IABG. Die Zukunft

Anwendungen von Extremwerttheorie im Risikomanagement

Wie statistische Methoden die Eintrittswahrscheinlichkeiten für selten auftretende Ereignisse prognostizieren können

Whitepaper 18.11.2020

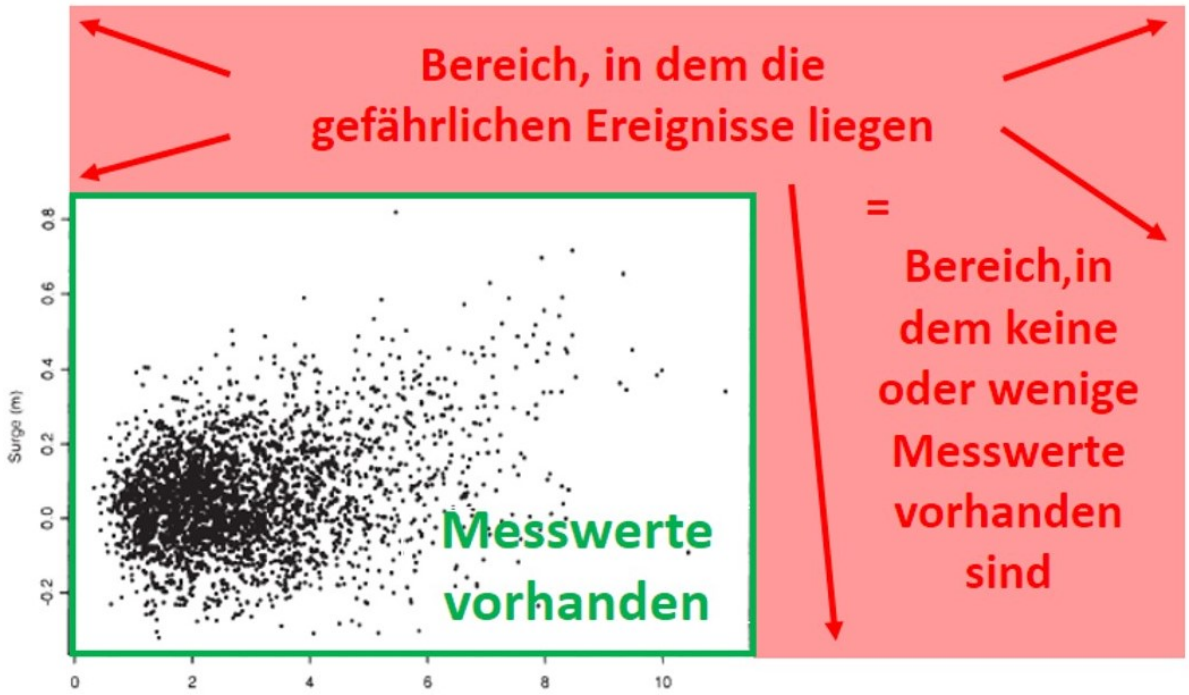
IABG

White Paper 03 der IABG Safety & Security Academy

Anwendungen der Extremwerttheorie im Risikomanagement

Wie statistische Methoden die Eintrittswahrscheinlichkeiten für selten auftretende Ereignisse prognostizieren können

Dr. Christine Schwarz-Hemmert und Dr. Lukas Höhndorf
schwarz-hemmert@iabg.de | hoehndorf@iabg.de



Zusammenfassung

Es gibt Ereignisse, die mit geringer Wahrscheinlichkeit eintreten aber katastrophale Auswirkungen haben können. Sie müssen identifiziert, analysiert und bewertet werden, um Risiken bestimmen und kontrollieren zu können, die in Zusammenhang mit den involvierten Systemen oder Phänomenen stehen. Um die Wahrscheinlichkeit für solche seltenen Ereignisse abzuschätzen, liefert die Extremwerttheorie hierzu geeignete Mittel. Die Extremwerttheorie ist eine auf der Auswertung von bereits auffälligen Beobachtungen und Messwerten basierende Methodik aus der Statistik für die Abschätzung solcher Wahrscheinlichkeiten.

In diesem Artikel wird illustriert, welchen Mehrwert eine auf der Extremwerttheorie basierende Prognose gegenüber Schätzungen aufgrund anderer statistischer Methoden bringen kann und wie

IABG Safety & Security Academy • Einsteinstr. 20 • 85521 Ottobrunn
Tel. +49 89 6088-4088 • academy@iabg.de • academy.iabg.de

das Vorgehen bei einer solchen Analyse prinzipiell aussieht. Des weiteren werden Anwendungen der Extremwerttheorie aus der Flugsicherheit und aus der satellitennavigationsdatenbasierten Fahrzeugnavigation vorgestellt, die im Rahmen von Projekten durch IABG-Mitarbeiter bearbeitet wurden.

