

# **Stochastik und Statistik**

## *Vorlesung WT 3 Verteilungsmodelle*

K.Gerald van den Boogaart  
<http://www.stat.boogaart.de>

# Wie geht es weiter?

- Welche Standardmodelle gibt es?
- Welches Modell gehört zu welcher Situation?
- Wie schätzt man die Parameter?
- Wie kann man mit den Modellen weiterrechnen?

# Was brauchen Sie?

- Die Fähigkeit einfache Spezialfälle nur mit Kenntnissen dieser Vorlesung durchrechnen zu können?
- Die Fähigkeit in einer konkreten Situation zu erkennen welche Bereiche der Stochastik eine Rolle spielen?
- ... , um die richtige Formel zu benutzen?
- ... , um das richtige Buch zu finden?
- ... , um den Richtigen zu fragen?
- ... , zu wissen, was da eine Rolle spielen könnte?
- ... , zu merken dass das was mit Stochastik zu tun hat?

# Herangehensweise

- 1. Überblick über die wichtigsten Konzepte,...
- 2. Motivation und Details zu einzelnen Konzepten,...
- und 3. Beispiele,...
- Aber bitte, bitte, bitte seien Sie nicht enttäuscht, wenn die Beispiele stark vereinfacht sind.
- ..., und lesen vielleicht sogar das passende Buch.

# **Erster Überblick: Standardkonzepte der stochastischen Modellierung**

# Konzept: Verteilungen

- Mathematik: Aus Annahmen Folgerungen ableiten.
- Die Mathematik liefert:  
In Situationen, in denen gewisse Annahmen zutreffen, liegen bestimmte Verteilungen vor.
- Diese Verteilungen haben Namen.  
z.B. Binomialverteilung
- und werden durch Notationen der folgenden Form:

AbgekürzterVerteilungsname(Parameter1,Parameter2,...)

notiert, z.B.

$$Bi(n, p)$$

# Informationen zu Verteilungen

- Zu jeder der Verteilungen finden Sie in Formelsammlungen, Stochastikbüchern, Lexika, ... und im Internet (z.B. Wikipedia):
- Verteilungsfunktion:  $F_{\text{Notation}}(x)$
- Dichtefunktion  $f_{\text{Notation}}(x)$  oder  $P_{\text{Notation}}(X = k)$
- Momente:  $E[X]$ ,  $\text{var}(X)$
- Tabellen mit Quantilen
- und vieles mehr,...

# ... Informationen zu Verteilungen

- seltener: eine Erklärung wozu sie da sind.
- seltener: was die Parameter bedeuten.
- fast nie: ein Rezept zur richtigen Verteilung zu kommen.
- fast nie: was macht man mit den Verteilungen.



# Konzept: Ereignisanzahlen

- z.B.  $X =$  “Anzahl der kaputten Geräte”
- Binomial  $Bi(n, p)$ :  
wenn  $n$  Geräte unabhängig ausfallen können.
- Hypergeometrisch  $Hyp(N, M, n)$   
wenn  $n$  Geräte aus  $N$  ausgewählt werden, von denen  $M$  kaputt sind.
- Poissonverteilung  $Po(\lambda)$   
wenn sehr viele gute Geräte unabhängig voneinander kaputt gehen können.

# Konzept: Abstraktion

- z.B.  $X = \text{“Irgendwelche Anzahl”}$
- **Binomial**  $Bi(n, p)$ :  
Anzahl der eingetretenen Ereignisse,  
wenn  $n$  unabhängige Ereignisse  
mit jeweils Eintrittswahrscheinlichkeit  $p$  betrachtet  
werden.
- **Hypergeometrisch**  $Hyp(N, M, n)$   
Anzahl der besonderen Dingen,  
wenn  $n$  Dinge aus  $N$  zufällig ausgewählt werden, von  
denen  $M$  besonders sind.
- **Poissonverteilung**  $Po(\lambda)$   
Anzahl der Ereignisse,  
wenn sehr viele Zufallsexperimente unabhängig  
voneinander durchgeführt werden.

