Extreme Value Distributions

The Weibull Distribution

© D. L. Bricker
Dept of Mechanical & Industrial Engineering
The University of Iowa

Asymptotic Distributions

From the *Central Limit Theorem* we know that, for "large" n,

$Y = \sum_{i=1}^{n} X_i$ has approximately a *Normal* distribution

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{y-\mu}{\sigma} \right)^2 \right\}$$

$Y = \prod_{i=1}^{n} X_i$ has approximately a *Lognormal* distribution

$$f_Y(y) = \frac{1}{y\sigma_X\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{\ln\left(y/\mu \right)}{\sigma} \right)^2 \right\}$$

Asymptotic Distributions

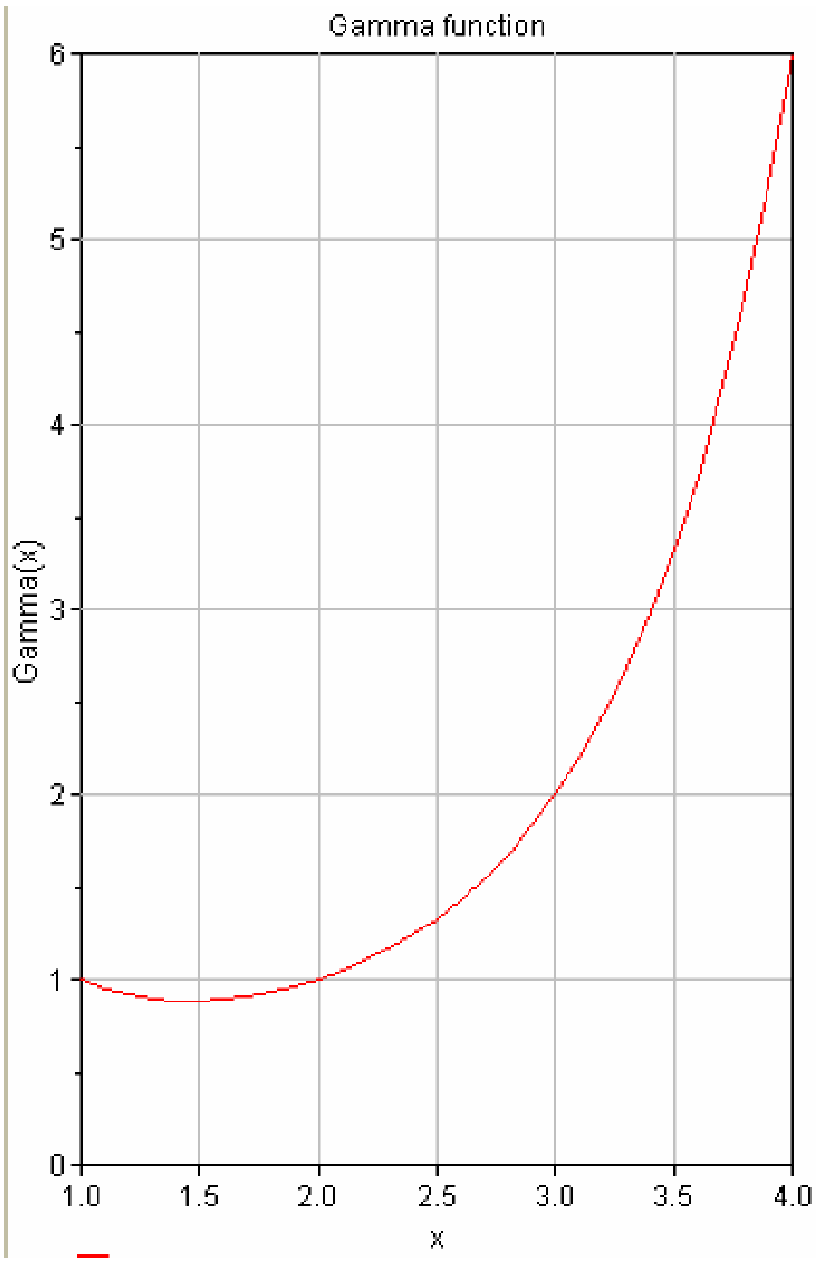
When the *minimum* value of X is 0,

i.e., X is nonnegative,

the limiting distribution for $T = \min\{X_i\}$ as $n \to \infty$

is the ***Weibull*** distribution:

CDF	$F_T(t) = 1 - e^{-\left(t/u\right)^k}$
Mean value	$\mu_T = u\Gamma\left(1 + \frac{1}{k}\right)$
Standard deviation	$\sigma_T = u\sqrt{\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)}$



The "Gamma" function Γ is a generalization of the *factorial* function, defined for *all x≥0*, (not necessarily integer) and satisfies

$$\Gamma(1+x) = x!$$

when x is a nonnegative integer.

