

# Qualitative Leistungsbewertung von Produktionschargen von Solarmodulen basierend auf Stichprobenmessungen

– Projektbericht für die TÜV Immissionsschutz und Energiesysteme GmbH –

Werner Herrmann\*

Ansgar Steland†

Henryk Zähle‡

---

\*TÜV Immissionsschutz und Energiesysteme GmbH, D-51101 Köln, werner.herrmann@de.tuv.com, Tel: 0221/806-2272

†RWTH Aachen, Institut für Statistik und Wirtschaftsmathematik, Wüllnerstr. 3, D-52056 Aachen steland@stochastik.rwth-aachen.de, Tel: 0241/80-94573

‡RWTH Aachen, Institut für Statistik und Wirtschaftsmathematik, Wüllnerstr. 3, D-52056 Aachen zaehle@stochastik.rwth-aachen.de, Tel: 0241/80-94612

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>3</b>
1.1	Grundlegendes Problem . . . . .	3
1.2	Vorbemerkungen und Fallunterscheidungen . . . . .	4
<b>2</b>	<b>Szenario ohne Flasherliste (keine Verteilungsannahme)</b>	<b>5</b>
2.1	Modellbeschreibung . . . . .	5
2.2	Wahl der Charakteristik der Entscheidungsregel . . . . .	6
<b>3</b>	<b>Szenario ohne Flasherliste (Normalverteilungsannahme)</b>	<b>10</b>
3.1	Modellbeschreibung . . . . .	10
3.2	Wahl der Charakteristik der Entscheidungsregel . . . . .	11
<b>4</b>	<b>Szenario mit Flasherliste (Normalverteilungsannahme)</b>	<b>15</b>
4.1	Modellbeschreibung . . . . .	15
4.2	Wahl der Charakteristik der Entscheidungsregel . . . . .	16
<b>5</b>	<b>Szenario mit Flasherliste (allgemeine Verteilung)</b>	<b>19</b>
5.1	Modellbeschreibung . . . . .	19
5.2	Wahl der Charakteristik der Entscheidungsregel . . . . .	20
<b>6</b>	<b>Zusammenfassung</b>	<b>23</b>
<b>7</b>	<b>Verteilungstafeln</b>	<b>24</b>
7.1	Quantile der Standardnormalverteilung . . . . .	24
7.2	Quantile von $\chi^2$ -Verteilungen . . . . .	25

# 1 Einführung

## 1.1 Grundlegendes Problem

Betreiber von photovoltaischen Anlagen (kurz: PV-Anlagen) werden regelmäßig mit der Tatsache konfrontiert, dass eine bestimmte Anzahl von verbauten Solarmodulen einem gewissen Leistungsanspruch nicht genügt. Dieses Problem tritt typischerweise bei Geräten aus Massenproduktionen auf, bei denen der Aufwand einer totalen Funktionalitätsprüfung des Produktionsausstoßes für den Produzenten zu zeitaufwändig und insbesondere zu kostspielig wäre.

Um das daraus für den Betreiber entstehende Risiko einzudämmen, sollte der Betreiber der Anlage die gelieferten Solarmodule vor dem Einbau wenigstens stichprobenartig überprüfen. Ganz natürlich ist hierbei der Wunsch nach einer möglichst kleinen Stichprobe, da dies den Zeit- und Kostenaufwand gering hält. Dieser Wunsch wird andererseits von dem Problem konterkariert, dass geringe Stichprobenumfänge nur relativ unsichere Aussagen über die Qualität der gesamten Lieferung zulassen. Die Wechselwirkung zwischen dem Umfang einer Stichprobe einerseits und dem Konfidenzlevel der aus der Stichprobenauswertung deduzierten Aussagen andererseits ist für die Evaluierung einer Produktionscharge von Solarmodulen somit von zentraler Bedeutung.

Der Betreiber einer PV-Anlage ist letztlich an einer Entscheidungsregel interessiert, die auf einer Stichprobe basiert und mit der er entscheiden kann, ob er die Solarmodullieferung annehmen oder reklamieren soll. Die Entscheidungsregel sollte dabei folgende Eigenschaften besitzen:

- (a) Besitzt die Lieferung eine „hohe Qualität“, dann sollte die Wahrscheinlichkeit, dass die Regel die Lieferung annimmt, einen bestimmten (möglichst großen) Wert  $\alpha$  nicht unterschreiten.
- (b) Ist die Lieferung dagegen von „geringer Qualität“, so sollte die Wahrscheinlichkeit, dass die Regel die Lieferung annimmt, einen bestimmten (möglichst kleinen) Wert  $\beta$  nicht überschreiten.

Die Werte  $\alpha$  und  $\beta$  legen gewissermaßen das Konfidenzlevel der Entscheidungsregel fest. Durch (b) wird das Risiko für den Betreiber der PV-Anlage beschränkt, Eigenschaft (a) garantiert, dass auch der Produzent von der Regel fair behandelt wird. Eine solche Konvention, insbesondere das gewissenhafte Festlegen von  $\alpha$  und  $\beta$ , ist für einen Vertragsabschluss letztlich von großer Bedeutung.

Um (unter bestimmten Modellannahmen) die Existenz einer entsprechenden Entscheidungsregel für vorgegebene Schwellenwerte  $\alpha$  und  $\beta$  zu garantieren, ist es u. a. notwendig, dass der Umfang der Stichprobe, auf der die Regel basieren soll, hinreichend groß ist. Wie bereits erwähnt sind die Vertragsparteien andererseits an kleinen Stichprobenumfängen interessiert. Dadurch entsteht für einen gegebenen Qualitätsanspruch und gegebene  $\alpha$  und  $\beta$  auf ganz natürliche Weise folgende Aufgabenstellung:

- Finde eine Entscheidungsregel und den Mindestumfang der Stichprobe, auf dem sie basiert, so dass das durch  $\alpha$  und  $\beta$  vorgegebene Konfidenzlevel gewährleistet ist!

## 1.2 Vorbemerkungen und Fallunterscheidungen

Um das am Ende von Kapitel 1.1 aufgeworfene Problem untersuchen zu können, muss der dargestellte Sachverhalt in ein Modell gefasst werden.

Betrachten wir die Situation aber zunächst noch einmal vom praktischen Standpunkt. Bei der Produktion von Solarmodulen mit einer gewünschten Leistung (Nennleistung)  $\mu_0$  weicht die tatsächliche Leistung der produzierten Module regelmäßig von dieser Nennleistung ab. Selbst der Mittelwert  $\mu$  der Leistungen der einzelnen Module stimmt in der Praxis oft nicht mit  $\mu_0$  überein. Bei einer „guten“ Produktion entspricht  $\mu$  der Nennleistung  $\mu_0$  (oder liegt zumindest „nahe“ bei ihr) und die systematische Abweichung der einzelnen Solarmodulleistungen von  $\mu$  fällt „klein“ aus. Aus statistischer Sicht ist die Varianz  $\sigma^2$  ein geeignetes Maß, um die genannte systematische Abweichung zu quantifizieren. Alternativ kann auch die Standardabweichung  $\sigma$  verwendet werden.

Der Hersteller der Module wird den Leistungsnennwert  $\mu_0$  möglicherweise einfach so angegeben, wie er es gerne hätte, d. h.  $\mu_0 :=$  gewünschte Leistung. Es ist aber nicht zuletzt von der Qualität der Produktionsgerätschaft abhängig, ob dieser Wert auch mit dem tatsächlichen Leistungsmittel  $\mu$  übereinstimmt. Ziel des Herstellers sollte es sein, den Produktionsprozess so zu organisieren, dass das tatsächliche Leistungsmittel  $\mu$  gerade auf  $\mu_0$  fällt. Erreicht er dies nicht, so müsste er fairer Weise  $\mu_0 := \mu$  als Nennwert ausschreiben. Zwar ist das tatsächliche Leistungsmittel  $\mu$  typischerweise unbekannt, es kann vom Hersteller aber (sehr) gut geschätzt werden, indem er eine (möglichst große) Stichprobe aus seiner Produktion testet und  $\mu$  als mittleren Leistungswert dieser Stichprobe definiert. Ob er sich in der Tat daran hält, kann vom Betreiber der PV-Anlage nur schwerlich überprüft werden. Selbst wenn der Hersteller Testmessungen durchführt und daraus  $\mu$  und somit  $\mu_0 := \mu$  fair spezifiziert, ist nicht klar, ob seine Messungen hinreichend genau sind. Somit kann aus Sicht des Betreibers der PV-Anlage nicht ad hoc  $\mu_0 = \mu$  angenommen werden.

Für die Modellbildung unterscheiden wir grundsätzlich zwischen den folgenden beiden Szenarien:

1. („ohne Flasherliste“). Der Betreiber der PV-Anlage erhält neben der Produktionscharge *keine* Angaben über die Leistungseigenschaften der gelieferten Module.
2. („mit Flasherliste“). Der Betreiber der PV-Anlage erhält neben der Produktionscharge eine Flasherliste, in der die individuellen Leistungen aller gelieferten Module angegeben sind. Die Liste wurde vom Hersteller der Module angefertigt und basiert auf seinen Messungen.

Für das erste Szenario betrachten wir sowohl einen verteilungsfreien Ansatz als auch ein Modell, in dem eine Normalverteilung als Verteilung der Module angenommen wird. Für das zweite Szenario legen wir zum einen ebenfalls eine Normalverteilungsannahme zugrunde. Und zum anderen studieren wir den Fall, in der die Leistungsverteilung durch die empirische Verteilungsfunktion geschätzt wird. Wir spezifizieren und studieren diese Modelle in den Kapiteln 2 bis 5 .

## 2 Szenario ohne Flasherliste (keine Verteilungsannahme)

### 2.1 Modellbeschreibung

Wir machen folgende Modellannahmen:

- (M1) Über die Wahrscheinlichkeitsverteilung der Solarmoduleleistungen liegt dem Konsumenten (Betreiber der PV-Anlage) keinerlei Information vor.
- (M2) Der Produzent der Solarmodule quantifiziert die Nennleistung auf  $\mu_0$ . Wie er auf diesen Wert kommt und ob dieser gerechtfertigt ist, spielt dabei keine Rolle.
- (M3) Ein Modul gilt als tolerierbar, wenn seine Leistung in einem vorgegebenen Intervall  $I_{\mu_0}$  liegt, das sinnvoller Weise  $\mu_0$  enthält. Die Festlegung dieses Toleranzbereiches basiert dabei auf der (subjektiven) Einschätzung des Betreibers der PV-Anlage und muss bei Vertragsabschluss ggf. mit dem Produzenten der Solarmodule abgestimmt worden sein. Technische Aspekte sollten hier die entscheidende Rolle spielen.

Das Intervall  $I_{\mu_0}$  nimmt typischerweise die Gestalt  $[\mu_0 - \epsilon, \mu_0 + \epsilon]$  oder  $[\mu_0 - \epsilon, \infty)$  an. In der Praxis wird in diesen Fällen regelmäßig mit  $\epsilon = 0.05\mu_0$  gearbeitet, d. h. mit einer Toleranz von 5%.

