Department of Geomatics Engineering,
University of Calgary

"Adjustment of Observations"

Lecture notes for ENGO 361

by Dr. Naser El-Sheimy (edits by Dr. Ivan Detchev)

Winter 2015 (edits W2017 & W2018)

1. INTRODUCTION TO ADJUSTMENT OF OBSERVATIONS

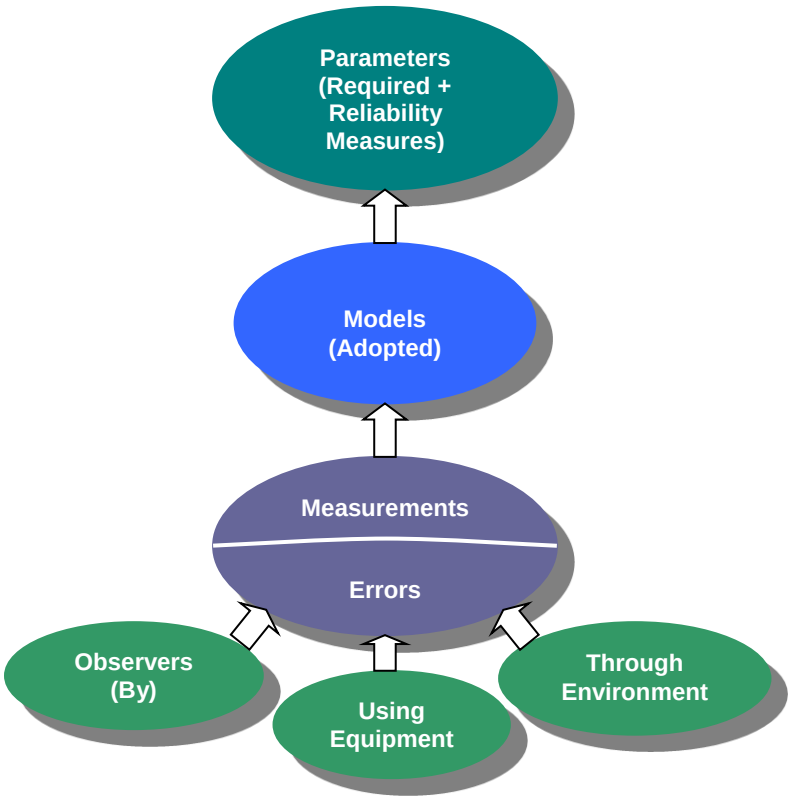
1.1. Why Observations?

Geomatics Engineers are usually faced with the problem of estimating some unknown quantities (parameters). This is done through the collection of measurement(s) of some kind (observations) – and then adopting the appropriate mathematical model relating both the *observations* and the *unknowns*.

- Observations are generally acquired using *instruments/devices* that are used by an *observer* in certain *environment*.
- The different conditions, under which the measurements are made, cause variations in measurements and, therefore, no measured quantity is completely determinable. A fixed value of a quantity may be conceived as its true value. The difference between the measured quantity and its true value t is known as error

$$\varepsilon = t - l$$

ε... *true error*
t... *true value*
l... *observed value*



- Since the true value of a measured quantity cannot be determined, the exact value of ε can never be found out. However, if a best estimate $\hat{x}$ which is known as the most probable value of *t*, can be determined, $\hat{x}$ can be used as a reference to express the variations in x. If we define υ as residual then

$$v = \hat{x} - l$$

v...*estimated error (estimated residual)*
$\hat{x}$...*estimated value*

1.2. Sources of Errors in Measurements

Personal	Instrumental	Natural
Limitation of the observer (the ability of the observer to repeat the same measurement) or a carelessness of the observer	Imperfect construction or incomplete calibration of the instrument being used (e.g. incorrect graduation of a measuring tape)	Changing environmental conditions during the measurement campaign (e.g. temperature variation causes expansion/contraction of steel measuring tapes)

1.3. Types of Errors

1) Gross Errors or blunders or mistakes

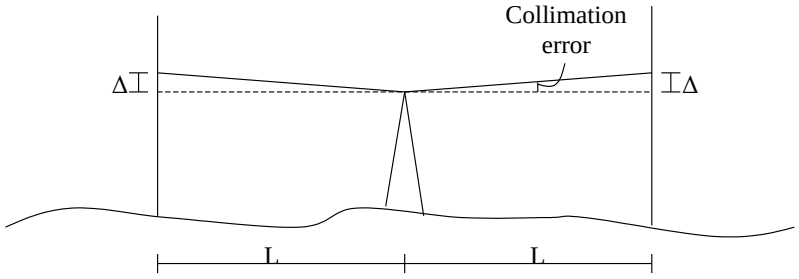
- Characteristics: Its magnitude is significantly large/small/different in comparison to the measured values (abnormal observation).
- Source: Personal errors (carelessness of the observer)
- Effect: Inhomogeneous set of observations
- Treatment: Must be detected and eliminated from the measurements by careful checking of the measurements.

e.g., measuring a distance – 4th measurement is a blunder
31.1 m, 31.3 m, 31.2 m, **13.1 m**, 31.15

2) Systematic Errors

- Characteristics: Occur according to a deterministic system, when known, can be expressed by functional/mathematical models.
- Source: Instrumental, natural, personal or all
- Effect: Shifting all the observations from the true value. It can be constant if its magnitude and sign remain the same throughout the measuring process.
- Treatment: Must be detected and corrected by, e.g. calibrating the instruments before being used.
- In some cases, can be eliminated by:
  1. Adjusting the instrument
  2. Using certain procedures during measurement, e.g. survey levels collimation errors can be minimized by taking measurements at equal distances from the level.

