Contents

Central Tendency

$$\boldsymbol{mean} = \frac{\sum x}{n} = \frac{Sum\ of\ all\ observations}{Totat\ number\ of\ observations}$$

$$\boldsymbol{sample\ mean} = \bar{\boldsymbol{x}} = \frac{\sum x}{n} = \frac{Sum\ of\ all\ samples}{Totat\ number\ of\ samples}$$

$$\boldsymbol{population\ mean} = \boldsymbol{\mu} = \frac{\sum X}{N} = \frac{Sum\ of\ all\ values\ within\ the\ populations}{Totat\ number\ of\ values\ within\ the\ population}$$

$\boldsymbol{Median} = \boldsymbol{Middle\ Value} =$ ~~185K~~, ~~197K~~, **230K**, ~~252K~~, ~~1.4M~~

$\boldsymbol{Median\ of\ Even\ Numbers} =$ ~~85K~~, **197K**, **230K**, ~~252K~~

$$\boldsymbol{Median} = \boldsymbol{M} = \boldsymbol{Mean\ of}\ \mathbf{197}\boldsymbol{K}\ \&\ 230\boldsymbol{K} = \frac{\sum \boldsymbol{x}}{\boldsymbol{n}} = \frac{\mathbf{197}\boldsymbol{K} + \mathbf{230}\boldsymbol{K}}{\mathbf{2}} = \frac{\mathbf{427}\boldsymbol{K}}{\mathbf{2}} = \mathbf{213.5}\boldsymbol{K}$$

Mode *is defined as the most frequently occurring value in a data set.*

Variance

$$Sample\ Variance = s^2 = \frac{\sum (x - \bar{x})^2}{n-1} = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$$

$$Population\ Variance = \sigma^2 = \frac{\sum (x - \bar{\mu})^2}{N}$$

$$Sample\ Standard\ Deviation:\ \ s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

$$Population\ Standard\ Deviation:\ \ s = \sqrt{\frac{\sum (x - \bar{\mu})^2}{N}}$$

$$Range = R = Max(x) - Min(x)$$

Probability

$$The\ \boldsymbol{Probability\ of\ A\ or\ B} = P(A\ U\ B) = P(A) + \ P(B) - \ P(A \cap B)$$

$$\text{The } \textbf{Intersection of A \& B } = \text{A and } \ \text{B} = \text{P}(\text{A} \cap \text{B})$$

$$The\ \boldsymbol{Probability\ of\ A}° = P(A°) = 1 - P(A)$$

$$For\ Mutually\ Exclusive\ Events: \ \ P(A \cap B) = 0$$

$$\boldsymbol{Probabiliy\ of\ A\ given\ B} = P(A\ |B) = \frac{P(A \cap B)}{P(B)} = \frac{The\ Intersection\ of\ A\ \&\ B}{The\ Probability\ of\ B}$$

$$\boldsymbol{For\ Independent\ Events}: \ \ \ P(A\ and\ B) = \ P((A \cap B) = P(A) * P(B)$$

$$The\ \boldsymbol{Union}\ of\ A\ \&\ B = \ A\ or\ B = (A\ U\ B)$$

$$The\ \boldsymbol{Intersection}\ of\ A\ and\ B = \ A\ \&\ B = (A \cap B)$$

$$The\ \boldsymbol{Multiplication}\ Rule\ for\ Dependent\ Events: \ \ \ P(A\ and\ B) = \ P(A \cap B) = P(A|B) * P(B)$$

$$The\ \boldsymbol{Addition}\ Rule\ for\ Two\ Events = P(A\ U\ B) = P(A) + \ P(B) - \ P(A \cap B)$$

$$For\ Mutually\ Exclusive\ Events: \ \ P(A \cap B) = 0$$

$$The\ \boldsymbol{Addition}\ Rule\ for\ Mutually\ Excluive\ Events = P(A\ U\ B) = P(A) + \ P(B)$$