Unter den Voraussetzungen (M1)-(M3) ist der Konsument der Solarmodule, also der Betreiber der PV-Anlage, prinzipiell in einer recht unvorteilhaften Position. Er hat keinerlei Informationen über die Wahrscheinlichkeitsverteilung der Solarmoduleleistungen und auch der Wert  $\mu_0$  trägt letztlich nur beschränkte Information, da er vom Hersteller beliebig angegeben werden kann. Ihm bleibt somit nichts weiter übrig, als aus den angelieferten Solarmodulen eine Stichprobe zu nehmen, die Module mit Leistungswerten außerhalb von  $I_{\mu_0}$  zu zählen und mit Hilfe dessen eine Entscheidung über Annahme oder Reklamation der Lieferung zu treffen. Die Anzahl der nicht tolerierbaren Module ist neben der Stichprobengröße und der Gesamtanzahl der gelieferten Module die einzige wirkliche Information, die der Konsument besitzt. Somit kann seine Entscheidung über Annahme oder Reklamation der Charge auch nur auf diesen Größen fußen. Letztlich sollte die Entscheidungsregel folgende Gestalt haben:

- (ER) Seien  $n$  der Stichprobenumfang und  $P_1, \dots, P_n$  die gemessenen Leistungen der Stichprobenmodule. Dann akzeptiert der Betreiber der PV-Anlage die Lieferung genau dann, wenn die Anzahl der nicht tolerierbaren Module einen bestimmten Wert  $c$  nicht übersteigt, d. h. wenn

$$T_n := \sum_{i=1}^n \mathbf{1}_{P_i \notin I_{\mu_0}} \leq c.$$

Die konkrete Wahl von  $c$  und  $n$  ist dabei der entscheidende Punkt. Soll die Entscheidungsregel für einen gegebenen Qualitätsanspruch ein bestimmtes Konfidenzlevel gewährleisten (vgl. Abschnitt 1.1 (a) und (b)), dann muss man hier sehr genau sein und einige Arbeit investieren. Diesen Sachverhalt behandeln wir nun in Abschnitt 2.2 ausführlich. Das Paar  $(c, n)$  nennen wir dabei auch Charakteristik der Entscheidungsregel (ER).

## 2.2 Wahl der Charakteristik der Entscheidungsregel

Bevor wir gleich zur Analyse des Modells kommen, sei an dieser Stelle herausgehoben, dass man im aktuellen Modell nicht mit außerordentlich kleinen Mindeststichprobenumfängen rechnen kann. Dies ist eine einfache Konsequenz daraus, dass der Konsument im Prinzip keinerlei Informationen über die Verteilung der Leistungswerte der Solarmodule hat. In unserem zweiten Modell (Abschnitt 3) wird sich die Situation bessern.

Es sei  $N$  die Anzahl der gelieferten Solarmodule. Modellannahme (M3) entsprechend klassifizieren wir diese  $N$  Module in tolerierbare und nicht tolerierbare Module. Zur Wiederholung: Ein Modul ist genau dann tolerierbar, wenn seine Leistung in einem vorgegebenen Toleranzbereich  $I_{\mu_0}$  liegt. Die unbekannte Anzahl der *nicht* tolerierbaren Module in der Lieferung sei  $N_1$ . Der relative Anteil der nicht tolerierbaren Module ist also

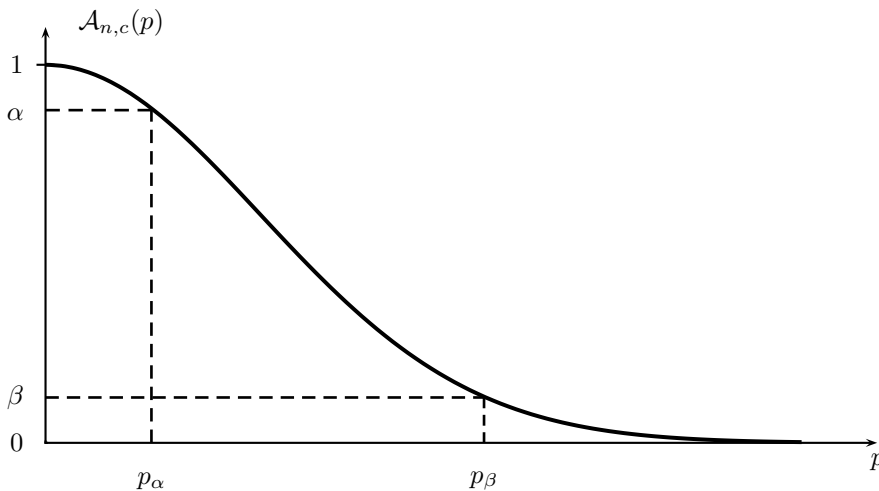
$$p := \frac{N_1}{N}.$$

Dieser Wert ist natürlich ebenfalls unbekannt. Nun wird eine Stichprobe mit Umfang  $n$  aus der gesamten Charge entnommen und geprüft ( $n \leq N$ ). Dabei bezeichne  $T_n$  die Anzahl der nicht tolerierbaren Module in dieser Stichprobe. Es gilt  $T_n \leq \min\{n, N_1\} \leq N$ . Laut der Entscheidungsregel (ER) akzeptiert der Betreiber der PV-Anlage die Charge genau dann, wenn  $T_n \leq c$  für einen Schwellenwert  $c \in \mathbb{N}$ . Der entscheidende Punkt ist nun die Wahl der Werte von  $c$  und  $n$ .

Wie in Abschnitt 1.1 (a) und (b) erläutert, sollte die Entscheidungsregel für einen gegebenen Qualitätsanspruch ein gewisses Konfidenzlevel garantieren. Um diesen Sachverhalt etwas mathematischer zu formulieren, führen wir zunächst folgende Funktion ein:

$$\mathcal{A}_{n,c}(p) := \mathbb{P}_p[T_n \leq c], \quad p \in [0, 1]. \quad (1)$$

Dabei ist  $\mathbb{P}_p[T_n \leq c]$  die Wahrscheinlichkeit des Ereignisses  $\{T_n \leq c\}$ , also der Akzeptanz der Charge, unter der Annahme, dass  $p$  der wahre relative Anteil der nicht tolerierbaren Module an der gesamten Charge ist. Wir nennen die Funktion  $p \mapsto \mathcal{A}_{n,c}(p)$  daher die Akzeptanzfunktion der Entscheidungsregel (ER). Es ist intuitiv klar und leicht zu beweisen, dass  $p \mapsto \mathcal{A}_{n,c}(p)$  monoton fallend ist. Folgende Abbildung zeigt den typischen Graphen.



Spezifizieren wir nun den Begriff „Qualitätsanspruch“. Wir sagen, eine Charge hat eine „hohe Qualität“, wenn der Anteil der nicht tolerierbaren Module an der gesamten Charge kleiner oder gleich einem vorgegebenen  $p_\alpha \in [0, 1]$  ist. Ferner reden wir von „geringer Qualität“ der Charge, wenn der genannte Anteil größer oder gleich einem vorgegebenen  $p_\beta \in [0, 1]$  ist. Natürlich macht dann nur die Konstellation  $p_\alpha < p_\beta$  Sinn. Insbesondere sollte neben  $p_\alpha$  auch  $p_\beta$  klein sein. Aus technischen Gründen fordern wir aber  $p_\beta - p_\alpha > 1/N$ . Unsere Notion von hoher und niedriger Qualität konkretisiert (a) und (b) in Abschnitt 1.1. Wie bereits erwähnt, sollte das in (a) eingeführte  $\alpha$  möglichst groß und das in (b) eingeführte  $\beta$  möglichst klein sein, zumindest aber  $\alpha > 1/2 > \beta$ . Im aktuellen Modell können die Bedingungen (a) und (b) nun auch wie folgt ausgedrückt werden:

$$(A) \mathcal{A}_{n,c}(p_\alpha) \geq \alpha$$

$$(B) \mathcal{A}_{n,c}(p_\beta) \leq \beta$$

Die Werte  $p_\alpha$  und  $p_\beta$  legen den Qualitätsanspruch fest,  $\alpha$  und  $\beta$  determinieren das Konfidenzlevel zu  $p_\alpha$  und  $p_\beta$ . Auf Grund der Monotonie von  $p \mapsto L_{n,c}(p)$  gelten (A) und (B) auch für alle  $p \leq p_\alpha$  bzw.  $p \geq p_\beta$  an Stelle von  $p_\alpha$  bzw.  $p_\beta$ .

Gesucht sind nun Paare  $(c, n)$ , für die die Bedingungen (A) und (B) erfüllt sind. Genauer suchen wir dasjenige Paar mit dem kleinsten  $n$ -Wert.

In Vorschrift 2.2 weiter unten ist beschrieben, wie man die Charakteristik  $(c, n)$  der Entscheidungsregel (ER) für gegebene  $p_\alpha$ ,  $\alpha$ ,  $p_\beta$  und  $\beta$  bestimmen kann. Auf dem Weg zu dieser Vorschrift sollte man offensichtlich zuerst versuchen,  $\mathcal{A}_{n,c}(p)$  für kleine  $p$  zu bestimmen. Diesbzgl. notieren wir folgende Bemerkung.

**Bemerkung 2.1** *Unter den Voraussetzungen*

$$n \leq N/10 \quad \text{und} \quad p := N_1/N \leq 0.1 \tag{2}$$

*gilt folgende gute Näherung*

$$\boxed{\mathcal{A}_{n,c}(p) \approx 1 - G_{2(c+1)}(2pn)} \tag{3}$$

*Hierbei ist  $G_{2(c+1)}$  die Verteilungsfunktion der  $\chi^2$ -Verteilung mit Freiheitsgrad  $2(c+1)$ :*

$$G_{2(c+1)}(x) = \int_0^x \frac{1}{2^{c+1}c!} y^c e^{-y/2} dy, \quad x \geq 0. \tag{4}$$

Die zweite Forderung in (2) ist bei der Produktion von Solarmodulen üblicherweise erfüllt. Eine Auswertung der Daten aus **Beispieldatensatz-2** ergibt für ein Toleranzintervall  $I_{\mu_0} = [\mu_0 - \epsilon, \mu_0 + \epsilon]$  mit  $\epsilon = 0.05\mu_0$  folgende Werte für  $p = N_1/N$ :

Produkttyp	SE160a	SE160b	SQ165	P220	7185	11007	I-155
$p = N_1/N$	0	0.015	0.007	0.008	0	0.466	0