# Konzept: Anwendung

- z.B.  $X =$  “Anzahl Anforderungen an Fuhrpark”
- **Binomial**  $Bi(n, p)$ : (Firmenfuhrpark)  
Anzahl der Anforderungen,  
wenn  $n$  Personen unabhängig mit gleicher  
Wahrscheinlichkeit anfordern.
- **Hypergeometrisch**  $Hyp(N, M, n)$   
Scheint nicht zu passen.
- **Poissonverteilung**  $Po(\lambda)$  (großer Fuhrpark)  
Anzahl der Anforderungen,  
wenn sehr viele Personen unabhängig voneinander ein  
Auto anfordern.
- Also: wird die Anzahl der Teilnehmer groß, muss ich  
weniger wissen und voraussetzen.

# Konzept: Asymptotik

Bei großen Systemen und vielen Versuchen

- verliert man den Überblick,
- kann man manche Sachen nur schwer berechnen: z.B.  $1000!$ ,
- kennt man oft keine Details,
- aber es gibt spezielle Modelle für große Systeme, die diese Details gar nicht mehr brauchen.

# Beispiel: Poissonverteilung

- Die Poissonverteilung ist eine asymptotische Verteilung für die Anzahl seltener Ereignisse in großen Populationen.
- Exakt: Binomial  $Bi(n, p)$   
Anzahl der Anforderungen,  
wenn  $n$  Personen unabhängig mit gleicher Wahrscheinlichkeit anfordern.
- Asymptotisch: Poissonverteilung  $Po(\lambda)$   
Anzahl der Anforderungen,  
wenn sehr viele unabhängig voneinander ein Auto anfordern.  
 $\lambda$  ist die mittlere Anzahl von Anforderungen
- weniger Voraussetzungen (gleiche Wahrscheinlichkeit),  
weniger benötigte Information (Anzahl Personen)

# Konzept: Versuchsanzahlen

Versuchsanzahlen:

$X$  = Anzahl der Versuche bis ein Ziel erreicht ist.

z.B.

- Wie oft muß ich Lotto spielen, bis ich gewinne?
- Wie oft muß ich explorieren, bis ich eine Lagerstätte finde?
- Wieviel Bauteile muß ich testen, bis ich ein kaputtes finde?
- Wie oft kann ich ein System benutzen, bis es ausfällt.

# Versuchsanzahl vs. Ereignisanzahl

Versuchsanzahlen unterscheiden sich konzeptionell von Ereignisanzahlen:

- **Ereignisanzahl:** Es passiert etwas, wir zählen wie oft. (passive, dem Zufalls folgend)
- **Versuchsanzahl:** Wir wiederholen der Vorgang so lange bis wir zufrieden sind. (aktiv, Regeln folgend)

# Verteilungen für Versuchszahlen

$p$  = “Wahrscheinlichkeit für Erfolg in jedem Versuch”

- (Volle) Geometrische Verteilung  $Geo(p)$   
Anzahl der Versuche bis zum ersten Erfolg.
- (Reduzierte) Geometrische Verteilung  $Geo'(p)$   
Anzahl der Misserfolge bis zum ersten Erfolg.
- Negativ Binomial Verteilung  $NBi(n, p)$   
Anzahl der Versuche bis zum  $n$ -ten Erfolg.
- Negativ Hypergeometrische Verteilung  
 $NHyp(N, M, n)$   
Anzahl der Züge bis zum  $n$ -ten Erfolg aus einer Gruppe von  $N$  Teilen von denen  $M$  als Erfolg gewertet werden.



# Konzept: Lebensdauern

Lebensdauer

$X$  = Zeitspanne, die das System funktioniert.

z.B.

- Lebensdauer einer Glühbirne
- Zeitdauer bis das System das nächste mal repariert werden muß.
- Zeitdauer, für welche die Telefonzelle betriebsbereit bleibt.

