_0$  auf die OC-Funktion  $L_{P_{oi}}^p$  identifiziert werden, da die drei Kurven nur einen nahezu identischen, optisch kaum voneinander unterscheidbaren, Kurvenverlauf aufweisen. Eine geringere Mindestlebensdauer  $t_0$  wirkt sich somit positiv auf die Güte der Abnahmeprüfung aus.

Zuletzt wird noch der Einfluss des maximal akzeptierten Losschlechtanteils  $p_0$  untersucht (vgl. Abbildung 5.4). Erneut ist zu erkennen, dass die Kurven der drei OC-Funktionen verschoben sind. Dies ist wiederum auf die Nullhypothese zurückzuführen, da diese ebenfalls von  $p_0$  abhängt. Je geringer das  $p_0$  gewählt wurde, desto weiter ist die Kurve nach rechts verschoben. In Abbildung 5.4(b) ist zu erkennen, dass eine Abnahmeprüfung mit  $p_0 = 0.05$  eine höhere Güte aufweist als die beiden Abnahmeprüfungen mit höherem maximal akzeptiertem Losschlechtanteil. Auch bei dieser Grafik wurde eine abweichende Skalierung verwendet.

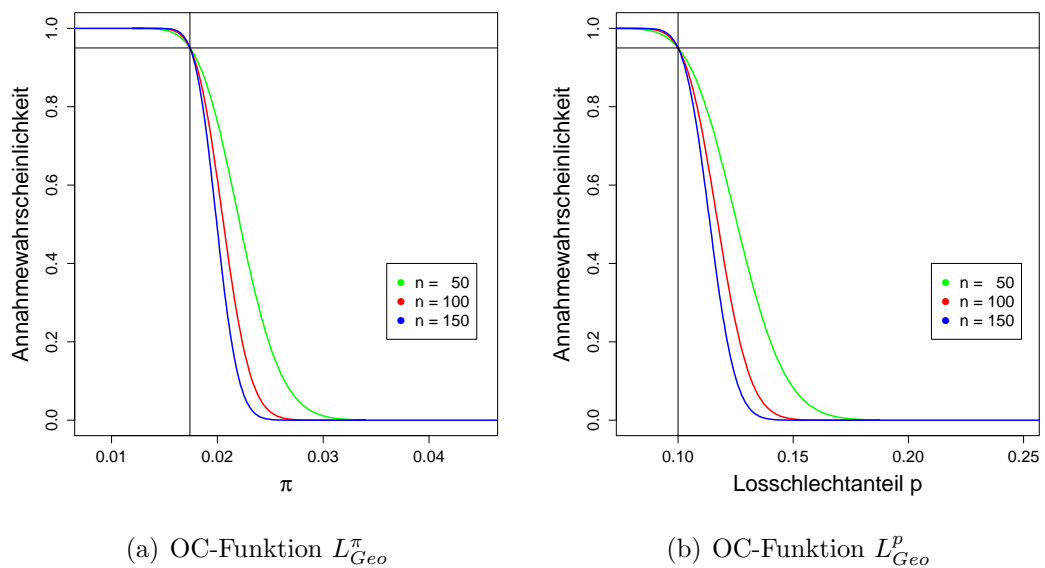
(a) OC-Funktion  $L_{Poi}^\lambda$ (b) OC-Funktion  $L_{Poi}^p$ 

**Abbildung 5.4:** OC-Funktion  $L_{Poi}$  abhängig von  $\lambda$  bzw.  $p$  mit den festen Parametern  $\alpha = 0.05$ ,  $n = 100$  und  $t_0 = 5$ . Der maximal akzeptierte Losschlechtanteil  $p_0$  wird variiert.

Die im Folgenden betrachteten Daten basieren auf geometrisch verteilten Lebensdauern. Bei der Simulation der in Abbildung 5.5 mittels OC-Funktion gezeigten Daten wurde zunächst der Parameter  $\alpha$  variiert. Bei Betrachtung der OC-Funktion  $L_{Geo}^\pi$  ist zu erkennen, dass die monoton fallenden Kurven identisch verlaufen (vgl. Abb. 5.5(a)). Je größer das Signifikanzniveau  $\alpha$  gewählt wurde, desto weiter sind die Kurven nach links verschoben. Dieselben Eigenschaften fallen auch beim Betrachten der Abbildung 5.5(b) auf. Die dort dargestellten Kurven sind monoton fallend. Da diese die gleiche Steilheit aufweisen, besitzen die OC-Funktionen der verschiedenen Signifikanzniveaus dieselbe Güte.



**Abbildung 5.5:** OC-Funktion  $L_{Geo}$  abhängig von  $\pi$  bzw.  $p$  mit den festen Parametern  $n = 100$ ,  $t_0 = 5$  und  $p_0 = 0.1$ . Das Signifikanzniveau  $\alpha$  wird variiert.



**Abbildung 5.6:** OC-Funktion  $L_{Geo}$  abhängig von  $\pi$  bzw.  $p$  mit den festen Parametern  $\alpha = 0.05$ ,  $t_0 = 5$  und  $p_0 = 0.1$ . Der Stichprobenumfang  $n$  wird variiert.

