Contents

Statistics

G.N.Dar

May 10, 2020

1 Error Propogation

The aim of this lecture is to know the uncertainity in function Q if we know the uncertainity in the varaible say x and y. The technique used is known as *Propogation of Errors.*

Varaince σ_Q^2 as a function of Varainces say σ_x^2 and σ_y^2 we are going to see if a random varaible X measures some quantity with varaince σ_x^2 and some other qauntity is measured by some other random varaible Y with a variance σ_y^2, then the total varaice is given as

$$\sigma_Q^2 = \sigma_x^2 \left(\frac{\partial Q}{\partial x}\right)^2\Bigg|_{\mu_x} + \sigma_y^2 \left(\frac{\partial Q}{\partial y}\right)^2\Bigg|_{\mu_y}$$

We know for measuring a quantity x say height of one single student in the class repeatedly we will get due to uncontrolable errors as

$$x = \{x_1, x_2, x_3, x_4, x_1, x_5, x_{n-1}, x_n\}$$

similarly for any other quantity say y

$$y = \{y_1, y_2, y_3, y_4, y_1, y_5, y_{n-1}, y_n\}$$

then the best value is taken as average given as

$$\bar{x} = \mu_x = \frac{1}{n}\sum_{i=1}^{n} x_i$$

And

$$\bar{y} = \mu_y = \frac{1}{n}\sum_{i=1}^{n} y_i$$

As the measurements varaied so is the variation or error propagated into the function say we have

$$f(x_1, y_1) \neq f(x_2, y_2) \neq \cdots \neq f(x_n, y_n) \neq f(\mu_x, \mu_y)$$

Therfore we define a parameter as

$$Q_i = f(x_i, y_i) = Q_1, Q_2, Q_3 ....Q_n$$

for $i = 1, 2, 3, ..., n$

And we also define to have a measurement

$$Q = f(\mu_x, \mu_y)$$

evaluated at average.

Using Taylor series we expand a function Q_i about the mean μ assuming the measured values are close to the mean and *neglect higher order ters*,therefore

$$Q_i = f(x_i, y_i) = f(\mu_x, \mu_y) + (x_i - \mu_x)\left(\frac{\partial Q}{\partial x}\right)\Big|_{\mu_x} + (y_i - \mu_y)\left(\frac{\partial Q}{\partial y}\right)\Big|_{\mu_y} + higher \quad order \quad terms$$

From the definition of variance (considering varaible here as Q instead of x) we know

$$\sigma_Q^2 = \frac{1}{n}\sum_{i=1}^{n}(Q_i - Q)^2$$

Now

$$Q_i - Q = (x_i - \mu_x)\left(\frac{\partial Q}{\partial x}\right)\Big|_{\mu_x} + (y_i - \mu_y)\left(\frac{\partial Q}{\partial y}\right)\Big|_{\mu_y} + higher \quad order \quad terms$$

in fact $i = 1, 2...n$

$$\sum_{i=1}^{n}(Q_i - Q) = (x_i - \mu_x)\left(\frac{\partial Q}{\partial x}\right)\Big|_{\mu_x} + (y_i - \mu_y)\left(\frac{\partial Q}{\partial y}\right)\Big|_{\mu_y} + neglect \quad higher \quad order \quad terms$$

Squring both sides

$$\frac{1}{n}\sum_{i=1}^{n}(Q_i - Q)^2 = \frac{1}{n}\sum_{i=1}^{n}\left((x_i - \mu_x)\left(\frac{\partial Q}{\partial x}\right)\Big|_{\mu_x}\right)^2 + \sum_{i=1}^{n}\frac{1}{n}\left((y_i - \mu_y)\left(\frac{\partial Q}{\partial y}\right)\Big|_{\mu_y}\right)^2$$

$$\sigma_Q^2 = \sigma_x^2\left(\frac{\partial Q}{\partial x}\right)^2\Big|_{\mu_x} + \sigma_y^2\left(\frac{\partial Q}{\partial y}\right)^2\Big|_{\mu_y} + \frac{2}{n}\sum_{i=1}^{n}(x_i - \mu_x)(y_- \mu_y)\left(\frac{\partial Q}{\partial x}\right)\Big|_{\mu_x}\left(\frac{\partial Q}{\partial y}\right)\Big|_{\mu_y}$$

Let us define

$$\sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y)$$

we have

$$\sigma_Q^2 = \sigma_x^2\left(\frac{\partial Q}{\partial x}\right)^2\Big|_{\mu_x} + \sigma_y^2\left(\frac{\partial Q}{\partial y}\right)^2\Big|_{\mu_y} + 2\sigma_{xy}\left(\frac{\partial Q}{\partial x}\right)\Big|_{\mu_x}\left(\frac{\partial Q}{\partial y}\right)\Big|_{\mu_y}$$

