



DIPLOMARBEIT

Bayes'sche Methoden in der Zuverlässigkeitstheorie

zur Erlangung des akademischen Grades
Diplom-Ingenieur

im Rahmen des Studiums
Statistik-Wirtschaftsmathematik

eingereicht von
Jutta Gamper BSc
Matrikelnummer 01026714

ausgeführt am Institut für Stochastik und Wirtschaftsmathematik
der Fakultät für Mathematik und Geoinformation der Technischen Universität Wien

Betreuung

Betreuer: Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Klaus Felsenstein

Wien, 17.04.2019

Unterschrift (Verfasserin)

Unterschrift (Betreuer)

Inhaltsverzeichnis

1	Einleitung	1
2	Bayes Statistik	3
2.1	Der Satz von Bayes	3
2.2	Die A-priori-Verteilung	5
2.2.1	Konjugierte A-priori	5
2.2.2	Nicht informative A-priori	8
2.3	Grundlagen der Entscheidungstheorie	10
2.3.1	Bayes'sche Parameterschätzung	11
2.4	Hypothesentests und Konfidenzbereiche	12
2.4.1	Bayes Test	12
2.4.2	HPD-Bereiche	14
3	Zuverlässigkeitstheorie	17
3.1	Konzepte der Zuverlässigkeitstheorie	17
3.1.1	Zensierte Daten	20
3.1.2	Likelihood bei zensierten Daten	20
3.2	Lebensdauerverteilungen	22
3.2.1	Exponentialverteilung	22
3.2.2	Weibullverteilung	25
3.2.3	Log-Normalverteilung	27
3.2.4	Gammaverteilung	28
3.2.5	Gumbelverteilung	30
3.3	Beschleunigte Lebensdauerversuche	31
3.3.1	Parameterschätzung	33
3.3.2	Ein einfaches Step-Stress Modell	36
4	Bayes'sche Analyse von parametrischen Modellen	39
4.1	Exponentialverteilung	39
4.2	Weibullverteilung	42
4.3	Log-Normalverteilung	48
4.4	Gammaverteilung	50

5	Bayes'sche Analyse von beschleunigten Lebensdauermodellen	51
5.1	Power Rule Modell	51
5.1.1	Lineares Modell	54
5.1.2	Beispiel mit simulierten Daten	55
5.2	Step-Stress Modell	58
6	Zusammenfassung	61

1 | Einleitung

Diese Arbeit gibt einen Überblick über die Anwendung Bayes'scher Methoden im Bereich der Zuverlässigkeitstheorie. Industrieunternehmen haben ein großes Interesse daran, sicherzustellen, dass ihre Produkte bestimmte Qualitätsstandards erfüllen. Um quantitative Aussagen über die Zuverlässigkeit von Systemen oder Komponenten zu treffen, ist es daher notwendig und zweckmäßig, statistische Methoden einzusetzen. In Zeiten, wo von bestimmten Produkten wie z.B. Halbleiterbauteilen eine immer höhere Zuverlässigkeit verlangt wird, ist es umso wichtiger, genaue Abschätzungen und Voraussagen betreffend dieser angeben zu können. Die Methoden der Bayes Statistik bieten in diesem Zusammenhang die Möglichkeit, subjektive Informationen, die gerade im Industriebereich in Form von Fachwissen von Produktionstechnikern vorhanden sind, in das statistische Modell einfließen zu lassen.

Nicht nur in technischen Bereichen, sondern auch in der Medizin oder in den Wirtschaftswissenschaften finden die Methoden der Zuverlässigkeits- oder Lebensdaueranalyse Anwendung. Im medizinischen Bereich ist es etwa von Interesse, die Auswirkungen bestimmter Medikamente auf die Überlebensrate von Patienten zu untersuchen.

Im Gegensatz zur klassischen oder frequentistischen Statistik, wo Wahrscheinlichkeit als relative Häufigkeit eines Ereignisses angesehen wird, basiert der Bayes'sche Wahrscheinlichkeitsbegriff auf subjektiven Wahrscheinlichkeiten. Namensgeber ist der englische Geistliche Thomas Bayes (1702-1761), der im erst nach seinem Tod 1764 erschienen Text „An essay towards solving a problem in the doctrine of chances“ erstmals eine einfache Form der Bayes'schen Formel beschreibt.

Die Prinzipien der Bayes Statistik ermöglichen es, subjektive Information in ein statistisches Modell einfließen zu lassen, indem das Vorwissen durch eine Wahrscheinlichkeitsverteilung (die A-priori Verteilung) modelliert wird. Durch neue Information, etwa den Daten aus einer Zufallsstichprobe, wird diese Wahrscheinlichkeitsverteilung aktualisiert.

Gerade im Bereich der Zuverlässigkeitsanalyse ist Expertenwissen über ein Produkt oder eine Komponente vorhanden, welches durch die A-priori Verteilung in der Modellierung berücksichtigt wird. Besonders wenn es wenig historische Daten gibt und noch keine Tests zu einem Produkt durchgeführt wurden, ist die Einschätzung von Experten oft die einzige vorhandene Information. Oder es gibt bereits genauere Schätzungen über einzelne Subkomponenten in einem System, welche für die Analyse des gesamten Systems miteinbezogen werden sollen. Durch das Zusammenführen von A-priori Wissen und Datenwissen auf mathematisch fundierte Weise ist es möglich, eine Bewertung über die Zuverlässigkeit von Systemen vorzunehmen, die auf allen vorhandenen Informationen basiert und in Form einer Wahrscheinlichkeitsverteilung zur Verfügung steht, welche für weitere Analysen und Einschätzungen herangezogen werden kann. Die Schätzmethode der klassischen Statistik bieten keine so flexible Möglichkeit

Vorinformation auf solche Art zu nutzen wie Bayes'sche Verfahren.

In den nächsten Kapiteln wird ein Blick auf die Methoden der Zuverlässigkeitstheorie aus Bayes'scher Sicht geworfen. Zunächst werden in Kapitel 2 grundlegende Konzepte der Bayes Statistik vorgestellt. Es werden der zentrale Satz von Bayes sowie verschiedene A-priori Verteilungen erläutert. Ebenso werden Bayes'sche Methoden der Parameterschätzung, Hypothesentests und Vertrauensbereiche beschrieben.

In Kapitel 3 werden wesentliche Begriffe aus der Zuverlässigkeitstheorie, wie z.B. Zuverlässigkeit, Lebensdauer und Ausfallrate definiert und einige parametrische Lebensdauerverteilungen charakterisiert. Ein besonderer Fokus dieser Arbeit, die Analyse der so genannten beschleunigten Lebensdauerexperimente, wird in diesem Kapitel aus Sicht der klassischen Statistik beschrieben.

Kapitel 4 führt die beiden Bereiche zusammen und behandelt die Anwendung der Bayes Statistik auf die zuvor vorgestellten parametrischen Lebensdauermodelle. Die Bayes'sche Analyse der beschleunigten Lebensdauerexperimente ist in Kapitel 5 beschrieben.

2 | Bayes Statistik

In diesem Kapitel werden Grundlagen der Bayes Statistik vorgestellt. Die folgenden Abschnitte basieren auf den Büchern [Robert, 2007], [Koch, 2000], [Berger, 2013] und [DeGroot, 2004]. Für grundlegende Begriffe aus Maß- und Wahrscheinlichkeitstheorie sei auf [Kusolitsch, 2014] verwiesen.

2.1 Der Satz von Bayes

Im Folgenden betrachten wir stets einen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Das ist eine Menge Ω , eine σ -Algebra \mathcal{A} auf Ω und eine Wahrscheinlichkeitsfunktion $P : \mathcal{A} \rightarrow [0, 1]$ mit den Eigenschaften:

1. $P(\Omega) = 1$,
2. (σ -Additivität) Für eine abzählbare Folge von disjunkten Ereignissen $A_i \in \mathcal{A}, i \in I$, $A_i \cap A_j = \emptyset, i \neq j$ gilt

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i).$$

Definition 2.1 (Bedingte Wahrscheinlichkeit). Seien $A, B \in \mathcal{A}$ und $P(B) > 0$. Die bedingte Wahrscheinlichkeit von A unter der Bedingung B ist definiert als

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Aus der obigen Definition 2.1 kann die folgende Multiplikationsregel abgeleitet werden: $P(A \cap B) = P(A|B)P(B)$.

Satz 2.1 (Bayes Formel). Seien H_1, H_2, \dots eine höchstens abzählbare Folge von disjunkten Ereignissen aus \mathcal{A} mit $\bigcup_i H_i = \Omega$. Weiters sei $A \in \mathcal{A}$ ein Ereignis mit $P(A) > 0$. Dann gilt

$$P(H_i|A) = \frac{P(A|H_i)P(H_i)}{\sum_j P(A|H_j)P(H_j)}.$$

Beweis. Aus der Definition 2.1 folgt

$$P(H_i|A) = \frac{P(A \cap H_i)}{P(A)} = \frac{P(A|H_i)P(H_i)}{P(A)}.$$

Nun gilt wegen der Definition der H_i , der σ -Additivität von P und der Multiplikationsregel

$$P(A) = P(A \cap \Omega) = P\left(A \cap \bigcup_i H_i\right) = \sum_i P(A \cap H_i) = \sum_i P(A|H_i)P(H_i). \quad \square$$

Der folgende Satz von Bayes ist eine Version dieses Resultats für stetige Zufallsvariablen. Eine n -dimensionale Zufallsvariable X ist eine reellwertige, messbare Funktion von Ω in den Messraum $(\mathbb{R}^n, \mathcal{B}^n)$, wobei \mathcal{B}^n die σ -Algebra der Borelmengen über \mathbb{R}^n bezeichnet.

Definition 2.2 (bedingte Dichte). Seien X und Y zwei stetige Zufallsvariablen in $(\mathbb{R}^n, \mathcal{B}^n)$ und $(\mathbb{R}^m, \mathcal{B}^m)$ mit gemeinsamer Dichte $f(x, y)$. Die Randdichten $f_X(x)$ und $f_Y(y)$ sind definiert als

$$f_X(x) = \int_{\mathbb{R}^m} f(x, y) \, dy \quad \text{und} \quad f_Y(y) = \int_{\mathbb{R}^n} f(x, y) \, dx.$$

Für ein x mit $f_X(x) > 0$ ist die bedingte Dichte von Y gegeben $X = x$ definiert durch

$$f(y|x) = \frac{f(x, y)}{f_X(x)}.$$

Für ein y mit $f_Y(y) > 0$ ist die bedingte Dichte von X gegeben $Y = y$ definiert durch

$$f(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

Satz 2.2 (Satz von Bayes). Seien X und Y zwei stetige Zufallsgrößen mit gemeinsamer Dichte $f(x, y)$. Für die bedingte Dichte von $Y|X$ gilt

$$f(y|x) = \frac{f(x|y)f_Y(y)}{\int f(x|y)f_Y(y) \, dy}.$$

Beweis. Aus der Definition 2.2 folgt

$$f(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{f(x|y)f_Y(y)}{f_X(x)}.$$

Dabei steht im Nenner die Randdichte von X :

$$f_X(x) = \int f(x, y) \, dy = \int f(x|y)f_Y(y) \, dy. \quad \square$$

Die Beobachtungen $X = (X_1, X_2, \dots, X_n)$ sind unabhängige und identisch verteilte Zufallsgrößen aus einem stochastischen Modell der Form $X \sim f(x|\theta)$. Der Parameter θ ist in der Bayes Statistik eine unbekannte Zufallsgröße, die die Verteilung von X durch das Modell $f(x|\theta)$ bestimmt. Die Unsicherheit über den Parameter wird durch eine Verteilung π auf dem Parameterraum Θ ausgedrückt. Diese Verteilung heißt *A-priori-Verteilung* von θ . Die bedingte Verteilung von θ nach Beobachtung der Daten X heißt *A-posteriori-Verteilung* und ist nach dem Satz von Bayes (2.2) gegeben durch

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int_{\Theta} f(x|\theta)\pi(\theta) \, d\theta}. \quad (2.1.1)$$

Die Randdichte von X ,

$$m(x) = \int_{\Theta} f(x|\theta)\pi(\theta) d\theta,$$

ist ein Normierungsfaktor, der nicht von θ abhängt. Daher wird (2.1.1) kurz als

$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta)$$

geschrieben. Nach Beobachtung der Daten wird die Dichte $f(x|\theta)$ als Funktion von θ interpretiert und als Likelihood bezeichnet:

$$l(\theta|x) = f(x|\theta).$$

Die Prognose für eine Zufallsgröße $Z \sim g(z|x, \theta)$ nach Beobachtung von $X \sim f(x|\theta)$ erfolgt durch die so genannte Prädiktivdichte.

Definition 2.3. Die (A-posteriori-) Prädiktivdichte von Z gegeben X ist definiert als

$$g(z|x) = \int_{\Theta} g(z|x, \theta)\pi(\theta|x) d\theta.$$

Die gemeinsame Dichte von θ und z unter x ist gegeben durch $\pi(\theta|x)g(z|x, \theta)$ und durch Integration über θ erhält man die Randdichte von z unter x .

2.2 Die A-priori-Verteilung

Die Wahl einer geeigneten A-priori-Verteilung ist ein zentraler Teil der Bayes Statistik. Die A-priori umfasst die gesamte Information, die über den unbekannt Parameter vor Beobachten der Daten vorhanden ist. Falls keine Vorinformation vorliegt, werden so genannte nicht informative A-priori-Verteilungen verwendet. Analytisch besonders einfach zu behandeln sind die konjugierten Verteilungen. Dabei gehören A-priori und A-posteriori zur selben Verteilungsfamilie.

2.2.1 Konjugierte A-priori

Definition 2.4 (Konjugierte Verteilungsfamilie). Eine Familie \mathcal{F} von Verteilungen auf Θ heißt konjugiert zur Likelihoodfunktion $f(x|\theta)$, wenn für alle $\pi \in \mathcal{F}$ gilt, dass die A-posteriori-Verteilung $\pi(\theta|x)$ ebenfalls in \mathcal{F} liegt.

Konjugierte Verteilungen sind mathematisch einfach zu behandeln, da nach Beobachten der Daten nur die Hyperparameter der A-priori angepasst werden müssen. Dabei bestimmt die Wahl der Hyperparameter die gesamte subjektive Information über den Parameter. Die Hyperparameter können als Ergebnis eines bereits durchgeführten, vergleichbaren Experimentes interpretiert werden.

Mischungen von mehreren konjugierten Verteilungen ergeben allgemeinere A-prioris, die immer noch analytisch leicht zu behandeln sind, aber mehr und flexiblere Möglichkeiten in der Modellierung von Vorinformation bieten.

Bei der Konstruktion von konjugierten Verteilungen spielen suffiziente Statistiken und die Exponentialfamilie von Verteilungen eine besondere Rolle. Eine suffiziente Statistik ist eine Funktion der Daten $T(X)$, welche die gesamte in der Stichprobe vorhandene Information über den Parameter enthält [Casella and Berger, 2002, Sec. 6.2].

Definition 2.5. Eine Funktion $T(X)$ heißt suffiziente Statistik für θ , falls die bedingte Verteilung von X unter $T(X)$ nicht von θ abhängt.

Mit dem folgenden Satz lässt sich eine suffiziente Statistik leicht konstruieren.

Satz 2.3 (Faktorisierungssatz). *Eine Statistik T ist genau dann suffizient für θ , wenn die Dichtefunktion $f(x|\theta)$ als Produkt von zwei nicht-negativen Funktionen g und h dargestellt werden kann:*

$$f(x|\theta) = g(T(x)|\theta)h(x).$$

Dabei hängt g von X nur über T ab und h ist von θ unabhängig.

Beweis. Ein Beweis ist z.B. in [DeGroot, 2004, S. 156] zu finden. \square

Definition 2.6 (Exponentialfamilie). Eine Verteilungsfamilie auf $(\Omega, \mathcal{A}, \mathbb{P})$ heißt Exponentialfamilie, wenn ihre Wahrscheinlichkeitsdichte geschrieben werden kann als

$$f(x|\theta) = C(\theta)h(x) \exp\{R(\theta)^\top T(x)\} \quad (2.2.1)$$

mit messbaren Funktionen $C : \mathcal{A} \rightarrow \mathbb{R}_+$, $h : \Theta \rightarrow \mathbb{R}_+$, $R : \Theta \rightarrow \mathbb{R}^k$ und $T : \mathcal{A} \rightarrow \mathbb{R}^k$. Ist $R(\theta) = \theta$, so spricht man von einer natürlich parametrisierten Exponentialfamilie.

Eine konjugierte Verteilung existiert genau dann, wenn die Beobachtungen aus einer Verteilungsfamilie stammen, für die eine suffiziente Statistik mit fester Dimension für jede Stichprobengröße existiert (vgl. [DeGroot, 2004, Sec. 9.2]). Aus dem Faktorisierungssatz folgt direkt, dass für eine Exponentialfamilie die Funktionen T für θ suffizient sind. Eine Umkehrung dieser Aussage liefert das folgende Lemma [Robert, 2007].

Satz 2.4 (Pitman-Koopmann Lemma). *Wenn für eine Verteilungsfamilie $f(\cdot|\theta)$ eine suffiziente Statistik existiert, deren Dimension unabhängig von der Stichprobengröße ist, dann ist die Familie eine Exponentialfamilie, wenn der Träger von $f(\cdot|\theta)$ nicht von θ abhängt.*

Den Abschnitt beschließen einige Beispiele für Familien mit konjugierten Verteilungen.

Beispiel 2.1 (Binomialverteilung). Sei $X \sim B(n, \theta)$ mit $n \in \mathbb{N}^+$ und $\theta \in [0, 1]$. Die Likelihood hat die Form

$$l(\theta|x_1, \dots, x_n) \propto \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{n - \sum_{i=1}^n x_i}.$$

Die konjugierte A-priori für dieses Modell ist eine Betaverteilung

$$\pi(\theta) \propto \theta^{a-1} (1 - \theta)^{b-1}, \quad a > 0, b > 0.$$

Die A-posteriori-Verteilung erhält man durch updaten der Hyperparameter:

$$\pi(\theta|x_1, \dots, x_n) \propto \theta^{a-1 + \sum_{i=1}^n x_i} (1 - \theta)^{b-1 + n - \sum_{i=1}^n x_i}.$$

Sie entspricht wieder einer Betaverteilung mit Parametern $a + \sum_{i=1}^n x_i$ und $b + n - \sum_{i=1}^n x_i$. \diamond

Im letzten Beispiel wurden die A-priori Hyperparameter a, b nicht nur auf die positiven ganzen Zahlen beschränkt, sondern können beliebige positive reelle Werte annehmen. Es ist üblich, konjugierte Verteilungsfamilien auf diese Art zu vergrößern bzw. zu vervollständigen.

Beispiel 2.2 (Normalverteilung, Varianz bekannt). Sei $X \sim N(\theta, \sigma_0^2)$, wobei $\sigma_0^2 \in \mathbb{R}_{>0}$ als bekannt vorausgesetzt wird. Die Likelihood ist

$$l(\theta|x_1, \dots, x_n) = \frac{1}{(2\pi\sigma_0^2)^{n/2}} \exp\left(-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \theta)^2\right).$$

Die konjugierte A-priori für θ ist eine Normalverteilung $N(m, d^2)$. Die A-posteriori-Verteilung ist eine Normalverteilung $N(m^*, d^{*2})$ mit

$$m^* = \frac{m\sigma_0^2 + nd^2\bar{x}}{\sigma_0^2 + nd^2} \quad \text{und} \quad d^{*2} = \frac{d^2\sigma_0^2}{\sigma_0^2 + nd^2}. \quad \diamond$$

Beispiel 2.3 (Normalverteilung, beide Parameter unbekannt). Sei $X \sim N(\mu, \sigma^2)$ mit unbekanntem Parametern $\mu \in \mathbb{R}$ und $\sigma^2 \in \mathbb{R}_{>0}$. Zur Vereinfachung verwenden wir eine Parametrisierung mit der Präzision $\tau = 1/\sigma^2$. Die Likelihood ist

$$l(\mu, \tau|x_1, \dots, x_n) = \frac{\tau^{n/2}}{(2\pi)^{n/2}} \exp\left(-\frac{\tau}{2} \sum_{i=1}^n (x_i - \mu)^2\right).$$

Die konjugierte A-priori für $\theta = (\mu, \tau)$ ist eine Normal-Gamma-Verteilung $NG(m, d^2, a, b)$, für die gilt, dass

$$\begin{aligned} \mu|\tau &\sim N(m, d^2/\tau), \\ \tau &\sim \gamma(a, b). \end{aligned}$$

Die A-posteriori-Verteilung ist nach dem Satz von Bayes

$$\pi(\mu, \tau|x_1, \dots, x_n) \propto \tau^{a+n/2-1} \tau^{1/2} \exp\left\{-\frac{\tau}{2} \left(2b + (\mu - m)^2 + \sum_{i=1}^n (x_i - \mu)^2\right)\right\}.$$

Durch Umformen und quadratisches Ergänzen kommt man wieder auf eine Normal-Gamma-Verteilung $NG(m^*, d^{*2}, a^*, b^*)$ mit

$$\begin{aligned} m^* &= \frac{m + d^2 n \bar{x}}{1 + d^2 n}, \\ d^{*2} &= \frac{d^2}{1 + d^2 n}, \\ a^* &= a + \frac{n}{2}, \\ b^* &= b + \frac{1}{2} \left(\frac{n(\bar{x} - m)^2}{1 + nd^2} + \sum_{i=1}^n (x_i - \bar{x})^2 \right), \end{aligned}$$

und $\bar{x} = 1/n \sum_{i=1}^n x_i$. Der A-posteriori Mittelwert ist ein gewichtetes Mittel aus dem A-priori Mittelwert m und dem Mittelwert der Daten \bar{x} . \diamond

Beispiel 2.4 (Exponentialverteilung). Eine exponentialverteilte Zufallsgröße X hat die Dichtefunktion $f(x|\theta) = \theta e^{-\theta x} \mathbb{1}_{(0, \infty)}(x)$. Die Likelihood ist

$$l(\theta|x_1, \dots, x_n) = \theta^n e^{-\theta \sum_{i=1}^n x_i}.$$

Die suffizienten Statistiken sind die Stichprobengröße n und die Summe der Daten $\sum_{i=1}^n x_i$. Die konjugierte A-priori ist von der Form

$$\pi(\theta) \propto \theta^{a-1} e^{-\theta b}, \quad a, b > 0.$$

Das ist eine Gammaverteilung $\gamma(a, b)$ mit dem Normierungsfaktor $b^a/\Gamma(a)$. Die A-posteriori ist eine $\gamma(a+n, b+\sum_{i=1}^n x_i)$ -Verteilung:

$$\pi(\theta|x) \propto \theta^{a+n} e^{-\theta(b+\sum_{i=1}^n x_i)}. \quad \diamond$$

Ganz allgemein ist die Likelihood einer Exponentialfamilie mit Dichte der Form (2.2.1) gleich

$$l(\theta|x_1, \dots, x_n) = \prod_{i=1}^n h(x_i) C(\theta)^n \exp \left\{ R(\theta)^\top \sum_{i=1}^n T(x_i) \right\}.$$

Der erste Faktor ist eine vernachlässigbare Konstante. Die suffizienten Statistiken sind n und $\sum_{i=1}^n T(x_i)$. Die konjugierte A-priori mit den Hyperparametern a, b hat die Form

$$\pi(\theta) \propto C(\theta)^a \exp(R(\theta)^\top b).$$

Die A-posteriori erhält man durch updaten:

$$\pi(\theta|x_1, \dots, x_n) \propto C(\theta)^{a+n} \exp \left\{ R(\theta)^\top \left(b + \sum_{i=1}^n T(x_i) \right) \right\}.$$

2.2.2 Nicht informative A-priori

Wenn keine Vorinformation vorhanden ist, bieten nicht informative A-priori-Verteilungen eine Möglichkeit, das Fehlen von Information abzubilden. Eine gewünschte Eigenschaft nicht informativer Verteilungen ist die Invarianz unter Parametertransformationen. Wenn keine Information über den Parameter vorhanden ist, dann soll die A-priori auch keine Information über eine Transformation des Parameters enthalten. [Robert, 2007]

Beispiel 2.5. Sei $X \sim F_\theta$, wobei F_θ eine Verteilungsfamilie mit Lageparameter θ und Verteilungsfunktion $F_\theta(x) = F(x - \theta)$ ist. Für eine Zufallsvariable $Y = X + c, c \in \mathbb{R}$ ist der Lageparameter $\theta + c$. Eine A-priori, die translationsinvariant ist, muss $\pi(\theta) = \pi(\theta - c)$ für jedes $c \in \mathbb{R}$ erfüllen. Daher muss $\pi(\theta)$ konstant sein und ist somit nicht normierbar. \diamond

Beispiel 2.6. Sei $X \sim F_\sigma$, wobei F_σ eine Verteilungsfamilie mit Skalenparameter $\sigma > 0$ und Dichte $f_\sigma(x) = \sigma^{-1} f(x/\sigma)$ ist. Die Dichte von $Y = cX$ ist $(1/c\sigma) f(y/c\sigma)$. Eine skaleninvariante A-priori muss $\pi(\sigma) = c^{-1} \pi(\sigma/c)$ für jedes $c > 0$ erfüllen. Daher muss $\pi(\sigma) \propto 1/\sigma$ sein. Diese Dichte ist ebenfalls nicht normierbar. \diamond

Nicht informative A-priori sind häufig uneigentliche Dichten. Eine nicht negative Dichte g heißt uneigentlich, wenn $\int_{\Theta} g(\theta) d\theta = \infty$ ist. Die A-priori im Beispiel 2.6 entspricht einer uneigentlichen $\gamma(0, 0)$ -Dichte. Das ist nicht weiter problematisch, solange die A-posteriori nach Beobachtung der Daten normierbar ist.