**Rechtfertigung von Bemerkung 2.1** Zunächst stellen wir fest, dass das Zufallsexperiment, auf dem die Entscheidungsregel (ER) basiert, zu der hypergeometrischen Verteilung  $H_{n;N_1,N}$  mit den Parametern  $n$ ,  $N_1$  und  $N$  korrespondiert („Ziehen ohne Zurücklegen“):

$$H_{n;N_1,N}(\{k\}) := \frac{\binom{N_1}{k} \binom{N - N_1}{n - k}}{\binom{N}{n}}$$

ist die Wahrscheinlichkeit, dass unter den  $n$  gleichzeitig gezogenen Stichprobenmodulen genau  $k$  nicht tolerierbare Module sind. Nun hat die hypergeometrische Verteilung den Nachteil, dass ihre relativ komplizierte Zähldichte analytische Betrachtungen ziemlich erschwert. Andererseits ist aus der Wahrscheinlichkeitstheorie bekannt, dass die hypergeometrische Verteilung für große  $N$  mit einer Binomialverteilung („Ziehen mit Zurücklegen“) approximiert werden kann. Genauer:

$$\lim_{N \rightarrow \infty} H_{n;N_1,N}(\{k\}) = B_{n,p}(\{k\}) \quad \forall k = 0, \dots, n$$

für  $n$  und  $p$  fest, wobei  $N_1 := Np$ . Dabei ist

$$B_{n,p}(\{k\}) := \binom{n}{k} p^k (1 - p)^{n-k}$$

die Wahrscheinlichkeit, dass unter  $n$  nacheinander (mit „Zurücklegen“ !!!) gezogenen Stichprobenmodulen genau  $k$  nicht tolerierbare Module sind. Ein numerischer Vergleich mit Hilfe von Tabellen der hypergeometrischen Verteilung und der Binomialverteilung zeigt, dass die Approximation hinreichend genau ist für  $n \leq N/10$ . Intuitiv kann man sich vorstellen, dass es bei großen Chargenumfängen  $N$  (relativ zur Stichprobengröße  $n$ ) im Prinzip egal ist, ob man getestete Module wieder „zurücklegt“ (und erneut ziehen könnte) oder nicht. Anstelle einer hypergeometrischen Verteilung  $H_{n;N_1,N}$  mit  $n \leq N/10$  und  $N_1 := Np$  kann man approximativ also mit der Binomialverteilung  $B_{n,p}$  arbeiten:

$$H_{n;N_1,N}(\{k\}) \approx B_{n,p}(\{k\}) \quad \forall k = 0, \dots, n. \quad (5)$$

Die Binomialverteilung lässt sich ihrerseits mit Hilfe der Poissonverteilung approximieren:

$$\lim_{n \rightarrow \infty} B_{n,p_n}(\{k\}) = \Pi_\lambda(\{k\}) := \frac{\lambda^k}{k!} e^{-\lambda} \quad \forall k = 0, 1, \dots$$

für Folgen  $(p_n)$ , für die  $\lim_{n \rightarrow \infty} np_n = \lambda$ . Ein numerischer Vergleich mit Hilfe von Tabellen der Binomialverteilung und der Poissonverteilung zeigt, dass die Approximation schon sehr genau ist für  $p \leq 0.1$ . Anstelle einer Binomialverteilung  $B_{n,p}$  mit  $p \leq 0.1$  kann man approximativ also mit der Poissonverteilung  $\Pi_\lambda$  mit  $\lambda := np$  arbeiten:

$$B_{n,p}(\{k\}) \approx \Pi_{np}(\{k\}) \quad \forall k = 0, \dots, n. \quad (6)$$

Aus (5) und (6) schließen wir, dass unter der Voraussetzung (2) als gute Näherung gilt:

$$H_{n;N_1,N}(\{k\}) \approx \Pi_{np}(\{k\}) \quad \forall k = 0, \dots, n. \quad (7)$$



Somit ist es unter Voraussetzung (2) gerechtfertigt, die Akzeptanzfunktion aus (1) mit Hilfe der Poissonverteilung zu spezifizieren:

$$\mathcal{A}_{n,c}(p) \approx \Pi_{np}(\{0, \dots, c\}) \left( = \sum_{k=0}^c \frac{(np)^k}{k!} e^{-np} \right), \quad (8)$$

zumindest für Argumente  $p \leq 0.1$ . Die rechte Seite von (8) stimmt nun aber gerade mit der rechten Seite von (3) überein, was eine Konsequenz von folgender Gleichung ist

$$\sum_{k=0}^c \frac{x^k}{k!} e^{-x} = 1 - \int_0^{2x} \frac{1}{2^{c+1} c!} y^c e^{-y/2} dy.$$

Letztere kann bewiesen werden durch Differenzieren auf beiden Seiten. Somit haben wir Bemerkung 2.1 gerechtfertigt.  $\square$

Unter Verwendung von (3) erhalten wir die Gültigkeit von (A) und (B) genau dann, wenn

$$\frac{1}{2p_\beta} G_{2(c+1)}^{-1}(1 - \beta) \leq n \leq \frac{1}{2p_\alpha} G_{2(c+1)}^{-1}(1 - \alpha). \quad (9)$$

Die Aufgabe ist nun also, ein  $c$  und ein möglichst kleines  $n$  zu finden, so dass (9) gilt. Da  $m \mapsto G_m^{-1}(\gamma)$  für jedes  $\gamma$  monoton wachsend ist, erhält man ein möglichst kleines  $n$ , wenn man ein möglichst kleines  $c$  so wählt, dass sich (9) gerade noch mit einer geeigneten natürlichen Zahl  $n$  erfüllen lässt. Nun scheint dies auf Grund der komplizierten Form von  $G_{2(c+1)}$  auf den ersten Blick sehr aufwendig zu sein. Die Situation ist aber besser als befürchtet, da die Quantile von  $\chi^2$ -Verteilungen sehr gut tabelliert sind; siehe Abschnitt 7.2. Somit können wir mittels folgender Vorschrift die gesuchten  $c$  und  $n$  bestimmen.

**Vorschrift 2.2** Setze für  $c$  nacheinander die natürlichen Zahlen  $0, 1, 2, \dots$  ein und ermittle mit Hilfe einer Tafel der  $\chi^2$ -Verteilung das kleinste  $c$ , zu dem es eine natürliche Zahl  $n_c > c$  gibt, die der Bedingung (9) mit  $n_c$  an Stelle von  $n$  genügt. Definieren wir  $n$  als die kleinste solche natürliche Zahl  $n_c$  und erfüllt dieses  $n$  die erste Bedingung in (2), d. h.  $n \leq N/10$ , dann haben wir eine geeignete Charakteristik  $(c, n)$  für die Entscheidungsregel (ER) gefunden.

Für gegebene  $p_\alpha, \alpha, p_\beta$  und  $\beta$  kann man mit Hilfe dieser Vorschrift also eine Charakteristik  $(c, n)$  der Entscheidungsregel (ER) bestimmen, die geeignet ist, wenn die Produktionscharge einen Umfang  $N$  von mindestens  $10n$  hat. Insbesondere hängt die Charakteristik  $(c, n)$  nur indirekt von  $N$  ab, nämlich über die letztgenannte Bedingung. In folgender Tabelle sind die Charakteristiken für typische Konstellationen von  $p_\alpha, p_\beta, \alpha$  und  $\beta$  angegeben.

$p_\alpha$	$\alpha$	$p_\beta$	$\beta$	$c$	$n$	geeignet, wenn $N \geq$
0.01	0.90	0.03	0.10	5	310	3100
0.01	0.90	0.05	0.10	2	107	1070

### 3 Szenario ohne Flasherliste (Normalverteilungsannahme)

#### 3.1 Modellbeschreibung

Wir machen folgende Modellannahmen:

- (M1) Bei der Produktion der Solarmodule sind die Leistungen der einzelnen Module unabhängig voneinander und identisch verteilt. Die Verteilung ist gegeben durch die Normalverteilung  $N(\mu, \sigma^2)$  mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$ . Allerdings sind  $\mu$  und  $\sigma^2$  unbekannt.
- (M2) Der Produzent der Solarmodule quantifiziert die Nennleistung auf  $\mu_0$ . Wie er auf diesen Wert kommt und ob dieser gerechtfertigt ist, spielt dabei keine Rolle.
- (M3) Ein Modul gilt als „schlecht“, wenn seine Leistung  $P$  die Ungleichung  $P < \mu_0 - \epsilon$  erfüllt. Die Festlegung der Toleranz  $\epsilon$  basiert dabei wieder auf der (subjektiven) Einschätzung des Konsumenten und muss bei Vertragsabschluss ggf. mit dem Hersteller abgestimmt worden sein.

Die Normalverteilungsannahme in (M1) kann in vielen Fällen sicher vertreten werden, in einigen Fällen jedoch eher nicht. Z. B. liefert R bzgl. der Daten aus **Beispieldaten** für den Shapiro-Wilk-Test auf Normalverteilung einen p-Wert von 0.429, bzgl. der Daten aus **Beispieldaten-2** (SE160a) einen p-Wert von 0.026. Zum Niveau 0.05 kann im ersten Fall die Normalverteilungsannahme also vertreten werden, im zweiten Fall hingegen nicht.

Laut (M3) ist ein Solarmodul also genau dann „gut“, wenn seine Leistung nicht um mehr als  $\epsilon$  nach unten von der Nennleistung  $\mu_0$  abweicht. In Bezug auf die Verteilung der Solarmodule heißt das: Eine Produktionscharge ist genau dann „akzeptabel“, wenn der tatsächliche (unbekannte) Erwartungswert  $\mu$  nicht um mehr als  $\epsilon$  nach unten von der Nennleistung  $\mu_0$  abweicht und die einzelnen Leistungen nicht „zu stark“ (relativ zu der Diskrepanz  $\mu - (\mu_0 - \epsilon) > 0$ ) um  $\mu$  streuen; letzteres bedeutet, dass die (unbekannte) Standardabweichung  $\sigma$  nicht „zu groß“ ist. Nun repräsentiert

$$\bar{P}_n := \frac{1}{n} \sum_{i=1}^n P_i \quad (10)$$

den besten Schätzer für den Erwartungswert  $\mu$  der Normalverteilung, wenn  $P_1, \dots, P_n$  die Leistungen der Stichprobenmodule angeben. Entsprechend ist  $\bar{P}_n$  eine geeignete Größe zur Evaluierung der Charge. Eine Charge könnte man genau dann als „akzeptabel“ betrachten, wenn  $\bar{P}_n$  oberhalb von

$$(\mu_0 - \epsilon) + c \frac{\sigma}{\sqrt{n}}$$

liegt, wobei  $\sigma/\sqrt{n}$  die Standardabweichung von  $\bar{P}_n$  angibt und der Faktor  $c$  noch geeignet gewählt werden muss. Das Problem ist allerdings, dass  $\sigma$  ebenso wie  $\mu$  unbekannt ist. Der Ausweg ist das Ersetzen von  $\sigma$  durch seinen besten Schätzer

$$S_n := \sqrt{\frac{1}{n-1} \sum_{i=1}^n (P_i - \bar{P}_n)^2} . \quad (11)$$

Dies rechtfertigt die Wahl folgender Entscheidungsregel.