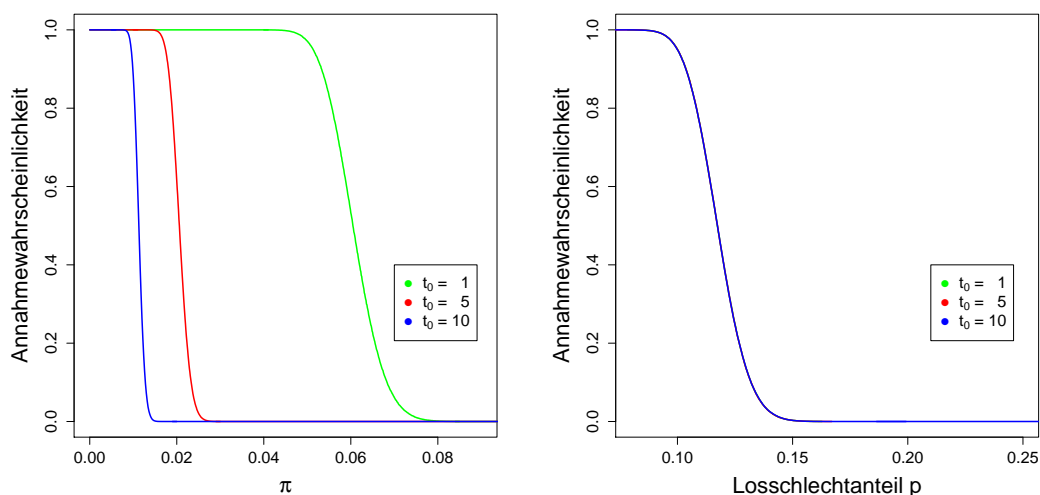
Abbildung 5.6 zeigt die OC-Funktionen  $L_{Geo}^\pi$  und  $L_{Geo}^p$ , welche mit unterschiedlichen Stichprobenumfängen simuliert wurden. Im Gegensatz zu den OC-Funktionen in Abbildung 5.5(a) weisen die in Abbildung 5.6(a) dargestellten Kurven nicht denselben Verlauf auf. Die mittels eines niedrigeren Stichprobenumfangs simulierten

Abnahmeprüfungen lassen eine geringere Güte erkennen, abgeleitet aus der Steilheit der Kurven. Durch die Parameter  $\alpha$ ,  $p_0$  und  $t_0$  ist der in Abbildung 5.6(a) zu sehende Schnittpunkt definiert:  $(1 - |{}^{t_0+1}\sqrt{1-p_0}|, 1 - \alpha) = (1 - |\sqrt[9]{0.9}|, 0.95) \approx (0.018, 0.95)$ . Die in Abbildung 5.6(b) dargestellten Kurven weisen einen Schnittpunkt in  $(p_0, 1 - \alpha) = (0.1, 0.95)$  auf. Auch kann diese Grafik die bereits gewonnene Erkenntnis über den Zusammenhang zwischen Güte und Stichprobenumfang bestätigen.

Bevor die OC-Funktionen in Abbildung 5.7(a) betrachtet werden, muss auf den unterschiedlichen Maßstab der Grafik hingewiesen werden. Mit Blick auf die Skala der x-Achse zeigt sich die enorme Verschiebung der Kurven nach rechts, abhängig von der zunehmenden Länge der geforderten Mindestlebensdauer. Diese Verschiebung steht in keinem linearen Zusammenhang mit dem Betrag der Mindestlebensdauer  $t_0$  und erklärt sich durch die von  $t_0$  abhängige Nullhypothese. Bei Betrachtung der Abbildung 5.7(b) scheint ein Einfluss von  $t_0$  auf die OC-Funktion  $L_{Geo}^p$  nicht vorzuliegen. Jedoch ist die exakte OC-Funktion

$$L_{Geo}^p(p) = 1 - \sum_{z=0}^c \binom{z+n-1}{z} \cdot (1 - |{}^{t_0+1}\sqrt{1-p}|)^n \cdot (|{}^{t_0+1}\sqrt{1-p}|)^z$$

von der Mindestlebensdauer  $t_0$  abhängig. Die drei identischen Kurvenverläufe sind dadurch zu erklären, dass der Unterschied zwischen den Kurven minimal ist und durch die Rechenungenauigkeit in der Grafik nicht sichtbar wird.

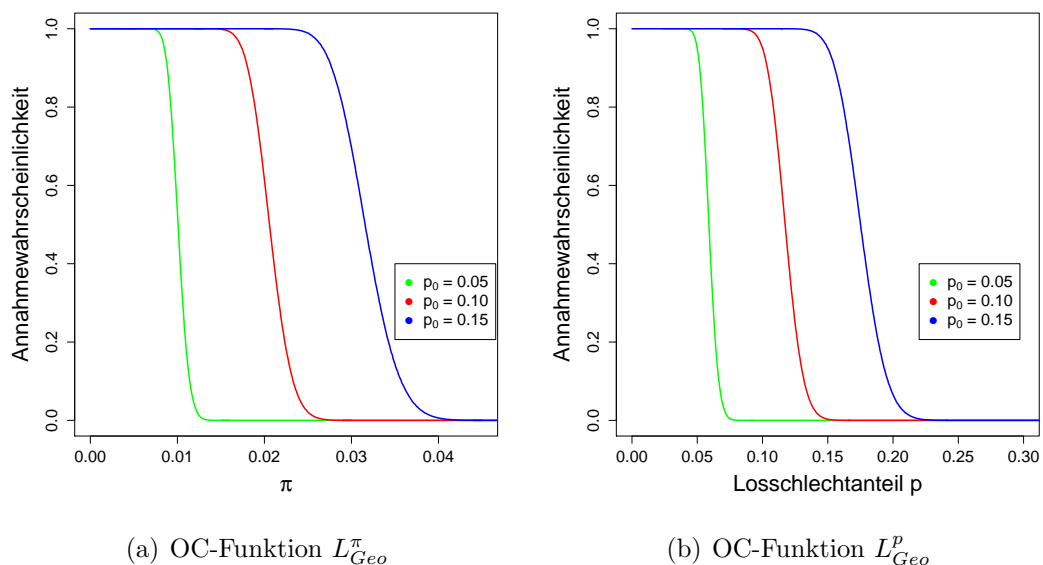


(a) OC-Funktion  $L_{Geo}^\pi$

(b) OC-Funktion  $L_{Geo}^p$

**Abbildung 5.7:** OC-Funktion  $L_{Geo}$  abhängig von  $\pi$  bzw.  $p$  mit den festen Parametern  $\alpha = 0.05$ ,  $n = 100$  und  $p_0 = 0.1$ . Die Mindestlebensdauer  $t_0$  wird variiert.

