



Gesundheitliche Bewertung unter Unsicherheit

Was leisten statistische Modelle?

Matthias Greiner

Bundesinstitut für Risikobewertung (BfR) und Tierärztliche Hochschule Hannover (TiHo)

DAGStat 2013: Was bedroht unser Leben wirklich? Statistische Bewertung von Gesundheitsrisiken, Berlin, 19.4.2013

Gliederung

Risikomodellierung

Kein Modell ohne Szenario

Parameter – ein Wert steht für viele

Ergebnisse und Erkenntnisgewinn

Zusammenfassung

Literatur

Risikomodellierung

Risikobegriff ■ Von der Fragestellung zum Modell ■ Beispiel
■ Modellierungsansätze ■ Unsicherheiten



Risiko = Funktion der Eintrittswahrscheinlichkeit (p) und den Konsequenzen (c) eines unerwünschten Ereignisses (h) in einem definierten Szenario



Hintergrund

Escherichia (E.) coli O157:H7 ist ein mikrobieller Hazard in Lebensmitteln wie Gehacktem. Gute Küchenhygiene und insbesondere vollständiges Durcherhitzen reduziert deutlich das Erkrankungsrisiko.

Fragestellung¹

Wie groß ist das Erkrankungsrisiko für **Hazard** → *E. coli* O157:H7 O157:H7, in Hinblick auf **Endpunkt** → die Anzahl von Erkrankungsfällen pro Jahr für ein Kind unter 6 Jahren auf Grund des **Szenario** → Verzehr von Gerichten, die aus Rinderhack hergestellt werden?

¹Beispielmodell modifiziert nach Pouillot et al. (2009)

Fragestellung und Modell

Gegeben

- ▶ Fragestellung: Hazard → Expositionsszenario → Endpunkt
- ▶ Daten und Informationen zu verschiedenen Abschnitten des Szenarios
- ▶ Plausibilität des Szenarios (Kausalkette)

Gesucht

Schätzung des definierten Endpunkts (Risiko-Prädiktion)



Methoden

Szenariomodell, deterministisches oder probabilistisches Modell, Simulation

Szenariomodell (Modellkonzept)



Vorkommen Konzentration c_0 des Hazards in rohem Hack vor Prozessierung

Prozess Konzentration am Ende der Prozessierung (Tiefrieren) über die Zeit t mit exponentieller Wachstumsrate k und Wachstumsfunktion $c_t = c_0 e^{kt}$

Exposition Verzehrsmenge g Hack pro Verzehrportion für Kinder
Anzahl s von Gerichten pro Kind und Jahr (Monat)
Anzahl n von infektiösen *E. coli* O157:H7 per Gericht
Fraktion i inaktiviert durch Erhitzen

Dosis-Wirkung Wahrscheinlichkeit P der Erkrankung bei einmaligem Verzehr, abhängig vom Parameter r einer binomialen Dosis-Wirkungs-Funktion

Konsequenz Jährliche Anzahl R der Erkrankung bedingt durch *E. coli* O157:H7 im Rinderhack bei einem Kind unter 6 Jahren

Modellierungsansätze

Deterministisches Modell

- ▶ Verwendung von Einzelwerten (z.B. Mittelwerte oder Perzentile) als Eingabegrößen

⇒ Ergebnis der Risikoabschätzung ist ein einzelner Wert

Probabilistisches Modell

- ▶ Verwendung von Wahrscheinlichkeitsverteilungen zur Abbildung der Informationen über Modellparameter, einschließlich Variabilität und Unsicherheit

⇒ Ergebnis der Risikoschätzung ist eine Verteilung

Unsicherheiten

Auf verschiedenen Ebenen — aus verschiedenen Quellen

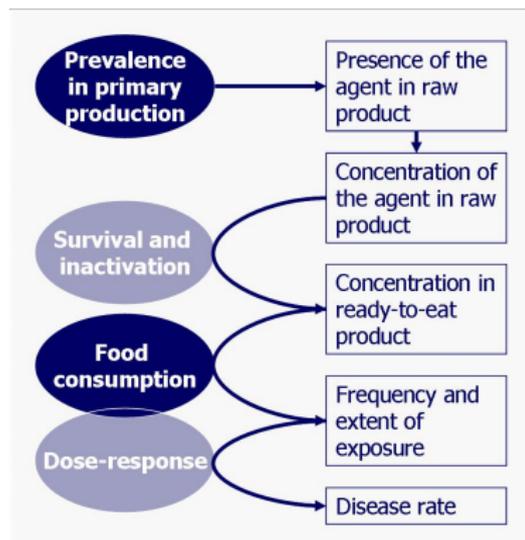
- ▶ Fragestellung
- ▶ Szenario
- ▶ Gesamtmodell
- ▶ Parameter des Modells
- ▶ Implementierung des Modells
- ▶ Ergebnisse

