Methods of Mathematical Physics-I

Dipankar Bhattacharya

IUCAA-NCRA Graduate School 2021

Case where the variance σ^2 is specified.

If $p(x|I)$ is normalisable and the variance is finite then the mean will also exist. Let us assume that the mean is μ, then

$$\langle (x-\mu)^2 \rangle = \int (x-\mu)^2 p(x|I)\, dx = \sigma^2$$

We then maximise

$$Q = -\sum_i p_i \ln\left[\frac{p_i}{m_i}\right] + \lambda_0 \left(1 - \sum_i p_i\right) + \lambda_1 \left(\sigma^2 - \sum_i (x_i - \mu)^2 p_i\right)$$

Hence $$p_i = m_i \mathrm{e}^{-(1+\lambda_0)} \mathrm{e}^{-\lambda_1 (x_i - \mu)^2}$$

Which for uniform measure yields, after normalisation and setting the variance,

$$p(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] ,$$

a Gaussian distribution.

In multi-dimensions, the entropy function to be maximised may be written as

$$S(\{x\}) = -\iint \cdots \int p(\{x\}) \ln \left[\frac{p(\{x\})}{m(\{x\})} \right] \, d^N x$$

where $\quad p(\{x\}) = p(x_1, x_2, \cdots, x_N | I)$ and so on.

The case where the variances of the individual x_k are specified to be σ_k^2, constrained maximisation yields

$$p(\{x\}|\{\mu, \sigma\}) = \prod_{k=1}^{N} \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left[-\frac{(x_k - \mu_k)^2}{2\sigma_k^2} \right],$$

a product of the individual Gaussians.

If we consider x_k to be individual data points D_k, with corresponding errors σ_k and predicted model values μ_k, then the above is the same as the least-squares likelihood function relevant to minimum χ^2 fits.

Case where the l_1-norm is specified:

If, instead of variance, the l_1-norms are specified for the pdf, namely

$$\langle |x_k - \mu_k| \rangle = \iint \cdots \int |x_k - \mu_k| p(\{x\})\, d^N x = \epsilon_k$$

then the constrained maximisation of entropy function results in

$$p(\{x\}|\{\mu, \epsilon\}) = \prod_{k=1}^{N} \frac{1}{2\epsilon_k} \exp\left[-\frac{|x_k - \mu_k|}{\epsilon_k} \right]$$

At times, instead of χ^2 fitting, this likelihood function is maximised for the purpose of fitting model to data. This method goes by the name *Maximum Likelihood fit,* Absolute Deviation fit or l_1-norm fit.

Statistics of Trials

Say M trials are conducted, resulting in N successes in a given instance. Repeated conduct of such trials will result in a probability distribution of N

It is specified that the average number of successes in M trials is μ, i.e.

$$\langle N \rangle = \sum_{N=0}^{M} N p(N|M,\mu) = \mu$$

Constrained maximisation of entropy in this case would yield

$$p(N|M,\mu) \propto m(N) \mathrm{e}^{-\lambda N}$$

In this case the measure $m(N)$ is not uniform.

In any trial there are two possible outcomes - success or failure. In M trials, there would be 2^M possible outcomes. In gross ignorance, all of them will have equal probability. Now the number of different ways N successes can be achieved in M trials is

$$^{M}C_N = \frac{M!}{N!(M-N)!}$$

which must be proportional to $m(N)$.

We then have $\quad p(N|M,\mu) = A \cdot {}^{M}C_N \cdot \mathrm{e}^{-\lambda N}$

Normalising, $\quad A \sum_{N=0}^{M} {}^{M}C_N \left(\mathrm{e}^{-\lambda}\right)^N = A\left(1+\mathrm{e}^{-\lambda}\right)^M = 1$

i.e. $\quad A = \left(1+\mathrm{e}^{-\lambda}\right)^{-M}$. Then,

$$\sum_{N=0}^{M} N p(N|M,\mu) = A \sum_{N=0}^{M} N \cdot {}^{M}C_N \cdot \mathrm{e}^{-\lambda N} = \frac{M\left(1+\mathrm{e}^{-\lambda}\right)^{M-1}}{\left(1+\mathrm{e}^{-\lambda}\right)^M}\mathrm{e}^{-\lambda} = \mu$$

giving $\quad 1+\mathrm{e}^{\lambda} = M/\mu$

and $\quad p(N|M,\mu) = \dfrac{M!}{N!(M-N)!}\left(\dfrac{\mu}{M}\right)^N\left(1-\dfrac{\mu}{M}\right)^{M-N}$

which is a Binomial Distribution. In the limit $M \to \infty$ this becomes

$$p(N|\mu) \approx \frac{M^N}{N!} \cdot \frac{\mu^N}{M^N} \cdot \left(1-\frac{\mu}{M}\right)^M = \frac{\mu^N \mathrm{e}^{-\mu}}{N!}$$: The Poisson Distribution