Schließlich werden die Daten der Abnahmeprüfungen analysiert, bei denen der maximal akzeptierte Losschlechtanteil  $p_0$  variiert wurde (vgl. Abb. 5.8). Bei den OC-Funktionen  $L_{Geo}^\pi$  fällt die Verschiebung nach rechts auf, je größer der Losschlechtanteil gesetzt wurde. Des Weiteren erkennt man in Abbildung 5.8(a), dass die Kurvenverläufe nicht identisch sind: Die Güte ist umso höher, je geringer der maximale Ausschussanteil ist. In Abbildung 5.8(b) zeigt sich dieser Aspekt ebenfalls. Die Kurven der OC-Funktionen  $L_{Geo}^p$  sind nach rechts verschoben, je größer der maximal akzeptierte Losschlechtanteil  $p_0$  ist. Dabei muss erneut die Skalierung beachtet werden, da diese von den anderen Grafiken abweicht.



**Abbildung 5.8:** OC-Funktion  $L_{Geo}$  abhängig von  $\pi$  bzw.  $p$  mit den festen Parametern  $\alpha = 0.05$ ,  $n = 100$  und  $t_0 = 5$ . Der maximal akzeptierte Losschlechtanteil  $p_0$  wird variiert.

Zusammenfassend kann festgestellt werden, dass sich die unterschiedlichen Parameter  $\alpha$ ,  $n$ ,  $t_0$  und  $p_0$ , unabhängig von der Verteilung der Lebensdauern, in ähnlicher Weise auf die OC-Funktionen  $L_{Poi}^p$  und  $L_{Geo}^p$  auswirken. Die Vergrößerung des Signifikanzniveaus  $\alpha$  bewirkt eine Verschiebung der OC-Funktionen nach links, wobei die Güte der Abnahmeprüfung unverändert bleibt. Die Vergrößerung des Stichprobenumfangs  $n$  bewirkt eine Verbesserung der Güte, wobei diese Verbesserung nicht in einem linearen Zusammenhang zu  $n$  steht. Die Erhöhung des maximal akzeptierten Losschlechtanteils bewirkt eine lineare Verschiebung der OC-Funktionen nach rechts, wobei die Güte der Abnahmeprüfungen sich leicht verschlechtert. Bei der Veränderung der Mindestlebensdauer  $t_0$  zeigen sich nur minimale Änderungen der OC-Funktionen  $L_{Poi}$  und  $L_{Geo}$ .

## 5.2 Vergleich der simulierten mit der exakten OC-Funktion

Die bisher betrachteten Daten wurden mittels der in Kapitel 4 vorgestellten Simulationsstudie erzeugt. Im Folgenden soll untersucht werden, ob die durch Simulation entstandenen OC-Funktionen der jeweiligen exakten OC-Funktion entsprechen. Falls starke Abweichungen auftreten, sind die Erkenntnisse der bereits durchgeführten Analyse nicht allgemeingültig. Diese Überprüfung findet beispielhaft an OC-Funktionen mit jeweils festen Parametern  $\alpha = 0.05$ ,  $n = 100$ ,  $t_0 = 5$  und  $p_0 = 0.1$  für Poisson- und geometrisch verteilte Lebensdauern statt.

Zunächst werden die simulierten und exakten OC-Funktionen  $L_{Poi}^\lambda$  und  $L_{Poi}^p$  mit oben genannten Parameterwerten verglichen. Für die exakten OC-Funktionen für Poisson-verteilte Lebensdauern wird zunächst  $\lambda_0$  bestimmt

$$\lambda_0 = \frac{1}{2} F_{\chi^2_{2(t_0+1)}}^{-1}(1 - p_0) = \frac{1}{2} F_{\chi^2_{12}}^{-1}(0.9) \approx 9.27$$

und anschließend mittels des kritischen Wertes  $k$

$$k = F_{Poi(n\lambda_0)}^{-1}(\alpha) \approx F_{Poi(927.47)}^{-1}(0.05) = 878$$