Kein Modell ohne Szenario

Beispiel ■ Quellen der Unsicherheit ■ Modellierungsansätze



Beispiel: Mikrobiologische Risikoanalyse



Freisetzung/Eintrag, Verbreitung, An- und Abreicherung, Kontakt/Exposition, Maßnahmen

Aspekte der Szenario-Unsicherheit²

- ▶ Freisetzungsquellen und -pfade
- ▶ Populationen
- ▶ Mikro-Umgebung
- ▶ Zeit
- ▶ Ort
- ▶ Maßnahmen
- ▶ Verhaltensweisen
- ▶ Expositionspfade

²Modifiziert nach WHO (2008, 2009).

Quellen der Unsicherheit

- ▶ Fehlende Kenntnisse über das Szenario
- ▶ Kausalzusammenhänge und Einflussfaktoren empirisch nicht belegt
- ▶ Externe Validität für den Anwendungsfall



Methoden

- ▶ Mit Daten: Schätzung von Szenarioparametern
- ▶ Ohne Daten: Sensitivitätsanalyse (alternative Szenarien)

Refsgaard et al. (2006)

Parameter – ein Wert steht für viele

- Variabilität und Unsicherheit ■ Beispiel
- Modellierungs-Ansätze ■ Probabilistik und Bayes

Warum keine festen (deterministischen) Parameter?

Abbildung von Variabilität

- ▶ Real existierende Variabilität eines Systems
- ⇒ Variabilität kann geschätzt und beschrieben, nicht aber reduziert werden

Abbildung von Unsicherheit

- ▶ Fehlende empirische Datengrundlage
 - ▶ Zufällige und systematische Fehler bei der Datenerhebung und Parameterschätzung
 - ▶ Qualitätseigenschaften der Daten
 - ▶ Kompensation von Szenariounsicherheit
- ⇒ Unsicherheiten können theoretisch durch weitere Studien reduziert werden



Modellparameter (1/3)

- $c_0 \sim N(10, 2)$ Mittlere Konzentration von *E. coli* O157:H7 in rohem Hack ist nicht sicher bekannt. Mikrobiologen nehmen an, dass die Unsicherheit über die mittlere Belastung durch eine durch eine Normalverteilung mit den Parametern $m = 10$ und $s = 2$ ausgedrückt werden kann.
- $k = 0$ Spezifische Wachstumsrate k von *E. coli* O157:H7 nach dem einfachen Modell $c_t = c_0 e^{kt}$, wobei c_t und c_0 die Anzahl von Bakterien zur Zeit t und $t = 0$ sind. Es wird angenommen, dass während der Prozessierung und Lagerung kein Wachstum auftritt da das Fleisch gefroren ist.

Beispiel Lebensmittelsicherheit



Modellparameter (2/3)

$t \sim \text{betapert}(6, 24, 12)$ Zeit unter Prozessierung und Lagerung.
Minimum, maximum und Most-likely Werte
basieren auf Expertenmeinungen und
beschreiben die Variabilität.

$i = \text{discrete}(X(1, 1/5, 1/50), p(0.027, 0.373, 0.600))$ Anteile
von überlebenden *E. coli* O157:H7 Zellen während
der Zubereitung. Keine Inaktivierung ($i = 1$) für
“roh” und $i = 1/5$ oder $i = 1/50$ überleben für
“medium” und “gut durchgebraten”.

