

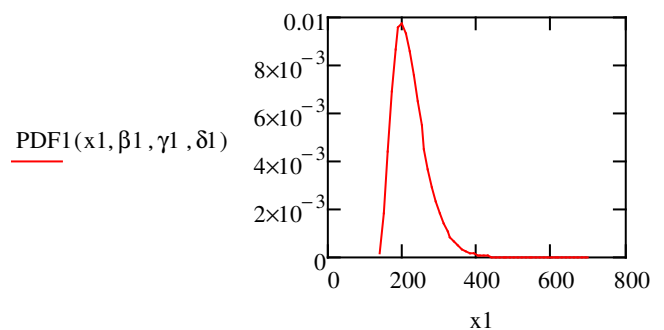
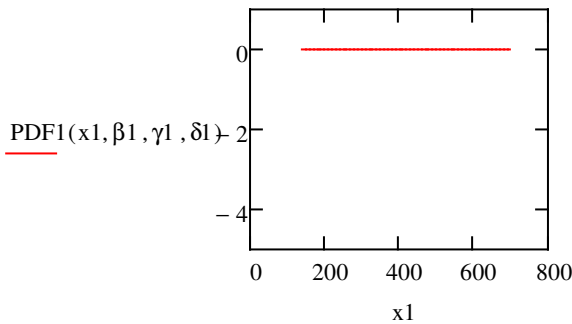
## The Fitted Gamma Function

This worksheet demonstrates the shape and behavior of the Gamma Distribution and fits data to a three parameter version of the function. Here is the PDF function that is to be fitted:

$$\beta_1 := 25 \quad \gamma_1 := 3.5 \quad \delta_1 := 135 \quad x_1 := 0, 10.. 700 \quad \text{Typical values}$$

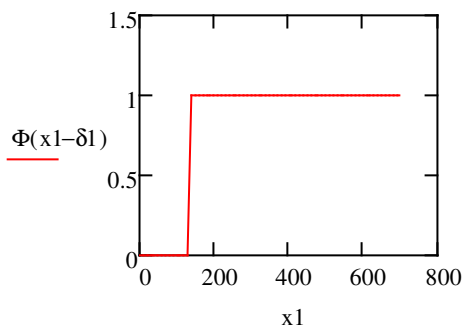
$$\text{PDF1}(x_1, \beta_1, \gamma_1, \delta_1) := \frac{\left(\frac{x_1 - \delta_1}{\beta_1}\right)^{\gamma_1 - 1} \cdot e^{-\frac{x_1 - \delta_1}{\beta_1}}}{\beta_1 \cdot \Gamma(\gamma_1)} \quad \text{Offset Gamma Function}$$

beta is a scale parameter; gamma is a shape parameter and delta is the offset from zero where the function "starts". Here is a typical example of this curve:



The PDF given above behaves incorrectly for x less than 135. The PDF function value needs to be set to zero for x less than delta (135). This construct is one way to do it:

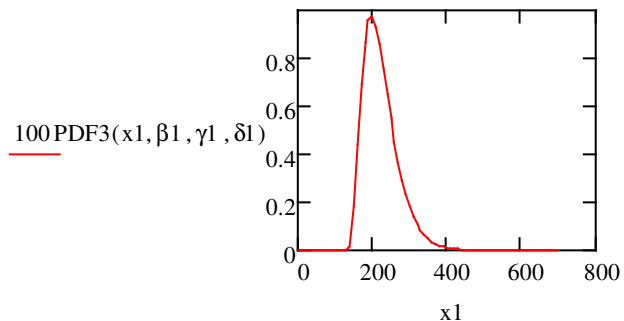
$$\text{PDF2}(x_1, \beta_2, \gamma_2, \delta_2) := \text{if}(x_1 > \delta_2, \text{PDF1}(x_1, \beta_1, \gamma_1, \delta_1), 0)$$



We can also use the "step function" to the left to force the PDF curve to equal zero for x less than delta:

$$\text{PDF3}(x_1, \beta_1, \gamma_1, \delta_1) := \text{PDF1}(x_1, \beta_1, \gamma_1, \delta_1) \cdot \Phi(x_1 - \delta_1)$$