Probability Distributions

$$\boldsymbol{Normal}\ Distribution\ Z\ Transformation: Z = \frac{X - \mu}{\sigma}$$

$$\boldsymbol{Uniform}\ Distribution\ Mean\ Value: \mu = \frac{a+b}{2}$$

$\boldsymbol{Exponential}\ Distribution\ Probability: P(X = x):\ f(x) = xe^{-\lambda x}$

$Exponential\ Distribution\ Cumulative\ Probability: P(X > x):\ f(x) = e^{-\lambda x}$

$Exponential\ Distribution\ Cumulative\ Probability: P(X < x):\ F(x) = 1 - e^{-\lambda x}$

$Exponential\ Distribution\ Mean\ Value = \theta = \frac{1}{\lambda}\quad where\quad \lambda = Occurrence\ Rate$

$$\boldsymbol{Weibull}\ Distribution\ Reliability:\quad R(t) = e^{-\left(\frac{t}{\theta}\right)^{\beta}}$$

$$\boldsymbol{Student\ T}\ Distribution: t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$\boldsymbol{F\ Distribution}: F = \frac{(S_1)^2}{(S_2)^2}$$

Binomial *Distribution: Mean (Expected Value):* $\mu = n*p$

Binomial Distribution: Standard Deviation: $\sigma = \sqrt{n * p(1-p)}$

$Binomial\ Distribution\ Probability: P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$

$$where\ \binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

Poisson *Distribution: Mean (Expected Value):* $\mu = \lambda$

Poisson Distribution: Standard Deviation: $\sigma = \sqrt{\lambda}$

$$Poisson\ Distribution\ Probability: f(x) = P(X = x) = \frac{e^{-\lambda} * \lambda^x}{x!}$$

$\boldsymbol{Hypergeometric}\ Distribution\ Probabibility: f(x) = \frac{\binom{A}{x} * \binom{N-A}{n-x}}{\binom{N}{n}}$

Point Estimates & Confidence Intervals

$$Interval\ Estimate\ of\ Population\ Mean\ (known\ variance)\ :\bar{x}\ \pm Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$

$$Interval\ Estimate\ of\ Population\ Mean\ (unknown\ variance):\ \bar{x}\ \pm t_{\frac{\alpha}{2}} * \frac{s}{\sqrt{n}}$$

$$Confidence\ Interval\ for\ Variance: \quad \frac{(n-1)s^2}{X^2_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{X^2_{\alpha/2}}$$

$$Confidence\ Interval\ for\ Standard\ Deviation: \quad \sqrt{\frac{(n-1)s^2}{X^2_{1-\alpha/2}}} < \sigma < \sqrt{\frac{(n-1)s^2}{X^2_{\alpha/2}}}$$

$$Confidence\ Interval\ (Proportion): p\ \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{p*(1-p)}{n}}$$

$$Variance\ of\ sample\ mean: V(\bar{x}) = \frac{\sigma^2}{n}$$

$$Standard\ Error\ of\ The\ Sample\ Mean: S.E. = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

Hypothesis Testing

$$Population\ Mean:\ z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$Population\ Mean:\ t - statistic = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$Population\ Variance -\ Chi\ Squared\ Test\ Statistic: X^2 = \frac{(N-1)s^2}{\sigma^2}$$

$$Two\ Population\ Variances -\ F - Test\ Statistic: F = \frac{s_1^2}{s_2^2}$$

$$Population\ Proportion: Z_O = \frac{\hat{p} - p_O}{\sqrt{\frac{p_O(1-p_O)}{n}}}$$

Type I & II Error in Hypothesis Testing

		The Truth	
		H_0 is True	**H_0 is False**
The Outcome of the Hypothesis Test	**Fail to Reject H_0**	**Correct Decision**	INCORRECT DECISION **(Type II Error)** **Beta (β) Risk**
	Reject H_0	INCORRECT DECISION **(Type I Error)** **Alpha (α) risk**	**Correct Decision** ***Power (1 - β)***

