relative frequency of category $i = \frac{\text{frequency of category } i}{n}$

relative frequency $= \frac{\text{frequency of each group}}{\text{total number of observations}}$

Mean: population of size N: $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$ sample of size n: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

Variance: population: $\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$ sample: $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

Standard deviation: $SD = \sqrt{\text{variance}}$

Process capability index: $C_p = \frac{\text{upper specification} - \text{lower specification}}{\text{total variation}}$

Standard value: Z-score for i^{th} observation: $Z_i = \frac{x_i - \bar{x}}{s}$

Coefficient of variation: $CV = \frac{\text{standard deviation}}{\text{mean}}$

Coefficient of Skewness: $CS = \frac{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^3}{\sigma^3}$

Coefficient of Kurtosis: $CK = \frac{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^4}{\sigma^4}$

Covariance: population: $cov(x,y) = \frac{\sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y)}{N}$, sample: $cov(x,y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n}$

Correlation: population: $\rho_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y}$ sample: $r_{xy} = \frac{cov(x,y)}{s_x s_y}$

Standard Error of the mean $= \frac{\sigma}{\sqrt{n}}$

Confidence interval for mean with know population SD: $\bar{x} \pm Z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right)$

Confidence interval for mean with unknown population SD: $\bar{x} \pm t_{\alpha/2, n-1} \left(\frac{\sigma}{\sqrt{n}} \right)$

Confidence interval for proportion: $\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Prediction interval: $\bar{x} \pm t_{\alpha/2, n-1} \left(s\sqrt{1 + \frac{1}{n}} \right)$

Sample size for mean: $n \geq (Z_{\alpha/2})^2 \frac{\sigma^2}{E^2}$ (E = margin of error)

Sample size of proportion: $n \geq (Z_{\alpha/2})^2 \frac{\pi(1-\pi)}{E^2}$ (E = sampling error, π = proportion)

One sample t-test for mean, σ known : $z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

One sample t-test for mean, σ unknown: $z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

One sample t-test on a proportion : $z = \frac{\hat{p} - \pi_0}{\sqrt{\pi_0 (1 - \pi_0)/n}}$

Smoothing $\text{window}_t = \frac{1}{k}\left(Y_t + Y_{t-1} + \dots + Y_{t-(k-1)}\right)$

Exponential smoothing model: $\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$

$$= \alpha Y_t + (1-\alpha)\left(\alpha Y_{t-1} + (1-\alpha)\hat{Y}_{t-1}\right)$$

$$= \alpha\left(Y_t + (1-\alpha) Y_{t-1} + (1-\alpha)^2 Y_{t-2} + \dots\right)$$

Root mean square error: $RMSE = \sqrt{\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2}$

= sqrt(mean(residuals(fit)^2))

Classification accuracy = $\frac{\text{True Positive} + \text{True Negative}}{TP + TN + FP + FN}$

Recall $= \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$

Precision $= \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$

F1 score $= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$