# Lebensdauer und Ausfall

Zwischen der Wahrscheinlichkeit für den Ausfall in einer gewissen Betriebsperiode und der Lebensdauerverteilung besteht ein einfacher Zusammenhang:

$$\begin{aligned} P(\text{"Teil fällt im Betrieb aus"}) &= P(\text{Lebensdauer} < \text{Betriebsdauer}) \\ &= F_{\text{Lebensdauer}}(\text{Betriebsdauer}) \end{aligned}$$

oder (für die nächste Betriebsperiode):

$$\begin{aligned} &P(\text{"Teil fällt im Zeitraum } [a, e] \text{ aus"} | \text{"Es war zu } a \text{ ganz"}) \\ &= P(\text{Lebensdauer} < e | \text{Lebensdauer} \geq a) \\ &= \frac{P^X([a, e])}{P^X([a, \infty))} = \frac{F_X(e) - F_X(a)}{1 - F_X(a)} \end{aligned}$$

Die Lebensdauerverteilungen hilft also Ausfallwahrscheinlichkeiten abhängig von der Betriebsdauer festzulegen.

# Lebensdauerverteilungen

- Exponentialverteilung  $Exp(\lambda)$   
Bei externen Ausfallursachen und nicht alternden Geräten.
- Weibull-Verteilung  $Weib(\alpha, \beta)$   
 $\alpha > 1$ , für alternde Geräte, die immer anfälliger werden.  
 $\alpha < 1$ , für Geräte, denen wir immer mehr vertrauen je länger sie schon fehlerfrei arbeiten.
- Log-Normalverteilung  $LN(\mu, \sigma^2)$   
allgemeine Verteilung für Geräte mit einer ungefähr definierten Lebensdauer, z.B. durch Abnutzung von Verschleißteilen.

# Konzept: Wartezeiten

$X =$  “Wartezeit bis etwas passiert”

z.B.

- Wann kommt die nächste Anforderung?
- Wann tritt der nächste Fehler auf?
- Wie lange dauert es bis der Vorgang abgelaufen ist?

# Wartezeitverteilungen

- Exponentialverteilung  $Exp(\lambda)$   
Ankunft externer Ereignisse
- Gammaverteilungen  $Gamma(\alpha, \lambda)$   
Wartzeiten bei Abläufen mit  $\alpha$   $Exp(\lambda)$ -verteilten  
Zwischenschritten.
- Normalverteilung  $N(\mu, \sigma^2)$  Wartzeiten bei Abläufen mit  
vielen Zwischenschritten.

# Lebensdauern und Wartezeiten

Wartezeiten und Lebensdauern sind sehr ähnlich:

Wartezeit = Lebensdauer des Wartevorgangs

Lebensdauer = Wartezeit auf Defekt

Je nach dem, wie man sich den zugrundeliegenden Vorgang vorstellt, wird man also eher die eine oder andere Verteilungsgruppe wählen.

# Konzept: Zensorierung

- Von Lebensdauern lassen sich oft keine Stichproben sammeln
- . . . , weil man einfach zu lange warten müßte, bis das letzte Teil kaputt ist. (Beispiel: Waschmaschine)
- Man kann also oft nur beobachten, ob das Gerät im Beobachtungszeitraum kaputt gegangen ist und wenn ja wann.
- Für die anderen Geräte kennt man nicht den Ausfallzeitpunkt, sondern nur den Beobachtungszeitraum,
- so als ob ein Zensor uns einen Teil der Information vorenthalten würde.
- Man spricht dann von zensurierten Daten.

# Motivation der Faltung

Wir hatten bei der Gamma-Verteilung:

$$\text{Gesamtwartezeit} = \text{Wartezeit}_1 + \dots + \text{Wartezeit}_n$$

Wir hatten bei der Binomial-Verteilung  $Bi(n, p)$

$$X_{\text{ges}} = X_1 + \dots + X_n$$

Wann ist das Teil repariert:

$$\text{Servicezeit} = \text{Wartezeit auf Monteur} + \text{Reparaturzeit}$$

Hierfür gibt es kein fertiges Modell



# Konzept: Faltung

- Wir interessieren uns also oft für die Verteilung von Summen von unabhängigen Zufallsgrößen mit bekannter Verteilung.
- Wir brauchen also allgemeine Formeln für die Berechnung dieser Verteilung aus den Einzelverteilungen.
- Diesen Vorgang nennen wir “Faltung” (warum auch immer)