Table of values of $\Gamma\left(1+\frac{1}{k}\right)$ for $k = 0.1$ *through* 9.9

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	∞	3.63E+06	120.000	9.2610	3.3230	2.0000	1.5050	1.2660	1.1330	1.0520
1	1.0000	0.9649	0.9407	0.9236	0.9114	0.9027	0.8966	0.8922	0.8893	0.8874
2	0.8862	0.8857	0.8856	0.8859	0.8865	0.8873	0.8882	0.8893	0.8905	0.8917
3	0.8930	0.8943	0.8957	0.8970	0.8984	0.8997	0.9011	0.9025	0.9038	0.9051
4	0.9064	0.9077	0.9089	0.9102	0.9114	0.9126	0.9137	0.9149	0.9160	0.9171
5	0.9182	0.9192	0.9202	0.9213	0.9222	0.9232	0.9241	0.9251	0.9260	0.9269
6	0.9277	0.9286	0.9294	0.9302	0.9310	0.9318	0.9325	0.9333	0.9340	0.9347
7	0.9354	0.9361	0.9368	0.9375	0.9381	0.9387	0.9394	0.9400	0.9406	0.9412
8	0.9417	0.9423	0.9429	0.9434	0.9439	0.9445	0.9450	0.9455	0.9460	0.9465
9	0.9470	0.9474	0.9479	0.9484	0.9488	0.9493	0.9497	0.9501	0.9505	0.9509

For example, $\Gamma\left(1+\frac{1}{2.5}\right)=\Gamma(1.4)=0.8873$

Table of values of $\Gamma\left(1+\frac{2}{k}\right)$ for $k = 0.1$ *through* 9.9

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0	2.43E+18	3628800	2594	120	24.0000	9.2610	5.0290	3.3230	2.4790
1	2.0000	1.7020	1.5050	1.3660	1.2660	1.1910	1.1330	1.0880	1.0520	1.0230
2	1.0000	0.9808	0.9649	0.9517	0.9407	0.9314	0.9236	0.9170	0.9114	0.9067
3	0.9027	0.8994	0.8966	0.8942	0.8922	0.8906	0.8893	0.8882	0.8874	0.8867
4	0.8862	0.8859	0.8857	0.8856	0.8856	0.8857	0.8859	0.8862	0.8865	0.8868
5	0.8873	0.8877	0.8882	0.8887	0.8893	0.8899	0.8905	0.8911	0.8917	0.8923
6	0.8930	0.8936	0.8943	0.8950	0.8957	0.8963	0.8970	0.8977	0.8984	0.8991
7	0.8997	0.9004	0.9011	0.9018	0.9025	0.9031	0.9038	0.9044	0.9051	0.9058
8	0.9064	0.9070	0.9077	0.9083	0.9089	0.9096	0.9102	0.9108	0.9114	0.9120
9	0.9126	0.9132	0.9137	0.9143	0.9149	0.9154	0.9160	0.9166	0.9171	0.9176

For example, $\Gamma\left(1+\frac{2}{2.5}\right)=\Gamma(1.8)=0.9314$

Computing Weibull Parameters

Given μ_Y and σ_Y, we wish to solve for the parameters u & k:

$$\begin{cases} \mu_T = u\Gamma\left(1+\frac{1}{k}\right) \\ \sigma_T = u\sqrt{\Gamma\left(1+\frac{2}{k}\right)-\Gamma^2\left(1+\frac{1}{k}\right)} \end{cases}$$

This is a system of two nonlinear equations in two unknowns *u* & *k*, which might be solved by (for example) the *Newton-Raphson* method.

