

# KAPITEL 16

---

## Herleitung einer Berechnung von Littlewood und Strigini

---

### 16.1 Bayessche Wahrscheinlichkeitsrechnung

Gesucht sei die Wahrscheinlichkeitsverteilung der ZV  $T$ , d.h. der Zeit bis zum nächsten Programmfehler. Sei  $\lambda$  die Rate bzw.  $\Theta = \lambda^{-1}$  die mittlere Zeit, mit bzw. in der ein Fehler im Programm auftritt. Damit ist  $T$  aufgrund des vorliegenden Poisson-Prozesses Erlang-verteilt mit Parameter  $n = 1$ :

$$p_T(k) = \text{Erl}(k; \lambda, n)|_{n=1} = \lambda e^{-\lambda k}$$

Dieses  $\lambda$  ist aber ebenfalls nicht genau bekannt und damit eine ZV. Man erhält also für  $T$  lediglich eine die bedingte Wahrscheinlichkeitsdichte  $p_{T|\lambda}$  für die Verteilung der ZV  $T$ . Die Wahrscheinlichkeitstheorie nach Bayes erlaubt es jedoch nun, auf der Basis von Beobachtungen, in diesem Fall der Zählung von aufgetretenen Programmfehlern  $x$  über einen bestimmten (längeren) Zeitraum  $t_0$ , eine zuvor *a priori* angenommene Verteilung, in diesem Fall der der Rate  $\lambda$ , dahingehend anzupassen. Aus dem Satz von Bayes erhält man

$$p_{\lambda|t_0, x} = \frac{p_{x, t_0|\lambda} \cdot p_{\lambda}}{\int_0^{\infty} p_{x, t_0|\lambda}(x', t'_0, r) \cdot p_{\lambda}(r) dr} \quad (16.1)$$

wobei  $p_{x, t_0|\lambda}$  die Likelihood Funktion von  $\lambda$  ist. Diese ist aufgrund des vorliegenden Poisson-Prozesses poisson-verteilt nach:

$$L(\lambda|x, t_0) = p_{x, t_0|\lambda}(x', t'_0, r) = \frac{(t'_0 r)^{x'}}{x'!} e^{-t'_0 r}$$

Bei der Wahl der apriori Verteilung werden in den folgenden Abschnitte nun zwei unterschiedliche Ansätze verfolgt.

## 16.2 Homogene Anpassung

Die Anpassung der Verteilung von  $\lambda$  aufgrund des Wissens durch die Beobachtungen sei *homogen*, d.h. in Beibehaltung der Familie der apriori Familie für die aposteriori Annahme und damit lediglich durch Änderung der der Verteilung zugehörigen *Hyper-Parameter*. Damit muss die apriori (und damit auch aposteriori) Verteilung *konjugiert* zu der Poisson Verteilung sein. Dies ist die sogenannte *Gamma-Verteilung*, gegeben durch

$$\text{Gam}_\lambda(r; a, b) = b^a r^{a-1} \frac{e^{-br}}{\Gamma(a)} \quad (16.2)$$

wobei  $\Gamma(\cdot)$  die wie folgt definierte *Gamma Funktion* sei:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

und  $a$  und  $b$  die Hyper-Parameter der Verteilung, auf die im folgenden noch näher eingegangen werden wird. Damit erhält man also aus Glg. (16.1) die a posteriori Verteilungsdichte für die Rate nach oben erwähnte Beobachtung

$$\begin{aligned} p_{\lambda|x, t_0}(r, x', t'_0) &= \frac{p_{x, t_0|\lambda} \cdot p_\lambda}{\int_0^\infty p_{x, t_0|\lambda}(x', t'_0, r) \cdot p_\lambda(r) dr} = \frac{p_{x, t_0|\lambda}(x', t'_0, r) \cdot \text{Gam}_\lambda(r; a, b)}{\int_0^\infty p_{x, t_0|\lambda}(x', t'_0, r) \cdot \text{Gam}_\lambda(r; a, b) d\lambda} \\ &= \frac{\frac{(t'_0 r)^{x'}}{x'!} e^{-t'_0 r} \cdot b^a r^{a-1} \frac{e^{-br}}{\Gamma(a)}}{\int_0^\infty \frac{(t'_0 r)^{x'}}{x'!} e^{-t'_0 r} \cdot b^a r^{a-1} \frac{e^{-br}}{\Gamma(a)} d\lambda} = \frac{r^{(a+x')-1} e^{-(b+t'_0)r}}{\int_0^\infty r^{(a+x')-1} e^{-(b+t'_0)r} d\lambda} \\ &= \frac{r^{(a+x')-1} e^{-(b+t'_0)r} (b+t'_0)^{a+x'}}{\Gamma(a+x')} = \text{Gam}_{\lambda|x, t_0}(r; a+x', b+t'_0) \end{aligned}$$