# Arten der Faltung

Sind  $X, Y$  unabhängig, so gilt:

- Faltungsformel für Dichten:

$$f_{X+Y}(s) = \int_{-\infty}^{\infty} f_X(s - y) f_Y(y) dy$$

- Faltungsformel für Verteilungen:

$$F_{X+Y}(s) = \int_{-\infty}^{\infty} F_X(s - y) f_Y(y) dy$$

- Faltungsformel für Momente:

$$\begin{aligned} E[X + Y] &= E[X] + E[Y] \\ \text{var}(X + Y) &= \text{var}(X) + \text{var}(Y) \end{aligned}$$

# Konzept: zentraler Grenzwertsatz

Zentraler Grenzwertsatz:

Überlagern sich viele kleine Störungen additiv, so ist das Ergebnis normalverteilt.

Auf Mathe-Sprache:

Zu einer stark mischenden Folge von Zufallsvariablen  $X_i$ , mit beidseitig beschränkter Varianz, konvergiert die Verteilung der Folge

$$M_i = \frac{R_i - E[R_i]}{\sqrt{\text{var}(R_i)}}, R_n := \sum_{i=1}^n X_i$$

gegen die Standardnormalverteilung.

# Grenzverteilungen

- Normalverteilung  $N(\mu, \sigma^2)$   
Für additiv überlagerte Störungen,  
z.B. Messfehler.
- Lognormalverteilung  $LN(\mu, \sigma^2)$   
Für multiplikativ überlagerte Störungen,  
z.B. Stoffkonzentrationen.
- Poissonverteilung  $Po(\lambda)$   
Für die kleine Anzahlen bei der Überlagerung vieler Chancen,  
z.B. Anzahl der Telefonanrufe in einem Callcenter.

# Konzept: Normalverteilung

Annähernd normalverteilt sind meist:

- Messfehler
- Fehler bei Schätzern:  $\hat{a} - a$
- Ergebnisse von “wackelnden” Prozessen
- $Po(\lambda)$  für große  $\lambda$
- $LN(\mu, \sigma)$  für kleine  $\sigma$
- $Gamma(\alpha, \beta)$  für große  $\alpha$
- ...

Deshalb wird oft eine Normalverteilung verwendet, solange man kein besseres Modell hat und sonst nichts dagegen spricht.

# Motivation: schwächstes Glied

- Additive und multiplikative Kombination vieler zufälliger Einzelteile führen also zu einfachen Verteilungen.
- Aber es gibt auch andere Kombinationen. (z.B. am Karusell)
- Die Kette hält nur so viel wie Ihr schwächstes Glied.

$$S_{\text{Kette}} = \min(S_1, \dots, S_n), \quad S_i = \text{Stärke von Glied } i$$

- Die Kette hält nur, wenn sie auch der stärksten auftretenden Belastung gewachsen ist.

$$B_{\text{max}} = \max(B_1, \dots, B_n), \quad S_i = \text{Belastungspitze von Nr. } i$$

- Ereignis  $B_{\text{max}} > S_{\text{Kette}}$  = "Kette reißt und ein Kind stirbt".

# Konzept: Verteilung von Maxima

- Kennt man die Verteilungsfkt.  $F_{X_i}(x)$  jedes einzelne  $X_i$
- und sind die  $X_i$  stochastisch unabhängig,
- so kann man leicht die Verteilung des Maximums ausrechnen:

$$\begin{aligned}F_{\max X_i}(x) &= P(\max_i X_i \leq x) \\&= P(X_1 \leq x \text{ und } \dots \text{ und } X_n \leq x) \\&= \prod_{i=1}^n P(X_i \leq x) \\&= \prod_{i=1}^n F_{X_i}(x)\end{aligned}$$

# Motivation: Extremwerttheorie

Frage: Wann kennt man das alles schon mal so genau?

Antwort: Sehr, sehr, selten kennt man die Verteilungsfunktion so genau, dass da etwas Vernünftiges herauskommt.