Beispiel Lebensmittelsicherheit

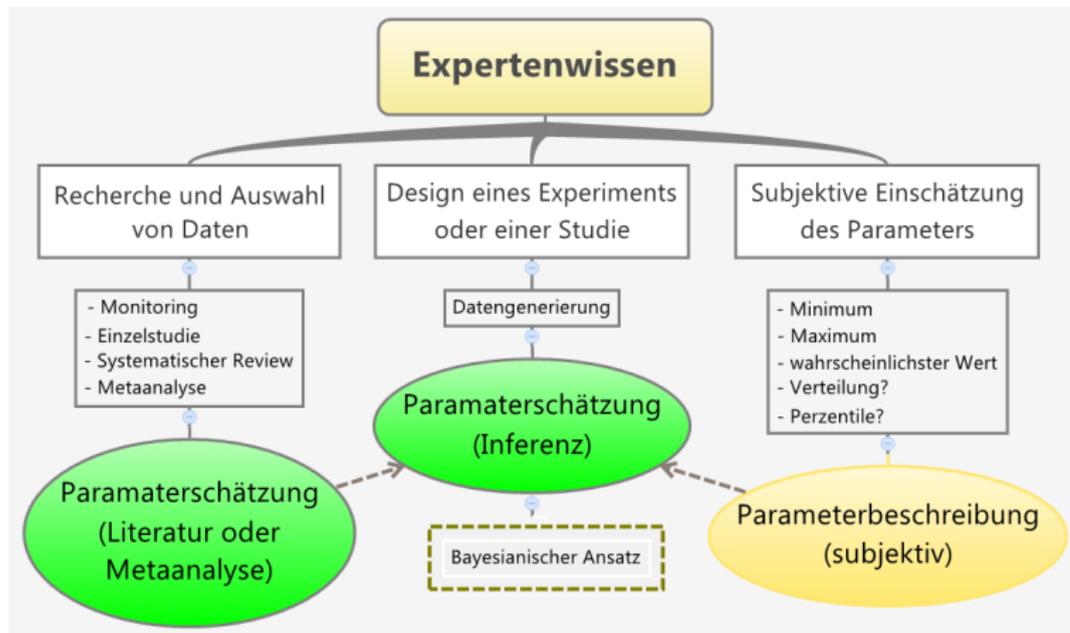


Modellparameter (3/3)

- $g \sim \text{gamma}(3.93, 0.0806)$ Verzehrsmenge (g) Rinderhack pro Verzehrportion für Kinder unter 6 Jahren, geschätzt aus Verzehrdaten.
- $s \sim \text{lognormal}(0.2, 0.5)$ Anzahl von Gerichten pro Monat und Kind unter 6 Jahren geschätzt aus Verzehrdaten.
- $r \sim U(0.0005, 0.0015)$ Experteneinschätzung des Parameters der Dosis-Wirkungs-Funktion. Der Minimum- und Maximum-Wert beschreibt die Unsicherheit.

Modellierungsansätze für Parameter

Rahmenbedingungen



Zentrale Rolle der Wahrscheinlichkeitsdichte (pdf)

In der statistischen Inferenz und in der Risikobewertung (RB)



Bayes-Statistik und RB

Interpretation der Populationsparameter als Zufallsvariablen.
Modellierung der verfügbaren Information durch W'keitsdichten.

Beta-Verteilung

Anwendung der Beta-Verteilung in der RB

- ▶ Repräsentation der Information über einen Anteilswert (Proportion)
- ▶ Repräsentation der Information über den Binomial-Parameter p unter Heterogenität

Notation

- ▶ $p \sim \text{Beta}(a, b) = f(p) = \frac{1}{B(a, b)} p^{a-1} (1 - p)^{b-1}$,
mit der Beta-Funktion $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$
- ▶ $E(p) = \frac{a}{a+b}$, $\text{var}(p) = \frac{ab}{(a+b+1)(a+b)^2}$

Beta-Verteilung

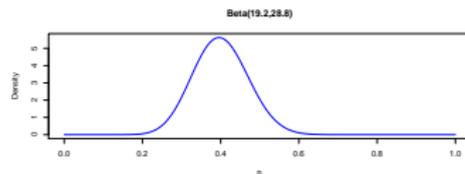
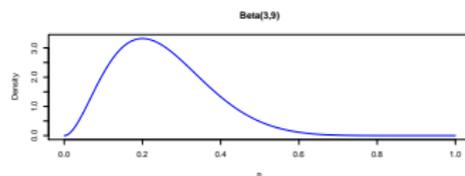
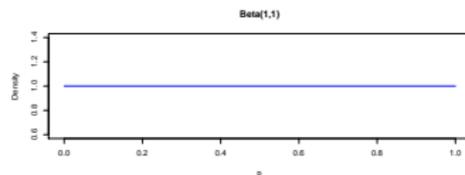
Beta-Verteilungsfunktion