- Example:



3) Random Errors

- Characteristics: It is the remaining errors after removing the gross and systematic errors. They have no functional relationship based upon a deterministic system, usually modelled by stochastic model (probability theory)
- Source: Personal, natural and instrumental.
- Cannot be generally eliminated however it can be minimized by taking redundant observations and applying the so called "Method of Least Squares"
- This process is referred to as the "***Adjustment of Observations***" or "Adjustment Computations" which is the main topic of this course.
- Based on the above fact, we cannot seek the 'true' value; all we can derive is an 'estimate' for the 'true' value.
- The small variations between the measurements and the 'true' value or its 'estimate' are regarded as "errors."

1.4. Notation Used in Least Squares

- Parameters: The unknown quantities, will be denoted by

$$\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_u \end{bmatrix}^T$$

u... number of unknowns

- Observations:

$$\boldsymbol{l} = \begin{bmatrix} l_1 & l_2 & \dots & l_n \end{bmatrix}^T$$

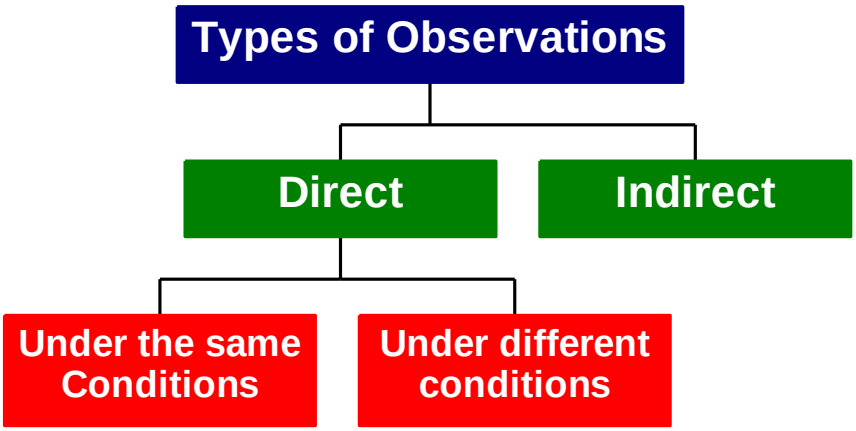
n... number of observations (measurements)

- Mathematical Model (discussed in further detail in chapter 3).

Ex : $\boldsymbol{0} = \mathbf{f}(\boldsymbol{x}, \boldsymbol{l})$

The function that relates $\boldsymbol{x}$ *and* $\boldsymbol{l}$.

1.5. Types of Observations



- **Direct observation under the same conditions**
  - The unknown is directly measured, and
  - Observer, instruments, and environment are all the same during the measurement campaign
  - Example: measuring an angle using a theodolite or total station:

Parameter:

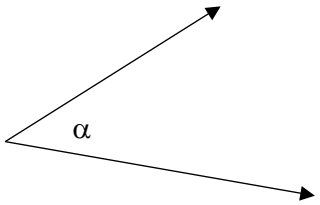
$$x = \alpha, u = 1$$

Observations:

$$l = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{bmatrix}^T$$

Math Model

$$\hat{x} = mean(l)$$



- **Direct observation under different conditions**
  - Same as above with the case that any of the observer, the instrument or the environment changes during the measurement campaign

- **Indirect observation**
  - **x** is measured indirectly (i.e. estimated)
  - Example: estimating the height of a building using vertical angle and horizontal distance measurement from a total station:

Parameter:

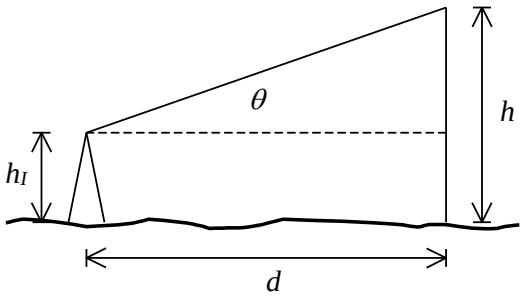
$$x = h, u = 1$$

Observations:

$$l = \begin{bmatrix} \theta & d \end{bmatrix}^T$$

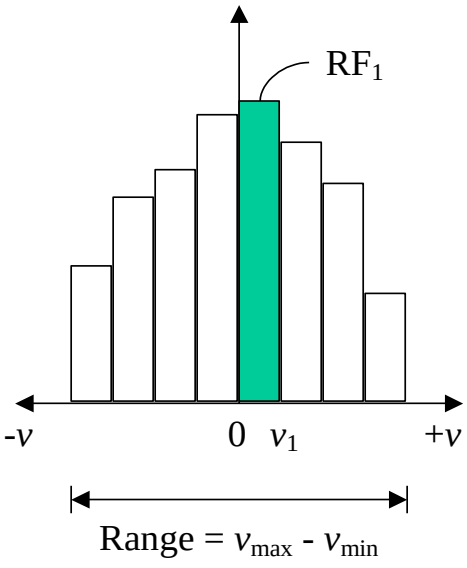
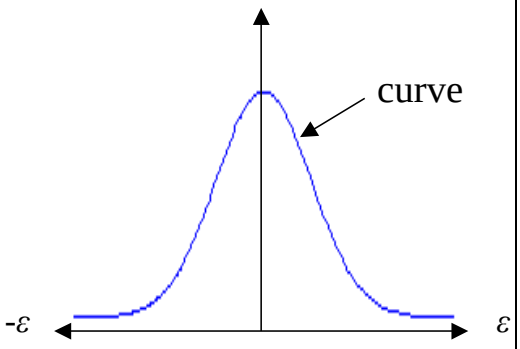
Math Model

$$\hat{x} = d \cdot tan\theta + h_I$$



1.6. Behaviour of Random Errors

- After removing the gross errors and correcting the systematic errors, the survey measurements will finally be left with random errors. In order to understand the random nature of a survey measurement, it may be measured repeatedly and its frequency distribution describing its random variation is studied.

	Practical Case	*Random Case* *Theoretical Case*
No. of obs.	$n \rightarrow$ finite number	$n \rightarrow \infty$
Variable	Random Sample $\boldsymbol{l} = [l_1\ l_2\ \ldots\ l_n]^T$	Random Variable
Estimate	$\hat{x} = \frac{\Sigma l_i}{n}$ (mean)	$\hat{x} \rightarrow t$ (true value)
Errors	Residuals $v_i = \hat{x} - l_i \quad (\Sigma v_i = o)$	True Errors $\varepsilon_i = t - l_i$
Spread of Errors (Distribution of errors)	Relative Frequency (RF)  Range = v_{max} - v_{min} RF_1 is the number of observations which have residuals between 0 and v_1 *Probability Distribution Histogram (PDH)*	Relative Frequency (RF)  *Probability Distribution Function (PDF)*

1.7. Understanding the Meaning of Residual Errors (v)

- Given $\boldsymbol{l} = [l_1\ l_2,\ \ldots\ l_n]^T$ and accepting, for the time being, the fact that the arithmetic mean $\bar{x}$ of all l_i is the best estimate $\hat{x}$ for the measured quantity – i.e.

$$\hat{x} = \bar{x} = \frac{\Sigma l_i}{n}$$

  we can compute v_i as $v_i = \hat{x} - l_i$

- The residuals express the degree of closeness of the repeated measurements of the same quantity to each other and therefore the (v) values can be used in expressing the precision of $\hat{x}$ (and also the precision of the observer who took/made these measurements.)

  Example: two observers A and B measure the same angle.

A	**B**
$v_1 = +2''$	$v_1 = +5''$
$v_2 = -1''$	$v_2 = -3''$
$v_3 = +1''$	$v_3 = +2''$
$v_4 = 0$	$v_4 = -3''$
$v_5 = -2''$	$v_5 = -4''$
	$v_6 = +3''$

  Now, we define the range over which the residuals change as,

  Range (v) = $v_{max} - v_{min}$

  Range A = 2″ - (-2″) = 4″

  Range B = 5″ - (-4″) = 9″

  Since Range A < Range B

  We conclude that the angle computed from the data set of (A) is more precise than that computed from the data set (B)

1.8. Drawing the Probability Distribution Histogram (PDH)

1) Calculate the estimated parameter:

$$\hat{x} = \bar{x} = \frac{\Sigma l_i}{n}$$

2) Calculate the residuals for each observation:

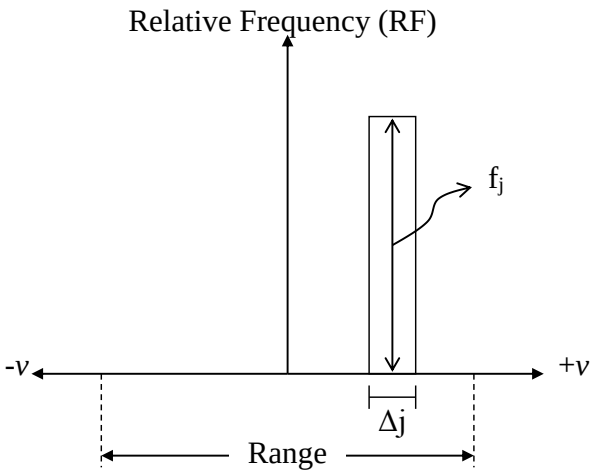
$$v_i = \hat{x} - l_i$$

3) Calculate the range of the residuals:

$$Range = v_{\max} - v_{\min}$$

4) Divide the range into k equal intervals Δ_j (j=1, 2, … k).

5) Calculate the relative frequency for each interval.