Goodness of Fit Testing

H_0: Sample Data Fits a Specific Distribution

H_a: Sample Data Does Not Fit that Specific Distribution

$$X^2 = \sum_{i}^{i=k} \frac{(O_i - E_i)^2}{E_i}$$

O_i = Observed Value from the Sample Data

E_i = Expected Value from the Population Distribution

ANOVA

Variation Source	Sum of Squares (SS)	Degrees of freedom (DF)	Mean Squares (MS)	F-Value
Treatment (Between)	Sum of Squares of the Treatment **(SS_t)**	DF of the Treatment **(DF_t)**	Mean Squares of the Treatment **(MST)**	**= MST / MSE**
Error (Within)	Sum of Squares of the Error **(SS_e)**	DF of the Error **(DF_e)**	Mean Squares of the Error **(MSE)**	
Total	Total Sum of Squares **(SS_{total})**	Total DF **(DF_{Total})**		

$$Total\ Sum\ of\ Squares: SS_{total} = \sum (X_i - GM)^2$$

$$Sum\ of\ Squares\ of\ the\ Error: SS_e = \sum (X_i - \bar{X})^2$$

$$Sum\ of\ Squares\ of\ the\ Treatment = SS_t = \sum n(\bar{X}_i - GM)^2$$

$$SS_{Total} = SS_e + SS_T$$

$$\text{Total Degrees of Freedom} = DF_{total} = DF_{error} + DF_{treatment}$$

$$DF_{total} = N - 1$$

$$DF_{treatment} = a - 1$$

$$DF_{error} = (N-1) - (a-1) = a(n-1)$$

$$Mean\ Square\ of\ the\ Error = \frac{Sum\ of\ Squares\ of\ the\ Error\ (SS_e)}{Degrees\ of\ Freedom\ of\ the\ Error (DF_{error})}$$

$$Mean\ Square\ of\ the\ Treatment = \frac{Sum\ of\ Squares\ of\ the\ Treatment (SS_t)}{Degrees\ of\ Freedom\ of\ the\ Treatment (DF_{treatment})}$$

$$F - statistic = \frac{MST}{MSE} = \frac{Mean\ Square\ of\ the\ Treatment}{Mean\ Square\ of\ the\ Error}$$

Contingency Tables

A **contingency table** is used to determine if the **two factors** being studied are statistically **independent** of each other, or if they share some level of **dependence**.

The contingency table uses the **chi-squared distribution (X^2) and statistic** to accept/reject our null hypothesis.

This chi-squared statistics is calculated as a comparison of the **observed values** of your data against their **expected values**.

$$X^2 = \sum_{i}^{i=k} \frac{(O_i - E_i)^2}{E_i}$$

Where

O_i = Observed Value from the sample data

E_i = Expected Value based on the assumption that the two factors are independent

The Observed values are the data you collect during your sampling.

The expected values can be calculated by multiplying the Row total by the Column total, and dividing by the grand total.

$$E_{ij} = \frac{Row_{total} * Column_{total}}{Grand\ Total} = \frac{R * C}{N}$$

To complete the hypothesis test, we must combine our **alpha risk** with our **degrees of freedom** to determine the **critical chi-squared value** that will establish our **rejection criteria** for the hypothesis test.

