RESTART: Geometric Sums in a Computer Reliability Problem

Søren Asmussen
Aarhus University, Denmark
`http://home.imf.au.dk/asmus/`

Heriot-Watt University, September 7, 2006

Joint work with
Lester Lipsky
Robert Sheahan
Pierre Fiorini

LIGHT TAILS

1) exponential $\overline{F}(x) = e^{-\beta x}$

2) Gamma-like $\overline{F}(x) \approx cx^{\alpha-1}e^{-\beta x}$

3) Rayleigh $\overline{F}(x) = e^{-\beta x^2}$

4) LT Weibull $\overline{F}(x) = e^{-\beta x^\gamma}$, $\gamma \geq 1$

General: $\int_0^\infty e^{\epsilon x} f(x)dx < \infty$, some $\epsilon > 0$

Failure rate $\dfrac{f(x)}{\overline{F}(x)} \longrightarrow \infty$ or $c > 0$

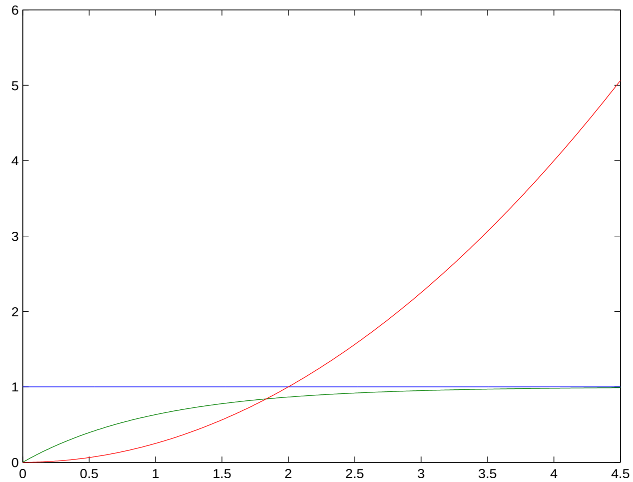


Figure 1:

HEAVY TAILS

General: $\int_0^\infty e^{\epsilon x} f(x) dt = \infty$, all $\epsilon > 0$

Failure rate $\dfrac{f(x)}{\overline{F}(x)} \to 0$

1) power tails $\overline{F}(x) = \dfrac{c}{x^\alpha}$

2) lognormal $\overline{F}(x) \approx ce^{-\alpha \log^2 x}$

3) HT Weibull $\overline{F}(x) = e^{-\beta x^\gamma}$, $\gamma < 1$

0) $\overline{F}(x) \approx \dfrac{c}{\log^\alpha x}$

2^δ) $\overline{F}(x) \approx e^{-\alpha \log^\delta x}$
$\delta = 1$: power tails

GEOMETRIC SUMS

Books by
Kalashnikov, Willmot & Liu
$N \in \{0, 1, 2, \ldots\}$
$P(N = n) = (1 - \rho)\rho^n$
$V_1, V_2, \ldots$ i.i.d. $\sim F$
$S_N \; = \; V_1 + \cdots + V_N$
Classical applications:
. . Pollazeck-Khinchine for M/G/1
. . Ruin probabilities

This lecture:
. . Tail asymptotics
. . Bounds
. . Renewal theory
. . Change of measure

Easy: asymptotics for heavy tails

Assume F subexponential

$$P(V_1 + \cdots + V_n > x) \sim n\,\overline{F}(x)$$

$$= P(S_N > x | N = n)$$

$$P(S_N > x | N) \sim N\,\overline{F}(x)$$

Guess: $\boxed{P(S_N > x) \sim EN\,\overline{F}(x)}$

Kesten's lemma:

For some $z > 1$ s.t. $\rho z < 1$ and some c

$$P(V_1 + \cdots + V_n > x) \le c z^n \overline{F}(x)$$

$$\frac{P(S_N > x)}{\overline{F}(x)}$$

$$= \sum_{n=0}^{\infty} (1-\rho)\rho^n \frac{P(V_1 + \cdots + V_n > x)}{\overline{F}(x)}$$

