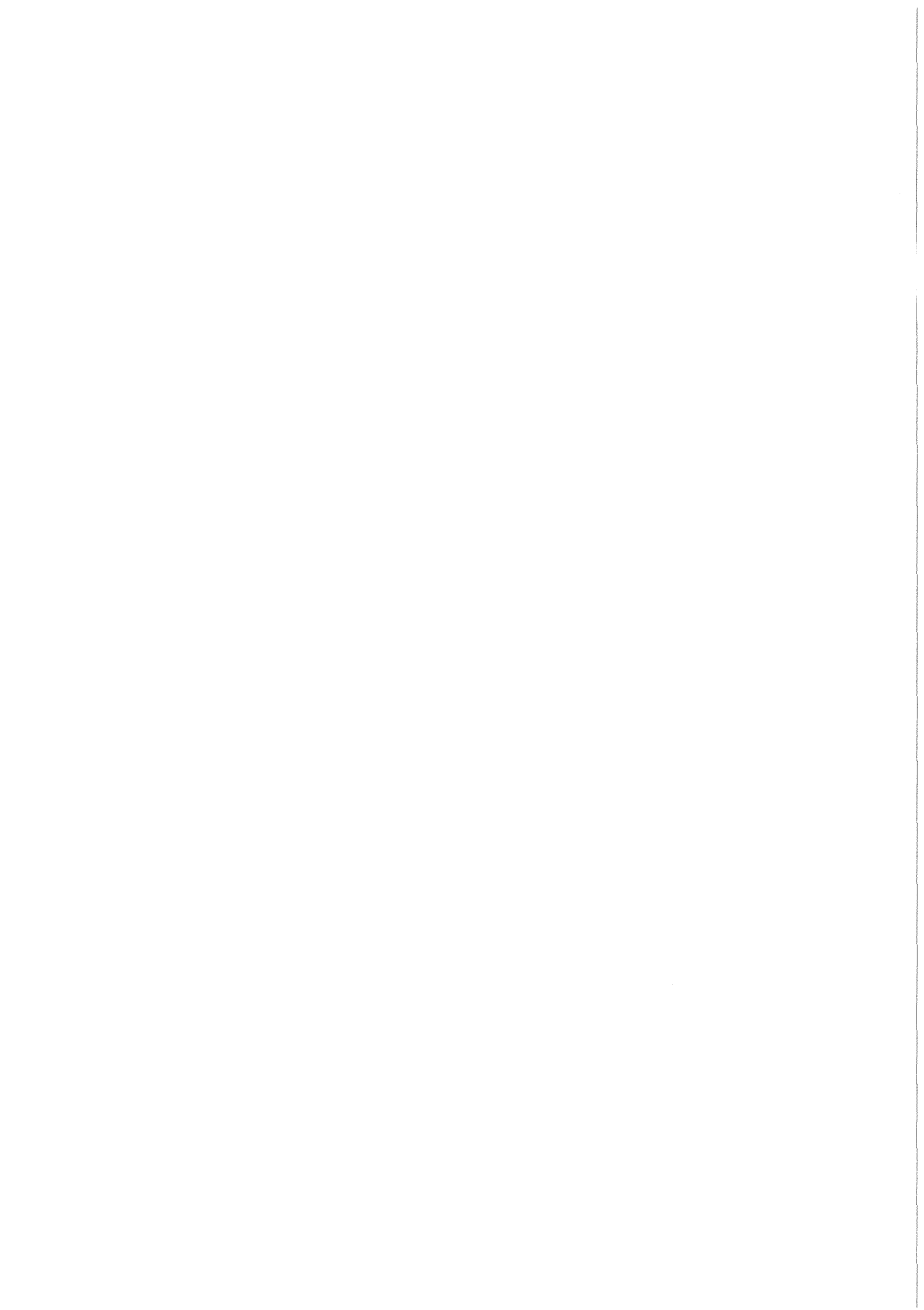


KfK 3661
Januar 1984

Bewertung von Test- und Schätzverfahren zur Entdeckung von Materialverlusten

E. Zerrweck
Projekt Kernmaterialüberwachung

Kernforschungszentrum Karlsruhe



KERNFORSCHUNGSZENTRUM KARLSRUHE

Projekt Kernmaterialüberwachung

KfK 3661

Bewertung von Test- und Schätzverfahren zur
Entdeckung von Materialverlusten

E. Zerrweck

Die Untersuchungen wurden im Auftrag des
Projekttes Kernmaterialüberwachung durch-
führt.

Kernforschungszentrum Karlsruhe GmbH, Karlsruhe

Als Manuskript vervielfältigt
Für diesen Bericht behalten wir uns alle Rechte vor

Kernforschungszentrum Karlsruhe GmbH
ISSN 0303-4003

Kurzfassung

Statistische Schätz- und Testverfahren zur Analyse sequentieller Materialbilanzdaten lassen sich für Zwecke der internationalen Kernmaterialüberwachung einsetzen. Insbesondere für den Fall der zeitkonstanten Abzweigung kleiner Mengen von spaltbarem Material ist ein Hypothesetest zu entwickeln, der bei kontrollierter Fehlalarmrate eine ausreichend hohe Entdeckungswahrscheinlichkeit ermöglicht.

Es wird gezeigt, daß in dieser Hinsicht der "Power-One-Test" für unabhängige Variable und unendlichen Zeithorizont optimale Eigenschaften besitzt, bei Anwesenheit eines systematischen Fehlers in der Bilanzermittlung jedoch die Anforderungen an Fehlalarmrate und Entdeckungswahrscheinlichkeit nicht mehr gleichzeitig zu erfüllen sind.

Weiterhin wird nachgewiesen, daß für endliche Zeithorizonte eine analytische Berechnung der Fehlalarmwahrscheinlichkeit nicht möglich ist; es werden daher verschiedene Abschätzmethoden auf ihre Eignung untersucht.

Aufgrund der kumulierten Meßfehler lassen sich die Nachweisgrenzen für eine Materialabzweigung durch die Einführung sequentieller Zwischentests oder komplexere mathematische Auswertemethoden prinzipiell nicht verändern. Wenn es jedoch gelingt, durch verstärkte Überwachungsmaßnahmen sicherzustellen, daß in einem dem Untersuchungszeitraum vorgelagerten Zeitintervall kein Materialverlust auftritt, läßt sich der Einfluß konstanter systematischer Fehler berücksichtigen. Damit lassen sich auch längere Zeiträume wirksam auf Materialabzweigungen hin kontrollieren.

Evaluation of Test- and Estimation Procedures for the Detection of
Material Losses

Abstract

The application of advanced statistical methods (estimation and test-procedures) in the analysis of sequential material balance data may improve the implementation of nuclear materials safeguards in bulk handling facilities. In particular a hypothesis-test has to be developed which allows the detection of a diversion constant in time of small amounts of fissile material over extended periods of time with a high detection probability and a controlled low false alarm probability.

It is demonstrated that for the case of independent variables and an infinite time horizon the Power-One-Test yields very satisfactory results. But in the presence of systematic errors of the balance data the requirements with respect to detection - and false alarm - probability can not be met simultaneously.

For a finite time horizon the analytical calculation of the false alarm probability proves not to be possible, therefore different methods to estimate the false alarm rate have been studied.

Due to the accumulated measurement error the limitations for detection of a protracted diversion can in principle not be influenced, neither by means of intermediate tests nor by application of more complex mathematical evaluation methods or a higher effort for recalibration or multiple measurements.

But if it can be assured that within a preceding accountancy interval no material losses will have occurred one can adjust for the systematic errors. By that way time periods much longer than that "clean calibration interval" can be effectively controlled for any significant material loss.

I N H A L T

	<u>Seite</u>
I. <u>Einführung und Zusammenfassung</u>	1
II. <u>Die Robbins-Schranken</u>	3
1. Eine allgemeine Abschätzung	3
2. Anwendung auf die Normalverteilung	3
3. Die Abschätzung ist streng	5
4. Das Gesetz des iterierten Logarithmus liefert eine Minimalbedingung für die Schrankenfolge	6
5. Beispiele zur Wahl von F	7
6. Weitere Beispiele für symmetrisches F	11
7. Kann für korrelierte Größen ein Power-One-Test konstruiert werden?	12
III. <u>Die Bestimmung der Test-Fehlalarmwahrscheinlichkeit α bei einer Folge von Materialbilanzen</u>	14
IV. <u>Protracted Loss und systematischer Fehler bei Anlagen mit hohem Durchsatz</u>	25
1. CUMUF ist ein Gleichgewichtstest	25
2. Die Minimalstrategie μ^*	27
3. Anlagen mit hohem Durchsatz	31
4. Die gestufte Zufallsstruktur der Meßfehler bei Transfermessungen	34
5. Die Sättigungsschranke für die Entdeckungscharakteristik $1-\beta$ bei "protracted loss"	42
6. Power-One-Test und Schätzung von σ_{res}	46
7. Programm, um das Ziel der Kontrolle einer Reprocessing-Anlage zu erreichen	54
V. <u>Referenzen</u>	56

I. Einführung und Zusammenfassung

In kerntechnischen Anlagen, in denen größere Mengen spaltbaren Materials in offener Form verarbeitet werden, wie z. B. Anlagen zur Wiederaufarbeitung bestrahlter Brennelemente, lassen die klassischen Verfahren der Kernmaterialbilanzierung, die sich jeweils über längere Zeiträume erstrecken, nur begrenzte Aussagen über mögliche Materialabzweigungen zu.

Der Einsatz von statistischen Schätz- und Testverfahren zur Analyse sequentieller Materialbilanzdaten läßt eine Verbesserung im Hinblick auf die Erreichung der zeit- und mengenbezogenen Entdeckungskriterien für eine Materialabzweigung erwarten. Solche Testverfahren lassen sich jeweils optimal auf ein spezifisches zugrundezulegendes Abzweigszenarium hin konstruieren. Es gibt aber keinen gleichzeitig für alle Abzweigszenarien optimalen Test.

Unter Berücksichtigung der endlichen Fehlerbreite der in die periodische Materialbilanzerstellung eingehenden Meßwerte bereitet die Entdeckung einer zeitkonstanten Abzweigung kleiner Mengen von spaltbarem Material über längere Zeiträume (protracted diversion) mit hoher Entdeckungswahrscheinlichkeit bei gleichzeitig kontrollierter niedriger Fehlalarmwahrscheinlichkeit besondere Schwierigkeiten. Das Ziel ist, einen Hypothesentest auszuwählen bzw. zu konstruieren, so daß dessen Anwendung auf Zeitreihen von Materialbilanzen - mit besonderer Berücksichtigung der Sequenz der CUMUFR-Werte (Cumulated sum if standardized MUF-residuals) - für ein solches Abzweigszenarium optimale Wirkung verspricht.

In der Arbeit wird gezeigt, daß der Power-One-Test für unabhängige Variablen und unendlichen Zeithorizont im starken Maße die mathematischen Möglichkeiten für einen sequentiellen Test

$\{S_i\}_{i=1}^n$ ausschöpft. Es wird bewiesen, daß in Anwesenheit eines systematischen Fehlers der Power-One-Test jedoch zusammenbricht,

d. h. nicht gleichzeitig die Fehlalarmrate kontrolliert und eine lineare Materialentnahme mit hoher Entdeckungswahrscheinlichkeit (genauer $P_{\text{det}} \rightarrow 1$) nachgewiesen werden kann. Dies gilt prinzipiell für alle Schrankenfolgen und nicht nur für den speziellen Power-One-Test mit den Robbins-Schranken.

Wegen der Unmöglichkeit, die Fehlalarmwahrscheinlichkeit für eine gegebene Schrankenfolge analytisch zu berechnen, werden daher für diese verschiedene Abschätzmethoden hergeleitet: mehrdimensionale Normalverteilungen, Wiederprozeßapproximation, Robbins-Gerade. Die einfach anwendbare Robbins-Gerade liefert hierbei die beste Abschätzung.

Es wird nachgewiesen, daß durch sequentielle Zwischentests und verschiedene in den vergangenen Jahren vorgeschlagene komplexere mathematische Auswertmethoden prinzipiell die Nachweisgrenzen der Materialbilanzierung wegen der kumulierten Meßfehler auch bei Ausnutzung der Möglichkeiten von Neukalibrierung und Mehrfachmessung nicht verbessert werden können. Andererseits besteht die Möglichkeit, durch Verwendung von Bilanzsequenzen einen Hauptteil des Gesamtfehlers zu reduzieren, wenn innerhalb einer vorgelagerten Teil-Bilanzsequenz sichergestellt wird, daß bei bestimmungsgemäßem Betrieb der Anlage keine Materialverluste eintreten können. Durch die Realisierung eines solchen "sauberen Eich-Intervalls" mit Hilfe eines erhöhten Meß- und Überwachungsaufwandes während dieser vorgelagerten Sequenz läßt sich mit mathematischen Methoden ein mehrfach längerer Zeitraum kontrollieren und eine Materialfehlmenge mit vorgegebener Größe bei vorgegebener Testcharakteristik (Fehlalarmwahrscheinlichkeit) entdecken. Lediglich kurzfristige systematische Fehler setzen der Erhöhung der Nachweismöglichkeit eine gewisse Schranke. In diesem Zusammenhang ergibt der vorgeschlagene Schätzer CUMUFR überraschend gute Ergebnisse; es lassen sich aber auch andere Schätzfunktionen mit hoher Entdeckungswahrscheinlichkeit und für die Praxis wichtigen Eigenschaften angeben.

II. Die Robbins-Schranken

Sei X_1, \dots eine Folge von Zufallsvariablen mit der Verteilung P .

Als Ziel soll eine Abschätzung von $P(S_n \geq f(n)) \leq \varepsilon$ der Überschreitungen durch die Summe $S_n = \sum_{i=1}^n X_i$ von geeignet gewählten Schrankenfunktionen $f(n)$ bei beliebig vorgegebenem ε gefunden werden.

1) Eine allgemeine Abschätzung

Wir setzen voraus, daß für alle $n \geq 1$ die Variablen X_1, \dots, X_n eine Dichte

$P(X_1, \dots, X_n \in B) = \int_B g_n dx_n$ bezüglich des Lebesgue-Maßes besitzen und ein weiteres Maß P' genauso darstellbar ist

$$P'(X_1, \dots, X_n \in B) = \int_B g_n' dx_n .$$

Auf $g_n > 0$ sei der Likelihood-Quotient definiert durch $z_n = \frac{g_n'}{g_n}$.

Dann gilt für $\varepsilon \geq 1$

$$(1) \quad P(z_n \geq \varepsilon \text{ für ein } n \geq 1) \leq \frac{1}{\varepsilon}$$

[Beweis: N sei das erste $n \geq 1$ mit $g_n' \geq \varepsilon g_n$ und $N = \infty$, falls kein solches n existiert. Es ist $P(g_n = 0 \text{ für ein } n \geq 1) = 0$ und

$$\begin{aligned} P(z_n \geq \varepsilon \text{ für ein } n \geq 1) &= P(N < \infty) = \sum_{n=1}^{\infty} \int_{\{N=n\}} g_n dx_n \\ &\leq \frac{1}{\varepsilon} \sum_{n=1}^{\infty} \int_{\{N=n\}} g_n' dx_n = \frac{1}{\varepsilon} P'(N < \infty) \leq \frac{1}{\varepsilon} \end{aligned}$$

2) Anwendung auf die Normalverteilung

Sei $P_\theta (-\infty < \theta < \infty)$ das Maß gebildet aus $N(\theta, 1)$ verteilten, unabhängigen X_1, \dots und $P := P_0$ in 1).

Die Dichte der Normalverteilung $N(0,1)$ werde mit $\phi(x)$ bezeichnet

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp -\frac{x^2}{2} \quad \phi(x) := \int_{-\infty}^x \phi(t) dt$$

Dann ist die Dichte von X_1, \dots, X_n unter P_θ gegeben durch

$$g_{\theta, n}(x_1, \dots, x_n) = \prod_1^n \phi(x_i - \theta) \text{ wegen der Unabhängigkeit.}$$

Wir nehmen P' als Mischung $P'(\cdot) = \int_{-\infty}^{\infty} P_\theta(\cdot) dF(\theta)$,

wobei F eine Verteilung auf $(-\infty, \infty)$ ist.

Dann läßt sich g'_n ausdrücken

$$g'_n(x_1, \dots, x_n) = \int_{-\infty}^{\infty} g_{\theta, n}(x_1, \dots, x_n) dF(\theta) = \int_{-\infty}^{\infty} \prod_1^n \phi(x_i - \theta) dF(\theta)$$

$$\begin{aligned} (2) \quad z_n &= \frac{g'_n}{g_{0, n}} = \int_{-\infty}^{\infty} \prod_1^n \frac{\phi(x_i - \theta)}{\phi(x_i)} dF(\theta) = \int_{-\infty}^{\infty} \prod_1^n \exp(\theta x_i - \frac{1}{2} \theta^2) dF(\theta) \\ &= \int_{-\infty}^{\infty} \exp(\theta S_n - \frac{1}{2} n \theta^2) dF(\theta). \end{aligned}$$

Nun definieren wir

$$f(x, t) := \int_{-\infty}^{\infty} \exp(xy - \frac{1}{2} y^2 t) dF(y) \text{ und setzen anstelle}$$

$F(\theta)$ die Verteilung $F(\theta\sqrt{m})$ ein ($m > 0$), so wird (2) zu ($y = \theta\sqrt{m}$)

$$\begin{aligned} z_n &= \int_{-\infty}^{\infty} \exp(\theta S_n - \frac{1}{2} n \theta^2) dF(\theta\sqrt{m}) = \int_{-\infty}^{\infty} \exp(y \frac{S_n}{\sqrt{m}} - \frac{1}{2} \frac{n}{m} y^2) dF(y) \\ &= f\left(\frac{S_n}{\sqrt{m}}, \frac{n}{m}\right). \end{aligned}$$

Also ergibt sich für die Normalverteilung in (1)

$$(3) \quad P\left(f\left(\frac{S_n}{\sqrt{m}}, \frac{n}{m}\right) \geq \varepsilon \text{ für ein } n \geq 1\right) \leq \frac{1}{\varepsilon} \quad (m > 0, \varepsilon \geq 1)$$

Dies wird nun in eine Abschätzung für S_n umgeformt, falls F auf $(0, \infty)$ konzentriert ist, also $f(\cdot, t)$ eine wachsende Funktion ist.

Dann wird für $t > 0$

$A(t, \varepsilon)$ als die positive Lösung x der Gleichung $f(x, t) = \varepsilon$ definiert;
damit gilt

$$f(x, t) \geq \varepsilon \text{ genau dann, wenn } x \geq A(t, \varepsilon).$$

In (3) ergibt sich dann

$$(4) P(S_n \geq \sqrt{m} A(\frac{n}{m}, \varepsilon) \text{ für ein } n \geq 1) \leq \frac{1}{\varepsilon} \quad (m > 0, \varepsilon \geq 1)$$

Dies ist besonders günstig, wenn sich für A ein geschlossener analytischer Ausdruck angeben läßt, so daß die Schranke sich dann in einfacher Weise für die sukzessiven n berechnen läßt. (siehe weiter unten)
Man hat also nun geeignete Verteilungen F zu wählen, erhält daraus Schrankenfunktionen - dies war das Ziel - und hat für das praktische Auswerten darauf zu achten, daß sich f und A explizit als Funktionen angeben lassen.

3) Die Abschätzung ist streng

Zunächst ergibt sich die Frage, wie gut die Abschätzung ist.

Ergebnis: Die Abschätzung mit $\frac{1}{\varepsilon}$ läßt sich nicht verbessern, die Schranke $\frac{1}{\varepsilon}$ ist genau. Denn es gilt

$$\lim_{m \rightarrow \infty} P\left[f\left(\frac{S_n}{\sqrt{m}}, \frac{n}{m}\right) \geq \varepsilon \text{ für ein } n \geq 1\right] = \frac{1}{\varepsilon}$$

Hierzu wird die Approximation durch den Wiener-Prozeß benutzt.