# Konzept: Extremwerttheorie

Satz:

Maxima (und Minima) vieler unabhängiger identisch verteilter stetiger Zufallsvariablen können asymptotisch nur 4 verschiedene Typen von Verteilungen haben.

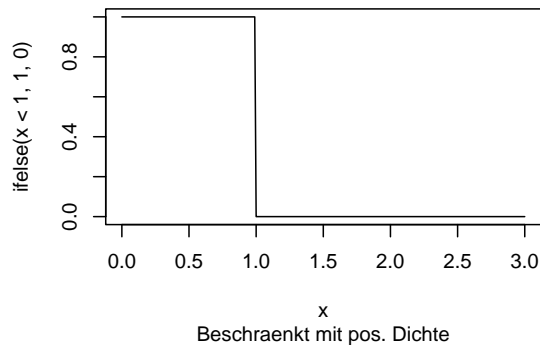
- Type 0: Einpunktverteilung  $\delta_{x_0}$
- Type I: Gumbel-Verteilung  $Gumble(\mu, \beta)$
- Type II: Fréchet-Verteilung  $Frechet(\mu, \sigma, \alpha)$
- Type III: Reverse Weibullverteilung  $RWei(\mu, \sigma, \alpha)$

und welcher es ist, hängt vom Verteilungstyp der einzelnen Zufallsvariable ab.

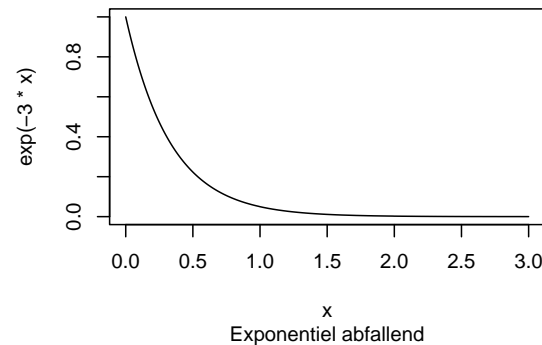
# Konzept: Anziehungsbereich

... und zwar von der Form der Dichte in Richtung auf extreme Werte hin.

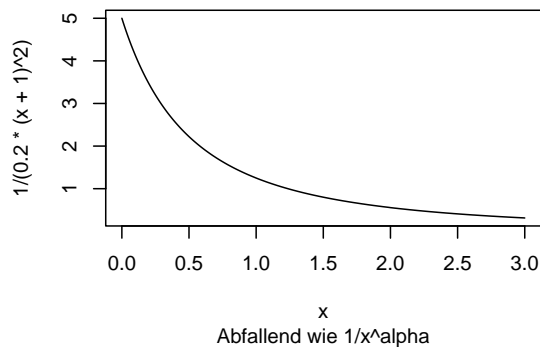
Type 0



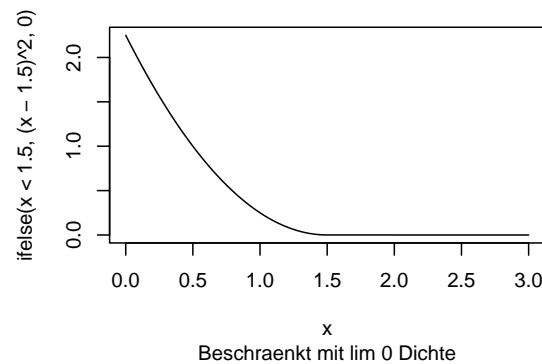
Type I



Type II



Type III



# Beispiel: Sicherheitsgrenze I

- Frage: Wieviel muß die Karusellkette aushalten?
- Die mögliche Maximalbelastung ist vermutlich nicht prinzipiell beschränkt und ....
- Wir modellieren sie also als Fréchetverteilt.
- Wir bauen in eine Kette einen Belastungsschreiber ein und messen jeweils 100mal die Maximalbelastung für 10 Minuten Karusellbetrieb (mit Dummy? oder mit Kind?).
- Wir schätzen aus dem Datensatz die Parameter der Fréchet-Verteilung.
- ...