Jeffreys A-priori

Die Jeffreys A-priori ist eine nicht informative A-priori-Verteilung, welche auf der Fisher-Information basiert. Sie ist invariant unter Parametertransformationen.

Definition 2.7 (Fisher-Information). Die Fisher-Informationsmatrix $\mathcal{I}(\theta) = (I_{ij}(\theta))_{i,j=1,\dots,k}$ hat die Einträge

$$I_{ij}(\theta) = \mathbb{E} \left(\frac{\partial}{\partial \theta_i} \log f(x|\theta) \frac{\partial}{\partial \theta_j} \log f(x|\theta) \right).$$

Eine alternative Darstellung der Fisher-Information ist

$$I_{ij}(\theta) = -\mathbb{E} \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(x|\theta) \right),$$

falls gewisse Regularitätsbedingungen¹ erfüllt sind.

Definition 2.8 (Jeffreys A-priori). Die Jeffreys A-priori ist proportional zur Wurzel der Determinante der Fisher-Information,

$$\pi_J(\theta) \propto |\mathcal{I}(\theta)|^{1/2}.$$

Beispiel 2.7. In einem Normalmodell $N(\mu, \sigma^2)$ seien Mittelwert und Varianz unbekannt und $\theta = (\mu, \sigma)$. Dann ist

$$\mathcal{I}(\mu, \sigma) = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/2\sigma^2 \end{pmatrix}$$

und die Jeffreys A-priori $\pi_J(\theta) \propto 1/\sigma^2$. Betrachtet man μ und σ unabhängig voneinander, ist die nicht informative A-priori $\pi(\mu, \sigma) \propto 1/\sigma$ (vgl. Beispiele 2.5 und 2.6). \diamond

Beispiel 2.8 (Exponentialverteilung). Sei $X \sim \text{Exp}(\theta)$. Die Fisher-Information ist

$$I(\theta) = -\mathbb{E} \left(\frac{\partial^2}{\partial \theta^2} \log f(x|\theta) \right) = -\mathbb{E} \left(-\frac{1}{\theta^2} \right) = \frac{1}{\theta^2}.$$

Damit ergibt sich die Jeffreys A-priori $\pi_J(\theta) \propto \theta^{-1}$, was einer uneigentlichen $\gamma(0, 0)$ -Dichte entspricht. \diamond

¹Die Bedingungen sind ([Lehmann and Casella, 2006]):

1. Der Parameterraum Θ ist ein offenes Intervall.
2. Die Dichten $f(x|\theta)$ haben einen gemeinsamen Träger für alle $\theta \in \Theta$.
3. $f(x|\theta)$ ist als Funktion von θ zwei Mal differenzierbar und Ableitung und Integral sind vertauschbar, d.h. für $k = 1, 2$ gilt:

$$\frac{\partial^k}{\partial \theta^k} \int_{-\infty}^{\infty} f(x|\theta) dx = \int_{-\infty}^{\infty} \frac{\partial^k}{\partial \theta^k} f(x|\theta) dx.$$

2.3 Grundlagen der Entscheidungstheorie

Das Schätzen von Parametern und das Testen von Hypothesen in einem statistischen Modell können als entscheidungstheoretische Probleme aufgefasst werden [Robert, 2007]. Die Aufgabe eines Statistikers ist es, eine Entscheidung $d \in \mathcal{D}$ über den Parameter $\theta \in \Theta$ zu treffen. Dies geschieht auf Basis von Beobachtungen $x \in \mathcal{X}$, die mit dem Parameter θ durch eine Modellverteilung $f(x|\theta)$ in Verbindung stehen. Eine Bewertung der Entscheidung findet durch eine Verlustfunktion $L : \Theta \times \mathcal{D} \rightarrow [0, \infty)$ statt.

Definition 2.9. Das a-posteriori zu erwartende Risiko ist definiert als

$$\rho(\pi, d|x) = \mathbb{E}_\pi[L(\theta, d)|x] = \int_{\Theta} L(\theta, d)\pi(\theta|x) d\theta.$$

Als Bayes Strategie bezeichnet man die Wahl jener Entscheidung d , die das a-posteriori Risiko minimiert, vorausgesetzt, es existiert eine Entscheidung mit endlichem Risiko.

In der klassischen Statistik wird das Risiko $R(\theta, \delta)$ als Erwartungswert der Verlustfunktion mit der Likelihoodfunktion $f(x|\theta)$ berechnet:

$$R(\theta, \delta) = \mathbb{E}_\theta[L(\theta, \delta(x))] = \int_{\mathcal{X}} L(\theta, \delta(x))f(x|\theta) dx.$$

Dabei ist $\delta : \mathcal{X} \rightarrow \mathcal{D}$ eine Entscheidungsregel, die üblicherweise als Schätzer bezeichnet wird. Es wird über die Beobachtungen x integriert und für festes δ ist $R(\theta, \delta)$ eine Funktion des unbekanntes Parameters θ . Wenn eine A-priori-Dichte $\pi(\theta)$ gegeben ist, kann man die Risikofunktion über θ gemäß der A-priori integrieren und erhält

$$r(\pi, \delta) = \mathbb{E}_\pi[R(\theta, \delta)] = \int_{\Theta} \int_{\mathcal{X}} L(\theta, \delta(x))f(x|\theta) d\theta \pi(\theta) d\theta.$$

Ein Schätzer, der $r(\pi, \delta)$ minimiert, minimiert auch $\rho(\pi, \delta(x)|x)$, denn es gilt

$$\begin{aligned} r(\pi, \delta) &= \int_{\Theta} \int_{\mathcal{X}} L(\theta, \delta(x))f(x|\theta) dx \pi(\theta) d\theta \\ &= \int_{\mathcal{X}} \int_{\Theta} L(\theta, \delta(x))f(x|\theta)\pi(\theta) d\theta dx \\ &= \int_{\mathcal{X}} \int_{\Theta} L(\theta, \delta(x))\pi(\theta|x) d\theta m(x) dx \\ &= \int_{\mathcal{X}} \rho(\pi, \delta(x)|x) m(x) dx. \end{aligned}$$

Definition 2.10 (Bayes Schätzer). Der Bayes Schätzer zu der A-priori-Verteilung π und der Verlustfunktion L ist eine Entscheidungsregel δ^π , die $r(\pi, \delta)$ minimiert. Der Wert $r(\pi) = r(\pi, \delta^\pi)$ heißt Bayes Risiko.

2.3.1 Bayes'sche Parameterschätzung

Der Bayes Schätzer minimiert unter allen möglichen Entscheidungen das Bayes Risiko. Für einige konkrete Verlustfunktionen wird nun der Bayes Schätzer hergeleitet.

Satz 2.5. Für die quadratische Verlustfunktion $L(\theta, d) = (\theta - d)^2$ ist der Bayes Schätzer δ^π gleich der a-posteriori Mittelwert

$$\delta^\pi(x) = \mathbb{E}_\pi[\theta|x].$$

Beweis. Das a-posteriori Risiko ist $\mathbb{E}_\pi[(\theta - d)^2|x] = \mathbb{E}_\pi[\theta^2|x] - 2d\mathbb{E}_\pi[\theta|x] + d^2$. Die Ableitung nach d ist $-2\mathbb{E}_\pi[\theta|x] + 2d$. Dieser Ausdruck ist gleich 0 für $d = \mathbb{E}_\pi[\theta|x]$. Die zweite Ableitung ist größer als Null und damit ist $d = \mathbb{E}_\pi[\theta|x]$ ein Minimum. \square

Dieses Ergebnis kann für gewichtete quadratische Verlustfunktionen verallgemeinert werden.

Satz 2.6. Für die gewichtete quadratische Verlustfunktion $L(\theta, d) = \omega(\theta)(\theta - d)^2$ ist der Bayes Schätzer

$$\delta^\pi(x) = \frac{\mathbb{E}_\pi[\omega(\theta)\theta|x]}{\mathbb{E}_\pi[\omega(\theta)|x]}.$$

Eine robustere Alternative zum quadratischen Verlust ist die absolute Verlustfunktion $L(\theta, d) = |\theta - d|$. Eine gewichtete absolute Verlustfunktion ist gegeben durch

$$L(\theta, d) = \begin{cases} k_2(\theta - d) & \text{falls } \theta > d, \\ k_1(d - \theta) & \text{falls } \theta \leq d. \end{cases} \quad (2.3.1)$$

Satz 2.7. Für die Verlustfunktion (2.3.1) ist der Bayes Schätzer das $k_2/(k_1 + k_2)$ -Quantil der A-posteriori Verteilung.

Beweis. Das $p = k_2/(k_1 + k_2)$ -Quantil der A-posteriori Verteilung, x_p erfüllt

$$P(\theta \leq x_p) \geq p \quad \text{und} \quad P(\theta \geq x_p) \geq 1 - p.$$

Zunächst soll gezeigt werden, dass $\mathbb{E}_\pi[L(\theta, x_p)|x] \leq \mathbb{E}_\pi[L(\theta, d)|x]$ ist für $d > x_p$. Für $d > x_p$ kann die Differenz der Verlustfunktionen $L(\theta, x_p) - L(\theta, d)$ abgeschätzt werden durch

$$L(\theta, x_p) - L(\theta, d) \leq k_2(d - x_p)\mathbb{1}_{(x_p, \infty)}(\theta) + k_1(x_p - d)\mathbb{1}_{(-\infty, x_p]}(\theta).$$

Damit gilt für den Erwartungswert

$$\begin{aligned} \mathbb{E}_\pi[L(\theta, x_p) - L(\theta, d)|x] &\leq k_2(d - x_p)P_\pi(\theta > x_p|x) + k_1(x_p - d)P_\pi(\theta \leq x_p|x) \\ &\leq k_2(d - x_p)(1 - p) + k_1(x_p - d)p \\ &= (d - x_p) \left(k_2 \frac{k_1}{k_1 + k_2} - k_1 \frac{k_2}{k_1 + k_2} \right) = 0. \end{aligned}$$

Es gilt also $\mathbb{E}_\pi[L(\theta, x_p)|x] \leq \mathbb{E}_\pi[L(\theta, d)|x]$. Analog zeigt man die Ungleichung auch für $d < x_p$. Der Bayes Schätzer ist daher genau das $p = k_2/(k_1 + k_2)$ -Quantil der A-posteriori Verteilung. \square

Für $L(\theta, d) = |\theta - d|$, d.h. $k_1 = k_2$, ist der Bayes Schätzer gleich dem Median der A-posteriori-Verteilung.

2.4 Hypothesentests und Konfidenzbereiche

2.4.1 Bayes Test

Hypothesentests sind ein großer Teilbereich der klassischen Statistik. Auch in der Bayes Statistik können Hypothesen darüber getestet werden, ob der wahre Wert des Parameters θ in einer bestimmten Menge $\Theta_0 \subset \Theta$ liegt. Es wird eine Nullhypothese der Form

$$H_0 : \theta \in \Theta_0$$

gegen eine Alternative

$$H_1 : \theta \in \Theta_1 = \Theta_0^c$$

getestet. Nicht immer muss $\Theta_0 \cup \Theta_1 = \Theta$ sein. Das Testproblem kann als spezielles Entscheidungsproblem aufgefasst werden. Man kann genau zwei Entscheidungen treffen, nämlich die H_0 akzeptieren oder sie verwerfen. Der Entscheidungsraum ist $\mathcal{D} = \{1, 0\}$. Jede Entscheidung $d \in \mathcal{D}$ ist ein Schätzer für die Indikatorfunktion $\mathbb{1}_{\Theta_0}(\theta)$. [Robert, 2007]

Als Verlustfunktion betrachten wir den gewichteten 0 – 1 Verlust, definiert durch

$$L(\theta, d) = \begin{cases} 0 & \text{falls } d = \mathbb{1}_{\Theta_0}(\theta), \\ a_0 & \text{falls } \theta \in \Theta_0 \text{ und } d = 0, \\ a_1 & \text{falls } \theta \notin \Theta_0 \text{ und } d = 1. \end{cases} \quad (2.4.1)$$

Die Werte von a_0 und a_1 geben an, wie groß der Verlust bei einem Fehler erster Art (fälschliches Verwerfen von H_0) und einem Fehler zweiter Art (fälschliches Annehmen der H_0) ist.

Satz 2.8. Für die Verlustfunktion (2.4.1) ist der Bayes Schätzer gegeben durch

$$\delta^\pi(x) = \begin{cases} 1 & \text{falls } P(\theta \in \Theta_0|x) > \frac{a_1}{a_0+a_1}, \\ 0 & \text{sonst.} \end{cases}$$

$P(\theta \in \Theta_0|x)$ bezeichnet die A-posteriori Wahrscheinlichkeit der Nullhypothese.

Beweis. Der A-posteriori zu erwartende Verlust ist

$$\begin{aligned} \mathbb{E}[L(\theta, d)|x] &= \int_{\Theta} L(\theta, d)\pi(\theta|x) d\theta \\ &= \int_{\Theta_0} a_0 \mathbb{1}_{\{0\}}(d)\pi(\theta|x) d\theta + \int_{\Theta_0^c} a_1 \mathbb{1}_{\{1\}}(d)\pi(\theta|x) d\theta \\ &= a_0 \mathbb{1}_{\{0\}}(d)P(\theta \in \Theta_0|x) + a_1 \mathbb{1}_{\{1\}}(d)P(\theta \notin \Theta_0|x). \end{aligned}$$

Die Entscheidung $d = 1$ minimiert das A-posteriori Risiko genau dann, wenn $\mathbb{E}[L(\theta, 1)|x] < \mathbb{E}[L(\theta, 0)|x]$ ist, oder wenn

$$a_1 P(\theta \notin \Theta_0|x) < a_0 P(\theta \in \Theta_0|x).$$

Das ist wiederum äquivalent zu

$$\frac{a_1}{a_0 + a_1} < P(\theta \in \Theta_0|x). \quad \square$$

Im Bayes'schen Sinne wird die Nullhypothese also verworfen, wenn die A-posteriori Wahrscheinlichkeit von H_0 kleiner ist als der Quotient $a_1/(a_0 + a_1)$, welcher in der Verlustfunktion festgelegt ist. [Robert, 2007]

Der Bayes Faktor ist die Teststatistik zum Vergleich der beiden Hypothesen und ist definiert als Quotient der A-posteriori Odds und der A-priori Odds der Hypothesen:

$$B_{01}(x) = \frac{P(\theta \in \Theta_0|x) \pi(\theta \in \Theta_1)}{P(\theta \in \Theta_1|x) \pi(\theta \in \Theta_0)}.$$

Sind π_0 und π_1 die A-priori Verteilungen unter H_0 und H_1 , so kann der Bayes Faktor geschrieben werden als

$$B_{01}(x) = \frac{\int_{\Theta_0} f(x|\theta) \pi_0(\theta) d\theta}{\int_{\Theta_1} f(x|\theta) \pi_1(\theta) d\theta} = \frac{m_0(x)}{m_1(x)}$$

und kann als bayes'scher Likelihoodquotient aufgefasst werden.

Satz 2.8 gibt die Bedingung an, unter der die H_0 akzeptiert wird:

$$P(\theta \in \Theta_0|x) > \frac{a_1}{a_0 + a_1}.$$

Man kann diese Bedingung umschreiben zu

$$\frac{a_0 P(\theta \in \Theta_0|x) \pi(\theta \in \Theta_0)}{a_1 P(\theta \in \Theta_1|x) \pi(\theta \in \Theta_1)} = \frac{a_0 m_0(x) \pi(\theta \in \Theta_0)}{a_1 m_1(x) \pi(\theta \in \Theta_1)} > 1.$$

Die Testentscheidung ist daher abhängig vom Verlustquotienten a_0/a_1 , dem Bayes Faktor und dem Quotienten der A-priori Wahrscheinlichkeiten von H_0 und H_1 .

Beim Test einer einpunktigen Nullhypothese der Form

$$H_0 : \theta = \theta_0 \quad \text{gegen} \quad H_1 : \theta \neq \theta_0$$

darf die A-priori von θ unter der Nullhypothese nicht stetig sein, denn sonst wäre die a-priori Wahrscheinlichkeit der H_0 gleich 0. Daher wird die A-priori für Θ_0 und Θ_1 getrennt definiert. Unter H_0 wird eine positive A-priori Wahrscheinlichkeit π_0 verwendet. Unter H_1 nimmt man eine A-priori Dichte $g_1(\theta)$ an, sodass die A-priori von θ aus einem diskreten und einem stetigen Teil besteht:

$$\pi(\theta) = \pi_0 \mathbb{1}_{\Theta_0}(\theta) + (1 - \pi_0) g_1(\theta) \mathbb{1}_{\Theta_1}(\theta).$$

Die A-posteriori von H_0 ist

$$\pi(\Theta_0|x) = \frac{f(x|\theta_0) \pi_0}{m(x)} = \frac{f(x|\theta_0) \pi_0}{f(x|\theta_0) \pi_0 + (1 - \pi_0) m_1(x)},$$

wobei $m_1(x) = \int_{\theta \neq \theta_0} f(x|\theta) g_1(\theta) d\theta$ die Randdichte von x unter g_1 ist. Der Bayes Faktor ist der Quotient der Likelihood bei θ_0 und der Randdichte $m_1(x)$ unter g_1 :

$$B_{01}(x) = \frac{f(x|\theta_0) \pi_0}{m_1(x) (1 - \pi_0)} \frac{(1 - \pi_0)}{\pi_0} = \frac{f(x|\theta_0)}{\int_{\theta \neq \theta_0} f(x|\theta) g_1(\theta) d\theta}.$$

Wenn man die zwei einpunktigen Hypothesen

$$H_0 : \theta = \theta_0 \quad \text{und} \quad H_1 : \theta = \theta_1$$

betrachtet, dann ist der Bayes Faktor gleich dem Quotienten der beiden Likelihoodfunktionen:

$$B_{01}(x) = \frac{f(x|\theta_0)}{f(x|\theta_1)}.$$

2.4.2 HPD-Bereiche

Neben der Punktschätzung und den Hypothesentests kann man auch in der Bayes Statistik einen Bereich angeben, in dem der Parameter mit einer möglichst hohen Wahrscheinlichkeit liegt. Unter allen solchen Bereichen wählt man jenen mit möglichst minimaler Größe oder Volumen.

Definition 2.11. Eine Menge $C_x \subset \Theta$ heißt $100(1 - \alpha)\%$ Vertrauensbereich, wenn

$$P(\theta \in C_x | x) = \int_{C_x} \pi(\theta | x) d\theta \geq 1 - \alpha.$$

Da in der Bayes Statistik der Parameter eine Wahrscheinlichkeitsverteilung hat, ist es sinnvoll, davon zu sprechen, dass θ mit einer gegebenen Wahrscheinlichkeit in einem festen Bereich C_x liegt. Das ist ein großer Unterschied zur frequentistischen Definition von Konfidenzbereichen, wo ein zufälliger Bereich eine gegebene Wahrscheinlichkeit hat, den unbekanntem Parameter zu enthalten. [Robert, 2007]

Definition 2.12. Ein $100(1 - \alpha)\%$ Vertrauensbereich C_x heißt Bereich mit der höchsten A-posteriori Dichte oder HPD-Bereich (engl. *highest posterior density region*), wenn

$$C_x = \{\theta | \pi(\theta | x) \geq k_\alpha\},$$

wobei k_α die größte Konstante ist, sodass $P(\theta \in C_x | x) \geq 1 - \alpha$.

Beispiel 2.9 (Normalverteilung). Bei einer Stichprobe aus einer Normalverteilung mit bekannter Varianz σ_0^2 und einer konjugierten $N(m, d^2)$ A-priori ist die A-posteriori wieder eine Normalverteilung mit Mittelwert

$$m^* = \frac{m\sigma_0^2 + nd^2\bar{x}}{\sigma_0^2 + nd^2}$$

und Varianz $d^{*2} = d^2\sigma_0^2/(\sigma_0^2 + nd^2)$. Das HPD-Intervall C_x^α mit Überdeckungswahrscheinlichkeit $1 - \alpha$ ist symmetrisch um den A-posteriori Mittelwert

$$C_x^\alpha = [m^* - k_\alpha d^*, m^* + k_\alpha d^*]$$

mit dem $\alpha/2$ -Quantil der Standardnormalverteilung k_α .

Unter Verwendung der Jeffreys A-priori ($d^2 \rightarrow \infty$) entspricht das HPD-Intervall dem klassischen $100(1 - \alpha)\%$ Konfidenzintervall:

$$C_x^\alpha = \left[\bar{x} - k_\alpha \frac{\sigma_0}{\sqrt{n}}, \bar{x} + k_\alpha \frac{\sigma_0}{\sqrt{n}} \right]. \quad \diamond$$

Beispiel 2.10 (Exponentialverteilung). Sei $X \sim \text{Exp}(\theta)$. Ein symmetrisches $100(1 - \alpha)\%$ Vertrauensintervall (θ_L, θ_U) für die Rate θ erfüllt

$$\int_0^{\theta_L} \pi(\theta | x) d\theta = \int_{\theta_U}^{\infty} \pi(\theta | x) d\theta = \frac{\alpha}{2}.$$

Ein $100(1 - \alpha)\%$ HPD Intervall für θ erfüllt $\pi(\theta_L|x) = \pi(\theta_U|x)$ und

$$\int_{\theta_L}^{\theta_U} \pi(\theta|x) d\theta = 1 - \alpha.$$

Konkret kann das HPD Intervall durch numerisches Lösen der beiden Gleichungen berechnet werden. \diamond

3 | Zuverlässigkeitstheorie

Zuverlässigkeit ist definiert als „die Wahrscheinlichkeit, dass ein Bauteil (eine Komponente, ein System) seine Funktion über einen bestimmten Zeitraum unter Betriebsbedingungen in angemessener Weise erfüllt.“ [Mann et al., 1974]

Komponenten der Zuverlässigkeits- oder Lebensdaueranalyse sind immer ein Zeitpunkt, ab dem eine Lebensdauer erfasst wird und ein Ereignis, dessen Eintreten das Ende dieser bestimmt. Startzeitpunkt kann z.B. der Betriebsbeginn eines Systems oder der Zeitpunkt des Eintrittes eines Patienten in eine klinische Studie sein. Das Ereignis kann ein Fehler oder der Ausfall des Systems, das Auftreten eines Tumors oder der Tod des Patienten sein. Die Methoden der Zuverlässigkeitsanalyse finden neben den Bereichen Industrie und Medizin auch in der Ökonomie und in den Sozialwissenschaften ihre Anwendung, z.B. zur Analyse der Dauer von Arbeitslosigkeit bei bestimmten Personengruppen. Das Ziel ist es, mittels stochastischer Modelle die Verteilung der Lebensdauer zu beschreiben und Parameter und Kennzahlen dieser Verteilung zu schätzen, um damit das zugrunde liegende System besser zu verstehen und Vorhersagen treffen zu können.

Eine Besonderheit von Daten aus Zuverlässigkeitsversuchen ist, dass der exakte Eintrittszeitpunkt des Ereignisses oft nicht beobachtet wird. Man spricht dann von zensierten Daten. Dies muss in der statistischen Analyse berücksichtigt werden. Des Weiteren sind Lebensdauern in der Regel nicht normalverteilt, weshalb andere Verteilungsmodelle zur Anwendung kommen. Dieses Kapitel behandelt einige grundlegende Konzepte der Zuverlässigkeitstheorie und basiert im Wesentlichen auf den Büchern [Mann et al., 1974], [Lawless, 2011] und [Meeker and Escobar, 2014]. Neben allgemeinen Definitionen werden einige konkrete Lebensdauerverteilungen sowie Modelle zu beschleunigten Lebensdauerexperimenten vorgestellt.

3.1 Konzepte der Zuverlässigkeitstheorie

Die Lebensdauer ist die Zeit vom Betriebsbeginn bis zum Ausfall eines Systems oder einer Komponente und ist eine nicht negative stetige Zufallsgröße T^1 . Die Verteilungsfunktion F gibt die Wahrscheinlichkeit an, dass bis zum Zeitpunkt $t, t \geq 0$ ein Ausfall passiert:

$$F(t) = P(T \leq t).$$

¹ T kann auch eine diskrete Zufallsvariable sein, etwa wenn ein System in Zyklen läuft und die Anzahl der durchlaufenen Zyklen vor dem Systemausfall gemessen wird. Die Konzepte in diesem Abschnitt können genauso für diskrete Zufallsvariablen definiert werden. In dieser Arbeit liegt der Fokus jedoch auf stetigen T .

