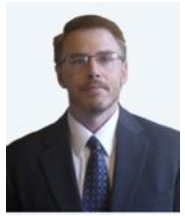
Probability Methods
Graphical Fitting Methods

Aircraft Airworthiness and Sustainment Conference

August 29, 2022






AA&S The Aircraft Airworthiness & Sustainment Conference
AUGUST 29 - SEPTEMBER 1, 2022

How to Select the Probability Distribution

- Given a set of data, there are several algorithms that can be employed to select the best fitting distribution.
  - Chi-squared, K-S, Anderson-Darling, Probability Plots
- We will focus on Probability Plots

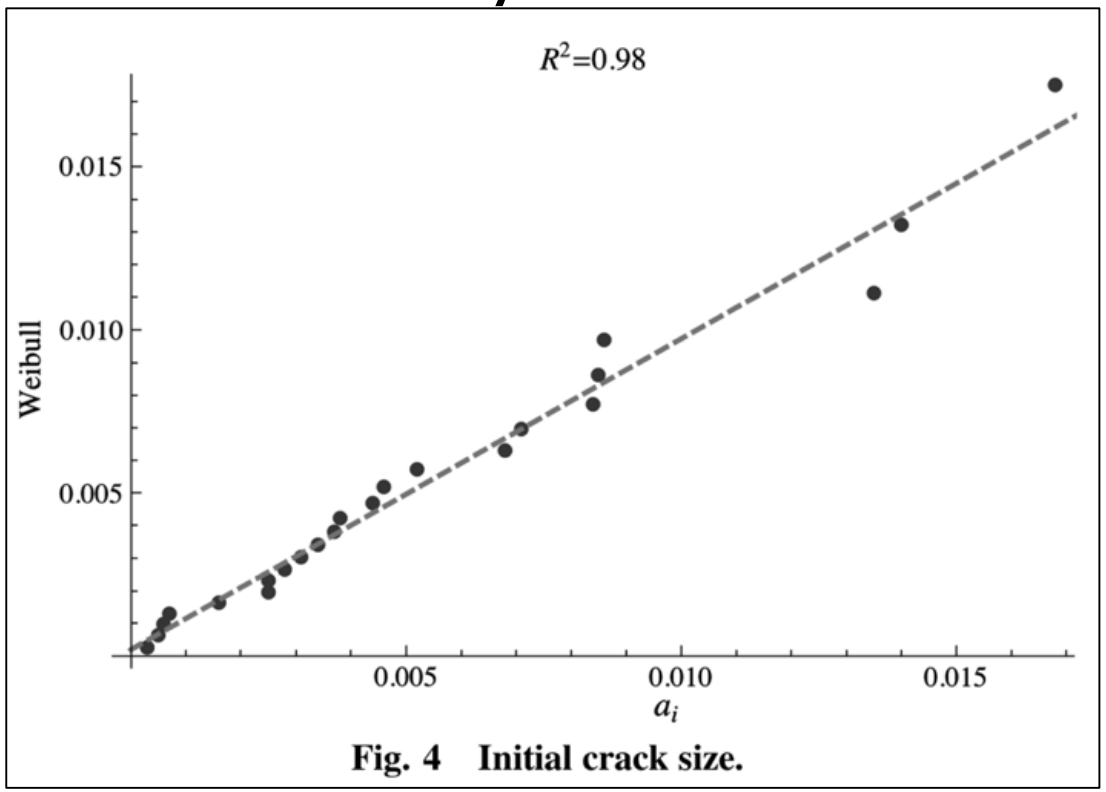


Fig. 4 Initial crack size.

H.R. Millwater, P. Wirsching - "Analysis Methods for Probabilistic Life Assessment", American Society of Metals Failure Analysis Handbook, V 11, 2003, pp. 250-268

Probability Plots

- Procedure
  1. Sort the data and create an empirical distribution
  2. Plot the data on a modified X-Y scale - modified for a particular distribution type.
  3. If the data plots as a straight line, that distribution type is a good fit.
  4. Repeat for other distribution types. Each has its own X-Y scale.

1. Empirical Distribution

- Sort the data and compute the CDF value for each data point as

$$F_X(x_i) = \frac{i - 1/2}{N}$$

Rank	x	F(x)
1	5.2	0.050
2	6.8	0.150
3	11.2	0.250
4	16.8	0.350
5	17.8	0.450
6	19.6	0.550
7	23.4	0.650
8	25.4	0.750
9	32.0	0.850
10	44.8	0.950

2. Scale Transformation

- Use an inverse CDF operation to develop a linear equation for each distribution type.
- Example: Gumbel max distribution

$$F(x) = \exp\left(-\exp\left(-a(x-b)\right)\right)$$

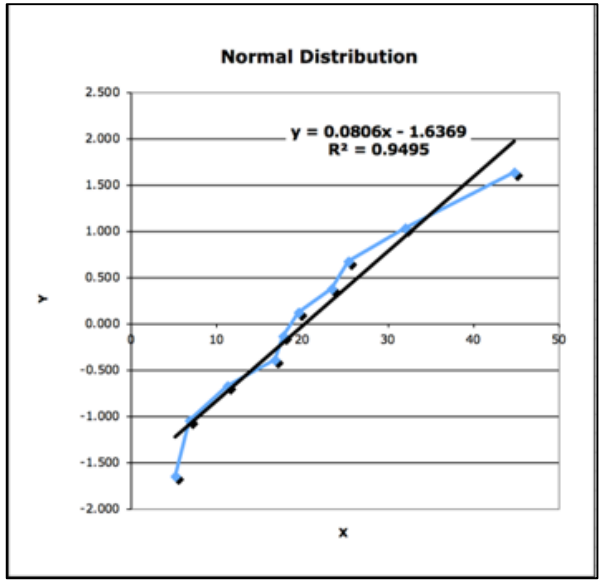
Taking the log twice of both sides leads to an equation

