Von der Zufallsvariablen zum Schätzwert

Statistische Begriffe in der Messtechnik

Hubert Schmidt

In der geodätischen Ausgleichungsrechnung besteht häufig die Auffassung, ein Schätzwert sei mit einer Varianz behaftet. Die Definitionen der mathematischen Statistik weisen die Varianz jedoch nicht dem Schätzwert, sondern der Schätzvariablen zu. Sollen (insbesondere bei der Berücksichtigung anderweitig ermittelter Schätzwerte bei Folgemessungen) Fehldeutungen vermieden werden, muss dieser Unterschied bei der Varianzzuweisung unbedingt beachtet werden. In diesem Beitrag werden die Eigenschaften von Zufallsvariablen und deren Realisierungen (Messwerte bzw. Schätzwerte) sowie von Konstanten diskutiert.

1 Einleitung

Die Modellbildung bei der induktiven (schließenden) Statistik, zu der die geodätische Ausgleichungsrechnung zählt, beruht auf Zufallsvariablen und Konstanten. Geodätische Messgrößen können nur dann als Zufallsvariable eingestuft werden, wenn sie

a) beim jeweiligen Messprozess tatsächlich gemessen werden und
b) den Gesetzen des Zufalls entsprechend zufällig streuende Messwerte (Realisierungen) annehmen können.

Werden diese Grundlagen nicht beachtet, entstehen Fehldeutungen, die unkorrekte Ergebnisse zur Folge haben können.

Beispielsweise wird zuweilen argumentiert, bei der Berichtigung von Messungen seien die Korrektionswerte als Zufallsgrößen zu behandeln, da sie als Schätzwerte mithilfe vorausgegangener Kalibriermessungen bestimmt worden und ihnen daher Varianzen zugeordnet seien. Wegen der Addition der Korrektionen seien die berichtigten Messwerte daher untereinander korreliert.

Die Messgrößen (Koordinaten, Höhen, usw.) der Anschlusspunkte einer Netzausgleichung sind ebenfalls durch vorausgegangene Messungen bestimmt worden. Diese werden aber in der Regel als Konstante behandelt, d. h. sie bilden eine sogenannte „varianzfreie Rechenbasis“. Bei der freien Netzausgleichung dagegen gibt es „bewegliche“ oder „freie“ Anschlusspunkte mit Varianzen.

Diese Beispiele verdeutlichen die unterschiedliche Behandlung vorab bestimmter Messgrößen. Die Tatsache, dass sie vorab bestimmt wurden, liefert keine Begründung dafür, dass sie in späteren Messprozessen als Zufallsgrößen aufzufassen sind. Es wird gezeigt, dass vorab bestimmten Messgrößen, deren Schätzwerte in späteren Messprozessen verwendet werden, nur dann Varianzen und Kovarianzen zugeordnet werden können, wenn diese Messgrößen bei ihrer Verwendung in späteren Messprozessen die Kriterien von Zufallsvariablen erfüllen.

2 Zufallsvariable, Schätzfunktion, Schätzwert

Definition einer Zufallsvariablen:
Nach Hartung et al. (2002, Kap. II.6) lautet die Definition einer Zufallsvariablen (auch zufällige Variable oder Zufallsgröße genannt):

> Allgemein sagt man, X sei eine Zufallsvariable, wenn die Werte dieser Größe reelle Zahlen sind, die durch ein Zufallsexperiment bestimmt werden und wenn für die Ereignisise, die man damit beschreiben kann, Wahrscheinlichkeiten angebbar sind. Den Wert x, den die Zufallsvariable X bei der Durchführung des Zufallsexperiments annimmt, nennt man Realisation von X.

Diese Definition entspricht sinngemäß den Definitionen z. B. in Rao (1973, Kap. 2a.5), Mood et al. (1974, Kap. II 2.2), Koch (1997, Kap. 221), Fahrmeir und Hamerle (1984, Kap. 1), Schwarze (2001, Kap. 8.1), Bock (2004, Kap. 7), Benning (2002, Kap. 3.2). In der Messtechnik wird die *Realisation* bzw. *Realisierung* auch *Messwert* oder *Beobachtung* genannt.

Der *Erwartungswert*

$$E(X) = \mu \tag{1}$$

der Zufallsvariablen X ist als gewogenes arithmetisches Mittel aller ihrer theoretisch möglichen Realisierungen x_i definiert.

