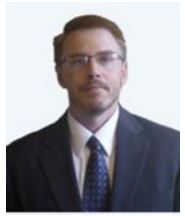
Probability Methods
Probability Distributions

Aircraft Airworthiness and Sustainment Conference

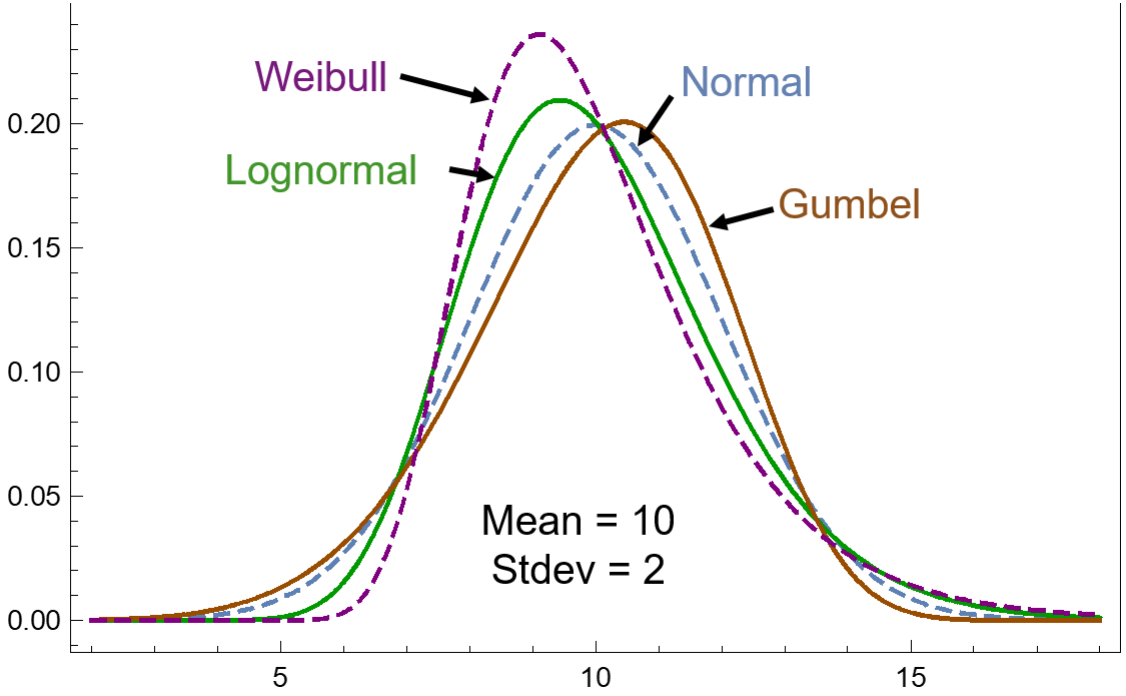
August 29, 2022






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

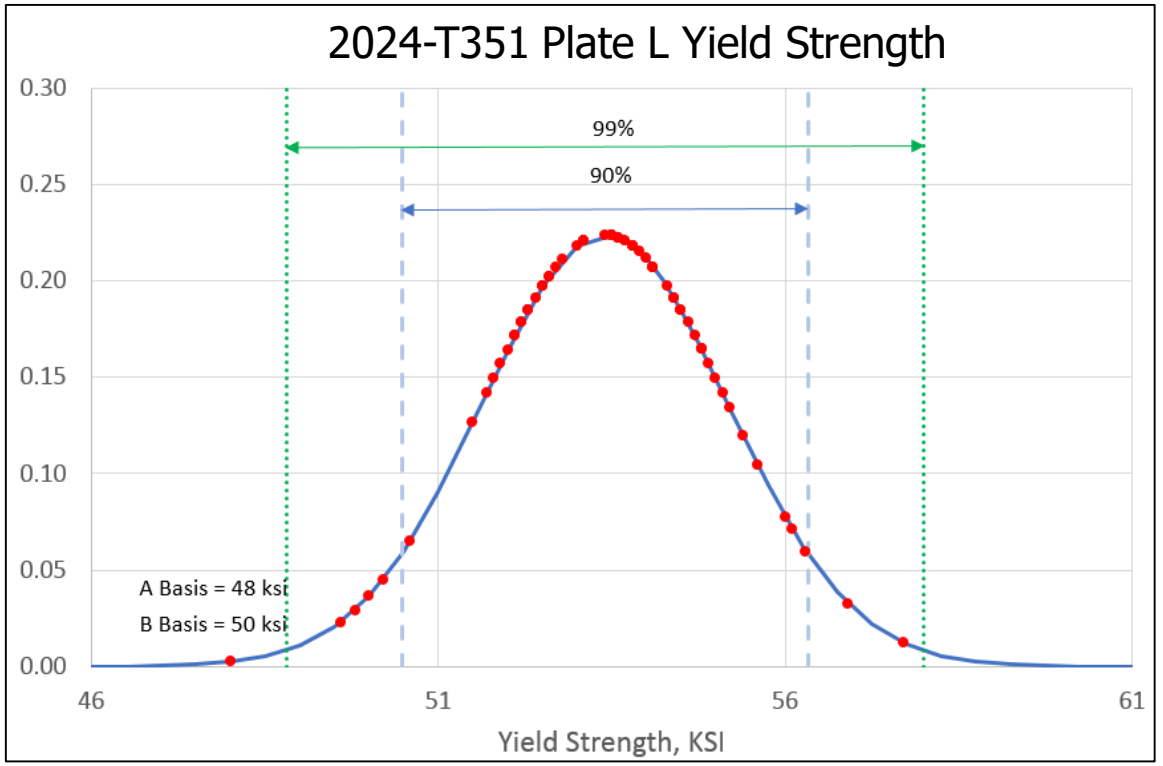
Common Engineering Distributions



- We will look at 4 common distributions used in engineering:
  - Normal, Lognormal, Weibull, Gumbel
- And for cases where a distribution is not available
  - Tabular