Denn einmal haben $\frac{S_1}{\sqrt{m}}, \frac{S_2}{\sqrt{m}}, \dots$ und

$w(\frac{1}{m}), w(\frac{2}{m}), \dots$ dieselbe Verteilung,

d.h. die Prozesse $X^m(t) := \frac{S_i}{\sqrt{m}}$ für $t \in [\frac{i}{m}, \frac{i+1}{m})$

konvergieren in Verteilung gegen den Wiener-Prozeß $w(t)$

$$X^m(t) \xrightarrow{D} w(t).$$

Dann gilt nach dem Invarianzprinzip für den Wiener-Prozeß, daß auch

ein stetiges Funktional diese Konvergenz erhält (unter Beachtung, daß

$m(t) := \frac{n}{m}$ für $t \in \left[\frac{n}{m}, \frac{n+1}{m}\right)$ gegen $\frac{n}{m}$ konvergiert)

also ergibt das Funktional

$(X^m(t), m(t)) \rightarrow \sup_t f(X^m(t), m(t))$ gerade die Beziehung

$P\left[\sup_{\frac{n}{m}} f\left(\frac{S_n}{\sqrt{m}}, \frac{n}{m}\right) \geq \varepsilon \text{ für ein } n \geq 1\right] \rightarrow P(f(w(t), t) \geq \varepsilon \text{ für ein } t \geq 0)$

für $m \rightarrow \infty$.

Zum anderen gilt $P(f(w(t), t) \geq \varepsilon \text{ für ein } t \geq 0) = \frac{1}{\varepsilon}$, was sich direkt mit Martingalsätzen ergibt ($f(w(t), t)$ ist ein Martingal, $t \in [0, \infty)$).

Insgesamt also das Ergebnis, daß $\frac{1}{\varepsilon}$ als Schranke genau ist.

4) Das Gesetz des iterierten Logarithmus liefert eine Minimalbedingung für die Schrankenfolge

a) F kann nun beliebig gewählt werden, so gilt das bisher Gesagte und man erhält eine Schrankenfunktion $f(n)$.

b) Jedoch ergibt sich nur bei geschickter Wahl von F ein direkter, expliziter Ausdruck für $f(x, t)$ und zudem noch für $A(t, \varepsilon)$. Es folgen anschließend wichtige Beispiele, die zeigen, daß eine sehr große Fülle von Funktionen $f(n)$ ohne weiteres erhalten werden kann.

c) Wünschenswert ist eine Schrankenfunktion, die so langsam wie möglich (für große n) anwächst.

Nach dem Gesetz des iterierten Logarithmus

$$P\left(\limsup_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = 1\right) = 1$$

gilt, daß

$$P(S_n \geq (1-\varepsilon) f(n) \text{ für unendl. viele } n) = 1$$

$$P(S_n \leq (1-\varepsilon) f(n) \text{ für unendl. viele } n) = 0$$

mit $f(n) := \sqrt{2n \log \log n}$.

Dies besagt,

α) Jede Funktion, die langsamer wächst (z.B. $c\sqrt{n}$, $c n^\alpha$ mit $0 < \alpha < \frac{1}{2}$, $c > 0$) als die nach diesem Gesetz gegebene $\sqrt{2 n \log \log n}$ wird unendlich oft (schließliches Verhalten!) von S_n überschritten. [Erst recht nur einmal.]

β) Jede Funktion, die schneller wächst (z.B. $c \sqrt{n \log n}$, $c n$, $c n^\alpha$ mit $\frac{1}{2} < \alpha < \infty$) wird schließlich für große n nicht mehr von S_n überschritten. (Trotzdem braucht daraus nicht zu folgen, daß $P(S_n \geq f(n) \text{ für ein } n) < 1$ für diese Schrankenfunktionen gilt.)

In diesem präzisierten Sinne ist $\sqrt{2 n \log \log n}$ die Schrankenfunktion mit dem langsamsten Anstieg. Es zeigt sich, daß sogar ein F gefunden wird, so daß $f(n)$ wie $\sqrt{2 n \log \log n}$ anwächst, also nach dem bisherigen dann gilt $P(S_n \geq f(n, \epsilon) \text{ für ein } n \geq 1) \leq \frac{1}{\epsilon}$, d.h. diese Wahrscheinlichkeit beliebig klein gemacht werden kann. Man hat damit für das schließliche Verhalten die strengste überhaupt denkbare Aussage:

- wächst $f(n)$ langsamer als $\sqrt{2 n \log \log n}$, so wird $f(n)$ von S_n unendlich oft mit Wahrscheinlichkeit 1 überschritten,
- wächst $f(n)$ gerade so wie diese Grenzfunktion, so wird $f(n)$ von S_n niemals (für kein n) überschritten mit Ausnahme einer Menge, deren Wahrscheinlichkeit beliebig klein gemacht werden kann [durch entsprechende Konstruktion von $f(n)$ ohne Änderung des Wachstumsverhaltens].

5) Beispiele zur Wahl von F (reines rechentechnisches Problem)

Man muß dabei zwei Schwierigkeiten beachten, nämlich daß sowohl $f(x, t)$ als auch $A(t, \epsilon)$ explizit angegeben werden können (für die praktische Verwendbarkeit zur Berechnung von $f(n)$).

5.1 F sei auf den Punkt $2a > 0$ konzentriert. Damit ist

$$f(x, t) = \exp(2a x - 2a^2 t) \geq \epsilon \text{ gleichwertig mit}$$

$$x \geq at + \frac{\log \epsilon}{2a} = at + b$$

Also ergibt sich in (4) mit $b = \frac{\log \epsilon}{2a}$

$$(5) \quad P(S_n \geq \frac{a}{\sqrt{m}} n + b \cdot \sqrt{m} \text{ für ein } n \geq 1) \leq e^{-2ab} \quad (a, b, m > 0)$$

gerade eine Gerade als Begrenzungsfunktion $f(n)$.

5.2 F sei gegeben durch die Dichte

$$\sqrt{\frac{2}{\pi}} e^{-\frac{1}{2} y^2} dy \quad \text{auf } (0, \infty)$$

Hier läßt sich $f(x, t)$ ausrechnen, nicht jedoch $A(t, \epsilon)$; eine weitere Abschätzung ergibt aber ein $\tilde{A} > A$, das sich explizit berechnen läßt; damit hat man

$$\left[\text{wegen } \frac{1}{\epsilon} \geq P(S_n \geq A(n, \epsilon) \text{ für ein } n \geq 1) > P(S_n \geq \tilde{A}(n, \epsilon) \text{ für ein } n \geq 1) \right]$$

nach einfacher Umformung

$$(6) \quad P(S_n \geq \sqrt{(n+m) [a^2 + \log(\frac{n}{m} + 1)]} \text{ für ein } n \geq 1) < \frac{1}{\epsilon} = \frac{e^{-\frac{1}{2} a^2}}{2\Phi(a)}$$

$$(a \geq 0, m > 0)$$

wobei $\epsilon = 2 e^{\frac{1}{2} a^2} \Phi(a)$ gesetzt wird.

5.3 Für $\delta > 0$ sei F definiert durch die Dichte

$$\frac{\delta}{y \log \frac{1}{y} (\log \log \frac{1}{y})^{1+\delta}} dy \quad \text{auf } (0, e^{-e}) \text{ und } 0 \text{ sonst.}$$

Dann läßt sich nicht einmal $f(x, t)$ explizit berechnen, jedoch gilt für $t \rightarrow \infty$, daß

$$A^2(t, \epsilon) = 2t \left[\log \log t + \left(\frac{3}{2} + \delta\right) \log \log \log t + \log \frac{\epsilon}{2\delta\sqrt{\pi}} + o(1) \right]$$

Dies ist sehr bemerkenswert, da man nun ein $f(n)$ gefunden hat (wie oben in 4.c) schon aufgeführt), mit $P(S_n \geq f(n) \text{ für ein } n \geq 1) \leq \frac{1}{\epsilon}$ und $f_n \sim \sqrt{2n \log \log n}$ für $n \rightarrow \infty$, also die bestmögliche Schranke überhaupt für das schließliche Verhalten. Für eine langsamer wachsende Schranke hat man ja gar keine Chance, diese Wahrscheinlichkeit kleiner als 1 zu machen, gleichgültig wie die Schranken am Anfang (bis zu einem beliebigen n) gewählt werden. Aber für dieses Wachstumverhalten (für große n) kann

schon die Wahrscheinlichkeit des Übertritts der Schranke beliebige Werte P ($0 < P \leq 1$) annehmen - diese Werte werden angenommen aufgrund des in 3) besprochenen Ergebnisses mit demselben Wachstumsverhalten für große n , da

$$\begin{aligned} \sqrt{m} A\left(\frac{n}{m}, \epsilon\right) &\sim \sqrt{m} \cdot \sqrt{2 \frac{n}{m} \log \log \frac{n}{m}} \\ &\sim \sqrt{2n \log \log n} \end{aligned}$$

- also durch entsprechende Wahl von ϵ insbesondere beliebig klein gemacht werden. Mit anderen Worten, man kann Schranken angeben mit

- a) dem bestmöglichen Anstiegsverhalten $\sqrt{2n \log \log n}$ und so, daß
- b) die Übertrittswahrscheinlichkeit beliebig klein wird und so, daß
- c) das Übertreten gerade eine vorgegebene Wahrscheinlichkeit P ergibt (genauer: beliebig nahe bei P liegt; wegen $\lim_{m \rightarrow \infty} \dots = P$). Man kann ja eine Konstante zu der Schranke addieren (oder gewisse Funktionen), so gilt a) und b) weiterhin, aber c) besagt, daß es sogar möglich ist (durch Wahl von m), die Schranke so "tief" zu legen, daß die vorgegebene Wahrscheinlichkeit P ausgeschöpft wird, die Abschätzung $P(\) \leq \frac{1}{\epsilon}$ also dann scharf ist.

5.4 Die bisher angeführten Beispiele zeigen, daß die Schrankenfunktionen $f(n)$ mit $P(S_n \geq f(n) \text{ für ein } n \geq 1) \leq \frac{1}{\epsilon}$ durch Wahl von F bereits das Wachstumsverhalten (für sehr große n)

(5.1) $f_n \sim c \cdot n$ Gerade

(5.2) $f_n \sim \sqrt{(n+m) \log \frac{n}{m}} \sim \sqrt{n \log n}$

(5.3) $f_n \sim \sqrt{2n \log \log n}$

aufweisen. Weitere F ergeben das Wachstumsverhalten

$$f_n \sim cn^\alpha, \quad \frac{1}{2} < \alpha < 1.$$

Damit ist praktisch jedes Wachstumsverhalten zwischen einer ansteigenden Geraden (Anstieg beliebig) und der Schranke mit dem geringstmöglichen Anstieg wählbar, so daß eine vorgegebene Wahrscheinlichkeit P ($0 < P < 1$) beliebig nahe durch die Übertrittswahrscheinlichkeit $P(S_n \geq f(n)$ für ein $n) \geq P - \delta$ ausgeschöpft wird. Nur die Aussage allein, daß diese Übertrittswahrscheinlichkeit abgeschätzt werden kann, wäre bei weitem nicht so wertvoll - man nehme etwa eine Gerade $f_n \sim c \cdot n$, wenn nicht die bestmögliche Schrankenfunktion $\sqrt{2n \log \log n}$ wählbar wäre. Der Power-One-Test

$$\begin{aligned} |S_n| \geq f(n) & \quad \text{so } H_1: \theta \neq 0 \\ |S_n| < f(n) & \quad \text{so } H_0: \theta = 0 \quad \text{keine Entscheidung} \end{aligned}$$

verwendet trotzdem die geringfügig schneller anwachsende Schranke wie in (b)

$$f(n) = \sqrt{(n+m) \left[a^2 + \log \left(\frac{n}{m} + 1 \right) \right]}$$

weil für das bestmögliche Anstiegsverhalten eine Auflösung von $f(x,t)$ und damit die nachfolgende Umformung in $A(n,\epsilon)$ erst recht scheitert. Man hat somit in aufsteigender Güte für die Schrankenfunktion $f(n)$

$$\sim ax^{\frac{1}{2} + \frac{1}{2} + d} \quad \text{gibt schlechtes Wachstumsverhalten für } n \rightarrow \infty$$

$$\sim x^{\frac{1}{2} + \epsilon}, \quad \epsilon > 0 \quad \text{wächst immer noch stärker als } \sqrt{x \log x}$$

$$\sim \sqrt{x \log x} \quad \text{die für den Power-One-Test gewählte Schranke, da } f(x,t) \text{ und } A(n,\epsilon) \text{ explizit berechenbar sind}$$

$$\sim \sqrt{x \log \log x} \quad \text{bestmögliche Schranke, jedoch weder } f(x,t) \text{ noch } A(n,\epsilon) \text{ explizit berechenbar}$$

$$\sim \sqrt{x} \quad \text{die für den nichtsequentiellen Neyman-Pearson-Test verwendete Schranke; wächst zu langsam wegen dem Gesetz des iterierten Logarithmus.}$$

Für das in einem endlichen Zeitabschnitt vorliegende Safeguardsproblem muß nun ein m gewählt werden, und damit kann sowohl für dieses $m < \infty$ als auch wegen des endlichen Zeitabschnitts $n \leq N$ die Abschätzung durch

$\frac{1}{\epsilon}$ sehr schlecht sein. Hiermit befaßt sich das Kapitel II.

Weiterhin wird sich zeigen, daß die zentrale Aussage des Power-One-Tests: daß sich die "lineare Entnahme" θ wegen $\frac{S_n}{n} \rightarrow \theta$ (Gesetz der großen Zahlen), $\frac{f_n}{n} \rightarrow 0$ also $P_\theta(S_n \geq f_n \text{ für ein } n) = 1 \forall \theta \neq 0$ mit Wahrscheinlichkeit eins entdecken läßt und dabei die Fehlalarme $P_0(S_n \geq f_n \text{ für ein } n) \leq \frac{1}{\epsilon}$ kontrolliert werden können, wie die genauen Betrachtungen in 1) bis 5) gerade beweisen - für die Safeguardsfragestellung nicht gültig ist. Dies wird ausführlich in Kapitel III. behandelt und ein mathematischer Teilaspekt in Abschnitt 7.

6) Weitere Beispiele für symmetrisches F

Falls F symmetrisch auf $(-\infty, \infty)$ verteilt ist, so ist $(f(-x, t) = f(x, t))$

$$f(x, t) \geq \epsilon \text{ genau dann, wenn } |x| \geq A(t, \epsilon)$$

und damit

$$P(|S_n| \geq \sqrt{m} A(\frac{n}{m}, \epsilon) \text{ für ein } n \geq 1) \leq \frac{1}{\epsilon} \quad (m > 0, \epsilon \geq 1)$$

In Beispiel 5.2 erhält man mit der Normalverteilung $F = \Phi$ (dort wurde diese auf $(0, \infty)$ konzentriert)

$$(7) \quad P(|S_n| \geq [(n+m)(a^2 + \log(\frac{n}{m} + 1))]^{\frac{1}{2}} \text{ für ein } n \geq 1) \leq \exp(-\frac{1}{2} a^2)$$

mit $\epsilon = \exp \frac{1}{2} a^2$.

Für andere symmetrische Verteilungen ergeben sich jeweils Schrankenfunktionen für $|S_n|$; es wird noch die Schranke für die Punktverteilung berechnet. Sei F auf den Punkt $\pm 2a$ jeweils mit $P = \frac{1}{2}$ konzentriert. Dann ist

$$f(x, t) = \exp(ax - a^2 t) + \exp(-ax - a^2 t) \geq \epsilon$$

genau dann, wenn

$$h(ax) := \exp(ax) + \exp(-ax) \geq \epsilon \exp(a^2 t)$$

$$|ax| \geq h^{-1}(\epsilon \exp(a^2 t)), \text{ also } |x| \geq \frac{1}{a} h^{-1}(\epsilon \exp(a^2 t))$$

und damit

$$P(|S_n| \geq \frac{1}{a} h^{-1}(\epsilon \exp(a^2 \frac{n}{m})) \sqrt{m} \text{ für ein } n \geq 1) \leq \frac{1}{\epsilon} .$$

7) Kann für korrelierte Größen ein Power-One-Test konstruiert werden?

Die für die Materialbilanzen relevanten Variablen MUF_1, \dots, MUF_n sind normalerweise vielfältig korreliert. Insofern ist es wichtig zu wissen, ob die Unabhängigkeitsforderung beseitigt werden kann. Dann ist für jedes n $X = (X_i)$, $i=1, \dots, n$ eine n -dimensionale Normalverteilung mit bekannter regulärer Kovarianzmatrix Φ .

Es sei wieder P die Wahrscheinlichkeit bestimmt durch die n -dimensionalen Randverteilungen $N(0, \Phi)$ mit der Dichte (bis auf Faktor)

$\exp(-\frac{1}{2} x' \Sigma^{-1} x)$ und P_θ sei die Wahrscheinlichkeit, so daß die Randverteilungen $N(\theta, \Phi)$ verteilt sind, wobei θ in allen Komponenten den Wert θ annimmt.

P' sei eine Mischung $P'(\cdot) = \int P_\theta(\cdot) dF(\theta)$, wobei F eine beliebige Verteilung auf $(-\infty, \infty)$ ist. Damit ist (bis auf Faktor)

$$g'(x) = \int \exp[-\frac{1}{2}(x-\theta)' \Phi^{-1}(x-\theta)] dF(\theta)$$

$$z_n = \frac{g'}{g} = \int \exp[-\frac{1}{2}(-2\theta' \Phi^{-1} x + \theta' \Phi^{-1} \theta)] dF(\theta)$$

Dies ergibt mit der Funktion $f(x, t)$ aus 2)

$$z_n = f(\sum_1^n a_i x_i, e' \Phi^{-1} e).$$

Mit der Funktion A erhält man damit die Abschätzung

$$P(\sum_1^n a_i x_i \geq A(e' \Phi^{-1} e, \epsilon) \text{ für ein } n \geq 1) \leq \frac{1}{\epsilon} .$$

Man hat damit für diese spezielle Linearkombination der X_i eine Schrankenfunktion, so daß die Wahrscheinlichkeit des Übertritts abgeschätzt werden kann. Im Falle unabhängiger Variablen hat man gerade $\sum_1^n a_i X_i = S_n$.

Zunächst notieren wir, daß sämtliche Bemühungen, eine Abschätzung für die Summen S_n (die einen erwartungstreuen Schätzer für den abgezweigten Betrag darstellen) daraus zu erhalten, zu keinem Erfolg führen:

- man kann dies nicht durch geschickte Wahl von $\theta = (\theta_1, \dots, \theta_n)$ und F (n-dimensional) erreichen;
- es ist nicht möglich, auf allgemeine Weise die Partialsummen $\text{Max}_t X^m(t)$ mittels des Wiener-Prozesses zu approximieren und damit Aussagen für den Wiener-Prozeß auf diese Summen anschließend zu übertragen, speziell um die Fehlalarmrate zu kontrollieren;
- es lassen sich n Variablen immer durch höchstens n orthogonale Basisvariablen darstellen, also MUF_1, \dots, MUF_n läßt sich in eine orthogonale Familie (auf vielfache Weise) linear transformieren, z.B. mit Hilfe der L^2 -Projektion, die gerade der bedingte Erwartungswert ist, oder z.B. $MUF = A \cdot MUF_R = [MUF_1, MUF_2 - E(MUF_2 | MUF_1), MUF_3 - E(MUF_3 | MUF_1, MUF_2), \dots]$. Für die orthogonalen, normalverteilten transformierten Größen läßt sich nun die Approximation durchführen. Für die ursprünglichen Größen MUF_i ergibt die Rücktransformation keine Schranken für einen Sequentialtest $\{ \sum_{i=1}^k MUF_i \}$, da in den sukzessiven Teilsummen laufend neue Koeffizienten von MUF_R_1, \dots, MUF_R_n auftreten:

$$\begin{aligned} MUF_1 &= \alpha_{11} MUF_R_1 \\ MUF_1 + MUF_2 &= \alpha_{21} MUF_R_1 + \alpha_{22} MUF_R_2 \\ MUF_1 + MUF_2 + MUF_3 &= \alpha_{31} MUF_R_1 + \alpha_{32} MUF_R_2 + \alpha_{33} MUF_R_3 \dots \end{aligned}$$

Es wird sich in Kapitel III. zeigen, daß der Power-One-Test für korrelierte Größen nicht konstruiert werden kann; es ergibt sich nämlich gerade in dem für den Safeguardsbereich vorliegenden Fall von systematischen Meßfehlern (was jedenfalls eine spezielle Korrelationsstruktur darstellt) die Aussage, daß nicht gleichzeitig die lineare Entnahme entdeckt und α kontrolliert werden kann (= Robbins-Test). Mit anderen Worten, es ist damit ein Gegenbeispiel für korrelierte Größen gefunden, das für die Nuklearmaterialkontrolle eintritt.