$\rho^n P(\cdot)/\overline{F}(x) \to \rho^n \cdot n$;

dominated by $c(\rho z)^n$

Sum has limit

$$\sum_{n=0}^{\infty} (1-\rho)\rho^n \cdot n = EN$$

Shneer 2004, Hartinger & Kortschak 2006: weaker conditions on N

Light-tailed asymptotics

$S_N = V_1 + \cdots + V_N$
$V_n \sim F$, N geom(ρ)
$Z(x) = P(S_N > x)$ solves

$$Z(x) = \rho\overline{F}(x) + \int_0^x Z(x-y)\,\rho F(dy) \qquad (1)$$

Defective renewal equation
Solution exponentially decaying
Lundberg inequality

$\rho F(dy)$ has mass $\rho < 1$
Choose γ as solution of

$$\int_0^\infty e^{\gamma x}\,\rho F(dx) = 1$$

Let $G(dx) = e^{\gamma x}\,\rho F(dx)$
. . $Z^*(x) = e^{\gamma x} Z(x)$
. . $z^*(x) = e^{\gamma x} \rho\overline{F}(x)$
Multiply (1) by $e^{\gamma x}$

$$
\begin{aligned}
& e^{\gamma x} Z(x-y)\,\rho F(dy) \\
&= e^{\gamma(x-y)} Z(x-y)\, e^{\gamma y} \rho F(dy) \\
&= Z^*(x-y) G(dy) \\
Z^*(x) &= z^*(x) + \int_0^x Z^*(x-y)\, G(dy)
\end{aligned}
$$

Know $Z^*(x) \to \int_0^\infty z^*(x)\, dx / \mu_G$

$$\boxed{Z(x) \sim C e^{-\gamma x}}$$

Cramér-Lundberg approximation

Lundberg's Inequality

$S_N = V_1 + \cdots + V_N$
$V_n \sim F$, N geom(ρ)

$$\boxed{Z(x) \; = \; P(S_N > x) \; \le \; e^{-\gamma x}}$$

$$Z(x) \; = \; \rho \overline{F}(x) + \int_0^x Z(x-y)\, \rho F(dy)$$

$Z_n(x) \; = \; P(S_N > x, N \le n)$
Induction: $Z_n(x) \; \le \; e^{-\gamma x}$

$$Z_{n+1}(x) = \rho \overline{F}(x) \; + \; \int_0^x Z_n(x-y)\, \rho F(dy)$$

$e^{\gamma x} \rho \overline{F}(x) \; \le \; \int_x^\infty e^{\gamma y} \rho F(dy)$
$e^{\gamma x} \int_0^x Z_n(x-y)\, \rho F(dy) \; \le \; \int_0^x e^{\gamma y} \rho F(dy)$

Heavy-tailed bounds: Kalashnikov,
Kalashnikov & Tsitsiashvili

Lower Lundberg bound
Use change of measure
$G(dx) = e^{\gamma x}\rho F(dx)$
LR similar to importance sampling

$$
\begin{aligned}
Ef(V_1) &= \int f(y)\,F(dy) \\
&= \rho^{-1}\int f(y)e^{-\gamma y}\,G(dy) \\
&= \rho^{-1}E_G[f(V_1)e^{-\gamma V_1}] \\
\tau &= \inf\{n :\ S_n > x\} \\
P(S_N > x) &= P(\tau \le N) \\
&= E_G[\rho^{-\tau}e^{-\gamma S_\tau}; \tau \le N] \\
&= E_G[\rho^{-\tau}e^{-\gamma S_\tau}\rho^{\tau}] \\
&= E_G e^{-\gamma S_\tau}
\end{aligned}
$$

Assume $V \le a$
$x\ <\ S_\tau\ <\ x + a$

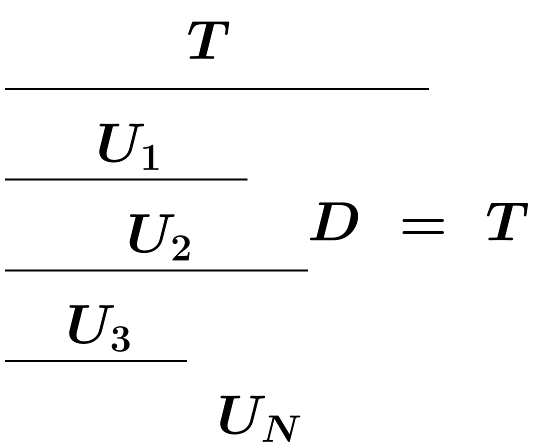
$$\boxed{P(S_N > x)\ \ge\ ce^{-\gamma x}}$$

$c = e^{-\gamma a}$

RESTART

T execution time of job $\sim\ F$
U failure time $\sim\ G$
D task duration $\sim\ H$

RESUME checkpoints
REPLACE start new job at failure
RESTART start same job



$$D = T + U_1 + \cdots + U_{N-1}$$

Target: tail $\overline{H}(x) = P(D > x)$ of D

$$\boxed{\boldsymbol{T}\ \textbf{bounded} \Rightarrow\ \overline{\boldsymbol{H}}(\boldsymbol{x}) \approx \boldsymbol{e}^{-\gamma \boldsymbol{x}}}$$

$$\boxed{\boldsymbol{T}\ \textbf{unbd} \Rightarrow\ \boldsymbol{H}\ \ \textbf{heavy-tailed}}$$