PDF3 is more tractable for subsequent use than is PDF2



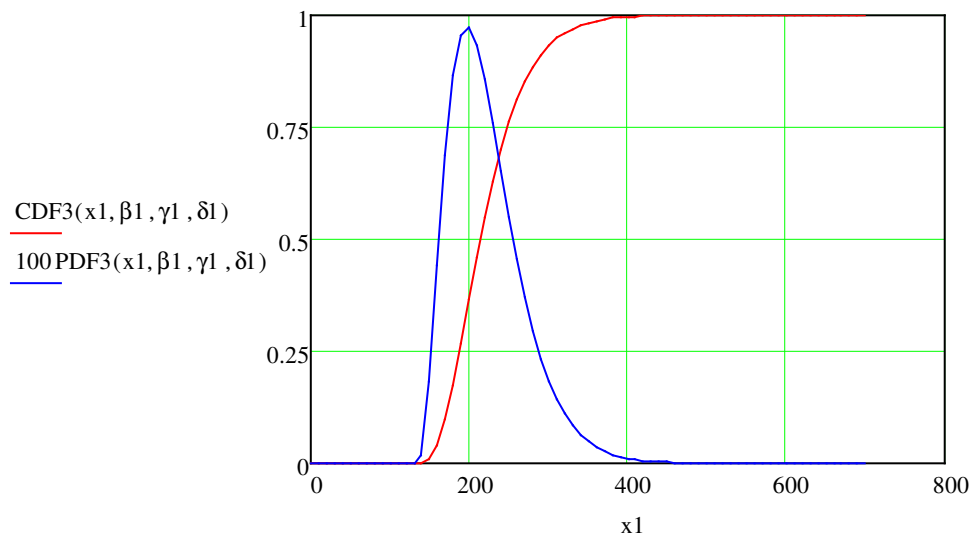
The next hurdle is that though the PDF to be fit is the PDF2 given above ... what we need to do is fit the integral of PDF2 ... the cumulative distribution function (CFD). The CFD is the data presented in available journal papers and The Book of Myrtle.

The CDF is the definite integral of PDF2 from 0 to x

k := 0, 1..59

$$\text{CDF3}(x1, \beta1, \gamma1, \delta1) := \int_0^{x1} \text{PDF2}(z, \beta1, \gamma1, \delta1) dz$$

The function to be fitted is the cumulative distribution function



The above work up uses typical values for the parameters that define the PDF and the CDF. Now we do it all over again ... this time with real data.

Here is the data to be fit. It is from The Book of Myrtle Fig 17

data :=  $\begin{pmatrix} 0 & 0 \\ 30 & 0 \\ 60 & 0 \\ 90 & 0 \\ 120 & 0 \\ 150 & 0 \\ 180 & 0 \\ 210 & 0 \\ 240 & 0 \\ 270 & 0 \\ 300 & 0 \\ 330 & 0.11 \\ 360 & 0.1 \\ 390 & 0.33 \\ 420 & 0.5 \\ 450 & 0.57 \\ 510 & 1 \\ 540 & 1 \\ 570 & 1 \\ 600 & 1 \\ 630 & 1 \\ 700 & 1 \end{pmatrix}$

We split the data vector into two vectors ... one for x and one for y:

$x := \text{data}^{\langle 0 \rangle}$

$y := \text{data}^{\langle 1 \rangle}$

$\text{length}(x) = 22$

$\text{Range} := \text{length}(x)$

$n1 := \text{length}(x) - 1$

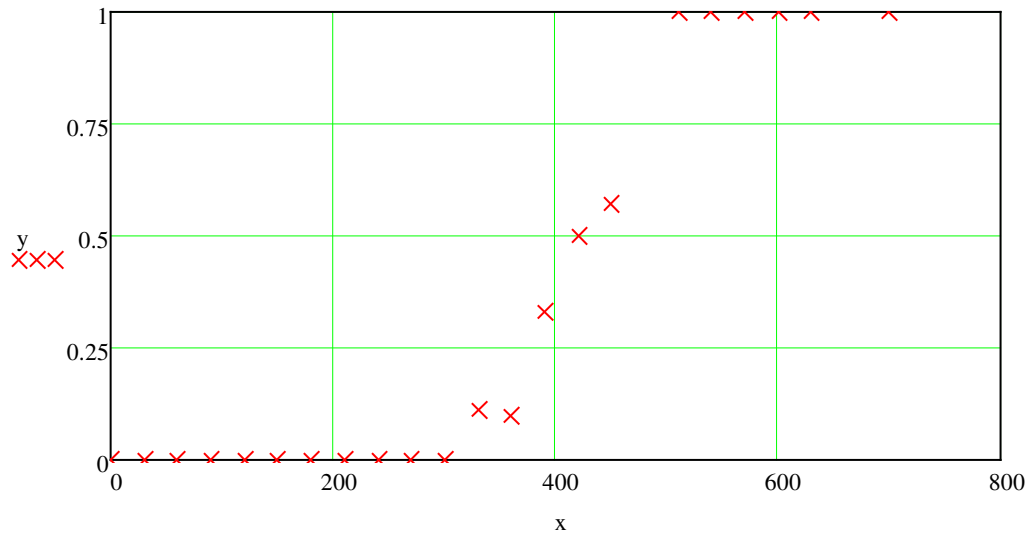
$n1 = 21$

$\text{Range} = 22$

	0	1
0	0	0
1	30	0
2	60	0
3	90	0
4	120	0
5	150	0
6	180	0
7	210	0
8	240	0
9	270	0
10	300	0
11	330	0.11
12	360	0.1
13	390	0.33
14	420	0.5
15	450	...

data =

Here is a plot of this data:



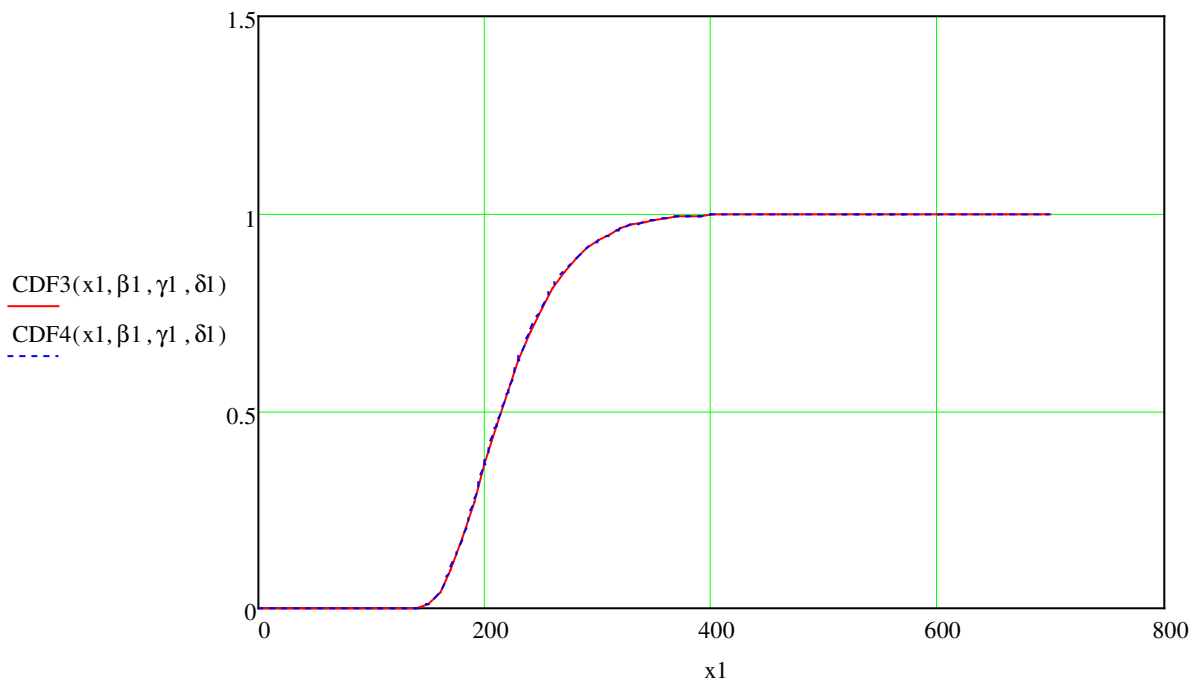
Now we set up for estimation of the three parameters beta, gamma, and delta

Here, from above, is equation that we need to fit the data to:

$$\text{CDF3}(x1, \beta1, \gamma1, \delta1) := \int_0^{x1} \text{PDF3}(z, \beta1, \gamma1, \delta1) dz$$

As it happens, the data is not a continuous function of x ... therefore we use the discrete version

$$\text{CDF4}(x1, \beta1, \gamma1, \delta1) := \sum_{z=0}^{x1} \text{PDF3}(z, \beta1, \gamma1, \delta1)$$



It seems that, when all is said and done, the two formulae are about the same.

In any case, we now move on to the estimation task.