III. Die Bestimmung der Test-Fehlalarmwahrscheinlichkeit α
bei einer Folge von Materialbilanzen

Es werden in diesem Teil die Methoden untersucht, um bei sequentiellen Techniken, angewandt auf die MUF-Zeitreihe, Aussagen über die Rate der auftretenden Fehlalarme zu erhalten, d.h. Alarme, die zu Unrecht angeben, daß eine Entwendung von überwachtem Material erfolgt ist. Es ist wichtig, hier garantierte Werte (d.h. strenge Abschätzungen) zu erhalten. Nur durch genaue Quantifizierung im Sinne konservativer Bestimmung von α kann sichergestellt werden, daß

- a) nicht laufend infolge der Datenanalyse weitere Nachforschungen (zur Prüfung, ob das Material ordnungsgemäß verwendet wird) angestellt werden, die lediglich aufgrund der schlechten Konstruktion des Entscheidungsverfahrens unnötigerweise veranlaßt sind (falls die tatsächliche Fehlalarmrate α zu hoch ist)
- b) aufgrund der Datenanalyse mit Hilfe von Entscheidungsverfahren mit Überzeugung vertreten werden kann, daß ein Verlust von überwachtem Material vorliegt und weitere klärende Schritte notwendig sind (falls die Fehlalarmrate genügend klein ist).

Es liege nun eine Folge von MUF-Werten vor, die eine zeitliche oder räumliche Aufgliederung von Materialbilanzen darstellen und alle eine Normalverteilung besitzen

$$\text{MUF}_i \sim N(\mu_i, \sigma_i^2)$$

In den Entscheidungsverfahren - deren Fehlalarmwahrscheinlichkeit also untersucht werden soll - werden diese Zufallsvariablen oder Transformationen (bisher lineare) davon betrachtet, die mit X_i bezeichnet werden sollen. Diese X_i sind wiederum normalverteilt. In jedem Fall ist also eine normalverteilte Folge

$$X_i \sim N(\mu_i, \sigma_i^2) \quad i = 1, \dots$$

gegeben. Insbesondere hat man damit für jedes endliche n einen normalverteilten Zufallsvektor $X = (X_1, \dots, X_n)' \sim N(\mu, \Phi)$, wobei Φ die Kovarianzmatrix bezeichnet. Unter Normalbedingungen sei X so normiert, daß dann $X \sim N(0, \Phi)$.

Es werden nun ausschließlich solche sequentiellen Testverfahren betrachtet, die auf jeder Zeitstufe nur zwischen

- a) Alarm
- b) kein Alarm und das Sequentialverfahren wird fortgeführt

entscheiden, wobei sowohl ein endlicher als auch ein unendlicher Zeithorizont zugelassen ist. In der Praxis kommen selbstverständlich nur endliche Zeithorizonte vor, jedoch hat der unendliche Horizont den wesentlichen Vorteil, daß der Kontrollzeitraum beliebig erhöht werden kann, ohne die Analyse dauernd wieder von neuem durchführen zu müssen (bezüglich der Festlegung von Signifikanzschranken, der Irrtumswahrscheinlichkeiten usw.).

Ein sequentielles Verfahren ist gegeben durch

$$\delta = (\delta_1(X_1), \delta_2(X_1, X_2), \dots, \delta_n(X_1, \dots, X_n))$$

mit $\delta_i(X_1, \dots, X_i) \in \{\text{Alarm, kein Alarm}\}$, wobei bei unendlichem Horizont $n = \infty$ (nicht abbrechende Folge) gesetzt wird. Hierfür wird $\alpha = P(\delta_i \in \{\text{Alarm}\})$ für ein $i, i=1, \dots, n$ definiert.

Es wird die Verabredung getroffen, daß das Entscheidungsverfahren nicht zu jedem Zeitpunkt i explizit definiert zu werden braucht, was praktisch heißt, daß entweder dort $\delta_i \equiv$ 'kein Alarm' gesetzt wird oder noch weitere Entscheidungsregeln hinzugefügt werden. Dies hat lediglich den Zweck, daß einmal für Beispiele vereinfachend nur zu gewissen Zeitpunkten δ_i erklärt zu werden braucht, um zu zeigen, daß dann α mindestens so groß wird. Zum anderen sind solche Verfahren mit drei Entscheidungsmöglichkeiten {Alarm, kein Alarm, kein Alarm und Abbruch des Verfahrens}, wie der Wald-Test, Page-Test, eingeschlossen, die die Zuordnung besitzen

$$\begin{aligned} &\{\text{Alarm, kein Alarm, kein Alarm und Abbruch}\} \\ &\rightarrow \{H_1, \text{Indifferenzzone}, H_0\}. \end{aligned}$$

Denn man kann dann dort, wo δ_i nicht definiert ist, wie oben beschrieben vorgehen (neu starten, neue Regeln hinzu, $\delta_i \equiv$ 'kein Alarm'). Auf jeden Fall wird dadurch α erhöht (oder bleibt günstigenfalls gleich).

Für ein zu konstruierendes Testverfahren ergeben sich die Probleme

- i) α darf einen vorgegebenen Wert nicht überschreiten
- ii) genaue Bestimmung von α .

Zum Beispiel ergibt ein hintereinander geschalteter Signifikanztest unabhängiger Variablen

$$X_1 \geq z_1 \quad \text{oder} \quad X_2 \geq z_2 \quad \text{jeweils zum Niveau } \alpha_i = 5 \%$$

schon eine Gesamtfehlalarmwahrscheinlichkeit $\alpha = 5 \% + 0,95 \cdot 5 \% \geq 9,7 \%$ und würde so fortfahrend laufend anwachsen. Unter den vielen Möglichkeiten, einzelne

Zeitabschnitte $\sum_{m_1}^{m_2} X_i$ auszuwählen, zu kombinieren, Testverfahren darauf anzuwenden,

dann Batterien zu bilden usw., interessieren jetzt solche Verfahren, die zum Zeitpunkt n aufgrund der kumulierten Summenfolge

$$S_n = \sum_{i=1}^n X_i$$

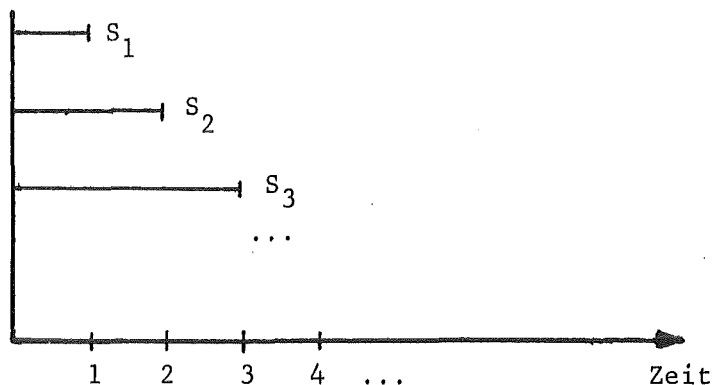
entscheiden und die Gestalt haben

$$S_n < z_n \quad \text{so führe den Test fort}$$

$$S_n \geq z_n \quad \text{Alarm}$$

oder dasselbe für $|S_n|$.

Im Diagramm:



Anzugeben ist hier nur noch die Folge der Schranken z_n .

A) Endlicher Zeithorizont n (n jedoch beliebig groß, fest gewählt)

Es läßt sich i) sofort erfüllen, d.h. $\bar{\alpha}$ vorgeben - kann extrem klein vorgegeben sein, etwa 0,01 % -, so gibt es Schranken z_k , so daß die Fehlalarmwahrscheinlichkeit den Wert $\bar{\alpha}$ nicht überschreitet. Indem man beispielsweise für jedes S_i , $i=1, \dots, n$ Signifikanztests zum Niveau $\frac{\bar{\alpha}}{n}$ wählt; dann ist die globale Fehlalarmwahrscheinlichkeit $\alpha \leq \sum_1^n \frac{\bar{\alpha}}{n} = \bar{\alpha}$, wie auch immer die X_i verteilt und korreliert sind.

Schwierig ist ii); es gibt keine Berechnungsmöglichkeit. Es ist lediglich möglich, (strenge) Abschätzungen anzugeben. Eine solche läßt sich sofort angeben für $(X_i) \sim N(0, \frac{1}{2})$, beliebig korreliert und beliebig gegebene Schranken z_k . Man berechne zuerst die Signifikanzniveaus α_k , dann gilt für das globale α

$$\alpha \leq \sum_1^n \alpha_k \quad \left[\alpha \leq \alpha_1 + (1-\alpha_1) \sum_{k=2}^n \alpha_k \text{ für } X_i \text{ unabhängig} \right]$$

Eine Bedeutung haben jedoch nur sehr gute Abschätzungen, die möglichst nahe an das globale α herankommen. Dies gerade ist das Ziel dieses Abschnitts und soll in erweiterter Form nochmals formuliert werden:

Ziel: Bestimme für gegebenes α Schrankenfolgen $\{z_k\}$ so, daß die globale Fehlalarmwahrscheinlichkeit streng durch α abgeschätzt wird und gleichzeitig dicht bei α liegt.

B) Unendlicher Zeithorizont

Wiederum ist es sehr einfach, für gegebenes α - kann extrem klein gewählt sein - Schranken z_k zu geben, so daß die Fehlalarmwahrscheinlichkeit den Wert α nicht überschreitet. Man kann jede Folge α_n wählen mit $\sum_1^\infty \alpha_n \leq \alpha$ und dazu die Signifikanzschranken z_k aus $\alpha_k = 1 - \Phi\left(\frac{z_k}{\sigma_k}\right)$, $\sigma_k^2 := \text{var } S_k$, berechnen. So gilt natürlich für das globale $\alpha_\infty \leq \sum_1^\infty \alpha_n \leq \alpha$. Das Hauptproblem besteht darin, diese Schrankenfunktion so tief wie möglich zu legen.

Wie schon in A) gibt es für den unendlichen Zeithorizont erst recht keine Berechnungsmöglichkeit von α . Schlimmer noch, es ist natürlich auch keine Monte-Carlo-Simulation durchführbar. Hier sind möglichst genaue Abschätzungen für das globale α der einzige Ausweg.

Möglichkeiten, um α zu bestimmen:

=====

1. Tabellen

Für die zwei- und dreidimensionale Normalverteilung gibt es ein Tabellenwerk (Owen/Ann. Math. Stat. 27, p. 1075-90). Da die Korrelationen in einfacher Weise berechnet werden können, wird hier ein Verfahren ausgearbeitet, für gegebene Schranken z_k

$$P(S_k \geq z_k \text{ für ein } k, k=1, \dots, n) = \alpha$$

wie folgt annähernd zu berechnen und gleichzeitig nach oben abzuschätzen:

$$P(S_k \geq z_k \text{ für ein } k, k=1, \dots, n) \leq P(S_1 \geq z_1) + [P(S_2 \geq z_2) - P(S_1 \geq z_1, S_2 \geq z_2)] + \dots + [P(S_n \geq z_n) - P(S_n \geq z_n, S_{n-1} \geq z_{n-1})]$$

Eine Abschätzung nach unten ergibt

$$\alpha \geq P(S_1 \geq z_1) + [P(S_2 \geq z_2) - P(S_1 \geq z_1, S_2 \geq z_2)] + \dots + [P(S_n \geq z_n) - \sum_{i=1}^{n-1} P(S_n \geq z_n, S_i \geq z_i)].$$

Das folgende Beispiel [3; p.80] mit 16 Perioden, $\sigma_i = \sqrt{\text{var } S_i} = \sqrt{\frac{i}{16}} \cdot 2,44 \text{ kg}$ (aus einem Test mit $\alpha_{16} = 5 \%$, $1-\beta = 95 \%$ für $M = 8 \text{ kg}$, $\sigma = 2,44 \text{ kg}$, X_i unabhängig) zeigt, daß damit α schon gut zwischen zwei Schranken eingekreist wird (die letzten beiden Spalten; die Schranken z_i sind implizit durch α_i gegeben).

i	σ_i	$\alpha_i \%$	$\Sigma \alpha_i$	benachbarte Korrelation berücksichtigt	Korrelation von 3 Variablen berücksichtigt
1	0,61	} $\leq 3 \cdot 10^{-3}$	$\leq 0,01\%$		
2	0,86				
3	1,06				
4	1,22				
5	1,36				
6	1,50	0,01%	0,02		
7	1,61	0,04	0,06		
8	1,73	0,13	0,19		
9	1,82	0,29	0,48		0,30
10	1,92	0,59	1,07	0,80	0,70
11	2,02	1,04	2,11	1,52	1,30
12	2,11	1,54	3,65	2,30	2,25
13	2,20	2,27	5,92	3,70	3,57
14	2,27	2,90	8,82	5,20	5,15
15	2,36	4,01	12,83		
16	2,44	5,00	17,83		

Die Berechnung durch Nachschauen in der Tafel ist allerdings wegen der Zwischenschaltung von weiteren tabellierten Hilfsfunktionen äußerst zeitaufwendig. Es gibt dazu fertige Subroutinen, die möglichst verwendet werden sollten. Weitere Verfeinerungen sind durch Berücksichtigen von drei korrelierten Variablen (mit Tafeln für vierdimensionale Normalverteilungen durch Berücksichtigung von Durchschnitten von je vier Mengen analog, usw.) wie folgt möglich:

$$\alpha \leq P(S_1 \geq z_1) + [P(S_2 \geq z_2) - P(S_1 \geq z_1, S_2 \geq z_2)] + \\ [P(S_3 \geq z_3) - \sum_{i=1}^2 P(S_3 \geq z_3, S_i \geq z_i) + P(S_3 \geq z_3, S_2 \geq z_2, S_1 \geq z_1)] + \dots \\ + [P(S_n \geq z_n) - \sum_{i=1}^{n-1} P(S_n \geq z_n, S_i \geq z_i) + \sum_{\substack{i,j=1 \\ i>j}}^{n-1} P(S_n \geq z_n, S_i \geq z_i, S_j \geq z_j)]$$

Gegebenenfalls können auch Abschätzungen für schwach korrelierte Variablen eingebaut werden, falls die Korrelation von S_i, S_j mit zunehmender Distanz schnell abnimmt. In diesem Fall liegt vermutlich die oben angeführte Abschätzung nach unten sehr nahe bei dem tatsächlichen Wert, was durch eine Simulationsstudie untersucht werden könnte.

2. Wiener-Prozeß

Für unabhängige Variablen X_i wurde in Zerrweck [3; p.34-35] gezeigt, indem die Varianzfunktion $\sigma^2(S_i)$ als neuer Zeitparameter eingeführt wird und die mit diesem Parameter indizierten Summen $\sum X_i$ in einen Brown'schen Prozeß eingebettet werden, daß dann mit den Bezeichnungen

$$\sigma_T^2 = \text{Varianz von } S_T \text{ für einen festen endlichen Horizont } T \\ ==$$

$(\sigma_{ij}^2)_{j=1, \dots, k_n}$ immer feinere Zerlegungen von σ_T^2 , d.h. $\sum_{j=1}^{k_n} \sigma_{ij}^2 = \sigma_T^2$ für alle i
 $==== n=1, \dots$

$$\text{und } \max_j \sigma_{ij}^2 \rightarrow 0 \text{ für } i \rightarrow \infty, \quad b_{ik} := \frac{\sum_{j=1}^k \sigma_{ij}^2}{\sigma_T^2}$$

$$\text{auf } [b_{ik}, b_{ik+1}) \text{ wird } S^i(t) = \sum_{j=1}^k X_{ij} \text{ gesetzt mit } X_{ij} \sim N(0, \frac{\sigma_{ij}^2}{\sigma_T^2})$$

die Aussage gilt:

$S^i(t) \xrightarrow{D} W(t)$, $W(t)$ der Wiener Prozeß auf $[0,1]$ und damit

$$\text{Prob}(\sum_{j=1}^k X_{ij} \geq a \text{ für ein } k, 1 \leq k \leq k_n) \rightarrow 2[1 - \Phi(a)] \text{ für } i \rightarrow \infty$$

Die Variablen X_1, \dots, X_n definieren eine solche Zerlegung $\sigma_1^2, \dots, \sigma_n^2$ von

$$\sigma_T^2 := \sum_1^n \sigma_i^2 \text{ und damit hat man mit der Identifizierung } X_{ij} := \frac{X_j}{\sigma_T}$$

$$P\left(\sum_{j=1}^k X_j \geq a \cdot \sigma_T \text{ f\u00fcr ein } k, 1 \leq k \leq n\right) \leq 2 \left[1 - \Phi\left(\frac{a}{\sigma_T}\right)\right] = 2 P(S_n \geq a \sigma_T)$$

Ergebnis: Die Wahrscheinlichkeit von Fehlalarmen f\u00fcr $z_K \equiv a$ l\u00e4\u00dft sich f\u00fcr den endlichen Horizont T absch\u00e4tzen durch

$$2 \left[1 - \Phi\left(\frac{a}{\sigma_T}\right)\right] \text{ mit } \sigma_T^2 = \text{var } S_T,$$

ist also maximal doppelt so hoch wie die Fehlalarmwahrscheinlichkeit f\u00fcr den Signifikanztest f\u00fcr S_T allein mit der Schranke a.

Nimmt man also eine Gerade $y \equiv a$ als Schranke, so liegt die Fehlalarmwahrscheinlichkeit zwischen α und 2α , falls $\alpha = 1 - \Phi\left(\frac{a}{\sigma_T}\right)$ die Signifikanz von S_T bezeichnet.

Bemerkenswert an dieser Absch\u00e4tzung ist, da\u00df beliebige Verfeinerungen des Bilanzierungszeitraums vorgenommen werden k\u00f6nnen (dann ist durch Simulationen die Aussage nicht zu erhalten) und der Proze\u00dfverlauf in σ_i stark schwankend sein kann.

Dieselbe Aussage gilt auch f\u00fcr beidseitige Schranken

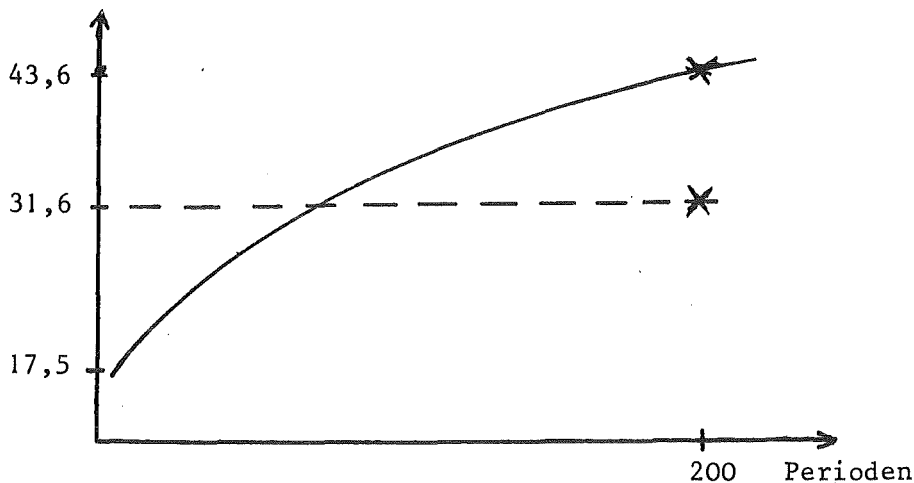
$$P\left(\left|\sum_{j=1}^K X_j\right| \geq a \sigma_T \text{ f\u00fcr ein } k, 1 \leq k \leq n\right) \leq 2 P(|S_n| \geq a \sigma_T)$$

und kann nun auf $MUFR_i$ oder $\frac{MUFR_i}{\sigma_i}$ oder MUF_i (falls unabh\u00e4ngig) angewandt werden.

Zum Vergleich ist wie in dem anschlie\u00dfenden Abschnitt 4. die Robbins-Schranke f\u00fcr $m = 50$, $\alpha = 5 \%$ und die f\u00fcr $\alpha = 5 \%$ erhaltene Gerade (beidseitig) aus

$$\alpha = 2 \cdot \left[1 - \Phi\left(\frac{a}{\sigma_T}\right)\right] \cdot 2, \sigma_T = \sqrt{200} \text{ also } a \approx 31,6$$

aufgetragen (nur die obere Schranke wurde im Schaubild gezeichnet).