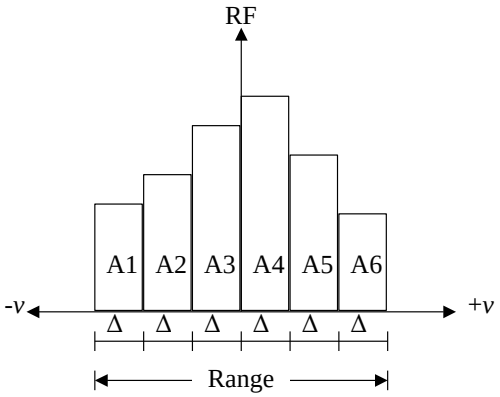
$$f_j = \frac{n_j}{n \cdot \Delta_j}$$

n_j ... *the number of residuals that fall within the boundaries of* Δ_j

Note:

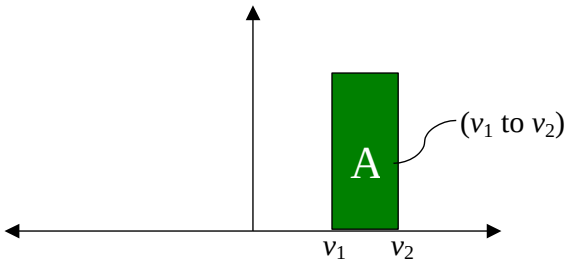
- The area under the histogram = 1

$$A_{PDH} = \sum_{j=1}^{k} A_i = \sum_{j=1}^{k} f_j \cdot \Delta = \sum_{j=1}^{k} \frac{n_j}{n \cdot \Delta} \cdot \Delta = \sum_{j=1}^{k} \frac{n_j}{n} = \frac{1}{n} \sum_{j=1}^{k} n_j = \frac{n}{n} = 1$$



- The histogram can be used for probability computation.

$$v_1 \le \text{Probability} \le v_2 = A(v_1 \text{ to } v2)$$

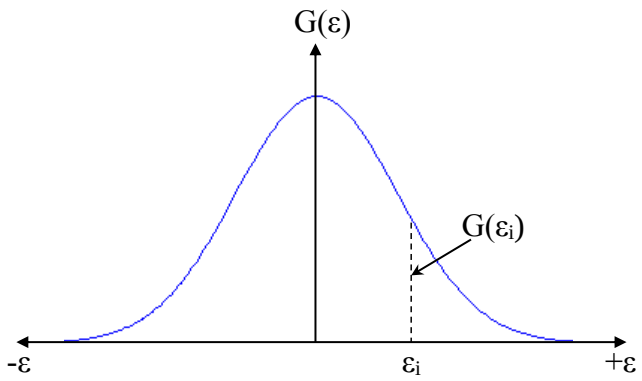


1.9. Characteristics of PDF

- The most commonly accepted model of PDF was given by Gauss (Gaussian PDF)

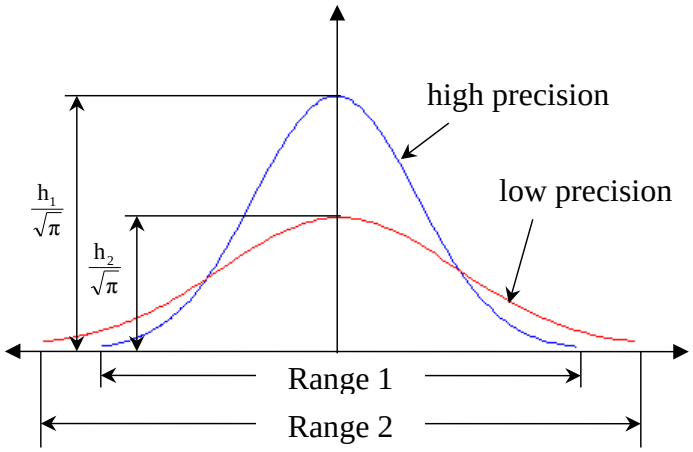
$$G(\varepsilon) = \frac{h}{\sqrt{\pi}} e^{-h^2\varepsilon^2}$$

e = 2.71828



Note: *h* is the only parameter that completely describes the shape of the Gauss PDF.

- $G(0) = \frac{h}{\sqrt{\pi}}$
- *h* is usually called "precision index" (precision $\propto$ h, the precision is proportional to h)
- Since the area under the curve = 1, then the higher the PDF (i.e. the larger the *h*), the narrower the curve must become – i.e. less range (more precise)



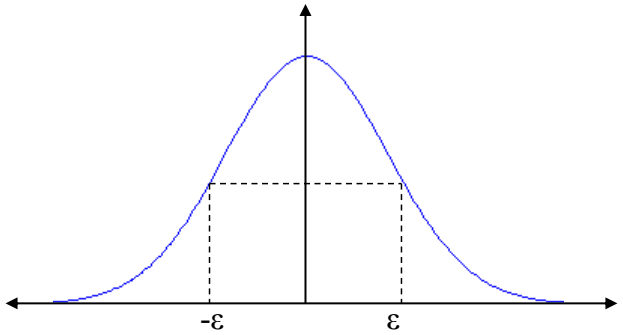
Properties of Gauss PDF

1) Area under the curve = unity

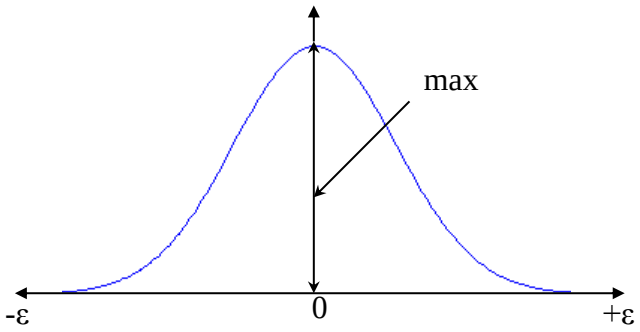
$$\int_{-\infty}^{\infty} G(\varepsilon)d\varepsilon = 1$$

2) Symmetric around ε = 0. Positive and negative errors have equal probability to occur (no systematic errors).

