PROBABILITY DISTRIBUTIONS USED, THEIR TRANSFORMATION, LINEARIZATION, AND THE APPLICATION OF LINEAR REGRESSIION TO FIND THE PARAMETERS

On website www.waterlog.info

(see also the notes 1 to 4 at the end)

Symbols used:

Fc = cumulative frequency

The function fitting procedure starts with ranking the data (X) in ascending order and assigning rank frequencies Fr=R/(N+1), where R is the rank number and N the total number of data. Fr is also called cumulative frequency (Fc) or "plotting position".

X = stochastic variable

E = exponent

Ft = transformed Fc

Xt = transformed X

^ = raised to the power of an exponent

* = multiplication

/ = division

Sr(y) = square root of y

y = a variable

Ln(y) = natural logarithm (with base e) of y

e = 2.71 . . . .

Exp(y) = e^y

pi = 3.141 . . . .

The parameters A and B are found from a linear regression

of Ft on X (or Xt), except in case 11 (Pareto, see note 3)

and in cases 1a, 1b, 1c, 1d, and 1e (normal and transformed

normal distributions) where a numerical method (Hastings)

is used:

Ft = A.Xt + B

<u>1a Normal distribution</u> (*symmetric*)

Numerical method of Hastings is used. Briefly:

Use Z = 1/(1+0.232X)

N = {1/Sr(2pi)}Exp(-X^2/2)

Fc = 1 - N(1 0.319 Z - 0.357 Z^2 + 1.781 Z^3

- 1.821 Z^4 + 1.330 Z^5)

<u>1b Normal distribution optimized</u> (*symmetric*)

Numerical method of Hastings is used. Briefly:

Use Z = 1/(1+0.232X)

N = {1/Sr(2pi)}Exp(-X^2/2)

Fc = 1 - N(1 0.319 Z - 0.357 Z^2 + 1.781 Z^3

- 1.821 Z^4 + 1.330 Z^5)

The mean and std. dev. are optimized. See also note 1.

<u>1c Log-normal distribution simple</u> (*skew to right*)

Numerical method of Hastings is used as in 1a

while replacing X by Ln(X)

<u>1d Log-normal distribution optimized</u> (*skew to right*)

Numerical method of Hastings is used as in 7a

while replacing X by Ln(X)

The mean and std. dev. are optimized. See also note 1.

<u>1e Root-normal distribution</u> (*skew to right*)

Numerical method of Hastings is used as in 7a

while replacing X by Sr(X)

The mean and std. dev. are optimized. See also note 1.

1f Square-normal distribution (skew to left)

Numerical method of Hastings is used as in 7a

while replacing X by X^2

The mean and std. dev. are optimized. See also note 1.

<u>1g Generalized normal distribution</u> (*any skewness*)

Numerical method of Hastings is used as in 7a

while replacing X by X^E

The mean and std. dev. are optimized. See also note 1.

E is to be optimized (see note 2 below)

2a Logistic distribution (*symmetrical*)

Fc = 1/(1+Exp(A*X+B)

Ft = Ln(-1+1/Fc)

Ft = A*X + B

2b Log-logistic distribution (*skew to right*)

Fc = 1/(1+Exp(A*Ln(X)+B)

Xt = Ln(X)

Ft = Ln(-1+1/Fc)

Ft = A*Xt + B

2c Generalized logistic distribution (*any skewness*)

Fc = 1/(1+Exp(A*X^E+B)

Xt = Ln(X^E) = E*Ln(X)

Ft = Ln(-1+1/Fc)

Ft = A*Xt + B

E is to be optimized (see note 2 below)

3a Cauchy distribution (*symmetrical*)

Fc = (1/pi)*arctan(A*X+B) + 0.5

Ft = tan{pi*(Fc-0.5)}

Ft = A*X + B

3b Cauchy generalized (*any skewness*)

Fc = (1/pi)*arctan(A*X^E+B) + 0.5

Xt = X^E

Ft = tan{pi*(Fc-0.5)}

Ft = A*Xt + B

E is to be optimized (see note 2 below)

4a Exponential distribution, Poisson-type (*Skew to right*)

Fc = 1 - Exp (- A*X)

Xt = Ln (X)

Ft = - Ln (1-Fc)