3. Monte-Carlo-Simulation

Damit kann für endlichen Horizont α hinreichend genau für vorgegebene Schranken z_K effektiv bestimmt werden. Die Nachteile sind, daß immer wieder von neuem simuliert werden muß,

- a) wenn der Zeithorizont erhöht wird oder die Bilanzierungsintervalle verfeinert/vergrößert werden
- b) wenn der Prozeßverlauf schwankt, d.h. nicht stationär in bezug auf σ_i^2 ist oder verschiedene Anlagen überwacht werden (falls die Variablen X_i im Test normiert vorkommen $X_i \sim N(0,1)$ und unabhängig sind, fällt dieser Punkt weg)
- c) wenn der Schrankenverlauf gegeben ist, aber erst noch auf das vorgegebene α justiert werden muß (Schranke noch von Parametern abhängig).

Ein wesentlicher Nachteil besteht darin, daß man mittels Simulation nicht feststellen kann, ob eine für beliebig lange Zeithorizonte vorgegebene Schrankenfolge (damit der Bilanzierungszeitraum beliebig erhöht werden kann) eine globale Fehlalarmwahrscheinlichkeit $\leq \alpha$ besitzt.

4. Robbins-Schranke

In Robbins, Siegmund [1] wurde die Schranke

$$z_n = \left[(n+m) \left\{ a^2 + \log\left(\frac{n}{m} + 1\right) \right\} \right]^{\frac{1}{2}}, \quad a \geq 0, \quad m > 0$$

- sowohl für S_n als auch $|S_n|$ - für unabhängige $N(0,1)$ verteilte X_i gegeben. Für vorgegebenes α wird a berechnet (siehe Teil I), m frei gewählt und dann gilt, daß die Fehlalarmwahrscheinlichkeit kleiner als α ist. Diese Abschätzung gilt für beliebige Zeithorizonte (unendlicher Horizont), umgeht also die Nachteile a) und b) in 3. Man weiß jedoch nicht, wie gut die Abschätzung für ein fest gewähltes m ist, erst recht nicht für den endlichen Horizont, für den ja die Fehlalarmwahrscheinlichkeit noch einmal kleiner ist.

So wurde in Beispielen mittels Monte-Carlo-Simulation in Sellinschegg [2] berechnet, daß selbst für große m (für $m \rightarrow \infty$ wird die Abschätzung genau) und große Horizonte n die Abschätzung sehr schlecht ist: $m = 50$, $n = 200$, $\alpha = 5\%$ ergibt ein $\alpha_{200} \approx 1\%$ und ähnliche Beispiele.

Im folgenden wird ein Verfahren angegeben, wie diese Abschätzung wesentlich verbessert werden kann.

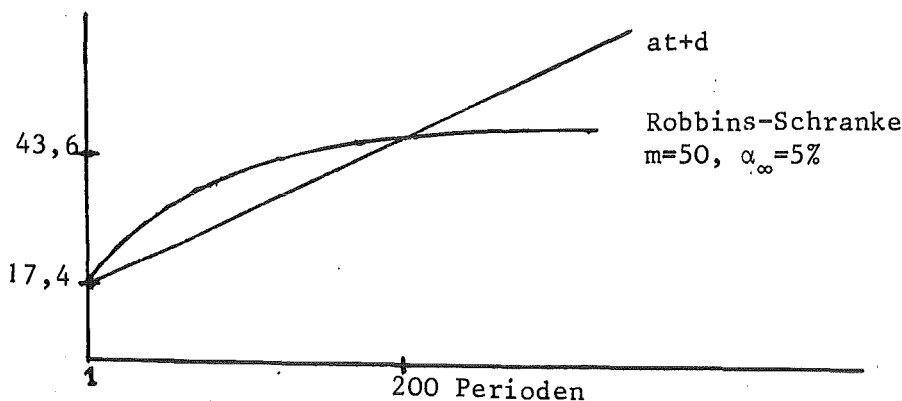
Für eine Gerade als Schrankenfunktion ergibt sich

$$P(S_n \geq an+d \text{ für ein } n \geq 1) \leq e^{-2ad} \quad (a, d > 0).$$

Bei einem endlichen Horizont können immer Geraden unterhalb der benutzten Schrankenfunktion eingeführt werden, die eine Abschätzung der Fehlalarmwahrscheinlichkeit ergeben

$$P(S_k \geq z_k \text{ für ein } k, 1 \leq k \leq n) \leq P(S_k \geq ak+d \text{ für ein } k, 1 \leq k \leq n) \leq e^{-2ad}$$

falls $z_k \geq ak+d$ für $1 \leq k \leq n$



Für das in [1] simulierte Beispiel wird die Abschätzung jetzt durchgeführt.

Die Steigung a ergibt sich aus $a = \frac{z_{200} - z_1}{199}$, wobei z_i aus der Formel der Robbins-Schranke berechnet wird, und d aus $d = z_1 - a$. Damit erhält man die Werte

$$z_{200} = \sqrt{250[\log 400 + \log 5]} = \sqrt{250 \cdot \log 2000} = \sqrt{1900} \approx 43,6$$
$$z_1 = \sqrt{51[\log 400 + \log(\frac{1}{200} + 1)]} = \sqrt{51 \cdot \log 402} = \sqrt{306} \approx 17,5$$

und daraus $a \approx 0,131$ $d \approx 17,4$.

Damit also die strenge Abschätzung

$$\alpha_{200} = P(S_k \text{ liegt über Robbins Kurve in } [1,200]) \leq P(S_k \text{ liegt über Gerade}) \leq e^{-2ad} = e^{-2 \cdot 0,13 \cdot 17,4} = \frac{1}{96} \approx 1,04 \%$$

für den einseitigen Test. Für den zweiseitigen Test hat man damit jedenfalls die sichere Abschätzung $\alpha_{200} \leq 2 \times 1,04 \%$.

Da die Steigung der Geraden für weniger Perioden anwächst, ergeben sich dort noch bessere Werte (z.B. nur alle 5, 10 Tage bilanzieren, ergibt 40 bzw. 20 Perioden). Für den sehr großen Periodenzeitraum $n = 1000$ wird ein Polygonzug anstelle der Geraden genommen (diese ergäbe $\alpha_{1000} \leq 6,6 \%$), d.h. die erste Gerade durch $z_1 = 17,5$ und $z_{200} = 43,6$ gelegt, ergibt Abschätzung in $[1,200]$, die zweite durch $z_{200} = 43,6$ und $z_{1000} = 95$, ergibt Abschätzung in $[200, 1000]$.

$$a_2 = \frac{95-43,5}{800} \approx 0,0645 \quad b_2 \approx 30,6 \quad , \quad \text{damit } \alpha_{1000} \leq \alpha_{200} + 1,9 \% = 2,94 \%$$

Ergebnis: Für den endlichen Horizont $[1,200]$ kann sowohl die Robbins-Schranke als auch die Gerade als Alarmschwelle gewählt werden. In beiden Fällen ist α_{200} besser abschätzbar als mit dem bekannten α_∞ Wert. Für eine Erhöhung des Horizontes kann in $[1,200]$ die Gerade, danach die Robbins-Schranke gewählt werden ($P(\text{Fehlalarme}) \leq \alpha_\infty + \alpha_{200}$).

IV. Protracted Loss und systematischer Fehler bei Anlagen mit hohem Durchsatz

1) CUMUF ist ein Gleichgewichtstest

Falls eine Reihe von MUF-Daten in den aufeinanderfolgenden Zeitabschnitten 1,2,... aus Messungen vorliegt, ergibt sich die Frage, welches statistische Entscheidungsverfahren δ anzuwenden ist um herauszufinden, ob eine Entnahme von Material stattfand.

Für eine bekannte Entwendungsstrategie $\mu=(\mu_i)$, wobei im Zeitabschnitt i die Menge μ_i entnommen wird, hat man somit zwischen den Hypothesen

$$H_0: E(MUF_i) = 0 \quad \text{für alle } i \quad | \quad \text{keine Entnahme}$$

$$H_1: E(MUF_i) = \mu_i \quad \text{für alle } i \quad \quad \quad i = 1, \dots, n$$

zu entscheiden. Dabei ist $X=(MUF_i) \sim N(E(X), \Sigma)$ eine n-dimensionale Normalverteilung mit bekannter Kovarianzmatrix Σ .

Der Neyman-Pearson-Test ist derjenige Test mit der höchsten Entdeckungswahrscheinlichkeit für fest vorgegebene Fehlalarmwahrscheinlichkeit α und wird gegeben durch den kritischen Bereich

$$(1) \quad X' \Sigma^{-1} \mu > c_\alpha$$

c_α ist so zu wählen, daß $P_0(X' \Sigma^{-1} \mu > c_\alpha) = \alpha$ wird. Dann ist die Entdeckungswahrscheinlichkeit gegeben durch

$$(2) \quad 1-\beta = \Phi(\sqrt{\mu' \Sigma^{-1} \mu} - U_{1-\alpha})$$

wobei Φ die Normalverteilungsfunktion und U die inverse Funktion bezeichnet. Dieser Wert gibt also eine obere Grenze an und für andere Testverfahren kann immer verglichen werden, um wieviel höher das beste Testverfahren für eine feste Entnahmestrategie wäre.

Es ist natürlich, verschiedene Testverfahren so zu vergleichen, daß eine Gesamtmenge M festgehalten wird, alle Verfahren eine gleichgroße Fehlalarmwahrscheinlichkeit besitzen und dann $1-\beta$ optimiert wird. $1-\beta$ ist nun sowohl von dem Verfahren δ als auch von der gewählten Strategie (μ_i)

mit $\sum_1^n \mu_i = M$ abhängig. In (2) wurde $\text{Max}_\delta 1-\beta(\mu, \delta)$ bestimmt und die

Minimierung dieses Ausdrucks ergibt also die für den Entwender beste Entnahmestrategie, wenn der Inspektor jeweils den besten Test für jedes μ wählen würde. Diese Minimierung ergibt mit der Lagrange-Methode

$$\mu^* = \frac{M}{\text{var}(e^T X)} \quad \text{Ze mit dem zugehörigen Neyman-Pearson-Test } \sum_1^n \text{MUF}_i$$

(Avenhaus [4]), d.h. falls die Wahl von μ durch den Entwender dem Inspektor bekannt ist, so wird mindestens mit $1-\beta_{NP}(\mu^*)$ entdeckt, indem der Inspektor immer den Neyman-Pearson-Test wählt.

Es liegt nun gerade der umgekehrte Sachverhalt vor, daß der Entwender das Testverfahren des Inspektors im voraus kennen kann, der Inspektor dagegen die Entwendung μ nicht kennt, d.h. wählt er irgendeinen Test δ - insbesondere also einen aus der Gesamtheit aller NP-Tests, die in der betrachteten $\text{Min}_\mu \text{Max}_\delta 1-\beta(\mu, \delta)$ Zielfunktion gerade auftreten - so hat der Entwender die Möglichkeit, die zugehörige Minimalstrategie $\text{Min}_\mu 1-\beta(\mu, \delta)$ zu wählen und damit deutlich unterhalb der Zielfunktion zu liegen, oder anders gesagt zu entwenden, ohne eine Entdeckung zu riskieren. Insofern muß umgekehrt vorgegangen werden. Den Inspektor interessiert also die garantierte Entdeckungswahrscheinlichkeit $\text{Min}_\mu 1-\beta(\mu, \delta)$ für ein festes Verfahren δ und damit das beste Verfahren, das dieses Minimum am größten macht, d.h. $\text{Max}_\delta \text{Min}_\mu 1-\beta(\delta, \mu)$. Dies wurde in [3] (p. 19) gelöst, indem die Gleichheit der Zielfunktionen gezeigt wurde ($\sum_1^n \text{MUF}$ tritt damit als Gleichgewichtspunkt auf) und man hat hiermit den

Satz: Für eine n-dimensionale Normalverteilung mit bekannter, regulärer Kovarianzmatrix Σ (insbesondere also für MUF_i) sei $1-\beta(\delta, \mu)$ die Entdeckungswahrscheinlichkeit zu einem Hypothesentest δ für (H_0, H_1) mit festem Niveau α , $\sum_1^n \mu_i = M$:

$$\text{unter } H_0 \text{ sei } E(X) = 0$$

$$\text{unter } H_1 \text{ sei } E(X) = (\mu_i), \quad \sum_1^n \mu_i = M.$$

Dann gilt

$$\text{Max}_{\delta \in \Delta_\alpha} \text{Min}_{\mu \in H_1} 1-\beta(\delta, \mu) = \text{Min}_{\mu \in H_1} \text{Max}_{\delta \in \Delta_\alpha} 1-\beta(\delta, \mu) = 1-\beta_\Sigma.$$

Hierbei ist der Sattelpunkt gegeben durch den Summentest $\sum_1^n X_i$, d.h.

$$\delta^* = \begin{cases} 1 & \sum_{i=1}^n X_i > s_\alpha \\ 0 & \text{sonst} \end{cases} \quad (s_\alpha \text{ Signifikanzschwelle})$$

Man hat damit das wichtige Ergebnis, daß der Summentest im Sinne garantierter Entdeckung (für beliebige Entnahmestrategien) das beste Verfahren darstellt. Eine Reihe von Folgerungen ist sehr wertvoll, um weitere Testverfahren T zu prüfen:

- a) Der Entwender kann die Strategie μ^* wählen. Somit ist $\sum X_i$ der Neyman-Pearson-Test, der Test δ mit größtem $1-\beta(\delta, \mu^*)$. Für das Testverfahren T ergibt sich notwendig ein kleinerer Wert $1-\beta(T, \mu^*) \leq 1-\beta(\sum, \mu^*)$.
- b) Die Minimalstrategie $\tilde{\mu}$ für das Verfahren T kann deutlich unter dem garantierten Wert $1-\beta_\Sigma$ liegen, da ja $1-\beta(T, \tilde{\mu}) < 1-\beta(T, \mu^*) \leq 1-\beta_\Sigma$ gilt.
- c) Andererseits kann für eine gegebene Strategie μ T bessere Werte ergeben als der $\sum X_i$ Test; der beste Wert wurde schon in (1) und (2) angegeben.

2. Die Minimalstrategie μ^*

Es wird das in [2] zugrunde gelegte Beispiel mit den Daten übernommen. Dort wurde eine Referenzanlage betrachtet mit 1000 Tonnen Durchsatz pro Jahr und 200 Betriebstagen pro Jahr.

Die Minimalstrategie μ^* ergibt sich mit den Daten für die randomisierten und systematischen Fehler zu (siehe [3] p. 38)

$$(3) \quad \mu^* = \frac{M}{\text{var } \sum X_i} \begin{pmatrix} \sigma^2(I_0) + \sigma^2(D_1) + (n-1)m^2 \sigma_s^2 \\ \sigma^2(D_2) + (n-1)m^2 \sigma_s^2 \\ \vdots \\ \sigma^2(I_n) + \sigma^2(D_n) + (n-1)m^2 \sigma_s^2 \end{pmatrix}$$

Dabei wurde die Notation benutzt

- $\sigma^2(I_0)$ die Varianz von dem Anfangsinventar I_0
- $\sigma^2(D_i)$ die Varianz von dem Transfer in der Periode i
- n die Anzahl der betrachteten Perioden

- m die Gewichtsmenge des Transfers in der Periode i (stationär)
- σ_s der systematische Fehler bei der Messung des Transfermaterials.

Die eingesetzten Daten ergeben eine Aufteilung von μ^* , die praktisch gleichverteilte μ_i bedeutet. Die Korrektur $\sigma^2(I_0)$ und $\sigma^2(I_n)$ ergibt z.B. bei 200 Perioden in der ersten und letzten Periode den ca. 1,1-fachen Betrag von den restlichen Perioden. Die Standardabweichung von $\sum_{i=1}^{200} X_i$ beträgt $\sigma_{200} = 103$ kg. Damit ergibt sich für einen Gesamtbetrag von $M = 40$ kg für μ^* eine Entdeckungswahrscheinlichkeit von $\Phi(\frac{40}{103} - 1,64) = 10,75$ %. Möchte man mit 95% entdecken (zu $\alpha=5\%$), so ist dies erst für eine Entnahme von 338 kg der Fall. Wegen der Neyman-Pearson-Eigenschaft läßt sich dies mit keinem Verfahren verbessern, denn jeder Test gibt eine Aufteilung von \mathbb{R}^n in die Mengen H und H^c mit $X \in H^c$ so Alarm, $X \in H$ so kein Alarm und der Neyman-Pearson Test gibt gerade die Menge H_0 mit dem größten $P_{\mu=\mu^*}(H_0^c) = 1-\beta$ unter allen solchen Zerlegungen mit $P_{\mu=0}(H^c) \leq \alpha$.

Insbesondere ergibt jedes sequentielle Verfahren keine Erhöhung der Entdeckungswahrscheinlichkeit, denn dann wird mit jeder Zeitstufe eine Alarmmenge $H_i^c \subset \mathbb{R}^i$ definiert (\tilde{H}_i sei definiert $\tilde{H}_i = H_i \times \mathbb{R}^{n-i}$) und $H := \bigcap_{i=1}^n \tilde{H}_i$ ergibt einen 'nichtsequentiellen' Test mit dem gleichen Alarmverhalten, der also keine Erhöhung bewirkt. Es ist offensichtlich, daß auch ein sequentieller Test mit Indifferenzmengen unter dieses Schema fällt (Vereinigung der Indifferenzmenge mit H_i ; H_n kann auch anders aufgeteilt werden). Das gleiche gilt für Testbatterien.

Es gilt sogar die schärfere Aussage, daß auch zu jedem beliebigen Zeitpunkt in $[1, n]$ irgendein Test zum Niveau α keine höhere Entdeckung liefert (siehe [3], p. 56).

Satz: Jeder Neyman-Pearson-Test (also bester Test) über einen kleineren Zeitraum ergibt eine schlechtere Entdeckung

$$\Phi\left(\frac{\sum_{i=1}^k \mu_i^*}{\sqrt{\text{var} \sum_{i=1}^k X_i}} - U_{1-\alpha}\right) \leq \Phi\left(\frac{\sum_{i=1}^n \mu_i^*}{\sqrt{\text{var} \sum_{i=1}^n X_i}} - U_{1-\alpha}\right)$$

$$\text{Beweis: } \sum_1^k \mu_i^* = \frac{M}{n} \cdot \text{cov} \left(\sum_1^k X_i, \sum_1^n X_i \right) = : M'$$

$$\frac{M'}{\sqrt{\text{var} \sum_1^k X_i}} = \frac{M}{\sqrt{\text{var} \sum_1^n X_i}} \cdot \frac{\text{cov}(\sum_1^k X_i, \sum_1^n X_i)}{\sqrt{\text{var} \sum_1^k X_i} \cdot \sqrt{\text{var} \sum_1^n X_i}} \leq \frac{M}{\sqrt{\text{var} \sum_1^n X_i}}$$

nach der Cauchy-Schwarzschen-Ungleichung. \square

Es muß nun der Betrag $\sum_1^k \mu_i^*$ in der Periode $[1, k]$ entsprechend der hierfür berechneten Minimalstrategie $\tilde{\mu}_1, \dots, \tilde{\mu}_k$ aufgeteilt werden, so weiß man damit, daß jeder Test schlechtere Entdeckung als der Neyman-Pearson-Test zu $\tilde{\mu}_1, \dots, \tilde{\mu}_k$ liefert und dieser ist kleiner als derjenige für $[1, n]$, wie gerade gezeigt wurde.

Selbst wenn die Strategie μ^* unverändert belassen wird, ergibt sich keine Möglichkeit, mit einem sequentiellen Test über einen kleineren Zeitraum eine höhere Entdeckung zu erzielen.

Satz: Jeder sequentielle Test ergibt für μ^* eine Entdeckungswahrscheinlichkeit

$$1 - \beta \leq \Phi \left(\frac{M}{\sqrt{\text{var} \sum_1^n X_i}} - U_{1-\alpha} \right)$$

über beliebigen Zeitintervallen in $[1, n]$.

