Curve fitting by method of least squares

Fitting of straight line

$y = ax + b$ $\qquad$ $y = a + bx$

$\Sigma y = a\Sigma x + bn$ $\qquad$ $\Sigma y = na + b\Sigma x$

$\Sigma xy = a\Sigma x^2 + b\Sigma x$ $\qquad$ $\Sigma xy = a\Sigma x + b\Sigma x^2$

Parabola. $y = ax^2 + bx + c$ $\qquad$ $y = a + bx + cx^2$

$\Sigma y = a\Sigma x^2 + b\Sigma x + nc$ $\qquad$ $\Sigma y = na + b\Sigma x + c\Sigma x^2$

$\Sigma xy = a\Sigma x^3 + b\Sigma x^2 + c\Sigma x$ $\qquad$ $\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$

$\Sigma x^2 y = a\Sigma x^4 + b\Sigma x^3 + c\Sigma x^2$ $\qquad$ $\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$

---

curve of form $y = ax^b$ take log on both sides

$\log y = \log a + b\log x \Rightarrow Y = A + BX$

$A = \log a$ $\quad \boxed{a = e^A}$ $\quad \boxed{b = B}$

$$V = \frac{\Sigma(x-\bar{x})^2}{n} \qquad V = \frac{\Sigma f(x-\bar{x})^2}{\Sigma f}$$

standard deviation (S.D) $\quad \sigma = \sqrt{V}$ or $\sigma^2 = V$

$$\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} = \sqrt{\frac{\Sigma X^2}{n}} \quad \text{where } X = x - \bar{x}$$

$$\boxed{\sigma^2 = \frac{\Sigma x^2}{n} - (\bar{x})^2}$$

or $\sigma^2 = \dfrac{\Sigma x^2}{n}$

Co-efficient of corelation

$$r = \frac{\Sigma XY}{n \sigma_x \sigma_y}$$

where $X = (x - \bar{x})$

$Y = (y - \bar{y})$

$$\boxed{r = \frac{\sigma_x^2 + \sigma_y^2 - \sigma_{x-y}^2}{2 \sigma_x \sigma_y}}$$

* $\sigma_x^2 = \frac{\Sigma x^2}{n} - (\bar{x})^2$

* $\bar{x} = \frac{\Sigma x}{n}$

* $\sigma_y^2 = \frac{\Sigma y^2}{n} - (\bar{y})^2$

* $\bar{y} = \frac{\Sigma y}{n}$

* $\sigma_{x-y}^2 = \frac{\Sigma (x-y)^2}{n} - (\overline{x-y})^2$

* $(\overline{x-y}) = \frac{\Sigma (x-y)}{n}$

Regression : estimation of one independent variable in terms of other

$$y - \bar{y} = r \frac{\sigma_y}{\sigma_x} (x - \bar{x})$$

$$x - \bar{x} = r \frac{\sigma_x}{\sigma_y} (y - \bar{y})$$

$$\boxed{\tan\theta = \left| \frac{m_1 - m_2}{1 + m_1 m_2} \right| \quad m_1 = \frac{r \sigma_y}{\sigma_x} \quad m_2 = \frac{\sigma_y}{r \sigma_x}}$$

or $r = \pm \sqrt{(\text{coef of } x)(\text{coef of } y)}$

or

$$r = \frac{n \Sigma xy - \Sigma x \Sigma y}{\sqrt{\{n \Sigma x^2 - (\Sigma x)^2\} \times \{n \Sigma y^2 - (\Sigma y)^2\}}}$$

Rank Corelation

$$P = 1 - \frac{6 \Sigma d_i^2}{(n^3 - n)}$$

Mean $\mu = \sum x_i p_i$

Variance $V = \sum x_i^2 p_i - \mu^2$

Standard Deviation $\sigma = \sqrt{V}$

Continuous probability distribution

Mean $\mu = \int_{-\infty}^{\infty} x f(x)\, dx$

Variance $V = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$

or $V = \int_{-\infty}^{\infty} x^2 f(x)\, dx - \mu^2$

S.D $\sigma = \sqrt{V}$

$f(x)$ is p.d.f if

$f(x) \geq 0$

$F(x) = P(X \leq x) = \int_{-\infty}^{\infty} f(x)\, dx$

Bionomial distribution

$p \rightarrow$ probability of success

$q \rightarrow$ probability of failure

$P(r) = {}^nC_r q^{n-r} p^r$

$r = 0, 1, \ldots n$ & $q = 1 - p$

Mean $\mu = np$

Variance $V = npq$

S.D $\sigma = \sqrt{V} = \sqrt{npq}$

Mean $\mu = \frac{\sum x_i f_i}{\sum f_i}$

$\boxed{F(x_i) = (\sum f_i) P(x_i)}$

Poisson Distribution

$P(r) = \frac{\lambda^r e^{-\lambda}}{r!}$, $r = 0, 1, 2, \ldots$

$\boxed{\lambda = np}$

Mean $\mu = \lambda$

Variance $V = \lambda$

S.D $\sigma = \sqrt{V} = \sqrt{\lambda}$

p