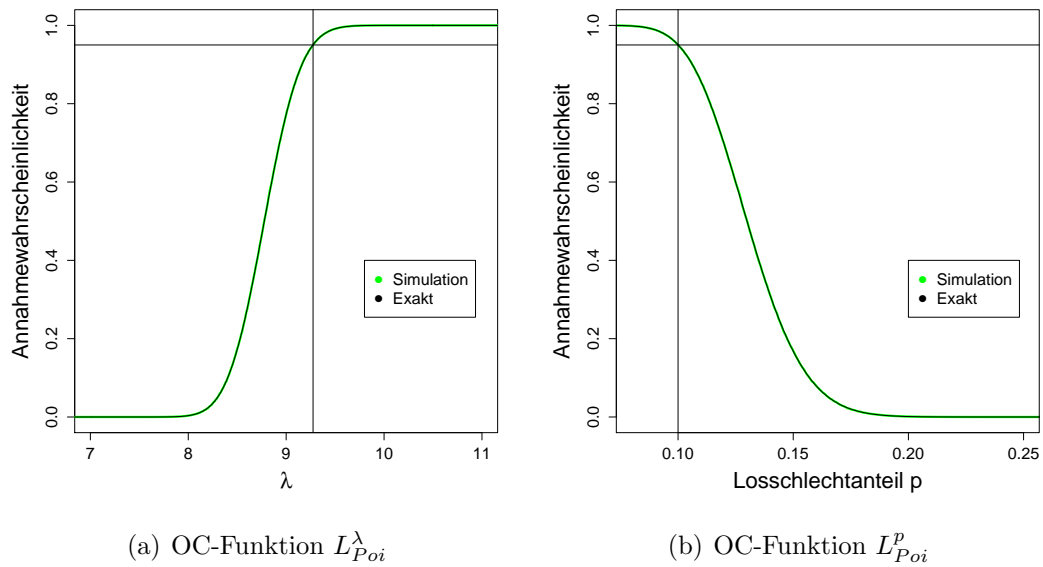
$c = k - 1 = 877$  bestimmt. Die exakte OC-Funktion  $L_{Poi}^\lambda$  besitzt die Form

$$\begin{aligned} L_{Poi}^\lambda(\lambda) &= 1 - \sum_{z=0}^c \frac{(n\lambda)^z \cdot e^{-(n\lambda)}}{z!} \\ &= 1 - \sum_{z=0}^{877} \frac{(100\lambda)^z \cdot e^{-(100\lambda)}}{z!} \end{aligned}$$

und die exakte OC-Funktion in Abhängigkeit von  $p$  lautet

$$\begin{aligned} L_{Poi}^p(p) &= 1 - \sum_{z=0}^c \frac{\left(\frac{n}{2} \cdot F_{\chi^2_{2(t_0+1)}}^{-1}(1-p)\right)^z \cdot e^{-\left(\frac{n}{2} \cdot F_{\chi^2_{2(t_0+1)}}^{-1}(1-p)\right)}}{z!} \\ &= 1 - \sum_{z=0}^{877} \frac{\left(50 \cdot F_{\chi^2_{12}}^{-1}(1-p)\right)^z \cdot e^{-50 \cdot F_{\chi^2_{12}}^{-1}(1-p)}}{z!} \end{aligned}$$

(vgl. Kapitel 3.3.3). Es ist in Abbildung 5.9 direkt zu erkennen, dass die beiden Kurven identisch verlaufen. Das bedeutet, dass die Simulation die exakte OC-Funktion sehr gut abbildet, unabhängig von dem abhängigen Parameter.



**Abbildung 5.9:** Vergleich der simulierten und der exakten OC-Funktion  $L_{P_{oi}}$  abhängig von  $\lambda$  bzw.  $p$  mit den Parametern  $\alpha = 0.05$ ,  $n = 100$ ,  $t_0 = 5$  und  $p_0 = 0.1$ .

Zum Vergleich der exakten OC-Funktionen  $L_{Geo}^\pi$  und  $L_{Geo}^p$  werden erneut zunächst  $\pi_0$  und der kritische Wert  $k$  bestimmt. Dabei ist  $\pi_0$

$$\pi_0 = 1 - |{}^{t_0+1}\sqrt{1-p_0}| = 1 - |\sqrt[6]{0.9}| \approx 1 - 0.98 = 0.02$$

und der kritische Wert

$$k = F_{negBin(\pi_0, n)}^{-1}(\alpha) \approx F_{negBin(0.02, 100)}^{-1}(0.05) = 4742.$$

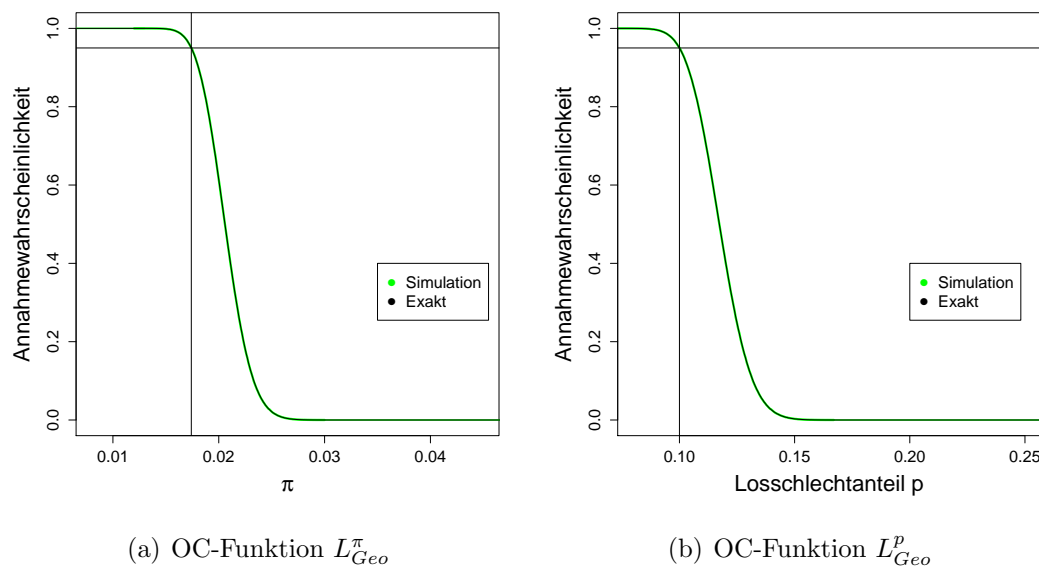
Folglich gilt  $c = k - 1 = 4741$  und die exakte OC-Funktion  $L_{Geo}^\pi$  hat die Form

$$\begin{aligned} L_{Geo}^\pi(\pi) &= 1 - \sum_{z=0}^c \binom{z+n-1}{z} \cdot \pi^n \cdot (1-\pi)^z \\ &= 1 - \sum_{z=0}^{4741} \binom{z+99}{z} \cdot \pi^{100} \cdot (1-\pi)^z. \end{aligned}$$