- ▶ Kontinuierlich für die Ereignisse $p \in [0, 1]$

Beispiel

Fall	Parameter
1	$a = 1, b = 1$
2	$a = 3, b = 9$
3	$a = 19.2, b = 28.8$

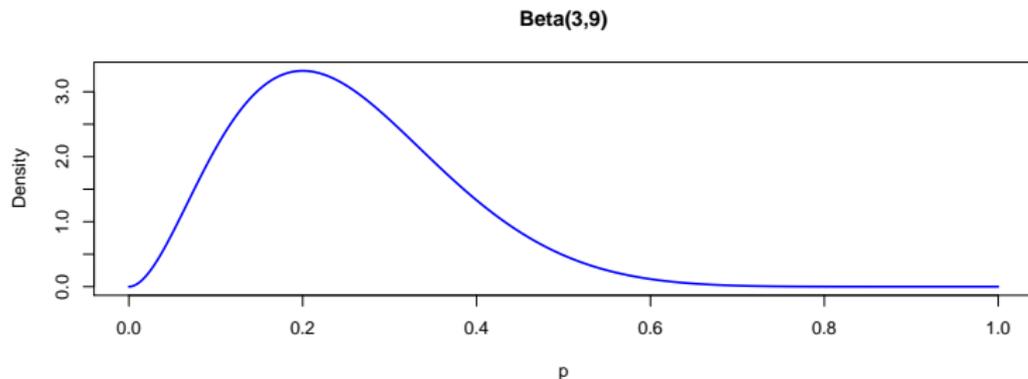
Visualisierung



Beta-Verteilung

Parameterisierung durch Studiendaten

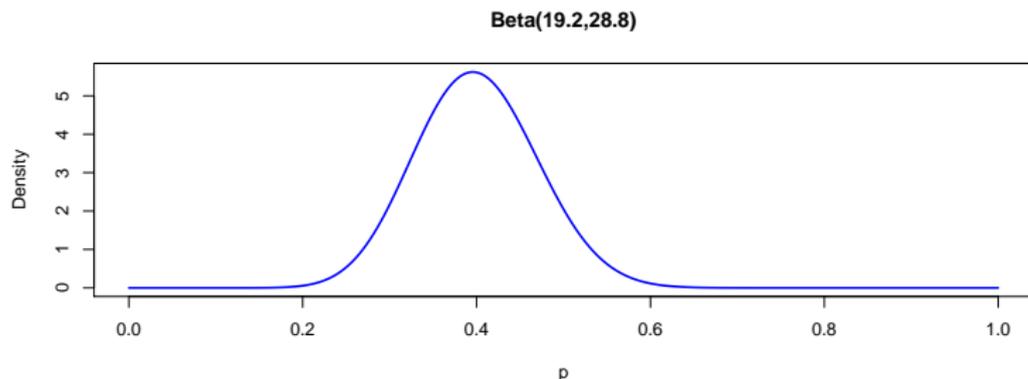
- ▶ Gegeben: $k = 2$ von $n = 10$ Tieren haben das Merkmal
- ▶ Parameter: $a = k + 1$, $b = n - k + 1$



Beta-Verteilung

Parameterisierung durch Studiendaten

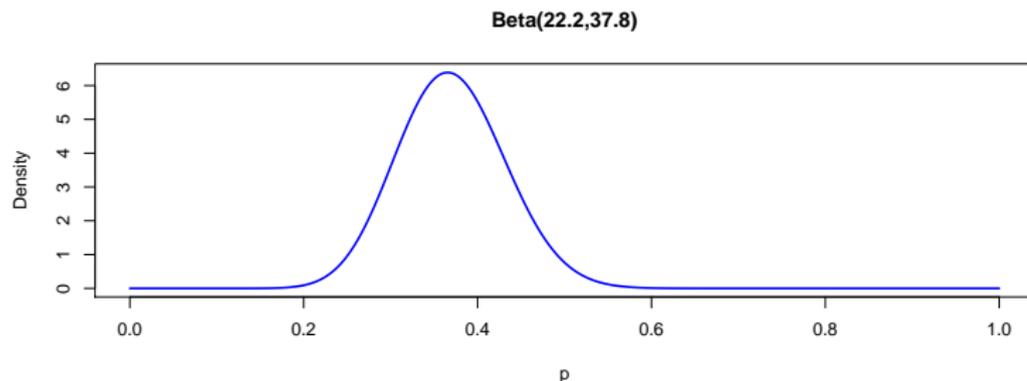
- ▶ Gegeben: $P = 0.4$ und Standardfehler $SE(P) = 0.070$
- ▶ Parameter: $a = 19.2$, $b = 28.8$ (durch numerische Methoden bestimmt)



