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

(see also the notes 1 to 5 at the end)

On www.waterlog.info public domain

Symbols used:
  Fc = cumulative frequency
       (see explantion in the introduction)
  X = stochastic variable
  A = distribution parameter
  B = distribution parameter
  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 . . . .

  Below, the parameters A and B are found from a linear regression of Ft on X (or Xt), except in case 7a and b (Student) and case 12 (GEV), where the ratio method is used to find parameter A, and in cases 1a, 1b, 1c, 1d, and 1e (normal and transformed normal distrunbutions), where a numerical method (Hastings) is used.

1a Normal distribution (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)

1b Normal distribution optimized (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.

1c Log-normal distribution simple (skew to right)
   Numerical method of Hastings is used as in 7a
   while replacing X by Ln(X)

1d Log-normal distribution optimized (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.

1e Root-normal distribution (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.

1g Generalized normal distribution (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

A and B are found from a linear regression

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

A and B are found from a linear regression

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

A and B are found from a linear regression

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

A and B are found from a linear regression

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

A and B are found from a linear regression

E is to be optimized (see note 2 below)

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

Fc = 1 - Exp (- A*X)

Xt = Ln (X)

Ft = - Ln (1-Fc)

Ft = A*Xt (use ratio method to find A)

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

A and B are found from a linear regression

E is to be optimized (see note 2 below)

4b Mirrored exponential distribution generalized
   (Skew to left)
   $Fc = Exp\{-(A*X^E+B)\}$
      $Xt = Ln(X^E) = E*Ln(X)$
      $Ft = -Ln\ (Fc)$
   $Ft = A*Xt + B$
   A and B are found from a linear regression
   E is to be optimized (see note 2 below)

5a Gumbel (Fisher-Tippett type I) distribution
   (skew to right)
   $Fc = Exp[-Exp\{-(A*X+B)\}]$
      $Ft = -Ln\{-Ln(Fc)\}$
   $Ft = A*X + B$
   A and B are found from a linear regression

5b Generalized Gumbel distribution (any skewness)
   $Fc = Exp[-Exp\{-(AX^E+B)\}]$
      $Xt = Ln(X^E) = E*Ln(X)$
      $Ft = -Ln\{-Ln(Fc)\}$
   $Ft = A*Xt + B$
   A and B are found from a linear regression
   E is to be optimized (see note 2 below)

5c Mirrored Gumbel distribution (skew to left)
   $Fc = 1 - Exp[-Exp\{-(AX+B)\}]$
      $Ft = -Ln\{-Ln(1-Fc)\}$
   $Ft = A*X + B$
   A and B are found from a linear regression

5d Generalized mirrored Gumbel distribution
   (any skewness)
   $Fc = 1 - Exp\{-Exp\{-A*X^E+B)\}$
      $Xt = Ln(X^E) = E*Ln(X)$
      $Ft = -Ln\{-Ln(1-Fc)\}$
   $Ft = A*Xt + B$
   A and B are found from a linear regression
   E is to be optimized (see note 2 below)

6. Generalized Kumaraswamy distribution
   (any skewness)
   $Fc = 1 - \{1 - (X/M)^A\}^B$
      $Xt = Ln\{(X/M)^A\} = A*Ln(X/M)$
      $Ft = Ln(1-Fc)$
   $Ft = B*Xt$ (use ratio method to find B
   $M > Xmax$ is to be optimized (see note 3)

7a Laplace distribution (composite, any skewness)
   if X<B
      Fc = 0.5*Exp{A1*(X-B)}
         Ft = ln(2Fc)
         C = -A1*B
      Ft = A1*X + C
      A1 and B1 are found from a linear regression
   if X>B
      Fc = 1 - 0.5*exp{A2*(X-B)
         Ft = ln(0.5) - ln(1-Fc)
      Ft = A2*X (use ratio method to find A2)
   B is to be optimized (see note 3)