(ER) Der Betreiber der PV-Anlage akzeptiert die Lieferung genau dann, wenn für einen bestimmten Wert  $c$ ,

$$T_n := \frac{\bar{P}_n - (\mu_0 - \epsilon)}{S_n} \sqrt{n} \geq c. \quad (12)$$

Wie im ersten Modell (Abschnitt 2) wird die konkrete Wahl der Charakteristik  $(c, n)$  der Entscheidungsregel dabei die entscheidende Rolle spielen. Die Spezifizierung von  $(c, n)$  hängt dabei sowohl von  $\mu_0$  und  $\epsilon$  als auch von dem noch zu definierenden Qualitätsanspruch und seinem Konfidenzlevel ab. Mit dieser Problematik beschäftigen wir uns nun in Abschnitt 3.2 im Detail.

### 3.2 Wahl der Charakteristik der Entscheidungsregel

Basis für die Entscheidung über Akzeptanz oder Reklamation der Charge ist eine Stichprobe mit Umfang  $n$ . Laut der Regel (ER) akzeptiert der Betreiber der PV-Anlage die Charge genau dann, wenn (12) erfüllt ist. Wie im ersten Modell soll die Entscheidungsregel für einen gegebenen Qualitätsanspruch ein gewisses Konfidenzlevel garantieren. Für die mathematisch exakte Formulierung dieses Sachverhaltes verwenden wir wieder eine sogenannte Akzeptanzfunktion, siehe (A) und (B) weiter unten.

In unserem ersten Modell in Abschnitt 2 war die Verteilung der tolerierbaren Module von den Parametern  $N_1$  und  $N$  abhängig. Die Akzeptanzfunktion wurde mit Hilfe eines Approximationsargumentes aber als eine Funktion von nur einem Argument, nämlich  $p$  betrachtet. Im aktuellen Modell wird die grundlegende Verteilung ebenfalls durch zwei Parameter, nämlich  $\mu$  und  $\sigma^2$  bestimmt. Wir wollen jedoch ähnlich wie im ersten Modell fortfahren und versuchen somit, die Akzeptanzwahrscheinlichkeit in Abhängigkeit von nur einem Parameter darzustellen. Die Rolle dieses Parameters wird von

$$p := \mathbb{P}_{\mu, \sigma^2}[P < \mu_0 - \epsilon] \quad (13)$$

gespielt. Hierbei bezeichnet  $\mathbb{P}_{\mu, \sigma^2}[P < \mu_0 - \epsilon]$  die Wahrscheinlichkeit, dass ein Modul „schlecht“ ist, unter der Annahme, dass  $\mu$  und  $\sigma^2$  die wahren Parameter sind. Der Parameter  $p$  gibt gerade den erwarteten relativen Anteil von „schlechten“ Modulen (im Sinne von (M3)) an.<sup>1</sup> Offensichtlich gilt für  $p$

$$p = \mathbb{P}_{\mu, \sigma^2} \left[ \frac{P - \mu}{\sigma} < \frac{\mu_0 - \epsilon - \mu}{\sigma} \right] = \Phi \left( \frac{\mu_0 - \epsilon - \mu}{\sigma} \right), \quad (14)$$

wobei  $\Phi$  die Verteilungsfunktion der Standardnormalverteilung ist. Nun kann die Testgröße  $T_n$  aus (12) mit Hilfe von (14) wie folgt umgeformt werden:

$$\begin{aligned} T_n &= \frac{\bar{P}_n - (\mu_0 - \epsilon)}{S_n} \sqrt{n} \\ &= \frac{\bar{P}_n - \mu + (\mu - (\mu_0 - \epsilon))}{S_n} \sqrt{n} \\ &= \frac{\bar{P}_n - \mu - \frac{\mu_0 - \epsilon - \mu}{\sigma} \sigma}{S_n} \sqrt{n} \\ &= \frac{(\bar{P}_n - \mu) \sqrt{n} - (\Phi^{-1}(p) \sqrt{n}) \sigma}{S_n}. \end{aligned}$$

---

<sup>1</sup> $\mathbb{E}_{\mu, \sigma^2} \left[ \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{P_i < \mu_0 - \epsilon} \right] = \frac{1}{n} n \mathbb{E}_{\mu, \sigma^2} [\mathbf{1}_{P_i < \mu_0 - \epsilon}] = \mathbb{P}_{\mu, \sigma^2}[P < \mu_0 - \epsilon].$

Aus der Wahrscheinlichkeitstheorie ist bekannt, dass der letzte Term unter  $\mathbb{P}_{\mu,\sigma^2}$  nach der nichtzentralen  $t$ -Verteilung  $t_{n-1,\delta(p)}$  mit Freiheitsgrad  $n - 1$  und Nichtzentralitäts-Parameter

$$\delta(p) := -\Phi^{-1}(p)\sqrt{n}$$

verteilt ist. Die Verteilung von  $T_n$  hängt also in dieser Darstellung nicht explizit von  $\mu$  und  $\sigma$  ab, sondern nur implizit via  $p$ . Die Akzeptanzwahrscheinlichkeit  $\mathbb{P}_{\mu,\sigma^2}[T_n \geq c]$  unsere Entscheidungsregel (ER) kann somit in der Tat als Funktion in nur einer Variablen,  $p$ , ausgedrückt werden:

$$\boxed{\mathcal{A}_{n,c}(p) \left( = \mathbb{P}_{\mu,\sigma^2}[T_n \geq c] \right) = 1 - F_{n-1,\delta(p)}(c)}, \quad p \in [0, 1]. \quad (15)$$

Dabei bezeichnet  $F_{n-1,\delta(p)}$  die Verteilungsfunktion von  $t_{n-1,\delta(p)}$ :

$$F_{n-1,\delta(p)}(x) := \int_{-\infty}^x f_{n-1,\delta(p)}(t) dt \quad \text{mit} \quad (16)$$

$$f_{n-1,\delta(p)}(t) := \frac{1}{2^{n/2}\Gamma((n-1)/2)\sqrt{\pi(n-1)}} \int_0^\infty y^{n/2} e^{-y/2 - (t\sqrt{y/(n-1)} - \delta(p))^2/2} dy.$$

Es sei noch bemerkt, dass  $\mathcal{A}_{n,c}(p)$  genauer die Wahrscheinlichkeit für die Akzeptanz der Charge angibt, wenn  $\mu$  und  $\sigma^2$  die wahren Parameter der Normalverteilung sind und  $p$  wie in (13) bzw. (14) gegeben ist. Umso größer der erwarteten Anteil  $p$  von „schlechten“ Modulen ist, desto weniger wahrscheinlich sollte die Annahme der Charge für feste  $c$  und  $n$  sein. Es ist somit intuitiv klar und kann leicht bewiesen werden, dass  $p \mapsto \mathcal{A}_{n,c}(p)$  monoton fallend ist; siehe Abschnitt 2.2 für den typischen Verlauf des Graphen von  $p \mapsto \mathcal{A}_{n,c}(p)$ .

Die Begriffe „Qualitätsanspruch“ und „Konfidenzlevel“ können nun ganz analog zu unserem ersten Modell (Abschnitt 2.2) verstanden werden. Wir sagen, eine Charge hat eine „hohe Qualität“, wenn der erwartete Anteil von „schlechten“ Modulen kleiner oder gleich einem vorgegebenen  $p_\alpha \in [0, 1]$  ist. Ferner reden wir von „geringer Qualität“ der Charge, wenn der genannte erwartete Anteil größer oder gleich einem vorgegebenen  $p_\beta \in [0, 1]$  ist. Sinnvollerweise nehmen wir wieder  $p_\alpha < p_\beta$  an. Unsere Notion von hoher und niedriger Qualität konkretisiert (a) und (b) in Abschnitt 1.1. Wie bereits erwähnt, sollte das in (a) eingeführte  $\alpha$  möglichst groß und das in (b) eingeführte  $\beta$  möglichst klein sein, zumindest aber  $\alpha > 1/2 > \beta$ . Im aktuellen Modell können die Bedingungen (a) und (b) nun wieder wie folgt ausgedrückt werden:

$$(A) \mathcal{A}_{n,c}(p_\alpha) \geq \alpha$$

$$(B) \mathcal{A}_{n,c}(p_\beta) \leq \beta$$

Die Werte  $p_\alpha$  und  $p_\beta$  legen den Qualitätsanspruch fest,  $\alpha$  und  $\beta$  determinieren das Konfidenzlevel zu  $p_\alpha$  und  $p_\beta$ . Auf Grund der Monotonie von  $p \mapsto \mathcal{A}_{n,c}(p)$  gelten (A) und (B) auch für alle  $p \leq p_\alpha$  bzw.  $p \geq p_\beta$  an Stelle von  $p_\alpha$  bzw.  $p_\beta$ .

Gesucht sind nun wieder Paare  $(c, n)$  mit möglichst kleinem  $n$ -Wert, für die die Bedingungen (A) und (B) erfüllt sind.

Aus analytischen Gründen beschränken wir uns auf den Fall  $\beta = 1 - \alpha$ . In folgender Vorschrift ist beschrieben, wie man die Charakteristik  $(c, n)$  der Entscheidungsregel (ER) für gegebene  $p_\alpha$ ,  $\alpha$ ,  $p_\beta$  (und  $\beta = 1 - \alpha$ ) bestimmen kann.

**Vorschrift 3.1** Für  $\alpha > 1/2$  und  $\beta = 1 - \alpha$  sind

$$c = -[\Phi^{-1}(p_\alpha) + \Phi^{-1}(p_\beta)]\sqrt{n}/2 \quad (17)$$

und

$$n = \left\lceil \frac{4 \Phi^{-1}(\alpha)^2 \{1 + [\Phi^{-1}(p_\alpha) + \Phi^{-1}(p_\beta)]^2/8\}}{[\Phi^{-1}(p_\beta) - \Phi^{-1}(p_\alpha)]^2} \right\rceil \quad (18)$$

gute Näherungen für die Charakteristik der Entscheidungsregel (ER). Um ganz sicher zu gehen, dass (A) und (B) auch wirklich (hinreichend genau) erfüllt sind, kann man  $c$  und  $n$  aus (17) respektive (18) in die jeweils linken Seiten von

$$1 - F_{n-1, \delta(p_\alpha)}(c) \geq \alpha \quad \text{und} \quad 1 - F_{n-1, \delta(p_\beta)}(c) \leq \beta \quad (19)$$

einsetzen und checken, ob diese zu (A) und (B) äquivalenten Ungleichungen (hinreichend genau) erfüllt sind.