COMPARISON RESULTS

$$F = G \;\Rightarrow\; \overline{H}(x) \approx \frac{1}{x}$$

$$F = G \;\Rightarrow\; ED = \infty$$

$$\overline{F} << \overline{G} \;\Rightarrow\; ED < \infty$$

$$\overline{F} >> \overline{G} \;\Rightarrow\; ED = \infty$$

$$\overline{F} << \overline{G}^2 \;\Rightarrow\; \mathrm{Var}\, D < \infty$$

$$\overline{F} >> \overline{G}^2 \;\Rightarrow\; \mathrm{Var}\, D = \infty$$

$$\overline{F}_1 << \overline{F}_2 \;\Rightarrow\; \overline{H}_{F_1,G} << \overline{H}_{F_2,G}$$

$$\overline{G}_1 << \overline{G}_2 \;\Rightarrow\; \overline{H}_{F,G_1} >> \overline{H}_{F,G_2}$$

TAIL OF H

Logarithmic asymptotics (large deviations)

$$h_1 \approx_{\log} h_2 : \quad \frac{\log h_1(x)}{\log h_2(x)} \to 1$$

$h(x) = 4x^3 e^{-\gamma x}$
$h(x) \approx_{\log} e^{-\gamma x}$, $h(x) \approx_{\log} 8x^7 e^{-\gamma x}$,...
Identifies γ but not x^3 or 4

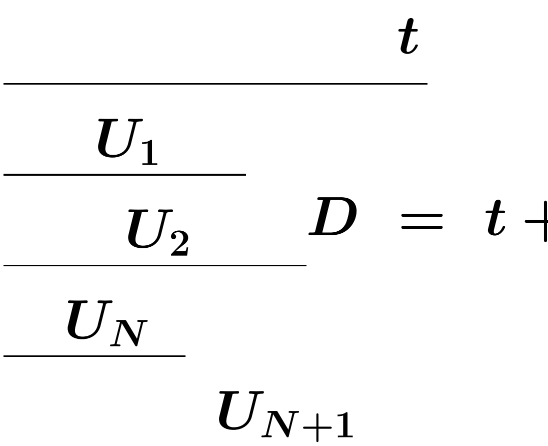
4 examples of each of F, G:
LT Weibull
exponential
HT Weibull
power

. $\overline{F}(t)$ / $\overline{G}(u)$	e^{-t^2}	e^{-t}	$e^{-t^{1/2}}$	$\frac{1}{t^\alpha}$
e^{-u^2}	$\frac{1}{x}$	$e^{-\log^{1/2} x}$	$e^{-\log^{1/4} x}$	$\frac{1}{\log^{\alpha/2} x}$
e^{-u}	$e^{-\log^2 x}$	$\frac{1}{x}$	$e^{-\log^{1/2} x}$	$\frac{1}{\log^{\alpha} x}$
$e^{-u^{1/2}}$	$e^{-\log^4 x}$	$e^{-\log^2 x}$	$\frac{1}{x}$	$\frac{1}{\log^{2\alpha} x}$
$\frac{1}{u^\alpha}$	$e^{-x^{\frac{2}{2+\alpha}}}$	$e^{-x^{\frac{1}{1+\alpha}}}$	$e^{-x^{\frac{1/2}{1/2+\alpha}}}$	$\frac{1}{x}$

Constants omitted $e^{-c\log^{1/2} x}$; $\dfrac{1}{x} = e^{-\log x}$
In some cases even log log asymptotics
Omitted case: T bounded (later)
Later: cases of exact asymptotics.

$T \equiv t$

Later use: conditioning, mixing



$$D = t + U_1 + \cdots + U_N$$

Geometric sum
Succes probability $\overline{G}(t)$
Summands: $U \mid U \leq t$; light tail
Cramér-Lundberg theory

$S_N = V_1 + \cdots + V_N$

$V_n \sim K$, N geom(ρ)

$Z(x) = P(S > x)$ solves

$$Z(x) = \rho\overline{K}(x) + \int_0^x Z(x-y)\,\rho K(dy)$$

Defective renewal equation

Solution exponentially decaying

Lundberg inequality

RESTART with $T \equiv t$:

$\gamma(t)$ solution of $1 = \int_0^t e^{\gamma(t)u} g(u)du$

$$\boxed{P(S > x) \sim C(t)e^{-\gamma(t)x}}$$

$$\boxed{e^{-\gamma(t)t}e^{-\gamma(t)x} \leq P(S > x) \leq e^{-\gamma(t)x}}$$

$$\boxed{\overline{H}(x) \approx e^{\gamma(t)t}C(t)e^{-\gamma(t)x}}$$

Bounded support:

$$f(t) \sim A(t_0 - t)^\alpha, \;\; t \uparrow t_0$$

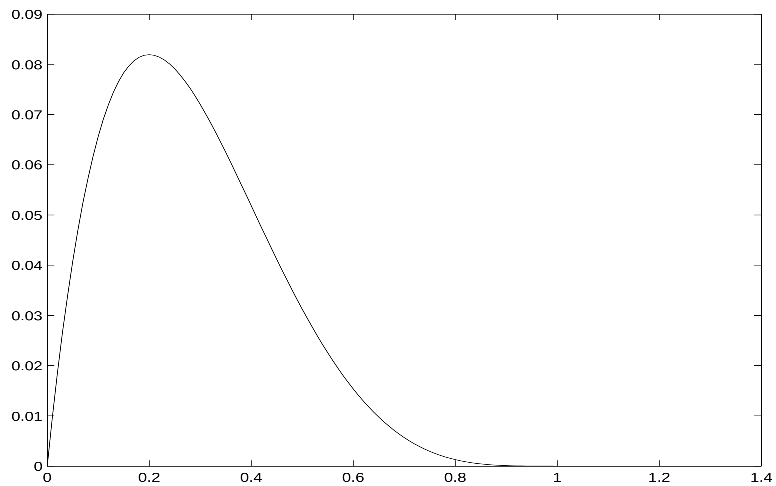


Figure 2:

$$\boxed{\overline{H}(x) \;\sim\; C_1(t_0)\,\frac{e^{-\gamma(t_0)x}}{x^\alpha}}$$

$C_1(t_0)$ involves

$$\gamma(t_0), C(t_0), \Gamma(\alpha+1), g(t_0)^{\alpha+1}$$

$\gamma(t)$ solution of $1 = \int_0^t e^{\gamma(t)u} g(u) du$

$$\boxed{\gamma(t) \sim \mu \overline{G}(t)} \quad \mu = 1/EU$$

Crucial lemma

for random, unbd. T:

$$\overline{H}(x) \sim \int_{t_0}^{\infty} \exp\{-\mu \overline{G}(t) x\} f(t)\, dt$$

Purely analytical problem;
non-trivial

F **Gamma-like:** $\overline{F}(x) \sim A x^{\eta} e^{-\delta x}$

$g(t) = \beta e^{-\beta t}$

$$\boxed{\overline{H}(x) \sim \frac{A \Gamma(\delta/\beta)}{\beta^{\delta/\beta - 1 - \eta}} \frac{\log^{\eta} x}{x^{\delta/\beta}}}$$

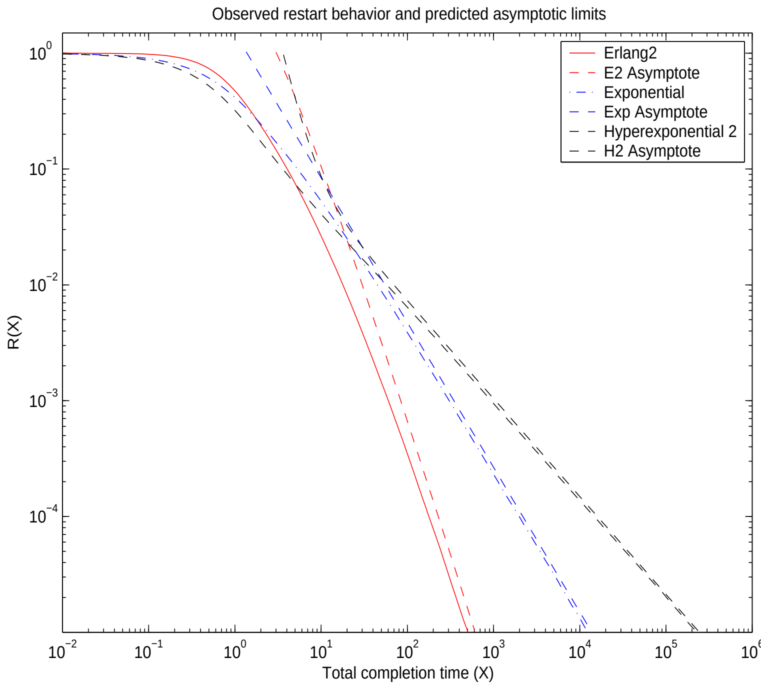


Figure 3:

Naive simulation, $R = 10^8$

Ongoing:

more sophisticated algorithm

Assumes exponential failures

Replaces

generation of $U_1 + \cdots + U_N$

by generation of 4 r.v.'s

Uniform order statistics

Importance sampling on T