Solve for u in terms of the mean μ_T:

$$\mu_T = u\Gamma\left(1+\frac{1}{k}\right) \Rightarrow u = \frac{\mu_T}{\Gamma\left(1+\frac{1}{k}\right)}$$

Use this to eliminate *u* from the second equation:

$$\sigma_T = u\sqrt{\Gamma\left(1+\frac{2}{k}\right) - \Gamma^2\left(1+\frac{1}{k}\right)}$$

$$\Rightarrow \quad \sigma_T = \frac{\mu_T}{\Gamma\left(1+\frac{1}{k}\right)}\sqrt{\Gamma\left(1+\frac{2}{k}\right) - \Gamma^2\left(1+\frac{1}{k}\right)}$$

This gives us a *single* (nonlinear) equation in a *single* variable k.

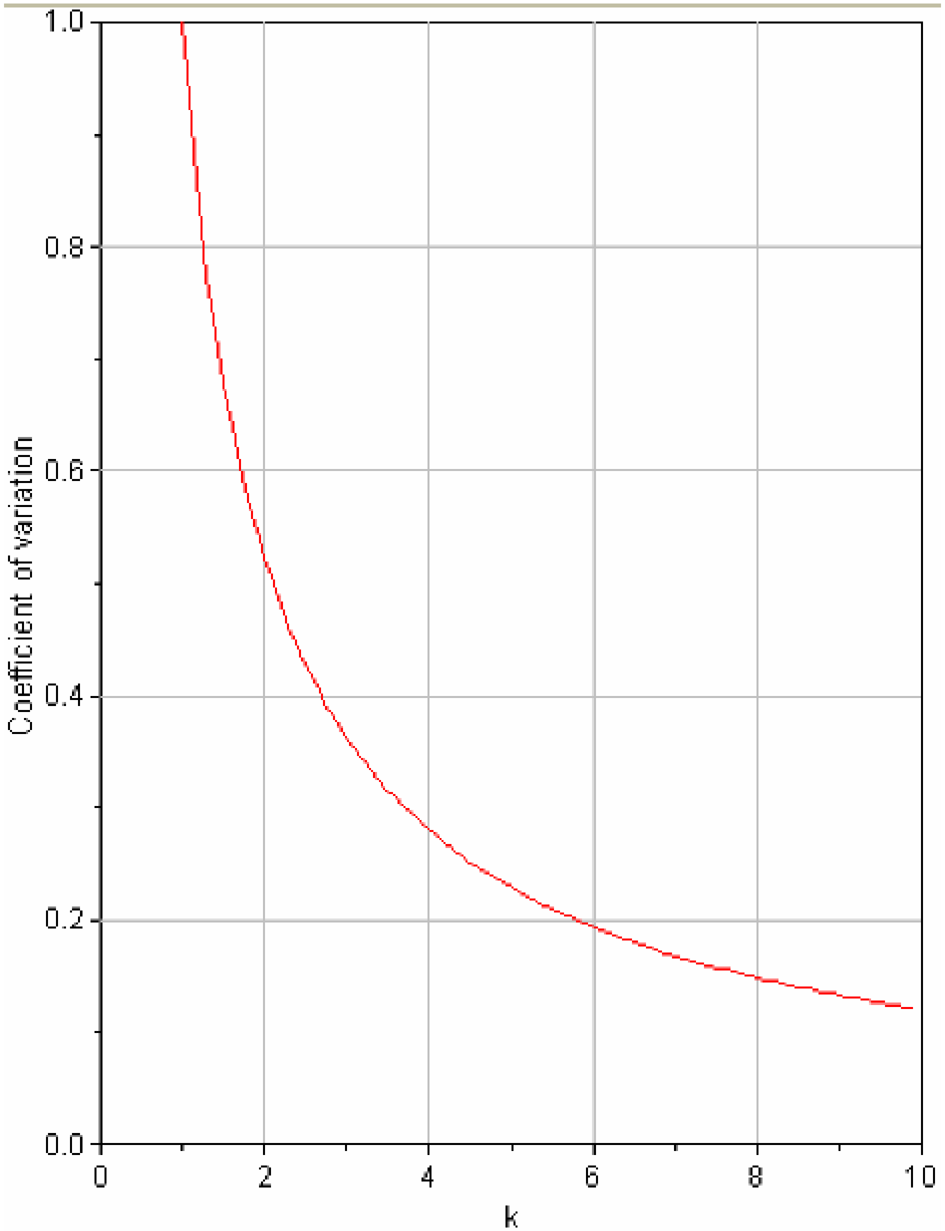
$$\sigma_T = \frac{\mu_T}{\Gamma\left(1+\frac{1}{k}\right)}\sqrt{\Gamma\left(1+\frac{2}{k}\right)-\Gamma^2\left(1+\frac{1}{k}\right)}$$