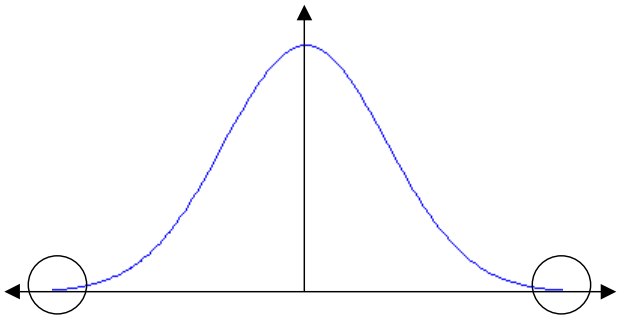
$$Probability(\varepsilon) = Probability(-\varepsilon)$$



3) Max ordinate at ε = 0. The probability of very small errors is very high.



4) Asymptotic to the ε axis at ε = ± ∞. The probability of very large errors is negligibly small. (i.e. no gross errors)



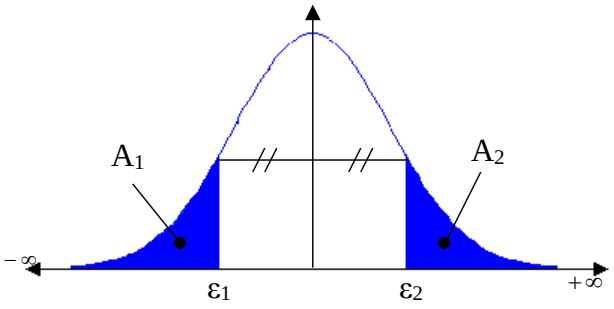
5) Statistical properties

$$\text{Prob}(\varepsilon \le \varepsilon_1) = \int_{-\infty}^{\varepsilon_1} G(\varepsilon)d\varepsilon = A_1$$

$$\text{Prob}(\varepsilon \ge \varepsilon_2) = \int_{\varepsilon_2}^{\infty} G(\varepsilon)d\varepsilon = A_2$$

$$= 1 - \int_{-\infty}^{\varepsilon_2} G(\varepsilon)d\varepsilon$$

$$\text{Prob}(\varepsilon_1 \le \varepsilon \le \varepsilon_2) = \int_{\varepsilon_1}^{\varepsilon_2} G(\varepsilon)d\varepsilon$$

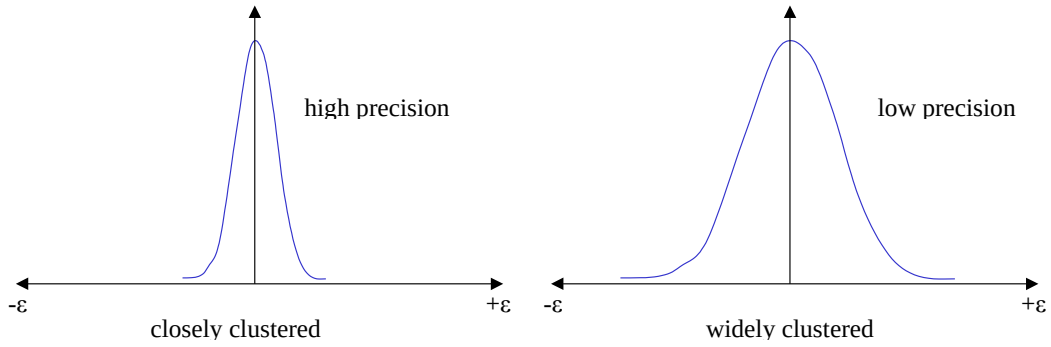


(More details to be discussed in Chapter 9 on Statistical Analysis)

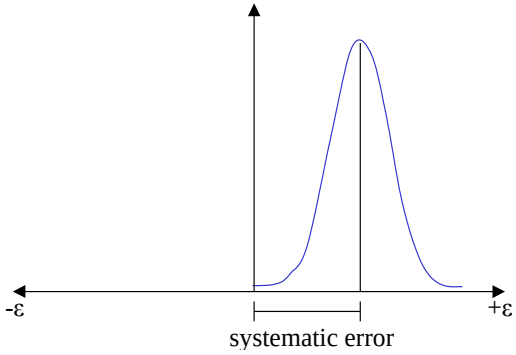
1.10. Reliability of Measurements

- Recall that the true value *t* of a certain unknown quantity can never be obtained. However, an estimate $\hat{x}$ can be determined. However, having an estimate $\hat{x}$ for this quantity will not be satisfactory, at least as far as its user is concerned, because it is influenced to a certain degree by the inherent random errors.
- Therefore, there should be a certain "measure" to express the existing random errors, associated with each estimate, to describe the degree of its "goodness" and "reliability".
- This measure should help, as well, in accepting or rejecting certain observations depending on the desired precision. Also this measure of errors is essential in comparing different sets of measurements, and hence accepting the more "precise" one.
- Three terms are commonly used in expressing the reliability of measurements:

1. *Precision*: the degree of closeness of repeated measurements of the same quantity to each other. Precision is affected only by random errors. Precision is indicated by the dispersion or spread of the probability distribution.