Für  $L_{Geo}^p$  gilt

$$\begin{aligned} L_{Geo}^p(p) &= 1 - \sum_{z=0}^c \binom{z+n-1}{z} \cdot (1 - |{}^{t_0+1}\sqrt{1-p}|)^n \cdot (|{}^{t_0+1}\sqrt{1-p}|)^z \\ &= 1 - \sum_{z=0}^{4741} \binom{z+99}{z} \cdot (1 - |\sqrt[6]{1-p}|)^{100} \cdot (|\sqrt[6]{1-p}|)^z. \end{aligned}$$

Die Betrachtung der Abbildung 5.10 zeigt, dass die exakten OC-Funktionen  $L_{Geo}^\pi$  und  $L_{Geo}^p$  genau mit den simulierten OC-Funktionen übereinstimmen.



**Abbildung 5.10:** Vergleich der simulierten und der exakten OC-Funktion  $L_{Geo}$  abhängig von  $\pi$  bzw.  $p$  mit den Parametern  $\alpha = 0.05$ ,  $n = 100$ ,  $t_0 = 5$  und  $p_0 = 0.1$ .

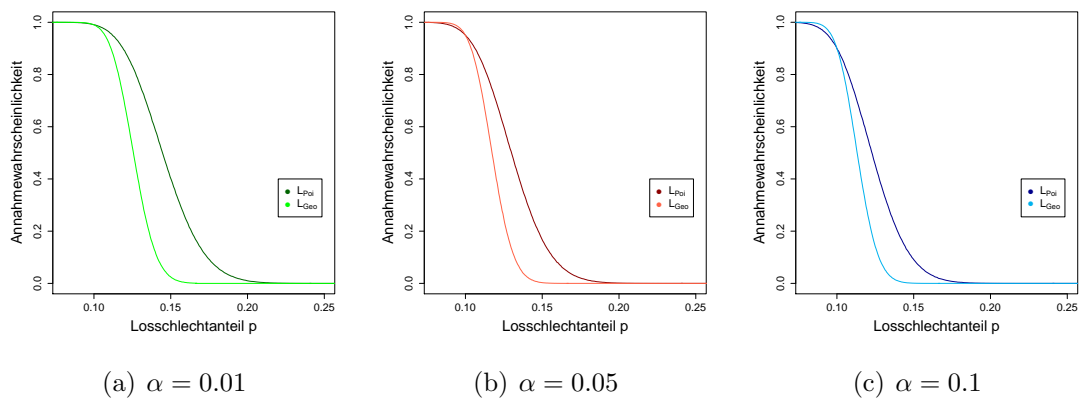
Zusammenfassend können die in Kapitel 5.1 festgestellten Auswirkungen der Parameter auf die simulierten OC-Funktionen auf die exakten OC-Funktionen verallgemeinert werden.

### 5.3 Vergleich der Poisson- und der geometrischen Verteilung als Lebensdauerverteilung

In dieser Bachelorarbeit wird als Qualitätsmerkmal von Produkten deren Lebensdauer betrachtet. Dabei wurde die Verteilung dieses Merkmals zum einen als geometrische Verteilung und zum anderen als Poisson-Verteilung aufgefasst. Im Folgenden soll nun untersucht werden, welche der beiden Verteilungen sich besser dazu eignet, den genannten Sachverhalt darzustellen. Dazu werden im Folgenden erneut die OC-Funktionen in Abhängigkeit vom Losschlechtanteil  $p$  auf Basis der in der Simulationsstudie erzeugten Daten betrachtet.

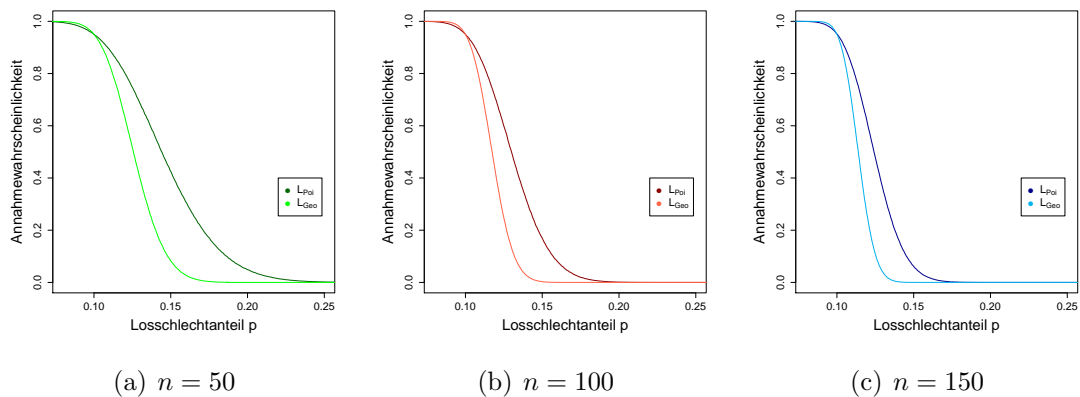
Abbildung 5.11 zeigt, dass die Güte der Abnahmeprüfung auf Basis von geometrisch verteilten Lebensdauern in jedem der drei Fällen höher ist. Da die Variation des Signifikanzniveaus keinen Einfluss auf die Güte der jeweiligen Abnahmeprüfungen hat, ist der mit größerem  $\alpha$  geringer werdende Unterschied der zwei Funktionen auf die unterschiedliche Verschiebung der OC-Funktionen zurückzuführen. Die Schnittpunkte der Kurven sind durch die Parameter  $p_0 = 0.1$  und  $\alpha$  gegeben.