$$\Rightarrow \qquad \frac{\sigma_T}{\mu_T} = \frac{1}{\Gamma\left(1+\frac{1}{k}\right)}\sqrt{\Gamma\left(1+\frac{2}{k}\right)-\Gamma^2\left(1+\frac{1}{k}\right)}$$

$$\Rightarrow \qquad \frac{\sigma_T}{\mu_T} = \sqrt{\frac{\Gamma\left(1+\frac{2}{k}\right)}{\Gamma^2\left(1+\frac{1}{k}\right)}-1}$$

Coefficient of Variation

Thus, the coefficient of variation of the Weibull distribution is determined by k alone.



Coefficient of Variation

as a function of parameter k

Coefficient of Variation, as a function of parameter k

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	------	429.8	15.843	5.4077	3.1409	2.2361	1.7581	1.4624	1.2605	1.1130
1	1.0000	0.9102	0.8369	0.7757	0.7238	0.6790	0.6399	0.6055	0.5749	0.5475
2	0.5227	0.5003	0.4798	0.4611	0.4438	0.4279	0.4131	0.3994	0.3866	0.3747
3	0.3635	0.3529	0.3430	0.3337	0.3248	0.3165	0.3085	0.3010	0.2939	0.2870
4	0.2805	0.2744	0.2684	0.2628	0.2573	0.2521	0.2471	0.2424	0.2378	0.2333
5	0.2291	0.2250	0.2210	0.2172	0.2135	0.2099	0.2065	0.2031	0.1999	0.1968
6	0.1938	0.1908	0.1880	0.1852	0.1826	0.1800	0.1774	0.1750	0.1726	0.1703
7	0.1680	0.1658	0.1637	0.1616	0.1596	0.1576	0.1557	0.1538	0.1519	0.1501
8	0.1484	0.1467	0.1450	0.1434	0.1418	0.1402	0.1387	0.1372	0.1357	0.1343
9	0.1329	0.1315	0.1302	0.1288	0.1275	0.1263	0.1250	0.1238	0.1226	0.1215

Several methods might be used to estimate the parameters u & k of the Weibull distribution

(a) method of moments, i.e., matching the mean and standard deviation

(b) linear regression (after transforming to linear form)

(c) maximum likelihood method

Method (a) requires that we have sufficient data to compute the mean & standard deviation.

Methods (b) & (c) require a sample of observations of T.

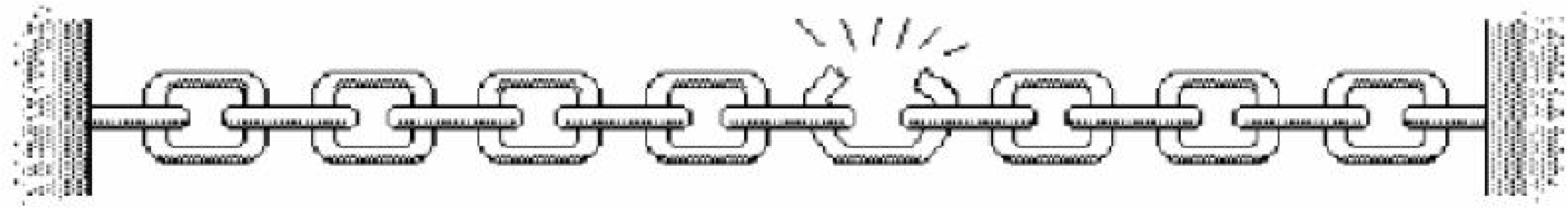
Method of Moments

Given the coefficient of variation (σ/μ), we can either

- approximate *k* through the use of the table (or graph), or
- use a numerical method, e.g., the Newton-Raphson or Secant method, to solve the *nonlinear* equation

$$\frac{\sigma_T}{\mu_T} = \sqrt{\frac{\Gamma\left(1+\frac{2}{k}\right)}{\Gamma^2\left(1+\frac{1}{k}\right)}-1}$$

(The Newton-Raphson method requires derivatives, while the secant method does not!)