Die Zuverlässigkeitsfunktion R oder Überlebensfunktion (engl. *reliability function*, *survival function*) ist das Komplement der Verteilungsfunktion:

$$R(t) = 1 - F(t) = P(T > t).$$

Definition 3.1 (Ausfallrate). Die Ausfallrate oder Fehlerrate (engl. *hazard rate*) ist definiert durch

$$h(t) = \frac{f(t)}{1 - F(t)}.$$

Dabei bezeichnet $f(t)$ die Dichtefunktion von T .

Die Ausfallrate wird auch als momentane Ausfallneigung bezeichnet. Es gibt einen Zusammenhang zwischen $h(t)$ und der bedingten Wahrscheinlichkeit, dass eine Komponente in einem kleinen Zeitintervall $(t, t + \Delta t)$ ausfällt, wenn sie bis zum Zeitpunkt t noch nicht ausgefallen ist. Es gilt

$$h(t)\Delta t \sim P(t < T \leq t + \Delta t | T > t) \text{ für } \Delta t \rightarrow 0,$$

denn:

$$\begin{aligned} h(t) &= \frac{f(t)}{1 - F(t)} \\ &= \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t(1 - F(t))} \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t)}{\Delta t(1 - F(t))} \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t)}{\Delta t P(T > t)} \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t}. \end{aligned}$$

Man kann daher $h(t)$ als die bedingte Dichte von T unter $T > t$ auffassen.

Das Konzept der Ausfallrate wird typischerweise als „Badewannenkurve“ mit drei charakteristischen Fehlertypen illustriert (Abb. 3.1). In der ersten Phase nach Betriebsbeginn (*early life*) treten gehäuft Frühausfälle auf, die mit der Zeit abnehmen. Das ist durch eine fallende Ausfallrate beschrieben. In der zweiten Phase ist die Ausfallrate konstant. Ausfälle treten zufällig durch externe Belastungen auf (*random failures*). Die dritte Phase gegen Ende der Betriebsdauer einer Komponente (*wear-out*) ist durch eine steigende Ausfallrate gekennzeichnet. Wegen zunehmender Abnutzung oder einer Häufung von externen Belastungen treten Ausfälle mit steigender Rate auf.

Die Verteilung von T kann durch jede der Funktionen $F(t)$, $f(t)$ und $h(t)$ spezifiziert werden. Aus

$$\frac{d \log R(t)}{dt} = \frac{dR(t)/dt}{R(t)} = -\frac{f(t)}{R(t)}$$

folgt

$$h(t) = \frac{f(t)}{R(t)} = -\frac{d \log R(t)}{dt}.$$

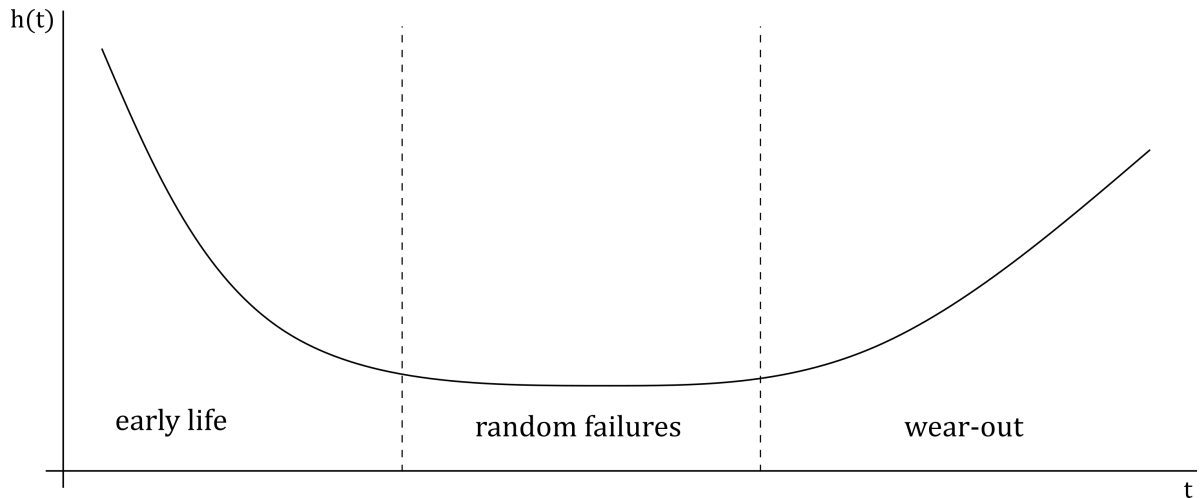


Abbildung 3.1: Badewannenkurve

Damit gilt für das Integral von $h(t)$

$$\int_0^t h(u) \, du = -\log R(t),$$

und somit

$$R(t) = \exp\left(-\int_0^t h(u) \, du\right) = \exp(-H(t)).$$

Die Funktion $H(t)$ wird als kumulierte Ausfallrate bezeichnet. Für die Dichtefunktion von T folgt nun aus Definition 3.1, dass

$$f(t) = h(t) \exp\left(-\int_0^t h(u) \, du\right) = h(t) \exp(-H(t)). \quad (3.1.1)$$

Häufig sind bestimmte Kennzahlen der Lebensdauerverteilung von Interesse, wie z.B. die mittlere Lebensdauer (*mean time to failure*, MTTF)

$$\mathbb{E}(T) = \int_0^{\infty} t f(t) \, dt = \int_0^{\infty} 1 - F(t) \, dt,$$

oder das γ -Quantil t_γ der Lebensdauerverteilung, das definiert ist durch $t_\gamma = F^{-1}(\gamma)$, wobei F^{-1} bei nicht absolut stetiger Verteilungsfunktion die verallgemeinerte Inverse von F bezeichnet:

$$F^{-1} : [0, 1] \rightarrow \mathbb{R} \\ p \mapsto \inf\{t \in \mathbb{R} \mid F(t) \geq p\}.$$

3.1.1 Zensierte Daten

In Zuverlässigkeitsversuchen werden die Daten in der Regel nicht vollständig beobachtet. Aufgrund von praktischen oder monetären Einschränkungen ist ein Experiment nicht bis zu dem Zeitpunkt durchführbar, zu dem alle Einheiten ausgefallen sind. Nur für einen Teil der getesteten Einheiten werden exakte Ausfallzeiten beobachtet, die restlichen Einheiten sind zensiert. Dabei können verschiedene Zensierungstypen auftreten.

Typ I-Zensierung: Zensierte Daten von diesem Typ treten dann auf, wenn der Versuch zu einem vorher festgelegten Zeitpunkt endet, zu dem eventuell noch nicht alle Versuchseinheiten ausgefallen sind. Die Anzahl der ausgefallenen Einheiten sowie die Ausfallzeiten sind stochastische Größen.

In bestimmten Anwendungsbereichen, z.B. in medizinischen Studien, hat jede Einheit eine eigene Zensurzeit. Dies kommt dadurch zustande, dass Patienten zu verschiedenen Zeitpunkten in eine Studie eintreten und unterschiedlich lange beobachtet werden. Die Zensurzeit kann auch eine stochastische Größe sein, man spricht dann von zufälliger Typ I-Zensierung. Diese Spezialfälle werden z.B. in [Lawless, 2011] ausführlicher behandelt.

Typ II-Zensierung: Wird im Vorhinein festgelegt, dass der Versuch nach Beobachtung der ersten r Ausfälle endet, spricht man von Typ II-Zensierung. Bei diesem Typ sind nur die Ausfallzeiten stochastische Größen, da die Anzahl der ausgefallenen Einheiten schon zu Beginn des Experimentes feststeht. [Mann et al., 1974]

Des Weiteren unterscheidet man rechts-, links- und intervallzensierte Daten. Rechtszensierte Daten sind dadurch gekennzeichnet, dass das Ereignis nach Versuchsende noch nicht eingetreten ist. Es ist also nur eine untere Grenze der Ausfallzeit bekannt, die so genannte Zensurzeit. Bei linkszensierten Daten ist nur bekannt, dass das Ereignis vor einem bestimmten Zeitpunkt eingetreten ist. In vielen Situationen können Ausfälle nur zu bestimmten Inspektionszeiten beobachtet werden. Dann weiß man über die Ausfallzeit nur, dass sie in einem Intervall $[t_l, t_u]$ liegt und spricht von intervallzensierten Daten.

Da bei Zuverlässigkeitsversuchen in der Regel rechtszensierte Daten auftreten, deren Zensurzeit für alle Versuchseinheiten gleich ist, wird dieser Zensierungstyp in weiterer Folge schwerpunktmäßig behandelt.

3.1.2 Likelihood bei zensierten Daten

Die Form der Likelihood ist abhängig vom gewählten Datenmodell und der Art der Zensierung. Die Lebensdauer von n Versuchseinheiten wird durch die Zufallsvariablen T_1, \dots, T_n repräsentiert. Eine Beobachtung ist entweder eine exakte Ausfallzeit oder zensiert. Die Zensurzeit $C > 0$ ist bei Typ I-Zensierung fest, für alle Beobachtungen gleich und unabhängig von T . Man definiert den Statusindikator δ_i , der angibt, ob es sich um eine exakte ($\delta_i = 1$) oder eine zensierte ($\delta_i = 0$) Ausfallzeit handelt. Dann sind die Beobachtungen Paare der Form $(x_1, \delta_1), \dots, (x_n, \delta_n)$ mit $X_i = \min(T_i, C)$ und $\Delta_i = \mathbb{1}_{T_i \leq C}$.

Im folgenden Satz wird die gemeinsame Dichte von (X, Δ) hergeleitet, wobei Δ die Zufallsvariable und δ eine Realisierung davon bezeichnet (vgl. [Lawless, 2011]).

Satz 3.1. *Bei Typ I-zensierten Daten mit gemeinsamer fester Zensurzeit C ist die gemeinsame Dichte von (X, Δ) gegeben durch*

$$f(x, \delta) = f(x)^\delta [1 - F(C)]^{1-\delta}$$

und wird als Dichte bezüglich des Produktes aus dem Lebesgue-Maß auf \mathbb{R}^+ und dem Zählmaß auf $\{0, 1\}$ verstanden.

Beweis. Die Zufallsvariable T ist absolut stetig bezüglich des Lebesgue-Maßes und hat eine Dichte $f(t)$. Der Indikator Δ hat eine diskrete Verteilung mit einer Dichte bzw. Wahrscheinlichkeitsfunktion p bezüglich des Zählmaßes auf $\{0, 1\}$. Die gemeinsame Verteilung von (X, Δ) hat eine Dichte bezüglich des Produktes aus Lebesgue-Maß und Zählmaß. Für die gemeinsame Dichte gilt

$$f(x, \delta) = f(x|\delta)p(\delta).$$

Im Fall $\delta = 1$ sind die Ausfallzeiten exakt. Daher gilt $X = T$ und $p(1) = P(T \leq C) = F(C)$. Daraus folgt

$$f(x|\delta = 1) = \frac{f(x)}{p(1)} = \frac{f(x)}{F(C)}$$

und

$$f(x, 1) = f(x|\delta = 1)p(1) = f(x).$$

Unter $\delta = 0$ sind die Beobachtungen zensiert und daher ist $X = C$ mit Wahrscheinlichkeit 1. Daraus folgt

$$f(x, 0) = f(x|\delta = 0)p(0) = p(0) = P(T > C) = 1 - F(C).$$

Insgesamt erhält man für die gemeinsame Dichte die Darstellung

$$f(x, \delta) = f(x)^\delta [1 - F(C)]^{1-\delta}. \quad \square$$

Die Likelihood von n unabhängigen Beobachtungen kann nach dem obigen Satz wie folgt dargestellt werden

$$l = \prod_{i=1}^n f(t_i)^{\delta_i} [1 - F(C)]^{\delta_i}. \quad (3.1.2)$$

Dabei ist zu beachten, dass $x = t$ ist, wenn $\delta = 1$. Nimmt man an, dass es $r \leq n$ exakte Ausfallzeiten und dementsprechend $n - r$ rechtszensierte Beobachtungen gibt, so kann man (3.1.2) auch schreiben als

$$l = \prod_{i=1}^r f(t_i) [1 - F(C)]^{n-r}.$$

Bei Typ II-zensierten Daten werden n Einheiten so lange getestet, bis die ersten $r \leq n$ Einheiten ausgefallen sind. Die Beobachtungen bestehen aus den ersten r Ordnungsstatistiken $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(r)}$. Die gemeinsame Zensurzeit der $n - r$ nicht ausgefallenen Einheiten ist $t_{(r)}$ und hängt von den Daten ab. Die Likelihoodfunktion ergibt sich aus der gemeinsamen Dichte der ersten r Ordnungsstatistiken, welche die folgende Form hat (vgl. [Casella and Berger, 2002])

$$l = \frac{n!}{(n-r)!} \prod_{i=1}^r f(t_i) [1 - F(t_{(r)})]^{n-r}.$$

Diese Likelihood kann auch in der Form (3.1.2) dargestellt werden, wenn man die Konstante $n!/(n-r)!$ weglässt und für die zensierten Beobachtungen $\delta_i = 0$ und $C = t_{(r)}$ setzt. [Lawless, 2011]

3.2 Lebensdauerverteilungen

Viele Probleme in der Zuverlässigkeitsanalyse lassen sich durch parametrische Modelle beschreiben. Die Lebensdauerverteilung ist dabei von der Form $F(t|\theta)$, wobei $\theta = (\theta_1, \dots, \theta_n)$ der Parametervektor ist. Der Vorteil von parametrischen Modellen ist, dass, im Gegensatz zu nichtparametrischen Modellen, nur eine endliche Anzahl von Parametern geschätzt werden muss. In diesem Abschnitt werden einige häufig verwendete Verteilungsfamilien vorgestellt.

3.2.1 Exponentialverteilung

Die Exponentialverteilung hat die Verteilungsfunktion

$$F(t|\lambda) = 1 - e^{-\lambda t}, \quad t \geq 0, \lambda > 0.$$

Die Dichtefunktion ist gegeben durch

$$f(t|\lambda) = \lambda e^{-\lambda t}, \quad t \geq 0, \lambda > 0.$$

Die Exponentialverteilung hat die konstante Ausfallrate $h(t) = \lambda$. Dadurch ist ihre Anwendung in der Praxis eingeschränkt auf die Modellierung von Systemen, bei denen keine Abnutzung oder Alterung auftritt.

Ist der Fehlermechanismus, der einem System zugrunde liegt, so dass die Ausfallrate eine Konstante λ ist, dann erhält man aus (3.1.1) genau die Dichtefunktion der Exponentialverteilung:

$$f(t) = \lambda \exp\left(-\int_0^t \lambda du\right) = \lambda e^{-\lambda t}, \quad t \geq 0.$$

Man kann die Exponentialverteilung auch über die Beziehung zur Poissonverteilung als Lebensdauerverteilung herleiten (vgl. [Mann et al., 1974]). Angenommen, in der Umgebung, in der ein System betrieben wird, treten zufällige externe Schocks auf. Diese Schocks seien poissonverteilt mit Rate λ . Sobald ein Schock auftritt, fällt das System aus. Die Zufallsvariable T bezeichnet die Zeit zwischen zwei aufeinanderfolgenden Schocks. Für die Verteilung von T gilt dann

$$P(T > t) = P(\text{kein Schock im Intervall } (0, t)) = e^{-\lambda t},$$

also ist T exponentialverteilt.

Parameterschätzung

In einem Lebensdauerexperiment werden n Einheiten unter Bedingungen, die möglichst den realen Betriebsbedingungen entsprechen, getestet.

Typ II-Zensierung: Der Versuch endet, nachdem die ersten $r \leq n$ Einheiten ausgefallen sind. Es werden also die kleinsten r Ausfallzeiten beobachtet, die restlichen $n - r$ Zeiten sind rechtszensiert. Die der Größe nach geordneten Ausfallzeiten sind $0 \leq t_1 \leq \dots \leq t_r$. Die

Zensurzeit ist die größte beobachtete Ausfallzeit t_r . Die Likelihoodfunktion für λ ist

$$\begin{aligned}
 l(\lambda|t_1, \dots, t_r) &= \prod_{i=1}^r f(t_i|\lambda) [1 - F(t_r|\lambda)]^{n-r} \\
 &= \prod_{i=1}^r \lambda \exp(-\lambda t_i) [\exp(-\lambda t_r)]^{n-r} \\
 &= \lambda^r \exp\left(-\lambda \sum_{i=1}^r t_i\right) \exp(-\lambda(n-r)t_r) \\
 &= \lambda^r \exp\left(-\lambda \left(\sum_{i=1}^r t_i + (n-r)t_r\right)\right). \tag{3.2.1}
 \end{aligned}$$

Die logarithmierte Likelihood ist

$$\log l(\lambda|t_1, \dots, t_r) = r \log \lambda - \lambda \left(\sum_{i=1}^r t_i + (n-r)t_r\right). \tag{3.2.2}$$

Ableiten und Null setzen von (3.2.2) ergibt den folgenden Maximum Likelihood Schätzer für λ ,

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^r t_i + (n-r)t_r}.$$

Im Fall $n = r$, d.h. wenn alle n Einheiten ausgefallen sind, ist der ML-Schätzer

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n t_i}.$$

Eine alternative Parametrisierung der Exponentialverteilung erhält man mit $\theta = 1/\lambda$. Der Parameter θ ist die mittlere Lebensdauer. Aufgrund der Invarianzeigenschaft ist der Maximum-Likelihood-Schätzer für θ gegeben durch

$$\hat{\theta} = \frac{1}{\hat{\lambda}} = \frac{\sum_{i=1}^r t_i + (n-r)t_r}{r}.$$

Um die Verteilung von $\hat{\theta}$ herzuleiten, wird die gemeinsame Dichte der ersten r Ordnungsstatistiken $T_{(1)}, \dots, T_{(r)}$ wie folgt geschrieben:

$$l(\theta|t_1, \dots, t_r) = \prod_{i=1}^r \frac{n-i+1}{\theta} \exp\left(-\frac{(n-i+1)(t_i - t_{i-1})}{\theta}\right),$$

$0 = t_0 \leq t_1 \leq \dots \leq t_r$. Für jedes $i \in \{1, \dots, r\}$ ist $(n-i+1)(t_i - t_{i-1})/\theta$ exponentialverteilt mit Mittelwert 1. Daher folgt die Größe

$$\frac{r\hat{\theta}}{\theta} = \frac{\sum_{i=1}^r (n-i+1)(t_i - t_{i-1})}{\theta}$$

als Summe von r exponentialverteilten Zufallsgrößen einer $\gamma(r, 1)$ -Verteilung und es gilt $\hat{\theta} \sim \gamma(r, r/\theta)$. Wegen $\hat{\lambda} = 1/\hat{\theta}$ hat $\hat{\lambda}$ eine inverse Gammaverteilung $\gamma^{-1}(r, r\lambda)^2$.

Aufgrund der Beziehung zwischen Gamma- und Chiquadratverteilung³ gilt

$$\frac{2\hat{\theta}r}{\theta} \sim \chi_{2r}^2.$$

Damit kann man ein $100(1 - \alpha)\%$ Konfidenzintervall für θ durch

$$[\theta_l, \theta_u] = \left[\frac{2\hat{\theta}r}{\chi_{1-\alpha/2, 2r}^2}, \frac{2\hat{\theta}r}{\chi_{\alpha/2, 2r}^2} \right] \quad (3.2.3)$$

angeben [Meeker and Escobar, 2014].

Typ I-Zensierung: In diesem Fall endet der Versuch nach einer im Vorhinein festgelegten Zeit t_0 . Angenommen, es werden bis zu diesem Zeitpunkt r Ausfälle mit Ausfallzeiten t_1, \dots, t_r beobachtet. Die log-Likelihood ist dann gegeben durch

$$\log l(\lambda | t_1, \dots, t_r) = r \log \lambda - \lambda \left(\sum_{i=1}^r t_i + (n - r)t_0 \right),$$

und der Maximum Likelihood Schätzer für λ ist

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^r t_i + (n - r)t_0}.$$

Der Ausdruck im Nenner, $\sum_{i=1}^r t_i + (n - r)t_0$, wird auch als „total time on test“ bezeichnet und ist die Summe der beobachteten Lebensdauern und der Zensurzeiten von allen n getesteten Einheiten.

Ein interessanter Spezialfall ist der Fall $r = 0$, in dem bis zum Zeitpunkt t_0 keine Ausfälle beobachtet werden. Dieser Fall tritt bei Tests von hochzuverlässigen Bauteilen in der Praxis häufig auf. Die log-Likelihood ist dann

$$\log l(\lambda | t_1, \dots, t_r) = -\lambda n t_0.$$

Der ML-Schätzer für λ (θ) ist $\hat{\lambda} = 0$ ($\hat{\theta} = \infty$). Da man davon ausgehen kann, dass bei Fortsetzung des Testes noch Ausfälle auftreten würden, sind diese Schätzer in der Praxis nicht brauchbar. Man kann jedoch eine untere Grenze für das $100(1 - \alpha)\%$ -Konfidenzintervall für λ (oder θ) angeben. Zur Motivation betrachtet man einen Typ I-zensierten Versuch, in dem jede ausgefallene Einheit sofort ersetzt wird (*testing with replacement*). Dann sind die beobachteten Ausfallzeiten die Zeiten zwischen den Ereignissen in einem Poisson-Prozess mit Rate λ (vgl. [Mann et al., 1974, Kapitel 5], [Lawless, 2011, Kapitel 4]). Die Anzahl der bis zur „total time on test“ nt_0 beobachteten Ausfälle ist poissonverteilt mit Mittelwert $\lambda n t_0$.

²Die inverse Gammaverteilung $\gamma^{-1}(\alpha, \beta)$ hat die Dichtefunktion

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right), \quad \alpha, \beta > 0.$$

Es gilt die Beziehung zur Gammaverteilung (Abschnitt 3.2.4): Ist $X \sim \gamma(\alpha, \beta)$, dann ist $1/X \sim \gamma^{-1}(\alpha, \beta)$.

³Ist $X \sim \gamma(r/2, 2)$, dann gilt $X \sim \chi^2(r)$ (vgl. [Casella and Berger, 2002] für eine Darstellung der Beziehungen zwischen gängigen Verteilungen).

Die Wahrscheinlichkeit, höchstens r Ausfälle im Intervall $[0, nt_0]$ zu beobachten, ist wegen der Beziehung zwischen Poisson- und Chiquadratverteilung gleich der Wahrscheinlichkeit, dass eine χ_{2r+2}^2 -verteilte Zufallsvariable größer als $2\lambda nt_0$ ist [Mann et al., 1974]. Damit ergibt sich eine konservative untere Grenze für ein $100(1 - \alpha)\%$ Konfidenzintervall für θ als

$$\theta_l = \frac{2nt_0}{\chi_{1-\alpha/2, 2r+2}^2}.$$

Im Fall $r = 0$ ergibt sich wegen $\chi_{1-\alpha/2, 2}^2 = -2 \log \alpha$ die untere Grenze für θ oder λ , die in [Meeker and Escobar, 2014] angeführt ist:

$$\theta_l = \frac{nt_0}{-\log \alpha} \quad \text{oder} \quad \lambda_l = \frac{-\log \alpha}{nt_0}.$$

3.2.2 Weibullverteilung

Die Weibullverteilung ist die am weitesten verbreitete parametrische Lebensdauerverteilung. Sie stellt eine Verallgemeinerung der Exponentialverteilung dar, die durch ihre Flexibilität in vielen Situationen Anwendung findet. In [Mann et al., 1974] werden unter anderem die Analyse von elektronischen Komponenten, Kugellagern und Halbleiterbauteilen als Anwendungsbeispiele genannt.

Die Verteilungsfunktion, Dichte und Ausfallrate der Weibullverteilung sind für $t \geq 0$ und $\eta, \beta > 0$ gegeben durch

$$\begin{aligned} F(t) &= 1 - \exp\left(-\left(\frac{t}{\eta}\right)^\beta\right), \\ f(t) &= \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{t}{\eta}\right)^\beta\right), \\ h(t) &= \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}. \end{aligned}$$

Der Parameter η ist der Skalenparameter und β der Formparameter der Verteilung. Mit der Weibullverteilung lassen sich konstante, steigende und fallende Ausfallraten modellieren. Für $\beta = 1$ erhält man die Exponentialverteilung mit konstanter Ausfallrate $1/\eta$. Ist $\beta > 1$, so ist die Ausfallrate steigend, bei $\beta < 1$ ist sie fallend. In Abbildung 3.2 ist die Ausfallrate einer Weibullverteilung mit $\eta = 1$ für verschiedene Werte von β dargestellt.

Für $t = \eta$ gilt $F(t) = 1 - e^{-1} \approx 0.6321$. Der Parameter η ist jener Zeitpunkt, an dem in etwa 63.21% der Einheiten ausgefallen sind und wird auch charakteristische Lebensdauer (engl. *characteristic life*) genannt.