Here we start a "solve block" to estimate  $\beta_1$  and  $\delta_1$

$k := 0, 1..n1$

$$SSE1(\beta_2, \gamma_2, \delta_2) := \sum_k (y_k - CDF4(x_k, \beta_2, \gamma_2, \delta_2))^2$$

the sum of the squared errors is to be minimized

$$SSE5(\beta_2, \gamma_2, \delta_2) := \sum_k |y_k - CDF4(x_k, \beta_2, \gamma_2, \delta_2)|$$

= scale      = shape       $\delta$  = offset

Given

$$SSE5(\beta_2, \gamma_2, \delta_2) = 0$$

$\beta_2 := 19.8$        $\gamma_2 := 8$        $\delta_2 := 267$       initial guess for parameter values

$10 < \beta_2 < 30$

$3 < \gamma_2 < 10$

$200 < \delta_2 < 320$

$$\begin{pmatrix} \beta_{1est} \\ \gamma_{1est} \\ \delta_{1est} \end{pmatrix} := \text{Minerr}(\beta_2, \gamma_2, \delta_2)$$

$\beta_{1est} = 19.882$        $\gamma_{1est} = 8.027$        $\delta_{1est} = 267.491$       Resultant estimated parameter values

$$PDF3(x, \beta_{1est}, \gamma_{1est}, \delta_{1est}) := PDF1(x, \beta_{1est}, \gamma_{1est}, \delta_{1est}) \cdot \Phi(x - \delta_{1est})$$

$\beta_{1est} = 19.882$

$$CDF5(x1, \beta_{1est}, \gamma_{1est}, \delta_{1est}) := \int_0^{x1} PDF3(z, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dz$$

$\gamma_{1est} = 8.027$

$$CDF4(x1, \beta_{1est}, \gamma_{1est}, \delta_{1est}) := \sum_{z=0}^{x1} PDF2(z, \beta_{1est}, \gamma_{1est}, \delta_{1est})$$

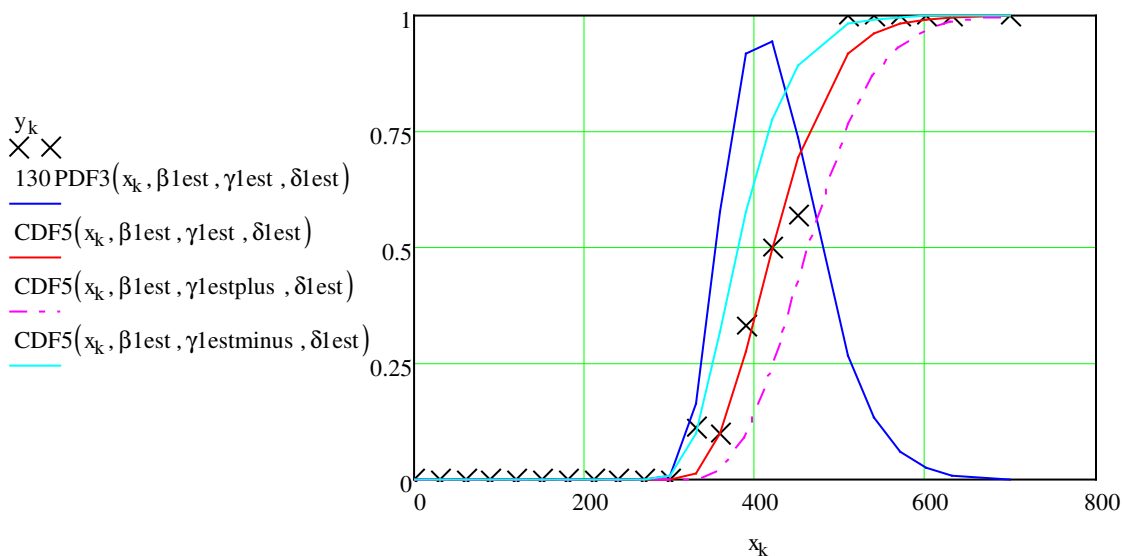
$\delta_{1est} = 267.491$

$\gamma_{1estplus} := 1.25 \gamma_{1est}$

$\gamma_{1estminus} := .75 \gamma_{1est}$

$\gamma_{1estplus} = 10.033$

$\gamma_{1estminus} = 6.02$



Now that we have basic curve fitting going we can calculate various parameters of the fitted curves.