7b Laplace distribution generalized (any skewness)
   if X<B
      Fc = 0.5*Exp{A1*(X^E1-B)}
         Ft = ln(2Fc)
         C = -A1*B
      Ft = A1*X^E1 + C
      A1 and B1 are found from a linear regression
   if X>B
      Fc = 1 - 0.5*exp{A2*(X^E2-B)
         Ft = ln(0.5) - ln(1-Fc)
      Ft = A2*X^E2 (use ratio method to find A2)
   B is to be optimized (see note 3)
   Exponents E1 and E2 are to be optimized (note 2)

8a 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

8b 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

9a Weibull distribution (skew to right)

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

with $C = Exp(-B/A)$

$Xt = Ln(X)$

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

$Bt = B/A$

$Ft = A*Xt + Bt$

A and Bt are found from a linear regression

9b Weibull generalized (any skewness)

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

with $C = Exp(-B/A)$

$Xt = Ln\{Ln(X)\}$

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

$Bt = B/A$

$Ft = A*Xt + Bt$

A and Bt are found from a linear regression

E is to be optimized (see note 2 below)

The next 7 are bounded:

10 Burr (generalized Pareto-Lomax) distribution (skew to right)

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

$Xt = Ln[\{B/(X+B)\}^A]$

$Ft = Ln(1-Fc)$

$Ft = E*Xt$ (use ratio method to find E)

B and A are to be optimized (see note 3 below)

11 Dagum (mirrored Burr) distribution (skew to left)

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

$Xt = Ln[\{B/(X+B)\}^A]$

$Ft = Ln(Fc)$

$Ft = E*Xt$ (use ratio method to find E)

B and A are to be optimized (see note 3 below)

12 Generalized extreme value (GEV) distribution

$Fc = exp[-\{1+K(X-A)/B\}^{(-1/K)}]$

K, A and B are to be optimized (see note 2)

13 Frechet (Fisher-Tippett type II) distribution
(skew to right)

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

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

$Ft = Ln\{-Ln(Fc)\}$

$Ft = A*Xt + B$

A and B are found from a linear regression

C is to be optimized (see note 4 below)

14 Fisher-Tippett type III distribution
   (skew to right)
     $Fc = Exp[-\{(C-X)/Exp(-B/A)\}^{\wedge}A]$
         $Xt = Ln(C-X)$ [X<C]
         $Ft = Ln\{-Ln(Fc)\}$
     $Ft = A*Xt + B$
     A and B are found from a linear regression
     C is to be optimized (see note 4 below)

15 Generalized Gompertz distribution (any skewness)
     $Fc = 1 - exp[A*\{exp(B*Z^{\wedge}E) - 1\}]$ with Z=ln(X)
         $Zt = exp(B*Z^{\wedge}E) - 1$
         $Ft = ln (1-F);$
     $Ft = A*Zt;$ (use ratio method to find A)
     B and E are to be optimized (see note 3 below)

16 Pareto-Lomax distribution (skew to right)
     $Fc = 1 - \{B/(X+B)\}^{\wedge}A$ [B>0, X>-B]
         $Xt = Ln\{B/(X+B)\}$
         $Ft = Ln(1-Fc)$
     $Ft = A*Xt$ (use ratio method to find A)
     B is to be optimized (see note 3 below)

NOTES

 1. The optimization of the standard deviation in a normal
 distribution is explained in: 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 Kumaraswamy (case 6), Laplace (case 7), Burr (case 10),
 Dagum (case 11), Gompertz (case 15), and Pareto (case 16)
 distributions use the ratio method instead of a linear regression
 to find the parameters B, B, B & A, B & E, B & E, and B respectively.
 The parameters A (Laplace), B (Pareto), A and B (Burr and Dagum),
 and B and E (Gompertz) are optimized in a similar way as explained
 under note 2.

 4. The parameter C in the Frechet (case 10) and F-T III distribution
 (case 12) are optimized in a similar way as explained under note 2.
 The same holds for the parameter M in the Kumaraswamy
 distribution (case 6)