Die Weibullverteilung kann durch einen Garantieparameter (*threshold parameter*) verallgemeinert werden. Dieser Parameter kann als Zeitpunkt interpretiert werden, vor dem garantiert kein Ausfall auftreten kann. Die Verteilungsfunktion hat die allgemeinere Form

$$F(t) = 1 - \exp\left(-\left(\frac{t - \gamma}{\eta}\right)^\beta\right), \quad \eta, \beta > 0, \gamma \geq 0, t \geq \gamma.$$

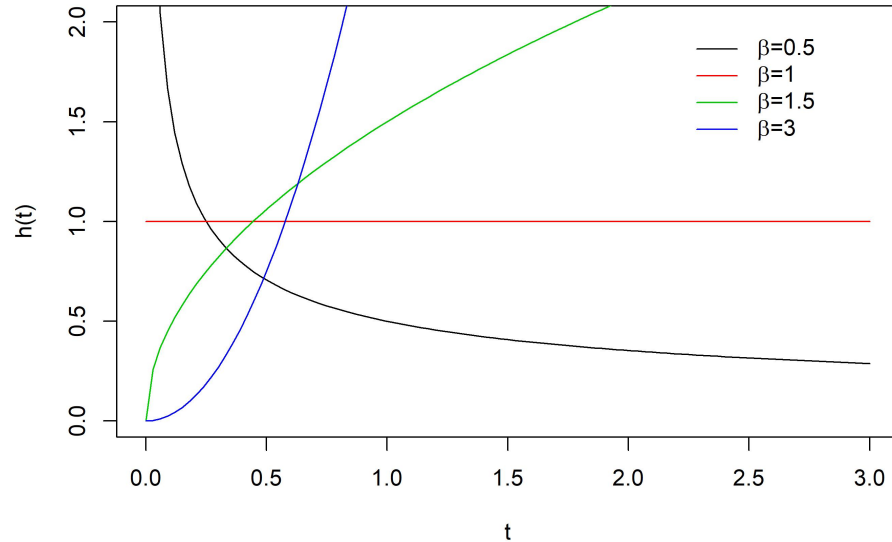


Abbildung 3.2: Die Ausfallrate $h(t)$ der Weibullverteilung mit $\eta = 1$ und verschiedenen Werten für β .

Parameterschätzung

Falls der Formparameter β bekannt ist, ist die Zufallsvariable T^β exponentialverteilt mit $1/\lambda = \eta^\beta$ und η kann, wie in Abschnitt 3.2.1 beschrieben, geschätzt werden.

Sind beide Parameter unbekannt, lassen sich Schätzer mit einer grafischen Methode gewinnen. Das p -Quantil der Weibullverteilung, t_p , ist gegeben durch

$$t_p = \eta(-\log(1-p))^{1/\beta}.$$

Durch Logarithmieren kommt man auf eine lineare Gleichung in $\log t_p$:

$$\log t_p = \log \eta + \frac{1}{\beta} \log(-\log(1-p)).$$

Trägt man in einer Grafik $\log(-\log(1-p))$ gegen $\log t_p$ auf, so ergibt sich bei einer Weibullverteilung näherungsweise eine Gerade, deren Steigung ein Schätzer für β ist. Einen Schätzer für η erhält man durch das $p = 0.6321$ -Quantil der Stichprobe, also dem Schnittpunkt mit der horizontalen Gerade bei $p = 0.6321$.

Bei unbekanntem β und η hat die Likelihood bei r beobachteten Ausfällen und $n - r$ rechtszensierten Beobachtungen mit Zensurzeit t_s die Form

$$l(\eta, \beta | t_1, \dots, t_r) = \prod_{i=1}^r \frac{\beta}{\eta} \left(\frac{t_i}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{t_i}{\eta}\right)^\beta\right) \left[\exp\left(-\left(\frac{t_s}{\eta}\right)^\beta\right)\right]^{n-r}.$$

Bei Typ I-Zensurung ist $t_s = t_0$ und bei Typ II-Zensurung ist $t_s = t_{(r)}$. Die log-Likelihood ist

$$\begin{aligned} \log l(\eta, \beta | t_1, \dots, t_r) = & r \log \beta - r \beta \log \eta - \sum_{i=1}^r \left(\frac{t_i}{\eta} \right)^\beta + \\ & (\beta - 1) \sum_{i=1}^r \log t_i - (n - r) \left(\frac{t_s}{\eta} \right)^\beta. \end{aligned} \quad (3.2.4)$$

Durch Ableiten nach η und β und Null setzen kommt man auf die folgenden beiden Gleichungen, die mit numerischen Methoden wie z.B. dem Newton-Raphson-Verfahren gelöst werden können:

$$\eta^\beta = \frac{1}{r} \left(\sum_{i=1}^r t_i^\beta + (n - r) t_s^\beta \right)$$

und

$$\left(\frac{\sum_{i=1}^r t_i^\beta \log t_i + (n - r) t_s^\beta \log t_s}{\sum_{i=1}^r t_i^\beta + (n - r) t_s^\beta} - \frac{1}{\beta} \right) - \frac{1}{r} \sum_{i=1}^r \log t_i = 0.$$

Beispiel 3.1. Für dieses Beispiel wurden $n = 100$ Beobachtungen aus einer Weibullverteilung mit $\eta = 1$ und Formparameter $\beta = 3$ und gemeinsamer Zensurzeit $t_0 = 1.2$ simuliert. Es gibt $r = 82$ beobachtete Ausfallzeiten und 18 rechtszensierte Beobachtungen. Die Log-Likelihood (3.2.4) kann mit dem Softwarepaket R und der Funktion `optim` direkt maximiert werden. Die ML-Schätzer für η und β sind $\hat{\eta} = 1.003$ und $\hat{\beta} = 3.08$. \diamond

3.2.3 Log-Normalverteilung

Die Lebensdauer T ist log-normalverteilt, $T \sim LN(\mu, \sigma^2)$, falls $X = \log T$ normalverteilt ist. Die Dichtefunktion von $X \sim N(\mu, \sigma^2)$ ist

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) = \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right), \quad x \in \mathbb{R},$$

wobei φ die Dichtefunktion der Standardnormalverteilung $N(0, 1)$ bezeichnet. Für T erhält man aus dem Transformationssatz für Dichten

$$f_T(t) = \frac{1}{\sqrt{2\pi\sigma^2}t} \exp\left(-\frac{(\log t - \mu)^2}{2\sigma^2}\right) = \frac{1}{\sigma t} \varphi\left(\frac{\log t - \mu}{\sigma}\right), \quad t > 0.$$

Die Verteilungsfunktion von T ist

$$F(t) = \Phi\left(\frac{\log t - \mu}{\sigma}\right), \quad t > 0,$$

mit der Verteilungsfunktion der Standardnormalverteilung Φ .

Die Ausfallrate $h(t)$ einer Log-Normalverteilung mit $\mu = 0$ und verschiedenen Werten für σ ist in Abbildung 3.3 dargestellt. Sie ist als Funktion von t steigend, bis sie ein Maximum erreicht, dann fallend und es gilt $\lim_{t \rightarrow \infty} h(t) = 0$ [Lawless, 2011, S. 22]. Anwendungen findet diese Verteilung z.B. bei der Modellierung von Ausfällen, die durch Materialermüdung entstehen.

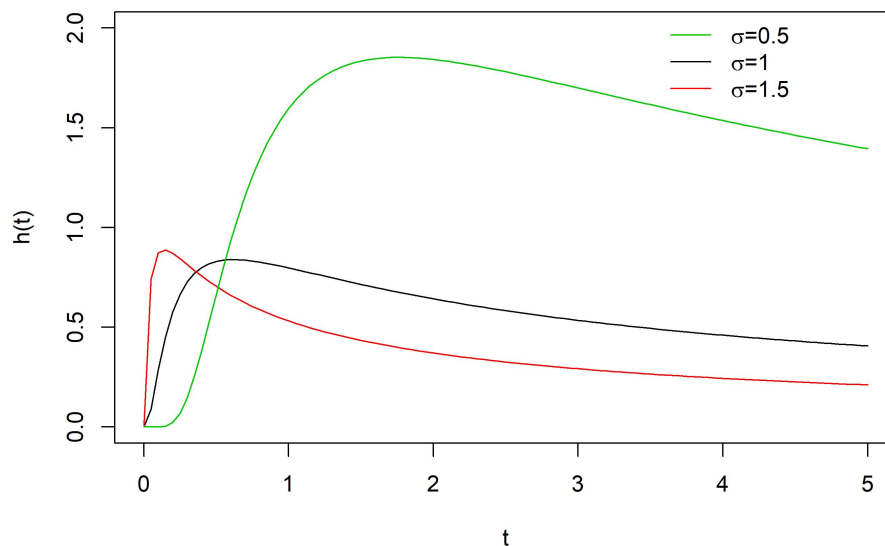


Abbildung 3.3: Die Ausfallrate $h(t)$ der Log-Normalverteilung mit $\mu = 0$ und verschiedenen Werten für σ .

Parameterschätzung

Bei vollständigen Daten sind die Maximum-Likelihood Schätzer für μ und σ^2 wegen der Beziehung zur Normalverteilung gegeben durch

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \log t_i \quad \text{und} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\log t_i - \hat{\mu})^2.$$

Wenn $r < n$ Ausfallzeiten beobachtet werden und die restlichen $n - r$ Beobachtungen zensiert sind mit gemeinsamer Zensurzeit t_s , dann ist die Likelihood gegeben durch

$$l(\mu, \sigma^2 | t_1, \dots, t_r) = \frac{1}{(2\pi\sigma^2)^{r/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^r (\log t_i - \mu)^2\right) \prod_{i=1}^r \frac{1}{t_i} \times \left[1 - \Phi\left(\frac{\log t_s - \mu}{\sigma}\right)\right]^{n-r}.$$

Die ML-Schätzer haben keine geschlossene Form und müssen numerisch bestimmt werden.

3.2.4 Gammaverteilung

Die Gammaverteilung ist eine weitere Verallgemeinerung der Exponentialverteilung. Ihre Dichtefunktion ist

$$f(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t}, \quad t \geq 0, \alpha, \beta > 0.$$

Dabei bezeichnet $\Gamma(\alpha)$ die Gammafunktion, die definiert ist als $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$. Die Verteilungsfunktion ist

$$F(t) = I(\alpha, \beta t) = \frac{1}{\Gamma(\alpha)} \int_0^{\beta t} u^{\alpha-1} e^{-u} du.$$

Die Funktion $I(\alpha, \beta)$ heißt unvollständige Gammafunktion. Der Parameter α ist der Formparameter und β ist der inverse Skalenparameter oder die Rate der Verteilung. Für den Fall $\alpha = 1$ erhält man eine Exponentialverteilung mit konstanter Ausfallrate β . Die Ausfallrate der Gammaverteilung lässt sich nicht in geschlossener Form darstellen. Sie ist, abhängig von α , monoton fallend ($\alpha < 1$), monoton wachsend ($\alpha > 1$) oder konstant ($\alpha = 1$) und es gilt $\lim_{t \rightarrow \infty} h(t) = \beta$ [Lawless, 2011, S. 25].

Zur Motivation der Gammaverteilung als Lebensdauerverteilung betrachtet man wieder eine Betriebsumgebung mit zufällig, mit einer Rate λ auftretenden, poissonverteilten Schocks. Das System fällt nach dem k -ten Schock aus und $T^{(k)}$ bezeichnet die Zeit bis zum k -ten Schock. Es gilt

$$T^{(k)} = T_1 + T_2 + \dots + T_k,$$

wobei T_i die Zeit zwischen dem i -ten und dem $i - 1$ -ten Schock bezeichnet. Die T_i sind alle exponentialverteilt mit Rate λ , daher ist die Zeit bis zum k -ten Schock als Summe von k exponentialverteilten Zufallsgrößen gammaverteilt mit Parametern k und λ .

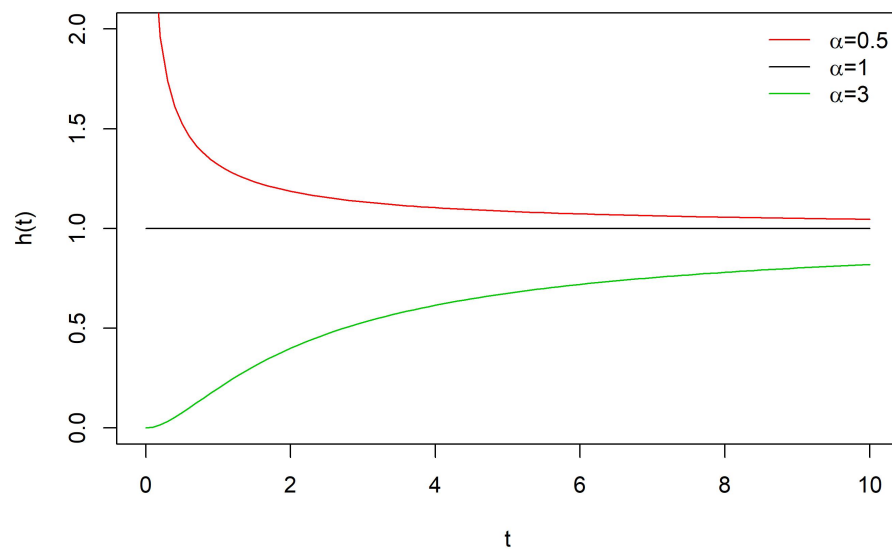


Abbildung 3.4: Die Ausfallrate $h(t)$ der Gammaverteilung mit $\beta = 1$ und verschiedenen Werten für α .

Parameterschätzung

Die Likelihood für $r \leq n$ Beobachtungen t_1, \dots, t_r mit einer gemeinsamen Zensurzeit t_s ist

$$l(\alpha, \beta) = \prod_{i=1}^n \frac{\beta^\alpha}{\Gamma(\alpha)} t_i^{\alpha-1} \exp(-\beta t_i) [1 - I(\alpha, \beta t_s)]^{n-r}. \quad (3.2.5)$$

Durch direktes Maximieren der log-Likelihood mit numerischen Methoden erhält man ML-Schätzer für α und β .

Bei vollständigen Daten ($n = r$) fällt der letzte Term mit der unvollständigen Gammafunktion in (3.2.5) weg, was die Berechnung der ML-Schätzer vereinfacht [Mann et al., 1974]. Die log-Likelihood ist dann

$$\log l(\alpha, \beta) = n\alpha \log \beta - n \log \Gamma(\alpha) + (\alpha - 1) \sum_{i=1}^n \log t_i - \beta \sum_{i=1}^n t_i,$$

Durch Ableiten und Null setzen erhält man nach einigen Umformungen die beiden Gleichungen

$$\frac{\alpha}{\beta} = \bar{t} \quad \text{und} \quad \log \alpha - \frac{d}{d\alpha} \log \Gamma(\alpha) = \log \bar{t} + \frac{1}{n} \sum_{i=1}^n \log t_i.$$

mit $\bar{t} = 1/n \sum_{i=1}^n t_i$. Eine Lösung der zweiten Gleichung, $\hat{\alpha}$, kann in die erste eingesetzt werden und ein Schätzer für β durch $\hat{\beta} = \hat{\alpha}/\bar{t}$ berechnet werden.

3.2.5 Gumbelverteilung

Die Gumbelverteilung oder Typ I-Extremwertverteilung ist eine von drei möglichen Grenzverteilungen des Maximums einer Folge von unabhängigen Zufallsgrößen⁴. Sie findet Anwendung, wenn der Ausfallmechanismus vom größten (oder kleinsten) Wert einer Zufallsgröße abhängt, deren Verteilung vom Typ einer Exponentialverteilung ist [Mann et al., 1974]. Die Verteilungs-

⁴Das Fisher-Tippett Theorem gibt eine Aussage über das Grenzverhalten des Maximums (oder des Minimums) einer Zufallsfolge. Für einen Beweis und weitere Resultate der Extremwerttheorie sei auf [Embrechts et al., 2013] verwiesen.

Satz 3.2. Sei (X_n) eine Folge von unabhängigen und identisch verteilten Zufallsgrößen und $M_n = \max(X_1, \dots, X_n)$. Falls Konstanten $c_n > 0$, $d_n \in \mathbb{R}$ und eine nicht degenerierte Verteilung H existieren, sodass $c_n^{-1}(M_n - d_n)$ in Verteilung gegen H konvergiert, dann gehört H zu einer von drei Verteilungen:

I. Gumbel: $F(x) = \exp(-e^{-x})$, $x \in \mathbb{R}$.

II. Fréchet: $F_\alpha(x) = \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases} \quad \alpha > 0.$

III. Weibull: $F_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha) & x \leq 0 \\ 1 & x > 0 \end{cases} \quad \alpha > 0.$

Wegen $m_n = \min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n)$ erhält man die entsprechenden Grenzverteilungen für das Minimum mit $F_{m_n}(x) = 1 - F_{M_n}(-x)$.

funktion, Dichtefunktion und Ausfallrate nach der Darstellung in [Mann et al., 1974] sind

$$\begin{aligned} F(t) &= 1 - \exp \left[- \exp \left(\frac{t - \mu}{\sigma} \right) \right], \\ f(t) &= \frac{1}{\sigma} \exp \left[\frac{t - \mu}{\sigma} - \exp \left(\frac{t - \mu}{\sigma} \right) \right], \\ h(t) &= \frac{1}{\sigma} \exp \left(\frac{t - \mu}{\sigma} \right). \end{aligned}$$

Der Parameter $\mu \in \mathbb{R}$ ist der Lageparameter und $\sigma > 0$ ist der Skalenparameter der Verteilung.

Die Beziehung der Weibullverteilung zur Gumbelverteilung ist dieselbe wie die Beziehung der Normalverteilung zur Log-Normalverteilung. Wenn $T \sim W(\eta, \beta)$, dann ist $\log T$ gumbelverteilt mit $\mu = \log \eta$ und $\sigma = 1/\beta$.

Die Gumbelverteilung als Lebensdauerverteilung folgt, wenn für die Ausfallrate h eine Exponentialfunktion angenommen wird:

$$h(t) = e^t, \quad -\infty < t < \infty.$$

Wegen (3.1.1) gilt für die Dichte- und Verteilungsfunktion

$$f(t) = \exp(t) \exp[-\exp(t)] \quad \text{und}$$

$$F(t) = 1 - \exp[-\exp(t)].$$

Aufgrund der exponentiell wachsenden Ausfallrate (Abb. 3.5) ist die Gumbelverteilung zur Modellierung von Komponenten, die nach einem bestimmten Zeitpunkt einen sehr hohen Verschleiß aufweisen, geeignet.

3.3 Beschleunigte Lebensdauerversuche

Oft ist es nicht möglich, Lebensdauerversuche unter realen Betriebsbedingungen durchzuführen. Viele Produkte sind hochzuverlässig und für sehr lange Betriebsdauern entwickelt, sodass solche Versuche viel zu lange dauern würden oder zu teuer wären. Daher werden so genannte beschleunigte Lebensdauerversuche (engl. *accelerated life tests*) durchgeführt. Dabei wird das Produkt unter höheren Belastungen als im normalen Gebrauch üblich getestet. Die Herausforderung besteht darin, aus dem Verhalten der Einheiten unter erhöhter Belastung auf das Verhalten unter normalen Betriebsbedingungen zu schließen. Belastungsfaktoren können z.B. elektrische Spannung, Temperatur oder Luftfeuchte sein. Die verschiedenen Stressfaktoren werden in einem Vektor \mathbf{S} zusammengefasst. Zum Einfluss von \mathbf{S} auf die Lebensdauerverteilung und ihre Parameter werden folgende Annahmen getroffen:

1. Das Ausmaß der Stressfaktoren beeinflusst die Parameterwerte von θ , hat aber keinen Einfluss auf die Form der Lebensdauerverteilung $f(t|\theta)$.
2. Die Beziehung zwischen \mathbf{S} und θ ist, bis auf endlich viele unbekannte Parameter $\tau = (a, b, c, \dots)$, bekannt: $\theta = g(\mathbf{S}; \tau)$, und diese Beziehung ist in einem gewissen Wertebereich von \mathbf{S} gültig.

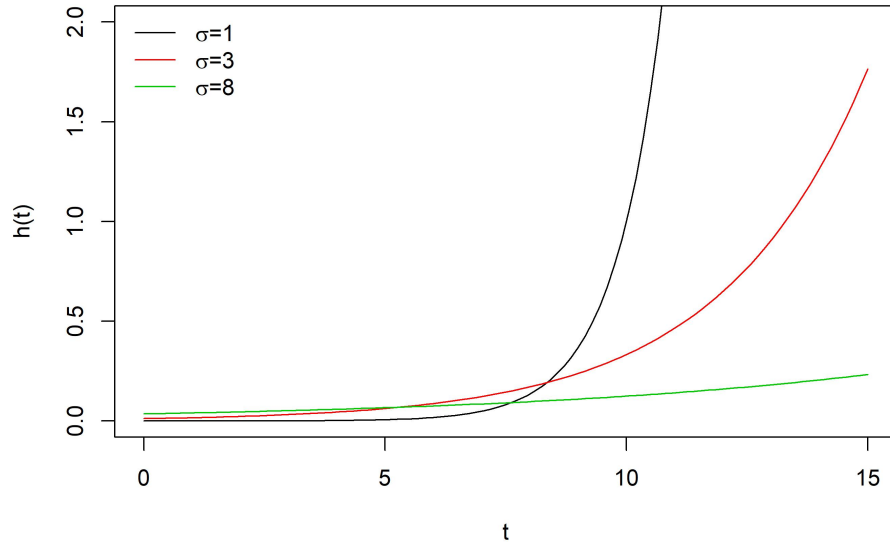


Abbildung 3.5: Die Ausfallrate $h(t)$ der Gumbelverteilung mit $\mu = 10$ und verschiedenen Werten für σ .

Die Parameter τ werden durch Daten aus einem Versuch unter erhöhtem Stress geschätzt und durch die Beziehung g wird dann ein Schätzer für θ unter normalen Betriebsbedingungen bestimmt. Dabei ist es wichtig, dass das gewählte Modell die Beziehung zwischen den Stressvariablen und dem Parameter der Lebensdauerverteilung möglichst gut beschreibt.

Am Beispiel der Exponentialverteilung werden in [Mann et al., 1974] einige klassische Modelle für die Beziehung zwischen einem Stressfaktor V und der Ausfallrate λ bzw. der mittleren Lebensdauer $\theta = \lambda^{-1}$ vorgestellt. Eine gute Übersicht über konkrete Anwendungsfälle dieser Modelle findet sich in [Nelson, 2009].

1. Das Power Rule Modell. Dieses Modell beschreibt θ als eine mit einer Potenz von V abnehmende Funktion:

$$\theta = \frac{C}{V^p}, \quad C, p > 0. \quad (3.3.1)$$

Die unbekannt Parameter sind C und p . Eine alternative Darstellung als lineares Modell ist

$$\log \theta = \log C - p \log V.$$

2. Das Arrhenius Reaction Rate Modell. Das Arrhenius Modell hat die Form

$$\lambda = \exp\left(A - \frac{B}{V}\right), \quad (3.3.2)$$

wobei V in der Regel die Temperatur beschreibt und A und B unbekannte Parameter sind. Diese Beziehung zwischen Lebensdauer und Stress ist geeignet zur Modellierung von Ausfällen,

die durch chemische Reaktionen ausgelöst werden, welche von der Betriebstemperatur abhängig sind. Konkrete Anwendungen findet das Modell bei Halbleiterbauteilen, Batteriezellen und anderen elektronischen Komponenten. [Nelson, 2009]

3. Das Eyring Modell. Für einen einzelnen Stressfaktor V ist das Eyring Modell gegeben durch

$$\lambda = V \exp\left(A - \frac{B}{V}\right). \quad (3.3.3)$$

Das Modell kann erweitert werden, indem man zwei Stressfaktoren T und V betrachtet:

$$\lambda = AT \exp\left(\frac{-B}{k_B T}\right) \exp\left(CV + \frac{DV}{k_B T}\right). \quad (3.3.4)$$

Dabei sind A, B, C, D die unbekannt Parameter und k_B ist die Boltzmann-Konstante. Die Grundlage für dieses Modell ist eine Gleichung für die Reaktionsgeschwindigkeit in einem chemischen Prozess, basierend auf der Theorie der Quantenmechanik [Mann et al., 1974], [Nelson, 2009].

3.3.1 Parameterschätzung

Wir betrachten hier beschleunigte Lebensdauerversuche, in denen die Komponenten unter k verschiedenen Belastungswerten $V_j, j = 1, \dots, k$ getestet werden und folgen im Wesentlichen der Darstellung aus [Mann et al., 1974, Kapitel 9]. Unter V_j wird angenommen, dass die Lebensdauer einer Exponentialverteilung mit Rate λ_j folgt. Um λ_j oder $\theta_j = \lambda_j^{-1}$ zu schätzen, werden k Tests auf folgende Weise durchgeführt:

1. Ein Wert V_i wird zufällig ausgewählt und
2. n_i Einheiten werden unter dem Stresslevel V_i getestet.
3. Der Test endet, wenn r_i Einheiten ausgefallen sind.