Normal Distribution

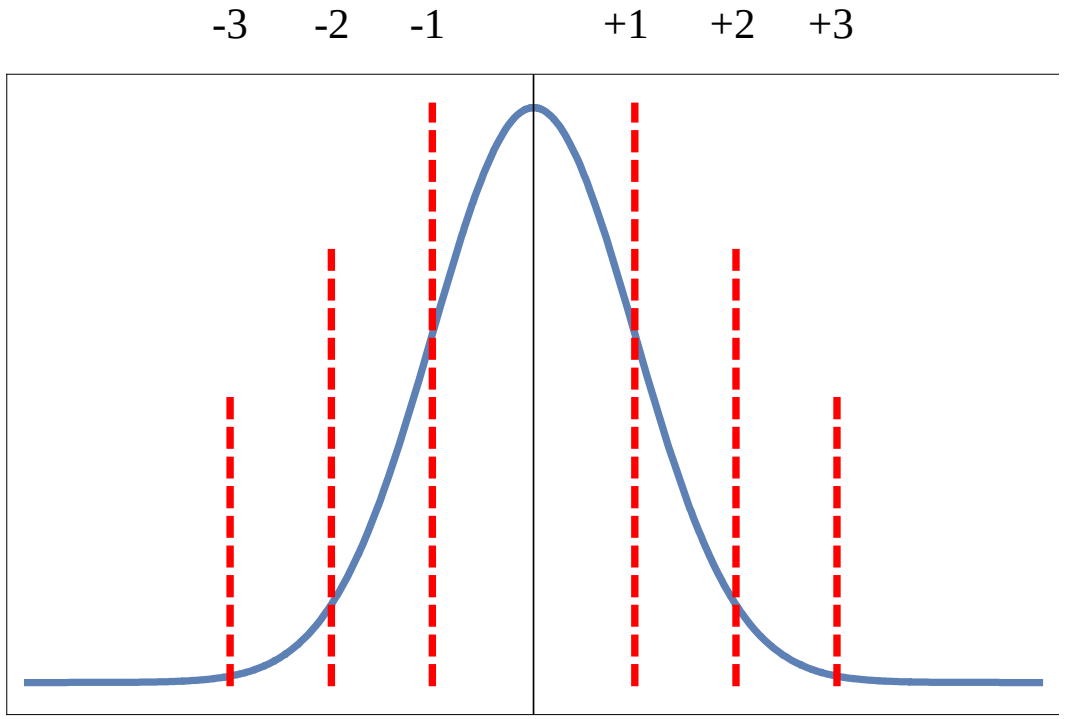
- Commonly used to represent statistical errors.
- Two parameters, mean (μ), standard deviation (σ).
- Some PDTA parameters are normal, e.g., fracture toughness, material strength.



Normal Distribution

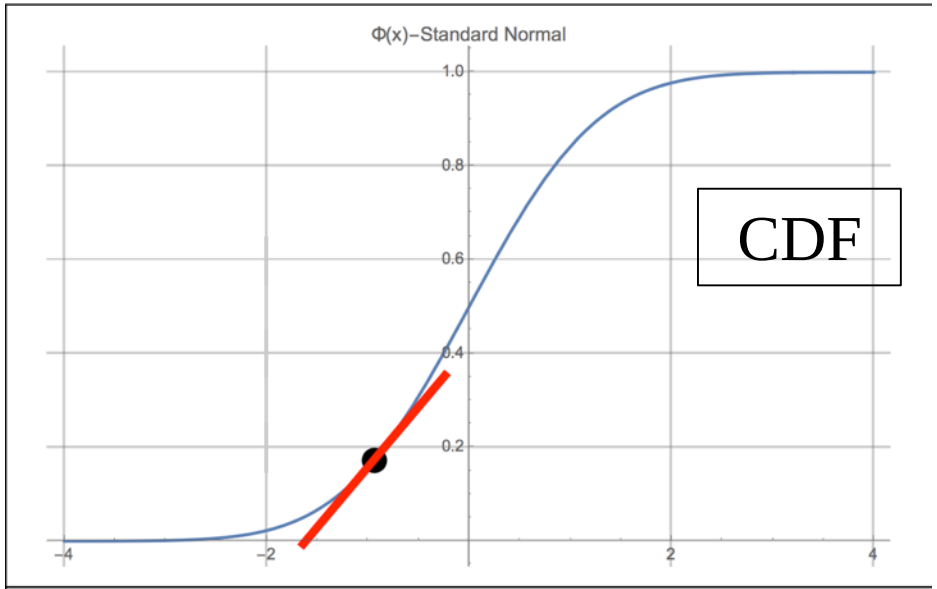
- Probability density function is symmetric
- Sum of Normal distributions is Normal

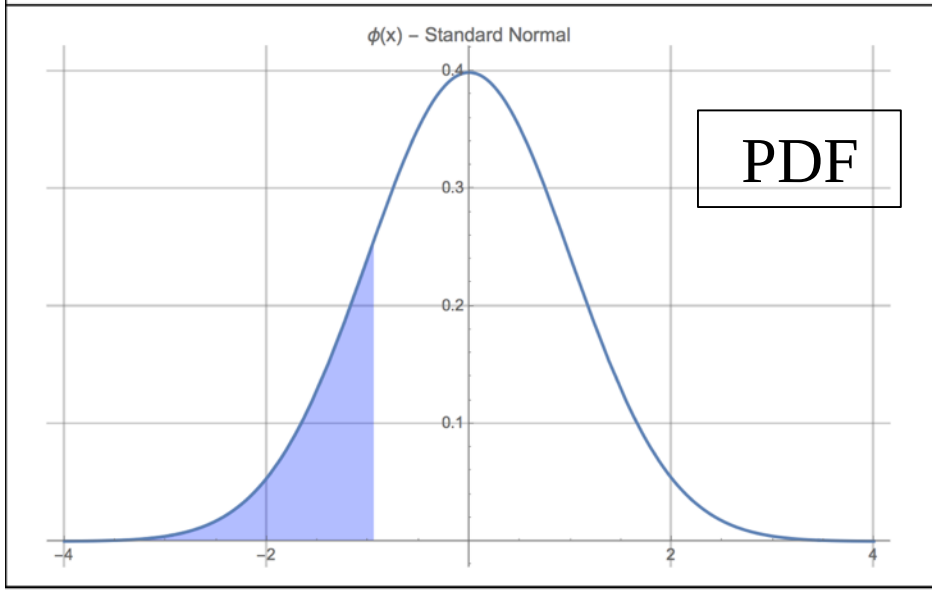
$$f_X(x) = \frac{\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma}$$



Standard Deviations of a Standard Normal PDF

Normal Distribution





- The slope of the CDF is the value of the PDF.
- The integral under the PDF is the value of the CDF

Normal – Probability Calculations

- The CDF for any normal distribution can be obtained from the standard normal:

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$x = \sigma\,\Phi^{-1}(F) + \mu$$

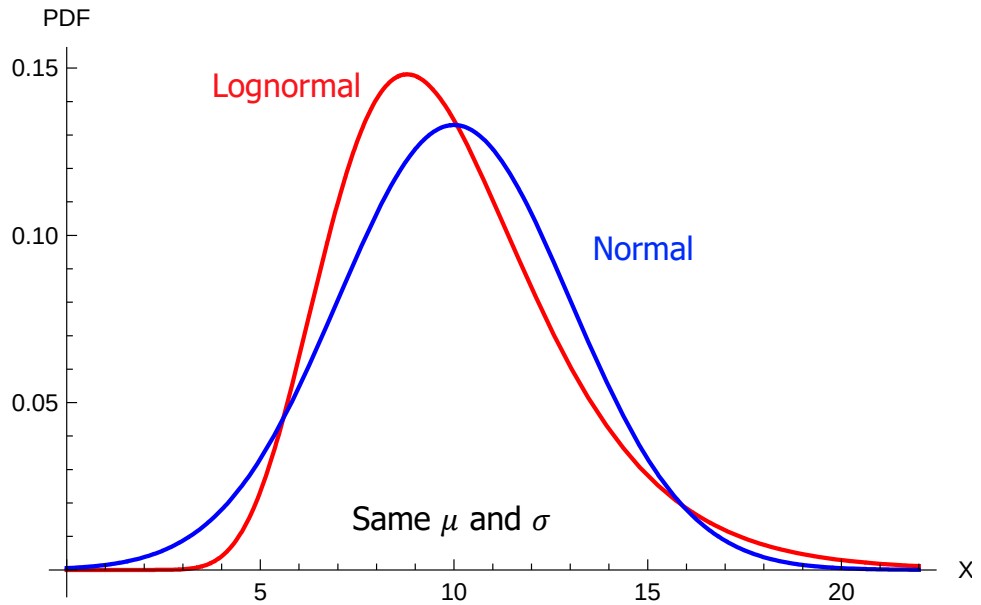
Φ represents the CDF of the normal distribution