Die *Varianz* $var(X) = \sigma^2$ kennzeichnet das Streuen, welches ebenfalls alle theoretisch möglichen Realisierungen x_i einer Zufallsvariablen X um ihren Erwartungswert $E(X) = \mu$ aufweisen. Sie ist als Erwartungswert aller quadrierten Zufallsabweichungen $(X-\mu)$ definiert:

$$var(X) = \sigma^2 = E[(X-\mu)^2] \tag{2}$$

Der Erwartungswert μ und die Varianz σ^2 sind Parameter der Zufallsvariablen X.[1]

Definition der Schätzfunktion und des Schätzwertes:

In Bock (2004, Kap. 11) wird der Zusammenhang zwischen Zufallsvariablen, Schätzfunktion und Schätzwert verdeutlicht:

> Es sei X eine Zufallsvariable, deren Wahrscheinlichkeitsverteilung P_ϑ einen unbekannten Parameter ϑ enthält. Eine Funktion $\hat{\vartheta}(\cdot)$, die den Realisierungen $x_1,\cdots,x_n$ von n unabhängigen Wiederholungen $X_1,\cdots,X_n$ von X einen festen Zahlenwert $\hat{\vartheta} = \hat{\vartheta}\ (x_1, \cdots, x_n)$ zuordnet, heißt Schätzfunktion, Schätzer oder Schätzstatistik. Der Wert $\hat{\vartheta}$ nach Einsetzen der Realisierungen $x_1,\cdots,x_n$ heißt *Schätzwert* oder *Punktschätzung* für ϑ.
>
> - *Vor* Beobachtung ist $\hat{\vartheta}(X_1 \cdots, X_n)$ eine Zufallsvariable.
> - *Nach* Beobachtung ist $\hat{\vartheta}(x_1, \cdots, x_n)$ eine feste Zahl.

Die Schätzfunktion ist eine Rechenvorschrift, mit der aus Eingangsvariablen die Ergebnisvariable abzuleiten ist, für die sich nach Einsetzen der Realisierungen der Ergebniswert ergibt. Sind beispielsweise alle Wiederholungen $X_1,\cdots,X_n$ von X paarweise stochastisch unabhängig und besitzen sie alle denselben Erwartungswert $E(X_i) = \mu$ und dieselbe Varianz $var(X_i) = \sigma^2$, liefert die Schätzfunktion die erwartungstreue Zufallsvariable (*Schätzvariable, Ergebnisvariable*)

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{3}$$

mit dem Erwartungswert

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$$
$$= \frac{1}{n}\cdot n \cdot \mu = \mu. \tag{4}$$

Die Schätzvariable $\bar{X}$ besitzt die Varianz

$$var(\bar{X}) = \sigma^2_{\bar{X}} = \frac{\sigma^2}{n}. \tag{5}$$

Durch Einsetzen der Messwerte liefert die Schätzfunktion den erwartungstreuen Schätzwert $\bar{x}$ der Schätzvariablen $\bar{X}$.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{6}$$

Die empirische Varianz

$$s^2_{\bar{X}} = \frac{s^2}{n} \tag{7}$$

ist der Schätzwert der a priori-Varianz $\sigma^2_{\bar{X}}$ der Zufallsvariablen $\bar{X}$.

Die a priori-Varianz $\sigma^2_{\bar{X}}$ und ihr Schätzwert $s^2_{\bar{X}}$ sind der Schätz-*variablen* $\bar{X}$ und nicht dem Schätz-*wert* $\bar{x}$ zugeordnet. Daher ist in Gl. (7) die Zufallsvariable $\bar{X}$ als Index angegeben. Die Varianz nach Gl. (2) kennzeichnet das Streuen *aller* wahrscheinlichkeitstheoretisch möglichen Werte der Zufallsvariablen um deren Erwartungswert. Die übliche Schreibweise der Standardabweichung $s_{\bar{x}}$ mit dem Schätzwert $\bar{x}$ als Index ist folglich missverständlich.

Der Schätzwert $\bar{x}$ ist *einer* der möglichen, entsprechend den Gesetzen des Zufalls sich ergebenden Werte (Realisierungen) der Schätzvariablen $\bar{X}$. Streuverhalten lässt sich bekanntlich nicht für *einen* Wert, sondern nur zwischen mehreren Werten konstatieren. Da die Varianz $\sigma^2_{\bar{X}}$ und ihr Schätzwert $s^2_{\bar{X}}$ den Streubereich *aller* theoretisch möglichen Werte der Zufallsvariablen $\bar{X}$ kennzeichnen, trifft die übliche Ausdrucksweise: „Der Schätzwert $\bar{x}$ besitzt die Standardabweichung $s_{\bar{x}}$“ diesen Sachverhalt nicht korrekt. Der Schätzwert $\bar{x}$ ist eine feste Zahl und kann daher nicht variabel sein.