Für $i = 1, \dots, k$ werden die Ausfallzeiten mit $t_{1i}, t_{2i}, \dots, t_{r_i i}$ bezeichnet. Der ML-Schätzer für θ_i ist (vgl. Abschnitt 3.2.1)

$$\hat{\theta}_i = \frac{\sum_{j=1}^{r_i} t_{ji} + (n_i - r_i)t_{r_i i}}{r_i}. \quad (3.3.5)$$

Nach k Versuchen erhält man die Daten $\{V_i, n_i, r_i, \hat{\theta}_i\}_{i=1}^k$. Durch das randomisierte Testverfahren ist gesichert, dass die $\hat{\theta}_i$ unabhängig sind. Die Verteilung von $\hat{\theta}_i$ wurde in Abschnitt 3.2.1 hergeleitet und ist eine Gammaverteilung mit Dichtefunktion

$$g(\hat{\theta}_i) = \frac{1}{\Gamma(r_i)} \left(\frac{\theta_i}{r_i}\right)^{r_i} \hat{\theta}_i^{r_i-1} \exp\left(-\frac{r_i \hat{\theta}_i}{\theta_i}\right), \quad 0 \leq \hat{\theta}_i < \infty, r_i \geq 1.$$

Power Rule Modell

Das Modell muss umgeschrieben werden, damit die ML-Schätzer für C und p asymptotisch unabhängig sind [Mann et al., 1974]:

$$\theta_i = \frac{C}{(V_i/\bar{V})^p},$$

wobei \dot{V} das gewichtete geometrische Mittel der V_i ist:

$$\dot{V} = \prod_{i=1}^k (V_i)^{r_i / \sum_{i=1}^k r_i}.$$

Die $\hat{\theta}_i$ sind unabhängig und gammaverteilt und die Likelihood für C und p ist gegeben durch

$$l(C, p | \hat{\theta}_1, \dots, \hat{\theta}_k) = \prod_{i=1}^k \frac{1}{\Gamma(r_i)} \left[\frac{r_i}{C} \left(\frac{V_i}{\dot{V}} \right)^p \right]^{r_i} (\hat{\theta}_i)^{r_i-1} \exp \left[-\frac{r_i \hat{\theta}_i}{C} \left(\frac{V_i}{\dot{V}} \right)^p \right].$$

Durch Ableiten und Null setzen der Log-Likelihood kommt man unter der Beachtung, dass

$$\sum_{i=1}^k r_i \log \left(\frac{V_i}{\dot{V}} \right) = 0$$

ist, auf die zwei Gleichungen

$$C = \frac{\sum_{i=1}^k r_i \hat{\theta}_i \left(\frac{V_i}{\dot{V}} \right)^p}{\sum_{i=1}^k r_i} \quad (3.3.6)$$

und

$$\sum_{i=1}^k r_i \hat{\theta}_i \left(\frac{V_i}{\dot{V}} \right)^p \log \frac{V_i}{\dot{V}} = 0. \quad (3.3.7)$$

Ein Schätzwert für p kann aus der zweiten Gleichung numerisch bestimmt werden und in die erste Gleichung eingesetzt werden, um einen Schätzwert für C zu erhalten.

Alternativ kann das Modell in ein lineares Modell umgeschrieben werden und die Parameter mit der Methode der kleinsten Quadrate geschätzt werden. Man nimmt an, dass ϵ_i ein multiplikativer Fehlerterm ist, sodass $\hat{\theta}_i = \theta_i \epsilon_i$ und erhält durch Logarithmieren

$$\log \hat{\theta}_i = \log C - p \log V_i + \log \epsilon_i.$$

In Matrixschreibweise hat das Modell die Form

$$Y = X\alpha + \epsilon,$$

mit $\alpha = (\log C, -p)$,

$$Y = \begin{pmatrix} \log \hat{\theta}_1 \\ \log \hat{\theta}_2 \\ \vdots \\ \log \hat{\theta}_k \end{pmatrix}, X = \begin{pmatrix} 1 & \log V_1 \\ 1 & \log V_2 \\ \vdots & \vdots \\ 1 & \log V_k \end{pmatrix} \text{ und } \epsilon = \begin{pmatrix} \log \epsilon_1 \\ \log \epsilon_2 \\ \vdots \\ \log \epsilon_k \end{pmatrix}.$$

Der Kleinste-Quadrate-Schätzer für α ist

$$\hat{\alpha} = (X'X)^{-1}X'Y.$$

Die bayes'sche Analyse dieses linearen Modells wird in Kapitel 5 vorgestellt.

Arrhenius Reaction Rate Modell

Um asymptotisch unabhängige Schätzer für A und B zu erhalten, wird das Modell mit dem gewichteten Mittel der V_i^{-1} ,

$$\bar{V} = \frac{\sum_{i=1}^k r_i / V_i}{\sum_{i=1}^k r_i}$$

umgeschrieben:

$$\lambda_i = \exp\left(A - B(V_i^{-1} - \bar{V})\right). \quad (3.3.8)$$

Die Verteilung von $\hat{\lambda}_i$ ist eine inverse Gammaverteilung $\gamma^{-1}(r_i, r_i \lambda_i)$. Setzt man in die Dichte dieser Verteilung (3.3.8) für λ_i ein, so ist die Likelihood für A und B gleich

$$l(A, B | \hat{\lambda}_1, \dots, \hat{\lambda}_k) = \prod_{i=1}^k \frac{1}{\Gamma(r_i)} \left[r_i \exp\left(A - B(V_i^{-1} - \bar{V})\right) \right]^{r_i} (\hat{\lambda}_i)^{-r_i-1} \times \exp\left(-\frac{r_i}{\hat{\lambda}_i} \exp\left(A - B(V_i^{-1} - \bar{V})\right)\right). \quad (3.3.9)$$

Zur Maximierung von (3.3.9) müssen zwei nichtlineare Gleichungen in A und B numerisch gelöst werden.

Mit der Annahme $\hat{\lambda}_i = \lambda_i \epsilon_i$ kann das Arrhenius Modell auch in linearer Form dargestellt werden:

$$\log \hat{\lambda}_i = A - \frac{B}{V_i} + \log \epsilon_i.$$

Mit $\alpha = (A, -B)$,

$$Y = \begin{pmatrix} \log \hat{\lambda}_1 \\ \log \hat{\lambda}_2 \\ \vdots \\ \log \hat{\lambda}_k \end{pmatrix}, X = \begin{pmatrix} 1 & \frac{1}{V_1} \\ 1 & \frac{1}{V_2} \\ \vdots & \vdots \\ 1 & \frac{1}{V_k} \end{pmatrix} \text{ und } \epsilon = \begin{pmatrix} \log \epsilon_1 \\ \log \epsilon_2 \\ \vdots \\ \log \epsilon_k \end{pmatrix}$$

kann man das Modell schreiben als

$$Y = X\alpha + \epsilon.$$

Der Kleinste-Quadrate-Schätzer für α ist $\hat{\alpha} = (X'X)^{-1}X'Y$.

Eyring Modell

Das einfache Eyring Modell (3.3.3) unterscheidet sich nur unwesentlich vom Arrhenius Modell (3.3.2) und wird daher nicht extra behandelt. Im verallgemeinerten Eyring Modell (3.3.4) wird λ_i als Funktion von zwei Stressvariablen V_i, T_i , den Parametern A, B, C, D und der Boltzmann Konstante k_B dargestellt:

$$\lambda_i = AT_i \exp\left(\frac{-B}{k_B T_i}\right) \exp\left(CV_i + \frac{DV_i}{k_B T_i}\right). \quad (3.3.10)$$

Um Maximum Likelihood Schätzer für die vier unbekannt Parameter zu berechnen, sind vier nichtlineare Gleichungen zu lösen.

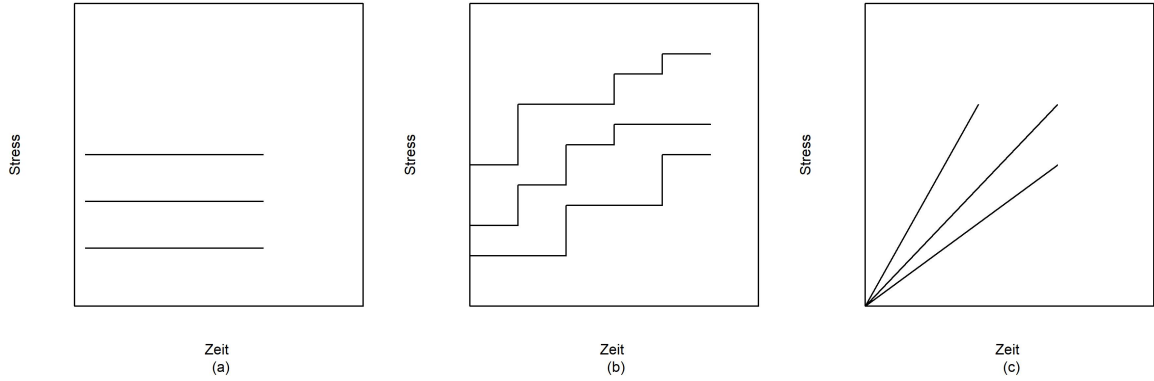


Abbildung 3.6: Illustration verschiedener Designs beschleunigter Lebensdauertests: (a) Konstanter Stress, (b) *Step-Stress* und (c) *Progressive Stress*.

Man kann das Eyring Modell, ähnlich wie die beiden Modelle zuvor, linearisieren. Angenommen ϵ_i ist ein multiplikativer Fehlerterm, sodass $\hat{\lambda}_i = \lambda_i \epsilon_i$ für $i = 1, \dots, k$, dann kann man die Beziehung (3.3.10) schreiben als

$$\log \hat{\lambda}_i = \log A + \log T_i - \frac{B}{k_B T_i} + C V_i + \frac{D V_i}{k_B T_i} + \log \epsilon_i.$$

In Matrixschreibweise ist

$$Y = X\alpha + \epsilon,$$

mit $\alpha = (\log A, -B, C, D)$,

$$Y = \begin{pmatrix} \log(\hat{\lambda}_1/T_1) \\ \log(\hat{\lambda}_2/T_2) \\ \vdots \\ \log(\hat{\lambda}_k/T_k) \end{pmatrix}, X = \begin{pmatrix} 1 & \frac{1}{k_B T_1} & V_1 & \frac{V_1}{k_B T_1} \\ 1 & \frac{1}{k_B T_2} & V_2 & \frac{V_2}{k_B T_2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{1}{k_B T_k} & V_k & \frac{V_k}{k_B T_k} \end{pmatrix} \text{ und } \epsilon = \begin{pmatrix} \log \epsilon_1 \\ \log \epsilon_2 \\ \vdots \\ \log \epsilon_k \end{pmatrix}.$$

Der Kleinste-Quadrate-Schätzer für α ist wieder $\hat{\alpha} = (X'X)^{-1}X'Y$.

3.3.2 Ein einfaches Step-Stress Modell

Bei dem in 3.3.1 beschriebenen Versuchsdesign wird jede Einheit unter einem konstanten Stresslevel getestet. Es gibt andere Versuchsdesigns, in denen das Stresslevel etwa mit der Zeit variiert. Zwei solche Designs sind in Abbildung 3.6 schematisch dargestellt. Bei einem *Step-Stress Accelerated Life Test* werden die Einheiten bis zu einem bestimmten Zeitpunkt unter einem Stresslevel, welches dann sukzessive erhöht wird, getestet. Wird das Stresslevel kontinuierlich und nicht schrittweise erhöht, so spricht man von einem *Progressive Stress Accelerated Life Test*. [Nelson, 2009]

Zur Illustration eines einfachen Step-Stress Modells mit zwei Stressleveln wird das folgende Testszenario unterstellt. Die Darstellung orientiert sich an [Balakrishnan et al., 2007], wo weitere Eigenschaften der Maximum Likelihood Schätzer beschrieben sind.

In dem Versuch mit Typ II-Zensierung werden insgesamt n Einheiten so lange getestet, bis $r \leq n$ Ausfälle auftreten. Zu Beginn werden die Einheiten unter dem Stresslevel V_1 getestet und zwar bis zu einem festen Zeitpunkt τ , an dem das Stresslevel auf V_2 erhöht wird. Es bezeichnen n_1 bzw. n_2 die Anzahl der Ausfälle bis τ bzw. nach τ und es ist $n_1 + n_2 = r$. Falls alle Einheiten unter V_1 ausfallen ($n_1 = r$), wird der Test beendet. Ansonsten wird unter V_2 weiter getestet, bis insgesamt r Ausfälle aufgetreten sind. Die geordneten beobachteten Ausfallzeiten sind $t_{(1)}, \dots, t_{(r)}$.

Für die Lebensdauer unter V_i wird eine Exponentialverteilung mit Mittelwert θ_i angenommen. Das Modell muss die Verteilung unter Step-Stress mit der Lebensdauerverteilung unter konstantem Stress bei üblichen Betriebsbedingungen in Beziehung setzen. Dabei wird ein so genanntes Cumulative Exposure Modell verwendet. In diesem Modell wird angenommen, dass die Restlebensdauer einer Einheit nur vom aktuellen Stresslevel und dem Anteil der bisher ausgefallenen Einheiten abhängt und unabhängig davon ist, wie dieser Anteil zustande gekommen ist. In [Nelson, 2009, Kapitel 10] sind weitere Details zu dem Cumulative Exposure Modell zu finden.

Die Verteilungsfunktionen unter V_1 und V_2 sind

$$F_1(t) = 1 - \exp\left(-\frac{t}{\theta_1}\right) \quad \text{und} \quad F_2(t) = 1 - \exp\left(-\frac{t}{\theta_2}\right).$$

Der Übergang von V_1 zu V_2 verändert die Verteilungsfunktion F_2 zu $F_2(t - h)$, wobei h so gewählt ist, dass $F_1(\tau) = F_2(\tau + h)$ ist. Unter Annahme einer Exponentialverteilung ist $h = \theta_2\tau/\theta_1 - \tau$. Damit ist die Verteilungsfunktion unter dem Step-Stress Modell gegeben durch

$$G(t) = \begin{cases} F_1(t) & 0 \leq t < \tau, \\ F_2\left(\frac{\theta_2}{\theta_1}\tau + t - \tau\right) & \tau \leq t < \infty. \end{cases} \quad (3.3.11)$$

Für die Likelihood von θ_1 und θ_2 nach n getesteten Einheiten mit r beobachteten Ausfallzeiten $t_{(1)} \leq \dots \leq t_{(r)}$ werden drei Fälle unterschieden:

$$l(\theta_1, \theta_2) = \begin{cases} \frac{1}{\theta_1^r} \exp\left(-\frac{1}{\theta_1}A\right) & n_1 = r, \\ \frac{1}{\theta_2^r} \exp\left(-\frac{1}{\theta_2}B(\theta_1, \theta_2)\right) & n_1 = 0, \\ \frac{1}{\theta_1^{n_1}\theta_2^{n_2}} \exp\left(-\frac{1}{\theta_1}C_1 - \frac{1}{\theta_2}C_2\right) & 1 \leq n_1 \leq r - 1, \end{cases} \quad (3.3.12)$$

mit

$$\begin{aligned} A &= \sum_{i=1}^r t_{(i)} + (n - r)t_{(r)}, \\ B(\theta_1, \theta_2) &= \sum_{i=1}^r \left(\frac{\theta_2}{\theta_1}\tau + t_{(i)} - \tau\right) + (n - r) \left(\frac{\theta_2}{\theta_1}\tau + t_{(r)} - \tau\right), \\ C_1 &= \sum_{i=1}^{n_1} t_{(i)} + (n - n_1)\tau, \\ C_2 &= \sum_{i=n_1+1}^r (t_{(i)} - \tau) + (n - r)(t_{(r)} - \tau). \end{aligned}$$

Im Fall $n_1 = r$ berechnet sich der Maximum Likelihood Schätzer für θ_1 wie in Abschnitt 3.2.1 für Typ II-zensierte Daten beschrieben. Für θ_2 existiert in diesem Fall kein ML-Schätzer. Im Fall $n_1 = 0$ treten alle Ausfälle unter dem Stresslevel V_2 auf und daher existiert kein ML-Schätzer für θ_1 . Der interessantere Fall mit jeweils mindestens einem Ausfall unter V_1 und unter V_2 ist $1 \leq n_1 \leq r - 1$. Die ML-Schätzer für θ_1 und θ_2 unter $1 \leq n_1 \leq r - 1$ sind

$$\hat{\theta}_1 = \frac{\sum_{i=1}^{n_1} t_{(i)} + (n - n_1)\tau}{n_1},$$
$$\hat{\theta}_2 = \frac{\sum_{i=n_1+1}^r (t_{(i)} - \tau) + (n - r)(t_{(r)} - \tau)}{r - n_1}.$$

Das Modell kann auch erweitert werden, indem für θ eine bekannte Funktion der Stressvariable eingesetzt wird, etwa eine der drei in den vorigen Abschnitten vorgestellten Funktionen. In Kapitel 5 wird ein Bayes'scher Ansatz für dieses Modell diskutiert.

4 | Bayes'sche Analyse von parametrischen Modellen

Die Bayes Statistik bietet die Möglichkeit, Vorinformation in das statistische Modell miteinzubeziehen. Gerade im Bereich der Zuverlässigkeitsanalyse gibt es Informationen über die zu testenden Komponenten, deren Fehlermechanismen oder Ergebnisse aus ähnlichen Experimenten, die in der Anwendung von klassischen statistischen Modellen nicht berücksichtigt werden. Bei der Verwendung von Bayes'schen Methoden gehen diese Informationen durch die A-priori Verteilung auf schlüssige Art und Weise in die Modellierung ein.

In diesem Kapitel geht es um die Bayes'sche Analyse der in Abschnitt 3.2 vorgestellten parametrischen Modelle. Den Anfang macht die Exponentialverteilung, die aufgrund ihrer Einfachheit eine detaillierte Darstellung der Bayes'schen Methodik ermöglicht. Bei der Analyse der weiteren Lebensdauermodelle wie der Weibullverteilung, der Log-Normalverteilung oder der Gammaverteilung hat die A-posteriori Verteilung keine geschlossene Form, weshalb in diesen Fällen numerische Verfahren angewendet werden müssen, um Stichproben aus der A-posteriori zu generieren. Die Bücher von [Hamada et al., 2008], [Ibrahim et al., 2013] und [Singpurwalla, 2006] dienen als wesentliche Quellen für die nächsten Abschnitte.

4.1 Exponentialverteilung

Vollständige Daten

Es ist $T \sim Exp(\lambda)$ mit $f(t|\lambda) = \lambda e^{-\lambda t}$, $t \geq 0, \lambda > 0$. Die Likelihood für n unabhängige beobachtete Ausfallzeiten t_1, \dots, t_n ist

$$l(\lambda|t_1, \dots, t_n) = \lambda^n \exp\left(-\lambda \sum_{i=1}^n t_i\right).$$

In Beispiel 2.4 wurde die Gammaverteilung als konjugierte A-priori für λ hergeleitet. Mit $\lambda \sim \gamma(a, b)$ ist die A-posteriori wieder eine Gammaverteilung:

$$\begin{aligned} \pi(\lambda|t_1, \dots, t_n) &\propto l(\lambda|t_1, \dots, t_n)\pi(\lambda) \\ &\propto \lambda^n e^{-\lambda \sum_{i=1}^n t_i} \lambda^{a-1} e^{-b\lambda} \\ &= \lambda^{a+n-1} e^{-\lambda(b + \sum_{i=1}^n t_i)}. \end{aligned}$$

Der Bayes Schätzer unter einer quadratischen Verlustfunktion ist der Erwartungswert der A-posteriori-Verteilung

$$\mathbb{E}_\pi[\lambda|t_1, \dots, t_n] = \frac{a+n}{b + \sum_{i=1}^n t_i}. \quad (4.1.1)$$

Betrachtet man in (4.1.1) den Fall $a \rightarrow 0$ und $b \rightarrow 0$, so stimmt der Bayes Schätzer mit dem Maximum-Likelihood Schätzer

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n t_i}$$

überein. Dieser entspricht auch dem Bayes Schätzer bei der nicht informativen Jeffreys A-priori $\pi_J(\lambda) \propto \lambda^{-1}$ (vgl. Beispiel 2.8). Die A-posteriori ist dann eine $\gamma(n, \sum_{i=1}^n t_i)$ -Verteilung und wird nur von den Beobachtungen bestimmt.

Die Prädiktivdichte für eine neue Beobachtung z mit $g(z|\lambda) = \lambda e^{-\lambda z} \mathbb{1}_{[0, \infty)}(z)$ berechnet sich nach Definition 2.3 durch

$$\begin{aligned} g(z|t_1, \dots, t_n) &= \int_0^\infty g(z|\lambda) \pi(\lambda|t_1, \dots, t_n) d\lambda \\ &= \frac{(b + \sum_{i=1}^n t_i)^{a+n}}{\Gamma(a+n)} \int_0^\infty \lambda^{a+n} e^{-\lambda(b + \sum_{i=1}^n t_i + z)} d\lambda \\ &= \frac{(b + \sum_{i=1}^n t_i)^{a+n}}{\Gamma(a+n)} \frac{\Gamma(a+n+1)}{(b + \sum_{i=1}^n t_i + z)^{a+n+1}} \\ &= \frac{(a+n)(b + \sum_{i=1}^n t_i)^{a+n}}{(b + \sum_{i=1}^n t_i + z)^{a+n+1}}. \end{aligned}$$

Die Prädiktivdichte entspricht der Dichte einer Paretoverteilung. Für Mittelwert μ und Varianz σ^2 der Prädiktivverteilung gilt

$$\mu = \frac{(a+n)(b + \sum_{i=1}^n t_i)}{a+n-1} \quad \text{und} \quad \sigma^2 = \frac{(a+n)(b + \sum_{i=1}^n t_i)}{(a+n-1)^2(a+n-2)}.$$

Die Zuverlässigkeitsfunktion ist gegeben durch

$$R(z) = P(T > z) = \left(\frac{b + \sum_{i=1}^n t_i}{b + \sum_{i=1}^n t_i + z} \right)^{a+n}.$$

Beispiel 4.1. In diesem Beispiel werden Ausfallzeiten für eine Komponente eines Raketenabwehrsystems analysiert. Die Daten sind aus [Hamada et al., 2008, Exercise 4.7] entnommen und enthalten die folgenden $n = 20$ Ausfallzeiten gemessen in 1000 Stunden: 14.4, 2.1, 0.4, 18.6, 1.2, 2.6, 11.5, 18.4, 14.0, 2.8, 7.6, 2.7, 35.4, 10.4, 19.8, 11.3, 2.6, 0.8, 11.3, 5.4.

Es wird angenommen, dass die Daten aus einer Exponentialverteilung stammen, deren Ausfallrate λ zu schätzen ist. Als A-priori wird die konjugierte Gammaverteilung mit den Hyperparametern $a = 5$ und $b = 90$ verwendet. Diese Wahl der Hyperparameter kann als Ergebnis eines ähnlichen Experimentes interpretiert werden, bei dem 5 Komponenten getestet wurden und die Summe der Ausfallzeiten gleich 90 war. Anhand dieser Werte ist die a-priori angenommene mittlere Ausfallrate etwa 0.05.

Aus den Daten erhält man die suffizienten Statistiken $n = 20$ und $S = \sum_{i=1}^n t_i = 193.3$. Die A-posteriori Verteilung ist eine $\gamma(25, 283.3)$ -Verteilung. Der Bayes Schätzer bei quadratischem

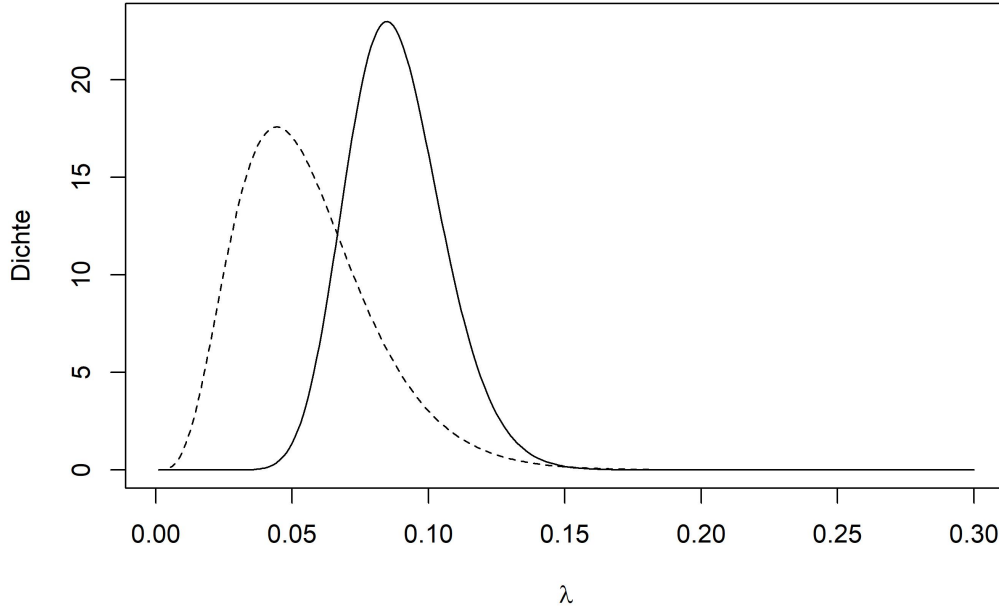


Abbildung 4.1: A-posteriori Verteilung (durchgehende Linie) für die Ausfallrate λ unter Annahme einer konjugierten $\gamma(5, 90)$ -A-priori (gestrichelte Linie).