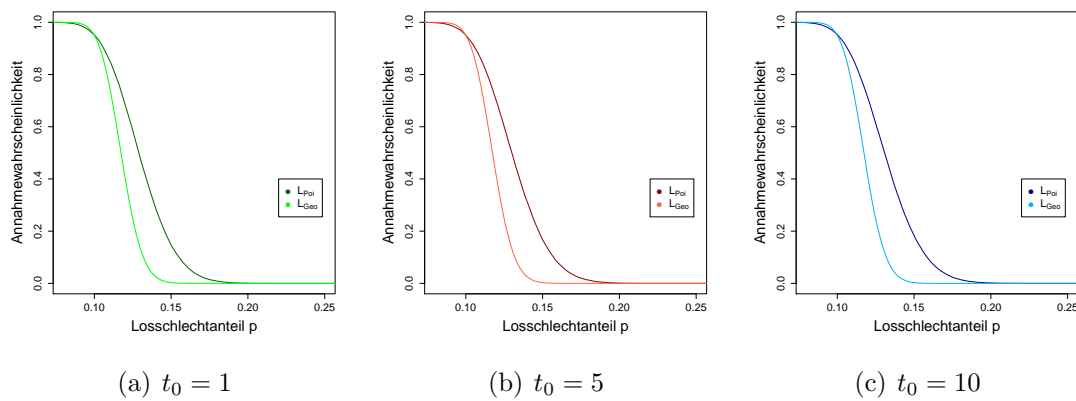
**Abbildung 5.11:** Vergleich der OC-Funktionen  $L^p_{Poi}$  und  $L^p_{Geo}$  mit den Parametern  $n = 100$ ,  $t_0 = 5$  und  $p_0 = 0.1$ .

Bei der Betrachtung der Auswirkungen von unterschiedlichen Stichprobenumfängen auf die OC-Funktionen erkennt man ein ähnliches Verhalten: In jeder der drei Datensituationen ist die Güte der Abnahmeprüfungen auf Grundlage von geometrischen Lebensdauern höher (Abbildung 5.12). Es ist anzumerken, dass die Güte von  $L^p_{Geo}$  nicht so stark unter einem niedrigen Stichprobenumfang leidet wie  $L^p_{Poi}$ . Mit wachsendem Stichprobenumfang  $n$  wird der Unterschied der OC-Funktionen  $L^p_{Poi}$  und  $L^p_{Geo}$  geringer.

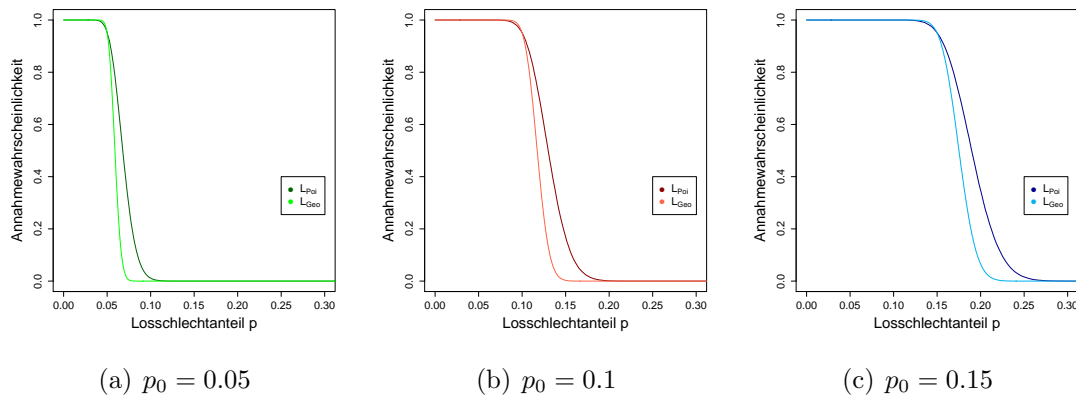


**Abbildung 5.12:** Vergleich der OC-Funktionen  $L^p_{Poi}$  und  $L^p_{Geo}$  mit den Parametern  $\alpha = 0.05$ ,  $t_0 = 5$  und  $p_0 = 0.1$ .

In Abbildung 5.13 ist deutlich zu erkennen, dass die Kurvenverläufe der jeweiligen zugrunde liegenden Verteilungen nahezu identisch sind. Insgesamt haben die Abnahmeprüfungen mit zugrunde liegenden geometrisch verteilten Lebensdauern eine höhere Güte. Jedoch ist der Unterschied nahezu identisch, da die Mindestlebensdauer  $t_0$  nur einen sehr geringen Einfluss auf die einzelnen OC-Funktionen hat (vgl. Kapitel 5.1).



**Abbildung 5.13:** Vergleich der OC-Funktionen  $L_{Poi}^p$  und  $L_{Geo}^p$  mit den Parametern  $\alpha = 0.05$ ,  $n = 100$  und  $p_0 = 0.1$ .



**Abbildung 5.14:** Vergleich der OC-Funktionen  $L_{Poi}^p$  und  $L_{Geo}^p$  mit den Parametern  $\alpha = 0.05$ ,  $n = 100$  und  $t_0 = 5$ .