In der Literatur der geodätischen Ausgleichungsrechnung wird die Zufallsvariable X zuweilen unkorrekterweise als Realisierung x bzw. Schätzwert $\bar{x}$ definiert. So heißt es in Pelzer (1985, Kap. 1.1.1):

> Eine Zufallsgröße X ist das durch eine Zahl ausgedrückte Ergebnis eines Experiments, dessen Ausgang innerhalb gewisser Grenzen ungewiss ist.

Diese Definition, welche eine reelle, feste Zahl als Zufallsvariable deklariert, steht im Widerspruch zu den oben zitierten Definitionen der mathematischen Statistik.

3 Linerarkombinationen von Zufallsvariablen

In diesem Kapitel soll die Behandlung der Messgrößen (Variablen oder Konstanten) von Anschlusspunkten verdeutlicht werden. Zuvor sei dazu die Varianz/Kovarianzfortpflanzung der Linearkombination von Zufallsvariablen rekapituliert, siehe Hartung et al. (2002, Kap. 8.4), Koch (1997, Kap. 233). Man beachte, dass die Zufallsvariablen mit großen Buchstaben und die Konstanten (reelle Zahlen) mit kleinen Buchstaben bezeichnet werden.

Für eine Linearkombination von m Zufallsvariablen $X_i (i = 1,...,m)$ mit den Erwartungswerten $E(X_i)$ und den Varianzen $\sigma^2_{X_i}$ gilt:

$$\text{Zufallsvariable } Y = a_1 \cdot X_1 + \ldots + a_m \cdot X_m + b \tag{8}$$

$$\text{Erwartungswert } E(Y) = a_1 \cdot E(X_1) + \ldots + a_m \cdot E(X_m) + b \tag{9}$$

Sind die X_i paarweise stochastisch unabhängig, so gilt außerdem:

$$\text{Varianz } \sigma^2_Y = a_1^2 \cdot \sigma^2_{X_1} + \ldots + a_m^2 \cdot \sigma^2_{X_m} \tag{10}$$

[1] Die wahrscheinlichkeitstheoretischen Begriffe der induktiven Statistik wie „Zufallsvariable“, „Erwartungswert“, „Erwartungstreue“, „Schätzwert“, „Freiheitsgrad“ usw. werden in der Literatur der Ausgleichungsrechnung zuweilen der deskriptiven Statistik zugeordnet, z. B. in Niemeier (2002, Kap. 1.3). Das führt zu Missverständnissen, da die deskriptive Statistik sich lediglich damit befasst, vorhandene Datenmengen in Diagrammen darzustellen und durch Kennzahlen zusammenzufassen (Anzahl PKW pro Haushalt). Wenn auch die Kennzahlen der deskriptiven Statistik teilweise die gleichen Bezeichnungen wie die Schätzwerte der induktiven Statistik tragen (Mittelwert, Varianz, Standardabweichung, Korrelationskoeffizient), so beruhen sie doch auf anderen statistischen Grundlagen und zum Teil abweichenden Berechnungsformeln. Parameter der Grundgesamtheit und wahrscheinlichkeitstheoretische Rückschlüsse (Konfidenzintervalle, Hypothesentests) sind in der deskriptiven Statistik nicht definiert.

Gl. (8) bis (10) verdeutlichen, dass eine additive Konstante b zwar den Erwartungswert, nicht aber die Varianz der Ergebnisgröße beeinflusst.
Werden mehrere Zufallsvariable Y_j durch Linearkombinationen gebildet und treten eine oder mehrere der Zufallsvariablen X_i in mehreren der Kombinationen gleichzeitig auf, ergeben sich bekanntlich Kovarianzen $\sigma_{Y_jY_k}$ und Korrelationen $\rho_{Y_jY_k}$ zwischen den Ergebnisgrößen Y_j und Y_k. Auch auf die Kovarianzen und Korrelationen haben additive Konstanten keinen Einfluss.
Nachfolgend soll die Deklarierung der Messgrößen von Anschlusspunkten am Beispiel eines Nivellements von einem Anschlusspunkt A aus über zwei Wechselpunkte W_1 und W_2 zu einem Neupunkt E (drei Rückblick- und Vorblickvariablen R_i und V_i) verdeutlicht werden. Die Höhe des Anschlusspunktes A wird zunächst als Konstante h_A aufgefasst. Die Konstanten a_i in Gl. (8) weisen hier die Werte $+1$ bzw. -1 auf.

$$\textit{Höhenvariable } H_E = h_A + (R_1 - V_1) + (R_2 - V_2) + (R_3 - V_3) \tag{11}$$

Mit der Annahme $\sigma^2_{R_i} = \sigma^2_{V_i} \Rightarrow \sigma^2$ folgt:

$$\textit{Varianz der Höhenvariablen } \sigma^2_{H_E} = 3(\sigma^2_R + \sigma^2_V) = 6\sigma^2 \tag{12}$$