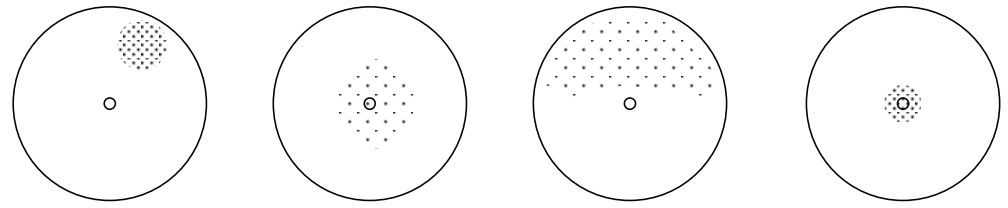


2. *Accuracy*: the degree of closeness of a measurement to the true value. Accuracy is affected by both random and systematic errors. The accuracy of measurements increases with the absence of systematic errors. In other words, the difference between the mean value of an estimate and the true value is a measure of "systematic error" in the measurements.



3. *Uncertainty*: the range within which it is expected that the error of a measurement will fall. A specified level of probability is generally associated with an uncertainty. The 95% uncertainty, for example, is the range of values within which it is 95% probable that the error of the measurements will fall.

- The above discussion can be illustrated by the following example: Consider a circular target with its centre representing the true position of shooting with a gun.



precise but not accurate | *not precise but accurate* | *not precise & not accurate* | *precise and accurate*

Precision ⟶ Internal Reliability
Accuracy ⟶ External Reliability

- In the case of absence of systematic errors, accuracy is equivalent to precision.

1.11. Measures of Precision

1) Average Error (a_e)

- The arithmetic mean of the absolute value of errors/residuals

True Value is Unknown	True Value is Known
$a_e = \frac{\Sigma\lvert v_i\rvert}{(n-1)}$ Note: we only have one unknown, say x, and hence we need a single observation to determine it. The reminder (n-1) is usually known as the number of "redundant observations" or "redundancy" or "degrees of freedom" which we will refer to as "*r*" $r = n - u$ *r... redundancy* *n... number of observations* *u... number of unknowns*	$a_e = \frac{\Sigma\lvert \varepsilon_i\rvert}{n}$ Since the true value of x is known, therefore all the observations are considered degrees of freedom
The estimated errors: $v_i = \hat{x} - l_i$	The true errors: $\varepsilon_i = t - l_i$
The best estimate: $\hat{x} = \bar{x} = \frac{\Sigma l_i}{n}$	

- Random variable

$$a_e = \int_{-\infty}^{\infty} |\varepsilon| \cdot G(\varepsilon)\, d\varepsilon = \frac{1}{h\sqrt{\pi}} \approx \frac{0.5642}{h}$$

2) Probable Error (p_e)

- p_e has the following property: Half the resulting errors/residuals are smaller in magnitude than p_e and the other half are larger than p_e.

$$50\%\ of\ |v_i| > p_e$$
$$50\%\ of\ |v_i| < p_e$$

- How: Arrange the absolute values of the errors/residuals into either ascending or descending order.

For odd numbers $p_e = v_{\left(\frac{n+1}{2}\right)}$

e.g. v_1 v_2 $\boxed{v_3}$ v_4 v_5

$p_e = v_3$

For even numbers $p_e = \frac{1}{2}\left[v_{\left(\frac{n}{2}\right)} + v_{\left(\frac{n}{2}+1\right)}\right]$

e.g. v_1 v_2 $\boxed{v_3}$ $\boxed{v_4}$ v_5 v_6

$$p_e = \frac{v_3 + v_4}{2}$$

- Problems with p_e

1″ 1″ 3″ 5″ 6″ 7″ 11″

1″ 1″ 3″ 5″ 6″ 7″ **100″**

p_e is the same although the 2nd group of observations contains a gross error.