Beweis: Wäre für einen Test in $[1, k]$ diese höher, so ergibt sich sofort ein "nichtsequentieller" Test in $[1, k]$ wie oben beschrieben mit $\tilde{H}_i = H_i \times \mathbb{R}^{k-i} \quad i=1, \dots, k$. Dann wäre der Test $H = \bigcap_1^k \tilde{H}_i \times \mathbb{R}^{n-k}$ ein Test für den gilt:

$P_{\mu=0}(H^c) = P_{\mu=0} \left[\left(\bigcap_1^k \tilde{H}_i \right)^c \right] \leq \alpha$ denn jeder betrachtete Test muß eine Fehlalarmwahrscheinlichkeit $P_F \leq \alpha$ von vornherein besitzen.

$$P_{\mu=\mu^*}(H^c) = P_{\mu=(\mu_1^*, \dots, \mu_k^*)} \left[\left(\bigcap_1^k \tilde{H}_i \right)^c \right] > \Phi \left(\frac{M}{\sqrt{\text{var} \sum_1^n X_i}} - U_{1-\alpha} \right),$$

d.h. H würde einen Test definieren, der besser als der Neyman-Pearson-Test zu μ^* wäre, was nicht sein kann. Die Gleichheiten bei den Zylindermengen ergeben sich wegen der Konsistenzeigenschaft bei Maßen für stochastische Prozesse. \square

Es ist sofort ersichtlich, daß hier jede beliebige Teilmenge $J \subset \{1, 2, \dots, n\}$ gewählt werden kann anstelle $\{1, \dots, k\}$, d.h. der sequentielle Test kann mit beliebiger Wahl der Daten X_1, \dots, X_n aufgebaut werden, seine Entdeckung ist jedenfalls geringer. Das Ergebnis besagt, daß bei optimaler Entnahme (für den großen Bereich) es nicht möglich ist, durch einen Zwischentest mit sequentiellen Verfahren selbst mit demselben α ein besseres Testergebnis zu bekommen.

Aus dem bisherigen Vorgehen ist sofort ersichtlich:

- a) Eine Verschiebung der Entscheidungspunkte bzw. Meßpunkte für die Bilanzen t_1, \dots, t_n ergibt lediglich andere Kovarianzmatrizen, aber keine Änderung des maßgeblichen Quotienten $\frac{M}{\sqrt{\text{var} \sum_{i=1}^n X_i}}$ für die Entdeckung (zu dem neuen μ^*). Diese ist nur bestimmt durch M , α und die aufsummierte Rauschvarianz aus den Messungen. Der Kalibrierungsplan muß selbstverständlich festgehalten werden.
- b) Jede Verfeinerung der Zwischenbilanzpunkte oder Reduzierung dieser d.h. Verfeinerung oder Vergrößerung der Partition $\{1, \dots, 200\}$ ergibt zu dem hierfür gewählten μ^* die gleiche (beste) Entdeckung, verbessert den Test damit nicht.
- c) Der zu dieser Minimalstrategie μ^* gehörige beste Test, der CUMUF-Test, ist invariabel gegenüber sämtlichen Parameteränderungen: außer den Punkten a), b) sei aufgezählt, daß CUMUF nicht von M , der Gesamtentnahme, α , dem Kalibrierungsplan, Durchflußverlauf (instationär, stationär), Meßsystem usw. abhängt.

Folgerung: Es ist damit gezeigt worden, daß die Minimalstrategie μ^* (protracted loss) mit $\Phi\left(\frac{\mu^*}{\sqrt{\text{var} \sum_{i=1}^n X_i}} - U_{1-\alpha}\right)$ bestenfalls entdeckt wird und resistent gegen sämtliche Verbesserungsversuche durch Schätzungen, sequentielle Zwischenteste und Testkombinationen ist. Die Schätzungen insbesondere können sowohl linear (Kalman-Filter, L^2 -Approximation, Zeitreihen, Smoother) als auch nichtlinear sein, notwendig sind diese aus den Variablen X_1, \dots, X_n aufgebaut, fallen mithin unter die angestellte Betrachtung mit den Zylindermengen. Eine Diskussion anderer Gesichtspunkte findet man in [3, Kapitel III].

3. Anlagen mit hohem Durchsatz

Es wird das in Karlsruhe ausgearbeitete Referenzbeispiel für eine Reprocessing Anlage übernommen mit einer Kapazität von 1000 Tonnen Metall pro Jahr und vorausgesetzt, daß die darin angegebenen Daten (diese sind für jedes zu entwickelnde statistische Kontrollsystem die entscheidende Determinante, auf welche Weise es schließlich zu konstruieren ist) die Praxis genau treffen.

Benötigt für die statistische Analyse werden genau folgende Daten

$\sigma^2(I_k)$ die Varianz des Lagers zu den Zeitpunkten k

$\sigma_r^2(T_k)$ die Varianz des Transfers im Zeitraum $[k, k+1]$, randomisierter Fehleranteil

$\sigma_s^2(T_k)$ die Varianz des Transfers im Zeitraum $[k, k+1]$, systematischer Fehleranteil.

Diese Daten sind gegeben (die Grundeinheit ist hier 1 kg und der Zeitabstand beträgt jeweils ein Tag)

$$\sigma^2(I_k) = 4,46 \quad \sigma_r^2(T_k) = 0,091 \quad \sigma_s^2(T_k) = 0,515^2 = 0,2645$$

Da sämtliche Meßfehlerquellen präsent sind, eignet sich dieser Datensatz nebenbei dazu, alle wesentlichen Aspekte für die dynamische Kontrolle beliebiger MUF-Sequenzen (nicht nur aus Reprocessing Anlagen) klarzumachen. Aber ein weiterer Grund hierzu besteht darin, daß die Daten außerordentlich schlecht für das Safeguardsziel gelagert sind, so daß das zu entwickelnde Entscheidungsverfahren die gegebene Dateninformation notwendig in optimaler Weise verwerten muß und sich für besser gelagerte Fälle erst recht eignet.

Die gesamte Standardabweichung beträgt nämlich nach Ablauf eines Referenzzeitraums von 200 Tagen (1 Arbeitsjahr) $\sigma_{200} = 200 \cdot 0,515 = 103$ kg, da keine Neukalibrierung der Transfermeßinstrumente vorgenommen wird. Falls also ein Test nach Ablauf von 200 Tagen durchgeführt wird, ergibt sich ein Betrag von $3,28 \times 103 = 338$ kg, der bei $\alpha=5\%$ mit Wahrscheinlichkeit 95% entdeckt wird; 40 kg werden mit 10,75% entdeckt, 8 kg mit 5,9%. Also selbst bei Zulassung einer erheblichen Fehlalarmwahrscheinlichkeit von 5 % völlig ungenügende Werte für die Kontrolle. Für eine angenommene Entnahme von $M = 40$ kg heißt dies praktisch, daß diese nicht entdeckt werden kann: es tritt keine signifikante Abweichung vom Normalzustand $M = 0$ auf.

Die Situation bleibt genauso prekär, wenn man sämtliche bisher vorgeschlagenen Testverfahren in Betracht zieht, sie ist es nämlich aus prinzipiellen Gründen, wie nun genauer im Anschluß an die Abschnitte 1. und 2. ausgeführt wird. In vielen Artikeln werden Verfahren hergeleitet, die schneller entdecken oder gewisse Strategien außerordentlich gut d.h. mit hoher Wahrscheinlichkeit entdecken (was im Grunde dasselbe ist, denn es sind immer im Vergleich zur gleichförmigen Entnahme zeitlich konzentrierte Strategien, die sobald sie durch das Verfahren mit hoher Wahrscheinlichkeit erfaßt werden, automatisch auch schnell entdeckt werden; dies wird noch weiter behandelt), indem aus den einzelnen MUF-Variablen sequentiell Verfahren aufgebaut werden, diese miteinander kombiniert werden, Schätzverfahren eingebaut werden usw. Aber dies ist für den Datensatz der Reprocessing Anlage (und für jeden Datensatz, der ein σ_{Gesamt} in dieser Größenordnung ergibt) gar nicht das Problem, es ist lediglich ein Unterproblem!

Das zentrale Problem besteht darin, eine einzige Strategie überhaupt zu entdecken, nämlich die Minimalstrategie μ^* . Denn es ist nicht nur vom praktischen Safeguardsstandpunkt aus sinnlos, möglichst viele Strategien immer besser zu entdecken, dann immer noch weitere Strategien hinzuzunehmen, die wiederum entdeckt werden usw., wenn schließlich ständig diese Strategie übrigbleibt, die nicht entdeckt wird. Dann hat man nur noch eine bedingte "Kontrolle", unter der Bedingung, daß der Entwender sich wohlverhält und ausschließlich die erfaßten Strategien wählt und zu allem hin noch auf diese feste (wohlbekannte) Strategie verzichtet! Sondern es ist von der statistischen Seite her gesehen der Kern des Problems, diese Strategie allein optimal zu erfassen:

a) jede andere Entnahme ist bezüglich $\frac{\text{MUF}_i}{\sigma_i}$, genauer $\frac{\text{MUF}_i}{\sqrt{\text{cov}(X_i, \Sigma X_i)}}$, zeitlich konzentrierter, ergibt anders gesagt - zu gewissen Zeitpunkten i jedenfalls - höhere Ausschläge vom Nullzustand aus und wird von einem vernünftigen Verfahren mit zumindest gleicher Güte entdeckt. Sämtliche anderen Strategien sind sozusagen für die Mindestentdeckung schon in μ^* mitberücksichtigt.

b) Hat man im ersten Schritt diese Aufgabe gelöst, so gibt es zumindest eine Möglichkeit, das Unterproblem der schnellen Entdeckung sofort ohne Komplikationen und für die Praxis völlig ausreichend zu lösen. Falls nämlich in dem als Lösung vorgeschlagenen Verfahren nicht schon sequentiell entschieden wird, so wird es

sequentiell aufgebaut, indem der Endpunkt über die Zeitabschnitte wandert, die somit erhaltenen Verfahren V_1, \dots, V_N jeweils die auf $[1, k]$ konzentrierte Entnahme von $M(\text{fest})$ mit der gleichen Wahrscheinlichkeit von z.B. 95% anzeigen und α mit den in Kapitel II. beschriebenen oder analogen Methoden bestimmt wird. Auf jeden Fall geht es mit Monte-Carlo-Simulation. Dieses Problem ist sicher pragmatisch zu bearbeiten und wird noch näher diskutiert. Hier kommt es auf einen ganz anderen Punkt an und zwar darzustellen, daß die Suche nach optimalen Lösungen oder auch nur guten - eine Optimalität kann für das Unterproblem in der bisherigen Form gar nicht definiert werden - überhaupt keinen Schritt das eigentliche Problem vorwärtsbringt.

Präziser und quantitativ ist in aller mathematischen Strenge in 1. und 2. gezeigt worden, daß für μ^* und für den gegebenen Datensatz der beste Test: CUMUF mit den schon beschriebenen Werten anspricht, also klar gesagt, 40 kg nicht entdeckt. Weiterhin ist gezeigt worden, daß sequentielle Verfahren nichts daran ändern, es gibt prinzipiell niemals ein Entscheidungsverfahren, das daran etwas ändert (Neyman-Pearson!). In dieser Situation sind eine große Zahl von Verfahren von verschiedenen an dem Problem arbeitenden Gruppen entwickelt worden, die schneller entdecken oder äquivalent einen Teil der Strategien sehr hoch entdecken, aber $|\mu^*| = 40$ kg mit noch geringerem $1-\beta$ erfassen, d.h. erst recht nicht entdecken (notgedrungen, wegen Neyman-Pearson). Hierbei wird also unterhalb dem besten Wert, der sozusagen schon die Schlacht verliert, optimiert. Es liegt jedoch das Kontrollproblem bereits mit einer höheren Schwierigkeitsstufe vor, in voller Allgemeinheit nämlich, d.h. die raum-zeitliche Gesamtbilanz ist schon für den Grundüberwachungszeitraum mit dem sie charakterisierenden σ in eine neue Größenordnung vorgestoßen.

Fazit: Es geht bei dem Datensatz nicht um den Zielkonflikt schnelle Entdeckung - hohe Entdeckung, sondern Entdeckung überhaupt. Alle Versuche, dies mit irgendwelchen sequentiellen Verfahren zu erreichen, sind prinzipiell fruchtlos; das zu optimierende Ziel so nicht zu finden. Die bisherige Vorgehensweise und statistische Analyse befaßt sich konzeptmäßig mit einer anderen Frage. Das Konzept, das alle Verfahren verfolgen (schnelle, hohe Entdeckung von gewissen Strategien, damit Unterteilung in MUF-Sequenzen, dann sequentielle Verfahren, obiger Zielkonflikt) gehört in seinem Kern in das Verifikationssystem und nicht in das Grundbilanzsystem.

Im Verifikationssystem wird nämlich immer genau dieses Problem gelöst: eine über einen raum-zeitlichen Bereich erhaltene MUF-Größe wird in Teilbereiche zerlegt bis hin zu Elementarbilanzen, die eine Messung an einem einzigen Elementarbatch bedeuten, das nicht mehr weiter zerlegt wird oder zerlegt werden kann, vergl. [3;p.50] . Dann wird optimal ausgewählt, welche dieser Elementarbilanzen zu welcher damit gestuften Signifikanz überprüft wird. Man kann es sich als ein Sieb vorstellen, wobei zuerst die hohe Signifikanz an vielen Elementarbilanzen (oder auch an allen) ausgeführt wird, dann die nächstniedere Signifikanz an einem ausgewählten Teil (wobei Teil auch Aggregation von Elementarbilanzen bedeuten kann) usw. und dieser Vorgang bedeutet in seiner zeitlichen Komponente das Unterproblem. Genauso ist die vorgeschlagene Zusammenkopplung von klassischen MUF-Sequenzen hier einzuordnen, wobei die einzelnen MUF-Größen zu der Hierarchie der Elementarbilanzen genommen werden und gefragt wird, ob eine solche MUF-Größe eine hohe Signifikanz ergibt oder erst mehrere zusammen.

4. Die gestufte Zufallsstruktur der Meßfehler bei Transfermessungen

Wie also schon in 3. ausgeführt, ist die Entdeckung von μ^* das zentrale Problem. Woran liegt es, daß diese eigentlich so schwierig ist? Zuerst einmal an dem hohen $\sigma_{\text{Gesamt}} = 103 \text{ kg}$. Ohne eine Verminderung ist die Entdeckung hoffnungslos wie in 2. gezeigt. Deswegen ein Brückenschlag zu einem anderen Fragenkomplex, der, wie sich zeigt, äquivalent zu dem gegebenen Problem ist. Man untersucht σ_{Gesamt} und sagt gewöhnlich: Verkleinere $\sigma^2(I_k)$ und $\sigma_r^2(T_k)$ durch wiederholtes Messen sowie $\sigma_s^2(T_k)$ [und eventuell $\sigma_s^2(I_k)$] durch laufendes Neukalibrieren und eventuell zuzüglich noch durch wiederholtes Neukalibrieren für dasselbe Meßobjekt auf einen verschwindend kleinen Pegel, was mittels dieses theoretischen Gedankengangs bis hin in beliebige Nähe von Null gehen würde. Allein schon tägliches Neukalibrieren ergibt mit dem obigen Datensatz $\sigma = 9 \text{ kg}$, also mit dem MUF-Test $1-\beta(\mu^*) = 95\%$ für $|\mu^*| = 29,5 \text{ kg}$ eine recht ansehnliche Kontrolle. Dieser Gedankengang bricht aber irgendwann ($\forall 0$) in der Realität zusammen, das Modell der Zersplitterung in σ_s/σ_r und deren fortgesetzte separate Reduzierung trifft ganz sicher irgendwann die Realität nicht mehr. [Man möchte damit den wahren Wert der Meßgröße durch wiederholtes Beobachten Y_i mit immer kleinerer Varianz einschließen, so daß nach dem Gesetz der

großen Zahlen $\frac{\sum_{i=1}^n Y_i}{n} \rightarrow E(Y)$ = wahrer unbekannter Wert geht, d.h. damit den Wert künstlich trotz nicht "präzisen" Instrumenten mit unbegrenzter Präzision erhalten. Tatsächlich geht aber $\frac{\sum_{i=1}^n Y_i}{n}$ mittels dieser Wiederholung zwar gegen einen Wert mit unbegrenzter Präzision (Varianz=0), dessen Abstand zum wahren Wert nun unbekannt ist, d.h. der wahre Wert ist dadurch nicht mit Präzision zu erhalten und der erhaltene Wert zusammen mit der Angabe "Präzision = genau" erweist sich als eine völlig wertlose, fiktive Größe]. Es ist zu vermuten, daß dies bei der Reprocessing Anlage bei durch diesen Verminderungsvorgang erhaltenen hohen σ -Werten bereits der Fall ist, sonst läge das Problem nicht dauernd zur Bearbeitung vor; es kann auch sein, daß der Meßaufwand zu groß würde.

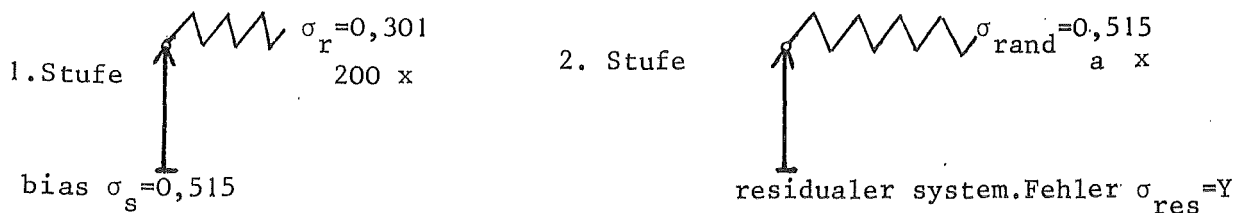
Jedenfalls wird jetzt dieser Fragenkomplex angegangen, irgendwann tritt er schließlich doch einmal auf, zumal zur Verminderung auf $\sigma = 2,44$ kg, also $1-\beta(\mu^*)=95\%$ für $|\mu^*|=8$ kg Pu ein 14-faches {Messen+Neukalibrieren} erforderlich ist (wegen $9/\sqrt{14}=2,44$) für jeden einzelnen Meßvorgang und jeden Tag à 5000 kg, wenn das Modell nicht schon zusammengebrochen ist! Man würde also jeden Tag 14x5000 kg (einmal) zu messen haben und für die Verifikation muß σ noch einmal grob $\sqrt{2}$ bis $\sqrt{3}$ -fach vermindert werden, also schon 42-fache(!) Messungen für den Operator $\hat{=}$ ungefähr 200 Tonnen/pro Tag. Und $\alpha=5\%$ ist auch noch zu hoch, $\alpha=1\%$ gibt $42 \times 1,5 = 63$ -faches Messen. Und schnell ist das System immer noch nicht, was nochmals eine Verminderung von σ erfordert. Er hätte also jeden Tag den äquivalenten Meßaufwand von etwa dem halben Jahresdurchsatz. Weiterhin wäre bei 60-100-fachem Neukalibrieren der Modellzusammenbruch schon längst erfolgt; es dürfte keine Meßapparatur geben, die eine Reduzierung des systematischen Fehlers auf den 10. Teil durch Kalibrierungsserien erlaubt. Die Anlage würde also fast ausschließlich damit beschäftigt sein, sich selbst zu messen und die Messungen wären nicht einmal verwertbar, wie gerade erklärt. Auf diesem Weg ist die Fehlerstruktur nicht in den Griff zu bekommen. Bei der angestellten Überlegung ist der beträchtliche Meßaufwand in den Vordergrund gestellt worden. Hierbei wurde zusätzlich klar, daß die Vermutung der nicht genügenden Reduzierbarkeit von σ_{Gesamt} effektiv bestätigt wird, was jetzt noch besser herausgearbeitet werden soll.

Im ersten Schritt wird jeden Tag kalibriert; hier tritt schon die Diskrepanz auf. Denn dies bedeutet die Verminderung des systematischen Fehlers auf den 14. Teil des ursprünglichen. Dies kann keineswegs erreicht werden, es bedeutet

nebenbei erwähnt für Transfergrößen dasselbe wie der Vorgang des 200-maligen Messens (mit Kalibrierung) desselben Objektes in der Absicht, damit den systematischen Fehler dieser Bestimmung auf den 14. Teil zu vermindern! Schon für den Fall, daß lediglich ein residualer system. Fehler von $\frac{\sigma_s}{\sqrt{200}}$ übrigbleibt, was durchaus plausibel ist - denn bei 200-fachem Messen hätte man diesen noch - ergäbe dies einen Beitrag von $\left[200 \times \frac{\sigma_s}{\sqrt{200}}\right]^2$ zusätzlich zur Gesamtbilanz $2\sigma^2(I) + 200[\sigma_r^2 + \sigma_s^2]$, also das ca. $\sqrt{2}$ -fache von $\sigma = 9$ kg.

Bei genauer Analyse hat man bei der Aufstellung der Gesamtbilanz eine Menge M_{200} - den (2-fachen) Jahresdurchsatz - sozusagen durch eine Messung (die aus vielen Teilmessungen besteht) bestimmt und dabei einen randomisierten Fehler $\sigma_R^2 = 2x\sigma^2(I) + 200x\sigma_r^2 = 9 + 18$ und einen systematischen Fehler von $\sigma_s^2 = 10\ 600$ erhalten.