# Konzept: Reskalierung

- Problem: Wir brauchen ja die Maximalbelastung für 30 Betriebsjahre.
- Idee: Es gibt eine Möglichkeit die Parameter auf vielfache Wiederholungsanzahlen (also längere Zeiträume) umzurechnen.
- Problem: Das ist reine Theorie, also bitte betrachten sie das bestenfalls als gewagte Abschätzung. Aber etwas Besseres haben wir nicht.

# Beispiel: Sicherheitsgrenze II

- ..., und berechnen aus den geschätzten Parametern die neuen Parameter für 30 Jahre,
- ... berechnen das 0.999%-Quantil
- ... und erhalten so eine geforderte Bruchfestigkeitsgrenze, für die wir 1 : 999 sicher sind, dass die Kette nicht durch Überlastung reißt.
- Ist das vernünftig? (1:999?)
- Könnte man nicht einfach die mittlere Belastung mit 1,5 multiplizieren????

# Beispiel: Kettenfestigkeit II

- Frage: Wieviel hält eine Karusellkette?
- Es geht um ein Minimum, das nicht unter 0 fallen kann,
- Theorie  $\Rightarrow$  Es kommen (also Type 0 oder) Type III in Frage.
- Da Ketten offenbar verschieden gut halten, modellieren wir  $X = -S$  also als Revers-Weibullverteilt.
- Aus Bruchversuchen mit 100 Ketten schätzen wir die Parameter dieser Verteilung.
- Aus den Parametern rechnen wir die Quantile für die Bruchfestigkeit aus.
- Bei längeren und kürzeren Ketten wenden wir Reskalierung an.

# Beispiel: Andere Konzepte einbringen

- Interessierendes Ereignis:  $D = S_{Kette} - B_{max} < 0$
- Vermutlich sind beide stochastisch unabhängig.
- Eine Faltung liefert die Verteilungsfunktion  $F_D$  von  $D$ .
- Die Auswertung  $F_D(0) = P(D < 0) = p$  liefert die W'keit für den Bruch einer Kette.
- Worst case:  $P(\text{eine Kette reißt}) \leq p * \text{“Anzahl Ketten”}$
- Und damit haben wir eine Abschätzung für die Sicherheit eines Aspekts unserer Karusellbaureihe.

# Problem: Naturphänomene

- Wir wollen in Tokio einen erdbebenstabilen Wolkenkratzer bauen.
- Frage: Welche Erdbebenstärken sollte er aushalten?
- Problem: Wir können nicht viele Versuche machen, wie stark die Erdbeben in der Gegend denn wohl so sind.
- Idee 1: Bestimme die maximale Erdbebenstärke in Dekadenblöcken und rechne die Verteilung durch “Reskalierung” auf die nächsten 200 Jahre hoch.
- Problem: Es gibt nicht viele Dekadenblöcke und kürzere Zeiträume folgen wahrscheinlich noch nicht unseren Gesetzen für das Maximum vieler Erdbeben.



# Konzept: Peak over threshold

- Idee: Es kommt doch für die Verteilung des Maximalwerts nur auf die Verteilung der großen Werte an, also werten wir doch diese statistisch aus.
- Wir benutzen also z.B. als Datensatz die Information, über die Stärke aller starken Erdbeben bei Tokio.
- Dazu legen wir eine Grenze fest, ab wo wir ein Erdbeben als stark betrachten. Diese Grenze heißt “threshold” (engl. Schwelle)
- Diese Beobachtungen über der Grenze bezeichnet man als “Peaks over threshold”.
- An diesen Datensatz müssen wir eine Verteilung anpassen.

# Generalisierte Paretoverteilung

**Begründete Beobachtung:**

Für einen ausreichend hohen “threshold”  $u$ , läßt sich an die Verteilung der “peaks over threshold” — bei praktisch jeder “natürlichen” Verteilung — eine generalisierte Paretoverteilung  $GPareto(u, \sigma, \xi)$  anpassen.

**Nutzen:** Aus den Parametern dieser Verteilung, zusammen mit der Häufigkeit des Auftretens von solchen Peaks kann man die Parameter der zugehörigen Extremwertverteilung bestimmen.