Die Verteilungsfunktion  $\Phi$  der Standardnormalverteilung ist in der Literatur sehr gut tabelliert, siehe auch Abschnitt 7.1. Tabellen für die Verteilungsfunktionen  $F_{\nu, \delta}$  von nicht-zentralen  $t$ -Verteilung findet man hingegen eher selten, z. B. in [2], Seiten 113-116. Seite 209 in [1] entnehmen wir aber folgende Näherung

$$F_{\nu, \delta}^{-1}(\alpha) \approx \frac{\delta + \Phi^{-1}(\alpha) \sqrt{1 + \frac{1}{2}[\delta^2 - \Phi^{-1}(\alpha)^2]/\nu}}{1 - \frac{1}{2}\Phi^{-1}(\alpha)^2/\nu}.$$

**Rechtfertigung von Vorschrift 3.1** Wegen (15) sind (A) und (B) gleichwertig zu (19). Die sehr komplizierte Gestalt in (16) lässt schnell erahnen, dass es schier unmöglich ist,  $F_{n-1, \delta(p_\alpha)}$  nach  $c$  und  $n$  aufzulösen und dadurch diese Bedingungen zu vereinfachen. Wir suchen uns einen Ausweg. Zunächst notieren wir

$$\begin{aligned} \mathcal{A}_{n,c}(p) &= \mathbb{P}_{\mu, \sigma^2}[T_n \geq c] \\ &= \mathbb{P}_{\mu, \sigma^2} \left[ \frac{\bar{P}_n - (\mu_0 - \epsilon)}{S_n} \sqrt{n} \geq c \right] \\ &= \mathbb{P}_{\mu, \sigma^2} \left[ \frac{\bar{P}_n - (\mu_0 - \epsilon)}{\sigma} \sqrt{n} \geq c \frac{S_n}{\sigma} \right] \\ &= \mathbb{P}_{\mu, \sigma^2} \left[ \frac{\bar{P}_n - \mu}{\sigma} \sqrt{n} - c \frac{S_n}{\sigma} \geq -\frac{\mu - (\mu_0 - \epsilon)}{\sigma} \sqrt{n} \right] \\ &= \mathbb{P}_{\mu, \sigma^2} \left[ \frac{\bar{P}_n - \mu}{\sigma} \sqrt{n} - c \frac{S_n}{\sigma} \geq \frac{\mu_0 - \epsilon - \mu}{\sigma} \sqrt{n} \right] \\ &= \mathbb{P}_{\mu, \sigma^2} \left[ \frac{\bar{P}_n - \mu}{\sigma} \sqrt{n} - c \frac{S_n}{\sigma} \geq \Phi^{-1}(p) \sqrt{n} \right], \end{aligned}$$

wobei der letzten Schritt aus (14) folgt. Aus der Wahrscheinlichkeitstheorie ist bekannt, dass  $\mathbb{E}_{\mu, \sigma^2}[S_n/\sigma] = \gamma_n$  und  $\text{Var}_{\mu, \sigma^2}[S_n/\sigma] = 1 - \gamma_n^2$ , wobei

$$\gamma_n := \sqrt{\frac{2}{n-1}} \frac{\Gamma(n/2)}{\Gamma((n-2)/2)}.$$

Es gilt

$$\frac{2n-3}{2n-2} \leq \gamma_n^2 \leq \frac{2n-2}{2n-1} \quad \text{und} \quad \frac{1}{2n-1} \leq 1 - \gamma_n^2 \leq \frac{1}{2n-2},$$

also  $\gamma_n \approx 1$  und  $1 - \gamma_n^2 \approx 1/2n$ . Für nicht zu kleine  $n$  ist  $cS_n/\sigma$  also eine Zufallsvariable mit sehr kleiner Streuung. Ferner ist  $((\bar{P}_n - \mu)/\sigma)\sqrt{n}$  unter  $\mathbb{P}_{\mu,\sigma^2}$  offensichtlich standard normalverteilt. Somit ist es approximativ gerechtfertigt, die Zufallsvariable

$$\frac{\bar{P}_n - \mu}{\sigma} \sqrt{n} - c \frac{S_n}{\sigma}$$

als normalverteilt anzusehen, und zwar mit ihrem tatsächlichen Erwartungswert  $-c\gamma_n$  und ihrer tatsächlichen Varianz  $1 + c^2(1 - \gamma_n^2)$ . Somit erhalten wir

$$\mathcal{A}_{n,c}(p) \approx 1 - \Phi\left(\frac{\Phi^{-1}(p)\sqrt{n} + c\gamma_n}{\sqrt{1 + c^2(1 - \gamma_n^2)}}\right).$$

Um diesen Ausdruck noch zu vereinfachen, ersetzen wir  $\gamma_n$  und  $1 - \gamma_n^2$  durch die sehr guten Näherungen 1 bzw.  $1/2n$  und erhalten

$$\mathcal{A}_{n,c}(p) \approx 1 - \Phi\left(\frac{\Phi^{-1}(p)\sqrt{n} + c}{\sqrt{1 + c^2/(2n)}}\right). \quad (20)$$

Hiermit kann man nun ein  $c$  und ein möglichst kleines  $n$  so bestimmen, dass die Bedingungen (A) und (B) erfüllt sind. Letztere sind näherungsweise äquivalent zu

$$(A)' \quad \Phi^{-1}(1 - \alpha)\sqrt{1 + c^2/(2n)} \geq \Phi^{-1}(p_\alpha)\sqrt{n} + c,$$

$$(B)' \quad \Phi^{-1}(1 - \beta)\sqrt{1 + c^2/(2n)} \leq \Phi^{-1}(p_\beta)\sqrt{n} + c.$$

Für ein möglichst kleines  $n$  sollten diese Ungleichungen „gerade noch so“ erfüllt sein, d. h.

$$(A)'' \quad +\Phi^{-1}(1 - \alpha)\sqrt{1 + c^2/(2n)} \approx \Phi^{-1}(p_\alpha)\sqrt{n} + c,$$

$$(B)'' \quad -\Phi^{-1}(1 - \alpha)\sqrt{1 + c^2/(2n)} \approx \Phi^{-1}(p_\beta)\sqrt{n} + c,$$

wobei wir für (B)'' unserer Konvention  $\beta = 1 - \alpha$  sowie die bekannte Relation  $\Phi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha)$  benutzt haben. Aus (A)'' und (B)'' erhalten wir sofort

$$0 \approx [\Phi^{-1}(p_\alpha)\sqrt{n} + c] + [\Phi^{-1}(p_\beta)\sqrt{n} + c]$$

und somit näherungsweise (17). Mit Hilfe von (20), (17) und  $\Phi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha)$  erreichen wir schließlich auch noch näherungsweise (18).  $\square$

In folgender Tabelle sind die Charakteristiken für typische Konstellationen von  $p_\alpha$ ,  $p_\beta$ ,  $\alpha$  (und  $\beta = 1 - \alpha$ ) angegeben.

$p_\alpha$	$\alpha$	$p_\beta$	$\beta$	$c$	$n$
0.01	0.90	0.03	0.10	21.7	107
0.01	0.90	0.05	0.10	12.9	43

## 4 Szenario mit Flasherliste (Normalverteilungsannahme)

### 4.1 Modellbeschreibung

Wir machen folgende Modellannahmen:

- (M1) Der Produzent der Solarmodule misst die Leistung jedes seiner Module und händigt diese Information dem Konsumenten auf einer sogenannten Flasherliste aus. Wir nehmen an, dass sich die gemessenen Leistungen  $P'_1, \dots, P'_N$  der Module seiner Produktionscharge wie folgt zusammensetzen:

$$P'_i = \Pi_i + \Delta + r'_i, \quad i = 1, \dots, N,$$

wobei  $\Pi_i$  die tatsächliche Leistung des  $i$ -ten Moduls ist,  $\Delta$  eine (möglicherweise gewollte) deterministische Verfälschung der Messungen angibt und  $r'_i$  den Messfehler spezifiziert. Es gelte:

- $\Pi_1, \dots, \Pi_N$  sind unabhängig identisch  $N(\mu, v)$ -verteilt,
- $r'_1, \dots, r'_N$  sind unabhängig identisch  $N(0, w)$ -verteilt
- $\Pi_1, \dots, \Pi_N$  und  $r'_1, \dots, r'_N$  sind unabhängig.

- (M2) Der Konsument (Betreiber der PV-Anlage) führt seinerseits Testmessungen an einer Stichprobe mit Umfang  $n$  durch. Die einzelnen gemessenen Leistungen setzen sich nun so zusammen:

$$P_i = \Pi_i + r_i, \quad i = 1, \dots, n,$$

wobei  $\Pi_i$  wie oben die tatsächliche Leistung des  $i$ -ten Moduls ist und  $r_i$  den Messfehler des Konsumenten spezifiziert.<sup>2</sup> Ferner nehmen wir an:

- $r_1, \dots, r_n$  sind ebenfalls unabhängig identisch  $N(0, w)$ -verteilt
- $r_1, \dots, r_n$  sind unabhängig von  $r'_1, \dots, r'_n$  und  $\Pi_1, \dots, \Pi_n$ .

- (M3) Die Nennleistung der Solarmodule wird auf  $\mu_0$  quantifiziert. Ob dieser Wert gerechtfertigt ist, spielt dabei keine Rolle.
- (M4) Ein Modul gilt als „schlecht“, wenn seine Leistung  $P$  die Ungleichung  $P < \mu_0 - \epsilon$  erfüllt. Die Festlegung der Toleranz  $\epsilon$  basiert dabei wieder auf der (subjektiven) Einschätzung des Konsumenten und muss bei Vertragsabschluss ggf. mit dem Hersteller abgestimmt worden sein.

Es sei bemerkt, dass die Zufallsvariablen  $P'_i - \Delta$  und  $P_i$  die gleiche Verteilung haben. Insbesondere besitzen die  $P'_i$  und die  $P_i$  also die gleiche Varianz  $\sigma^2$ , nämlich  $\sigma^2 = v + w$ .

Laut (M4) ist ein Solarmodul also genau dann „gut“, wenn seine Leistung nicht um mehr als  $\epsilon$  nach unten von der Nennleistung  $\mu_0$  abweicht. Somit sind wir wieder prinzipiell in dem Setting von Kapitel 3. Allerdings haben wir nun den Vorteil, die Standardabweichung  $\sigma$  der  $P_1, \dots, P_n$  als bekannt annehmen zu können. Sie ist zwar nicht wirklich bekannt,

---

<sup>2</sup>Ohne Einschränkung nehmen wir also an, die Module mit Indizes  $1, \dots, n$  sind die Stichprobenmodule.

sie kann aber mit Hilfe der Leistungsdaten  $P'_1, \dots, P'_N$  von der Flasherliste sehr genau geschätzt werden, und zwar durch

$$S'_N := \sqrt{\frac{1}{N-1} \sum_{i=1}^N (P'_i - \bar{P}'_N)^2}, \quad (21)$$

wobei  $\bar{P}'_N := \frac{1}{N} \sum_{i=1}^N P'_i$ . Der entscheidende Punkt ist dabei, dass  $N$  sehr viel größer sein wird als  $n$  und somit  $S'_N$  sehr viel näher an  $\sigma^2$  liegen wird als  $S_n$  aus (11); man beachte, dass  $S_n \rightarrow \sigma^2$  für  $n \rightarrow \infty$ . Wie bereits in Abschnitt 3 ausgeführt, kann man dann eine Charge als „akzeptabel“ betrachten, wenn  $\bar{P}_n$  oberhalb von

$$(\mu_0 - \epsilon) + c \frac{\sigma}{\sqrt{n}}$$

liegt, wobei

$$\sigma = \hat{\sigma} := S'_N$$

angenommen wird und der Faktor  $c$  noch geeignet gewählt werden muss. Unsere Entscheidungsregel hat also wieder folgende Gestalt:

(ER) Der Betreiber der PV-Anlage akzeptiert die Lieferung genau dann, wenn für einen bestimmten Wert  $c$ ,

$$T_n := \frac{\bar{P}_n - (\mu_0 - \epsilon)}{\sigma} \sqrt{n} \geq c. \quad (22)$$

Wie in den beiden ersten Modellen (Abschnitte 2 und 3) wird die konkrete Wahl der Charakteristik  $(c, n)$  der Entscheidungsregel dabei die entscheidende Rolle spielen. Die Spezifizierung von  $(c, n)$  hängt wieder sowohl von  $\mu_0$  und  $\epsilon$  als auch von dem noch zu definierenden Qualitätsanspruch und seinem Konfidenzlevel ab. Mit dieser Problematik beschäftigen wir uns nun in Abschnitt 4.2 im Detail.