Schließlich zeigt Abbildung 5.14, dass die Güte der OC-Funktionen  $L_{Geo}^p$ , unabhängig von der Höhe des gewählten maximal akzeptierten Losschlechtanteil, höher ist. Jedoch ist der Unterschied bei niedrigerem  $p_0$  geringer, als bei höherem maximal akzeptiertem Ausschussanteil.

Zusammenfassend ist festzustellen, dass die Abnahmeprüfungen auf Grundlage von geometrisch verteilten Lebensdauern in allen Situationen eine höhere Güte aufweisen und folglich zu bevorzugen sind. Demzufolge scheint es sinnvoll, nach Möglichkeit die abgeschlossenen Zeiteinheiten, bevor ein Ausfall eintritt, mittels der geometrischen Verteilung zu modellieren.

# Kapitel 6

## Zusammenfassung

Die vorliegende Bachelorarbeit untersucht das Verhalten von OC-Funktionen bei diskreten Lebensdauerverteilungen, denen unter anderem im Qualitätsmanagement eine hohe Bedeutung zukommt. Dies erfolgt beispielhaft anhand der Poisson-Verteilung und der geometrischen Verteilung als Lebensdauerverteilungen durch die Durchführung einer Simulationsstudie. Von Interesse ist das Verhalten von OC-Funktionen bei unterschiedlichen Datenlagen, da diese auf Grund von Kompromissen zwischen den Ansprüchen des Konsumenten und denen des Produzenten entstehen. Ein Konsument bevorzugt beispielsweise einen niedrigeren Stichprobenumfang, da diese Einheiten nach der Prüfung keine Verwendung mehr finden, wohingegen der Produzent möglicherweise aus Kostengründen, in Form der Vermeidung von weiteren Investitionen, einen höheren Ausschussanteil akzeptiert.

Zunächst wurde in Kapitel 2 eine Einführung in die Thematik der Qualitätssicherung gegeben, welche neben der Definition von Qualität und Zuverlässigkeit die Motivation zur Untersuchung von OC-Funktionen liefert. Nachdem in Kapitel 3 die mathematischen und statistischen Grundlagen eingeführt und die zu untersuchenden Abnahmeprüfungen hergeleitet wurden, stellt das Kapitel 4 den Ablauf der Simulationsstudie und die dazu benötigten programmierten R-Funktionen vor. Die daraus resultierenden Daten liefern die Basis für die anschließende Auswertung. Dazu wurden zunächst jeweils für die zugrunde liegenden Poisson- bzw. geometrisch verteilten Lebensdauern die Variation des Signifikanzniveaus  $\alpha$ , des Stichprobenumfangs  $n$ , der Mindestlebensdauer  $t_0$  und des maximal akzeptierten Losschlechtanteils  $p_0$  analysiert. Dabei konnte festgestellt werden, dass diese Parameter sich in ähnlicher Weise auf die OC-Funktionen  $L_{Poi}^p$  und  $L_{Geo}^p$  auswirken. Im Anschluss wurde untersucht, inwieweit Unterschiede zwischen den simulierten OC-Funktionen und den exakten, in Kapitel 3 hergeleiteten, OC-Funktionen vorliegen. Die Untersuchung zeigt, dass

die Erkenntnisse der Simulationsstudie auf exakte OC-Funktionen übertragen werden können. Die zu Grunde liegenden Verteilungen modellieren beide die Anzahl an abgeschlossenen Einheiten, bevor der Ausfall eintritt. Daher war von Interesse, welche der beiden Verteilungen zur Modellierung eben genannter Sachverhalte besser geeignet ist: Die auf geometrisch verteilten Lebensdauern basierenden Abnahmeprüfungen weisen in allen untersuchten Situationen eine höhere Güte auf.

Auf Grund der Tatsache, dass die Untersuchung auf einer an zwei Verteilungen exemplarisch durchgeführten Simulationsstudie basiert, sollten die Ergebnisse unter Vorbehalt betrachtet werden. Insoweit ist zu empfehlen, die Ergebnisse in Form einer unabhängigen Studie nochmals zu verifizieren. Auch kann oder sollte im Zuge dessen eine Erweiterung der Variationsstufen der einzelnen Parameter erfolgen. Weiterhin sollte der Vergleich der auf Grundlage der Simulationen erstellten OC-Funktionen mit den exakten OC-Funktionen näher untersucht werden.