The **degrees of freedom** for a contingency table is calculated as such:

$$df_{contingency} = (r - 1)(c - 1)$$

Relationship Between Variables

$$\boldsymbol{Simple\ Linear\ Regression\ Model : y = \beta_1 x + \beta_0}$$

$$\boldsymbol{Slope = \beta_1 = \frac{S_{xy}}{S_{xx}}}$$

Where

$$\boldsymbol{S_{xy} = \sum (X_i * Y_i) - \frac{(\sum X_i)(\sum Y_i)}{n}}$$

$$\boldsymbol{S_{xx} = \sum X_i^2 - \frac{(\sum X_i)^2}{n}}$$

$$\boldsymbol{Y - Intercept = \beta_0 = \bar{Y} - \beta_1 \bar{X}}$$

Where

$$\boldsymbol{\bar{X} = \frac{\sum X_i}{n}}$$

$$\boldsymbol{\bar{Y} = \frac{\sum Y_i}{n}}$$

$$\boldsymbol{SSE = Sum\ of\ Squares\ of\ the\ Error = \sum \varepsilon rror^2 = \sum (y - \hat{y})^2 = S_{yy} - \beta_1 S_{xy}}$$

$$\boldsymbol{S_{yy} = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}}$$

$$\boldsymbol{Sample\ Variance\ of\ the\ Error: s_e^2 = \frac{SSE}{n-2}}$$

The **null hypothesis** associated with linear regression is that the **β_1 coefficient, for the slope of the line, is zero**. The **alternative hypothesis** is that your slope is not equal to zero, thus proving a linear relationship between the two variables.

H_0: $\beta_1 = 0$ & H_A: $\beta_1 \neq 0$

$$t = \frac{\hat{\beta}_1 - \beta_1}{S_e / \sqrt{S_{xx}}} \qquad Where\ the\ degrees\ of\ Freedom\ d.f. = n - 2$$

The **Pearson Correlation Coefficient, r**, measures the strength of the linear relationship between two variables which means that a change in one variables (the independent variable) also implies a change in the other (dependent variable).

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} * \sqrt{S_{yy}}}$$

The **coefficient of determination**, **R^2**, reflects the proportion of the total variability in the Y variable that can be explained by the regression line, and thus it also reflects your models adequacy for a particular sample of data.

$$\textbf{Coefficient of Determination } (\mathbf{R^2}) = \mathbf{r^2}$$

Process and Performance Capability

$$Process\ Capability\ Analysis = \frac{Process\ Specification}{Process\ Performance}$$

$$C_p = \frac{USL - LSL}{6\sigma}$$

$$C_{pk} = Min\left(C_{p,Lower}, C_{p,Upper}\right) = Min\left(\frac{USL - \tilde{x}}{3s}, \frac{\tilde{x} - LSL}{3s}\right)$$

$$C_{pm} = \frac{C_p}{\sqrt{1 + \frac{(\mu - T)^2}{s^2}}} = \frac{USL - LSL}{6\sqrt{s^2 + (\tilde{x} - T)^2}}$$

$$C_r = \frac{1}{C_p}$$

$$P_p = \frac{USL - LSL}{6\boldsymbol{s_{p_p}}}$$

$$C_p = \frac{USL - LSL}{6\boldsymbol{s_{c_p}}}$$

$$P_{pk} = Min\left(\frac{USL - \tilde{x}}{3\boldsymbol{s_{p_p}}}, \frac{\tilde{x} - LSL}{3\boldsymbol{s_{p_p}}}\right)$$

$$C_{pk} = Min\left(\frac{USL - \tilde{x}}{3\boldsymbol{s_{c_p}}}, \frac{\tilde{x} - LSL}{3\boldsymbol{s_{c_p}}}\right)$$

Reliability

Definitions

Reliability is formally defined as *the probability that a product will perform successfully under specified operating conditions for a given period of time.*

Maintainability is formally defined as *the ability of an item to be retained or restored to specified conditions when maintenance action is performed by personnel having specified skill levels and using prescribed procedures and resources at each prescribed level of maintenance and repair.*

MTTF (Mean Time To Failure) - MTTF is the most common Reliability measurement for *non-repairable units* that represents exactly what it says it does, the *mean time* to product failure.

MTBF (Mean Time Between Failure) - This Reliability Indices is similar to MTTF except it applies to *repairable units* and again represents the mean time between product failures.