Durch Einsetzen der Messwerte (Ablesungen r_i und v_i) in Gl. (11) ergibt sich der Höhenwert h_E des Neupunktes E als reelle Zahl:

$$\textit{Höhenwert } h_E = h_A + (r_1 - v_1) + (r_2 - v_2) + (r_3 - v_3) \tag{13}$$

Fasst man das in Gl. (11) geschlossen dargestellte Nivellement als separate Einzelnivellements zwischen den Höhenpunkten auf, ergeben sich die Höhenvariablen der Wechselpunkte und des Endpunktes mit ihren Varianzen:

$$H_{W_1} = h_A + (R_1 - V_1), \quad \sigma^2_{H_{W_1}} = 2\sigma^2 \tag{14}$$

$$H_{W_2} = H_{W_1} + (R_2 - V_2), \quad \sigma^2_{H_{W_2}} = \sigma^2_{H_{W_1}} + 2\sigma^2 = 4\sigma^2 \tag{15}$$

$$H_E = H_{W_2} + (R_3 - V_3), \quad \sigma^2_{H_E} = \sigma^2_{H_{W_2}} + 2\sigma^2 = 6\sigma^2 \tag{16}$$

Die Wechselpunkte W_1 und W_2 sollen hier jeweils die Anschlusspunkte für die nachfolgenden Einzelnivellements darstellen. Die geschlossene Darstellung nach Gl. (11) und die gestaffelte Modellierung nach Gl. (14) bis (16) liefern für die Höhenvariable des Neupunktes E die identische Varianz.

- Dieses Beispiel verdeutlicht, dass die Messgrößen von Anschlusspunkten als Zufallsvariable aufgefasst werden können, wenn zwei separate Messvorgänge unmittelbar miteinander verknüpft werden, sodass die beiden als *ein* Messvorgang darstellbar sind. Die Zufallsvariablen der Anschlusspunkte beinhalten den Varianzumfang (Präzisionsinformation) der vorhergehenden Messgrößen.
- Ist die Einstufung einer zuvor bestimmten Messgröße bei Folgemessungen als Zufallsvariable zulässig, folgt als nächstes die Entscheidung darüber, ob diese Deklarierung sachlich auch erforderlich bzw. erwünscht ist. Diese Frage lässt sich mithilfe des Höhenanschlusspunktes A des Beispielnivellements diskutieren.
  - *Messgröße des Anschlusspunktes A = Zufallsvariable*: Will man die Präzisionsinformationen der Bestimmungsmessungen des Anschlusspunktes in die Folgemessungen einbeziehen, fasst man die Messgröße des Anschlusspunktes bei der Modellbildung als Zufallsvariable (hier H_A) auf. Gl. (11) und (12) zeigen sich dann in der Form:

$$\textit{Höhenvariable} \quad H_E = H_A + (R_1 - V_1) + (R_2 - V_2) + (R_3 - V_3) \tag{17}$$

$$\textit{Varianz der Höhenvariablen } \sigma^2_{H_e} = \sigma^2_{H_A} + 6\sigma^2 \tag{18}$$

  - *Messgröße des Anschlusspunktes A = Konstante*: Will man hingegen die Präzisionsinformationen der Bestimmungsmessungen des Anschlusspunktes *nicht* einbeziehen, fasst man dessen Messgröße als Konstante auf, wie mit der Anschlusshöhe h_A in Gl. (11) geschehen. Diese Einstufung ist beispielsweise dann sinnvoll, wenn zur Beurteilung der relativen Höhenbewegung der Punkte eines Ingenieurnetzes hochgenaue Nivellements (mm-Genauigkeit) an den Höhenbolzen des amtlichen Höhennetzes (cm-Genauigkeit) angeschlossen werden und der Höhenanschluss nur der höhenmäßigen Einpassung ins amtliche Netz dient.

Wird die Messgröße eines Anschlusspunktes als Zufallsvariable deklariert, muss weiterhin beachtet werden, ob sie bei der Varianzberechnung von abgeleiteten Messgrößen zu berücksichtigen ist oder nicht. Das soll das folgende Beispiel verdeutlichen:
Von einem Anschlusspunkt P_1 aus, dessen Höhe als Zufallsvariable H_1 deklariert wird, werden jeweils die Höhen H_2 des Punktes P_2 und H_3 des Punktes P_3 durch die Messgrößen L_{12} und L_{13} bestimmt.