Software	CDF function	inverse CDF function
Excel	norm.dist(x, μ, σ, TRUE)	norm.inv(x, μ, σ)
Python	scipy.stats.norm.cdf(x, μ, σ)	scipy.stats.norm.ppf(x, μ, σ)
Matlab	normcdf(x, μ, σ)	norminv(x, μ, σ)
R	stats.pnorm(x, μ, σ)	stats.qnorm(x, μ, σ)

LogNormal Distribution

- Closely related to Normal distribution. If ln(X) is Normal, X is Lognormal.
- Cannot go negative. Unsymmetric - Long right tail.

$$f_X(x) = \frac{1}{\sqrt{2\pi}(\xi x)} \exp\left(-\frac{1}{2}\left(\frac{\ln(x)-\lambda}{\xi}\right)^2\right)$$



LogNormal Distribution

- Defined by 2 parameters: mean (λ) and stdev (ξ) of ln(x).
- λ and ξ are related to μ and λ by

$$\xi = \sqrt{1 + \left(\frac{\sigma}{\mu}\right)^2} \; ; \qquad \lambda = \ln(\mu) - \frac{1}{2}\xi^2$$

Caution! Most codes (including Excel) want the mean and stdev of ln(x), That is, they want λ and ξ.

LogNormal – Probability Calculations

- The CDF of the LogNormal distribution can be computed from the Normal distribution:

$$F_X(x) = \Phi\left(\frac{\ln(x) - \lambda}{\xi}\right)$$

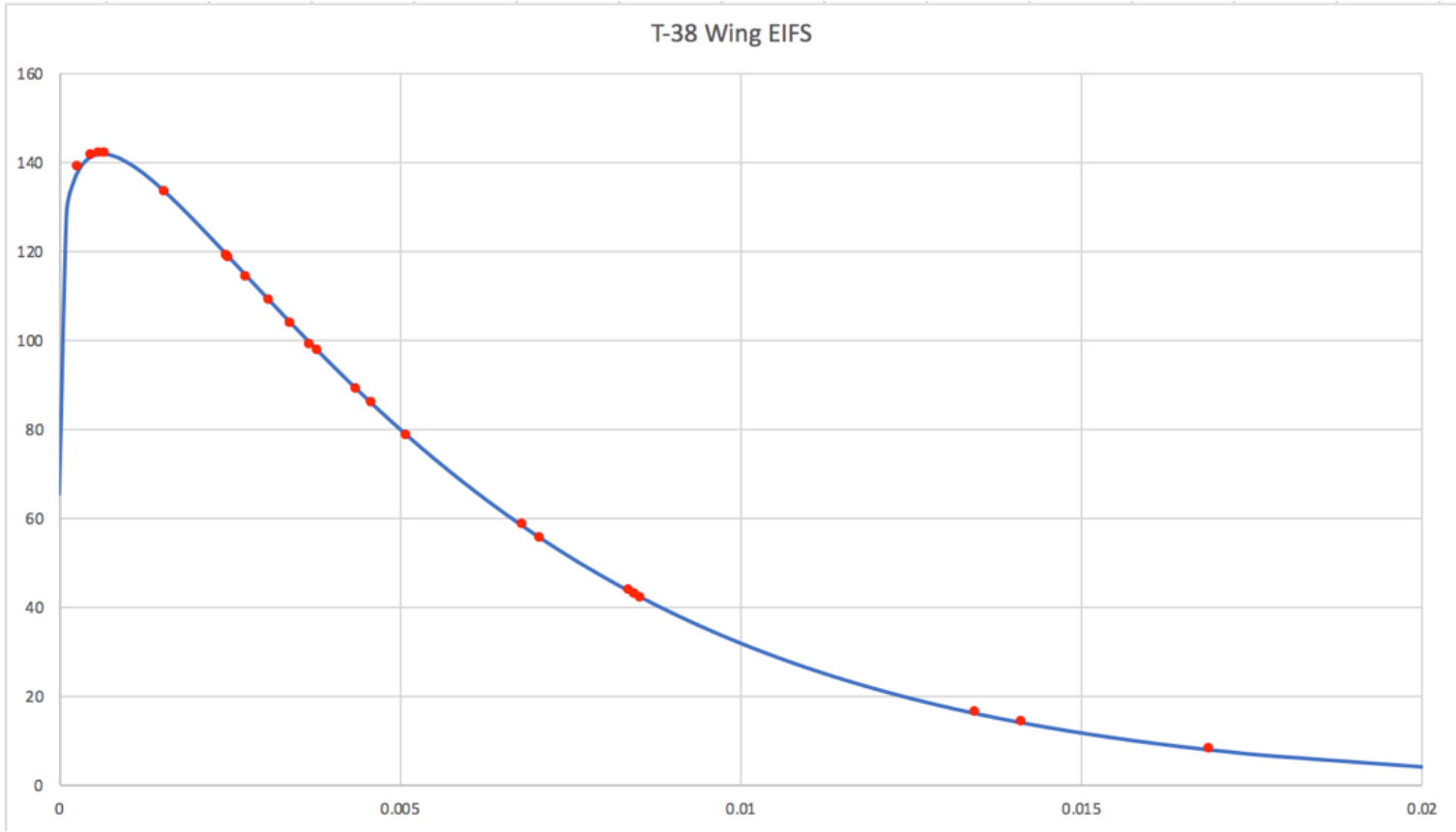
$$x = \exp(\xi\ \Phi^{-1}(F) + \lambda)$$

Φ represents the CDF of the normal distribution

Software	CDF function	inverse CDF function
Excel	lognorm.dist(x, λ, ξ, TRUE)	lognorm.inv(x, λ, ξ)
Python	scipy.stats.lognorm.cdf(x, ξ, exp(λ))	scipy.stats.lognorm.ppf(x, ξ, exp(λ))
Matlab	logncdf(x, λ, ξ)	logninv(x, λ, ξ)
R	stats.plnorm(x, λ, ξ)	stats.qlnorm(x, λ, ξ)

Weibull Distribution

- Very flexible distribution. Often used to model equivalent initial flaw size, fatigue life (time to failure).