# Einschub

Wie hoch wäre eigentlich die Wahrscheinlichkeit, dass der Wolkenkratzer in den nächsten 100 Jahre von einem Erdbeben zerstört wird, wenn er so gebaut wird, dass der genau dem stärksten Erdbeben der letzten 100 Jahre standhalten könnte.

# Konzept: Schätzung

- Die theoretischen Überlegungen liefern eine Verteilungsmodell,
- . . . , sie liefern aber meist höchstens einen Teil der Parameter.
- Wir schätzen also die Parameter durch Schätzer.
- Schätzer geben aber nur ungefähre Größen.
- Frage: Wie rechnet man mit ungenauen Werten weiter?

# Beispiele: Weiterrechnen

- Wir haben den Parameter der Extremwertverteilung der Erdbebenstärken ungefähr geschätzt.
- Daraus wollen wir das 0.999-Quantil ausrechnen,
- ..., weil das die Stärke des Erdbeben ist, auf das wir uns einrichten müssen, wenn unser Wolkenkratzer das stärkste Erdbeben der nächsten 100 Jahre, mit einer Wahrscheinlichkeit von 0.999 überstehen soll.

# Konzept: Transformation

**Abstrakte Sichtweise:**

Bei jeder Art von Weiterrechnen kann die zu berechnende Größe durch eine Funktion beschrieben werden:

egal, wie schwierig es vielleicht sein mag die Funktion aufzuschreiben.

# Weiterrechnen mit Unsicherheit

Es gibt vier Konzepte zum Weiterrechnen mit Unsicherheit

- Vertrauensintervalle (nicht immer vorhanden/gut)
- Transformation der Verteilung (exakt, aber schwierig)
- Fehlerrechnung (einfacher, aber ungenau)
- Simulation (rechenaufwendig)

# Konzept: Vertrauensbereich

Statt nur eines Schätzers  $\hat{\lambda}$  kann man manchmal auch einen Bereich angeben, der den wahren Parameterwert mit einer hohen Wahrscheinlichkeit enthält: z.B. ein Intervall

$$[u_{\lambda}(\text{Daten}), o_{\lambda}(\text{Daten})]$$

wobei die Funktionen  $u_{\lambda}(\dots)$  und  $o_{\lambda}(\dots)$  so geschickt gemacht sind, dass für den wahren Parameter  $\lambda$  gilt:

$$P(\lambda \in [u_{\lambda}(\text{Daten}), o_{\lambda}(\text{Daten})]) =$$

Diese zufälligen Intervalle heißen  $1-\alpha$ -Vertrauensintervalle oder  $1-\alpha$ -Konfidenzintervalle für  $\lambda$ .



# Einfache Vertrauensbereiche

Für erwartungstreue Schätzer  $\hat{\lambda}$  für einen Parameter  $\lambda$  mit

- ungefähr normalverteiltem Schätzfehler  $\hat{\lambda} - \lambda$
- mit bekannter Varianz  $\text{var}(\hat{\lambda} - \lambda)$

ist

$$\left[ \hat{\lambda} - 2\sqrt{\text{var}(\hat{\lambda})}, \hat{\lambda} + 2\sqrt{\text{var}(\hat{\lambda})} \right]$$

ein 0.95-Vertrauensintervall für  $\lambda$

# Übertragener Vertrauensbereich

Bei Transformationen lassen sich Vertrauensbereiche übertragen:

Es gilt

$$\begin{aligned} P(G \in V_g) &= P(g(\lambda) \in \{g(x) : x \in [u_\lambda, o_\lambda]\}) \\ &\geq P(\lambda \in [u_\lambda, o_\lambda]) \\ &\geq 1 - \alpha \end{aligned}$$

mit dem neuen Vertrauensbereich

$$V_G := \{g(x) : x \in [u_\lambda, o_\lambda]\}$$

für  $G = g(\lambda)$ .