Here are the first four moments of the fitted PDF3

From CRC Statistics Handbook p 77and 96

First moment - Mean

$$M1 := \int_{\delta_{1est}}^{\infty} x \cdot PDF3(x, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dx \quad m1 := \gamma_{1est} \cdot \beta_{1est} + \delta_{1est}$$

$$M1 = 427.076 \quad m1 = 427.076$$

Second moment - Variance

$$M2 := \int_{\delta_{1est}}^{\infty} (x - M1)^2 \cdot PDF3(x, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dx \quad m2 := \gamma_{1est} \cdot \beta_{1est}^2$$

$$M2 = 3.173 \times 10^3 \quad m2 = 3.173 \times 10^3$$

Standard Deviation

$$StdDev := \sqrt{M2} \quad StdDev = 56.328 \quad Mode := \beta_{1est} \cdot (\gamma_{1est} - 1) + \delta_{1est}$$

$$Mode = 407.194$$

Third moment - Skew

$$M3 := \int_{\delta_{1est}}^{\infty} (x - M1)^3 \cdot PDF3(x, \beta_{1est}, \gamma_{1est}, \delta_{1est}) \cdot \left(\frac{1}{StdDev}\right)^3 dx \quad m3 := \frac{2}{\sqrt{\gamma_{1est}}}$$

$$M3 = 0.706 \quad m3 = 0.706$$

$$CoefVariation := \frac{1}{\sqrt{\gamma_{1est}}} \quad CoefVariation = 0.353$$

Fourth moment - Kurtosis

$$M4 := \int_{\delta_{1est}}^{\infty} (x - M1)^4 \cdot PDF3(x, \beta_{1est}, \gamma_{1est}, \delta_{1est}) \cdot \left(\frac{1}{StdDev}\right)^4 dx \quad m4 := 3 \left(1 + \frac{2}{\gamma_{1est}}\right)$$

$$M4 = 3.748 \quad m4 = 3.748$$

Some Quartiles:

$$\int_0^{177} PDF3(z, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dz = 0 \quad \int_0^{188} PDF3(z, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dz = 0$$

$$\int_0^{211} PDF3(z, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dz = 0 \quad \int_0^{243} PDF3(z, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dz = 0$$

$$\int_0^{285} PDF3(z, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dz = 3.875 \times 10^{-6} \quad \int_0^{330} PDF3(z, \beta_{1est}, \gamma_{1est}, \delta_{1est}) dz = 0.015$$

Now we use the previously determined shape function to estimate the PDF from the given mean and StdDev

$$\text{Shape factor} = \gamma = \gamma_{\text{lest}} \quad \beta_{\text{lest}} = 19.882 \quad \gamma_{\text{lest}} = 8.027 \quad \delta_{\text{lest}} = 267.491$$

For the gamma function we have the these relationships

$$\text{Mean}_{\text{est}} := \gamma_{\text{lest}} \cdot \beta_{\text{lest}} + \delta_{\text{lest}} \quad \text{Variance}_{\text{est}} := \gamma_{\text{lest}} \cdot \beta_{\text{lest}}^2 \quad \text{StdDev}_{\text{est}} := \beta_{\text{lest}} \cdot \sqrt{\gamma_{\text{lest}}}$$

These equations can be solved for  $\beta_2$  and  $\delta_2$  when the Mean, StdDev2 and shape factor  $\gamma_2$  are known

$$\text{Mean}_{\text{est}} = 427.076 \quad \text{StdDev}_{\text{est}} = 56.328$$

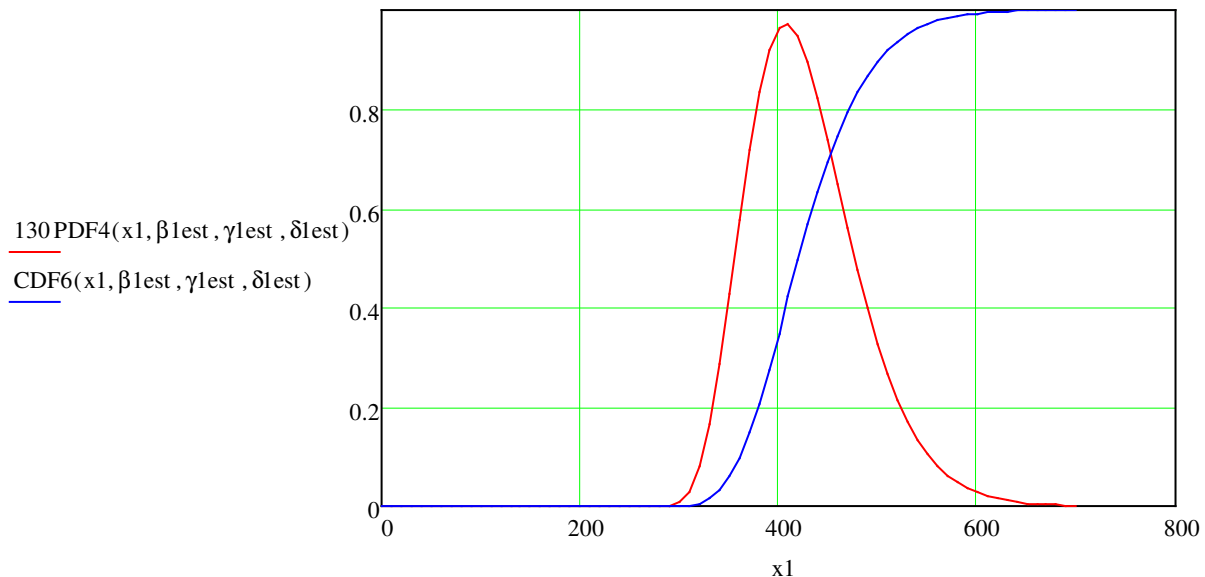
$$\beta_4 := \frac{\text{StdDev}_{\text{est}}}{\sqrt{\gamma_{\text{lest}}}} \quad \delta_4 := \text{Mean}_{\text{est}} - \text{StdDev}_{\text{est}} \cdot \left( \frac{\gamma_{\text{lest}}}{\sqrt{\gamma_{\text{lest}}}} \right)$$