Weibull Distribution

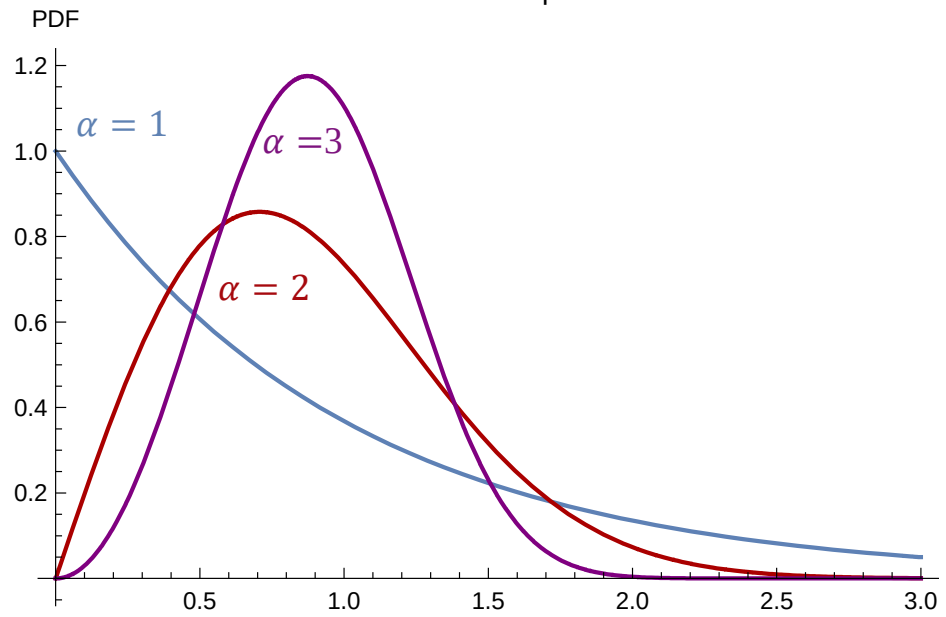
- Two parameters: shape (α) and scale (β)

$$f_X(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)$$

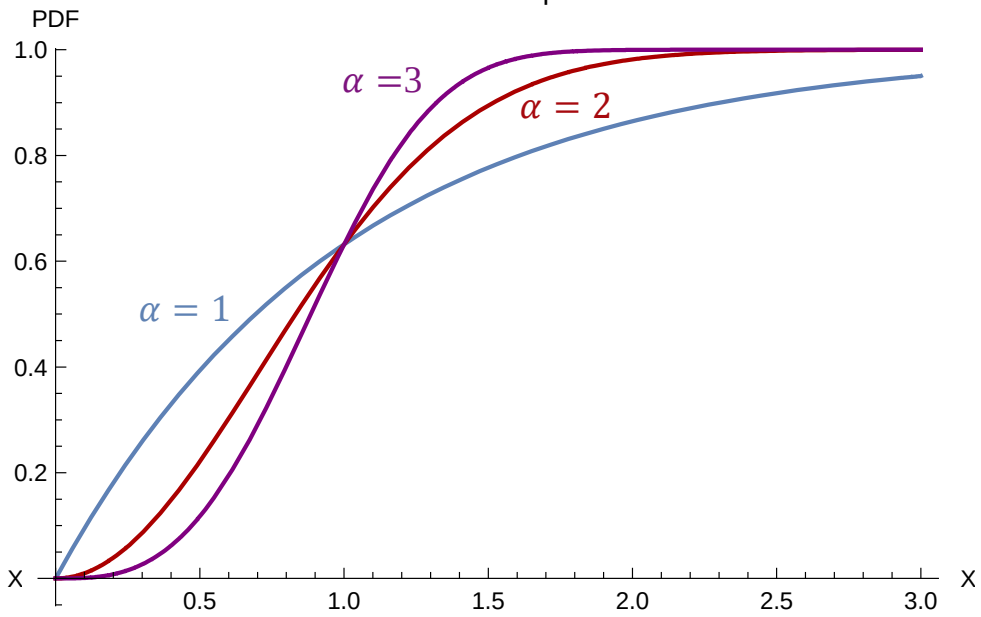
- Calculating the Mean and Standard Deviation of Weibull distributions requires evaluating the gamma function, Γ

$$\mu = \beta\,\Gamma(1+1/\alpha); \qquad \sigma = \beta\sqrt{\Gamma(1+2/\alpha) - \left(\Gamma(1+1/\alpha)\right)^2}$$

Weibull PDF plots



Weibull CDF plots



Weibull – Probability Calculations

- Excel does not have an inverse Weibull CDF, but it is simple to construct analytically, where *F* is a probability

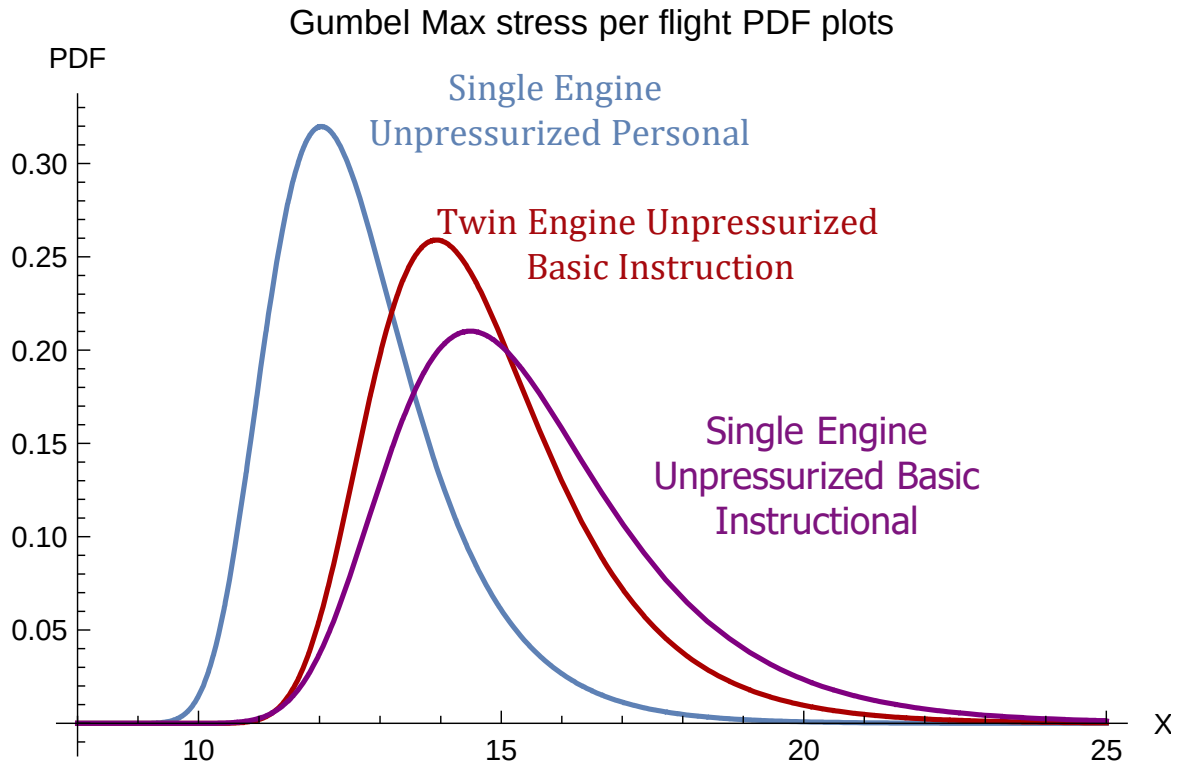
$$F_X(x) = 1 - \exp(-(x/\beta)^{\alpha})$$