Ft = A*Xt

4a Generalized (negative) exponential distribution (*Poisson-type, skew to right*)

Fc = 1 - Exp{-(A*X^E+B)}

Xt = Ln(X^E) = E*Ln(X)

Ft = - Ln (1-Fc)

Ft = A*Xt + B

E is to be optimized (see note 2 below)

4b Mirrored exponential distribution generalized (*Skew to left*)

Fc = Exp{-(A*X^E+B)}

Xt = Ln(X^E) = E*Ln(X)

Ft = -Ln (Fc)

Ft = A*Xt + B

E is to be optimized (see note 2 below)

5a Gumbel (Fisher-Tippett type I) distribution (*skew to right*)

Fc = Exp[-Exp{-(A*X+B)}]

Ft = -Ln{-Ln(Fc)}

Ft = A*X + B

5b Generalized Gumbel distribution (*any skewness*)

Fc = Exp[-Exp{-(AX^E+B)}]

Xt = Ln(X^E) = E*Ln(X)

Ft = -Ln{-Ln(Fc)}

Ft = A*Xt + B

E is to be optimized (see note 2 below)

5c Mirrorred Gumbel distribution (*skew to left*)

Fc = 1 - Exp[-Exp{-(AX+B)}]

Ft = -Ln{-Ln(1-Fc)}

Ft = A*X + B

5d Generalized mirrorred Gumbel distribution (*any skewness*)

Fc = 1 - Exp{-Exp{-A*X^E+B)}

Xt = Ln(X^E) = E*Ln(X)

Ft = -Ln{-Ln(1-Fc)}

Ft = A*Xt + B

E is to be optimized (see note 2 below)

6a Student's t-distribution with 1 degree of freedom (*symmetrical*)

Fc = 0.5 + arctan{(X-AvX)/StD}/pi

AvX = Average of X

StD = Standard deviation of X

6b Student's t-distribution with 2 degrees of freedom (*symmetrical*)

Fc = 0.5{1+(RedX)}/Sr(2+RedX^2)

RedX = (X-AvX)/StD (reduced X)

for other symbols: see 8a

7a Weibull distribution (*skew to right*)

Fc = 1 - Exp{-(X/C)^A}

with C = Exp(-B/A)

Xt = Ln(X)

Ft = Ln{-Ln(1-Fc)}

Bt = B/A

Ft = A*Xt + Bt

7b Weibull generalized (*any skewness*)

Fc = 1 - Exp{-(X^E/C)^A}

with C = Exp(-B/A)

Xt = Ln{Ln(X)}

Ft = Ln{-Ln(1-Fc)}

Bt = B/A

Ft = A*Xt + Bt

E is to be optimized (see note 2 below)

The next 3 distributions are bounded:

8. Frechet (Fisher-Tippett type II) distribution (*skew to right*)

Fc = Exp[-{(X-C)/Exp(-B/A)}^A]

Xt = Ln(X-C) [X>C]

Ft = Ln{-Ln(Fc)}

Ft = A*Xt + B

C is to be optimized (see note 4 below)

9. Fisher-Tippett type III distribution (*skew to right*)

Fc = Exp[-{(C-X)/Exp(-B/A)}^A]

Xt = Ln(C-X)   [X<C]

Ft = Ln{-Ln(Fc)}

Ft = A*Xt + B

C is to be optimized (see note 4 below)

10 Pareto-Lomax distribution (*skew to right*)

Fc = 1 - {B/(X+B)}^A   [B>0, X>-B]

Xt = Ln{B/(X+B)}

Ft = Ln(1-Fc)

Ft = A*Xt

B is to be optimized (see note 3 below)

NOTES

1. The optimization of the standard deviation in a normal distribution is explained in: https://www.waterlog.info/pdf/stdev.pdf

2. The exponent E in the generalized distributions is optimized using a range of values and selecting the value giving the minimum sum of squares of deviations of calculated and observed cumulative frequencies.

3. The Pareto (case 14) distribution uses the ratio method instead of a linear regression to find the parameter A, while the parameter B is optimized in a similar way as explained under note 2.

4. The parameter C in the Frechet (case 14) and F-T III distribution (case 17) is optimized in a similar way as explained under note 2.