## 4.2 Wahl der Charakteristik der Entscheidungsregel

Wir verfahren ganz analog zu Abschnitt 3.2. Die Entscheidungsregel soll für einen gegebenen Qualitätsanspruch wieder ein gewisses Konfidenzlevel garantieren. Und für die mathematisch exakte Formulierung dieses Sachverhaltes verwenden wir einmal mehr eine sogenannte Akzeptanzfunktion, siehe (A) und (B) weiter unten.

Wir versuchen erneut, die Akzeptanzwahrscheinlichkeit in Abhängigkeit von dem Parameter

$$p := \mathbb{P}_\mu[P < \mu_0 - \epsilon] \quad (23)$$

darzustellen. Hierbei bezeichnet  $\mathbb{P}_\mu[P < \mu_0 - \epsilon]$  die Wahrscheinlichkeit, dass ein Modul „schlecht“ ist, unter der Annahme, dass  $\mu$  der wahre Parameter ist. Der Parameter  $p$  gibt gerade den erwarteten relativen Anteil von „schlechten“ Modulen (im Sinne von (M4)) an. Wir haben wieder

$$p = \mathbb{P}_\mu \left[ \frac{P - \mu}{\sigma} < \frac{\mu_0 - \epsilon - \mu}{\sigma} \right] = \Phi \left( \frac{\mu_0 - \epsilon - \mu}{\sigma} \right), \quad (24)$$



wobei  $\Phi$  die Verteilungsfunktion der Standardnormalverteilung ist. Entsprechend gilt auch

$$\mu_0 - \epsilon - \mu = \Phi^{-1}(p)\sigma. \quad (25)$$

Nun ist die Testgröße  $T_n$  aus (22) ganz offensichtlich  $N((\mu - \mu_0 + \epsilon)\frac{\sqrt{n}}{\sigma}, 1)$ -verteilt, d. h. unter Beachtung von (25),

$$T_n \sim N\left(-\Phi^{-1}(p)\sqrt{n}, 1\right).$$

Die Verteilung von  $T_n$  kann also in Abhängigkeit allein von dem Parameter  $p$  angegeben werden. Insbesondere ist es möglich, die Akzeptanzwahrscheinlichkeit  $\mathbb{P}_\mu[T_n \geq c]$  unserer Entscheidungsregel (ER) als Funktion von  $p$  auszudrücken:

$$\boxed{\mathcal{A}_{n,c}(p) \left( = \mathbb{P}_\mu[T_n \geq c] \right) = 1 - \Phi\left(\Phi^{-1}(p)\sqrt{n} + c\right)}, \quad p \in [0, 1]. \quad (26)$$

Es sei noch bemerkt, dass  $\mathcal{A}_{n,c}(p)$  genauer die Wahrscheinlichkeit für die Akzeptanz der Charge angibt, wenn  $\mu$  der wahre Erwartungswert der Normalverteilung ist und  $p$  wie in (23) bzw. (24) gegeben ist. Umso größer der erwartete Anteil  $p$  von „schlechten“ Modulen ist, desto weniger wahrscheinlich sollte die Annahme der Charge für feste  $c$  und  $n$  sein. Es ist somit intuitiv klar und kann leicht bewiesen werden, dass  $p \mapsto \mathcal{A}_{n,c}(p)$  monoton fallend ist; siehe Abschnitt 2.2 für den typischen Verlauf des Graphen von  $p \mapsto \mathcal{A}_{n,c}(p)$ .

Die Begriffe „Qualitätsanspruch“ und „Konfidenzlevel“ können nun wieder ganz analog zu unseren ersten beiden Modellen (Abschnitte 2.2 und 3.2) verstanden werden. Wir sagen, eine Charge hat eine „hohe Qualität“, wenn der erwartete Anteil von „schlechten“ Modulen kleiner oder gleich einem vorgegebenen  $p_\alpha \in [0, 1]$  ist. Ferner reden wir von „geringer Qualität“ der Charge, wenn der genannte erwartete Anteil größer oder gleich einem vorgegebenen  $p_\beta \in [0, 1]$  ist. Sinnvollerweise nehmen wir wieder  $p_\alpha < p_\beta$  an. Unsere Notion von hoher und niedriger Qualität konkretisiert (a) und (b) in Abschnitt 1.1. Wie bereits erwähnt, sollte das in (a) eingeführte  $\alpha$  möglichst groß und das in (b) eingeführte  $\beta$  möglichst klein sein, zumindest aber  $\alpha > 1/2 > \beta$ . Im aktuellen Modell können die Bedingungen (a) und (b) nun wieder wie folgt ausgedrückt werden:

$$(A) \mathcal{A}_{n,c}(p_\alpha) \geq \alpha$$

$$(B) \mathcal{A}_{n,c}(p_\beta) \leq \beta$$

Die Werte  $p_\alpha$  und  $p_\beta$  legen den Qualitätsanspruch fest,  $\alpha$  und  $\beta$  determinieren das Konfidenzlevel zu  $p_\alpha$  und  $p_\beta$ . Auf Grund der Monotonie von  $p \mapsto \mathcal{A}_{n,c}(p)$  gelten (A) und (B) auch für alle  $p \leq p_\alpha$  bzw.  $p \geq p_\beta$  an Stelle von  $p_\alpha$  bzw.  $p_\beta$ .

Gesucht sind nun wieder Paare  $(c, n)$  mit möglichst kleinem  $n$ -Wert, für die die Bedingungen (A) und (B) erfüllt sind.

Aus analytischen Gründen beschränken wir uns auf den Fall  $\beta = 1 - \alpha$ . In folgender Vorschrift ist beschrieben, wie man die Charakteristik  $(c, n)$  der Entscheidungsregel (ER) für gegebene  $p_\alpha$ ,  $\alpha$ ,  $p_\beta$  (und  $\beta = 1 - \alpha$ ) bestimmen kann.

**Vorschrift 4.1** Für  $\alpha > 1/2$  und  $\beta = 1 - \alpha$  bilden

$$c = -[\Phi^{-1}(p_\alpha) + \Phi^{-1}(p_\beta)]\sqrt{n}/2 \quad (27)$$

und

$$n = \left\lceil \frac{4 \Phi^{-1}(\alpha)^2}{[\Phi^{-1}(p_\beta) - \Phi^{-1}(p_\alpha)]^2} \right\rceil \quad (28)$$

die gesuchte Charakteristik der Entscheidungsregel (ER).

**Rechtfertigung von Vorschrift 4.1** Wegen (26) sind (A) und (B) gleichwertig zu

$$1 - \Phi(\Phi^{-1}(p_\alpha)\sqrt{n} + c) \geq \alpha \quad \text{und} \quad 1 - \Phi(\Phi^{-1}(p_\beta)\sqrt{n} + c) \leq \beta. \quad (29)$$

Hiermit kann man nun ein  $c$  und ein möglichst kleines  $n$  so bestimmen, dass die Bedingungen (A) und (B) erfüllt sind. Die Bedingungen in (29) sind äquivalent zu

$$(A)' \quad \Phi^{-1}(1 - \alpha) \geq \Phi^{-1}(p_\alpha)\sqrt{n} + c,$$

$$(B)' \quad \Phi^{-1}(1 - \beta) \leq \Phi^{-1}(p_\beta)\sqrt{n} + c.$$

Für ein möglichst kleines  $n$  sollten diese Ungleichungen „gerade noch so“ erfüllt sein, d. h.

$$(A)'' \quad +\Phi^{-1}(1 - \alpha) \approx \Phi^{-1}(p_\alpha)\sqrt{n} + c,$$

$$(B)'' \quad -\Phi^{-1}(1 - \alpha) \approx \Phi^{-1}(p_\beta)\sqrt{n} + c,$$

wobei wir für (B)'' unserer Konvention  $\beta = 1 - \alpha$  sowie die bekannte Relation  $\Phi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha)$  benutzt haben. Aus (A)'' und (B)'' erhalten wir sofort

$$0 \approx [\Phi^{-1}(p_\alpha)\sqrt{n} + c] + [\Phi^{-1}(p_\beta)\sqrt{n} + c]$$

und somit (27). Mit Hilfe von (26), (27) und  $\Phi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha)$  erreichen wir schließlich auch noch (28).  $\square$

In folgender Tabelle sind die Charakteristiken für typische Konstellationen von  $p_\alpha, p_\beta, \alpha$  (und  $\beta = 1 - \alpha$ ) angegeben.

$p_\alpha$	$\alpha$	$p_\beta$	$\beta$	$c$	$n$
0.01	0.90	0.03	0.10	12.1	34
0.01	0.90	0.05	0.10	7.5	15

## 5 Szenario mit Flasherliste (allgemeine Verteilung)

### 5.1 Modellbeschreibung

Wir machen folgende Modellannahmen:

- (M1) Der Produzent der Solarmodule misst die Leistung jedes seiner Module und händigt diese Information dem Konsumenten auf einer sogenannten Flasherliste aus. Wir nehmen an, dass sich die gemessenen Leistungen  $P'_1, \dots, P'_N$  der Module seiner Produktionscharge wie folgt zusammensetzen:

$$P'_i = \Pi_i + \Delta + r'_i, \quad i = 1, \dots, N,$$

wobei  $\Pi_i$  die tatsächliche Leistung des  $i$ -ten Moduls ist,  $\Delta$  eine (möglicherweise gewollte) deterministische Verfälschung der Messungen angibt und  $r'_i$  den Messfehler spezifiziert. Es gelte:

- $\Pi_1 + r'_1, \dots, \Pi_N + r'_N$  sind unabhängig identisch verteilt nach einer strikt steigenden, stetigen Verteilungsfunktion  $F_\mu$  aus einer Lokationsfamilie

$$\mathbb{F} := \{F_\nu(x) = F(x - \nu) : \nu \in \mathbb{R}\}$$

von Verteilungen mit zweitem Moment, wobei  $\nu$  den Erwartungswert angibt. Insbesondere ist  $F = F_0$  also die „zentrierte Verteilung“ aus  $\mathbb{F}$ .