Verlust für λ ist $\lambda^B = 0.088$ und ein 95% HPD Intervall ist $(0.05, 0.12)$ (vgl. Beispiel 2.10). Die Dichtefunktionen der A-priori und der A-posteriori Verteilung sind in Abbildung 4.1 dargestellt. Es ist ersichtlich, dass die neuen Daten die Annahmen aus der A-priori beeinflussen. Die Ausfallrate nach Beobachtung der Daten ist deutlich höher, als aufgrund der Ergebnisse ähnlicher Tests a-priori vermutet wurde.

Die A-posteriori der mittleren Lebensdauer $\theta = \lambda^{-1}$ ist eine inverse Gammaverteilung $\gamma^{-1}(25, 283.3)$. Der Bayes Schätzer ist $\theta^B = \frac{283.3}{24} = 11.8$ Tausend Stunden. Die A-priori Hyperparameter lassen mit einer mittleren Lebensdauer von 22.5 Tausend Stunden einen wesentlich höheren Wert annehmen. Die A-posteriori Wahrscheinlichkeit, dass θ den Wert 22.5 übersteigt, ist 0.0013. Die Annahmen, welche in der A-priori stecken, werden also durch die Daten upgedatet.

Unter Annahme einer nicht informativen Jeffreys A-priori ist die A-posteriori eine Gammaverteilung mit den Parametern 20 und 193.3. Der Bayes Schätzer, welcher mit dem ML-Schätzer übereinstimmt, ist gleich 0.103. Ein 95% HPD Intervall ist $(0.06, 0.14)$.

Die Daten aus diesem Beispiel werden in der Folge zum Vergleich auch mit einer Weibull-, einer Lognormal- und einer Gammaverteilung modelliert. \diamond

Zensierte Daten

Im Fall von zensierten Daten mit $r < n$ beobachteten Ausfallzeiten und $n - r$ rechtszensierten Einheiten ist die Likelihood gegeben durch

$$\begin{aligned} l(\lambda|t_1, \dots, t_r) &= \lambda^r \exp\left(-\lambda\left(\sum_{i=1}^r t_i + (n-r)t_s\right)\right) \\ &= \lambda^r e^{-\lambda S}. \end{aligned}$$

Bei Typ I-Zensierung ist t_s gleich der vorher festgelegten Zensurzeit t_0 und bei Typ II-Zensierung ist t_s die größte beobachtete Ausfallzeit $t_{(r)}$. Es sind r und die „total time on test“

$$S = \sum_{i=1}^r t_i + (n-r)t_s$$

die suffizienten Statistiken. Die Gammaverteilung $\gamma(a, b)$ ist die konjugierte A-priori für λ und es gilt

$$\pi(\lambda|t_1, \dots, t_r) \propto \lambda^{a+r-1} e^{-\lambda(b+S)}.$$

Die A-posteriori ist wieder eine Gammaverteilung mit Parametern $a + r$ und $b + S$. Der Bayes Schätzer bei quadratischem Verlust ist

$$\mathbb{E}_\pi[\lambda|t_1, \dots, t_r] = \frac{a+r}{b+S}.$$

Die A-posteriori Prädiktivdichte ist ähnlich wie im Fall vollständiger Daten eine Pareto-dichte, denn für eine Beobachtung z mit $g(z|\lambda) = \lambda e^{-\lambda z} \mathbb{1}_{[0, \infty)}(z)$ ist

$$\begin{aligned} g(z|t_1, \dots, t_r) &= \int_0^\infty g(z|\lambda) \pi(\lambda|t_1, \dots, t_r) d\lambda \\ &= \frac{(a+r)(b+S)^{a+r}}{(b+S+z)^{a+r+1}}. \end{aligned}$$

Die Überlebensfunktion von z ist

$$R(z) = P(T > z) = \left(\frac{b+S}{b+S+z}\right)^{a+r}.$$

4.2 Weibullverteilung

In diesem Abschnitt betrachten wir eine Reparametrisierung der Weibullverteilung mit $\lambda = \eta^{-\beta}$. Die Dichtefunktion ist dann

$$f(t|\lambda, \beta) = \beta \lambda t^{\beta-1} \exp(-\lambda t^\beta), \quad t \geq 0, \lambda > 0, \beta > 0.$$

Vollständige Daten

Bei einer vollständigen Stichprobe mit n beobachteten Ausfallzeiten t_1, \dots, t_n ist die Likelihood

$$\begin{aligned} l(\lambda, \beta | t_1, \dots, t_n) &= \prod_{i=1}^n \beta \lambda t_i^{\beta-1} \exp(-\lambda t_i^\beta) \\ &= \beta^n \lambda^n \prod_{i=1}^n t_i^{\beta-1} \exp\left(-\lambda \sum_{i=1}^n t_i^\beta\right) \\ &= \beta^n \lambda^n \exp\left(-\lambda \sum_{i=1}^n t_i^\beta + (\beta-1) \sum_{i=1}^n \log t_i\right). \end{aligned} \quad (4.2.1)$$

Wenn β bekannt ist, dann ist $T^\beta \sim \text{Exp}(\lambda)$ und die Gammaverteilung ist die konjugierte A-priori Verteilung für λ . Sind beide Parameter unbekannt, so existiert für die Weibullverteilung keine gemeinsame konjugierte A-priori. Es existiert nämlich keine suffiziente Statistik mit fester Dimension für jede Stichprobengröße. In [Soland, 1969] wird eine konjugierte A-priori für (λ, β) hergeleitet, die aus einem stetigen und einem diskreten Teil zusammengesetzt ist. Der diskrete Teil betrifft den Parameter β . Es wird vorausgesetzt, dass β einen von nur endlich vielen Werten β_1, \dots, β_m mit A-priori Wahrscheinlichkeiten

$$p_i = P(\beta = \beta_i), \quad i = 1, \dots, m, \quad \sum_{i=1}^m p_i = 1$$

annehmen kann. Die bedingte Verteilung von λ gegeben $\beta = \beta_i$ ist eine Gammaverteilung mit Dichtefunktion

$$f(\lambda | \beta_i) = \frac{b_i^{a_i}}{\Gamma(a_i)} \lambda^{a_i-1} \exp(-\lambda b_i).$$

Diese Verteilung hängt von β_i nur durch den Index i der Hyperparameter a_i, b_i ab. Die A-posteriori Verteilung hat dann dieselbe Form wie die A-priori Verteilungen [Soland, 1969]. Diese diskret-stetige A-priori wird von manchen Autoren jedoch kritisiert, da die Anwendung für reale Probleme als schwierig angesehen wird, insbesondere bezogen auf den diskreten Teil der Verteilung [Kaminskiy and Krivtsov, 2005], [Joarder et al., 2011].

In [Hamada et al., 2008] werden λ und β als a-priori unabhängig angenommen, sodass $\pi(\lambda, \beta) = \pi(\lambda)\pi(\beta)$ ist, mit den bedingten konjugierten A-priori Verteilungen

$$\begin{aligned} \pi(\lambda) &\propto \lambda^{a_\lambda-1} e^{-\lambda b_\lambda}, \\ \pi(\beta) &\propto \beta^{a_\beta-1} e^{-\beta b_\beta}, \end{aligned} \quad (4.2.2)$$

und Hyperparametern $a_\lambda, b_\lambda, a_\beta$ und b_β .

Mit der Likelihood (4.2.1) und der A-priori (4.2.2) ist die A-posteriori Verteilung von λ und β proportional zu

$$\pi(\lambda, \beta | t_1, \dots, t_n) \propto \lambda^{n+a_\lambda-1} \beta^{n+a_\beta-1} \exp\left(-\lambda \left(b_\lambda + \sum_{i=1}^n t_i^\beta\right) - \beta \left(b_\beta - \sum_{i=1}^n \log t_i\right)\right).$$

Da diese Verteilung keine geschlossene Form hat, müssen numerische Verfahren angewendet werden, um Realisierungen aus der A-posteriori zu generieren. Dazu ist festzuhalten, dass

die bedingte Verteilung von $\lambda|\beta, t_1, \dots, t_n$ eine Gammaverteilung mit Parametern $n + a_\lambda$ und $b_\lambda + \sum_{i=1}^n t_i^\beta$ ist, denn

$$\pi(\lambda|\beta, t_1, \dots, t_n) \propto \lambda^{n+a_\lambda-1} \exp\left(-\lambda\left(b_\lambda + \sum_{i=1}^n t_i^\beta\right)\right).$$

Realisierungen aus dieser bedingten Verteilung sind einfach zu generieren. Die bedingte Verteilung von $\beta|\lambda, t_1, \dots, t_n$ hat keine geschlossene Form, sie ist proportional zu

$$\pi(\beta|\lambda, t_1, \dots, t_n) \propto \beta^{n+a_\beta-1} \exp\left(-\beta\left(b_\beta - \sum_{i=1}^n \log t_i\right) - \lambda \sum_{i=1}^n t_i^\beta\right).$$

Durch eine Kombination des Metropolis-Hastings Algorithmus und des Gibbs Sampling Verfahrens, wie in Algorithmus 4.1 beschrieben, können Realisierungen der A-posteriori Verteilung $\pi(\lambda, \beta|t_1, \dots, t_n)$ gewonnen werden. In [Robert and Casella, 2013] ist diese Art von Algorithmus als „Metropolis within Gibbs“ Algorithmus beschrieben.

Algorithmus 4.1.

1. Lege Startwerte λ_0 und β_0 fest.

Für $i = 1, 2, \dots$:

2. generiere $\lambda_i \sim \gamma\left(n + a_\lambda, b_\lambda + \sum_{i=1}^n t_i^{\beta_{i-1}}\right)$,
3. generiere $\beta' \sim q(\beta'|\beta_{i-1})$,
4. berechne

$$\rho = \min\left\{1, \frac{\pi(\lambda_i, \beta'|t_1, \dots, t_n)q(\beta_{i-1}|\beta')}{\pi(\lambda_i, \beta_{i-1}|t_1, \dots, t_n)q(\beta'|\beta_{i-1})}\right\}$$

und setze $\beta_i = \beta'$ mit Wahrscheinlichkeit ρ .

Die bedingte Verteilung mit der Dichte $q(\cdot|\beta)$ wird als Vorschlagsdichte (eng. *proposal distribution*) bezeichnet und ist in der Regel eine einfach zu erzeugende Verteilung. In Anlehnung an [Robert, 2007, Example 6.3.2] wird in den folgenden Beispielen zur Weibullverteilung $q(\beta'|\beta)$ festgelegt als

$$q(\beta'|\beta) = \frac{1}{\beta} \exp\left(-\frac{\beta'}{\beta}\right).$$

Beispiel 4.2 (Fortsetzung von Beispiel 4.1). Die Ausfallzeiten der Komponente eines Raketenabwehrsystems aus Beispiel 4.1 können auch durch eine Weibullverteilung modelliert werden. Dies ist eine alternative Analysemöglichkeit, wenn z.B. nicht ausgeschlossen werden kann, dass nicht doch eine steigende oder fallende Ausfallrate vorliegt.

Als A-priori für λ und β werden zwei unabhängige Gammaverteilungen angenommen. Für λ dieselbe $\gamma(5, 90)$ -Verteilung aus Beispiel 4.1 und für β eine $\gamma(1, 1)$ -Verteilung. Diese hat den Erwartungswert 1, was einer konstanten Ausfallrate entspricht, sie ermöglicht aber genauso eine fallende oder eine steigende Ausfallrate (vgl. [Hamada et al., 2008]).

Unter Anwendung von Algorithmus 4.1 werden Stichproben aus der A-posteriori Verteilung beider Parameter der Weibullverteilung gezogen. Abbildung 4.2 zeigt die 10000 MCMC Iterationen nach einem Burn-In Sample von 1000 sowie Kerndichteschätzer der A-posteriori

Randverteilungen von λ und β . Die Abbildungen zeigen, dass die simulierten Werte gut durchmischt sind und sie lassen keine Trends oder Auffälligkeiten erkennen.

Die Betrachtung der A-posteriori Randverteilung von β zeigt, dass eine steigende Ausfallrate bei diesen Daten durchaus zutreffen könnte. Der A-posteriori Mittelwert von β ist 1.14 und die A-posteriori Wahrscheinlichkeit, dass β größer als 1 ist, ist

$$P_\pi(\beta > 1 | t_1, \dots, t_{20}) = 0.883.$$

Diese große Wahrscheinlichkeit lässt darauf schließen, dass die Weibullverteilung zur Modellierung dieser Daten gegenüber der Exponentialverteilung vorzuziehen ist.

Die A-posteriori der mittleren Lebensdauer der Weibullverteilung, die sich durch

$$\mathbb{E}[T] = \lambda^{-1/\beta} \Gamma\left(1 + \frac{1}{\beta}\right)$$

berechnet, wird durch Einsetzen der Stichproben der A-posteriori von λ, β approximiert. Der A-posteriori Mittelwert der mittleren Lebensdauer liegt mit 10.8 Tausend Stunden um etwa 1 Tausend Stunden unter dem mit der Exponentialverteilung geschätzten Wert. \diamond

Jeffreys A-priori

Die Jeffreys A-priori für die Weibullverteilung wird in [Sun, 1997] ausführlich hergeleitet. Die Fisher-Informationsmatrix für die Parametrisierung mit η und β ist

$$\mathcal{I}(\eta, \beta) = \begin{pmatrix} \frac{\beta^2}{\eta^2} & -\frac{\gamma_1+1}{\eta} \\ -\frac{\gamma_1+1}{\eta} & \frac{\gamma_2+2\gamma_1+1}{\beta^2} \end{pmatrix}.$$

Dabei bezeichnet

$$\gamma_i = \int_0^\infty [\log(y)]^i e^{-y} dy$$

das i -te Moment von $\log(Z)$ für eine exponentialverteilte Zufallsvariable Z mit $\mathbb{E}[Z] = 1$.

Die Determinante von $\mathcal{I}(\eta, \beta)$ ist proportional zu $1/\eta^2$ und somit ist die Jeffreys A-priori für (η, β) gegeben durch $\pi_J(\eta, \beta) = \eta^{-1}$.

Für die Parametertransformation $(\eta, \beta) \rightarrow (\lambda, \beta)$ mit $\lambda = \eta^{-\beta}$ lässt sich die Fisher-Information mit Hilfe der Jacobi-Matrix H als $H^\top \mathcal{I} H$ berechnen. Es ist

$$H = \begin{pmatrix} -\beta^{-1} \lambda^{-1/\beta-1} & \beta^{-2} \lambda^{-1/\beta} \log(\lambda) \\ 0 & 1 \end{pmatrix},$$

und

$$H^\top \mathcal{I} H = \begin{pmatrix} \frac{1}{\lambda^2} & -\frac{\gamma_1+1-\log(\lambda)}{\beta\lambda} \\ -\frac{\gamma_1+1-\log(\lambda)}{\beta\lambda} & \frac{(\log(\lambda)+\gamma_1+1)^2 + \gamma_2 - \gamma_1^2}{\beta^2} \end{pmatrix}.$$

Die Determinante dieser Matrix ist $(\gamma_2 - \gamma_1^2)/(\lambda\beta)^2$ und die Jeffreys A-priori ist dementsprechend

$$\pi_J(\lambda, \beta) \propto (\lambda\beta)^{-1}.$$

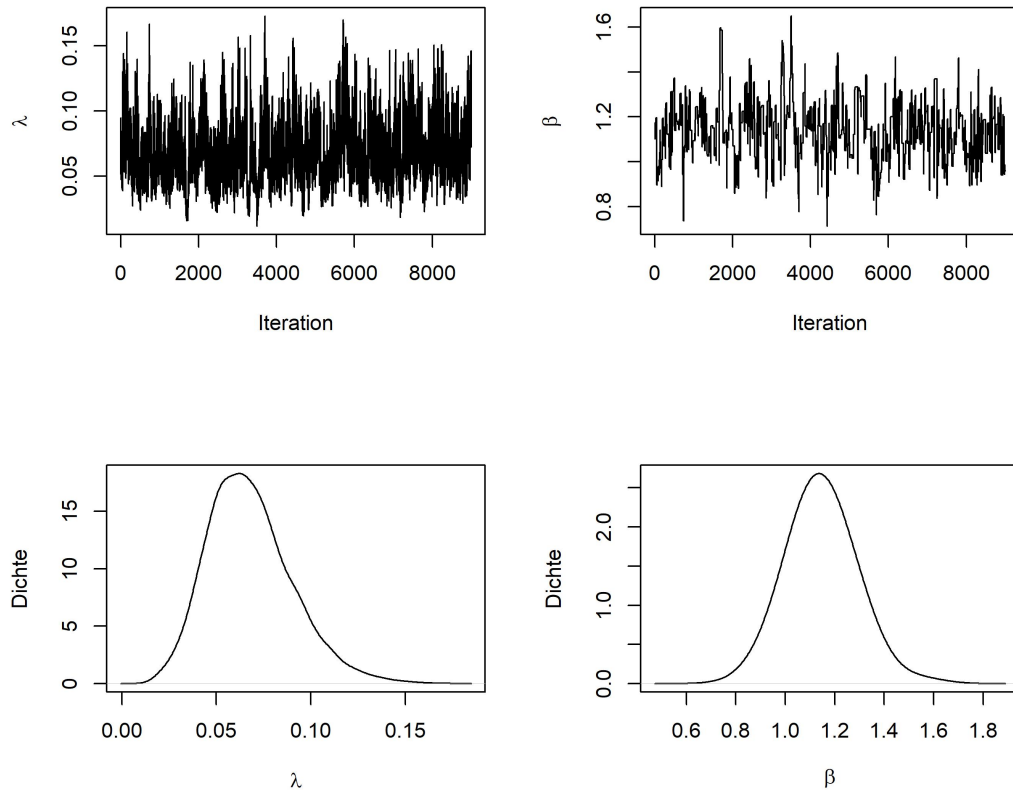


Abbildung 4.2: MCMC Iterationen und Kerndichteschätzer der A-posteriori Verteilungen der beiden Parameter der Weibullverteilung λ (links) und β (rechts).

Für die A-posteriori Verteilung gilt

$$\pi(\lambda, \beta | t_1, \dots, t_n) \propto \beta^{n-1} \lambda^{n-1} \exp\left(-\lambda \sum_{i=1}^n t_i^\beta + \beta \sum_{i=1}^n \log t_i\right).$$

Wegen

$$\lambda | \beta, t_1, \dots, t_n \sim \gamma\left(n, \sum_{i=1}^n t_i^\beta\right)$$

lassen sich Realisierungen der A-posteriori Verteilung nach Algorithmus 4.1 generieren, indem die Gammaverteilung in Schritt 2 entsprechend angepasst wird.

Zensierte Daten

Bei zensierten Daten in Form von $r < n$ beobachteten Ausfallzeiten t_1, \dots, t_r und $n - r$ rechtszensierten Beobachtungen, die eine gemeinsame Zensurzeit t_s haben, ist die Likelihood

gegeben durch

$$\begin{aligned}
l(\lambda, \beta | t_1, \dots, t_r) &= \prod_{i=1}^r \beta \lambda t_i^{\beta-1} \exp(-\lambda t_i^\beta) \left[\exp(-\lambda t_s^\beta) \right]^{n-r} \\
&= \beta^r \lambda^r \prod_{i=1}^r t_i^{\beta-1} \exp\left(-\lambda \left(\sum_{i=1}^r t_i^\beta + (n-r)t_s^\beta \right)\right) \\
&= \beta^r \lambda^r \exp\left(-\lambda \left(\sum_{i=1}^r t_i^\beta + (n-r)t_s^\beta \right) + (\beta-1) \sum_{i=1}^r \log t_i\right).
\end{aligned}$$

Nimmt man wie in (4.2.2) λ und β als a-priori unabhängig und gammaverteilt an, so erhält man die A-posteriori

$$\begin{aligned}
\pi(\lambda, \beta | t_1, \dots, t_r) &\propto \lambda^{r+a_\lambda-1} \beta^{r+a_\beta-1} \exp(-\lambda b_\lambda - \beta b_\beta) \\
&\quad \times \exp\left(-\lambda \left(\sum_{i=1}^r t_i^\beta + (n-r)t_s^\beta \right) + \beta \sum_{i=1}^r \log t_i\right).
\end{aligned}$$

Die bedingte Verteilung von $\lambda | \beta, t_1, \dots, t_r$ ist wieder eine Gammaverteilung

$$\lambda | \beta, t_1, \dots, t_r \sim \gamma\left(r + a_\lambda, b_\lambda + \sum_{i=1}^r t_i^\beta + (n-r)t_s^\beta\right).$$

Realisierungen der A-posteriori Verteilung können, ähnlich wie in Algorithmus 4.1 beschrieben, erzeugt werden. Im folgenden Beispiel soll dies illustriert werden.

Beispiel 4.3 (Fortsetzung von Beispiel 3.1). In Beispiel 3.1 wurden $n = 100$ Beobachtungen einer Weibullverteilung mit $\eta = \lambda = 1$ und $\beta = 3$ simuliert. Die Hyperparameter der A-priori (4.2.2) werden festgelegt als $a_\lambda = b_\lambda = 2$ und $a_\beta = 50, b_\beta = 20$.

Realisierungen der A-posteriori Verteilung werden, wie in Algorithmus 4.1 beschrieben, generiert. Es werden 10000 MCMC Samples erzeugt, wobei die ersten 1000 zur Burn-In Phase gezählt werden und nicht für die Schätzung verwendet werden. Die Vorschlagsdichte $q(\beta' | \beta)$ wird festgelegt als (vgl. [Robert, 2007, Example 6.3.2])

$$q(\beta' | \beta) = \frac{1}{\beta} \exp\left(-\frac{\beta'}{\beta}\right).$$

Die Histogramme in Abbildung 4.3 zeigen die A-posteriori Verteilungen von λ und β . Die A-posteriori Mittelwerte (Bayes Schätzer) und Varianzen sind

$$\lambda^B = 0.995, \text{Var}(\lambda) = 0.011 \quad \text{und} \quad \beta^B = 2.865, \text{Var}(\beta) = 0.048.$$

Die A-posteriori der Zuverlässigkeitsfunktion $R(t)$ kann approximiert werden, indem für eine Stichprobe von (λ, β)

$$R(t) = \exp(-\lambda t^\beta)$$

berechnet wird. Ebenso kann die A-posteriori der Ausfallrate $h(t)$ approximiert werden, indem man in

$$h(t) = \lambda \beta t^{\beta-1}$$

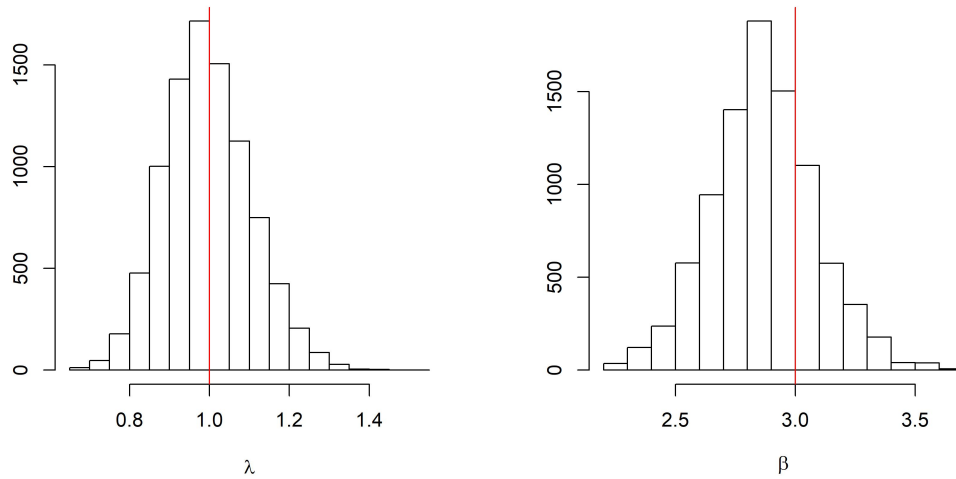


Abbildung 4.3: Histogramm der simulierten Stichprobe der A-posteriori von λ (links) und β (rechts). Die roten Linien kennzeichnen jeweils die wahren Werte.

einsetzt (vgl. [Hamada et al., 2008]). Abbildung 4.4 zeigt die A-posteriori Mittelwerte von $R(t)$ und $h(t)$. Zum Vergleich sind auch die beiden Kurven eingezeichnet, die man erhält, wenn man für λ und β die jeweiligen Bayes Schätzer in die Formeln für $R(t)$ und $h(t)$ einsetzt, d.h. die Kurven

$$R(t) = \exp\left(\lambda^B t^{\beta^B}\right) \quad \text{und} \quad h(t) = \lambda^B \beta^B t^{\beta^B - 1}.$$

Man sieht, dass die beiden Kurven in diesem Beispiel nicht sehr stark voneinander abweichen. \diamond

4.3 Log-Normalverteilung

Die Bayes'sche Analyse der Log-Normalverteilung hängt eng mit jener der Normalverteilung zusammen. Man ersetzt die Werte x_i einfach durch die logarithmierten Werte $\log x_i$.