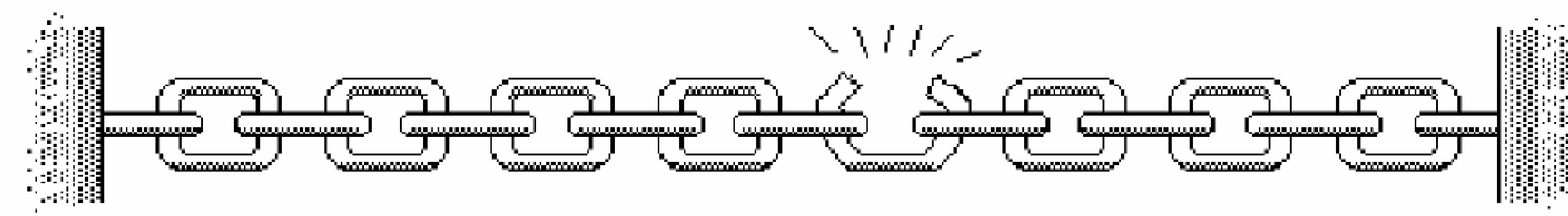
EXAMPLE

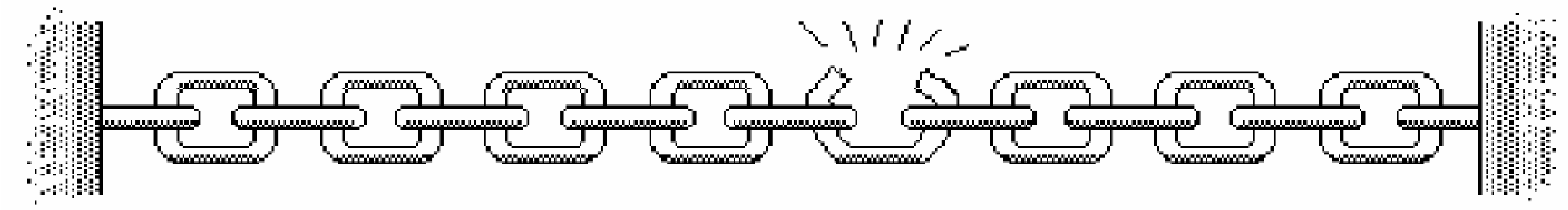
Let X_i = lifetime of link i of a chain ($X_i \geq 0$)

$$T = \min\{X_1, X_2, \ldots X_n\} = \text{lifetime of the chain}$$

For large n (long chains) the distribution of T should be approximately Weibull.

Suppose that we have estimates for the mean μ_T=150 hours and standard deviation σ_T=50 hours.





What is the probability that…

- the chain fails during its first 100 hours of use?
- the chain has not yet failed after 200 hours?

Since the lifetime of the chain is the minimum of the lifetimes (times until failure) of the individual links of the chain, and these lifetimes are each bounded below by zero, we will assume the Weibull distribution.

Computation of the Parameter k

The *coefficient of variation* of the lifetime T is $\sigma/\mu = 50/150 = 1/3$.