Beta-Verteilung

Parameterisierung durch multiple Studiendaten

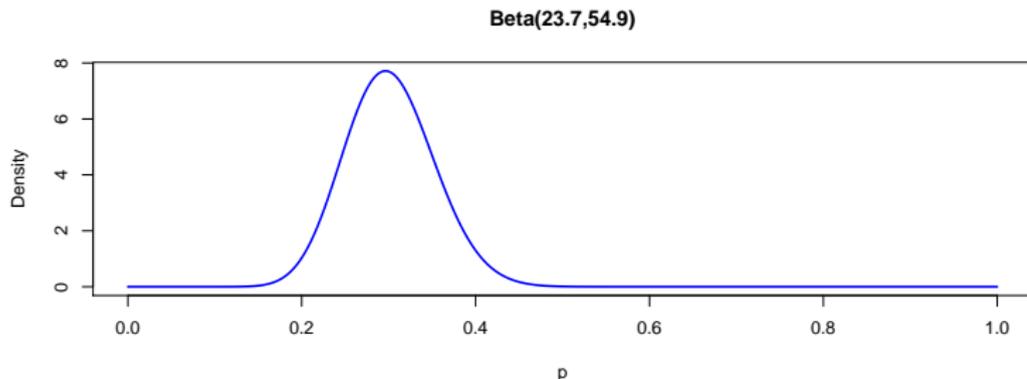
- ▶ Gegeben: Informationen der letzten beiden Folien
- ▶ Parameter: $a = 22.2$, $b = 37.8$ (additive Eigenschaft der Beta-Verteilung)



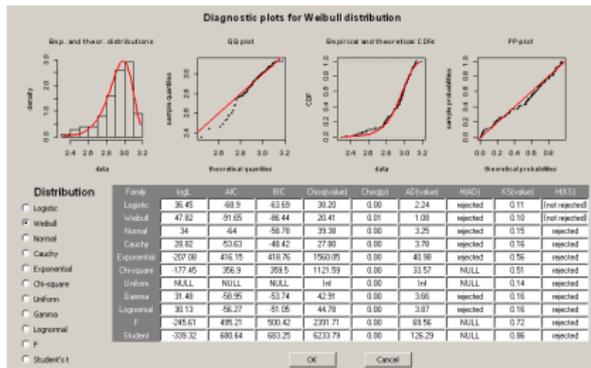
Beta-Verteilung

Parameterisierung durch Expertenmeinung

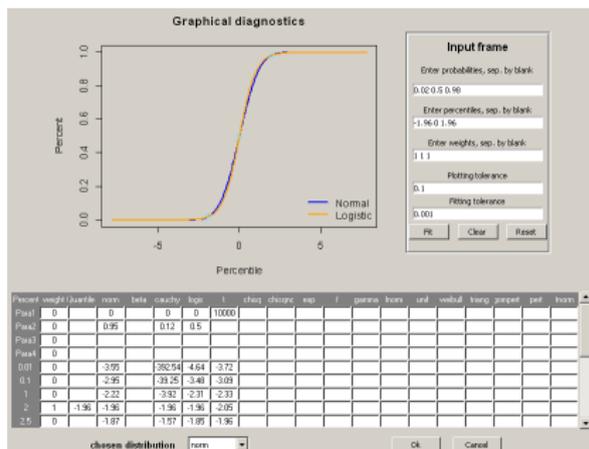
- ▶ Gegeben: Wahrscheinlichster Wert ist 0.3, zu 95% sicher, dass der Wert zwischen 0.2 und 0.4 liegt
- ▶ Parameter: $a = 23.7$, $b = 54.9$ (numerische Methoden)



Daten



Quantile



Methoden

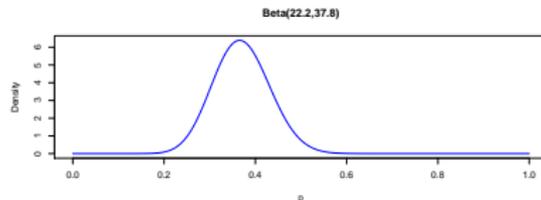
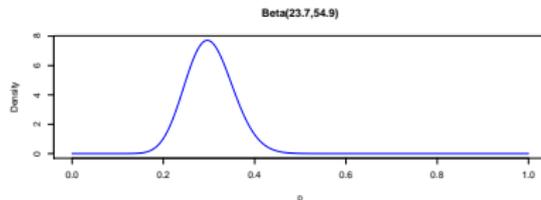
R-Paket `riskDistributions` zur Anpassung von Verteilungen an Daten und Quantile