Die konjugierte A-priori für die Parameter μ und $\tau = 1/\sigma^2$ einer Log-Normalverteilung ist eine Normal-Gamma-Verteilung $NG(m, d^2, a, b)$ (vgl. Beispiele 2.2 und 2.3). Die A-posteriori ist wieder eine Normal-Gamma-Verteilung mit den Hyperparametern

$$\begin{aligned} m^* &= \frac{m + d^2 n \bar{x}}{1 + d^2 n} \\ d^{*2} &= \frac{d^2}{1 + d^2 n} \\ a^* &= a + \frac{n}{2} \\ b^* &= b + \frac{1}{2} \left(\frac{n(\bar{x} - m)^2}{1 + nd^2} + \sum_{i=1}^n (\log x_i - \bar{x})^2 \right), \end{aligned}$$

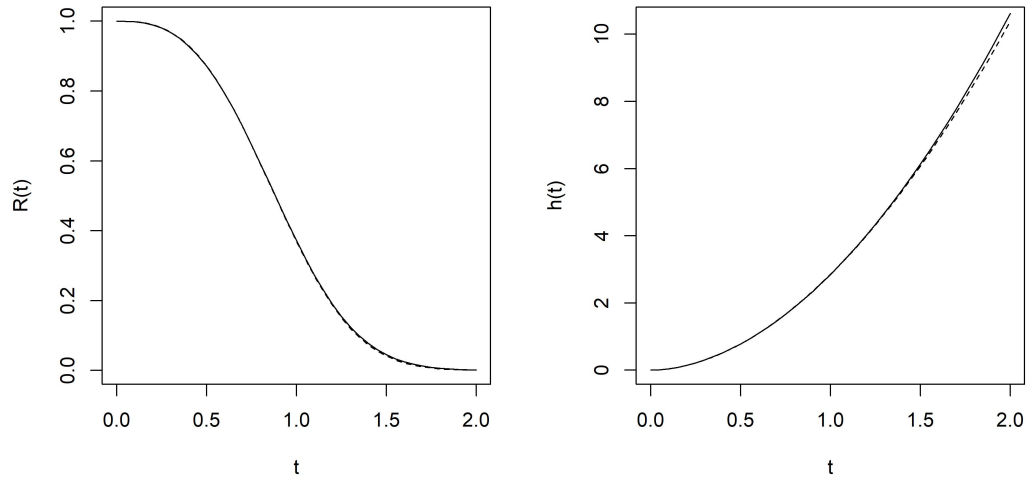


Abbildung 4.4: A-posteriori Mittelwert der Ausfallrate $R(t)$ (links) und der Zuverlässigkeitsfunktion $h(t)$ (rechts). Die gestrichelte Linien zeigen die beiden Kurven nach Einsetzen der jeweiligen Bayes Schätzer von λ und β .

wobei hier

$$\bar{x} = \sum_{i=1}^n \log x_i$$

ist. Der Bayes-Schätzer unter einer quadratischen Verlustfunktion ist gleich der A-posteriori Mittelwert m^* und ist ein gewichtetes Mittel aus dem A-priori Mittelwert m und dem Mittelwert der (logarithmierten) Daten \bar{x} .

Die Jeffrey's A-priori für (μ, σ^2) ist (vgl. Beispiel 2.7)

$$\pi_J(\mu, \sigma^2) \propto 1/\sigma^2.$$

Bei quadratischem Verlust ist der Bayes-Schätzer gleich dem Grenzwert von m^* für $d^2 \rightarrow \infty$ und ist daher gleich \bar{x} .

Zensierte Daten

Die Likelihood bei zensierten Daten mit Beobachtungen t_1, \dots, t_r und $n - r$ zensierten Daten mit Zensurzeit t_s hat die Form

$$l(\mu, \tau | t_1, \dots, t_r) = \frac{\tau^{r/2}}{(2\pi)^{r/2}} \exp\left(-\frac{\tau}{2} \sum_{i=1}^r (\log t_i - \mu)^2\right) \prod_{i=1}^r \frac{1}{t_i} \\ \times [1 - \Phi(\sqrt{\tau}(\log t_s - \mu))]^{n-r}.$$

Mit einer Normal-Gammaverteilung $NG(a, b, m, d^2)$ als A-priori ergibt sich die A-posteriori

$$\begin{aligned} \pi(\mu, \tau | t_1, \dots, t_r) &= \frac{b^a}{\Gamma(a)} \tau^{a-1} e^{-b\tau} \frac{\sqrt{\tau}}{\sqrt{2\pi d^2}} \exp\left(-\frac{\tau}{2d^2}(\mu - m)^2\right) \frac{\tau^{r/2}}{(2\pi)^{r/2}} \\ &\times \exp\left(-\frac{\tau}{2} \sum_{i=1}^r (\log t_i - \mu)^2\right) \prod_{i=1}^r \frac{1}{t_i} [1 - \Phi(\sqrt{\tau}(\log t_s - \mu))]^{n-r}. \end{aligned}$$

Mit der nicht informativen A-priori $\pi(\mu, \sigma^2) \propto 1/\sigma^2$ ist die A-posteriori Verteilung

$$\begin{aligned} \pi(\mu, \sigma^2 | t_1, \dots, t_r) &\propto \frac{1}{\sigma^2} \frac{1}{(2\pi\sigma^2)^{r/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^r (\log t_i - \mu)^2\right) \prod_{i=1}^r \frac{1}{t_i} \\ &\times \left[1 - \Phi\left(\frac{\log t_s - \mu}{\sigma}\right)\right]^{n-r}. \end{aligned}$$

4.4 Gammaverteilung

Die Likelihood bei vollständigen Daten t_1, \dots, t_n ist

$$\begin{aligned} l(\alpha, \beta | t_1, \dots, t_n) &= \prod_{i=1}^n \frac{\beta^\alpha}{\Gamma(\alpha)} t_i^{\alpha-1} \exp(-\beta t_i) \\ &= \frac{\beta^{\alpha n}}{\Gamma(\alpha)^n} \prod_{i=1}^n t_i^{\alpha-1} \exp\left(-\beta \sum_{i=1}^n t_i\right). \end{aligned}$$

Eine konjugierte A-priori ist definiert durch

$$\pi(\alpha, \beta) \propto \frac{\beta^{\alpha s-1}}{\Gamma(\alpha)^r} p^{\alpha-1} \exp(-\beta q) \quad (4.4.1)$$

mit Hyperparametern $p, q, r, s > 0$ [Miller, 1980]. Die A-posteriori ist proportional zu

$$\pi(\alpha, \beta | t_1, \dots, t_n) \propto \frac{\beta^{\alpha(s+n)-1}}{\Gamma(\alpha)^{r+n}} (pP)^{\alpha-1} \exp\{-\beta(q+S)\},$$

dabei bezeichnet $P = \prod_{i=1}^n t_i$ und $S = \sum_{i=1}^n t_i$. Die bedingte A-posteriori von β gegeben α ist die Dichte einer $\gamma(\alpha(s+n), q+S)$ -Verteilung. Die A-posteriori Randdichte von α ist proportional zu

$$\frac{\Gamma(\alpha(s+n))(pP)^{\alpha-1}}{\Gamma(\alpha)^{r+n}(q+S)^{\alpha(s+n)}}.$$

In [Miller, 1980] sind auch zwei nicht informative A-priori für die Gammaverteilung angegeben:

$$\pi(\alpha, \beta) \propto \frac{1}{\alpha\beta} \quad \text{und} \quad \pi(\alpha, \beta) \propto \frac{1}{\beta}.$$

Letztere entspricht in (4.4.1) dem Grenzfall $p \rightarrow 1, q \rightarrow 0, r \rightarrow 0$ und $s \rightarrow 0$. Die A-posteriori ist dann

$$\pi(\alpha, \beta | t_1, \dots, t_n) \propto \frac{\beta^{\alpha n-1}}{\Gamma(\alpha)^n} P^{\alpha-1} \exp\{-\beta S\}.$$

5 | Bayes'sche Analyse von beschleunigten Lebensdauermodellen

Dieses Kapitel beschreibt Bayes'sche Ansätze zu den beschleunigten Lebensdauermodellen aus Abschnitt 3.3.1. Der Fokus liegt wieder auf der Exponentialverteilung, da sie analytisch einfach zu handhaben ist. Das Power Rule Modell wird mit einem direkten Ansatz und als lineares Modell analysiert und in einem ausführlichen Beispiel illustriert. Genauso wird das einfache Step-Stress Modell aus 3.3.2 einer Bayes'schen Analyse unterzogen.

5.1 Power Rule Modell

In diesem Abschnitt betrachten wir das Power Rule Modell (3.3.1) in der Form

$$\lambda_i = \frac{V_i^p}{C}, \quad C, p > 0. \quad (5.1.1)$$

Es werden k Tests unter den Stressleveln V_1, \dots, V_k durchgeführt und ein Test unter V_i läuft so lange, bis $r_i \leq n_i$ Einheiten ausgefallen sind, sodass $n_i - r_i$ Ausfallzeiten rechtszensiert sind.

Die Ausfallrate λ_i ist abhängig vom Wert des Stressfaktors V_i . Die Daten für den Versuch unter dem i -ten Stresslevel werden als $D_i = \{V_i, n_i, r_i, t_{ij}\}$ zusammengefasst. Dann ist die Likelihood für λ_i unter V_i gleich

$$\begin{aligned} l(\lambda_i | D_i) &= \prod_{j=1}^{r_i} f(t_{ij} | \lambda_i) [1 - F(t_{ir_i} | \lambda_i)]^{n_i - r_i} = \prod_{j=1}^{r_i} \lambda_i \exp(-\lambda_i t_{ij}) \exp(-\lambda_i t_{ir_i})^{n_i - r_i} \\ &= \lambda_i^{r_i} \exp\left(-\lambda_i \left(\sum_{j=1}^{r_i} t_{ij} + (n_i - r_i)t_{ir_i}\right)\right) \\ &= \lambda_i^{r_i} \exp(-\lambda_i S_i). \end{aligned}$$

Es bezeichnet dabei

$$S_i = \sum_{j=1}^{r_i} t_{ij} + (n_i - r_i)t_{ir_i}$$

die „total time on test“ unter dem i -ten Stresslevel.

Nach k unabhängigen Versuchen unter den Stressleveln V_1, \dots, V_k erhält man die Daten $D = \{V_i, n_i, r_i, t_{ij}\}_{i=1}^k$ und die Likelihood für $\lambda = (\lambda_1, \dots, \lambda_k)$ ist gleich

$$l(\lambda|D) = \prod_{i=1}^k l(\lambda_i|D_i) = \prod_{i=1}^k \lambda_i^{r_i} \exp(-\lambda_i S_i).$$

Setzt man nun für λ_i die Beziehung (5.1.1) ein, so erhält man die Likelihood für C und p

$$\begin{aligned} l(C, p|D) &= \prod_{i=1}^k \frac{V_i^{pr_i}}{C^{r_i}} \exp\left(-\frac{V_i^p}{C} S_i\right) \\ &= \frac{1}{C^r} \prod_{i=1}^k V_i^{pr_i} \exp\left(-\frac{1}{C} \sum_{i=1}^k V_i^p S_i\right) \end{aligned}$$

mit $r = \sum_{i=1}^k r_i$.

Nun soll die Fisher-Information für dieses Modell hergeleitet werden, um daraus die Jeffreys A-priori abzuleiten. Die logarithmierte Likelihood ist

$$\log l(C, p|D) = -r \log C + p \sum_{i=1}^k r_i \log V_i - \frac{1}{C} \sum_{i=1}^k V_i^p S_i.$$

Die partiellen Ableitungen nach C und p sind

$$\begin{aligned} \frac{\partial}{\partial C} \log l(C, p|D) &= -\frac{r}{C} + \frac{1}{C^2} \sum_{i=1}^k V_i^p S_i, \\ \frac{\partial}{\partial p} \log l(C, p|D) &= \sum_{i=1}^k r_i \log V_i - \frac{1}{C} \sum_{i=1}^k V_i^p S_i \log V_i. \end{aligned}$$

Die zweiten partiellen Ableitungen sind

$$\begin{aligned} \frac{\partial^2}{\partial C^2} \log l(C, p|D) &= \frac{r}{C^2} - \frac{2}{C^3} \sum_{i=1}^k V_i^p S_i, \\ \frac{\partial^2}{\partial p^2} \log l(C, p|D) &= -\frac{1}{C} \sum_{i=1}^k V_i^p S_i (\log V_i)^2, \\ \frac{\partial^2}{\partial C \partial p} \log l(C, p|D) &= \frac{\partial^2}{\partial p \partial C} \log l(C, p|D) = \frac{1}{C^2} \sum_{i=1}^k V_i^p S_i \log V_i. \end{aligned}$$

Die Ausfallzeiten t_{ij} sind unabhängig und exponentialverteilt mit Erwartungswert $1/\lambda_i$. Daher gilt

$$\begin{aligned} \mathbb{E}[S_i] &= \mathbb{E} \left[\sum_{j=1}^{r_i} t_{ij} + (n_i - r_i) t_{ir_i} \right] = \sum_{j=1}^{r_i} \mathbb{E}[t_{ij}] + (n_i - r_i) \mathbb{E}[t_{ir_i}] \\ &= \frac{r_i}{\lambda_i} + \frac{n_i - r_i}{\lambda_i} = \frac{n_i}{\lambda_i} = \frac{n_i C}{V_i^p}. \end{aligned}$$

Die Fisher-Informationsmatrix ist nach Definition 2.7 gleich ($n = \sum_{i=1}^k n_i$)

$$\mathcal{I}(C, p) = \begin{pmatrix} \frac{n-r}{C^2} & -\frac{1}{C} \sum_{i=1}^k n_i \log V_i \\ -\frac{1}{C} \sum_{i=1}^k n_i \log V_i & \sum_{i=1}^k n_i (\log V_i)^2 \end{pmatrix}.$$

Die Jeffreys A-priori ist proportional zur Wurzel der Determinante von $\mathcal{I}(C, p)$:

$$\pi_J(C, p) \propto C^{-1}.$$

Die A-posteriori bei Verwendung von $\pi_J(C, p)$ ist

$$\pi(C, p|D) \propto C^{-r-1} \prod_{i=1}^k V_i^{p r_i} \exp\left(-\frac{1}{C} \sum_{i=1}^k V_i^p S_i\right).$$

Eine andere Möglichkeit ist es, die Parameter als a-priori unabhängig anzunehmen, sodass für die gemeinsame A-priori Verteilung gilt

$$\pi(C, p) = \pi(C)\pi(p).$$

Wenn man für C eine inverse Gammaverteilung und für p eine Gammaverteilung als A-priori ansetzt, dann ist

$$\pi(C, p) \propto C^{-a_C-1} p^{a_p-1} \exp\left(-\frac{b_C}{C} - p b_p\right).$$

Die Werte der Hyperparameter $a_C, b_C, a_p, b_p > 0$ sollen so gewählt werden, dass die a-priori vorhandene Information geeignet widerspiegelt wird. Die A-posteriori Verteilung hat dann die Form

$$\pi(C, p|D) \propto C^{-r-a_C-1} p^{a_p-1} \prod_{i=1}^k V_i^{p r_i} \exp\left(-\frac{1}{C} \left(b_C + \sum_{i=1}^k V_i^p S_i\right) - p b_p\right).$$

Die vollständige bedingte Verteilung von $C|p, D$ ist eine inverse Gammaverteilung mit den Parametern $r + a_C$ und $b_C + \sum_{i=1}^k V_i^p S_i$. Zur Simulation der gemeinsamen A-posteriori bietet sich, ähnlich wie in Beispiel 4.3 zur Weibullverteilung, eine Kombination aus Gibbs Sampling und Metropolis-Hastings Algorithmus an. Für die Vorschlagsdichte $q(p'|p)$ verwenden wir eine Normalverteilung. Damit ist $q(p'|p) = q(p|p')$ und die Akzeptanzwahrscheinlichkeit ρ berechnet sich nur als Quotient der A-posterioris. Man spricht von einem Random Walk Metropolis Algorithmus [Robert and Casella, 2013].

Algorithmus 5.1.

1. Lege Startwerte C_0 und p_0 fest.

Für $j = 1, 2, \dots$:

2. generiere $C_j \sim \gamma^{-1}\left(r + a_C, b_C + \sum_{i=1}^k V_i^{p_{j-1}} S_i\right)$,
3. generiere $p' \sim N(p_{j-1}, \sigma^2)$,

4. berechne

$$\rho = \min \left\{ 1, \frac{\pi(C_j, p' | D)}{\pi(C_j, p_{j-1} | D)} \right\}$$

und setze $p_j = p'$ mit Wahrscheinlichkeit ρ .

Die A-posteriori der Ausfallrate λ_u unter üblichen Betriebsbedingungen V_u kann approximiert werden, indem man Realisierungen von C und p aus $\pi(C, p | D)$ generiert und damit $\lambda_u = V_u^p / C$ berechnet.

5.1.1 Lineares Modell

In Abschnitt 3.3.1 wurde die alternative lineare Darstellung des Modells durch

$$Y = X\alpha + \epsilon$$

beschrieben. Für eine detaillierte Beschreibung der Bayes'schen Analyse linearer Modelle sei auf [Koch, 2000, Kapitel 4] verwiesen.

Der Fehlerterm ϵ soll die folgenden drei Bedingungen erfüllen:

- i. $\mathbb{E}[\epsilon] = 0$,
- ii. $\text{Cov}[\epsilon] = \mathbb{E}[\epsilon\epsilon^\top] = \sigma^2 I_k$ und
- iii. $\epsilon \sim N(0, \sigma^2 I_k)$.

Dann folgt Y einer k -dimensionalen Normalverteilung mit Mittelvektor $X\alpha$ und Kovarianzmatrix $\sigma^2 I_k$. Die Likelihoodfunktion ist

$$l(\alpha, \sigma^2 | Y) = \frac{1}{(2\pi\sigma^2)^{k/2}} \exp \left(-\frac{1}{2\sigma^2} (Y - X\alpha)'(Y - X\alpha) \right).$$

Die unbekannt Parameter in dem Regressionsmodell sind α und σ^2 , wobei wir zur Vereinfachung statt der Varianz σ^2 wieder die Präzision $\tau = 1/\sigma^2$ verwenden.

Die konjugierte A-priori für α und τ ist eine Normal-Gamma-Verteilung $NG(m, \Sigma, a, b)$ der Form

$$\alpha | \tau \sim N(m, \tau^{-1}\Sigma) \quad \text{und} \quad \tau \sim \gamma(a, b), \quad (5.1.2)$$

mit den Hyperparametern $m \in \mathbb{R}^2$, $\Sigma \in \mathbb{R}^{2 \times 2}$ und $a, b \in \mathbb{R}^+$. Die A-priori Dichte ist proportional zu

$$\pi(\alpha, \tau) \propto \tau^{a-1} \tau \exp \left(-\frac{\tau}{2} \left[2b + (\alpha - m)' \Sigma^{-1} (\alpha - m) \right] \right).$$

Die A-posteriori ist proportional zu

$$\pi(\alpha, \tau | Y) \propto \tau^{k/2+a-1} \tau \exp \left(-\frac{\tau}{2} \left[2b + (\alpha - m)' \Sigma^{-1} (\alpha - m) + (Y - X\alpha)'(Y - X\alpha) \right] \right).$$

Der Term in eckigen Klammern lässt sich umformen zu

$$\begin{aligned} 2b + (\alpha - m)' \Sigma^{-1} (\alpha - m) + (Y - X\alpha)'(Y - X\alpha) &= \\ &= \alpha'(X'X + \Sigma^{-1})\alpha - \alpha'(X'Y + \Sigma^{-1}m) - (Y'X + m'\Sigma^{-1})\alpha + (2b + Y'Y + m'\Sigma^{-1}m) \\ &= (\alpha - m^*)' \Sigma^{*-1} (\alpha - m^*) + (2b + Y'Y + m'\Sigma^{-1}m - m^*'\Sigma^{*-1}m^*) \end{aligned}$$

Die A-posteriori ist wieder eine Normal-Gamma-Verteilung $NG(m^*, \Sigma^*, a^*, b^*)$ mit den Parametern

$$\begin{aligned} m^* &= (X'X + \Sigma^{-1})^{-1}(X'Y + \Sigma^{-1}m), \\ \Sigma^* &= (X'X + \Sigma^{-1})^{-1}, \\ a^* &= a + \frac{k}{2}, \\ b^* &= 2b + Y'Y + m'\Sigma^{-1}m - m^{*\prime}\Sigma^{*-1}m^*. \end{aligned}$$

Der Bayes Schätzer für α unter einer quadratischen Verlustfunktion ist der A-posteriori Mittelwert, $\alpha^B = m^*$. Diesen kann man als gewichtetes Mittel des Kleinstquadrat-Schätzers $\hat{\alpha} = (X'X)^{-1}X'Y$ und des A-priori Mittelwertes m schreiben, denn

$$\begin{aligned} \alpha^B &= (X'X + \Sigma^{-1})^{-1}(X'Y + \Sigma^{-1}m) \\ &= (X'X + \Sigma^{-1})^{-1}[(X'X)(X'X)^{-1}X'Y + \Sigma^{-1}m] \\ &= (X'X + \Sigma^{-1})^{-1}[(X'X)\hat{\alpha} + \Sigma^{-1}m]. \end{aligned}$$

Falls keine Vorinformation vorhanden ist, so kann man die nicht informative A-priori $\pi(\alpha, \tau) \propto 1/\tau$ anwenden. Der Bayes Schätzer stimmt dann mit dem Kleinst-Quadrate Schätzer $\hat{\alpha}$ überein. Formal kann man diese A-priori durch den Grenzwert $\Sigma^{-1} \rightarrow 0$ aus der konjugierten A-priori (5.1.2) erhalten.

5.1.2 Beispiel mit simulierten Daten

Beispiel 5.1. Die vorgestellten Methoden zur Analyse des Power Rule Modells bei exponentialverteilten Daten werden in diesem Beispiel auf simulierte Daten angewendet. Es wird angenommen, dass ein beschleunigter Lebensdauer-versuch unter $k = 5$ verschiedenen Stress-leveln durchgeführt wird, wobei $n_i = n = 45$ Komponenten so lange getestet werden, bis r_i Ausfälle aufgetreten sind. Die Parameter der Beschleunigungsfunktion werden festgelegt als $C = 200$ und $p = 2$ und die λ_i werden nach dem Power Rule Modell als

$$\lambda_i = \frac{V_i^p}{C}, i = 1, \dots, 5$$

berechnet. Eine Zusammenfassung der Versuchsparmeter ist in Tabelle 5.1 gegeben.

V_i	5	7	10	20	25
n_i	45	45	45	45	45
r_i	15	15	25	30	35
λ_i	0.125	0.245	0.500	2.000	3.125

Tabelle 5.1: Anordnung eines simulierten beschleunigten Lebensdauer-versuches.

Das Ziel des Versuches ist es, die Ausfallrate λ_u bei üblicher Belastung $V_u = 1.5$ zu schätzen. Dazu werden zum einen klassische Schätzmethode wie Maximum Likelihood Schätzung und die Methode der kleinsten Quadrate bei Formulierung als lineares Modell herangezogen. Zum anderen werden Bayes Schätzer mit informativer A-priori, mit nicht informativer A-priori und aus dem linearen Modell berechnet.

1. Maximum Likelihood Schätzung

Für die Maximum Likelihood Schätzung aus Abschnitt 3.3.1 wurden die Werte $\hat{\theta}_i$ nach Formel (3.3.5) berechnet. Durch Lösen der Gleichung (3.3.7) erhält man den Maximum Likelihood (ML) Schätzer von p . Bei der Schätzung von C ist zu beachten, dass im modifizierten Modell

$$\theta_i = \frac{C_0}{(V_i/\dot{V})^p}$$

der Wert von C beeinflusst wird, es gilt $C_0 = C/\dot{V}^p$, nicht jedoch der Wert von p (vgl. [Singpurwalla, 1971]). Einsetzen in Gleichung (3.3.6) ergibt für den ML-Schätzer von C_0 den Wert $\hat{C}_0 = 1.17$. Um daraus den Wert des ML-Schätzers für C zu erhalten, muss dieser noch transformiert werden. Die ML Schätzwerte für C und p sind in Tabelle 5.2 angegeben.