$$H_2 = H_1 + L_{12}, \quad \sigma^2_{H_2} = \sigma^2_{H_1} + \sigma^2_{L_{12}} \tag{19}$$

$$H_3 = H_1 + L_{13}, \quad \sigma^2_{H_3} = \sigma^2_{H_1} + \sigma^2_{L_{13}} \tag{20}$$

Berechnet man für den Höhenunterschied

$$\Delta H_{23} = H_3 - H_2 \tag{21}$$

dessen Varianz mithilfe der zuvor berechneten Varianzen der Höhenpunkte

$$\sigma^2_{\Delta H_{23}} = \sigma^2_{H_3} + \sigma^2_{H_2} = 2\sigma^2_{H_1} + \sigma^2_{L_{13}} + \sigma^2_{L_{12}}, \tag{22}$$

ergibt sich ein um den Term $2\sigma^2_{H_1}$ falsches Ergebnis. Es wird nicht beachtet, dass in der Linearkombination, welche der Varianzberechnung zugrunde liegt, identische Zufallsvariablen (hier H_1) nicht mehrfach aufgeführt sein dürfen, da sie ansonsten als unterschiedliche, eigenständig Zufallsvariablen in die Berechnung eingehen. Die korrekte Modelierung lautet daher:

$$\Delta H_{23} = H_3 - H_2 = (H_1 + L_{13}) - (H_1 + L_{12}) = L_{13} - L_{12} \tag{23}$$

$$\sigma^2_{\Delta H_{23}} = \sigma^2_{L_{13}} + \sigma^2_{L_{12}} \tag{24}$$

4 Berücksichtigung von Korrektionsgrößen

Durch Kalibriermessungen lassen sich Korrektionsfunktionen für geodätische Instrumente bestimmen, mit denen Korrekturwerte zur Berichtigung der Messwerte anderweitiger Messgrößen berechnet werden können. Die Korrektur dient der (mehr oder weniger vollständigen) Eliminierung systematischer Messabweichungen und wird üblicherweise von der Instrumentensoftware bei der Messwerterfassung an den „rohen“ Messwerten programmgesteuert angebracht. Wegen der Vielzahl der Umrechnungen und Korrekturen, beginnend bei den Eingangssignalen der Messwerterfassung bis zur Ausgabe der berichtigten Messwerte (zum Teil mit Wechsel der physikalischen Einheiten), lässt sich ein separater Korrekturvorgang zum Teil überhaupt nicht erkennen, geschweige denn als Zufallsereignis definieren; zumal die nachträgliche Berücksichtigung einer Korrektur „von Hand“ die Ausnahme geworden ist.

In SCHMIDT (2004) wurde aus der Sichtweise der Regressions- und Prognosemodelle dargelegt, dass die Anbringung vom Korrektionen an Messwerten nur deren Berichtigung (Eliminierung systematischer Messabweichungen) bewirkt, aber keine Korrelation der berichtigten Messwerte verursachen kann. Nachfolgend wird diese Thematik mithilfe der in Kap. 2 angegebenen Definitionen der Wahrscheinlichkeitstheorie beleuchtet.

Die bei der Kalibrierung ermittelten Varianzen und Kovarianzen kennzeichnen die Präzision des Kalibriervorganges. Bei der Berichtigung anderweitiger Messgrößen lässt sich eine Korrektionsgröße prinzipiell *nicht* als Zufallsvariable auffassen, weil die oben genannten Kriterien der Zufallsvariablen nicht zutreffen:

Kriterium 1: *Tatsächliche Ermittlung von Messwerten* (siehe Kap. 2).

Nur bei den Kalibriermessungen stellen die zu ermittelnden Korrektionsgrößen Zufallsvariablen dar. Während des Messprozesses zur Ermittlung der anderweitigen (zu korrigierenden) Messgrößen lassen sich für die Korrektionsgrößen keine Messwerte ermitteln. Kriterium 1 wird nicht erfüllt.

Kriterium 2: *Zufallsstreuen bei Wiederholungsmessungen* (siehe Kap. 2).

Bei mehrfacher Wiederholung des Messprozesses für die zu korrigierenden anderweitigen Messgrößen nimmt die Korrektionsgröße keine zufällig streuenden Werte an. Sie ist im wahrsten Sinne des Wortes eine Konstante, weil bei Wiederholungsmessungen stets der gleiche konstante Wert berücksichtigt wird. Kriterium 2 wird nicht erfüllt.

Kriterium 3: *Anschlusspunkt als Zufallsvariable* (siehe Kap. 3).