Die 1. Stufe des Zufalls, die man bei dieser Einzelmessung modelliert hat, ist das Zufallselement bei einem gegebenen Meßinstrument, an dem nichts geändert wird, und mit dem wiederholt eine Menge M gemessen wird mit einer Zufallsstreuung von $0,301^2 = 0,09$ [das Lager wird jetzt zur Vereinfachung weggelassen] pro Messung an einem Tagesdurchsatz. Ein weiterer Fehler σ_s^2 ist nicht als variabel modelliert. Durch die Aufspaltung der Gesamtmessung in 200 Tagesmessungen ist dieser Zufall mit dem "Wert" 18 vollständig beseitigt als Komponente der Meßgenauigkeit $10\ 600 + 18$, sie taucht unter; obwohl in der Einzelmessung von bedeutendem Gewicht (fast gleiches Gewicht)



Folgerung a): es hat absolut keinen Wert, durch wiederholtes Messen pro Tag diesen Zufall der 1. Stufe (mit Verminderungsindex 200 x) noch mehr beseitigen zu wollen. Das Modell ist zu verfeinern, da man die Technik der Wiederholung zur Zufallsbeseitigung, um damit höhere Präzision zu erzielen, für den Vorgang "Messen" längst erschöpft hat. Für die Zielmessung M_{200} , aus der ja eine Entnahme einer gewissen Menge $||\mu^*||$ (z.B. 8 kg) herausgefiltert werden soll,

ist es im Kontext "protracted loss" - wie wir gesehen haben - völlig unerheblich, Einzelmessungen noch genauer zu bestimmen, etwa hier durch wiederholtes Messen, was durchaus möglich wäre: es zählt lediglich der Varianzbeitrag zur Gesamtmessung M_{200} und diese ist bezüglich dem Vorgang "Messen" bereits präzise bestimmt, d.h. man ist bis zum wahren Wert, der durch "Messen" bestimmt wird, vorgezogen.

Die 2. Stufe des Zufalls, um die Wiederholungstechnik anzuwenden, besteht in der Modellierung einer weiteren Komponente des Meßvorgangs: Kalibrierung des Meßinstrumentes. Beseitigung dieses Zufallselementes heißt Wiederholung des Kalibrierungsvorganges, der eine Zufallsstreuung von (höchstens) $0,515^2$ besitzt. |Ob diese tiefer angesetzt werden kann, muß aus einer Analyse des Vorganges erfolgen.| Ein wichtiger Unterschied zur 1. Stufe besteht bezüglich des Aufwandes, da die 1. Stufe automatisch durch den sowieso durchgeführten Meßvorgang auf den wahren Wert reduziert wurde. Hier muß dagegen a gewählt werden. Wiederum ist ein nicht variabler Fehler σ_{res}^2 vorhanden, von dessen Größe es abhängt, wie a bei gegebenem Gesamtaufwand für das Meßsystem zu wählen ist. Für $\sigma_{res} \ll \sigma_{rand}$ wird a groß sein, um dadurch σ_{rand} durch Ausspielen des Zufalls untertauchen zu lassen als Komponente der Einzelmessung M_{200} , weil damit der gegenüber σ_{rand}^2 (Gesamt) = 10 600 auf σ_{res}^2 (Gesamt) verminderte Anteil der Gesamtvarianz für die Messung M_{200} übrig bleibt. Hat dagegen σ_{res} fast gleiches Gewicht wie σ_{rand} , so hat es absolut keinen Wert, diesen Vorgang durch Wiederholung auf den wahren Wert zu reduzieren.

Folgerung b): Jede weitere Stufe kann nur teilweise reduziert werden, kein Meßsystem ist frei von einem schließlichen residualen Fehler, der dann als systematischer Fehler in dem Modell wiedergespiegelt werden muß.

Die Reduzierung hat höchstens bis zur Beseitigung der Zufallsschwankung durch a Wiederholungen einen Sinn, darüber hinaus überhaupt keinen Wert. Der Aufwand begrenzt zusätzlich a . Im Falle, daß der residuale Fehler fast gleich groß wie derjenige des bisher betrachteten Gesamtsystems ohne Wiederholung des spezifischen Vorganges ist, ist die Wiederholung ohne Wert. Vorgänge für weitere Stufen: Standardgrößen zur Eichung vervielfachen, mehrere Instrumente gleichen Typs (z.B. mehrere Analyselaboratorien), mehr Proben.

Gemäß den Daten der Referenzanlage ist die Größe $\sigma_s^2 = 10\ 600$ fast allein durch den system. Fehler der Inputmessung mit Varianz 10 080 bestimmt, der somit eine Verbesserung strikt blockiert. Es ist von zentraler Bedeutung zu wissen, ob $\sigma_\delta = 1\%$ schon einen residualen system. Fehler darstellt, ob also nach sorgfältiger Kalibrierung mit genauen Standardgrößen usw. dieses σ_δ evaluiert wurde. Dann ist der Zufall schon eliminiert und die Einführung weiterer Stufen nach b) sinnlos, mit $\sigma_{\text{Gesamt}} = 103\text{ kg}$ die Sensibilität des Meßsystems erschöpft. Insbesondere falls schon die besten existierenden Meßinstrumente diesen residualen Fehler 1 % aufweisen, kann die Schranke nicht mehr weiter zurückgeschoben werden.

Folgerung: Im Zusammenhang der Entdeckung von μ^* - die eine Grenze des Kontrollsystems definiert, insofern es kein statistisches Verfahren gibt, das besser als $\phi\left(\frac{|\mu^*|}{\sigma_{\text{Gesamt}}} - U_{1-\alpha}\right)$ entdeckt - wird eine zweite prinzipielle Grenze durch nicht weiter elimierbare systematische Fehler gegeben. Es wird deswegen als wesentliche, primäre Fragestellung für die Referenzanlage festgehalten

Problem I: Analyse des Meßsystems mit dem Ziel, ob der systematische Fehler noch weiter verkleinert werden kann, in erster Linie σ_δ des Inputs, der sonst alles blockiert.

Es bestehen ausschließlich folgende Auswege, um $\sigma = 103\text{ kg}$ zu vermindern

- i) Reduzierung auf einen residualen Fehler durch die Wiederholungstechnik, wie soeben beschrieben, unter Beibehaltung des bisherigen Meßsystems
- ii) Entwicklung von präziseren Instrumenten
- iii) Aufteilung der Mengen eines einzelnen Meßvorgangs, falls dann präzisere Instrumente genommen werden können für die kleineren Mengen (bzw. Aggregation)
- iv) Eliminierung system. Fehlerkomponenten, falls in Input und Output für Teilmengen dieselben Instrumente verwendet werden können.

Denn der Meßvorgang stellt sich dar $M_{200}(\sigma_R, \sigma_S) = \sum M_i(\sigma_r, \sigma_s)$, also verändert werden kann nur der Index (iii), σ_r mit i) und ii), σ_s mit i), ii) und iv). Bei allen Möglichkeiten bestehen prinzipielle Grenzen, so daß man zu einem schließlichen σ_R, σ_S gelangt. Insbesondere falls die

Konzentrationsbestimmung im Input einen residualen Fehler von 1 % aufweist (mit dem präzisest möglichen Instrument), was wegen der 1 % Pu-Konzentration eine Genauigkeit für die Probenkonzentration von $\sigma_{\text{res}} = 0,01$ % Pu bedeutet, erweist sich σ_{S} resistent gegen jegliche Mengenaufteilung, Wiederholung, d.h. jede Möglichkeit der Verbesserung. Durch den einmaligen präzisen Meßvorgang (mit σ_{res}) der Konzentrationsbestimmung wird notwendigerweise ein $\sigma_{\text{S}} = 1$ % eingeführt, der eine prinzipielle Schranke bildet, in diesem Fall $M_{200} \cdot 0,0001 = 100$ kg Pu; ganz gleichgültig mit welchen Massen, mit wieviel Proben aus sich unterscheidenden Konzentrationen, mit wieviel Analysevorgängen man es zu tun hat, man bleibt auf dem Wert 0,01 % Genauigkeit für die Konzentrationsbestimmung sitzen, der sich ausweglos auf den Gesamtdurchsatz M_{200} weitervererbt. Wegen dem verschiedenen Material in Input und Output muß man notgedrungen eine absolute Mengenbestimmung (keine Differenz) mit Einschluß des residualen system. Fehlers durchführen und weil dies lediglich via Konzentrationsbestimmung geht, blockiert diese Genauigkeit in linearer Weise die Genauigkeit für die Gesamtmenge.

Schlußfolgerung: Läßt sich mit jeder Anstrengung bei einer einzelnen Konzentrationsbestimmung/Input nur eine Genauigkeit von $\sigma_{\text{res}} = 0,01$ % Pu erreichen, so bedeutet dies für σ_{Gesamt} eine prinzipielle Grenze von 100 kg Pu bei der Referenzanlage, die ausschließlich durch die Entwicklung eines Instrumentes mit kleinerem σ_{res} verbessert werden könnte. Diese prinzipielle Grenze wiederum bedeutet im Kontext protracted loss, daß eine Entnahme von 40 kg/Jahr prinzipiell durch kein statistisches Verfahren unter Anwendung des Materialbilanzprinzips entdeckt werden kann: Bei $\alpha = 1$ % bekäme man in 3 % der Fälle zufällig einen Alarm, der selbst dann noch nicht unbedingt eine Entdeckung (second action level) bewirkt. Bei fortgesetzter Entnahme über Jahre zu je 40 kg kann an diesem Sachverhalt wiederum prinzipiell nichts verbessert werden (σ_{Gesamt} wächst linear wegen σ_{res}), es sei denn, die Entnahme wird "abrupt" beendet und σ_{res} bleibt für alle MUF-Bilanzen verbindlich fest gewählt. Zusätzliche Schwierigkeiten, wie nicht meßbare Verluste infolge des reprocessing, nicht totales cleaning, nicht repräsentative Proben sind noch nicht berücksichtigt (d.h. irgendwelche kurzfristige systematische Fehler bei der Konzentrationsbestimmung).

Man beachte, daß zwar die Blockierung in linearer Weise bei der Konzentrationsmessung für sich allein ein Vorteil wäre, da man nur "einmal" eine repräsentative Probe aus M_{200} mit größtmöglicher Präzision σ_{res} messen muß (oder äquivalent viele Proben mit laufender Kalibrierung, womit auch noch die geschichtete Konzentrationsdichte gleichzeitig erledigt ist), aber es heißt datenmäßig dasselbe wie aus 1 Million kg eine Konzentrationsänderung von z.B. 8, 10, 40 kg erkennen, d.h. eine $\sigma = 0,0002\%$ genaue Konzentrationsbestimmung vorzunehmen, um die man in keiner Weise herumkommt (für $\alpha=5\%$, $1-\beta=95\%$, $M=8kg$).

Es gibt lediglich zwei Auswege

Problem II: Der Input muß in einer größeren MBA so eingebettet werden, daß die neuen Bilanzvariablen mit nun in transformierter Form vorliegendem Material ein wesentlich besseres σ_{res} aufweisen. Hier gibt es viele Wege, die von dem detaillierten Prozeßgeschehen abhängen, um die Kontrolle zu verbessern; nur hat man beim Output im Bereich $\sigma_G = 20$ kg dieselbe Blockade, über deren Beseitigung man sich die analogen Gedanken machen muß, und schließlich kommt man zu dem primären Input und endgültigen Output, wo sowieso nichts mehr zu ändern ist. Der bisherige Input würde damit ins Lager aggregiert und müßte in der Periodeneinheit immer in möglichst geringer Menge vorliegen.

Problem III: Das strenge Materialbilanzprinzip muß fallengelassen werden, um dem residualen systematischen Fehler doch noch beizukommen.

Abschließend wird darauf hingewiesen, daß die Frage der Reduzierung (Problem I, II) nicht umgangen werden kann. Skizze: Jedes mit einem zu hoch geschätzten σ_s gebildete Modell läßt für μ^* höchstens die Entdeckung $\Phi\left(\frac{|\mu^*|}{\sigma_{Gesamt}} - U_{1-\alpha}\right)$

mit dem NP-Test zu, wobei in σ_{Gesamt} das hohe σ_s eingeht,
 $\sigma_{Gesamt}^2 = \sum_i M_i \sigma_r^2 + \sum_{i,j} M_i M_j \cdot \sigma_s^2$. Denn für einen mit dem Wert σ_s versehenen system. Fehler ist dies die Entdeckungswahrscheinlichkeit des bestmöglichen Tests. Würde nun irgendein sequentieller Test konstruiert, der ja nach Voraussetzung erst einmal für σ_s eine Fehlalarmrate $\leq \alpha$ haben muß - ich muß ja so tun, als wäre $\sim N(0, \sigma_s^2)$ für den syst. Fehler möglich - so würde dieser

Test (auch falls ein Filter eingebaut ist) die Gestalt $Y_i = \sum_{k=1}^i \alpha_k X_k$ mit Alarmmengen $Y_i > z_i$ bzw. $|Y_i| > z_i$ haben. Nun würde also ein σ , es sei gleich σ_{res} genannt, mit $\sigma_{res} < \sigma_s$ merklich kleiner die tatsächliche Streuung charakterisieren. Nach den Beweisen in 2. ergibt dieser sequentielle Test für $N(0, \sigma_s^2)$ jedenfalls eine kleinere Entdeckung für μ^* als $\Phi\left(\frac{|\mu^*|}{\sigma_{Gesamt}} - U_{1-\alpha}\right)$.

Kann er - etwa durch Schätzung von σ_{res} - unter $N(0, \sigma_{res}^2)$ eine höhere Entdeckung automatisch produzieren? Es interessiert gerade der Fall, daß etwa $|\mu^*| = 40$ kg oder eine andere Menge mit Prob $\leq 50\%$ entdeckt wird. Damit liegt immer $E\left(\sum_{k=1}^i \beta_k X_k\right)$ unterhalb z_i (sonst hätte man für σ_s einen besseren Test) und weiterhin wird die Varianz $\left(\sum_{k=1}^i \beta_k X_k\right)$ durch Übergang von σ_{res} zu σ_s auch für den Filter erhöht, d.h. für jedes i würde $\sum \beta_k X_k$ mit höherer Wahrscheinlichkeit über z_i springen, also insgesamt unter σ_s für den Gesamttest über alle Perioden auch mit höherer Wahrscheinlichkeit als für σ_{res} , das selbst höher als Φ angenommen wurde. Dies kann aber wegen Neyman-Pearson nicht sein.

Anders gesagt, kein Verfahren hat bezüglich μ^* bei der Bestimmung von σ_s etwas zu verschenken; es gehen entdeckbare Entnahmen verloren. Der andere Extremfall bei der Modellbildung, Unterschätzung von σ_s - also $\sigma_s < \sigma_{res}$ wird dem Modelldesign zugrunde gelegt - wäre dagegen sehr gefährlich, denn es würden Entnahmen mit sehr hoher Signifikanz angezeigt werden (ungerechtfertigt); also gerade durch den Einbau in das Modell von häufigem Kalibrieren würde diese unerwünschte Erscheinung schließlich zwangsläufig auftreten. Wobei natürlich häufiges Kalibrieren selbst nur verbessert (Fehlalarme werden vermindert), nicht aber der unsorgfältige Einbau in das Modell.

Es scheint weiterhin bei der Anlage eine Inkonsistenz vorhanden zu sein: denn wenn der Output mit 0,2 % Genauigkeit gemessen werden kann, müßte der Input indirekt mit derselben Genauigkeit meßbar sein, indem die Instrumente für die Konzentrationsmessung des Inputs durch einen Referenzoutput (nicht unbedingt aus der Anlage) geeicht werden, es sei denn, es würden unmeßbare Verluste σ_L beim Transformationsverfahren auftreten: $\sigma_{In} = \sigma_{Out} + \sigma_L$.
Damit würden aber diese auch in der Referenzanlage auftreten $\sigma_L = 0,01 - 0,002 = 0,008$, also einen 0,8 % prinzipiell unbestimmbaren Fehler einführen, der die Kontrolle wiederum erschwert und ein höheres $\sigma_s^2 = \sigma_{In}^2 + \sigma_L^2 + \sigma_{Out}^2$ ergäbe. Oder auch σ_{In}

ist gar nicht streng systematisch. Die Eichung gehört unter iv) eingeordnet; es müßte jedenfalls noch mehr Genauigkeit aus der Anlage zu erhalten sein. Es kann lediglich als Hindernis auftreten a) hohe Kosten: dann muß man sich mit protracted Entwendungen von 40 kg abfinden, b) ungenaue Modellbildung: σ_L müßte eingeführt werden oder σ_s ist nicht konstant. Diese Fragen müssen anhand des detaillierten Prozesses systematisch überlegt werden (Problem I, II).

Nachdem die sowohl notwendige (wie oben skizziert) als auch aussichtsreich erscheinende Reduzierung von $\sigma=103$ kg auf einen gleich σ_{res} genannten Wert erfolgt ist, bleibt nunmehr der Fall übrig: σ_{res} ist weiterhin dominierend (sonst vermindere man σ_r durch Wiederholung) und ergibt keinen zufriedenstellenden Wert σ_{Gesamt} . Die Reduzierung wäre nur dann überflüssig, falls die damit verbundenen Kosten ein später zu definierendes Gleichgewicht bereits erreicht haben.

5. Die Sättigungsschranke für die Entdeckungscharakteristik $1-\beta$ bei "protracted loss"

Nachdem in 4. alle Möglichkeiten für die Eliminierung des system. Fehlers, der für die Nichtentdeckbarkeit von μ^* in signifikanter Größe verantwortlich ist, analysiert wurden, und gezeigt wurde, daß man prinzipiell schließlich auf ein σ_{res} stößt, so daß nur noch eine Menge $||\mu^*|| \geq 3,28 M \sigma_{res}$ mit der üblichen Testcharakteristik entdeckt werden kann, und $M\sigma_{res}$ bei der Referenzanlage bis auf weiteres im Bereich 100 kg Pu liegen könnte (sonst sollte man umgehend den Wert 1 % ändern), bleibt nichts anderes übrig, als zu versuchen, den unbekanntem Wert σ_{res} direkt zu schätzen.

Zuvor nochmal die genaue Quantifizierung, damit man sieht, an welcher Stelle man sich beim Restproblem befindet (es wird vereinfachend ein Material - hier Input - betrachtet, das Lager verändert die Betrachtung unwesentlich)

$M(t) = t \cdot m = \sum_{i \in \mathcal{Z}_0} M_i$ M wächst linear mit der Zeit und wird für die Messung in Grundteilbeträge M_i - vereinfachend gleiche M_i , d.h. befinden sich alle in demselben Größenordnungsbereich - aufgespalten. Für gleiche M_i ist es gleichgültig, ob das additive oder multiplikative Fehlermodell zugrunde liegt. Diese Aufspaltung definiert die Grundverfeinerung \mathcal{Z}_0 .

$$\sigma_{\text{Gesamt}}^2 = \sigma^2(\mathcal{Z}_0, t) = \sum_{i \in \mathcal{Z}_0(t)} M_i^2 \sigma_r^2 + \sum_{i,j \in \mathcal{Z}_0(t)} M_i M_j (\sigma_{\text{res}}^2 + \sigma_{\Delta}^2)$$

Für den random. Teil $R(\mathcal{Z}_0, t_0) = \sum M_i^2 \sigma_r^2$ kann ab \mathcal{Z}_0 weiter verfeinert werden durch mehrfaches Messen $n\mathcal{Z}_0$. Dadurch schrumpft dieser auf $R(n\mathcal{Z}_0, t_0) = \frac{1}{n} R(\mathcal{Z}_0, t_0)$ durch n-fachen Meßaufwand, also σ_R auf den $\frac{1}{\sqrt{n}}$ ten Teil.

Für den system. Teil $S(\mathcal{Z}, t_0)$ kann ab $||\mathcal{Z}||=1$ mit $S(\mathcal{Z}, t_0) = M^2(t_0)(\sigma_{\text{res}}^2 + \sigma_{\Delta}^2)$ verfeinert werden durch mehrfaches Kalibrieren z.B. Dadurch schrumpft dieser auf $S(n\mathcal{Z}, t_0) = \text{RES}(t_0) + \text{KAL}(n\mathcal{Z}, t_0) = M^2(t_0)\sigma_{\text{res}}^2 + \frac{1}{n} \text{KAL}(\mathcal{Z}, t_0)$ durch n-faches gleichförmiges Verteilen der Kalibrierung, also RES bleibt konstant, σ_{Δ} geht auf den $\frac{1}{\sqrt{n}}$ ten Teil. Die Entnahme $\mu^*(t)$ erfolgt hier linear $\mu^*(t) = t \cdot \mu^*$. Die Betrachtung beginnt bei einem festen Zeitintervall $[0, t]$. Die beste Entdeckung charakterisiert durch $Q(\mathcal{Z}_0, t) = \frac{t\mu^*}{\sigma(\mathcal{Z}_0, t)}$ ist aus irgendeinem Grund nicht zufriedenstellend, sei es daß die Zielmenge eingesetzt schon einen schlechten Wert ergibt oder daß man μ^* noch weiter verkleinern möchte [μ^* ist die Entnahmedichte].