$$x = \beta(-\ln(1-F))^{1/\alpha}$$

Software	CDF function	inverse CDF function
Excel	weibull.dist(x, α, β, TRUE)	must be manually defined (see above)
Python	scipy.stats.weibull_min.cdf(x, α, β)	scipy.stats.weibull_min.ppf(x, α, β)
Matlab	wblcdf(x, β, α)	wblinv(x, β, α)
R	stats.pweibull(x, α, β)	stats.qweibull(x, α, β)

Gumbel Distribution

- An extreme value distribution often used to model loads. There are Gumbel minimum and maximum distributions. Gumbel max is used for loads. (Also called *Extreme Value Distribution Type-I*.)



Excel Implementation

- Excel does NOT have the Gumbel (EVD) distribution.
- This can be easily programmed.

$$F(x) = e^{-e^{-\frac{x-a}{b}}}$$ Evaluate probabilities

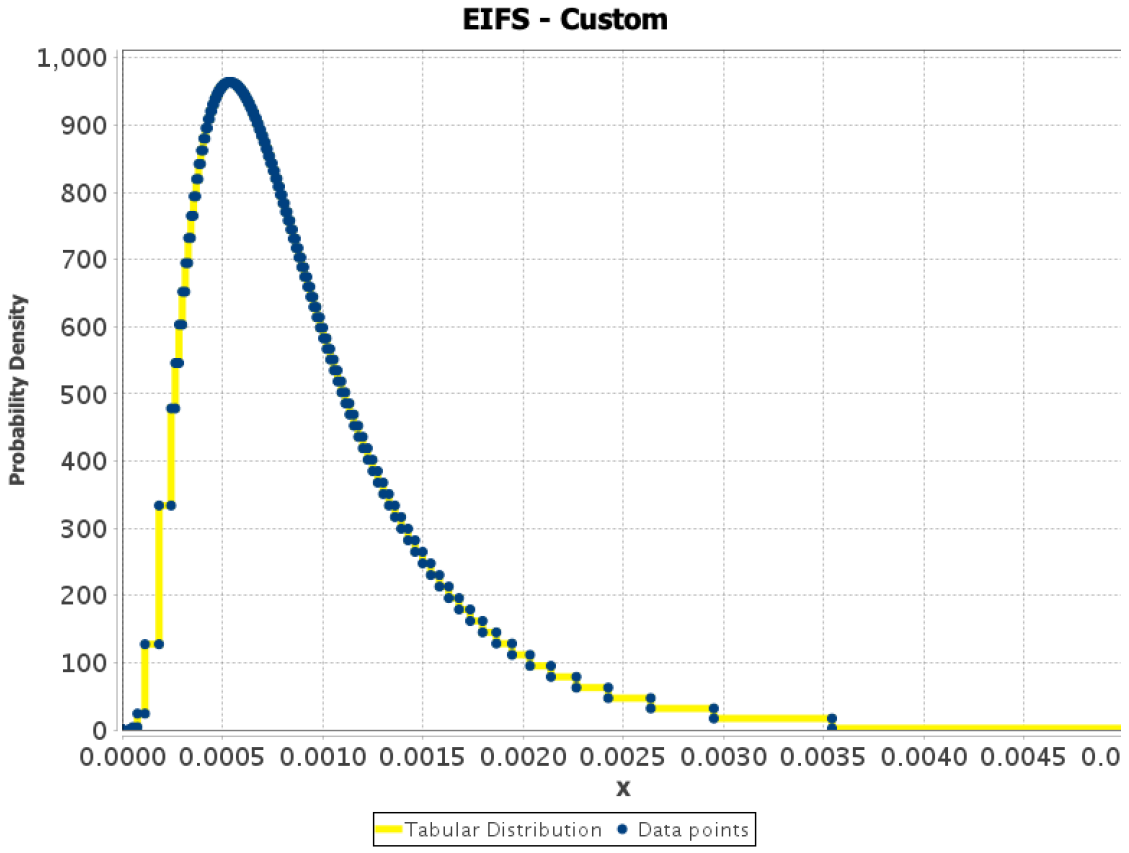
$$x = a + b(-\ln(F(-\ln)))$$ Generate samples

(where F is a probability between 0 and 1)

Software	CDF function	inverse CDF function
Excel	must be manually defined	must be manually defined
Python	scipy.stats.genextreme.cdf(x, 0, a, b)	scipy.stats.genextreme.ppf(x, 0, a, b)
Matlab	gevcdf(x, 0, b, a)	gevinv(x, 0, b, a)
R	evd.pgev(x, a, b, 0)	evd.qgev(x, a, b, 0)

Tabular Distribution

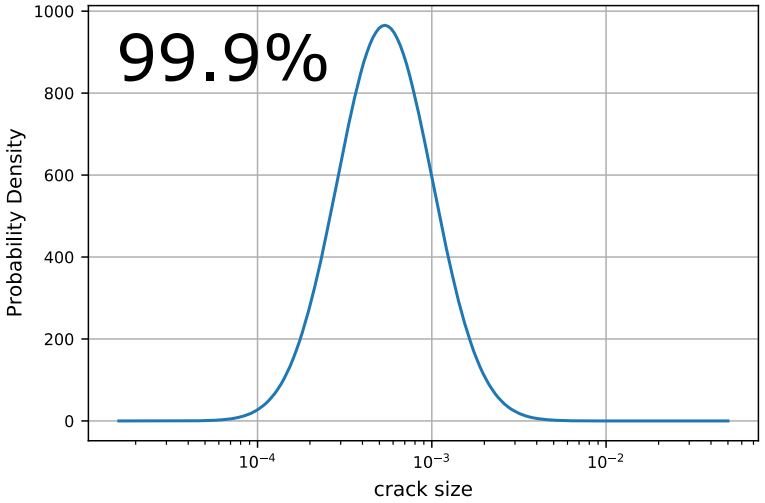
! crksz	CDF
0.003093	0.000000000
0.003094	0.000000001
0.003094	0.00000001
0.003095	0.0000001
0.002506	0.000001
0.001629	0.00001
0.000537	0.0001
0.000184	0.001
0.000185	0.01
0.000245	0.03
⋮	⋮
0.002956	0.98
0.003550	0.99
0.025939	0.999
0.048000	0.9999
0.049800	0.99999
0.049980	0.999999
0.049998	0.9999999
0.050000	0.99999999
0.050000	0.999999999
0.050000	1.000000000