Die Korrektionsgröße lässt sich auch aus räumlichen und sachlichen Gründen nicht als „Messgröße eines Anschlusspunktes“ deklarieren, über den sich die Kalibriermessungen und die zu korrigierenden anderweitigen Messgrößen verknüpfen ließen. Wegen der räumlichen Trennung zwischen dem Kalibriernetz und dem zu bestimmenden Netz der Neupunkte und den unterschiedlichen Messzwecken ist ein direkter Anschluss beider Messprozesse und die Einstufung als *ein* geschlossener Messprozess nicht möglich. Kriterium 3 wird nicht erfüllt. Der Sachverhalt sei an nachfolgendem Nivellierbeispiel verdeutlicht.

4.1 Berücksichtigung der Nullpunktabweichung einer Nivellierlatte

Sollen der Höhenunterschied H_{AB} zwischen zwei Bodenpunkten A und B sowie der Höhenunterschied H_{AC} zwischen A und einem Deckenpunkt C nivellitisch bestimmt werden, stellen die Messgrößen L_A und L_B sowie (mit umgedrehter Latte) L_C Zufallsvariablen dar (Abb. 1), deren Messwerte mit l_A, l_B und l_C bezeichnet werden sollen. H_{AB} und H_{AC} sind die Ergebnisvariablen, für die sich nach Einsetzen der Messwerte in die Linearkombinationen die Ergebniswerte h_{AB} und h_{AC} zeigen. Die a priori-Varianzen der Messgrößen werden gleich groß angenommen:

$$\sigma^2_{L_A} = \sigma^2_{L_B} = \sigma^2_{L_C} \Rightarrow \sigma^2_L \tag{25}$$

Die Lattenteilung (Skalierung) hat auf das Zufallsstreuen der Ablesungen keinerlei Einfluss. Eventuelle systematische Abweichungen in der Lattenteilung stellen für die jeweilige Ablesestelle funktional berechenbare Konstanten dar. Die Nullpunktabweichung verdeutlicht den einfachsten Fall systematischer Abweichungen in der Lattenteilung.

Bekanntermaßen wird die Nullpunktabweichung bei der Auswertung eines Nivellements durch die Differenzbildung „Rückblick minus Vorblick“ vollständig eliminiert. Das könnte prinzipiell nicht der Fall sein, wenn die Nullpunktabweichung eine Zufallsvariable wäre. Nur weil sie eine Konstante ist, gelingt die vollständige Elimination.

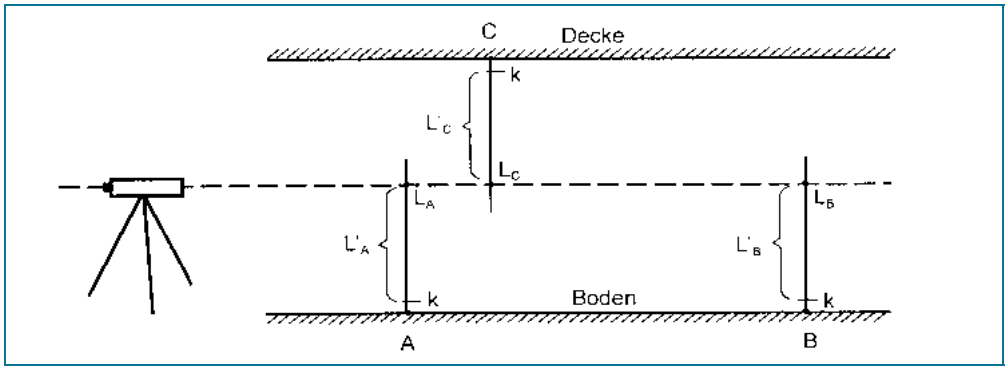

Abb. 1: Ablesungen an einer Nivellierlatte

Es wird hier davon ausgegangen, dass die Messgrößen L_A, L_B und L_C eine Nullpunktabweichung k beinhalten, sodass die berichtigten Messgrößen bezeichnet sein sollen in der Form $L'_A = L_A - k$, $L'_B = L_B - k$ und $L'_C = L_C - k$. Mit diesen Bezeichnungen ergeben sich die Linearkombinationen:

$$H_{AB} = L_A - L_B = (L'_A + k) - (L'_B + k) = L'_A - L'_B \quad (26)$$

$$H_{AC} = L_A + L_C = (L'_A + k) + (L'_C + k) = L'_A + L'_C + 2k \quad (27)$$

Die Berücksichtigung der Nullpunktkorrektion k beeinflusst die *Richtigkeit* des Höhenunterschiedes H_{AC}, aber nicht dessen *Präzision*. Die Anbringung einer konstanten Korrektion kann das Streuverhalten der Messwerte nicht beeinflussen. Es gilt folglich:

$$\sigma^2_{L_A} = \sigma^2_{L'_A}, \quad \sigma^2_{L_B} = \sigma^2_{L_B}, \quad \sigma^2_{L_C} = \sigma^2_{L'_C} \quad (28)$$