Belgorodski et al. (2012b)

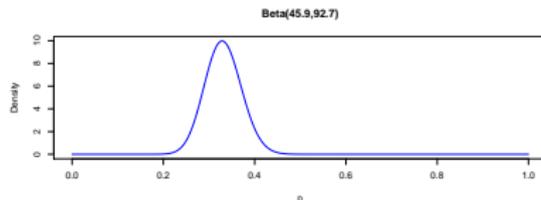
Beta-Verteilung

Parameterisierung durch Bayes-Ansatz

- ▶ Gegeben: Expertenmeinung $\text{Beta}(23.7, 54.9)$ und Studiendaten $\text{Beta}(22.2, 37.8)$ ³



- ▶ ergibt Parameter: $a = 45.9$, $b = 92.7$ (Bayes Posterior)



³Die Priori-Verteilung ist durch die Experten definiert; die Beta-Verteilung für die Studiendaten hat die gleiche Form wie die Likelihoodfunktion.

Modellierung von Parametern, die Unsicherheit ausdrücken

Full model

Verteilung, wie aus den “unsicheren” Daten oder Expertenmeinungen abgeleitet

Relaxed model

Uniforme Verteilung mit den Grenzen eines absoluten, plausiblen Minimums und Maximums

Probabilistik = Rechnen mit Verteilungen

Mikrobiologisches Beispiel

$c_t = c_0 \exp(kt)$ Die Anzahl von *E. coli* O157:H7 nach Zeit t .

$n \sim \text{Pois}(c_t i g)$ Anzahl von infektiösen *E. coli* O157:H7 in einem verzehrsfertigen Gericht.

$P = 1 - (1 - r)^n$ Wahrscheinlichkeit der Erkrankung in einem Kind unter 6 Jahren nach Verzehr eines *E. coli* O157:H7 kontaminierten Gerichts aus Rinderhack.

Anzahl der Fälle pro Jahr (R) wird aus P abgeleitet.

Die Parameter t , i und g sind Verteilungen. Die Funktionen P und R ebenso.

Typische Parameter in der RB

- ▶ Prävalenz mit diagnostischer Missklassifikation
- ▶ Prävalenz mit gepoolten Daten
- ▶ Prävalenzschätzung mit Nachweisgrenze für Zähldaten (Null-Inflationsmodell)



Methoden

R-Paket `riskBayes`. Mit Bayes-Modellen können Korrelationen zwischen Parametern und ein Informations-Feedback abgebildet werden.

Belgorodski et al. (2012a); Greiner et al. (2013)

Prototyp einer Open-Source Software für quantitative Risikobewertung

1. Reproduzierbarkeit, Versionskontrolle, Portabilität
2. Standardisierte Modelldokumentation
3. Implementierung identisch mit Modellbeschreibung
4. Umfangreiche Funktionalität
5. Freie Verfügbarkeit

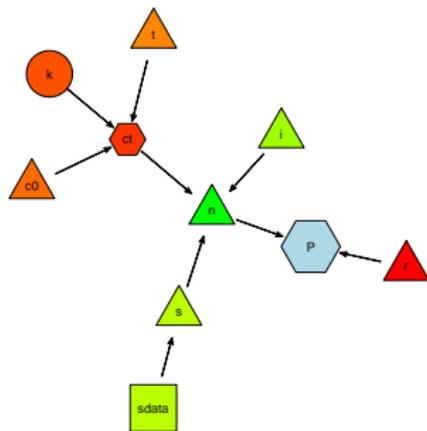
$$\text{R} + \text{Risiko} = \text{rrisk}$$

Ergebnisse und Erkenntnisgewinn

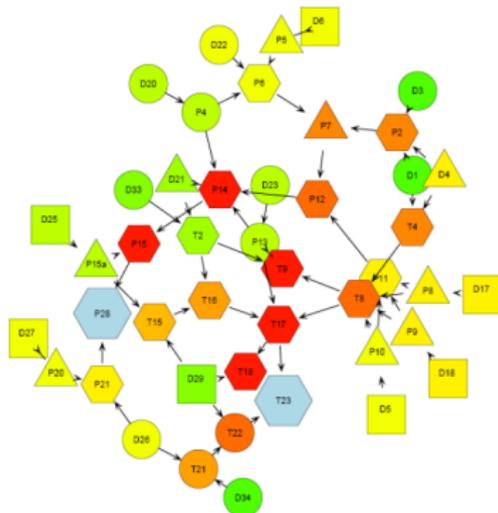
Visualisierungen ■ 2D-Simulation ■ Sensitivität ■ Bewertung unter Unsicherheit

Strukturelle Abhängigkeit der Zielgröße von den Parameter?