Failure Rate - This Reliability indices represents the rate at which your product fails. This can be thought of as the number of Failures per Unit of Time or number of Failures per Cycle.

Availability - This important indices combines reliability & maintainability into one single indices to indicate the overall probability that your product or process equipment will be ready for use when you're ready to use it.

Exponential Distribution Equations

$$\boldsymbol{PDF}:\ f(t) = \lambda e^{-\lambda t}$$

$$\boldsymbol{CDF}:\ F(t) = 1 - e^{-\lambda t}$$

$$\boldsymbol{Reliability}:\ R(t) = e^{-\lambda t} = e^{\frac{-t}{\theta}}$$

$$\boldsymbol{Failure\ Rate} = \lambda = \frac{1}{\theta}$$

$$\boldsymbol{\lambda} = \frac{Number\ of\ Failures}{Operating\ Time}\ or\ \frac{Number\ of\ Failures}{Number\ of\ Cycles}$$

$$\boldsymbol{\lambda} = \frac{1}{MTBF}$$

$$\boldsymbol{MTBF} = \frac{1}{Failure\ Rate\ (\lambda)} = \theta$$

$$\boldsymbol{MTTR} = \frac{Total\ Time\ Spent\ in\ Maintenance}{Total\ Number\ of\ Maintenance\ Activities\ conducted} = Maintainability$$

$$\boldsymbol{Availability} = \frac{MTBF}{MTBF + MTTR}$$

Weibull Distribution Equations

Weibull Distribution Equations

$PDF: \quad f(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta} e^{-\left(\frac{t}{\theta}\right)^{\beta}}$	$Reliability: \quad \mathrm{R(t)} = \mathrm{e}^{-\left(\frac{t}{\theta}\right)^{\beta}}$
$CDF: \quad \mathbf{F(t)} = \mathbf{1} - \mathbf{e}^{-\left(\frac{t}{\theta}\right)^{\beta}}$	$Failure\ Rate: \quad \mathrm{h(t)} = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1}$

- **β(Beta)** - the Weibull Shape or Slope Parameter, and **θ(Theta)** - the Weibull Scale Parameter

when **β < 1**, the weibull distribution represents a system with a decreasing failure rate like the early life failures on the bathtub curve.

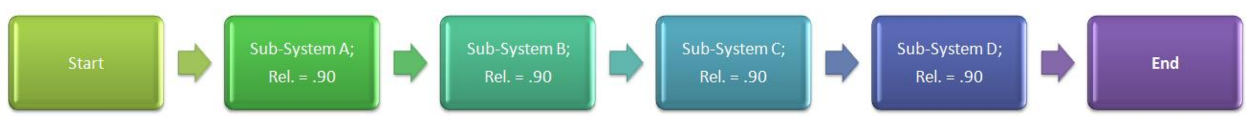
when **β = 1**, the weibull distribution is approximately equal to the **exponential distribution** and the failure rate is constant.

when **β > 1**, the weibull distribution represents a system with an increasing failure rate like the end of life failures on the bathtub curve.

when **β = 3.5**, this is another important beta value where the weibull distribution is approximately equal to the **normal distribution** and the failure rate is increasing.

System Reliability (Series & Parallel)

$$\boldsymbol{Series\ System\ Reliability} = \boldsymbol{R_{system}} = \boldsymbol{R_1} \times \boldsymbol{R_2} \times \boldsymbol{R_3} \times \boldsymbol{R_4} \times \dots \dots \boldsymbol{R_n}$$



$$\boldsymbol{Parallel\ System\ Reliability} = \boldsymbol{R_{system}} = \boldsymbol{1} - (\boldsymbol{U_1} \times \boldsymbol{U_2} \times \boldsymbol{U_3} \times \boldsymbol{U_4} \times \dots \dots \boldsymbol{U_n})$$

$$\boldsymbol{Where\ U_1} = \boldsymbol{1} - \boldsymbol{R_1}$$