# Konzept: Verteilungstransformation

Kennt man die Verteilung einer Zufallsgröße (z.B. Bruchfestigkeit  $B$  des Hebelmaterials), so kann man daraus die Verteilung transformierter Größen (z.B.  $G = g(B)$  die Belastbarkeit des Hebels) ausrechnen.  
z.B. für streng monotone  $g$ :

$$F_G(x) = F_B(g^{-1}(x))$$

da  $P(G < x) = P(g(B) < x) = P(B < g^{-1}(x))$   
und durch Ableiten erhalten wir:

$$f_G(x) = F'_B(g^{-1}(x))(g^{-1})'(x) = \frac{f_B(g^{-1}(x))}{g'(g^{-1}(x))}$$

# Konzept: Simulation

- Wenn ich die Zufallszahlen  $X_1, \dots, X_n$  mit der Verteilung von  $X$  erzeugen kann, ...
- ..., dann kann ich auch Zufallszahlen  $G_1, \dots, G_n$  mit der Verteilung von  $G = g(X)$  erzeugen, ...
- ... indem ich einfach  $G_i = g(X_i)$  setze.
- Und diese Verteilung kann ich dann wieder statistisch analysieren.

# Motivation Fehlerrechnung

- Motivation: Bei Normalverteilung genügt für die Bestimmung eines Vorhersagebereichs die Kenntnis von Erwartungswert und Varianz.
- Oft kennt man Erwartungswert und Varianz von  $X$ .
- Wenn man daraus Erwartungswert und Varianz von  $G = g(X)$  bestimmen könnte, hätte man einen Bereich in dem der Wert liegen sollte.
- Das führte C.F. Gauss ca. 1820 zum wahrscheinlich älteste stochastische Verfahren des Planeten: der Fehlerrechnung

# Idee der Fehlerrechnung

- Aufgabe: Bestimme Erwartungswert und Varianz von  $G = g(X)$  aus Erwartungswert und Varianz von  $X$ .
- Ist  $G = g(X)$  wie eine Geradengleichung, z.B.

$$G = a + bX$$

so gilt einfach:

$$E[G] = E[a + bX] = a + bE[X]$$

und

$$\text{var}(G) = \text{var}(a + bX) = \text{var}(bX) = b^2 \text{var}(X)$$

# Ist $g(X)$ keine Gerade,...

dann hoffe, dass es fast eine ist:

# Fehlerrechnung (univariat)

- geg:  $\mu \approx E[X]$ ,  $\sigma^2 \approx \text{var}(X)$
- Geradenapproximation  
 $g(x) \approx g(\mu) + g'(\mu)(x - \mu)$
- $E[g(X)] \approx g(\mu)$
- $\text{var}(g(X)) \approx g'(\mu)^2 \sigma^2$
- $\text{sd}(g(X)) \approx g'(\mu) \sigma$



# ..., mit n unabhängigen Eingaben

- geg:  $\mu_i \approx E[X_i]$ ,  $\sigma_i^2 \approx \text{var}(X_i)$

- mehrdimensionale lineare Approximation

$$g(x_1, \dots, x_n) \approx g(\mu_1, \dots, \mu_n) + \sum_{i=1}^n \frac{\partial g(x_1, \dots, x_n)}{\partial x_i} (x_i - \mu_i)$$

- $E[g(X_1, \dots, X_n)] \approx g(\mu_1, \dots, \mu_n)$

- $\text{var}(g(X_1, \dots, X_n)) \approx \sum_{i=1}^n \left( \frac{\partial g(x_1, \dots, x_n)}{\partial x_i} \right)^2 \sigma_i^2$

- $\text{sd}(g(X)) \approx \sqrt{\sum_{i=1}^n \left( \frac{\partial g(x_1, \dots, x_n)}{\partial x_i} \right)^2 \sigma_i^2}$

# Zusammenfassung

- Eine Frage des Überblicks:  
16 Verteilungen in 7 Anwendungsgruppen,  
13 theoretische Konzepte,  
3 allg. Rechenformeln
- Um es verwenden zu können, müssen wir es nochmal detaillierter ansehen.
- Um sich das merken zu können, müssen wir es nochmal detaillierter ansehen.