# Anhang A

## Programmiercode

```
#Eingabeparameter:
# V          - Verteilung der Lebensdauern
# n          - Stichprobengroesse
# theta      - wahrer Parameterwert der Stichprobe
# theta_Null - Parameterwert unter H_0
# alpha      - Signifikanzniveau

#Ausgabe:
# Testentscheidung: 1 - Los wird verworfen
#                   0 - Los wird nicht verworfen

Test <- function(V, n, theta, theta_Null, alpha){
  if(V == "Poi"){
    TS <- sum(rpois(n,theta))
    k <- qpois(alpha,n*theta_Null)
  }else{
    TS <- sum(rgeom(n,theta))
    k <- qnbinom(alpha,n,theta_Null)
  }
  return(as.numeric(TS < k))
}
```

```
#Eingabeparameter:
# V          - Verteilung der Lebensdauern
# n          - Stichprobengroesse
# theta      - wahrer Parameterwert der Stichprobe
# theta_Null - Parameterwert unter H_0
# alpha      - Signifikanzniveau
# W          - Anzahl der Wiederholungen

#Ausgabe:
# Anteil der nicht verworfenen Lose von den W untersuchten.

Wiederholung <- function(V, n, theta, theta_Null, alpha, W){
  x <- numeric()
  for(i in 1:W){
    x[i] <- Test(V, n, theta, theta_Null, alpha)
  }
  return(1-mean(x))
}
```

```
#Eingabeparameter:
# V          - Verteilung der Lebensdauern
# n          - Stichprobengroesse
# p_Null     - maximal akzeptierter Ausschussanteil
# t_Null     - geforderte Mindestlebensdauer
# alpha      - Signifikanzniveau
# W          - Anzahl der Wiederholungen
# theta_a    - kleinster interessierender Parameterwert
# theta_b    - groesster interessierender Parameterwert
# Delta_theta - Sequenzlaenge zwischen den interessierenden Parameterwerten

#Ausgabe:
# Annahmewahrscheinlichkeiten der OC-Funktion im Bereich [theta_a,theta_b]
```

```
OCFunktion <- function(V, n, p_Null, t_Null, alpha, W, theta_a,
                      theta_b, Delta_theta){
  if(V == "Poi"){
    theta_Null <- 1/2 * qchisq(1-p_Null, 2*(t_Null+1))
  }else{
    if(V == "Geo"){
      theta_Null <- 1 - abs((1-p_Null)^(1/(t_Null+1)))
    }else{stop("Fehler bei Verteilungsangabe")}
  }
  w_thetas <- seq(theta_a,theta_b,Delta_theta)
  OC <- numeric()
  for(j in 1:length(w_thetas)){
    OC[j] <- Wiederholung(V, n, w_thetas[j], theta_Null, alpha, W)
  }
  return(OC)
}
```





# Literaturverzeichnis

- Bamberg, G., Baur, F., Krapp, M. (2008): *Statistik*, 14. Auflage, Oldenbourg Verlag, München, Wien.
- Dehling, H., Haupt, B. (2004): *Einführung in die Wahrscheinlichkeitstheorie und Statistik*, 2. Auflage, Springer Verlag, Berlin, Heidelberg.
- DIN Deutsches Institut für Normung e.V. (2000): *Qualitätsmanagementsysteme: Grundlagen und Begriffe (ISO 9000:2000)*, Beuth Verlag GmbH, Berlin.
- Fahrmeir, L., Künstler, R., Pigeot, I., Tutz, G. (2010): *Statistik*, 7. Auflage, Springer Verlag, Berlin Heidelberg.
- Hartung, J., Elpelt, B., Klösener, K.-H. (2009): *Statistik*, 15. Auflage, Oldenbourg Verlag, München.
- Meyna, A., Pauli, B. (2010): *Zuverlässigkeitstechnik*, 2. Auflage, Carl Hanser Verlag, München, Wien.
- Mosler, K., Schmid, F. (2006): *Wahrscheinlichkeitsrechnung und schließende Statistik*, 2. Auflage, Springer Verlag, Berlin, Heidelberg.
- R Core Team (2013): *R: A language and environment for statistical computing.*, R Foundation for Statistical Computing, Wien.
- Rasch, D. (1995): *Mathematische Statistik: Eine Einführung für Studenten der Mathematik, Statistik, Biometrie und Naturwissenschaften*, 1. Auflage, Johann Ambrosius Barth Verlag, Heidelberg, Leipzig.
- Rinne, H., Mittag, H.-J. (1995): *Statistische Methoden der Qualitätssicherung*, 3. Auflage, Carl Hanser Verlag, München, Wien.
- Roman, D. (2010): „*Ein schlechtes Teil kann eine Lawine auslösen*“, 20 Minuten AG, Zürich, <http://www.20min.ch/tools/suchen/story/31828595> (21.5.2014).

Timischl, W. (2002): *Qualitätssicherung: Statistische Methoden*, 3. Auflage, Carl Hanser Verlag, München, Wien.

TOMORROW FOCUS Media GmbH (2010): *Dudenhöffer: Toyota-Rückruf hat langfristige Folgen*,  
[http://www.focus.de/auto/news/dudenhoeffler-toyota-rueckruf-hat-langfristige-folgen\\_aid\\_478472.html](http://www.focus.de/auto/news/dudenhoeffler-toyota-rueckruf-hat-langfristige-folgen_aid_478472.html) (21.5.2014).

Toyota Deutschland GmbH (2014): *Für ihre Zufriedenheit: Unser Qualitätsversprechen*, <http://www.toyota.de/rueckruf> (21.5.2014).

Vogel, F. (2000): *Beschreibende und schließende Statistik: Formeln, Definitionen, Erläuterungen, Stichwörter und Tabellen*, 12. Auflage, Oldenbourg Verlag, München.

# Eidesstattliche Versicherung

Aschenbruck, Rabea Dagmar

Matrikel-Nr.: 140431

Ich versichere hiermit an Eides statt, dass ich die vorliegende Bachelorarbeit mit dem Titel

## OC-Funktion bei diskreten Lebensdauerverteilungen anhand der Poisson-Verteilung und der geometrischen Verteilung

selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinn-gemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

---

Dortmund, den

---

Unterschrift

### Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden (§ 63 Abs. 5 Hochschulgesetz -HG-).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafen bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

---

Dortmund, den

---

Unterschrift

## **Einverständniserklärung des Urhebers**

Ich erkläre mich hiermit einverstanden, dass meine Bachelorarbeit nach § 6 (1) des URG der Öffentlichkeit durch die Übernahme in die Bereichsbibliotheken zugänglich gemacht wird. Damit können Leser der Bibliothek die Arbeit einsehen und zu persönlichen wissenschaftlichen Zwecken Kopien aus dieser Arbeit anfertigen. Weitere Urheberrechte werden nicht berührt.

---

Dortmund, den

---

Unterschrift