Und so lässt sich nun die angepasste Verteilung  $p(t|x, t_0)$  von  $T$  nach den Beobachtungen angeben:

$$\begin{aligned}
 p_{T|x, t_0}(t, x', t'_0) &= \int_0^{\infty} p_{T|\lambda}(t, r) p_{\lambda|x, t_0}(r, x', t'_0) dr \\
 &= \int_0^{\infty} r e^{-rt} \cdot \frac{(b + t'_0)^{a+x'} \lambda^{a+x'-1} e^{-(b+t'_0)\lambda}}{\Gamma(a+x')} dr \\
 &= \frac{(b + t'_0)^{a+x'}}{\Gamma(a+x')} \int_0^{\infty} \lambda^{a+x'} e^{-(b+t'_0+t)\lambda} d\lambda \\
 &\stackrel{\text{Subst. } z=(\dots)r}{=} \frac{(b + t'_0)^{a+x'}}{\Gamma(a+x')} \int_0^{\infty} \left(\frac{z}{b + t'_0 + t}\right)^{a+x'} e^{-z} dz \\
 &= \frac{(b + t'_0)^{a+x'}}{(b + t'_0 + t)^{a+x'+1} \Gamma(a+x')} \underbrace{\int_0^{\infty} z^{a+x'} e^{-z} dz}_{\Gamma(a+x'+1) = \Gamma(a+x') \cdot (a+x')} \\
 &= \frac{(b + t'_0)^{a+x'} (a+x')}{(b + t'_0 + t)^{a+x'+1}}
 \end{aligned}$$

Daher folgt nach der Beobachtung von  $x$  Fehlern im Zeitraum  $t_0$  für anzunehmende Zuverlässigkeit des zu untersuchenden Programms, d.h. der Wahrscheinlichkeit, dass zum Zeitpunkt  $t$  noch kein Fehler aufgetreten ist ( $T > t$ )

$$\begin{aligned}
 R(t|x, t_0) &= P(T > t|x, t_0) = \int_t^{\infty} p_{T|x, t_0}(t', x, t_0) dt' = \int_t^{\infty} \frac{(b + t_0)^{a+x} (a+x)}{(b + t_0 + t')^{a+x+1}} dt' \\
 &= \left[ \frac{(b + t_0)^{a+x} (a+x)}{(a+x)(b + t_0 + t')^{a+x}} \right]_t^{\infty} = \frac{(b + t_0)^{a+x}}{(b + t_0 + t)^{a+x}}
 \end{aligned}$$

Im Falle von Ultra-hoch zuverlässigen Systemen ist in der Regel kein Fehler in realisierbaren Testzeiträumen zu beobachten, d.h mit  $x = 0$  vereinfacht sich die Formel zu

$$R(t|x = 0, t_0; a, b) = \left( \frac{b + t_0}{b + t_0 + t} \right)^a \quad (16.3)$$

Offen gelassen wurde bisher, wie man die Hyper-Parameter der apriori-Verteilung, in diesem Fall die beiden in Glg. (16.2), zu wählen hat. Sie drücken unweigerlich bereits ein so notwendiges Vorwissen über die Verteilung der Fehlerrate  $\lambda$  und damit auch von  $T$ . aus. Die Forderung nach einer Unvoreingenommenheit ist nicht trivial umzusetzen wie man am vorliegenden Problem sehen kann. So ist  $\lambda$  Gamma-verteilt, d.h. aposteriori ist sie  $\text{Gam}_\lambda(r; a + x, b + t_0)$ . Zur Minimierung des Einflusses der Parameter  $a$  und  $b$  aus der apriori Verteilung, müssen diese zu den beobachteten Größen  $x$  und  $t_0$  vergleichsmäßig klein sein, d.h. man erhält im Limes  $a, b \rightarrow 0$  eine vorannahmenfreie Verteilung:

$$\begin{aligned} \lim_{a, b \rightarrow 0} \text{Gam}_{\lambda|x, t_0}(r; a + x', b + t'_0) &= \lim_{a, b \rightarrow 0} \frac{(b + t'_0)^{a+x'} r^{a+x'-1} e^{-(b+t'_0)r}}{\Gamma(a + x')} \\ &= \frac{(t'_0)^{x'} r^{x'-1} e^{-(t'_0)r}}{\Gamma(x')} \quad (x' \neq 0) \\ &= \text{Gam}_{\lambda|x, t_0}(r; x', t'_0) \quad (x' \neq 0) \end{aligned}$$

Die vorannahmsfreie Dichtefunktion nach obigem Limes ist allerdings (bedingt durch die  $\Gamma$ -Funktion) nicht definiert für oben angenommenen Wert für  $x = 0$  und damit auch die entsprechende Herleitung für die von  $T$  und daraus folgend die Formel der Zuverlässigkeit nicht mehr gültig.