- (M2) Der Konsument (Betreiber der PV-Anlage) führt seinerseits Testmessungen an einer Stichprobe mit Umfang  $n$  durch. Die einzelnen gemessenen Leistungen setzen sich nun so zusammen:

$$P_i = \Pi_i + r_i, \quad i = 1, \dots, n,$$

wobei  $\Pi_i$  wie oben die tatsächliche Leistung des  $i$ -ten Moduls ist und  $r_i$  den Messfehler des Konsumenten spezifiziert.<sup>3</sup> Ferner nehmen wir an:

- $\Pi_1 + r_1, \dots, \Pi_n + r_n$  sind ebenfalls unabhängig identisch verteilt nach der Verteilungsfunktion  $F_\mu$ .

- (M3) Die Nennleistung der Solarmodule wird auf  $\mu_0$  quantifiziert. Ob dieser Wert gerechtfertigt ist, spielt dabei keine Rolle.

- (M4) Ein Modul gilt als „schlecht“, wenn seine Leistung  $P$  die Ungleichung  $P < \mu_0 - \epsilon$  erfüllt. Die Festlegung der Toleranz  $\epsilon$  basiert dabei wieder auf der (subjektiven) Einschätzung des Konsumenten und muss bei Vertragsabschluss ggf. mit dem Hersteller abgestimmt worden sein.

Es sei bemerkt, dass die Zufallsvariablen  $P'_i$  und  $P_i$  die gleiche Varianz,  $\sigma^2$ , besitzen.

Laut (M4) ist ein Solarmodul also genau dann „gut“, wenn seine Leistung nicht um mehr als  $\epsilon$  nach unten von der Nennleistung  $\mu_0$  abweicht. Somit sind wir wieder prinzipiell in dem Setting der Kapitel 3 und 4. Allerdings liegt diesmal nicht zwingend eine Normalverteilung, sondern eine allgemeine Verteilung  $F$  zugrunde. Wie in Kapitel 4 können wir die

---

<sup>3</sup>Ohne Einschränkung nehmen wir also an, die Module mit Indizes  $1, \dots, n$  sind die Stichprobenmodule.

Standardabweichung  $\sigma$  der  $P_1, \dots, P_n$  als bekannt annehmen. Sie ist zwar nicht wirklich bekannt, sie kann aber mit Hilfe der Leistungsdaten  $P'_1, \dots, P'_N$  von der Flasherliste mittels  $S'_N$  aus (21) sehr genau geschätzt werden. Der entscheidende Punkt ist dabei wieder, dass  $N$  sehr viel größer sein wird als  $n$  und somit  $S'_N$  sehr viel näher an  $\sigma^2$  liegen wird als  $S_n$  aus (11); man beachte, dass  $S_n \rightarrow \sigma^2$  für  $n \rightarrow \infty$ . Wie bereits in Abschnitt 3 ausgeführt, kann man dann eine Charge als „akzeptabel“ betrachten, wenn  $\bar{P}_n$  oberhalb von

$$(\mu_0 - \epsilon) + c \frac{\sigma}{\sqrt{n}}$$

liegt, wobei

$$\sigma = \hat{\sigma} := S'_N$$

angenommen wird und der Faktor  $c$  noch geeignet gewählt werden muss. Unsere Entscheidungsregel hat also wieder folgende Gestalt:

(ER) Der Betreiber der PV-Anlage akzeptiert die Lieferung genau dann, wenn für einen bestimmten Wert  $c$ ,

$$T_n := \frac{\bar{P}_n - (\mu_0 - \epsilon)}{\sigma} \sqrt{n} \geq c. \quad (30)$$

Wie in den vorangegangenen Modellen wird die konkrete Wahl der Charakteristik  $(c, n)$  der Entscheidungsregel dabei die entscheidende Rolle spielen. Die Spezifizierung von  $(c, n)$  hängt neben  $\mu_0$  und  $\epsilon$  erneut auch von dem noch zu definierenden Qualitätsanspruch und seinem Konfidenzlevel ab. Damit beschäftigen wir uns nun im Detail.

## 5.2 Wahl der Charakteristik der Entscheidungsregel

Die Entscheidungsregel soll für einen gegebenen Qualitätsanspruch wieder ein gewisses Konfidenzlevel garantieren. Für die mathematisch exakte Formulierung dieses Sachverhaltes verwenden wir erneut eine sogenannte Akzeptanzfunktion, siehe (A) und (B) weiter unten.

Das erste Ziel ist, die Akzeptanzwahrscheinlichkeit in Abhängigkeit von dem Parameter

$$p := \mathbb{P}_\mu[P < \mu_0 - \epsilon] = F_\mu(\mu_0 - \epsilon) = F(\mu_0 - \epsilon - \mu) \quad (31)$$

darzustellen. Hierbei bezeichnet  $\mathbb{P}_\mu[P < \mu_0 - \epsilon]$  die Wahrscheinlichkeit, dass ein Modul „schlecht“ ist, unter der Annahme, dass  $\mu$  der wahre Parameter ist. Der Parameter  $p$  gibt gerade den erwarteten relativen Anteil von „schlechten“ Modulen (im Sinne von (M4)) an. Wir können  $p$  auch wie folgt ausdrücken

$$p = \mathbb{P}_\mu \left[ \frac{P - \mu}{\sigma} < \frac{\mu_0 - \epsilon - \mu}{\sigma} \right] = G \left( \frac{\mu_0 - \epsilon - \mu}{\sigma} \right), \quad (32)$$

wobei  $G$  die Verteilungsfunktion der standardisierten Variablen  $X_i := (P_i - \mu)/\sigma$  bezeichnet. Es gilt  $G(x) = F(\sigma x)$  für alle  $x \in \mathbb{R}$ . Mit Hilfe von (32) erhalten wir

$$\begin{aligned} \mathbb{P}_\mu[T_n \geq c] &= \mathbb{P}_\mu \left[ \frac{\bar{P}_n - (\mu_0 - \epsilon)}{\sigma} \sqrt{n} \geq c \right] \\ &= \mathbb{P}_\mu \left[ \frac{\bar{P}_n - \mu}{\sigma} \sqrt{n} \geq c - \frac{\mu - \mu_0 + \epsilon}{\sigma} \sqrt{n} \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}_\mu \left[ \frac{\bar{P}_n - \mu}{\sigma} \sqrt{n} \geq c + G^{-1}(p) \sqrt{n} \right] \\
&= \mathbb{P}_\mu \left[ \frac{\bar{P}_n - \mu}{\sigma} \geq \frac{c}{\sqrt{n}} + G^{-1}(p) \right] \\
&= 1 - G^{*n} \left( \frac{c}{\sqrt{n}} + G^{-1}(p) \right).
\end{aligned}$$

Die Verteilung von  $T_n$  kann also in Abhängigkeit allein von dem Parameter  $p$  angegeben werden. Insbesondere ist es möglich, die Akzeptanzwahrscheinlichkeit  $\mathbb{P}_\mu[T_n \geq c]$  unserer Entscheidungsregel (ER) als Funktion von  $p$  auszudrücken:

$$\boxed{\mathcal{A}_{n,c}(p) \left( = \mathbb{P}_\mu[T_n \geq c] \right) = 1 - G^{*n} \left( \frac{c}{\sqrt{n}} + G^{-1}(p) \right)}, \quad p \in [0, 1]. \quad (33)$$

Es sei noch bemerkt, dass  $\mathcal{A}_{n,c}(p)$  genauer die Wahrscheinlichkeit für die Akzeptanz der Charge angibt, wenn  $\mu$  der wahre Erwartungswert ist und  $p$  wie in (32) gegeben ist. Umso größer der erwartete Anteil  $p$  von „schlechten“ Modulen ist, desto weniger wahrscheinlich sollte die Annahme der Charge für feste  $c$  und  $n$  sein. Es ist somit intuitiv klar und kann leicht bewiesen werden, dass  $p \mapsto \mathcal{A}_{n,c}(p)$  monoton fallend ist; siehe Abschnitt 2.2 für den typischen Verlauf des Graphen von  $p \mapsto \mathcal{A}_{n,c}(p)$ .

Die Begriffe „Qualitätsanspruch“ und „Konfidenzlevel“ können nun wieder ganz analog zu unseren vorangegangenen Modellen verstanden und die Bedingungen (a) und (b) in Abschnitt 1.1 wie folgt ausgedrückt werden:

$$(A) \quad \mathcal{A}_{n,c}(p_\alpha) \geq \alpha$$

$$(B) \quad \mathcal{A}_{n,c}(p_\beta) \leq \beta$$

Die Werte  $p_\alpha$  und  $p_\beta$  legen den Qualitätsanspruch fest,  $\alpha$  und  $\beta$  determinieren das Konfidenzlevel zu  $p_\alpha$  und  $p_\beta$ . Auf Grund der Monotonie von  $p \mapsto \mathcal{A}_{n,c}(p)$  gelten (A) und (B) auch für alle  $p \leq p_\alpha$  bzw.  $p \geq p_\beta$  an Stelle von  $p_\alpha$  bzw.  $p_\beta$ .

Gesucht sind nun wieder Paare  $(c, n)$  mit möglichst kleinem  $n$ -Wert, für die die Bedingungen (A) und (B) erfüllt sind.

Aus analytischen Gründen beschränken wir uns erneut auf den Fall  $\beta = 1 - \alpha$ . In folgender Vorschrift ist beschrieben, wie man die Charakteristik  $(c, n)$  der Entscheidungsregel (ER) für gegebene  $p_\alpha$ ,  $\alpha$ ,  $p_\beta$  (und  $\beta = 1 - \alpha$ ) bestimmen kann.

**Vorschrift 5.1** Setze für  $i = 1, \dots, N$

$$X'_i := \frac{P'_i - \bar{P}'_N}{S'_N}$$

und betrachte die empirische Verteilungsfunktion von  $X'_1, \dots, X'_N$

$$G_N(x) := \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(-\infty, x]}(X'_i), \quad x \in \mathbb{R}.$$

Ferner definiere  $G_N^{-1}(y) := \inf\{x \in \mathbb{R} : G(x) = y\}$ . Dann bilden

$$c = -[G_N^{-1}(p_\alpha) + G_N^{-1}(p_\beta)]\sqrt{n}/2 \quad (34)$$

und

$$n = \left\lceil \frac{4 \Phi^{-1}(\alpha)^2}{[G_N^{-1}(p_\beta) - G_N^{-1}(p_\alpha)]^2} \right\rceil \quad (35)$$

die gesuchte Charakteristik der Entscheidungsregel (ER).