Risikomanagement erfordert die Abschätzung von Eintrittswahrscheinlichkeiten

Wie häufig wird ein Sensormesswert um mehr als eine erlaubte Obergrenze vom tatsächlichen Wert abweichen? Mit welcher Wahrscheinlichkeit treten ggf. verkettete kritische Zustände in einer realen Systemumgebung auf? Oft gibt es nur sehr wenige oder überhaupt keine Feldbeobachtungen für die hinsichtlich ihrer Auftretenswahrscheinlichkeit einzuschätzenden kritischen Zustände, was für eine verlässliche Prognose solcher Wahrscheinlichkeiten mithilfe von Methoden aus der klassischen Datenanalyse problematisch ist.

Bei klassischen Datenanalysen für die Schätzung alltäglicher Ereignisse können sehr selten auftretende extreme Messwerte – also meist sehr große oder sehr kleine Messwerte – oftmals als Ergebnis einer fehlerhaften Messung betrachtet werden und für die Modellbildung unberücksichtigt bleiben. Bei Risikoanalysen hat jedoch neben der Betrachtung von Datenverteilungen immer auch die „Auswirkung“, die mit den jeweiligen Datenwerten in Zusammenhang steht, direkten Einfluss auf das Analyseergebnis. Deshalb sind hierbei die selten auftretenden Extremwerte keine zu ignorierenden Ausnahmen, sondern hoch relevante Fälle, die detailliert untersucht werden müssen.

Die Extremwerttheorie ist ein Teilgebiet der Statistik, das sich mit seltenen Situationen beschäftigt und beispielsweise bei Versicherungen und im Finanzsektor [3] sowie für Sicherheitsanalysen technischer Systeme [4] seit Jahren erfolgreich angewandt wird. In der Literatur findet man Anwendungen auf Fragestellungen wie z.B. „Wie hoch sollte ein Damm gebaut werden, um höchstens einmal in hundert Jahren überschwemmt zu werden?“ oder „Wie groß ist die Wahrscheinlichkeit, dass ein versicherter Schadensfall einen bestimmten Wert überschreitet?“ oder „Welche Auswirkungen könnte ein Börsencrash morgen auf ein bestimmtes Finanzprodukt haben?“ oder „Bei welchen Krafteinwirkungen wird ein Glasfaserkabel selten genug brechen?“. Ein aktuelles Anwendungsgebiet ist auch die Analyse von Infektionsgeschehen, siehe hierzu beispielsweise [5].

Bessere Schätzung von Eintrittswahrscheinlichkeiten extrem seltener Ereignisse

Die Abbildung 1 zeigt die Häufigkeit aufgetretener Messwerte eines realistischen fiktiven Beispiels anhand eines Histogramms (blau) und die Dichtefunktion einer Normalverteilung zur Beschreibung der Messdatenverteilung (rot). Die Modellierung der Messdaten ist an den Enden der roten Kurve – also für die Prognose extremer Ereignisse – problematisch, da sie sehr schnell gegen Null tendiert .

Abbildung 2 zeigt die Modellierung der Daten aus Abbildung 1 anhand einer Pareto-Verteilung (grün) im Vergleich mit einer Normalverteilung (rot). Die gewählte Verteilungsfunktion modelliert den Verlauf der Daten am rechten Rand besser als die Normalverteilung.

Ziel einer Extremwertanalyse ist die möglichst ak-

IABG Safety & Security Academy • Einsteinstr. 20 • 85521 Ottobrunn
Tel. +49 89 6088-4088 • academy@iabg.de • academy.iabg.de