### 16.3 Vorurteilsfreie $\lambda$ -Verteilung

Ein Weg, das Problem mit der Wahl von Hyper-Parametern zu umgehen, ist die Wahl einer (parameter-freien) Gleichverteilung der Form

$$p_\lambda^{\text{ign.}} \equiv 1$$

Diese Funktion liefert wohl keine wohldefinierte Wahrscheinlichkeitsdichte, da das Integral der Gesamtwahrscheinlichkeit nicht existiert und damit die Dichte nicht normierbar ist. Jedoch liefert sie eine wohldefinierte aposteriori Verteilung (für  $x = 0$ )

$$\begin{aligned}
 p_{\lambda|x,t_0}^{\text{ign.}}(r, 0, t'_0) &= p_{\lambda|x,t_0}^{\text{ign.}}(r, x, t'_0)|_{x=0} \stackrel{\text{Glg. (16.1)}}{=} \frac{p_{x,t_0|\lambda} \cdot p_{\lambda}^{\text{ign.}}}{\int_0^{\infty} p_{x,t_0|\lambda}(x', t'_0, r) \cdot p_{\lambda}^{\text{ign.}}(r) dr} \Big|_{x=0} \\
 &= \frac{\frac{(t'_0 r)^{x'}}{x'!} e^{-t'_0 r}}{\int_0^{\infty} \frac{(t'_0 r)^{x'}}{x'!} e^{-t'_0 r} dr} \Big|_{x=0} = \frac{e^{-t'_0 r}}{\int_0^{\infty} e^{-t'_0 r} dr} = t'_0 e^{-t'_0 r}
 \end{aligned}$$

und hieraus nun eine ebenfalls wohldefinierte Verteilung für  $T$

$$\begin{aligned}
 p_{T|x,t_0}^{\text{ign.}}(t, 0, t'_0) &= \int_0^{\infty} p_{T|\lambda}(t, r) p_{\lambda|x,t_0}(r, 0, t'_0) dr \\
 &= \int_0^{\infty} r e^{-\lambda t} t'_0 e^{-t'_0 r} dr = t'_0 \int_0^{\infty} r e^{-\lambda(t+t'_0)} dr = \frac{t'_0}{(t+t'_0)^2}
 \end{aligned}$$

Die letztendlich gesuchte Zuverlässigkeitsfunktion lautet nun

$$\begin{aligned}
 R^{\text{ign.}}(t|0, t_0) &= P(T > t|0, t_0) = \int_t^{\infty} p_{T|x,t_0}^{\text{ign.}}(t', 0, t_0) dT = \int_t^{\infty} \frac{t'_0}{(t' + t'_0)^2} dt' \\
 &= \frac{t'_0}{(t + t'_0)}
 \end{aligned}$$

## 16.4 Fazit

Betrachte man die Zuverlässigkeitsfunktion des zweiten Ansatzes, dargestellt im Graphen 16.1, so erkennt man, dass man nach einem fehlerfreien Testzeitraum  $t_0$  für einen gleichlangen Folgezeitraum nur in der Hälfte der Fälle keinen Fehler erwartet. Somit muss man hier einsehen, dass diese Methode zur Verifizierung hoch-zuverlässiger Systeme unbrauchbar ist, da der Testzeitraum unrealisierbar gross sein müsste.

Als Abschluss soll nun noch das Ergebnis des ersten Ansatzes nochmal aufgegriffen werden, um zu überlegen, welche apriori Annahme in Form der Hyper-Parameter man

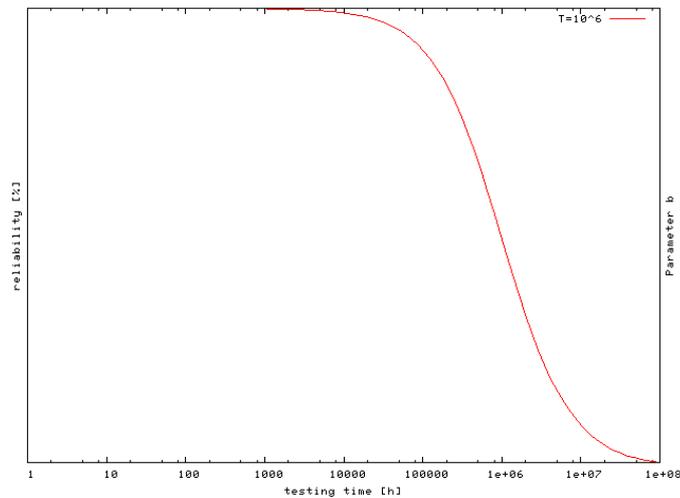


Abbildung 16.1: Zuverlässigkeitsfunktion  $R^{\text{ign}}$ .