Modell als Graph



rrisk Demo model



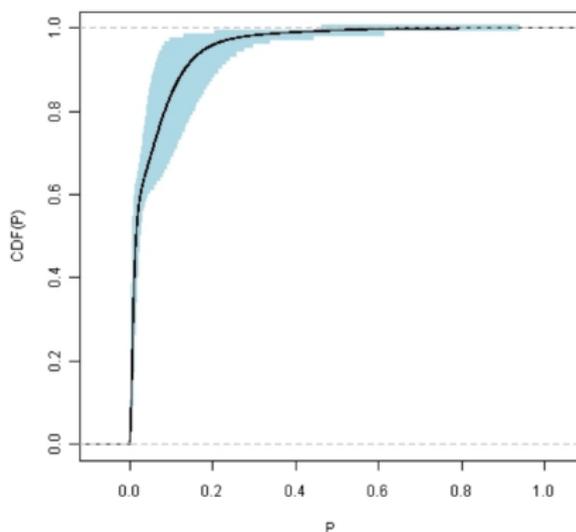
EFSA (2011)

Effekte der Variabilität und Unsicherheit auf die Zielgröße?



2D-Simulation

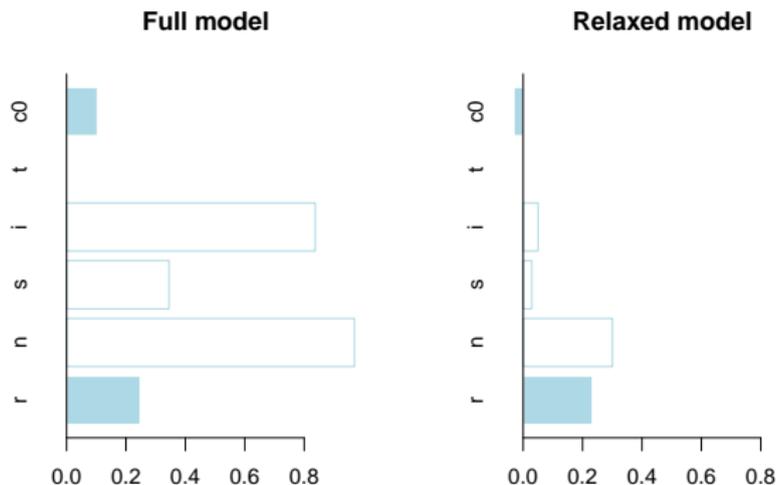
- ▶ Quantile der der Zielgröße unter Variabilität: Ablesen an kumulativen Verteilung (Linie)
- ▶ Effekt der der Unsicherheit: Ablesen am Konfidenzband



Effekte der Parameter auf die Zielgröße?



Tornadochart

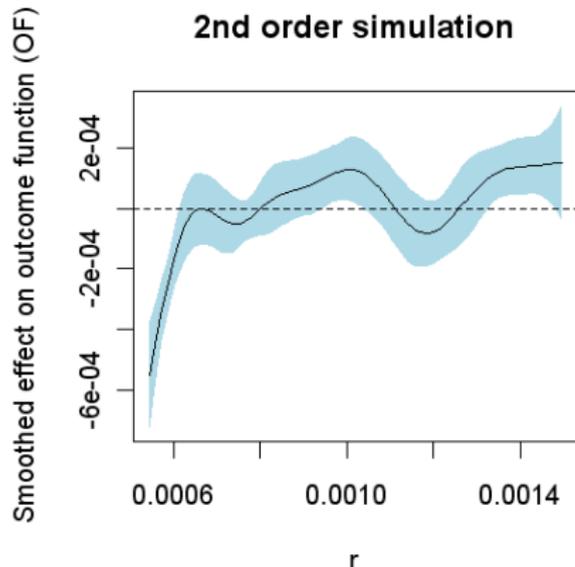


Correlation coefficients using rank transformed data

Effekte eines Parameters auf die Zielgröße?