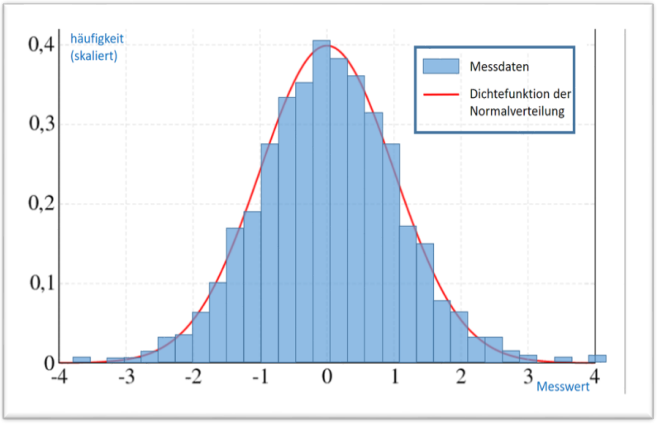


Abbildung 1: Modellierung der Häufigkeit von Messwerten durch Normalverteilung

kurate Modellierung der Wahrscheinlichkeiten für extreme Ereignisse. Daher spielt ein Großteil der „gewöhnlichen" Daten (also i.A. Daten in der Nähe eines Mittelwerts) in der Analyse keine Rolle. Es gehen ausschließlich Daten in die Analyse ein, die selbst bereits auffällige aber ggf. noch unkritische Werte enthalten. Im Kontext der Fragestellung „Wie hoch sollte ein Damm gebaut werden, um höchstens einmal in hundert Jahren überschwemmt zu werden?", könnten solche „auffällige" Daten beispielsweise Wasserstandmessungen sein, die mindestens 1m über dem Durchschnittswert der Wasserstände von 2010 liegen.

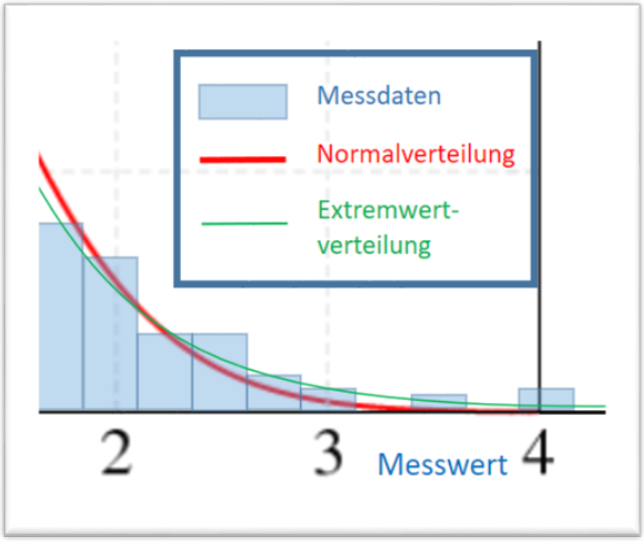


Abbildung 2: Modellierung der Häufigkeit von Messwerten mit Pareto-Verteilung

Während kritische Extremwerte normalerweise sehr selten im Rahmen von Messungen aufgezeichnet werden können, ist man in Zeiten von „Big Data" dennoch häufig in der Situation, viele für eine Extremwertanalyse geeignete „auffällige" Daten verfügbar zu haben. Jedoch müssen bei der Auswahl und Nutzung solcher Daten verschiedene methodische Herausforderungen gemeistert werden.

Herausforderungen bei der Anwendung der Extremwerttheorie

Für die Wahl der Thresholds (Schwellwerte, ab dem die Daten als „auffällig" gelten) gibt es beispielsweise kein eindeutig vorgegebenes Vorgehen. Eine der Herausforderungen bezüglich der Wahl eines Thresholds ist, dass ein höherer Threshold meistens zu einem besseren Fit der vorhandenen Daten führt, dass jedoch damit auch die Wahrscheinlichkeitsprognosen weniger verlässlich werden. Denn der höhere Schwellwert bedeutet, dass der Umfang der für die Analyse verwendeten Daten geringer ist und somit die Konfidenz der Abschätzung in Folge kleiner wird.

Eine weitere Herausforderung ist die Auswahl der Daten, wenn die vorhandenen Datensätze Abhängigkeiten voneinander aufweisen. Beispielsweise können zwei kurz nacheinander von einem Sensor durchgeführte Messungen aufgrund der zeitlichen Korrelation ähnliche Messwerte ergeben. Bei Abhängigkeiten in den Datensätzen sollten vor der

IABG Safety & Security Academy • Einsteinstr. 20 • 85521 Ottobrunn
Tel. +49 89 6088-4088 • academy@iabg.de • academy.iabg.de

Durchführung der statistischen Analyse Maßnahmen zur Entkoppelung der Datensätze umgesetzt werden, siehe z.B. „Block Maxima" Methode in [2].

Die IABG unterstützt ihre Kunden bei der Bewältigung dieser Herausforderungen und darüber hinaus, z.B. bei der Formulierung von Anforderungen an nachzuweisende Wahrscheinlichkeiten nach dem Stand der Technik und konform zu regulatorischen Vorgaben sowie bei der Generierung auswertbarer Daten.