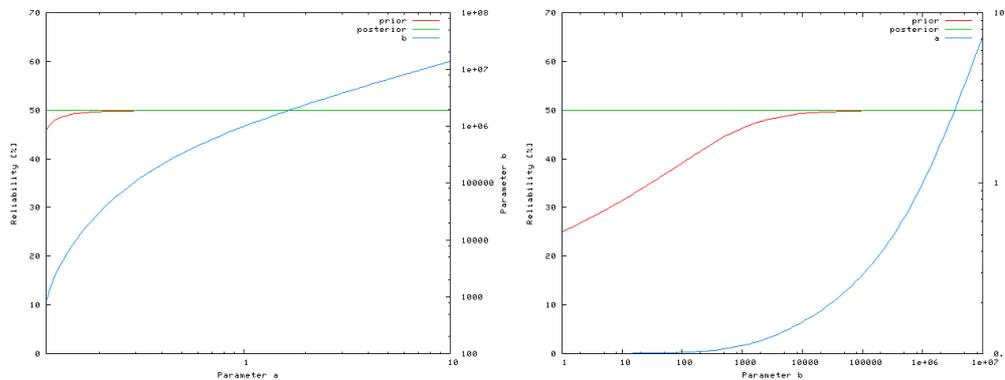
eigentlich treffen muss, um nach einer noch realisierbaren Testzeit, hier  $t_0 = 10^3 h$ , um eine um Größenordnungen größeren, mittlere Laufzeit bis zum Auftreten eines Fehlers des Programms, z.B.  $T = 10^6$ , zu verifizieren, d.h. eine Zuverlässigkeit von 50% für diesen Zeitraum zu erreichen. Setzt man diese Vorgaben in die Zuverlässigkeitfunktion aus Glg. (16.2) ein, erhält man als Beziehung zwischen den Parametern

$$R(t = 10^6 | x = 0, t_0 = 10^3; a, b) \Big|_{t=10^6, x=0, t_0=10^3} = \left( \frac{b + t_0}{b + t_0 + t} \right)^a \Big|_{t=10^6, t_0=10^3} \stackrel{!}{=} \frac{1}{2}$$

$$\Leftrightarrow a = \frac{\ln 2}{\ln(b + t_0 + t) - \ln(b + t_0)} \Big|_{t=10^6, t_0=10^3} = \frac{\ln 2}{\ln(b + 10^3 + 10^6) - \ln(b + 10^3)}$$

$$\Leftrightarrow b = \frac{1}{\sqrt[2]{2} - 1} t - t_0 \Big|_{t=10^6, t_0=10^3} = \frac{1}{\sqrt[2]{2} - 1} 10^6 - 10^3$$

Davon ausgehend, dass wir zumindest annehmen müssen, wann wir mal einen Fehler beobachtet haben, erhalten wir für die Zuverlässigkeit dieser Annahme



**Abbildung 16.2:** Zuverlässigkeits-  $R$  und Parameterfunktion  $a(\cdot)$  (links) bzw.  $b(\cdot)$  (rechts) ( $T = 10^6, t_0 = 10^3$ )

$$\begin{aligned}
 R^{\text{prior}}(t; a = 1, b) &= \left( \frac{b}{b+t} \right)^a \Big|_{a=1, t} \\
 \Rightarrow R^{\text{prior}}(t; a = 1) &= \left( \frac{\frac{1}{\sqrt[2]{2}-1} 10^6 - 10^3}{\frac{1}{\sqrt[2]{2}-1} 10^6 - 10^3 + t} \right)^a \Big|_{a=1} \\
 &= \frac{10^6 - 10^3}{10^6 - 10^3 + t}
 \end{aligned}$$

Abb. 16.2 zeigt die zugehörigen Diagramme

Daraus folgt, dass wir bereits eine Zuverlässigkeit in der Größenordnung annehmen müssen, die wir dann auch verifizieren wollen. Damit ist dieser Ansatz ebenso unbrauchbar.