Some Cases
Case (i)

If x and y are uncorrelated that is $\sigma_{xy} = 0$ we have

$$\sigma_Q^2 = \sigma_x^2 \left(\frac{\partial Q}{\partial x}\right)^2\Bigg|_{\mu_x} + \sigma_y^2 \left(\frac{\partial Q}{\partial y}\right)^2\Bigg|_{\mu_y}$$

OR

$$\sigma_Q^2 = \sigma_x^2 f'^2(\mu_x) + \sigma_y^2 f'^2(\mu_y)$$

Case (ii)

If x and y are correlated that is then we have

$$\sigma_Q^2 = \sigma_x^2 \left(\frac{\partial Q}{\partial x}\right)^2\Bigg|_{\mu_x} + \sigma_y^2 \left(\frac{\partial Q}{\partial y}\right)^2\Bigg|_{\mu_y} + 2\sigma_{xy} \left(\frac{\partial Q}{\partial x}\right)\Bigg|_{\mu_x} \left(\frac{\partial Q}{\partial y}\right)\Bigg|_{\mu_y}$$

Case (iii)

If there is only one varaible involved say x then we have

$$Q_i = f(x_i)$$
$$Q = f(\mu_x)$$
$$\mu_x = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})$$

then varaince or error in the varaible is

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

then the spread in the function or error propagated in the function is given as

$$\sigma_Q^2 = \sigma_x^2 \left(\frac{\partial Q}{\partial x}\right)^2\Bigg|_{\mu_x} = \sigma_x^2 f'^2(\mu_x)$$

1.1 Some problems on Error propagation

Example Power in an electric circuit is given as $P = I^2R$ and $I = 1.0 \pm 0.1$ Ampere, $R = 10 \pm 1 Ohm$. Calculate the varaince in power using Propogation of error method.
Solution

$$\sigma_P^2 = \sigma_I^2 \left(\frac{\partial P}{\partial I}\right)^2\Bigg|_{\mu_I} + \sigma_R^2 \left(\frac{\partial P}{\partial R}\right)^2\Bigg|_{\mu_R}$$

$$\sigma_P^2 = \sigma_I^2 \left(\frac{\partial P}{\partial I}\right)^2 \Bigg|_{1.0} + \sigma_R^2 \left(\frac{\partial P}{\partial R}\right)^2 \Bigg|_{10}$$

$$\sigma_P^2 = (0.1)^2 (2IR) + (1)^2 (I^2)^2$$

$$(0.1)^2 (2.1.10)^2 + 1^2 (1^2)^2 = 5watt$$

Therfore

$$P = 10 \pm \sigma_P = 10 \pm 2$$

True power $P = 10watt$
But on repeated measurements we get with uncertainity of σ_P as per Gaussian distribution

68% measurement within $(8, 12)$ watts due to $(P \pm 1\sigma_P)$
95% measurement within $(6, 14)$ watts due to $(P \pm 2\sigma_P)$
99.7% measurement within $(4, 16)$ watts due to $(P \pm 3\sigma_P)$
Relative Error

$$\frac{\sigma_P^2}{P^2} = \frac{\sigma_I^2}{P^2} 4.I^2R^2 + \frac{\sigma_R^2}{P^2} I^2 = (0.1)^2$$

$$= 4\frac{\sigma_I^2}{I^2} + \frac{\sigma_R^2}{R^2 I^2} = 4(\frac{0.1}{1})^2 + (\frac{1}{10 * 1})^2 = (0.1)^2$$

uncertainity in power is dominated by uncertainity in current as it involved as squre so current has to be measured very carefully in order to reduce errors. *Example*
Error in measuring area of a table
Let

$$x = 95.0 \pm 0.5cm$$

$$y = 190.0 \pm 0.5cm$$

$$\sigma_x = 0.5cm$$

$$\sigma_y = 0.5cm$$

so the error in area is

$$\sigma_A^2 = \sigma_x^2 \left(\frac{\partial A}{\partial x}\right)^2 \Bigg|_{\bar{x}} + \sigma_y^2 \left(\frac{\partial A}{\partial y}\right)^2 \Bigg|_{\bar{y}}$$