- Table defining the value, x, and the CDF, probability that $x \leq \text{value}$
- Useful in software where it is not practical to implement large set of seldom used distributions

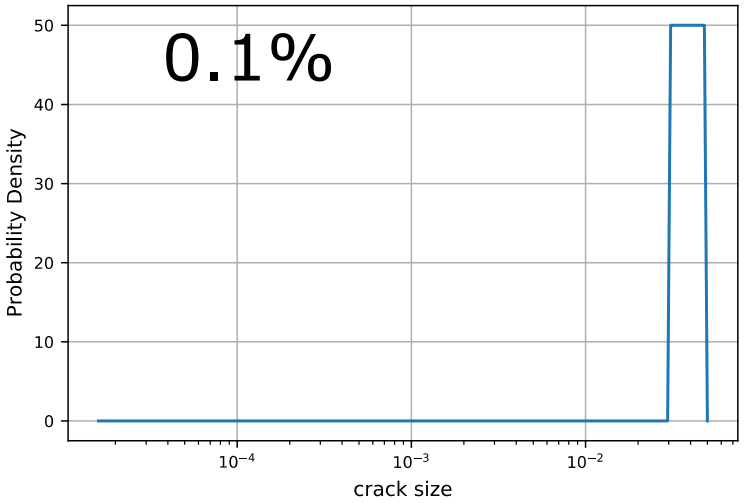
Tabular Distribution

Lognormal PDF



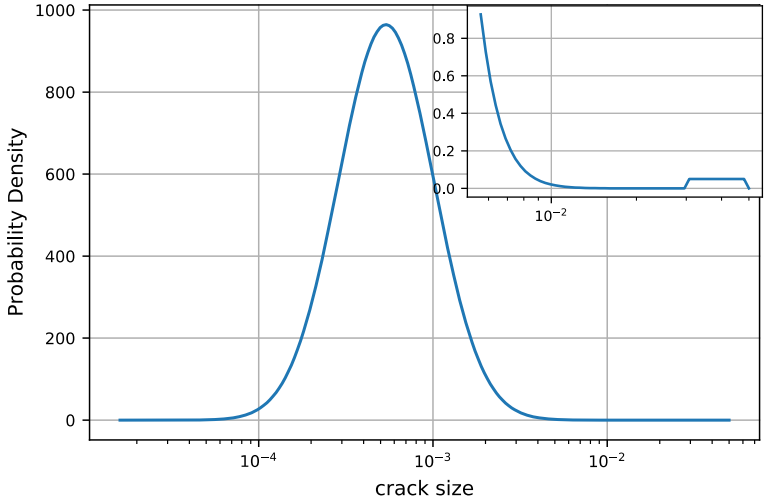
+

Uniform PDF



=

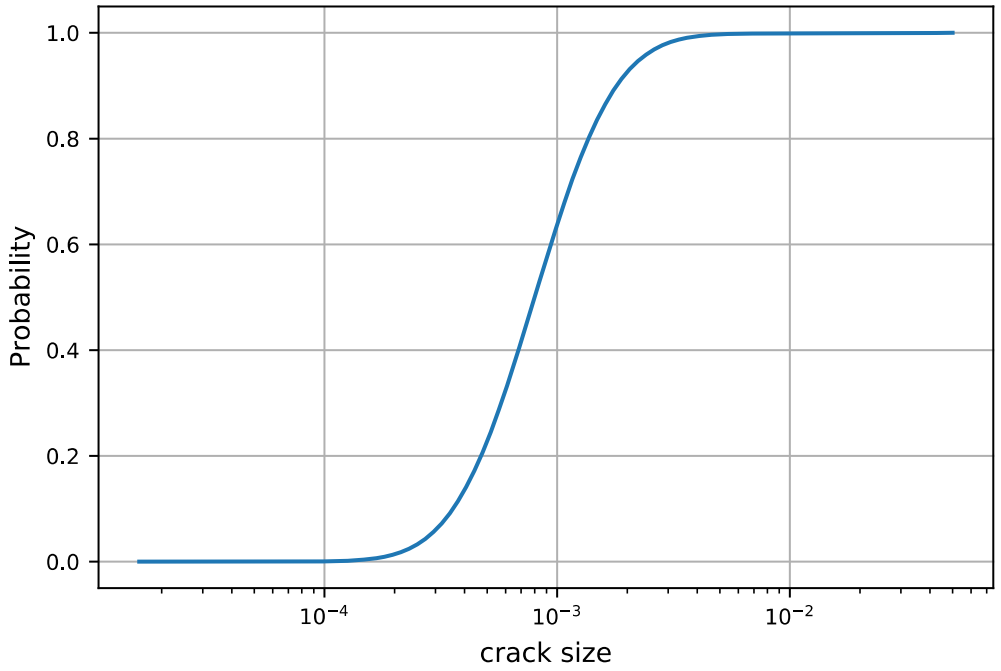
User Defined PDF



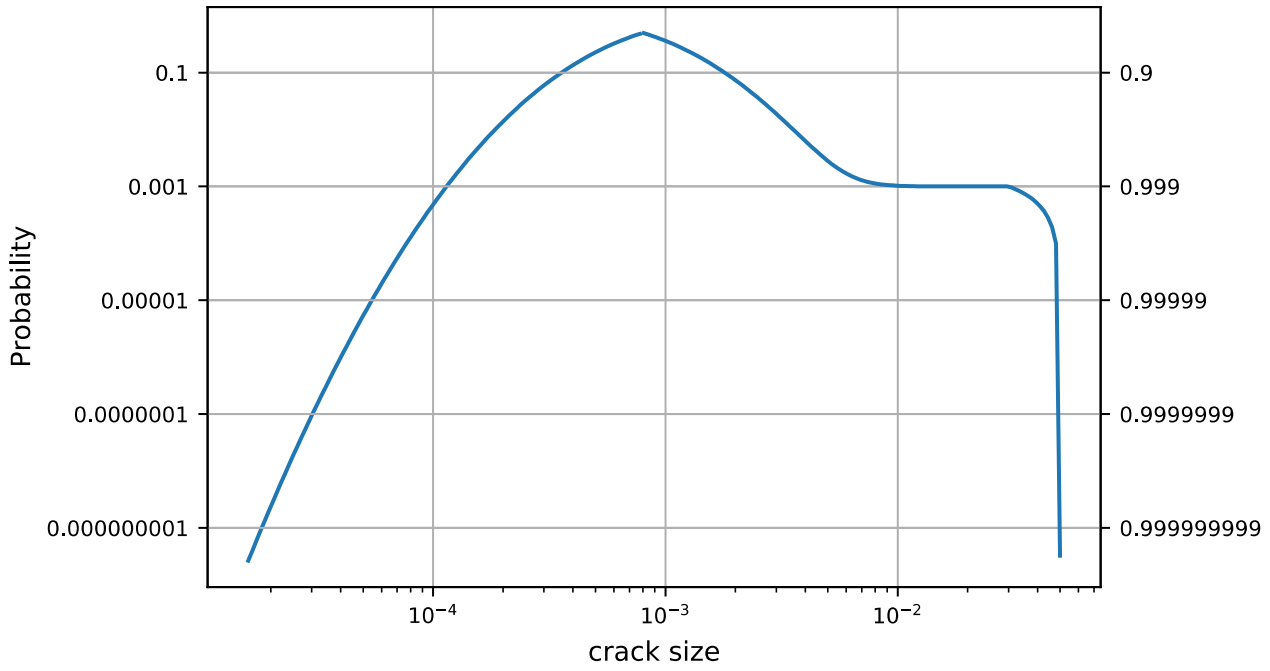
- Example: lognormal + uniform to give 0.001 probability of 0.05” rogue flaw
- The rogue flaw probability density is not visible unless one zooms in on the right tail
- What does the CDF look like?