We will use the **Secant Method**

(which, unlike the Newton-Raphson method,

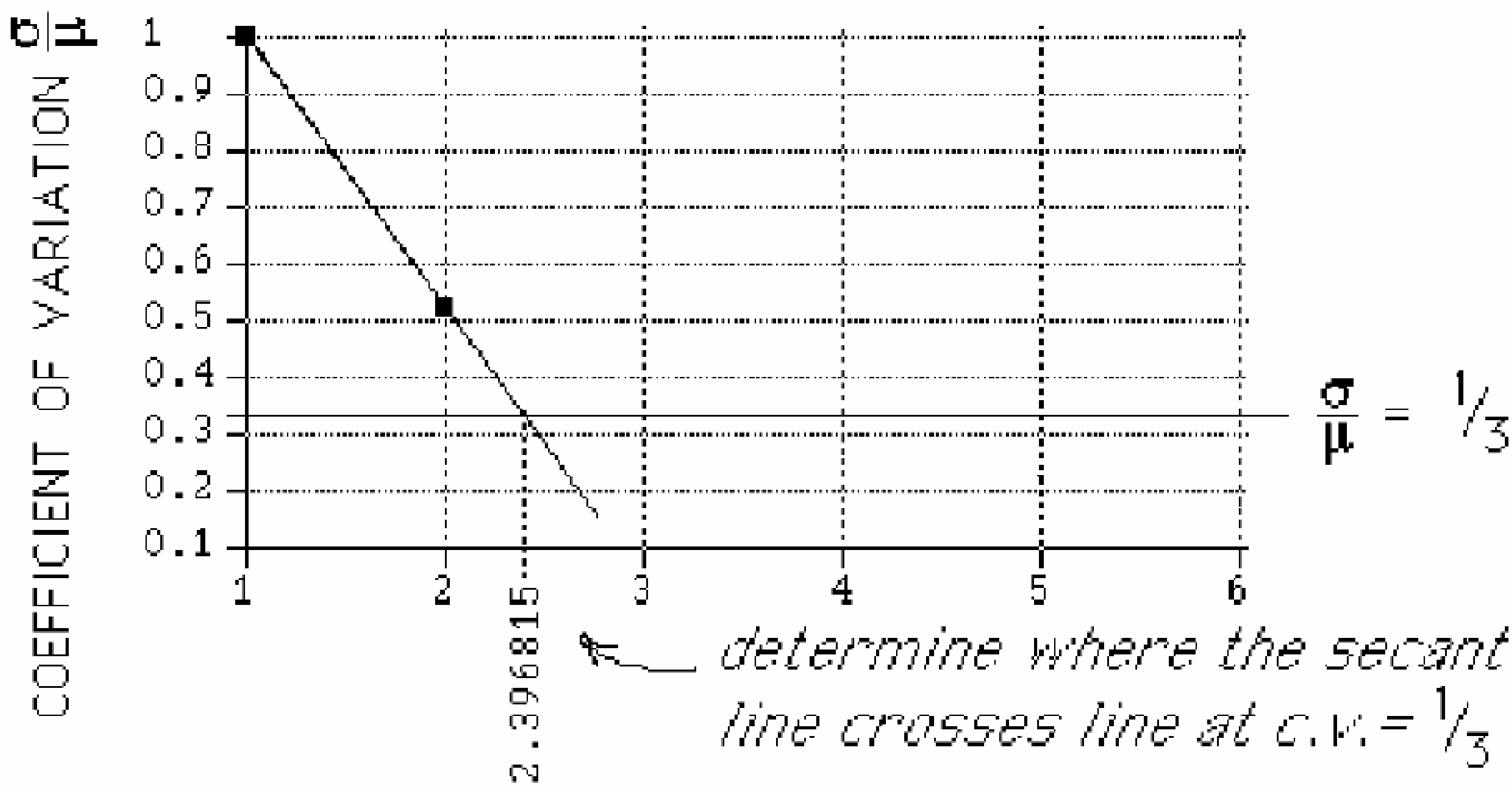
does *not* require derivatives)

to solve the nonlinear equation $$\frac{\sigma_T}{\mu_T} = \sqrt{\frac{\Gamma\left(1+\frac{2}{k}\right)}{\Gamma^2\left(1+\frac{1}{k}\right)} - 1}$$