A) σ_R dominierend.

$||n\mathcal{Z}_0|| \downarrow$ verfeinern, so $R(n\mathcal{Z}_0) \ll S(\mathcal{Z})$

System wird besser

Dann

$||\mathcal{Z}||_{\Delta} \downarrow$ verfeinern, so $R + \text{KAL} \ll \text{RES}$

System wird besser

Q wächst mit \sqrt{n}

Nun ist $M^2(t_0)\sigma_{\text{res}}^2$ dominierend

$||\mathcal{Z}|| \uparrow$ gibt höheres $\sigma(\mathcal{Z}, t)$, weniger Aufwand, falls Entdeckung zufriedenstellend

$t \uparrow$ gibt wachsendes $\sigma^2(\mathcal{Z}, t)$, lineares Wachstum für $R(\mathcal{Z}, t)$ also mit \sqrt{t} für $Q(\mathcal{Z}, t)$, hierhin gehört genau CUMUF(t), der mit \sqrt{t} besser wird. Schließlich für

$t \uparrow \uparrow$ wird RES dominierend.

$t \downarrow$ $Q(t, \mathcal{Z})$ nimmt mit \sqrt{t} ab, also zeitliche Unterteilung immer schlechter, und σ_R wird noch dominierender.

Man gelangt notwendig für $t \uparrow$ oder auch dadurch, daß man σ verkleinern wollte (für t fest, um die Entdeckung zu erhöhen) zu

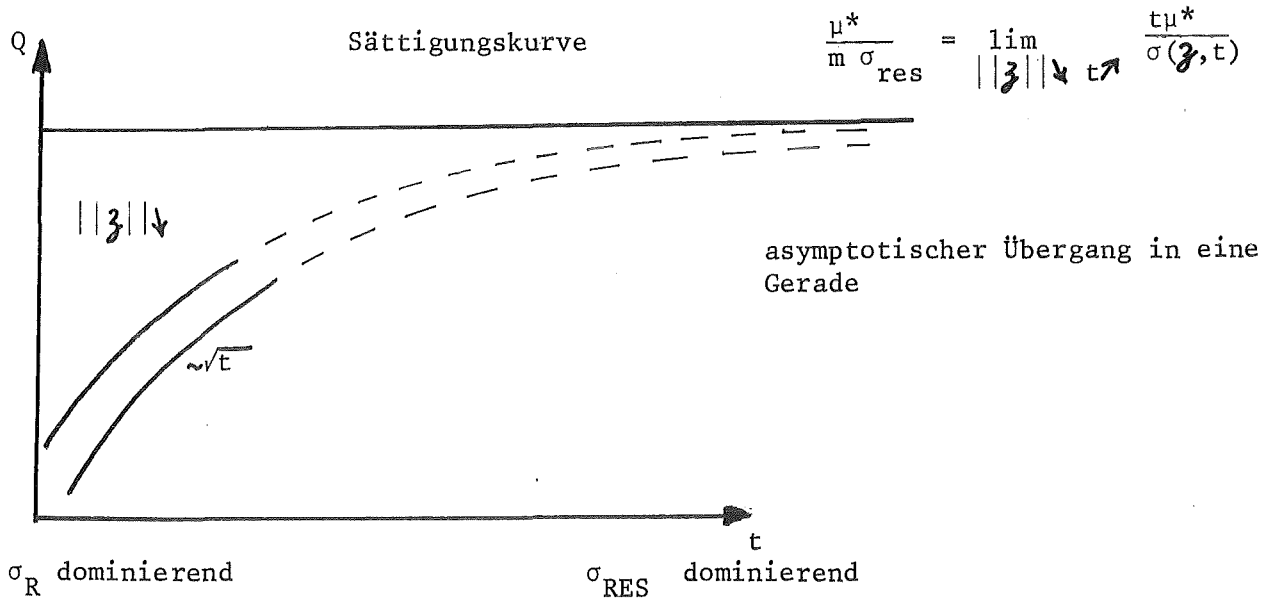
B) RES(t) dominierend.

Für $||\mathcal{Z}|| \uparrow$, $||\mathcal{Z}|| \downarrow$ keine Verbesserung von $Q(\mathcal{Z}, t)$ möglich (abgesehen von einer vernachlässigbaren für $||\mathcal{Z}|| \downarrow$)

Für $t \uparrow$ bleibt $Q(\mathcal{Z}, t)$ konstant, d.h. der Entnahmebetrag nimmt linear zu.

Für $t \downarrow$ wird $Q(\mathcal{Z}, t)$ schlechter, sobald $R + \text{KAL}$ ins Gewicht fallen.

Man hat damit immer einen Bereich, wo $Q(z, t) \sim c \cdot \sqrt{t} \cdot \sqrt{n}$ sich verbessern läßt durch zeitliches Datensammeln und einen schließlichen Bereich, wo $Q(z, t)$ zu einer Konstanten wird (Sättigungskurve)



Das Resultat dieser Analyse wird festgehalten als wichtiges

Ergebnis: Jede konstante Entnahme pro Zeiteinheit wird bestenfalls mit

$\Phi\left(\frac{\mu^*}{m \sigma_{res}} - U_{1-\alpha}\right)$ durch Kumulieren der MUF-Variablen entdeckt.

Im Grunde besagt dies, daß das Modell berücksichtigt, daß kein Meßsystem schließlich jeden noch so kleinen Betrag μ^* diskriminieren kann, also Messungen mit totaler Präzision durchführen kann (gemäß dem Gedankengang in 4.). Wesentlich sind nun die tatsächlichen Daten, die vorerst ein $\sigma_{RES} = 100$ kg zulassen für die Zeiteinheit $t = 1$ Jahr, und die daraus folgende Höhe der Sättigungskurve mit μ^* als Parameter. Während bei konventionellen Bilanzen bisher wegen $\frac{G/t}{M/t \cdot \sigma_{res}} \geq \frac{G}{\sigma_{Gesamt}} \geq 3.28$ die Sättigungsgerade für G sehr hoch lag, damit erst nach vielen Perioden für $G =$ Zielmenge in Erscheinung tritt, ergibt sie jetzt für die Entdeckung innerhalb eines Jahreszeitraums wegen dem hohen M/t immer die Grenze und ist also für die interessierenden Werte nicht genügend hoch (ergibt beispielsweise für lineare Entnahme von 40 kg/Jahr keinen Alarm, gerade auch in Zukunft nicht). Anders gesagt, das Zusammenkoppeln konventioneller MUF-Systeme hat genau im Bereich, wo $\sigma_R(t)$ noch dominiert, einen Sinn, danach muß von vorne begonnen werden (wegen der Sättigung), es sei

denn man kann den residualen Fehler schätzen.

Wie kann dieser überhaupt geschätzt werden? Der einmal realisierte system. Fehler s (Einstellung des Meßsystems) verhält sich wie eine lineare posit./negative Entnahme und kann deswegen in Präsenz einer linearen Entnahme des Entwenders nicht besser diskriminiert werden als die Entnahme μ^*+s im Laufe der Zeit gegenüber dem Rauschen; diese wird durch den CUMUF(n)-Test $\Phi\left(\frac{n(\mu^*+s)}{\sigma_G} - U_{1-\alpha}\right)$ mit dem beschriebenen Sättigungsverhalten $\Phi\left(\frac{\mu^*+s}{\sigma} - U_{1-\alpha}\right)$ am besten entdeckt, also wegen der Symmetrie im Mittel mit $\Phi\left(\frac{\mu^*}{\sigma} - U_{1-\alpha}\right)$ maximal. Der Entwender kann nur durch totale Kontrolle in einem Zeitabschnitt gezwungen werden, nicht-linear zu entnehmen, was dann gleichbedeutend mit gar keiner Entnahme ist. Dann kann aber dieser Zeitabschnitt gleichzeitig dazu benutzt werden, den konstanten syst. Fehler gegenüber dem Rauschen σ_R zu evaluieren, da er eine lineare Abzweigung gegenüber dem nun sichergestellten Nullzustand bewirkt. Mit anderen Worten, da keine Masse verloren geht, wird hierdurch das Meßsystem durch sich selbst geeicht und dies geht bei genügend langem Abschnitt ohne Entnahme mit unbegrenzt zunehmender Präzision [wenn sowohl s konstant gehalten werden soll, als auch keine Änderung der nicht meßbaren auftreten]. Sobald aber mit der Kontrolle allein mit dem Materialbilanzprinzip begonnen wird, hat man aus der Schätzung einen (stark verminderten) neuen Fehler σ_{res} im System, der prinzipiell nach längerem Zeitablauf wieder als Sättigungslinie erscheint. Für die Praxis hängt alles von den Daten ab.

Es bleibt als Frage bestehen, ob eine zeitweise totale Kontrolle akzeptabel sowie vom technischen Standpunkt aus durchführbar ist.

Problem IIIa: Ist eine totale physische Kontrolle möglich?

Zusammenfassung: Liegt eine konventionelle Materialbilanz vor, so ist diese für den Zielbetrag entweder zufriedenstellend $(1-\beta)$. Wird wie bisher jeweils konsekutiv ein MUF-Test nach dem anderen für jede Inventarperiode durchgeführt, dann ist die zusätzliche Zusammenkopplung mehrerer Bilanzen im Bereich, wo σ_R dominiert, auf jeden Fall vorzuziehen; im Bereich, wo $\sigma_{SYST.}$ dominierend wird, abzulehnen. [Denn für nicht-lineare Entnahmen hat man in einer der ursprünglichen MUF-Bilanzen $\Phi\left(\frac{M_1}{\sigma} - U_{1-\alpha}\right)$ und $\Phi\left(\frac{M_2}{\sigma} - U_{1-\alpha}\right)$ eine höhere Entdeckung

als mit $\Phi\left(\frac{M_1+M_2}{2\sigma} - U_{1-\alpha}\right)$ als zusätzlichem Test. Dagegen für die lineare Entnahme keine Erhöhung gegenüber den Grundbilanzen, aber immer eine Zunahme der Fehlalarme; eine Schätzung ist hier vorläufig weggelassen, da diese den protracted loss nur bei nicht-konventioneller Kontrolle besser entdecken würde.]

Oder die Entdeckungswahrscheinlichkeit ist nicht genügend. Dann bestehen nur die beiden Möglichkeiten, σ zu reduzieren (wie in 4. beschrieben) oder nach Aufteilung in Teilbilanzen σ_{RES} zu schätzen mit einer zeitweiligen totalen physischen Kontrolle bzw. beide Methoden zu kombinieren. Wird die 2. Möglichkeit ausgeschlossen und stößt die (mit dem Aufwand als Parameter versehene) Reduzierung an die Grenze σ_{RES} ohne zu genügender Entdeckung zu gelangen, so nützt weder Aufteilung noch Zusammenkoppeln von Bilanzen etwas. Die Zulassung der totalen physischen Kontrolle ergibt ein zu optimierendes Aufwandverteilungsproblem (Aufwand für phys. Kontrolle x Anzahl der Tage gegen Aufwand für σ -Reduzierung), wobei im ersten Schritt die maximale Entdeckung des protracted loss als Zielfunktion gewählt wird.

6. Power-One-Test und Schätzung von σ_{res}

Die für das Safeguards-Problem relevanten Eigenschaften werden diskutiert.

a) Für korrelierte Größen (insbesondere für durch einen systematischen Fehler verknüpfte Werte)

Eine wesentliche Aussage des Power-One-Tests, die Entdeckung einer linearen Entnahme Θ in den Periodenabschnitten mit Power-one für $n \rightarrow \infty$, ist für den Safeguards-Fall nicht erfüllt. Hierzu müßte $E(MUFR_1)$ als Transformationsgröße auch linear anwachsen, was von vornherein nicht zu gelten braucht. Auf direkte Weise wurde in 5. gezeigt, daß für die lineare Entnahme (bis auf Korrektur in der 1. Periode) die Sättigungslinie $\frac{\Theta}{\sigma_{RES}}$ eine strikte obere Schranke $\Phi\left(\frac{\Theta}{\sigma_{RES}} - U_{1-\alpha}\right)$ für den Test bildet, d.h. der Power-One-Test entdeckt lineare Entnahmen nicht mit Power-one, müßte für den Safeguardsbereich besser umbenannt werden. In dem in [2] behandelten Referenzfall entdeckt der Test maximal mit $\Phi\left(\frac{\Theta}{0,515} - U_{1-\alpha}\right)$ für $n \rightarrow \infty$, wobei Θ die lineare Entnahmemenge/Tag bezeichnet. Für das Beispiel $M=40$ kg/pro Jahr entdeckt er also maximal mit $\Phi(0,388-1,64) = 10,5$ % diese fortgesetzte Entnahme für $\alpha = 5\%$ bzw. $\Phi = 2,6$ % für $\alpha = 1\%$. Maximal heißt über den unendlichen Resthorizont, hört man vorher mit dem Test auf, so ist die Entdeckung noch geringer, jedoch die Gesamtentnahme bis dahin $n \times 40$ kg mit $n =$ Anzahl der Jahre. Er entdeckt für noch so große lineare(!) Entnahmen niemals (zeitlich)

mit Power-one, unabhängig von der Wahl der Schranken z_n . Der Hintergrund hierfür wurde ausführlich in 1. bis 5. dargestellt, es ist gerade die im Safeguards-Problem liegende prinzipielle Schranke, die bei der Überwachung einer Reprocessing-Anlage nun schon stark ins Gewicht fällt; bei den bisherigen Anlagen würde sie wegen Präzision der Mengenmessung/ Zeiteinheit in Relation zu der Zielmenge erst in größerer zeitlicher Entfernung auftreten. Man kann es auch so erklären, daß für Power-one σ_{sys} randomisiert werden müßte; dies geht nur bis hinab zu σ_{RES} , dann schleppt der sequentielle Test immer das nicht mehr variable, konstant vorhandene σ_{RES} mit - wobei in [2] vorerst $\sigma_{\text{sys}} = \sigma_{\text{RES}}$ als nicht weiter reduzierbar gesetzt wurde - und diskriminiert die Entnahme Θ zu der konstanten Meßabweichung σ_{RES} mit dem Einperiodentest $\Phi\left(\frac{\Theta}{\sigma_{\text{RES}}} - U_{1-\alpha}\right)$ und sämtliche späteren Perioden lassen keine bessere Diskriminierung durch Neuinformation zu wegen dem strikt gleichförmigen Anwachsen $\frac{n \cdot \Theta}{n \cdot \sigma_{\text{RES}}}$.

Hätte man nur randomisierte Fehler, so würde der CUMUF(n) Test eine lineare Entnahme selbstverständlich mit $\Phi=1$ für $n \rightarrow \infty$ entdecken, sogar die langsamere Entnahme proportional $\sqrt{n \log n}$. Der Power-one-Test sagt dann lediglich aus, daß man sequentiell mit der dadurch zusätzlich entstehenden Fehlalarmbelastung auch noch $\Phi=1$ bewahren kann, indem die Schranken geschickt gewählt werden. Es gilt wie in 2. gezeigt immer für die Entdeckung der linearen Entnahme Robbins $(n) \uparrow \leq \text{CUMUF}(n) \uparrow \text{CUMUF}(\infty) = 1$. Für systematische Fehler, diese Beziehung gilt ja immer - denn sequentielle Verfahren haben immer $1-\beta(n) \uparrow \leq \text{CUMUF}(n)$ und $\text{CUMUF}(n)$ wächst wegen dem abnehmenden Einfluß von σ_R , siehe 5. - ist aber $\text{CUMUF}(\infty)$ schon durch $\Phi\left(\frac{\Theta}{\sigma_{\text{RES}}} - U_{1-\alpha}\right)$ begrenzt, erst recht also ein zu konstruierender "Power-one-Test" für die (abhängigen) Summen $\sum_{i=1}^n \text{MUF}_i$. Anders gesagt heißt dies, daß die Robbins-Schranken für die MUF_i keine Kontrolle der Fehlalarme ergeben würden: $\alpha \uparrow 1$. [Angenommen $\alpha < 1$, so würde $\text{CUMUF}(n)$ zu diesem α gegen 1 gehen].

b) Allgemeingültige Aussagen für die zusätzliche Verwendung eines Schätzers

Während also CUMUFR und die Robbins-Schranken nichts an den Schwierigkeiten zur Entdeckung des "protracted loss" ändern, ist CUMUFR ein Verfahren, um den systematischen Fehler zu schätzen. Dies geht immer nur dann, wie in 5. erklärt, wenn die Entnahme M_1 nichtlinear erfolgt, also z.B. auch, wenn zwei Periodenabschnitte vorliegen, in denen jeweils linear mit verschiedener Pegelhöhe entnommen wird, wobei einer davon $M_1 = 0$ /keine Entnahme aufweist.

Für solche Schätzverfahren gelten einige allgemeine Aussagen, falls sie überhaupt dem Problem angepaßt konstruiert sind:

i) Für $\frac{\sigma_R^2(T)}{\sigma_S^2(T)} \gg 1$ ist die Verbesserung eines Testverfahrens mit Schätzer/ohne

Schätzer gering, tritt erst auf der Zeitachse in Erscheinung, wo das Gesamttrauschen sowieso schon zu groß ist, um die interessierenden Mengen zu entdecken. Dies gilt erst recht, wenn in der konventionellen Periode

$\frac{\sigma_R^2(T)}{\sigma_S^2(T)} \gg 1$. Denn der Schätzer reduziert $\sigma_S^2(T)$ geringfügig

auf einen residualen Fehler $\sigma_{RES}^2(T)$ wegen der geringen Korrelation. Dann benötigt man viele Grundbilanzperioden, damit sich das Verhältnis $[\sigma_R^2 + (\sigma_S^2 - \sigma_{RES}^2)] + \sigma_{RES}^2$ zum ursprünglichen $[\sigma_R^2] + \sigma_S^2$ - wobei immer in den randomisierten und system. Teil aufgespalten wurde - in einer geringeren Gesamtvarianz mit Schätzung niederschlägt; oder entsprechend für die Aufteilung der Grundperiode.

ii) Für den Übergang in den anderen Fall bis hin zu $\sigma_R^2(T) \ll \sigma_S^2(T)$ wird durch Schätzung des system. Fehlers σ_S (unter der Nullhypothese) ein stark verminderter residualer Fehler erreicht, so daß die Gesamtvarianz wesentlich verkleinert wird. Dadurch wird eine außerhalb des Schätzbereichs konzentrierte Entnahme, selbst wenn sie dort protracted entnommen wird, mit sehr hoher Wahrscheinlichkeit entdeckt im Vergleich zu ohne Schätzung. Hier ist zusätzlich noch entscheidend, wie das Verhältnis $\sigma^2(I)/\sigma_S^2(T)$ ist, je kleiner $\sigma^2(I)$ desto kleiner kann der Schätzbereich gewählt werden.

iii) Durch unbegrenztes Vergrößern des Schätzbereichs kann theoretisch der residuale Fehler beliebig klein gemacht werden, so daß die Sättigungslinie immer höher nach oben verschoben wird. Jedoch erreicht man nie Power-one, da dort wo die lineare Entnahme beginnt, ein residualer Fehler dann gleichzeitig auftritt. Das beliebige Verkleinern des residualen Fehlers infolge der Schätzung würde außerdem jedes reale Meßsystem überfordern, insofern der system. Fehler strikt konstant gehalten werden müßte. Das bedeutet gleichzeitig eine wesentliche Grenze jedes Schätzverfahrens, sobald ein Bereich mit verändertem system. Fehler beginnt - etwa durch Kalibrieren - muß dies sofort im Modell berücksichtigt werden, d.h. die

Schätzung von vorne beginnen. Damit ist also, sei es durch Kalibrieren oder durch natürliche Veränderung der Skalen, nur eine begrenzte Verbesserung des Testverfahrens durch Schätzung möglich. Letztere ist sowieso nur, wie schon in 1-5 bewiesen, außerhalb des Schätzbereichs möglich, also es darf nicht ein protracted loss über den Gesamtbereich eintreten, der in keiner Weise jemals besser entdeckt werden kann.