Tabular Distribution

CDF Plot



Mountain Plot



- Example: lognormal + uniform to give 0.001 probability of 0.05” rogue flaw
- The 1/1000 rogue flaw bump is not visible on a CDF plot
- … but can be seen clearly by plotting exceedance on a logarithmic scale as shown by the mountain plot
- When given as a tabular distribution for EIFS, the POF will be highly dependent on the random samples generated from

Questions

Supplemental Material

Normal Distribution

NORM.DIST function

Excel for Office 365, Excel for Office 365 for Mac, Excel 2019, Excel 2016, More...

Evaluate CDF & PDF

Returns the normal distribution for the specified mean and standard deviation. This function has a very wide range of applications in statistics, including hypothesis testing.

Syntax

NORM.DIST(x,mean,standard_dev,cumulative)

The NORM.DIST function syntax has the following arguments:

- **X** Required. The value for which you want the distribution.
- **Mean** Required. The arithmetic mean of the distribution.
- **Standard_dev** Required. The standard deviation of the distribution.
- **Cumulative** Required. A logical value that determines the form of the function. If cumulative is T NORM.DIST returns the cumulative distribution function; if FALSE, it returns the probability density function.

Remarks

- If mean or standard_dev is nonnumeric, NORM.DIST returns the #VALUE! error value.
- If standard_dev ≤ 0, NORM.DIST returns the #NUM! error value.
- If mean = 0, standard_dev = 1, and cumulative = TRUE, NORM.DIST returns the standard normal distribution, NORM.S.DIST.
- The equation for the normal density function (cumulative = FALSE) is:

$$f(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

- When cumulative = TRUE, the formula is the integral from negative infinity to x of the given formula.

Example

Copy the example data in the following table, and paste it in cell A1 of a new Excel worksheet. For formulas to show results, select them, press F2, and then press Enter. If you need to, you can adjust the column widths to see all the data.

Data	Description	
42	Value for which you want the distribution	
40	Arithmetic mean of the distribution	
1.5	Standard deviation of the distribution	
Formula	**Description**	**Result**
=NORM.DIST(A2,A3,A4,TRUE)	Cumulative distribution function for the terms above	0.9087888
=NORM.DIST(A2,A3,A4,FALSE)	Probability mass function for the terms above	0.10934

Normal Distribution

NORM.INV function

Excel for Office 365, Excel for Office 365 for Mac, Excel 2019, Excel 2016, More...

Inverse fn. – generate samples
X=norm.inv(p, μ, σ)

Returns the inverse of the normal cumulative distribution for the specified mean and standard deviation.

Syntax

NORM.INV(probability,mean,standard_dev)

The NORM.INV function syntax has the following arguments:

- **Probability** Required. A probability corresponding to the normal distribution.
- **Mean** Required. The arithmetic mean of the distribution.
- **Standard_dev** Required. The standard deviation of the distribution.

Remarks

- If any argument is nonnumeric, NORM.INV returns the #VALUE! error value.
- If probability <= 0 or if probability >= 1, NORM.INV returns the #NUM! error value.
- If standard_dev ≤ 0, NORM.INV returns the #NUM! error value.
- If mean = 0 and standard_dev = 1, NORM.INV uses the standard normal distribution (see NORMS.INV).

Given a value for probability, NORM.INV seeks that value x such that NORM.DIST(x, mean, standard_dev, TRUE) = probability. Thus, precision of NORM.INV depends on precision of NORM.DIST.

Example

Copy the example data in the following table, and paste it in cell A1 of a new Excel worksheet. For formulas to show results, select them, press F2, and then press Enter. If you need to, you can adjust the column widths to see all the data.

Data	Description	
0.908789	Probability corresponding to the normal distribution	
40	Arithmetic mean of the distribution	
1.5	Standard deviation of the distribution	
Formula	**Description**	**Result**
=NORM.INV (A2,A3,A4)	Inverse of the normal cumulative distribution for the terms above (42)	42.000002

LogNormal – Excel Options

LOGNORM.DIST function

Returns the lognormal distribution of **x**, where ln(x) is normally distributed with parameters **mean** and **standard_dev**
Use this function to analyze data that has been logarithmically transformed.

Syntax

LOGNORM.DIST(x,mean,standard_dev,cumulative)

Argument	Description	Remarks
x	The value at which to evaluate the function.	• If this argument is equal to or less than zero, this function returns the #NUM! error value. • If this argument is nonnumeric, this function returns the #VALUE! error value.
mean	A mean of ln(x).	• If this argument is nonnumeric, this function returns the #VALUE! error value.
standard_dev	The standard deviation of ln(x).	• If this argument is equal to or less than zero, this function returns the #NUM! error value. • If this argument is nonnumeric, this function returns the #VALUE! error value.
cumulative	A logical value that determines the form of the function.	• If this argument is TRUE, this function returns the cumulative distribution function; if FALSE, it returns the probability density function. • If this argument is nonnumeric, this function returns the #VALUE! error value.

Evaluate CDF & PDF

General remarks

The equation for the lognormal cumulative distribution function is:

$LOGNORM.DIST(x,\mu,\sigma) = NORM.S.DIST(1n(x)-\mu / \sigma)$

Example

To make the following example easier to understand, you can copy the data to a blank sheet and then enter the function underneath the data. Do not select the row or column headings (1, 2, 3... A, B, C...) when you copy the sample data to a blank sheet.

	A	B
1	Data	Description
2	4	Value at which to evaluate the function (x)
3	3.5	Mean of ln(x)
4	1.2	Standard deviation of ln(x)
	Formula	Description (Result)
	=LOGNORM.DIST(A2,A3,A4,TRUE)	Cumulative lognormal distribution at 4 with the terms in the table (0.039084)
	=LOGNORM.DIST(A2,A3,A4,FALSE)	Probability lognormal distribution at 4 with the terms in the table (0.017618)

LOGNORM.INV function

Returns the inverse of the lognormal cumulative distribution function of **x**, where ln(x) is normally distributed with parameters **mean** and **standard_dev**. If p = **LOGNORM.DIST**(x,...) then **LOGNORM.INV**(p,...) = **x**.