**Rechtfertigung von Vorschrift 5.1** Die Form der Akzeptanzfunktion in (33) lässt analytische Betrachtungen kaum zu. Wir suchen uns einen Ausweg. Zunächst erinnern wir uns an

$$\mathcal{A}_{n,c}(p) = \mathbb{P}_\mu \left[ \frac{\bar{P}_n - \mu}{\sigma} \sqrt{n} \geq G^{-1}(p)\sqrt{n} + c \right].$$

Da  $\sqrt{n}(\bar{P}_n - \mu)/\sigma$  nach dem Zentralen Grenzwertsatz approximativ standardnormalverteilt ist, schließen wir

$$\mathcal{A}_{n,c}(p) \approx 1 - \Phi(G^{-1}(p)\sqrt{n} + c). \quad (36)$$

Ferner ist  $G_N$  eine sehr gute Approximation von  $G$ ; entscheidend ist dabei  $N \gg n$ . Daher sind (A) und (B) approximativ gleichwertig zu

$$(A)' \quad \Phi^{-1}(1 - \alpha) \geq G_N^{-1}(p_\alpha)\sqrt{n} + c,$$

$$(B)' \quad \Phi^{-1}(1 - \beta) \leq G_N^{-1}(p_\beta)\sqrt{n} + c.$$

Für ein möglichst kleines  $n$  sollten diese Ungleichungen „gerade noch so“ erfüllt sein, d. h.

$$(A)' \quad +\Phi^{-1}(1 - \alpha) \approx G_N^{-1}(p_\alpha)\sqrt{n} + c,$$

$$(B)' \quad -\Phi^{-1}(1 - \alpha) \approx G_N^{-1}(p_\beta)\sqrt{n} + c.$$

wobei wir für (B)' unserer Konvention  $\beta = 1 - \alpha$  sowie die bekannte Relation  $\Phi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha)$  benutzt haben. Aus (A)' und (B)' erhalten wir sofort

$$0 \approx [G_N^{-1}(p_\alpha)\sqrt{n} + c] + [G_N^{-1}(p_\beta)\sqrt{n} + c]$$

und somit (34). Mit Hilfe von (36), (34) und  $\Phi^{-1}(\alpha) = -\Phi^{-1}(1 - \alpha)$  erreichen wir schließlich auch noch (35).  $\square$

In folgender Tabelle sind die Charakteristiken für typische Konstellationen von  $p_\alpha$ ,  $p_\beta$ ,  $\alpha$  (und  $\beta = 1 - \alpha$ ) angegeben, wobei der Datensatz *Beispieldaten-2* (Produkt P220) zugrunde gelegt wurde.

$p_\alpha$	$\alpha$	$p_\beta$	$\beta$	$c$	$n$
0.01	0.90	0.03	0.10	12.62	210
0.01	0.90	0.05	0.10	6.97	74

## 7 Verteilungstafeln

### 7.1 Quantile der Standardnormalverteilung

Tabellarisiert sind die Werte  $\Phi(z + \Delta z)$ .

Für  $z < 0$  benutze man  $\Phi(z) = 1 - \Phi(-z)$ .

$\Delta z \backslash z$	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
0,0	,5000	,5040	,5080	,5120	,5160	,5199	,5239	,5279	,5319	,5359
0,1	,5398	,5438	,5478	,5517	,5557	,5596	,5636	,5675	,5714	,5753
0,2	,5793	,5832	,5871	,5910	,5948	,5987	,6026	,6064	,6103	,6141
0,3	,6179	,6217	,6255	,6293	,6331	,6368	,6406	,6443	,6480	,6517
0,4	,6554	,6591	,6628	,6664	,6700	,6736	,6772	,6808	,6844	,6879
0,5	,6915	,6950	,6985	,7019	,7054	,7088	,7123	,7157	,7190	,7224
0,6	,7257	,7291	,7324	,7357	,7389	,7422	,7454	,7486	,7517	,7549
0,7	,7580	,7611	,7642	,7673	,7704	,7734	,7764	,7794	,7823	,7852
0,8	,7881	,7910	,7939	,7967	,7995	,8023	,8051	,8078	,8106	,8133
0,9	,8159	,8186	,8212	,8238	,8264	,8289	,8315	,8340	,8365	,8389
1,0	,8413	,8438	,8461	,8485	,8508	,8531	,8554	,8577	,8599	,8621
1,1	,8643	,8665	,8686	,8708	,8729	,8749	,8770	,8790	,8810	,8830
1,2	,8849	,8869	,8888	,8907	,8925	,8944	,8962	,8980	,8997	,9015
1,3	,9032	,9049	,9066	,9082	,9099	,9115	,9131	,9147	,9162	,9177
1,4	,9192	,9207	,9222	,9236	,9251	,9265	,9279	,9292	,9306	,9319
1,5	,9332	,9345	,9357	,9370	,9382	,9394	,9406	,9418	,9429	,9441
1,6	,9452	,9463	,9474	,9484	,9495	,9505	,9515	,9525	,9535	,9545
1,7	,9554	,9564	,9573	,9582	,9591	,9599	,9608	,9616	,9625	,9633
1,8	,9641	,9649	,9656	,9664	,9671	,9678	,9686	,9693	,9699	,9706
1,9	,9713	,9719	,9726	,9732	,9738	,9744	,9750	,9756	,9761	,9767
2,0	,9772	,9778	,9783	,9788	,9793	,9798	,9803	,9808	,9812	,9817
2,1	,9821	,9826	,9830	,9834	,9838	,9842	,9846	,9850	,9854	,9857
2,2	,9861	,9864	,9868	,9871	,9875	,9878	,9881	,9884	,9887	,9890
2,3	,9893	,9896	,9898	,9901	,9904	,9906	,9909	,9911	,9913	,9916
2,4	,9918	,9920	,9922	,9925	,9927	,9929	,9931	,9932	,9934	,9936
2,5	,9938	,9940	,9941	,9943	,9945	,9946	,9948	,9949	,9951	,9952
2,6	,9953	,9955	,9956	,9957	,9959	,9960	,9961	,9962	,9963	,9964
2,7	,9965	,9966	,9967	,9968	,9969	,9970	,9971	,9972	,9973	,9974
2,8	,9974	,9975	,9975	,9977	,9977	,9978	,9979	,9979	,9980	,9981
2,9	,9981	,9982	,9982	,9983	,9984	,9984	,9985	,9985	,9986	,9986
3,0	,9987	,9987	,9987	,9988	,9988	,9989	,9989	,9989	,9990	,9990
3,1	,9990	,9991	,9991	,9991	,9992	,9992	,9992	,9992	,9993	,9993
3,2	,9993	,9993	,9994	,9994	,9994	,9994	,9994	,9995	,9995	,9995
3,3	,9995	,9995	,9995	,9996	,9996	,9996	,9996	,9996	,9996	,9997
3,4	,9997	,9997	,9997	,9997	,9997	,9997	,9997	,9997	,9997	,9998

## 7.2 Quantile von $\chi^2$ -Verteilungen

Tabellarisiert sind die  $(1 - \alpha)$ -Quantile  $q_{1-\alpha}$  für die Freiheitsgrade  $m$ .<sup>4</sup>

$m \backslash \alpha$	0,995	0,990	0,950	0,900	0,500	0,100	0,050	0,010	0,005
1	0,0 <sup>4</sup> 39	0,0 <sup>3</sup> 16	0,0 <sup>2</sup> 39	0,016	0,455	2,71	3,84	6,63	7,9
2	0,010	0,020	0,103	0,211	1,39	4,61	5,99	9,21	10,6
3	0,072	0,115	0,352	0,584	2,37	6,25	7,81	11,3	12,8
4	0,207	0,297	0,711	1,06	3,36	7,78	9,49	13,3	14,9
5	0,412	0,554	1,15	1,61	4,35	9,24	11,1	15,1	16,7
6	0,676	0,872	1,64	2,20	5,35	10,6	12,6	16,8	18,5
7	0,989	1,24	2,17	2,83	6,35	12,0	14,1	18,5	20,3
8	1,34	1,65	2,73	3,49	7,34	13,4	15,5	20,1	22,0
9	1,73	2,09	3,33	4,17	8,34	14,7	16,9	21,7	23,6
10	2,16	2,56	3,94	4,87	9,34	16,0	18,3	23,2	25,2
11	2,60	3,05	4,57	5,58	10,3	17,3	19,7	24,7	26,8
12	3,07	3,57	5,23	6,30	11,3	18,5	21,0	26,2	28,3
13	3,57	4,11	5,89	7,04	12,3	19,8	22,4	27,7	29,8
14	4,07	4,66	6,57	7,79	13,3	21,1	23,7	29,1	31,3
15	4,60	5,23	7,26	8,55	14,3	22,3	25,0	30,6	32,8
16	5,14	5,81	7,96	9,31	15,3	23,5	26,3	32,0	34,3
17	5,70	6,41	8,67	10,1	16,3	24,8	27,6	33,4	35,7
18	6,26	7,01	9,39	10,9	17,3	26,0	28,9	34,8	37,2
19	6,84	7,63	10,1	11,7	18,3	27,2	30,1	36,2	38,6
20	7,43	8,26	10,9	12,4	19,3	28,4	31,4	37,6	40,0
21	8,03	8,90	11,6	13,2	20,3	29,6	32,7	38,9	41,4
22	8,64	9,54	12,3	14,0	21,3	30,8	33,9	40,3	42,8
23	9,26	10,2	13,1	14,8	22,3	32,0	35,2	41,6	44,2
24	9,89	10,9	13,8	15,7	23,3	33,2	36,4	43,0	45,6
25	10,5	11,5	14,6	16,5	24,3	34,4	37,7	44,3	46,9
26	11,2	12,2	15,4	17,3	25,3	35,6	38,9	45,6	48,3
27	11,8	12,9	16,2	18,1	26,3	36,7	40,1	47,0	49,6
28	12,5	13,6	16,9	18,9	27,3	37,9	41,3	48,3	51,0
29	13,1	14,3	17,7	19,8	28,3	39,1	42,6	49,6	52,3
30	13,8	15,0	18,5	20,6	29,3	40,3	43,8	50,9	53,7
40	20,7	22,2	26,5	29,1	39,3	51,8	55,8	63,7	66,8
50	28,0	29,7	34,8	37,7	49,3	63,2	67,5	76,2	79,5
60	35,5	37,5	43,2	46,5	59,3	74,4	79,1	88,4	92,0
70	43,3	45,4	51,7	55,3	69,3	85,5	90,5	100	104
80	51,2	53,5	60,4	64,3	79,3	96,6	102	112	116
90	59,2	61,8	69,1	73,3	89,3	108	113	124	128
100	67,3	70,1	77,9	82,4	99,3	118	124	136	140

<sup>4</sup> $q_{1-\alpha}$  ist implizit definiert via  $1 - G_m(q_{1-\alpha}) = \alpha$ .



## Literatur

- [1] N.I. Johnson and S. Kotz, *Continuous univariate distributions - 2*. Houghton Mifflin Company, Boston, 1970.
- [2] D.B. Owen, *Handbook of Statistical Tables*. Reading Mass.-Palo Alto, London, 1962.