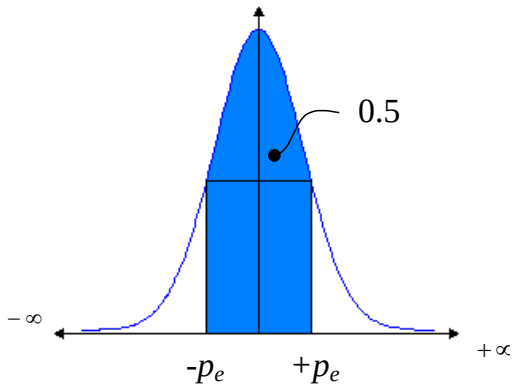
- p_e for random variable

$$\int_{-p_e}^{p_e} G(\varepsilon)d\varepsilon = 0.5$$

and, due to the symmetry of Gaussian PDF:

$$2\int_{0}^{p_e} G(\varepsilon)d\varepsilon = 0.5$$

$$p_e \approx \frac{0.4769}{h}$$



3) Standard Deviation (σ)

- σ is defined as the square-root of the arithmetic mean of the sum of square of the errors/residuals.
- The square of the standard deviation – σ^2 – is known as the "variance" or "mean square error" and consequently σ is sometimes referred to as the Root Mean Square Error (RMSE)

Random Sample

t = unknown

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n} v_i^2$$

t = known

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i^2$$

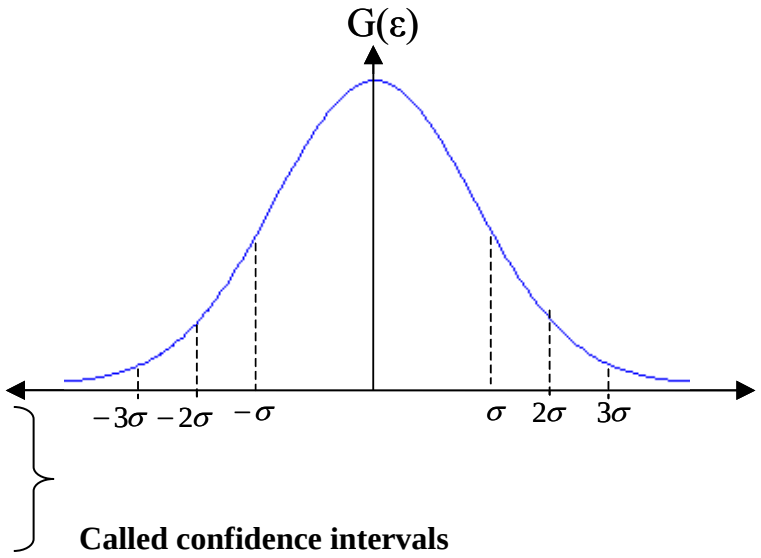
- σ for a random variable

$$\sigma^2 = \int_{-\infty}^{\infty} \varepsilon^2 G(\varepsilon) d\varepsilon = \frac{1}{2h^2}$$

$$\sigma = \frac{1}{h\sqrt{2}} \approx \frac{0.7071}{h}$$

Unlike the previous two measures of precision, σ has some distinguished properties which make it a preferred measure of precision, they are:

1. It has a physical interpretation related to the PDF



$$\text{Probability}(-\sigma \le \varepsilon \le \sigma) = 0.683$$
$$\text{Probability}(-2\sigma \le \varepsilon \le 2\sigma) = 0.954$$
$$\text{Probability}(-3\sigma \le \varepsilon \le 3\sigma) = 0.997$$

Called confidence intervals

- The 3σ limit is usually used in practice as a criterion for rejecting bad observations.
- |Residuals| greater than 3σ are usually treated as gross errors.
- $\text{Probability}(|\varepsilon| \ge 3\sigma) = 0.003 = 0.3\%$

2. Because of the squaring process of the residuals in the calculation of σ, the magnitude of large errors is completely represented and reflected in the computed σ value

1.12. Uni-variate Statistics

- Uni-variate: Single variable (i.e. we deal with "measurements" of a single quantity – e.g. a distance, angle, etc.)

  Given: $\boldsymbol{l} = l_1 , l_2 , l_3 , \ldots\ldots\ldots l_n \ldots$ (n observations of the unknown quantity x)

  Required: Best estimate of x, i.e. $\hat{x}$

- Best estimate: $\hat{x}$

$$\hat{x} = \bar{x} = \frac{\sum_{i=1}^{n} l_i}{n}$$

- Variance of x: σ_x^2

$$\sigma_x = \sqrt{\sigma_x^2}$$ σ_x units of x (e.g. m or degrees)

$$\sigma_x^2 = \frac{\sum v^2}{n-1}$$ σ^2_x square units of x (m^2 or $degree^2$)

- Variance of best estimate: $\sigma_{\bar{x}}^2$

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n}$$

Note: Since $\hat{x}$ is the best estimate of the unknown parameter out of a group of measurements (having a variance of σ_x^2), therefore $\sigma_{\bar{x}}^2$ should be better than σ_x^2

$$\sigma_{\bar{x}} < \sigma_x$$

$\sigma_{\bar{x}}$... *precision of the mean*

σ_x... *precision of a single measurement*

- Proof that $\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n}$:

To determine the standard deviation of the mean, we begin with the expression for the mean:

$$\bar{x} = \frac{\sum_{i=1}^{n} l_i}{n} = \frac{l_1}{n} + \frac{l_2}{n} + \ldots\ldots\ldots\ldots\ldots + \frac{l_n}{n}$$

Then, we use the *Law of Propagation of Variances*:

- Given x = **f** (a, b, c) with σ_a ,σ_b ,σ_c

$$\sigma_x^2 = \left(\frac{\partial \boldsymbol{f}}{\partial a}\right)^2 \sigma_a^2 + \left(\frac{\partial \boldsymbol{f}}{\partial b}\right)^2 \sigma_b^2 + \left(\frac{\partial \boldsymbol{f}}{\partial c}\right)^2 \sigma_c^2$$

- e.g. $x = x_1 + x_2 = \boldsymbol{f}(x_1, x_2)$

$$\sigma_x^2 = \left(\frac{\partial \boldsymbol{f}}{\partial x_1}\right)^2 \sigma_{x1}^2 + \left(\frac{\partial \boldsymbol{f}}{\partial x_2}\right)^2 \sigma_{x2}^2$$

$$\sigma_x^2 = (1)\sigma_{x1}^2 + (1)^2 \sigma_{x2}^2 = \sigma_{x1}^2 + \sigma_{x2}^2$$

♦ Applying the law of propagation of variances to the equation of the mean

$$\sigma_{\bar{x}}^2 = \left(\frac{1}{n}\right)^2 \sigma_{\ell 1}^2 + \left(\frac{1}{n}\right)^2 \sigma_{\ell 2}^2 + \ldots\ldots\ldots\ldots + \left(\frac{1}{n}\right)^2 \sigma_{\ell n}^2$$

$$= \left(\frac{1}{n}\right)^2 \left(\sigma_{\ell 1}^2 + \sigma_{\ell 2}^2 + \ldots\ldots\ldots\ldots + \sigma_{\ell n}^2\right)$$

$$= \frac{1}{n^2}\left(n\sigma^2\right)$$

$$= \frac{\sigma^2}{n}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

♦ This yields the final expression for the standard deviation of the mean

$$\boxed{\sigma_{\bar{x}} = \sqrt{\frac{\sum v_i^2}{(n-1)n}}}$$

1.13. Direct Observations with Different Conditions

- Consider that two observers A and B are measuring the same quantity X (e.g. a distance or angle) and that it is required to estimate the best estimate of X by making use of the two sets of observations

Observer A	*Observer B*
α_1	α_1
α_2	α_2
α_3	α_3
⋮	⋮
α_n	α_m

- Note that each group of observations is considered as a uni-variate problem with its own best estimate and standard deviations.

	Observer A	Observer B
Best estimate:	$\bar{\alpha}_A = \frac{\sum_{i=1}^{n} \alpha_i}{n}$	$\bar{\alpha}_B = \frac{\sum_{i=1}^{m} \alpha^i}{m}$
Precision:	$\sigma_{\alpha A} = \sqrt{\frac{\sum v_i^2}{n-1}}$ $\text{for} \rightarrow i = 1.....n$ $v_i = \bar{\alpha}_A - \alpha_i$	$\sigma_{\alpha B} = \sqrt{\frac{\sum v_i^2}{m-1}}$ $\text{for} \rightarrow i = 1.....m$ $v_i = \bar{\alpha}_B - \alpha_i$

- Now, $\bar{\alpha}_A$ and $\bar{\alpha}_B$ can be considered as measurements with different precision – they are the best estimate of two groups of measurements for the same quantity (x).
- Therefore our original problem can be reformulated as follows:

  Given: $\bar{\alpha}_A, \sigma_{\bar{\alpha}A}$ and $\bar{\alpha}_B, \sigma_{\bar{\alpha}B}$

  Required: The best estimate of α, $\hat{\alpha}$

- Could we calculate the new estimate, $\hat{\alpha}$, by simply doing the methodical mean: $\alpha = \frac{\bar{\alpha}_A + \bar{\alpha}_B}{2}$

  ? The answer should be **NO** - because this will not take into account the difference in precision of the two sets of observations expressed by $\sigma_{\bar{\alpha}A}$ *and* $\sigma_{\bar{\alpha}B}$

- To account for the difference in precision between the two sets of observations, we introduce a new quantity called the weight (P), where: $P \propto \frac{1}{\sigma^2}$

- Recall

$$\sigma_x = \sqrt{\frac{\sum v_i^2}{n-1}}$$ where v_i is the estimated error (residual) for observation (i), and

$$\sigma_{\bar{x}} = \sqrt{\frac{\sum v_i^2}{(n-1)n}}$$

- Therefore, a measurement of high precision will have a small variance and vice versa.
- Since the value of variance goes in opposite direction to that of precision, another measure of precision is often used.
- We usually call this quantity "weight".

$$\bar{\alpha} = \frac{\sum P_k \alpha_k}{\sum P_k} = \frac{P_A \bar{\alpha}_A + P_B \bar{\alpha}_B}{P_A + P_B}$$

$$P_A = \frac{1}{\sigma^2_{\bar{\alpha}A}} \quad P_B = \frac{1}{\sigma^2_{\bar{\alpha}B}}$$

- In general:

$$\hat{x}_{WM} = \frac{\sum P_k \bar{x}_k}{\sum P_k}$$ weighted mean

$$\sigma_{\hat{x}_{WM}} = \sqrt{\frac{\sum_{k=1}^{N} P_k v_k^2}{(N-1)\sum P_k}}$$ standard deviation of the weighted mean

- Note: if all the observations have the same variance (i.e. equal weights), the above formulas will be the same as the mean and variance equations discussed before.