Use the lognormal distribution to analyze logarithmically transformed data.

Syntax

LOGNORM.INV(probability,mean,standard_dev)

Argument	Description	Remarks
probability	A probability associated with the lognormal distribution.	• If this argument is equal to or less than zero, this function returns the #NUM! error value. • If this argument is equal to or greater than 1, this function returns the #NUM! error value. • If this argument is nonnumeric, this function returns the #VALUE! error value.
mean	A mean of ln(x).	• If this argument is nonnumeric, this function returns the #VALUE! error value.
standard_dev	The standard deviation of ln(x).	• If this argument is equal to or less than zero, this function returns the #NUM! error value. • If this argument is nonnumeric, this function returns the #VALUE! error value.

Inverse fn. – generate samples

Example

To make the following example easier to understand, you can copy the data to a blank sheet and then enter the function underneath the data. Do not select the row or column headings (1, 2, 3... A, B, C...) when you copy the sample data to a blank sheet.

	A	B
1	Data	Description
2	0.039084	Probability that is associated with the lognormal distribution
3	3.5	Mean of ln(x)
4	1.2	Standard deviation of ln(x)
	Formula	Description (Result)
	=LOGNORM.INV(A2,A3,A4)	Inverse of the lognormal cumulative distribution function for the terms in the table (4.000025)

LogNormal – Excel Options

LOGNORM.DIST function

Excel for Office 365, Excel for Office 365 for Mac, Excel 2019, Excel 2016, More...

Returns the lognormal distribution of x, where ln(x) is normally distributed with parameters Mean and Standard_dev.

Use this function to analyze data that has been logarithmically transformed.

Syntax

LOGNORM.DIST(x,mean,standard_dev,cumulative)

The LOGNORM.DIST function syntax has the following arguments:

- **X** Required. The value at which to evaluate the function.
- **Mean** Required. The mean of ln(x).
- **Standard_dev** Required. The standard deviation of ln(x).
- **Cumulative** Required. A logical value that determines the form of the function. If cumulative is TRUE, LOGNORM.DIST returns the cumulative distribution function; if FALSE, it returns the probability density function.

Remarks

- If any argument is nonnumeric, LOGNORM.DIST returns the #VALUE! error value.
- If x ≤ 0 or if standard_dev ≤ 0, LOGNORM.DIST returns the #NUM! error value.
- The equation for the lognormal cumulative distribution function is:

  LOGNORM.DIST(x,μ,o) = NORM.S.DIST($1n(x)-\mu / o$)

Evaluate CDF & PDF

Example

Copy the example data in the following table, and paste it in cell A1 of a new Excel worksheet. For formulas to show results, select them, press F2, and then press Enter. If you need to, you can adjust the column widths to see all the data.

Data	Description	
4	Value at which to evaluate the function (x)	
3.5	Mean of ln(x)	
1.2	Standard deviation of ln(x)	
Formula	**Description**	**Result**
=LOGNORM.DIST (A2,A3,A4,TRUE)	Cumulative lognormal distribution at 4, using the arguments in A2:A4.	0.0390836
=LOGNORM.DIST (A2,A3,A4,FALSE)	Probability lognormal distribution at 4, using the same arguments.	0.0176176

LogNormal – Excel Options

LOGNORM.INV function

Excel for Office 365, Excel for Office 365 for Mac, Excel 2019, Excel 2016, More...

This article describes the formula syntax and usage of the **LOGNORM.INV** function in Microsoft Excel.

Inverse fn. – generate samples
x=lognorm.inv(p, λ, ξ)

Description

Returns the inverse of the lognormal cumulative distribution function of x, where ln(x) is normally distributed with parameters Mean and Standard_dev. If p = LOGNORM.DIST(x,...) then LOGNORM.INV(p,...) = x.

Use the lognormal distribution to analyze logarithmically transformed data.

Syntax

LOGNORM.INV(probability, mean, standard_dev)

The LOGNORM.INV function syntax has the following arguments:

- **Probability** Required. A probability associated with the lognormal distribution.
- **Mean** Required. The mean of ln(x).
- **Standard_dev** Required. The standard deviation of ln(x).

Remarks

- If any argument is nonnumeric, LOGNORM.INV returns the #VALUE! error value.
- If probability <= 0 or probability >= 1, LOGNORM.INV returns the #NUM! error value.
- If standard_dev <= 0, LOGNORM.INV returns the #NUM! error value.

Example

Copy the example data in the following table, and paste it in cell A1 of a new Excel worksheet. For formulas to show results, select them, press F2, and then press Enter. If you need to, you can adjust the column widths to see all the data.

Data	Description	
0.039084	Probability associated with the lognormal distribution	
3.5	Mean of ln(x)	
1.2	Standard deviation of ln(x)	
Formula	**Description**	**Result**
=LOGNORM.INV(A2, A3, A4)	Inverse of the lognormal cumulative distribution function for the terms above	4.0000252

WEIBULL.DIST function

Returns the Weibull distribution. Use this distribution in reliability analysis, such as calculating a device's mean time to failure.

Syntax

WEIBULL.DIST(x,alpha,beta,cumulative)

Argument	Description	Remarks
x	The value at which to evaluate the function.	• If this argument is less than zero, this function returns the #NUM! error value. • If this argument is nonnumeric, this function returns the #VALUE! error value.
alpha	A parameter of the distribution.	• If this argument is nonnumeric, this function returns the #VALUE! error value. • If this argument is equal to or less than zero, this function returns the #NUM! error value.
beta	A parameter of the distribution.	• If this argument is nonnumeric, this function returns the #VALUE! error value. • If this argument is equal to or less than zero, this function returns the #NUM! error value.
cumulative	Determines the form of the function.	• None.

General remarks

- The equation for the Weibull cumulative distribution function is:
  $F(x;\alpha,\beta) = 1 - e^{-(x/\beta)^{\alpha}}$
- The equation for the Weibull probability density function is:
  $f(x;\alpha,\beta) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-(x/\beta)^{\alpha}}$
- When **alpha** = 1, **WEIBULL.DIST** returns the exponential distribution with:
  $\lambda = \frac{1}{\beta}$

Evaluate CDF & PDF

Example

To make the following example easier to understand, you can copy the data to a blank sheet and then enter the function underneath the data. Do not select the row or column headings (1, 2, 3... A, B, C...) when you copy the sample data to a blank sheet.

	A	B
1	Data	Description
2	105	Value at which to evaluate the function
3	20	Alpha parameter to the distribution
4	100	Beta parameter to the distribution
	Formula	Description (Result)
	=WEIBULL.DIST(A2,A3,A4,TRUE)	Weibull cumulative distribution function for the terms in the table (0.929581)
	=WEIBULL.DIST(A2,A3,A4,FALSE)	Weibull probability density function for the terms in the table (0.035589)

Excel does not have an inverse Weibull CDF, but it is simple to construct analytically, where *F* is a probability

$$x = \beta(-\ln(1-F))^{1/\alpha}$$