Area is given as

$$A = x * y$$

$$(\frac{\partial A}{\partial x})|_{\bar{x}} = y = 190$$

$$(\frac{\partial A}{\partial y})|_{\bar{y}} = x = 95$$

therefore error in area is

$$\sigma_A^2 = \sigma_x^2 \left(\frac{\partial A}{\partial x}\right)^2 \Bigg|_{\bar{x}} + \sigma_y^2 \left(\frac{\partial A}{\partial y}\right)^2 \Bigg|_{\bar{y}}$$

$$= (0.5)^2[190^2 + 95^2] = 0.011cm$$

Example
Uncertainity in simple pendulum

$$T = 2\pi\sqrt{(\frac{L}{g})}$$

$$g = \frac{4.\pi^2 L}{T^2}$$

Therfore the error in g is given as a function of error in L and T as

$$\sigma_g^2 = \sigma_L^2 \left(\frac{\partial g}{\partial L}\right)^2\Bigg|_{\bar{L}} + \sigma_T^2 \left(\frac{\partial g}{\partial T}\right)^2\Bigg|_{\bar{T}}$$

substitute the value of a partial differentiaiton we have the Relative error as

$$\frac{\sigma_g}{g} = \sqrt{(\frac{\sigma_L^2}{L^2} + 4\frac{\sigma_T^2}{T^2}}$$

Home Work

a) If

$$a = b + c$$

show that

$$\sigma_a^2 = \sigma_b^2 + \sigma_c^2$$

Hint:

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

b) show the error in

$$a = \frac{b}{c}$$

is

$$\frac{\sigma_a^2}{a^2} = \frac{\sigma_b^2}{b^2} + \frac{\sigma_c^2}{c^2}$$

Hint:

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

c) Show the error in m in the expression

$$m = -2.5ln_{10}(\frac{F}{F_0})$$

is

$$\sigma_m^2 = (1.087)^2(\frac{\sigma_F}{F})^2$$

Hint:

$$\sigma_m^2 = \sigma_F^2 \left(\frac{\partial m}{\partial F}\right)^2_{\bar{F}}$$

where

$$\left(\frac{\partial m}{\partial F}\right)_{\bar{F}} = -\frac{2.5}{F}$$

d) If $F = x + k$ show

$$\sigma_F = \sigma_x$$

e) If $F = x * y$ show

$$(\frac{\sigma_F}{F})^2 = (\frac{\sigma_x}{x})^2 + (\frac{\sigma_y}{y})^2$$

f) If $F = kx$ show

$$(\frac{\sigma_F}{F}) = (\frac{\sigma_x}{x})$$

g) If $F = x^n$ show

$$(\frac{\sigma_F}{F})^2 = n^2(\frac{\sigma_x}{x})^2$$

1.2 True Value

Although the best value is considered and taken as mean of the observations. However even the mean value devietes from the true or exact value. That is mean value is not the final one and there is scope to improve even that. Let us have measurements of a quantity X as

$$X = X_1, X_2, X_3, X_4, ...X_n$$

AND

$$\bar{X} = \mu = \frac{1}{n}(X_1 + X_2 + X_3 + \cdots + X_n)$$

Deviation of each X_i from $\bar{X}$ is given as

$$\sigma_{X_i^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

for example

$$\sigma_{X_1^2} = \frac{1}{n}(X_1 - \bar{X})^2$$

$$\sigma_{X_2^2} = \frac{1}{n}(X_2 - \bar{X})^2$$

etc Now let us find the Deviation associated to mean value . we have seen mean is a function of $X - 1, X_2, ...X - n$ that is

$$\mu = \frac{1}{n}(X_1 + X_2 + X_3 + \cdots + X_n)$$

therfore

$$\sigma_\mu^2 = \sigma_{X_1^2}\left(\frac{\partial\mu}{\partial X_1}\right)^2 + \sigma_{X_2^2}\left(\frac{\partial\mu}{\partial X_2}\right)^2 + \cdots + \sigma_{X_1^n}\left(\frac{\partial\mu}{\partial X_n}\right)^n$$

Assume

$$\sigma_{X_1^2} = \sigma_{X_2^2} = \sigma_{X_3^2} .... = \sigma_{X_n^2} = \sigma^2$$

therfore

$$\sigma_\mu^2 = \sigma^2 \left[\left(\frac{\partial \mu}{\partial X_1} \right)^2 + \left(\frac{\partial \mu}{\partial X_2} \right)^2 + \cdots + \left(\frac{\partial \mu}{\partial X_n} \right)^n \right]$$

substitute the required value we get

$$\sigma_\mu^2 = \sigma^2 \left[\frac{1}{n^2} + \frac{1}{n^2} + \cdots + \frac{1}{n^2} \right] = n \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$\sigma_\mu = \frac{\sigma}{\sqrt{n}}$$

Hence from the above expression it is clear that there is Deviation or spread associated to the mean with an amount equal to $\sigma_\mu = \frac{\sigma}{\sqrt{n}}$, which can be improved upon if the number of observations n for the given quantity is increased. And clearly the spread or uncertainity $\sigma_\mu \to 0$ if $n \to 0$.