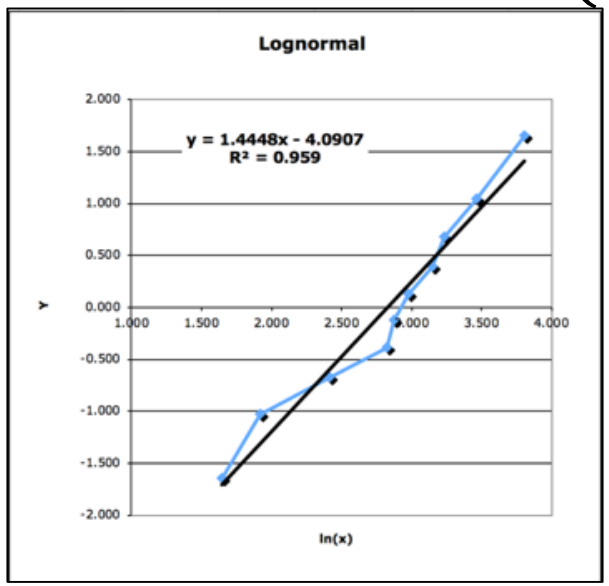
$$-\ln\left(-\ln\left(F(x)\right)\right) = ax - ab$$

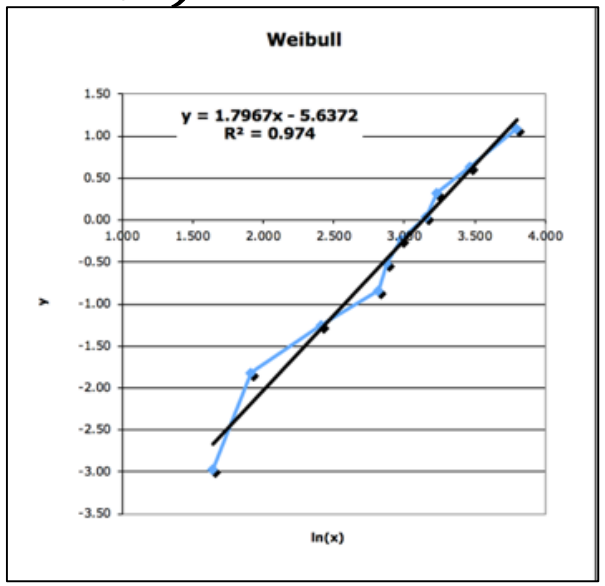
Therefore we let $y = -\ln\left(-\ln\left(F(x)\right)\right)$ and plot x vs. y

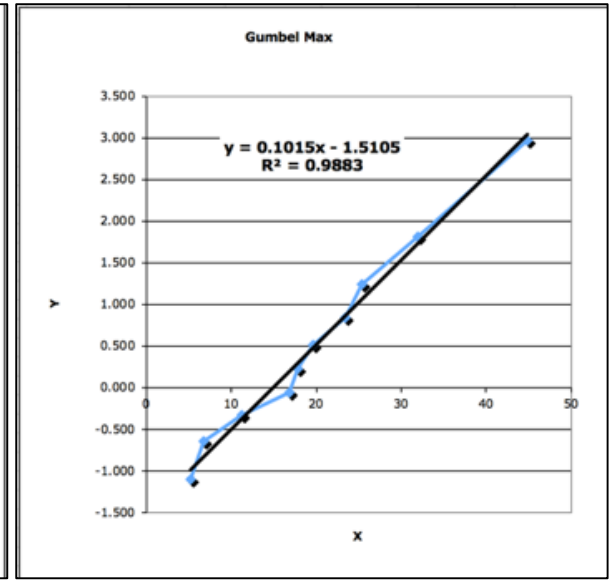
Transformation Eqns.

- Normal: x vs. $\Phi^{-1}\left(F(x_i)\right)$ (use norm.s.inv in Excel)
- Lognormal: ln(x) vs. $\Phi^{-1}\left(F(x_i)\right)$
- Weibull: ln(x) vs. $\ln(-\ln\left(1-F(x_i)\right))$
- Gumbel: x vs. $-\ln\left(-\ln\left(F(x_i)\right)\right)$
- Exponential: x vs. $-\ln\left(F(x_i)\right)$









Excel Example

- Use an inverse CDF operation to develop a linear equation for each distribution type.
- The spreadsheet "probPlotting.xlsx" is available for viewing and editing. The user only needs to provide the data.

Rank	x	F(x)	Exponential	EV Min	Gumbel (max)	ln(x)	Weibull	Normal	Lognormal
1	5.2	0.050	0.051	2.970	-1.097	1.649	-2.97	-1.645	-1.645
2	6.8	0.150	0.163	1.817	-0.640	1.917	-1.82	-1.036	-1.036
3	11.2	0.250	0.288	1.246	-0.327	2.416	-1.25	-0.674	-0.674
4	16.8	0.350	0.431	0.842	-0.049	2.821	-0.84	-0.385	-0.385
5	17.8	0.450	0.598	0.514	0.225	2.879	-0.51	-0.126	-0.126
6	19.6	0.550	0.799	0.225	0.514	2.976	-0.23	0.126	0.126
7	23.4	0.650	1.050	-0.049	0.842	3.153	0.05	0.385	0.385
8	25.4	0.750	1.386	-0.327	1.246	3.235	0.33	0.674	0.674
9	32	0.850	1.897	-0.640	1.817	3.466	0.64	1.036	1.036
10	44.8	0.950	2.996	-1.097	2.970	3.802	1.10	1.645	1.645
r^2			0.96	0.88	0.98		0.98	0.94	0.96

Questions