- iv) Es wird durch Schätzen nicht nur der system. Fehler reduziert, sondern auch die Varianz in den mit der Schätzung korrigierten Teilbereichen, so daß dort eine abrupte Entnahme signifikanter entdeckt werden kann als ohne Schätzung. Wesentlicher ist jedoch die Verbesserung für den protracted loss über Teilbereiche, die hauptsächlich durch Reduzierung des system. Fehlers zustande kommt.
- v) Tritt schon eine Entnahme im Schätzbereich auf, so wird die Verbesserung wesentlich reduziert bis zum Grenzfall der linearen Entnahme. Dies kann ausschließlich durch die Forderung der totalen physischen Kontrolle vermieden werden. Die Ungenauigkeit infolge der Nichtmeßbarkeit des Hold-up bei laufender Anlage bedeutet genau dasselbe wie ein randomisierter Fehler $\sigma(H)$, d.h. wird durch Erhöhung der Lagervarianz zu $\sigma^2(I) + \sigma^2(H)$ berücksichtigt. [Sonst gäbe dieser bei Schwankungen Fehlalarme]. Realistischer ist eine nichtrandomisierte Schwankung, die bei den Schätzern in der Regel durch Kompensation noch weniger ins Gewicht fällt, analog der system. Fehlerreduzierung. Genau dasselbe gilt für unmeßbare Verluste. Mit anderen Worten, der Schätzer hat die sehr günstige Eigenschaft, die Hold-up-Schwankungen und unmeßbare Verluste genauso wie die system. Fehler zu reduzieren bezügl. der Berücksichtigung im Modell. Würde aber zwischen Schätzbereich und Kontrollbereich eine sprunghafte Veränderung eintreten, die danach konstant bleibt, so wäre genau das Gegenteil der Fall, nämlich eine starke Erhöhung aller dieser system. Fehlereinflüsse. Dies gerade wurde oben schon bezügl. des system. Fehlers im engeren Sinne unterstrichen, daß keine Neukalibrierung vorgenommen werden darf, ansonsten muß anders modelliert werden; was im Ergebnis nichts anderes heißt, als daß die Einführung eines Schätzers nichts mehr groß verbessert für die betrachteten Abschnitte.

Es ist also ein wesentlicher Punkt, daß keine sprunghaften Änderungen der Gesamteinflüsse zwischen dem Schätzbereich und Kontrollbereich stattfinden und die nichtrandomisierten Fehlereinflüsse beträchtlich sind gegenüber dem random. Fehler. Genau dann bewirkt der Schätzer eine starke Verbesserung, ist aber im Gegenzug sehr empfindlich gegen Verletzung dieser Bedingungen, die deswegen eine sorgfältige konservative Modellierung erfordern. Nichts anderes ist von vornherein zu erwarten, denn der Schätzer bewirkt ja präziseres Messen als die Instrumente erlauben durch die Annahme der Konstanz dieser Einflüsse, womit gerade das Meßsystem durch den Prozeß selbst geeicht wird. Falls der Entwender keine Manipulation bei den unmeßbaren Verlusten sowie dem Hold-up vornehmen kann, was zutreffend erscheint, bleibt die Frage bestehen, warum das Meßsystem nicht direkt so gut geeicht werden kann als durch den wiederholten Durchsatz von Massen in indirekter Weise unter Ausnutzung der Massenerhaltung, denn nichts anderes bedeutet die Schätzung. Sind es tragbare Instrumente, können diese nicht einmalig gleich sorgfältig geeicht werden? Sind es fest im Prozeß installierte Instrumente, für die das gleiche gilt? Nun ist lediglich die Schwierigkeit der Eichung auf diejenige einer totalen Kontrolle der Probedurchläufe und perfektem Meßsystemverhalten - damit indirekt durch die Massenerhaltung auf Null geeicht werden kann - verschoben. Andernfalls würde ohne totale Kontrolle die Schätzung zwar in gleicher Weise funktionieren aber unter der Annahme, daß nichtlinear (besser: im Schätzbereich $M_1=0$ ist) entnommen wird, was für die strenge Kontrolle von zweifelhaftem Wert ist.

Ist tatsächlich das Prozeßverhalten (unmeßbare Verluste/Hold-up nicht sprunghafte, system. Änderungen) sowie die Konstanz der Instrumente (system. Fehler) um z.B. einen Faktor 10 präziser als der gewöhnliche Eichvorgang, so daß in der Periode [20, 200] von $\sigma=90$ kg auf $\sigma=12$ kg reduziert werden kann? Vermutlich hat man im Grunde immer schließlich die gleiche Schwierigkeit vorliegen: Konstanz des Meßvorgangs. Es kann aus organisatorischen Gründen der Anlage durchaus sein, daß nur die Schätzung ein Zurückschieben der Grenzen der Konstanz erlaubt.

c) Verbesserung der Entdeckung mit CUMUFR

CUMUFR gibt mit dem zugrunde gelegten Datensatz (extrem großes $\sigma_S^2(T) \gg \sigma_R^2(T+I)$) starke Reduzierungen der Gesamtvarianz von Teilbilanzen.

- eine Tagesbilanz wird von $2 \times 4,46 + 0,09 + 0,26 = 9,27$ auf $\sigma^2=5$ reduziert; damit ist eine abrupte Entwendung signifikanter entdeckbar
- eine Summe von Teilbilanzen wird wegen der Unabhängigkeit z.B. in dem Zeitintervall von $[20,200]$ auf $\sigma^2=180 \times 5$ reduziert gegenüber $\sigma_{180}^2=90^2$ ohne Schätzung; damit wäre eine Entnahme von 40 kg einmal mit $\Phi\left(\frac{40}{90} - 1,64\right) = 11,5\%$ durch den Summentest zu erfassen, während diese mit Schätzung approximativ durch $\Phi\left(\frac{40}{30} - 1,64\right) = 38\%$ entdeckbar wäre im Falle protracted loss. Jedoch ist dies eine sehr grobe Approximation, da die Entnahme in nichtlinearer Weise in $E(MUF_{i-1})$ transformiert wird und für diese Transformation der Schätzbereich $[0,20]$ ohne Entnahme entscheidend ist. Einerseits nehmen $MUF_{i-1} - E(MUF_{i-1} | \dots)$ infolge der eigenen Kompensation durch die lineare Entnahme, die in MUF_1, \dots, MUF_{i-1} immer stärker Einfluß gewinnt, ab, so daß die Transformation unterhalb der Geraden $i \cdot E(MUF_{i-1})$ liegen würde, andererseits ergibt der Vorlauf in $[0,20]$ eine gute Schätzung des system. Fehlers, was sich technisch im Filter durch Addition einer positiven Korrekturgröße $\delta [MUF_{i-1} - E(MUF_{i-1} | \dots) + \delta_{i-1}]$ ab dem Entnahmbereich $i=21$ niederschlägt. Diese Größe würde gerade bei Entnahme ab $n=1$ sofort durch $E(MUF_{i-1} | \dots)$ kompensiert werden.
- mit 10-tägiger Bilanz ist die Entnahme $M=40$ kg in $[20,200]$ sogar mit 95% entdeckbar
- es darf aber weiterhin nicht vergessen werden, daß die lineare Entnahme $M=40$ kg in $[1,200]$ von keinem Verfahren entdeckt wird und die Schätzung deswegen so gravierend verbessert, weil gerade die schätzbaren Einflüsse (system. Fehler σ_S) extrem hoch sind; sobald hier eine zeitliche Inkonzanz auftritt, würden ohne Änderung des Modells Fehlalarme auftreten, also muß dann der Filter von vorne beginnen bzw. müssen kurzfristige system. Fehler modelliert werden, was die Gesamtverbesserung wesentlich begrenzt. Der Filter vergleicht infolge der starren Verbindung zu den Vorabschnitten immer kleinste Schwankungen innerhalb einer großen Abweichung vom Nullzustand (system. Fehler), diese starre Verbindung muß aber von der Realität des Prozesses erfüllt werden. In jedem Fall ist $\sigma_R^2=9+18$ nach 200 Tagen eine Schranke, so daß $3,3 \times \sigma_R \approx 17$ kg noch überwachbar wären; die Bedeutung der random. Fehler [Reduzierung, lineare Entnahme] ist in

5. erläutert. Denn nur randomisierte Fehler würden den "Filter" CUMUF ergeben, dessen endgültige Schätzgrenze genau $3,28\sigma_R$ (nicht sequentieller Test) ist. Ob die system. Fehler über lange Perioden derart starr sind, damit diese Eliminierung so weitgehend durchgeführt werden kann, ist die äquivalente Frage der realistischen Modellbildung, der Einführung von kurzfristigen system. Fehlern, ob es besser ist, mit dieser Methode oder durch häufiges Kalibrieren - falls möglich - auf σ_{res} zu reduzieren, ob die Empfindlichkeit überfordert ist, was zu Fehlalarmen durch Meßsystemschwankungen führt.

- die Hold-up-Schwankungen dürfen bei dieser adaptiven Schätzung nicht einfach randomisiert eingeführt werden. Es könnte in den Perioden 1 bis 20 z.B. immer die gleiche Abweichung in einer Richtung von 0,5 kg auftreten, insbesondere falls der Betreiber dies manipulieren kann. (Hat er keinen Einfluß, so kann er eventuell nicht an eine Toleranz gebunden werden.) Damit würde der tägliche Filter auf diese 0,5 kg als Nullzustand geeicht werden. Dann könnte ab der Periode 21 die Abweichung auf Null gehen oder randomisiert weitergehen, was bedeutet, daß entweder der Entwender über die restlichen 180 Perioden durch Entnahme à 0,5 kg ohne Alarm kompensieren kann, d.h. 90 kg entnimmt oder bei keiner Entnahme nur die Abweichung à 90 kg einen Fehlalarm produziert. Abgesehen von der Fehlalarmmöglichkeit wäre dies bei Manipulation keine zusätzliche Schwierigkeit, denn es wäre das gleiche, wie eine sofort ab $n=1$ beginnende lineare Entnahme. Falls zur Verhinderung dieser jedoch die Totalkontrolle in $[1,20]$ eingeführt wird, ergäbe dies eine Möglichkeit, manipulativ den Hold-up in $[1,20]$ zu erhöhen - falls "Alarm" eintritt in $n=20$ wäre bei clean-out noch alles Material da, falls keiner eintritt, kann risikolos danach nun eine tatsächliche Entnahme linear fortgeführt werden. Oder diese Erhöhung des Hold-up kann eventuell unbeabsichtigt von alleine auftreten und die unerwünschten Folgen haben.

d) Weitere Schätzfunktionen mit verbesserter Entdeckung

Mit den gleichen Daten der Referenzanlage lassen sich neben CUMUFR weitere Schätzer entwickeln

- eine Schätzfunktion, die den protracted loss von $M=40$ kg über $[20,200]$ am Ende des Referenzzeitraumes von einem Jahr mit 65% entdeckt, über $[25,200]$

mit 84%, sowie eine beliebig verteilte Entnahme von nur $M=10$ kg in $[1,20]$ mit 95%. (Protracted würde ohne Schätzer mit 11,5% entdeckt, die Entnahme von 10 kg mit $\Phi(\frac{10}{10} - 1,64) = 26\%$ und selbst bei abrupter Entnahme in einem Tag höchstens mit $\Phi(\frac{10}{3,05} - 3,18) = 53\%$, wobei die Fehlalarme für Tagesbilanzen in $[1,20]$ gleichmäßig verteilt wurden, also $\alpha_1 = 0,25\%$, damit auch am 20. Tag eine abrupte Entnahme erkannt wird. Der Schätzer CUMUFR würde eine auf $[1,20]$ verteilte oder auf den Tag 1 konzentrierte Entnahme von 10 kg ebenfalls mit $1-\beta \leq 25\%$ entdecken.) Eine abrupte Entnahme von 5,4 kg in einem Tag irgendwo im ganzen Zeitraum wird mit 99% entdeckt zu dem Signifikanzniveau 4σ , d.h. keine auftretenden Fehlalarme, während für den Test CUMUF ohne Schätzung lediglich $\Phi(\frac{5,4}{3,05} - 4) = 1,3\%$, also keine Entdeckung, erreicht werden kann. CUMUFR müßte, um eine hohe Entdeckung für den protracted loss zu erreichen, mit einer Bilanzfrequenz von 10 Tagen ausgeführt werden; dann ergibt die doppelt so hohe abrupte Entnahme $M=10$ kg eine Entdeckung von 25%.

Dieser Schätzer ist sehr einfach zu berechnen [es ist nur kurze Kopfrechnung nötig], konzeptmäßig ein in gewissem Sinne optimaler Schätzer. Ein wesentlicher Vorteil ist die leichte Implementierung ohne Simulationen.

- eine weitere komplizierter strukturierte Schätzfunktion, die jedoch numerisch nur geringen Aufwand erfordert. Diese ergibt für den protracted loss über $[20,200]$ von $M=28$ kg eine Entdeckung von $1-\beta = 95\%$, für den protracted loss über $[25,200]$ von $M=20$ kg eine Entdeckung von 95%. (Hier ergibt CUMUF $1-\beta = \Phi(\frac{20}{87,5} - 1,64) = 7,9\%$ praktisch keine Entdeckung; CUMUFR würde 40 kg mit 95% entdecken.) Weiterhin eine Entdeckung der auf 20 Perioden irgendwo verteilten abrupten Entnahme von 30 kg mit $1-\beta = 82\%$ jedoch zum sicheren Signifikanzniveau 4σ , d.h. kein Fehlalarm analog dem Attributtest. (CUMUF ergibt $1-\beta = \Phi(\frac{30}{10} - 4) = 15,8\%$).

Die abrupte Entnahme von 5,4 kg in einem Tag würde genauso zu 4σ mit 99% auf sichere Weise entdeckt. Hier wird also gleichzeitig der protracted loss von $M=20$ kg mit 95% und diverse abrupte Entnahmen wie $M=5,4$ kg an einem Tag oder $M=30$ kg in 20 Tagen mit sicherer Signifikanz entdeckt.

Diese Schätzfunktionen befinden sich gegenwärtig in Entwicklung zusammen mit einer systematischen Untersuchung des Fragenkreises mit dem Ziel, eine bestmögliche Kombination von Schätzung/Test zu geben, so daß die Schätzung nicht zu extensiv eingesetzt wird und dann bei jeder geringen unbekanntem Änderung der system. Einflüsse des realen Meßsystems empfindlich reagiert [Fehlalarme]. Auf Prozeßänderungen selbst [≙ Entnahmen, wegen der Differenzbildung] soll ja so empfindlich wie möglich reagiert werden, was genau die beiden nicht gleichzeitig erfüllbaren widerstrebenden Forderungen für die Entdeckung insbesondere des protracted loss bei großen Durchsätzen sind.

7. Programm, um das Ziel der Kontrolle einer Reprocessing Anlage zu erreichen

Es wird in großen Zügen ein Programm angegeben, wie dieses Ziel (und zugleich die Kontrolle jeder beliebigen Anlage bzw. Grundbilanz mit hoher Standardabweichung σ_{Gesamt} , d.h. auch Zusammenkopplung von MUF-Sequenzen) sicher erreicht werden kann, wenn auch dafür noch beträchtliche Mühe erforderlich ist. Mit allem Nachdruck muß nochmals betont werden, daß die seit ca. 1977 laufend in immer größerer Zahl entwickelten sequent. Verfahren am Kern der Sache primär vorbeigehen (d.h. immer in dem gegebenen Beispiel $1-\beta(\mu^*) \approx 5\%$, $|\mu^*| = 8$ kg ergeben), möglicherweise auch für ein schnelles System keine gangbare Lösung bei diesem σ_{Gesamt} bieten.

Eine wichtige Erkenntnis für eine letzte Grenze des Meßsystems kann vorweg gegeben werden. Weder sind es die randomisierten Fehler, die ein unüberwindbares Hindernis bieten (Wiederholung der Messung), noch keineswegs die systematischen Fehler, die gerade notfalls durch Schätzung beliebig weit eliminiert werden könnten zusammen mit einer sorgfältigen Kalibrierung zu Beginn, sondern genau die kurzfristigen system. Fehler, die im Modell wandernde Skalen, Inkonstanz der Meßinstrumente usw. abbilden und eine beliebig verfeinerte Schätzung verhindern. Diese Inkonstanz wäre also am stärksten zu vermindern von der technischen Seite. Die kurzfristigen systematischen Fehler sind durch keine statistische Technik mehr reduzierbar (wohlgemerkt bei dem Problem der laufenden Anlage, wobei der system. Fehler durch Differenzbildung eliminiert wird), denn der Prozeß selbst kann nicht parallel wiederholt werden. Der wandernde kurzfristige systemat. Fehler ergibt nach Beseitigung des random.+ system. Fehlers aufgrund der Differenzbildung eine langgestreckte unreduzierbare Wellenbewegung um einen letzten nichtkontrollierbaren (äußere Einflüsse) residualen Bias.

Programm

- A) Qualitative Analyse des Meßsystems mit grob geschätzten σ_r , σ_s , σ_k (kurzfristige Fehler) - bisher $\sigma_R^2 = 27$ $\sigma_S^2 = 10\ 600$ - um im Detail des Prozesses zu sehen, wo eine Reduzierung den größten Nutzen bringt.

- B) Meßinstrumentalternativen (Präzision $[\sigma_r, \sigma_s, \sigma_k]$ gegen Aufwand/Kosten) auflisten und analysieren; Angabe, bis zu welcher Präzision Messungen benötigt würden, um die Kontrolle in gewünschter Form zu erhalten [Beispiel: ist die Konzentrationsbestimmung mit 0,00 02% erreichbar?].
- C) Die wechselseitige Analyse von A) und B) ergibt eine Wahl [event. Neuentwicklung] des schließlichen Meßsystems sowie die Angabe von $[\sigma_r, \sigma_s, \sigma_k]$ der Messungen mit dem Ziel, σ_{Gesamt} zu minimieren unter den vielfältigen restriktiven Bedingungen der Kosten, der vorhandenen Instrumente; eventuelle beschränkte Wiederholung der Messungen sowie Kalibrierung schon eingeschlossen.
- D) Fehlerstruktur modellieren aufgrund dieser Wahl. Dies ist der schwierige Angelpunkt, da $[\sigma_r, \sigma_s, \sigma_k]$ immer eine Idealisierung darstellen. Deswegen konservative Werte annehmen, was σ_{Gesamt} erhöht.
- Bis hierher hat man dann den Datensatz zusammen mit der Fehlerstruktur unter Berücksichtigung von Verbesserungen [Probleme I, II] erhalten. Falls mit diesem σ_{Gesamt} die Anlage nicht kontrollierbar ist für den protracted loss, was auch am Kostenparameter aus C) liegen kann und zeitlich irgendwann sowieso eintritt
- E) Schätzverfahren einführen gemäß Teil 6. Hierbei muß berücksichtigt werden, daß zwar wesentliche Verbesserungen erzielbar sind, deswegen sollte auf jeden Fall ein Schätzer genommen werden, aber noch keine sichere Kontrolle [Problem IIIa]
- F) das Aufwandverteilungsproblem, Verteilung auf C) und E) falls totale physische Kontrolle möglich; für gegebenen Aufwand größtmögliches garantiertes $1-\beta$ für einen gegebenen Referenzzeitraum; rechtzeitige Entdeckung aufgrund der Verfahren in suboptimaler Weise sowieso erfüllt.
- G) Verifikation einbauen mit dem Aufwand als Parameter Gesamtaufwandverteilungsproblem für Verifikation einerseits und die Verteilung auf Wiederholungen oder physische Kontrolle zur Reduzierung von σ_{Gesamt} andererseits.

V. Referenzen

- 1 Robbins, H.; Siegmund, D.; Confidence Sequences and Interminable Test, Internat. Statistical Institute Bull., Vol. 43 (1969), p. 379-87
- 2 Sellinschegg, D.; A Statistic Sensitive to Deviations from the Zero-Loss Condition in a Sequence of Material Balances, Journal of the Institute of Nuclear Materials Management, XI, 4 (1982), p. 48-59
- 3 Zerrweck, K.E.; unveröffentlichter Bericht, 1981
- 4 Avenhaus, R.; Jaech, J.; On Subdividing Material Balances in Time and/or Space, Journal of the Institute of Nuclear Materials Management, X, 3 (1981), p. 24-33

Adresse des Autors: E. Zerrweck, Friedrichstr. 48, 7853 Höllstein