$$\beta_4 = 19.882 \quad \delta_4 = 267.491 \quad \gamma_{\text{lest}} = 8.027$$

The resultant PDF is:

$$\text{PDF4}(x_1, \beta_{\text{lest}}, \gamma_{\text{lest}}, \delta_{\text{lest}}) := \frac{\left[ \left( \frac{x_1 - \delta_{\text{lest}}}{\beta_{\text{lest}}} \right)^{\gamma_{\text{lest}} - 1} \cdot e^{-\frac{x_1 - \delta_{\text{lest}}}{\beta_{\text{lest}}}} \right]}{\beta_{\text{lest}} \cdot \Gamma(\gamma_{\text{lest}})} \cdot (\Phi(x_1 - \delta_{\text{lest}}))$$

$$\text{CDF6}(x_1, \beta_{\text{lest}}, \gamma_{\text{lest}}, \delta_{\text{lest}}) := \int_0^{x_1} \text{PDF4}(z, \beta_{\text{lest}}, \gamma_{\text{lest}}, \delta_{\text{lest}}) dz$$



Guess value:

$$Q001 := 300 \quad Q10 := 350 \quad Q25 := 350 \quad Q50 := 350 \quad Q75 := 350 \quad Q90 := 350$$

Now we use a "Solve Block" to find all the Quantiles of interest

Quantile .001:      Guess value: 200      Q001 = 300

Given

$$\text{CDF}(x1, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) := \int_0^{x1} \text{PDF4}(z, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) dz$$

$$\text{CDF}(Q001, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) = .001 \quad T001 := \text{Find}(Q001) \quad T001 = 306.922$$

Quantile .1:      Guess value: 250      Q10 = 350

Given

$$\text{CDF}(x1, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) := \int_0^{x1} \text{PDF4}(z, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) dz$$

$$\text{CDF}(Q10, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) = .1 \quad T10 := \text{Find}(Q10) \quad T10 = 360.471$$

Quantile .25:      Guess value: 250      Q25 = 350

Given

$$\text{CDF}(x1, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) := \int_0^{x1} \text{PDF4}(z, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) dz$$

$$\text{CDF}(Q25, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) = .25 \quad T25 := \text{Find}(Q25) \quad T25 = 386.375$$

Quantile .5:      Guess value: 275      Q50 = 350

Given

$$\text{CDF}(x1, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) := \int_0^{x1} \text{PDF4}(z, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) dz$$

$$\text{CDF}(Q50, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) = .5 \quad T50 := \text{Find}(Q50) \quad T50 = 420.5$$

Quantile .75:      Guess value: 300      Q75 = 350

Given

$$\text{CDF}(x1, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) := \int_0^{x1} \text{PDF4}(z, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) dz$$

$$\text{CDF}(Q75, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) = .75 \quad T75 := \text{Find}(Q75) \quad T75 = 460.631$$

Quantile .9:      Guess value: 400      Q90 = 350

Given

$$\text{CDF}(x1, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) := \int_0^{x1} \text{PDF4}(z, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) dz$$

$$\text{CDF}(Q90, \beta1\text{est}, \gamma1\text{est}, \delta1\text{est}) = .9 \quad T90 := \text{Find}(Q90) \quad T90 = 502.172$$