Expectation, Variance, Skew, Kurtosis (and Covariance) are *statistical measures* which help characterize the *probability distribution* associated with a *random variable*.

Expectation (or "mean" value) $\mu = E[x]$

For many distributions, the mean value of a random variable x corresponds (or is near to) the maximum of the probability distribution plotted vs x

$$\boxed{E[x] = \mu = \sum_x xP(x)}$$

Discrete e.g. *Binomial Distribution*

$$x = 0.....N$$

$$P(x) = \frac{N!}{(N-x)!x!} p^x (1-p)^{N-x}$$

$$\boxed{E[x] = \mu = \int_{-\infty}^{\infty} xp(x)dx}$$

Continuous e.g. *Normal Distribution*

$$p(x) = \frac{\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)}{\sqrt{2\pi\sigma^2}}$$

Expectation of a *linear combination* of random variables

$$\boxed{E[aX + bY + cZ + ... + d] = aE[X] + bE[Y] + cE[Z] + ... + d}$$

Covariance is proportional to the correlation between two random variables. If it is zero, there is no correlation and the variables are said to be *independent*.

$$\boxed{\begin{aligned} &\mathrm{Cov}[X,Y] = \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] \\ &\mu_X = E[X] \\ &\mu_Y = E[Y] \\ &\mathrm{Cov}[X,Y] = \mathrm{E}[XY] - \mu_X \mu_Y \end{aligned}}$$

... Which also means

$$P(X,Y) = P(X)P(Y)$$

$$\mathrm{Cov}[X,Y] = 0$$

$$\Rightarrow \boxed{E[XY] = E[X] \times E[Y]}$$

Variance $\sigma^2 = V[x]$ is a measure of spread of a random variable about the mean $\mu = E[x]$

$$V[X] = \mathrm{E}\left[(X-\mu)^2\right] = \mathrm{E}[X^2 - 2\mu X + \mu^2]$$

$$V[X] = \mathrm{E}[X^2] - 2\mu E[X] + \mu^2$$

$$\boxed{V[X] = \mathrm{E}[X^2] - \mu^2}$$

Discrete

$$V[x] = \sum_x x^2 P(x) - \mu^2$$

$$V[x] = \int_{-\infty}^{\infty} x^2 p(x)dx - \mu^2$$

Continuous

$$\boxed{V[aX \pm bY] = a^2 V[X] + b^2 V[Y] \pm 2ab\mathrm{Cov}(X,Y)}$$

$$\boxed{\mathrm{Cov}[X,Y] = 0 \Rightarrow V[aX \pm bY] = a^2 V[X] + b^2 V[Y]}$$

In summary, if *X,Y,Z*... are *independent* (i.e. uncorrelated) random variables i.e. Cov[*X,Y*] = 0

$$\boxed{V[aX + bY + cZ + d] = a^2 V[X] + b^2 V[Y] + c^2 V[Z]}$$

If random variables *X,Y* are both *independent* and *normally distributed*, then *X+Y* is also normally distributed

$$\boxed{\begin{aligned} &X \sim \mathrm{N}(\mu_X, \sigma_X^2) \\ &Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2) \\ &\alpha X + \beta Y \sim \mathrm{N}(\alpha\mu_X + \beta\mu_Y, +\alpha^2\sigma_X^2 + \beta^2\sigma_Y^2) \end{aligned}}$$

The same is true for *aX+b*

$$\boxed{\begin{aligned} &X \sim \mathrm{N}(\mu, \sigma^2) \\ &\mu = E[X], \quad \sigma^2 = V[X] \\ &aX + b \sim \mathrm{N}(a\mu + b^2\sigma^2) \end{aligned}}$$

If two random variables are *independent*, and both derive from a *Poisson Distribution*

$$\boxed{\begin{aligned} &x \sim \mathrm{Po}(\lambda), \quad y \sim \mathrm{Po}(\mu) \\ &x + y \sim \mathrm{Po}(\lambda + \mu) \end{aligned}}$$

Proof

Start by proving that the MGF of the sum of two random variables is the product of their MGFs

$$M(x,t) = E[e^{xt}]$$

$$\boxed{M(x+y,t) = E[e^{(x+y)t}] = M(x,t)M(y,t)}$$

MGF = Moment Generating Function

The MGF for a Poisson distribution is:

$$\boxed{M(x,t) = e^{\lambda(e^t - 1)}}$$

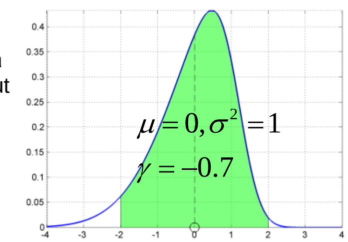
Hence: $M(x+y,t) = e^{\lambda(e^t-1)} \times e^{\mu(e^t-1)} = \boxed{e^{(\lambda+\mu)(e^t-1)}}$

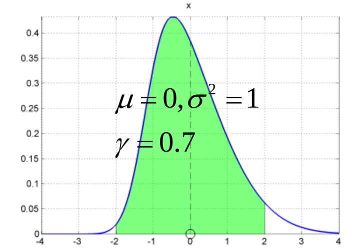
i.e. the MGF of a Poisson Distribution of mean (and variance) $\lambda + \mu$

Skew

$$\mathrm{skew}[x] = \gamma = E\left[\left(\frac{x-\mu}{\sigma}\right)^3\right] = \frac{E[x^3] - 3\mu\sigma^2 - \mu^3}{\sigma^3}$$

"Asymmetry of a distribution about the mean"





Kurtosis

$$\mathrm{kurt}[x] = E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right] - 3$$

"Fatness of tail" i.e. propensity for extreme values

$$\mathrm{kurt}[x] = \frac{E[x^4] - 4\mu E[x^3] + 6\mu^2 E[x^2] - 4\mu^3 E[x] + \mu^4}{\sigma^4} - 3$$

$$\mathrm{kurt}[x] = \frac{E[x^4] - 4\mu E[x^3] + 6\mu^2 E[x^2] - 3\mu^4}{\sigma^4} - 3$$