Anwendungsbeispiel aus der Luftfahrt

Als erstes Beispiel betrachten wir Daten von landenden Flugzeugen. Kontext ist die Prognose möglicher Unfälle in Zusammenhang mit möglicherweise erforderlichen Landestrecken, die die Länge einer verfügbaren Landebahn überschreiten, siehe [1]. Die zu analysierenden Daten beschreiben die „Actual Landing Distance" (ALD), also die Landedistanz, die das Flugzeug tatsächlich am Boden zurückgelegt hat. Wir interessieren uns hierbei für besonders lange und somit sicherheitskritische Landungen. In Abbildung 3 ist daher ein Histogramm der Daten mit einer ALD von mehr als 1900 Metern abgebildet.

Dieser Randbereich der Daten wird zum Vergleich auf zwei verschiedene Arten modelliert: Die gepunktete Kurve gibt eine Beschreibung der Daten anhand einer Normalverteilung wieder. Als zweites Modell wurde eine Extremwertverteilung mithilfe der Peaks-Over-Threshold (POT) Methode geschätzt und als durchgezogene Kurve in Abbildung 3 eingetragen. Man kann erkennen, dass die Extremwertverteilung die Verteilung der Daten im betrachteten Randbereich wesentlich besser widerspiegelt als die Normalverteilung.

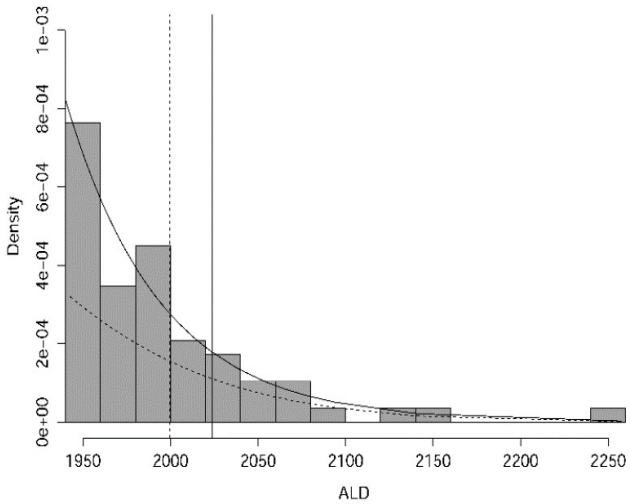


Abbildung 3: Randbereich von „Actual Landing Distance" Daten in Metern mit geschätzten

Anwendungsbeispiel aus der Automobilindustrie

Hochautomatisiertes und autonomes Fahren erfordern unter anderem eine präzise und zeitgenaue Positionsbestimmung. Die Absicherung der Korrektheit und die zuverlässige Erkennung von Fehlern in der Positionsbestimmung ist von erheblicher Bedeutung für verschiedene aktuelle und viele zukünftige Sollfunktionen für Fahrzeuge. Beispielsweise muss jederzeit eindeutig feststellbar sein, ob sich ein Fahrzeug auf oder doch neben einer Autobahn befindet. In einem unserer Kundenprojekte kam deshalb im Rahmen des „Safety of the intended Functionality" (SOTIF) Sicherheitsnachweises für ein GPS-Empfangsgerät eine Extremwertanalyse zum Einsatz. Ziel war es hierbei, in aufgezeichneten Navigationsdatensätzen die Abweichungen der gemessenen GPS-

IABG Safety & Security Academy • Einsteinstr. 20 • 85521 Ottobrunn
Tel. +49 89 6088-4088 • academy@iabg.de • academy.iabg.de

Positionen von den wahren Positionen zu analysieren und die Wahrscheinlichkeiten abzuschätzen, mit denen diese Abweichungen vorgegebene Toleranzwerte überschreiten könnten. In den umfassenden Datensätzen, die für die Analyse zur Verfügung standen, waren viele insignifikante Abweichungen der Messdaten von den wahren Positionen enthalten, aber keine Überschreitungen der genannten (sicherheitskritischen) Toleranzwerte. Die Analyse der Positionsabweichungen mithilfe der Methoden aus der Extremwerttheorie ermöglichte schließlich den Nachweis der Erfüllung der vorgegebenen Anforderungen an die maximal zulässigen Wahrscheinlichkeiten für das Überschreiten der Toleranzwerte.

Fazit

Die Methoden klassischer Datenanalysen sind häufig für die Schätzung extrem selten auftretender Ereignisse methodisch ungeeignet und können zu falschen Einschätzungen von Risiken führen mit signifikanten Auswirkungen auf Leib und Leben oder in Bezug auf finanzielle Verluste.