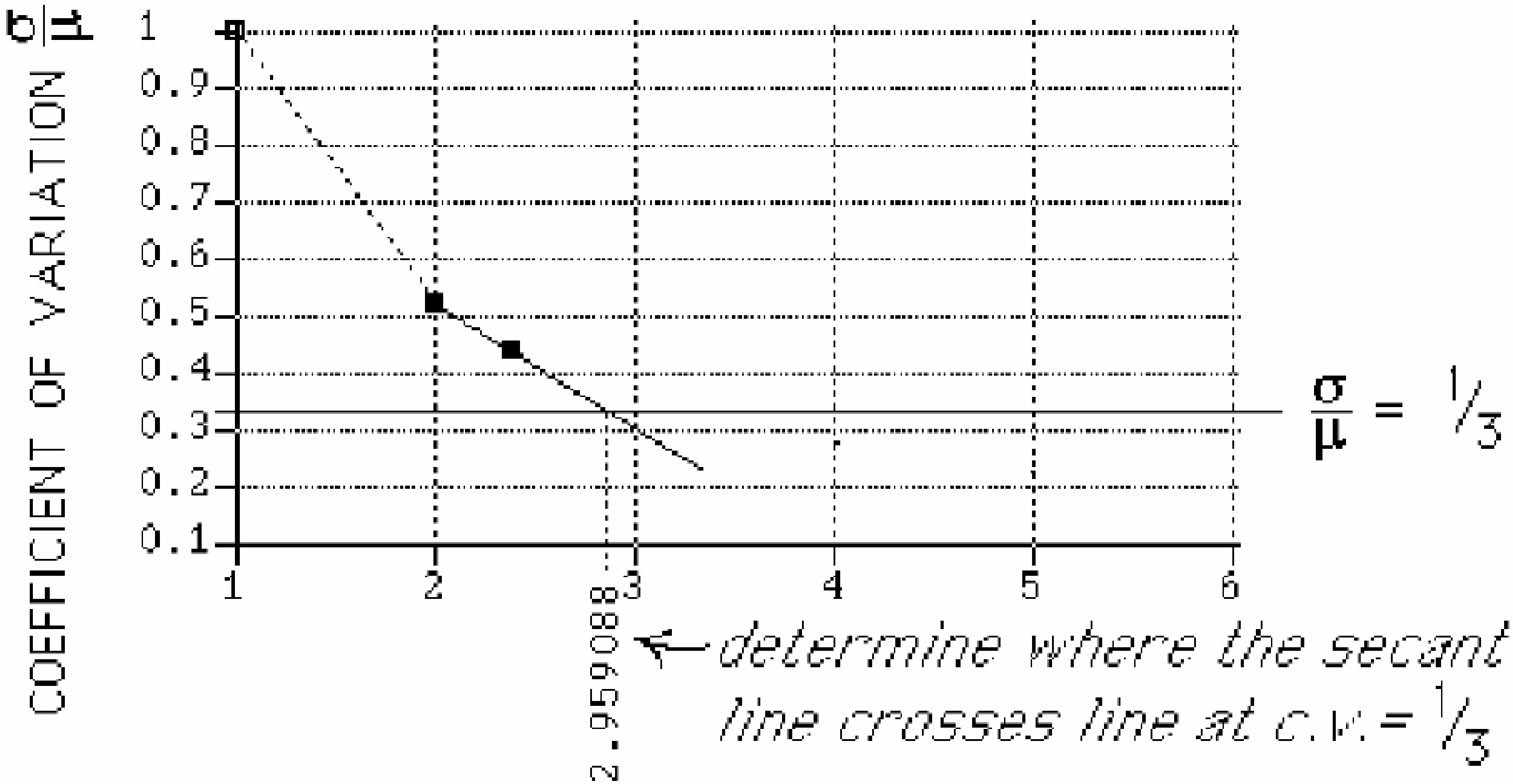
for k.