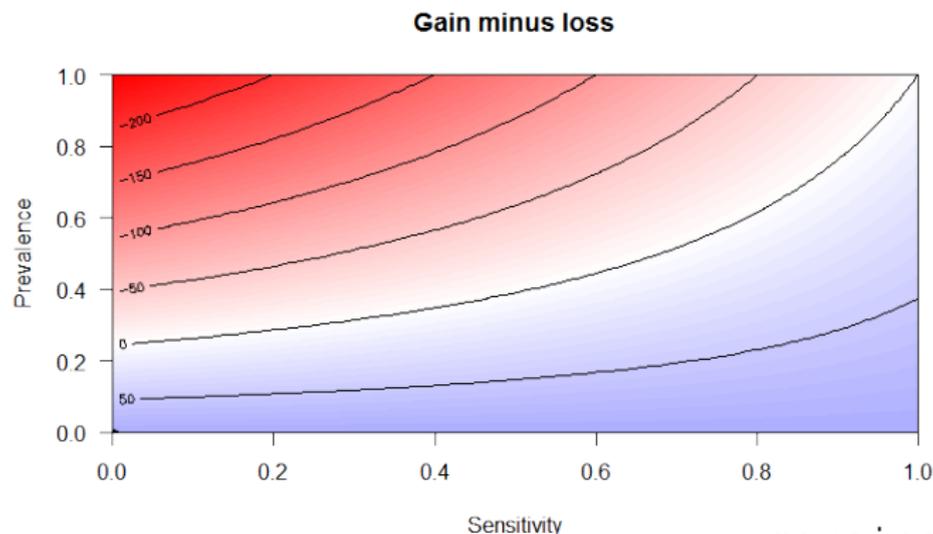


Nicht-parametrische Regression



Welche Parameterkombination ist günstig?

Break-Even Analyse



Möglicher Gewinn (gain) und Verlust (loss) ist abhängig von den unsicheren Parametern Prävalenz und Sensitivität.

Zusammenfassung



Risikoschätzung unter Unsicherheit ist möglich.

Das zugrunde gelegte Szenario ist entscheidend für das Ergebnis.

Verschiedene Informationsquellen können integriert und müssen vollständig dokumentiert werden.

Der Effekt von unsicheren Parameter oder anderen Setzungen sollte untersucht werden.

- Belgorodski, N., Greiner, M., and Engelhardt, A. (2012a). *riskBayes: Predefined Bayes models fitted with Markov chain Monte Carlo (MCMC) (related to the 'risk' project)*. R package version 2.3.
- Belgorodski, N., Greiner, M., Tolksdorf, K., and Schueller, K. (2012b). *riskDistributions: Fitting distributions to given data or known quantiles*. R package version 1.8.
- EFSA (2011). Scientific Opinion on the revision of the quantitative risk assessment (QRA) of the BSE risk posed by processed animal proteins (PAPs). *EFSA Journal*, 9:1947.
- Greiner, M., Smid, J., Havelaar, A. H., and Müller-Graf, C. (2013). Graphical models and Bayesian domains in risk modelling: Application in microbiological risk assessment. *Preventive Veterinary Medicine*, 110:4–11.
- Pouillot, R., Delignette-Muller, M., Kelly, D., and Dennis, J. (2009). The mc2d package. Technical report.
- Refsgaard, J. C., Van der Sluijs, J. P., Brown, J., and Van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, 29(11):1586–1597.
- WHO (2008). Uncertainty and data quality in exposure assessment. part 1: Guidance document on characterizing and communicating uncertainty in exposure assessment. part 2: Hallmarks of data quality in chemical exposure assessment. Technical report, World Health Organization and International Labour Organization and the United Nations Environment Programme.
- WHO (2009). Risk characterization of microbiological hazards in food. Technical report, World Health Organization (WHO).

Vielen Dank

- ... den Mitarbeitenden des `risk`-Projekts
- ... den Mitgliedern des Statistik-Ausschusses der BfR-Kommission für Exposition
- ... Ihnen für Ihre Aufmerksamkeit

Prof. Dr. Matthias Greiner (Dipl ECVPH, MSc)

Bundesinstitut für Risikobewertung (BfR)

Abteilung für Wissenschaftliche Querschnittsaufgaben

Max-Dohrn-Straße 8–10, 10589 Berlin, Germany

Tel. +49 30 18412–3297, Matthias.Greiner@bfr.bund.de

Stiftung Tierärztliche Hochschule Hannover (TiHo)

Professur für Expositionsschätzung und Quantitative Risikobewertung