Die Vorhersage der Wahrscheinlichkeit selten auftretender Ereignisse hoher Signifikanz bedarf einer spezifischen Herangehensweise. Hier liefert die Extremwerttheorie erprobte Methoden für die Beantwortung von Fragestellungen, die derartige Prognosen erfordern.

Referenzen

[1] Chong Wang, Ludwig Drees, Nadine Gissibl, Lukas Höhndorf, Javensius Sembiring, Florian Holzapfel, Quantification of Incident Probabilities Using Physical and Statistical Approaches, 6th International Conference on Research in Air Transportation (ICRAT), Januar 2014, https://www.researchgate.net/publication/275887604_Quantification_of_Incident_Probabilities_Using_Physical_and_Statistical_Approaches

[2] Stuart Coles, An Introductin to Statistical Modeling of Extreme Values, Springer Verlag, 3rd edition, 2004

[3] Paul Skiba, „Wie man Ausreißer am Finanzmarkt besser bewertet“, Private Banking Magazin, 30.01.2017, https://www.private-banking-magazin.de/aus-der-naturkatastrophenforschung-wie-man-ausreisser-am-finanzmarkt-besser-bewertet/

[4] Dr. Valerie Chavez-Demoulin, ETH Zürich, „Was ist Extremwerttheorie?“, RISKNEWS, 05/2004, http://www.risknet.de/fileadmin/eLibrary/RISKNEWS-2004-05-Extremwerttheorie.pdf

[5] Ulf von Rauchhaupt, „Das dicke Ende der Häufigkeitsverteilung“, FAZ-Onlie Artikel, 02.06.2020, https://www.faz.net/aktuell/wissen/computer-mathematik/pandemie-statistik-das-dicke-ende-der-haeufigkeitsverteilung-16793371.html?GEPC=s5

Die Extremwerttheorie ist eine auf der Auswertung von bereits auffälligen Beobachtungen und Messwerten basierende Methodik aus der Statistik für die Abschätzung solcher Wahrscheinlichkeiten, die bereits in einer Fülle von Anwendungen aus verschiedenen Domänen und Branchen erfolgreich eingesetzt wird.

IABG Safety & Security Academy • Einsteinstr. 20 • 85521 Ottobrunn
Tel. +49 89 6088-4088 • academy@iabg.de • academy.iabg.de

2020-04 • © IABG

Die **Industrieanlagen-Betriebsgesellschaft mbH (IABG)** wurde 1961 auf Initiative der Bundesrepublik Deutschland als zentrale Analyse- und Testeinrichtung für die Luftfahrt und das Verteidigungsministerium gegründet. Die **IABG** wurde 1993 privatisiert und ist heute ein eigentümergeführtes europäisches Technologie-Unternehmen mit den Kernkompetenzen Analyse, Simulation & Test und Anlagenbetrieb. Der Begriff „**Sicherheit**“ bildet dabei das thematische Dach des Portfolios: Funktionssicherheit neu entwickelter High-Tech-Produkte und Verkehrssysteme (**Safety**) sowie Sicherheit von Staat, Wirtschaft und Gesellschaft (**Security**). In diesem Kontext erbringt die **IABG** technisch-wissenschaftliche Dienstleistungen für private und öffentliche Kunden in den Hauptgeschäftsfeldern: **Automotive, InfoKom, Mobilität & Energie, Umwelt & Geodaten Services, Luftfahrt, Raumfahrt, Verteidigung & Sicherheit**.

Die **IABG** beschäftigt rund 1.000 hochqualifizierte Mitarbeiter am Stammsitz in Ottobrunn sowie in kundennahen in- und ausländischen Niederlassungen. Die **IABG** ist konzernunabhängig und produktneutral und vertritt – oftmals bewertend tätig – ausschließlich die Interessen ihrer Kunden. Die Kundenstruktur setzt sich im Wesentlichen zusammen aus renommierten Industrieunternehmen (Hersteller von Luft- und Raumfahrtsystemen, von Automobilen und deren Zulieferer) sowie aus wichtigen Bundesministerien (z.B. Bundesministerium der Verteidigung) und verschiedenen Länderministerien (hier insbesondere Ministerien des Innern).

IABG
Safety & Security Academy

IABG Safety & Security Academy
Einsteinstr. 20 • 85521 Ottobrunn
Tel. +49 89 6088-4088 • academy@iabg.de
academy.iabg.de

IABG