Secant method

Start with "guesses" k=1 & 2

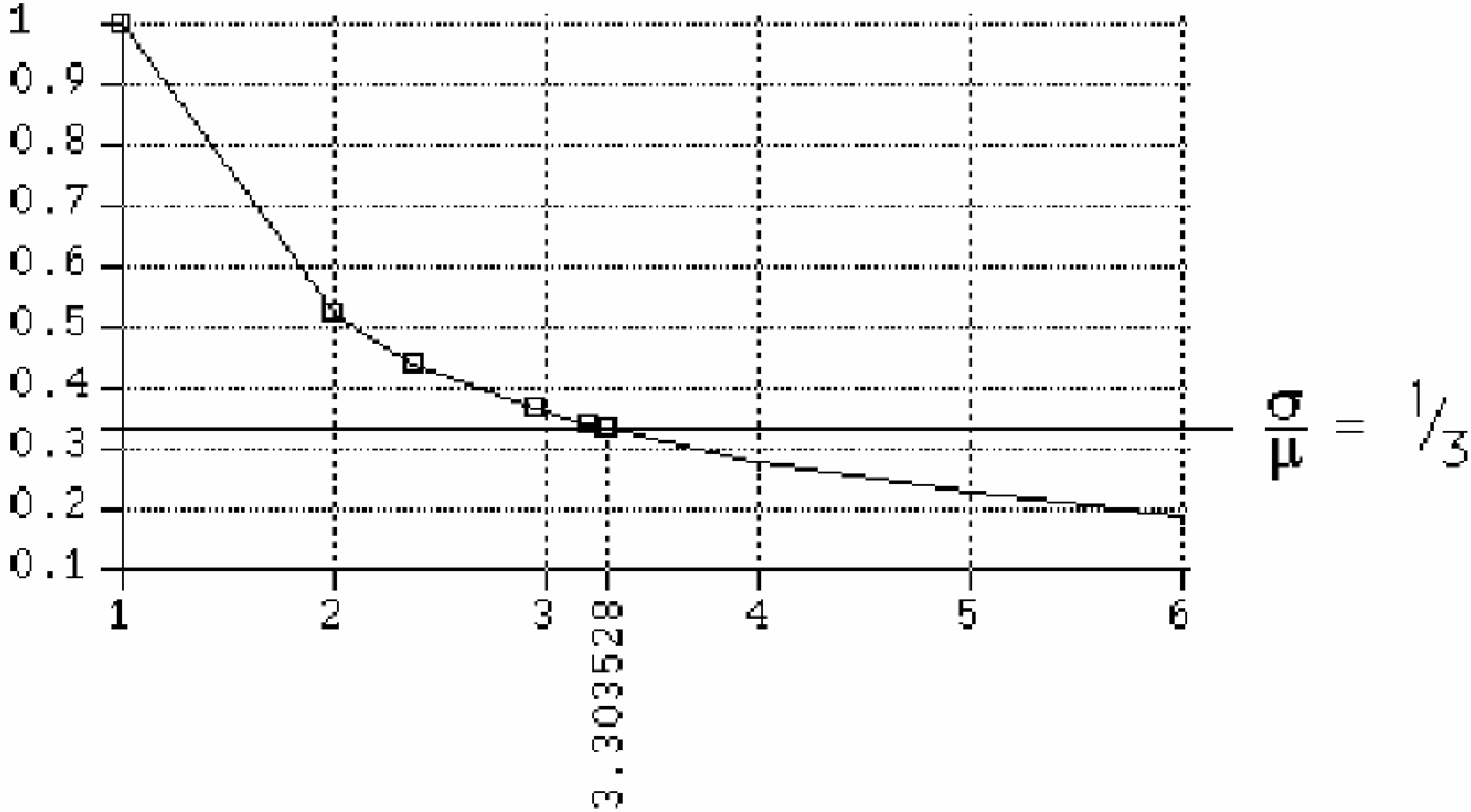


Using last 2 values of k, draw a new secant line



← determine where the secant line crosses line at c.v. = $\frac{1}{3}$

Continue, until the procedure converges to
k = 3.303528



Using the secant method, we get the following approximations to the value of k:

k	error
0.250000	7.973291
2.000000	0.189390
2.042579	0.179577
2.821777	0.050615
3.127600	0.016787
3.279356	0.002206
3.302318	0.000109
3.303517	0.000001
3.303525	0.000000

Given k=3.3035 and mean μ_T = 150, we can now solve for the parameter u:

$$u = \frac{\mu_T}{\Gamma\left(1+\frac{1}{k}\right)} = \frac{150}{\Gamma(1.3027)} = \frac{150}{0.8971} = 167.21$$

What is the probability that chain fails during its first 100 hours of use?

i.e., $P\{T \leq 100\} = F_T(100) = ?$

What is the probability that it has not yet failed after 200 hours of use?

i.e., $P\{T \geq 200\} = 1 - F_T(200) = ?$

$$F_T(t) = 1 - e^{-\left(t/u\right)^k} = 1 - e^{-\left(t/167.21\right)^{3.3025}}$$

Therefore,

$$P\{T \le 100\} = F_T(100) = 1 - \exp\left(100/167.21\right)^{3.3025}$$

$$= 1 - \exp(0.60529)^{3.3025} = 1 - \exp(0.19052) = 1 - 0.82653 = 0.17347$$

That is, the chain is about 17% likely to fail during its first 100 hours of use.

$$F_T(t) = 1 - e^{-\left(t/u\right)^k} = 1 - e^{-\left(t/167.21\right)^{3.3025}}$$

Therefore,

$$P\{T \geq 200\} = 1 - F_T(200) = e^{-\left(200/167.21\right)^{3.3025}} = 0.16423$$

That is, the chain is about equally likely (16%) to survive its first 200 hours of use.