

Central Tendency

$$\text{mean} = \frac{\sum x}{n} = \frac{\text{Sum of all observations}}{\text{Total number of observations}}$$

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

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

Median = Middle Value = 185K, 197K, 230K, 252K, 1.4M

Median of Even Numbers = 85K, 197K, 230K, 252K

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

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

Variance

$$\text{Range} = R = \text{Max}(x) - \text{Min}(x)$$

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

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

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

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

Probability

*The **Probability of A or B** = $P(A \cup B) = P(A) + P(B) - P(A \cap B)$*

*The **Intersection of A & B** = A and B = $P(A \cap B)$*

*The **Probability of A°** = $P(A^\circ) = 1 - P(A)$*

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

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

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

*The **Union of A & B** = A or B = $(A \cup B)$*

*The **Intersection of A and B** = A & B = $(A \cap B)$*

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

*The **Addition Rule for Two Events** = $P(A \cup B) = P(A) + P(B) - P(A \cap B)$*

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

*The **Addition Rule for Mutually Exclusive Events** = $P(A \cup B) = P(A) + P(B)$*

Probability Distributions

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

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

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 $= \theta = \frac{1}{\lambda}$ where $\lambda = \text{Occurrence Rate}$

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

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

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

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

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: $\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\alpha/2}}$

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

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

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

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

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

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

$$\text{Population Proportion: } Z_o = \frac{\hat{p} - p_o}{\sqrt{\frac{p_o(1 - p_o)}{n}}}$$

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

H₀: 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})		

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

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

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

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

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

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

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

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

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

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

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

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 (χ^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**.

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

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.

$$E_{ij} = \frac{\text{Row}_{total} * \text{Column}_{total}}{\text{Grand Total}} = \frac{R * C}{N}$$

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:

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

Process and Performance Capability

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

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

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

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

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

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

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

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

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

Reliability

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

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

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

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

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

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

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

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

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

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

Reliability Continued

Weibull Distribution Equations

Weibull Distribution Equations

$$PDF: f(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1} e^{-\left(\frac{t}{\theta}\right)^{\beta}}$$

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

$$CDF: F(t) = 1 - e^{-\left(\frac{t}{\theta}\right)^{\beta}}$$

$$Failure Rate: h(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1}$$

- **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)

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



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

$$Where U_1 = 1 - R_1$$