Ein Schätzer für die Ausfallrate unter der Belastung $V_u = 1.5$ ist gegeben durch

$$\hat{\lambda}_u = \frac{(V_u/\dot{V})^{\hat{p}}}{\hat{C}_0} = 0.009.$$

2. Kleinste-Quadrate Schätzer

Der KQ-Schätzer für α im Modell $Y = X\alpha + \epsilon$ mit $\alpha = (\log C, -p)$,

$$Y = \begin{pmatrix} \log \hat{\theta}_1 \\ \log \hat{\theta}_2 \\ \vdots \\ \log \hat{\theta}_k \end{pmatrix}, X = \begin{pmatrix} 1 & \log V_1 \\ 1 & \log V_2 \\ \vdots & \vdots \\ 1 & \log V_k \end{pmatrix} \text{ und } \epsilon = \begin{pmatrix} \log \epsilon_1 \\ \log \epsilon_2 \\ \vdots \\ \log \epsilon_k \end{pmatrix},$$

ist $\hat{\alpha} = (X'X)^{-1}X'Y = (5.55, -2.06)$. Für C und p ergeben sich die Schätzwerte $\tilde{C} = 259.36$ und $\tilde{p} = 2.06$. Der Schätzwert für λ_u nach diesem Modell ist $\tilde{\lambda}_u = V_u^{\tilde{p}}/\tilde{C} = 0.0089$. Diese Ergebnisse sind vergleichbar mit jenen der Maximum Likelihood Methode.

3. Bayes Schätzung

Für die Bayes Schätzung werden die beiden in Abschnitt 5.1 dargestellten A-priori Verteilungen verwendet. Realisierungen der gemeinsamen A-posteriori von C und p werden unter Anwendung von Algorithmus 5.1 generiert. Im Fall der Jeffreys A-priori ist zu beachten, dass C_j aus einer $\gamma^{-1}(r, \sum_{i=1}^k V_i^{p_j-1} S_i)$ -Verteilung gezogen wird.

Die Werte für die a-priori Hyperparameter der inversen Gammaverteilung sind $a_C = 25$ und $b_C = 3500$ und für die Gammaverteilung werden $a_p = b_p = 1$ festgelegt.

Die A-posteriori Mittelwerte von C und p aus einer Stichprobe von 10000 und einer Burn-In Periode von 1000 sind in Tabelle 5.2 angegeben. In Abbildung 5.1 sind die Kerndichteschätzer der A-posteriori Randverteilungen für beide A-priori Verteilungen dargestellt.

Mit den klassischen Methoden wurden in diesem Beispiel nur Punktschätzer für λ_u berechnet. Im Gegensatz dazu bringt die Bayes'sche Analyse, neben dem Miteinbeziehen von Vorinformation, den Vorteil, dass eine Wahrscheinlichkeitsverteilung für λ_u ausgegeben wird. Durch Berechnen von $\lambda_u = V_u^p/C$ mit den simulierten a-posteriori Werten von C und p kann eine Stichprobe der A-posteriori Verteilung von λ_u gewonnen werden. Das Histogramm einer solchen Stichprobe ist in Abbildung 5.2 zu sehen.

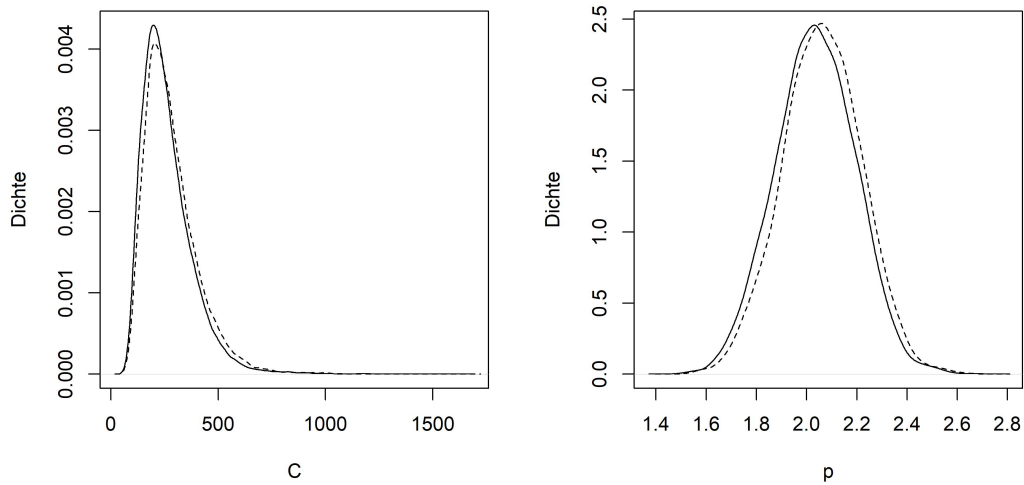


Abbildung 5.1: Kerndichteschätzung der A-posteriori Randverteilungen von C (links) und p (rechts) unter einer Jeffreys A-priori (gestrichelte Linie) und einer informativen A-priori (durchgehende Linie).

	MLE	KQ	Jeffreys	Informativ	Bayes Linear	wahrer Wert
C	252.03	259.36	242.11	229.79	224.53	200
p	2.05	2.06	2.01	2.03	1.99	2.01
λ_u	0.009	0.0089	0.0105	0.0110	0.0101	0.0112

Tabelle 5.2: Schätzwerte für C , p und λ_u unter Anwendung verschiedener Schätzmethoden.

4. Bayes Lineares Modell

Für die Normal-Gamma-Verteilung werden die folgenden Hyperparameter verwendet:

$$m = (5, -1), \quad \Sigma = \text{diag}(10, 100), \quad a = 1, \quad b = 1.$$

Die A-posteriori ist eine Normal-Gamma-Verteilung, deren Hyperparameter man durch updaten, wie in Abschnitt 5.1.1 beschrieben, erhält:

$$m^* = (5.41, -2.00), \quad \Sigma^* = \begin{pmatrix} 3.06 & -1.19 \\ -1.19 & 0.49 \end{pmatrix}, \quad a^* = 3.5, \quad b^* = 21.77.$$

Die Bayes Schätzer von C und p sind $e^{5.41} \approx 224.53$ und 2.00. Mit diesen Schätzwerten ist der Bayes Schätzer von λ_u gleich 0.010. Bei Anwendung einer Jeffreys A-priori erhält man die Kleinste-Quadrate Schätzer.

Zusammenfassung

Aus Tabelle 5.2 ist ersichtlich, dass es bei allen fünf Methoden keine starken Abweichungen der Schätzwerte vom wahren Wert von p gibt. Die Schätzwerte für C liegen zwischen 224.53

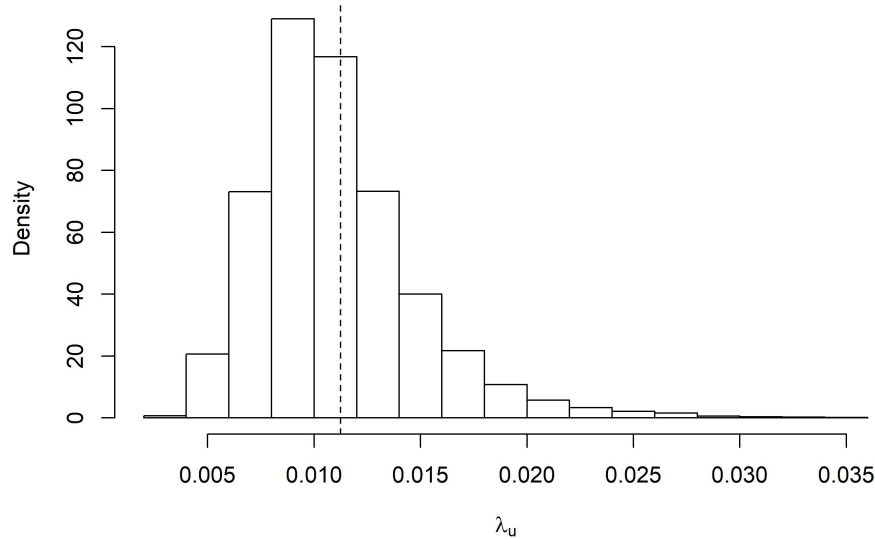


Abbildung 5.2: Histogramm einer Stichprobe aus der A-posteriori Verteilung von λ_u bei Verwendung einer informativen A-priori. Die gestrichelte Linie kennzeichnet den wahren Wert $\lambda_u = 0.1125$.

und 259.36, wobei die beiden Methoden mit informativen A-prioris („Informativ“ und „Bayes Linear“) im Vergleich zu den klassischen Methoden und der Jeffreys A-priori näher am wahren Wert von C liegen. Der Schätzer aus dem Bayes'schen linearen Modell ist dem wahren Wert von C am nächsten.

Ziel des beschleunigten Lebensdauerexperimentes ist es, die Ausfallrate unter normalen Bedingungen zu schätzen. Hier sind die Bayes'schen Methoden jene, die den wahren Wert λ_u am besten schätzen, während die beiden klassischen Varianten am weitesten daneben liegen. \diamond

5.2 Step-Stress Modell

Das einfache Step-Stress Modell aus Abschnitt 3.3.2 wird nun aus Bayes'scher Sicht betrachtet. Um die Berechnungen zu vereinfachen, wird die Exponentialverteilung mit $\lambda = 1/\theta$ parametrisiert. Die Likelihood im Cumulative Exposure Modell ist dann

$$l(\lambda_1, \lambda_2) = \begin{cases} \lambda_1^r \exp(-\lambda_1 A) & n_1 = r, \\ \lambda_2^r \exp(-\lambda_2 B(\lambda_1, \lambda_2)) & n_1 = 0, \\ \lambda_1^{n_1} \lambda_2^{n_2} \exp(-\lambda_1 C_1 - \lambda_2 C_2) & 1 \leq n_1 \leq r-1, \end{cases} \quad (5.2.1)$$

mit

$$\begin{aligned}
 A &= \sum_{i=1}^r t_{(i)} + (n-r)t_{(r)}, \\
 B(\lambda_1, \lambda_2) &= \sum_{i=1}^r \left(\frac{\lambda_1}{\lambda_2} \tau - \tau + t_{(i)} \right) + (n-r) \left(\frac{\lambda_1}{\lambda_2} \tau - \tau + t_{(r)} \right), \\
 C_1 &= \sum_{i=1}^{n_1} t_{(i)} + (n-n_1)\tau, \\
 C_2 &= \sum_{i=n_1+1}^r (t_{(i)} - \tau) + (n-r)(t_{(r)} - \tau).
 \end{aligned}$$

Die gemeinsame A-priori für λ_1 und λ_2 setzt sich aus zwei unabhängigen konjugierten Gammaverteilungen zusammen:

$$\begin{aligned}
 \pi(\lambda_1, \lambda_2) &= \pi_1(\lambda_1)\pi_2(\lambda_2) \\
 &\propto \lambda_1^{a_1-1} \lambda_2^{a_2-1} e^{-b_1\lambda_1 - b_2\lambda_2}, \quad a_i, b_i > 0, i = 1, 2.
 \end{aligned}$$

Im Fall $1 \leq n_1 \leq r-1$ ist die A-posteriori proportional zu

$$\pi(\lambda_1, \lambda_2 | D) \propto \lambda_1^{n_1+a_1-1} \lambda_2^{n_2+a_2-1} \exp(-\lambda_1(b_1 + C_1) - \lambda_2(b_2 + C_2)).$$

Die A-posteriori Verteilungen der beiden Spezialfälle $n_1 = r$ und $n_1 = 0$ ergeben sich einfach durch entsprechendes Einsetzen. Die A-posteriori Randverteilungen von λ_1 und λ_2 sind Gammaverteilungen. Es gilt für $1 \leq n_1 \leq r-1$

$$\begin{aligned}
 \pi(\lambda_1 | D) &\propto \lambda_1^{n_1+a_1-1} e^{-\lambda_1(b_1+C_1)} \quad \text{und} \\
 \pi(\lambda_2 | D) &\propto \lambda_2^{n_2+a_2-1} e^{-\lambda_2(b_2+C_2)}.
 \end{aligned}$$

Bei quadratischem Verlust sind die Bayes Schätzer von λ_1 und λ_2 gegeben durch

$$\lambda_1^B = \frac{n_1 + a_1}{b_1 + C_1}, \quad \lambda_2^B = \frac{n_2 + a_2}{b_2 + C_2}.$$

Beispiel 5.2. Zur Illustration dieses Step Stress Modells wurde ein Datensatz nach dem Modell (3.3.11) simuliert. Die Parameter werden auf folgende Werte festgesetzt

$$n = 30, \quad r = 22, \quad \lambda_1 = 1/15, \quad \lambda_2 = 1/5 \quad \text{und} \quad \tau = 4.5.$$

Die Ausfallzeiten sind in Tabelle 5.3 angegeben.

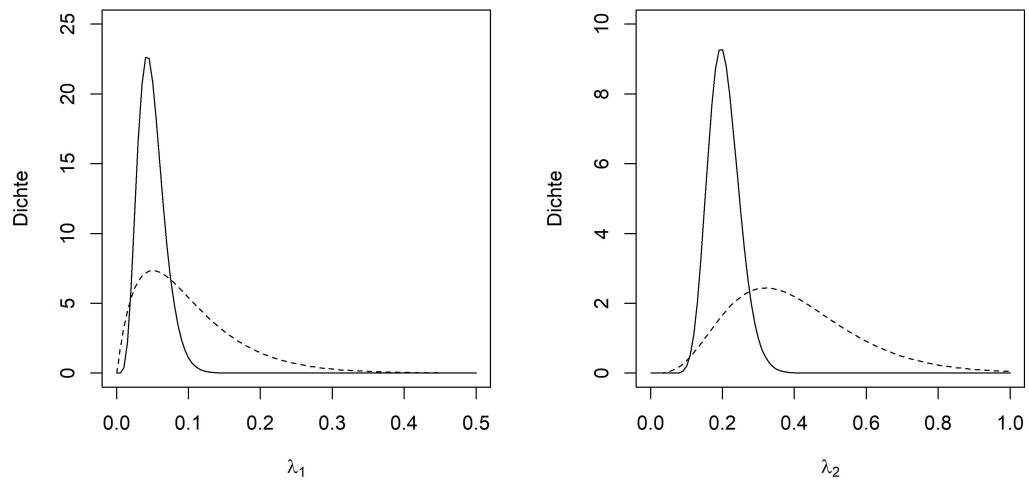
Die Maximum Likelihood Schätzer werden, wie in Abschnitt 3.3.2 beschrieben, berechnet und sind gleich $\hat{\lambda}_1 = 1/\hat{\theta}_1 = 0.040$ und $\hat{\lambda}_2 = 1/\hat{\theta}_2 = 0.178$.

Für die Bayes'sche Analyse werden zunächst die A-priori Hyperparameter der beiden Gammaverteilungen $\pi_1(\lambda_1)$ und $\pi_2(\lambda_2)$ festgelegt. Aus Vorinformation ist bekannt, dass die Ausfallrate unter dem ersten Stresslevel, λ_1 , im Mittel bei 0.1 liegt mit einer Varianz von 0.005. Daraus ergeben sich die Hyperparameter $a_1 = 2$ und $b_1 = 20$. Für die Ausfallrate unter dem zweiten Stresslevel wird angenommen, dass der Erwartungswert von λ_2 bei 0.4 liegt und die Varianz 0.032 ist. Daraus folgt $a_2 = 5$ und $b_2 = 12.5$.

Die A-posteriori Randverteilungen von λ_1 bzw. λ_2 sind eine $\gamma(7, 142)$ - bzw. eine $\gamma(22, 107)$ -Verteilung und sind in Abbildung 5.3 dargestellt. Die Bayes Schätzer sind $\lambda_1^B = 0.049$ und $\lambda_2^B = 0.204$. Sie sind im Vergleich zu den ML-Schätzern näher an den wahren Werten von λ_1 und λ_2 . \diamond

Stresslevel	Ausfallzeit						
1	0.64	0.69	1.62	2.38	4.23		
2	4.69	4.70	4.98	5.62	6.01	6.04	6.75
	6.92	7.00	7.25	7.50	8.11	8.33	8.65
	8.90	9.16	10.76				

Tabelle 5.3: Beispieldatensatz

Abbildung 5.3: A-priori (gestrichelte Linie) und A-posteriori (durchgehende Linie) Randverteilungen der Ausfallraten unter den beiden Stressleveln λ_1, λ_2 .

6 | Zusammenfassung

In zahlreichen Industriebereichen werden, auch bedingt durch den technologischen Fortschritt, immer höhere Qualitätsanforderungen an Produkte und einzelne Komponenten gestellt. Um diese Anforderungen garantieren zu können, werden häufig statistische Methoden eingesetzt. Eine Herausforderung für die Statistik tritt insbesondere bei hochzuverlässigen Produkten wie etwa Halbleiterbauteilen auf, bei denen in Zuverlässigkeitstests oft wenige bis gar keine Ausfälle beobachtet werden. Deshalb ist es umso wichtiger, dass abseits von beobachteten Daten jede verfügbare Art von Information über das Produkt zur Beurteilung der Zuverlässigkeit verwendet wird.

In dieser Arbeit werden Bayes'sche Methoden im Bereich der Zuverlässigkeitstheorie betrachtet, die eine Möglichkeit bieten, externe Informationen auf kohärente Weise in das statistische Modell einzubringen. Im Gegensatz zu klassischen statistischen Methoden ist es in der Bayes Statistik möglich, Vorinformation, die von Produktingenieuren oder aus anderen Zuverlässigkeitstests vorhanden ist, zu nutzen.

Eine häufige Anwendung in der Zuverlässigkeitstheorie ist die Schätzung von Parametern bestimmter Lebensdauerverteilungen. In der Bayes Statistik betrachtet man den Parameter als zufällig und weist ihm eine Wahrscheinlichkeitsverteilung zu. Die Vorinformation wird in der A-priori Verteilung zusammengefasst und enthält die gesamte Information über den Parameter, bevor Daten aus einem Lebensdauertest vorhanden sind. Die Beobachtung von Daten liefert neue Information und das gesamte Wissen über den Parameter wird in der A-posteriori Verteilung dargestellt.

Die Wahl der A-priori Verteilung stellt einen zentralen Punkt der Analyse dar. Es werden in der Arbeit sowohl konjugierte Verteilungen, welche die gleiche Struktur wie die Verteilung der Daten aufweisen, als auch nicht informative A-priori Verteilungen, bei denen kaum oder keine Information über den Parameter vorliegt, betrachtet.

Das Bayes'sche Verfahren wird am Beispiel der Exponentialverteilung konkret dargestellt. Die Familie der Gammaverteilungen stellen die zur Exponentialverteilung konjugierte Verteilungsfamilie dar. Durch geeignete Wahl der Hyperparameter wird die a-priori vorhandene Information über den Parameter nach Beobachtung der Daten upgedatet. Statt eines einzelnen Punktschätzers, wie in der klassischen Statistik, erhält man a-posteriori eine Wahrscheinlichkeitsverteilung für den Parameter.

Die Weibullverteilung ist eine häufig verwendete Lebensdauerverteilung, die im Vergleich zur Exponentialverteilung flexibler ist, weil sie nicht nur die Modellierung konstanter, sondern auch steigender oder fallender Ausfallraten zulässt. Für die Parameter der Weibullverteilung gibt es keine gemeinsame stetige konjugierte A-priori Verteilung. Daher werden sogenannte bedingte konjugierte Verteilungen verwendet, konkret zwei unabhängige Gammaverteilungen.

Die A-posteriori Verteilung lässt sich dann nicht mehr in geschlossener Form darstellen, weshalb Markov Chain Monte Carlo (MCMC) Methoden angewendet werden, um Stichproben aus der A-posteriori Verteilung der Parameter zu ziehen.

Bei besonders zuverlässigen Bauteilen oder Produkten ist ein Test unter normalen Betriebsbedingungen zeitlich oder wirtschaftlich nicht machbar. Um mehr Ausfälle in angemessener Zeit herbeizuführen werden daher beschleunigte Lebensdauerversuche durchgeführt, in welchen die Bauteile erhöhten Belastungen ausgesetzt werden. Die Bayes Statistik bietet auch für die Analyse von solchen Versuchen Methoden, um a-priori Information im Modell zu berücksichtigen. Dabei wird angenommen, dass der Parameter der Lebensdauerverteilung durch eine Funktion von einer oder mehreren Stressvariablen abhängt. Diese Funktion ist bekannt bis auf bestimmte Parameter, für die es gilt, eine A-priori Verteilung festzulegen. Mit den a-posteriori Schätzern können dann die a-posteriori Parameter der Lebensdauerverteilung unter normalen Werten der Stressvariablen durch die funktionelle Beziehung zwischen Stress und Lebensdauer approximiert werden. Eine einfache Beziehung dieser Art stellt das Power Rule Modell dar, auf dem ein Fokus der Arbeit liegt. Dabei wird die mittlere Lebensdauer als eine Funktion modelliert, die mit einer Potenz des Stressfaktors abnimmt.

Das Power Rule Modell kann auch als lineares Modell dargestellt werden und die Parameter können somit auch durch ein Bayes'sches lineares Modell geschätzt werden. Die Vorinformation wird dabei durch eine (konjugierte) Normal-Gamma-Verteilung modelliert und der a-posteriori Mittelwert ergibt den Bayes Schätzer. Diese Methode hat in Beispiel 5.1 im Vergleich zur Schätzung mittels eines klassischen linearen Modells bessere Ergebnisse gezeigt.

Zusammengefasst lässt sich sagen, dass sich die Methoden der Bayes Statistik für Anwendungen in der Zuverlässigkeitsanalyse gut eignen. Eine große Herausforderung in der Praxis ist das Übersetzen der vorhandenen Information in eine passende A-priori Verteilung. Sobald dieser Schritt getan ist, kann z.B. mit Hilfe der vorgestellten Algorithmen eine umfangreiche Analyse der A-posteriori Verteilung vorgenommen werden. Ein großer Vorteil gegenüber frequentistischen Methoden ist, dass Fragen zur Wahrscheinlichkeit, mit der der Parameter einen bestimmten Wert übersteigt oder in einem bestimmten Intervall liegt, direkt beantwortet werden können.

Literaturverzeichnis

- [Balakrishnan et al., 2007] Balakrishnan, N., Kundu, D., Ng, H. K. T., and Kannan, N. (2007). Point and interval estimation for a simple step-stress model with Type-II censoring. *Journal of Quality Technology*, 39(1):35–47.
- [Berger, 2013] Berger, J. O. (2013). *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media.
- [Casella and Berger, 2002] Casella, G. and Berger, R. L. (2002). *Statistical inference*. Duxbury Pacific Grove, CA.
- [DeGroot, 2004] DeGroot, M. H. (2004). *Optimal Statistical Decisions*, volume 82. John Wiley & Sons.
- [Embrechts et al., 2013] Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media.
- [Hamada et al., 2008] Hamada, M. S., Wilson, A., Reese, C. S., and Martz, H. (2008). *Bayesian reliability*. Springer Science & Business Media.
- [Ibrahim et al., 2013] Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2013). *Bayesian survival analysis*. Springer Science & Business Media.
- [Joarder et al., 2011] Joarder, A., Krishna, H., and Kundu, D. (2011). Inferences on Weibull parameters with conventional Type-I censoring. *Computational Statistics & Data Analysis*, 55(1):1–11.
- [Kaminskiy and Krivtsov, 2005] Kaminskiy, M. P. and Krivtsov, V. V. (2005). A simple procedure for Bayesian estimation of the Weibull distribution. *IEEE Transactions on Reliability*, 54(4):612–616.
- [Koch, 2000] Koch, K.-R. (2000). *Einführung in die Bayes-Statistik*. Springer-Verlag.
- [Kusolitsch, 2014] Kusolitsch, N. (2014). *Maß- und Wahrscheinlichkeitstheorie: Eine Einführung*. Springer-Verlag.
- [Lawless, 2011] Lawless, J. F. (2011). *Statistical models and methods for lifetime data*, volume 362. John Wiley & Sons.
- [Lehmann and Casella, 2006] Lehmann, E. L. and Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.

- [Mann et al., 1974] Mann, N. R., Schafer, R. E., and Singpurwalla, N. D. (1974). *Methods for statistical analysis of reliability and life data*. Wiley.
- [Meeker and Escobar, 2014] Meeker, W. Q. and Escobar, L. A. (2014). *Statistical methods for reliability data*. John Wiley & Sons.
- [Miller, 1980] Miller, R. B. (1980). Bayesian analysis of the two-parameter gamma distribution. *Technometrics*, 22(1):65–69.
- [Nelson, 2009] Nelson, W. B. (2009). *Accelerated testing: statistical models, test plans, and data analysis*, volume 344. John Wiley & Sons.
- [Robert, 2007] Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media.
- [Robert and Casella, 2013] Robert, C. and Casella, G. (2013). *Monte Carlo statistical methods*. Springer Science & Business Media.
- [Singpurwalla, 1971] Singpurwalla, N. D. (1971). A problem in accelerated life testing. *Journal of the American Statistical Association*, 66(336):841–845.
- [Singpurwalla, 2006] Singpurwalla, N. D. (2006). *Reliability and risk: a Bayesian perspective*. John Wiley & Sons.
- [Soland, 1969] Soland, R. M. (1969). Bayesian analysis of the Weibull process with unknown scale and shape parameters. *IEEE Transactions on Reliability*, 18(4):181–184.
- [Sun, 1997] Sun, D. (1997). A note on noninformative priors for Weibull distributions. *Journal of Statistical Planning and Inference*, 61(2):319–338.