Die Kovarianzberechnung ergibt:

$$\begin{aligned} &\text{Varianz} && \sigma^2_{H_{AB}} = \sigma^2_{L_A} + \sigma^2_{L_B} \\ &\text{Varianz} && \sigma^2_{H_{AC}} = \sigma^2_{L_A} + \sigma^2_{L_C} \\ &\text{Kovarianz} && \sigma_{H_{AB}H_{AC}} = \sigma^2_{L_A} \end{aligned} \quad (29)$$

Mit $\sigma^2_{L_A} = \sigma^2_{L_B} = \sigma^2_{L_C} \Rightarrow \sigma^2$ folgt:

$$\begin{aligned} &\text{Varianzen} && \sigma^2_{H_{AB}} = \sigma^2_{H_{AC}} = 2\sigma^2 \\ &\text{Kovarianz} && \sigma_{H_{AB}H_{AC}} = \sigma^2 \\ &\text{Korrelationskoeffizient} && \rho_{H_{AB}H_{AC}} = 0,5 \end{aligned} \quad (30)$$

Eingesetzt in die Gl. (26) und (27) ergeben sich mit den Messwerten l_A, l_B, l_C und der Korrektion k die Höhenwerte:

$$h_{AB} = l_A - l_B \quad (31)$$

$$h_{AC} = l_A + l_C - 2k, \quad (32)$$

Es wird die bekannte Tatsache bestätigt, dass der berechnete Höhenunterschied h_{AB} in Gl. (31) von der Nullpunktabweichung k unbeeinflusst ist und zwar unabhängig davon, ob der Korrekturwert an den Messwerten angebracht wird oder nicht. Der Höhenunterschied h_{AC} in Gl. (32) hingegen wäre um den doppelten Betrag der Nullpunktabweichung verfälscht, wenn er nicht berücksichtigt würde. Da Korrektionen Konstante sind, haben sie auf die Varianz/Kovarianzberechnung prinzipiell keinen Einfluss.

4.2 Unkorrekte Einstufung von Korrektionen als Zufallsvariable

Bei der theoretischen Diskussion geodätischer Auswertungen wird zuweilen argumentiert, Korrekturen würden Korrelationen zwischen den berichtigten Messwerten verursachen. Da die Nullpunktabweichung als Unbekannte in der Ausgleichung der Kalibriermessungen bestimmt worden sei, sei ihr „Schätzwert eine Zufallsvariable“. Die Aussage „Schätzwert = Zufalls*variable*“ ist in Kap. 2 widerlegt worden. Wie dort bereits nachgewiesen, ist ein Schätzwert eine reelle Zahl und kann folglich keine Zufallsvariable sein. Konstante Korrektionen werden bei der Differenzbildung der berichtigten Messwerte vollständig eliminiert, was bei Zufallsvariablen gemäß Definition nicht möglich ist.

Folgte man trotzdem der angegebenen Argumentation und würde die Korrektur k als Zufallsvariable K einstufen, zeigten sich Gl. (26) und (27) in der Form:

$$H_{AB} = (L'_A + K) - (L'_B + K) = L'A - L'B \quad (33)$$

$$H_{AC} = (L'_A + K) + (L'_C + K) = L'_A + L'_C + 2K \quad (34)$$

Bezogen auf die Terme der rechten Seite in den Gl. (33) und (34) würde die Kovarianzberechnung liefern:

$$\begin{aligned} &\text{Varianz} && \sigma^2_{H_{AB}} = \sigma^2_{L_A} + \sigma^2_{L_B} \\ &\text{Varianz} && \sigma^2_{H_{AC}} = \sigma^2_{L_A} + \sigma^2_{L_C} + 4\sigma^2_K \\ &\text{Kovarianz} && \sigma_{H_{AB}H_{AC}} = \sigma^2_{L_A} \end{aligned} \quad (35)$$

Mit $\sigma^2_{L_A} = \sigma^2_{L_B} = \sigma^2_{L_C} \Rightarrow \sigma^2$ ergäben sich hier:

$$\begin{aligned} &\text{Varianz} && \sigma^2_{H_{AB}} = 2\sigma^2 \\ &\text{Varianz} && \sigma^2_{H_{AC}} = 2\sigma^2 + 4\sigma^2_K \\ &\text{Kovarianz} && \sigma_{H_{AB}H_{AC}} = \sigma^2 \\ &\text{Korrelationskoeffizient} && \rho_{H_{AB}H_{AC}} = \frac{0,5}{\sqrt{1 + 2\frac{\sigma^2_K}{\sigma^2}}} \end{aligned} \quad (36)$$

In die Varianz $\sigma^2_{H_{AC}}$ ginge demnach die Varianz σ^2_K mit dem *vierfachen* Betrag ein. Die Höhen H_{AB} und H_{AC} wären mit $\rho_{H_{AB}H_{AC}} < 0,5$ korreliert, d. h. bei der Argumentation „Korrektion = Zufallsvariable“ würde die Anbringung der Korrektionen anstelle einer Vergrößerung sogar eine Verkleinerung der Korrelation bewirken.

Literatur

[1] Benning, W.: Statistik in Geodäsie, Geoinformation und Bauwesen. Herbert Wichmann Verlag, Hüthig GmbH & Co. KG, Heidelberg 2002.

[2] Bock, H. H.: Formelsammlung zur Vorlesung Statistik I/II. Vorlesungskript SS 2004 und WS 2004/05, Institut für Statistik und Wirtschaftsmathematik der RWTH Aachen, 2004.

[3] Fahrmeir, L., Hamerle, A.: Multivariate statistische Verfahren. Walter de Gruyter, Berlin 1984.

[4] Hartung, J., Elpelt, B., Klösener, K.-H.: Statistik. R. Oldenbourg Verlag, München 2002.

[5] Koch, K. R.: Parameterschätzung und Hypothesentests. Ferd. Dümmlers Verlag, Bonn 1997.

[6] Mood, A. M., Graybill, F. A., Boes, D. C.: Introduction to the Theorie of Statistics. McGraw-Hill, 1974.

[7] Niemeier, W.: Ausgleichungsrechnung. Walter de Gruyter, Berlin 2002.

[8] Pelzer, H.: Grundlagen der mathematischen Statistik und der Ausgleichungsrechnung. In: Pelzer, H. (Hrsg.): Geodätische Netze in Landes- und Ingenieurvermessung II. Konrad Wittwer, Stuttgart 1985, S. 3 – 120.

[9] Rao, C. R.: Lineare statistische Methoden und ihre Anwendungen. Akademie-Verlag, Berlin 1973.

[10] Schmidt, H.: Zur Bestimmung und Anwendung von Kalibrier- und Korrektionsfunktionen. Allgemeine Vermessungs-Nachrichten Heft 4/2004, S. 141 – 146.

[11] SCHWARZE, J.: Grundlagen der Statistik II; Wahrscheinlilchkeitsrechnung und induktive Statistik. Verlag Neue Wirtschafts-Briefe, Herne/Berlin, 7. Auflage 2001.

Anschrift des Autors:
Dr.-Ing. HUBERT SCHMIDT,
Geodätisches Institut der RWTH Aachen,
Templergraben 55, D-52062 Aachen,
Tel.: 02 41-8 09 52 81,
hubert.schmidt@gia.rwth-aachen.de

Zusammenfassung

Alle in einem Messprozess ermittelten Messgrößen, die den Gesetzen der Wahrscheinlichkeit entsprechend zufällig streuende Werte annehmen können, werden in Linearkombinationen und Schätzfunktionen als Zufallsvariablen berücksichtigt. Die mithilfe des Varianz/Kovarianzfortpflanzungsgesetzes bzw. bei der Ausgleichungsrechnung ermittelten Varianzen und Kovarianzen sind nicht den Ergebnis*werten*, sondern den Ergebnis*variablen* zugeordnet. Systematische Messeinflüsse verfälschen die Messwerte (Realisierungen) der Zufallsvariablen und beeinflussen dadurch ihre Richtigkeit. Wenn die systematischen Messeinflüsse im jeweiligen Messprozess nicht direkt ermittelbar sind (was überwiegend der Fall ist), müssen die erforderlichen Korrektionen durch separate Kalibriermessungen bestimmt werden. Die zur Berichtigung der (späteren) Messwerte angebrachten Korrektionen können in den Linearkombinationen und Schätzfunktionen *nicht* als Zufallsvariablen berücksichtigt werden, da im konkreten Messprozess für sie keine Messwerte ermittelt werden. Ihrer systematischen Natur gemäß beeinflussen sie nicht das Zufallsstreuen der zu korrigierenden Messgrößen. Daher können sie in die Linearkombinationen und Schätzfunktionen nur als Konstante eingehen. Eine korrelierende Wirkung der Korrektionen ist definitionsgemäß ausgeschlossen.

Schlagwörter:

Messgröße, Korrektionsgröße, Zufallsvariable, Konstante

Abstract

The discussion of statistical terms and methods in geodetic adjustment techniques is often stamped by misunderstandings. It is to distinguish, if measurands have to understand as random variables or as constants. In this paper the behaviour of random variables and their realisations (measured values and estimated values respectively) on the one hand and of constants on the other hand is treated.