Distribution where the log of a variable is normally distributed. Similar in appearance to the gamma distribution.

Given $0 \le x < \infty; \quad -\infty < \mu < \infty; \quad \sigma > 0$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\frac{e^{-(\log x - \mu)^2/(2\sigma^2)}}{x}$$
$$\mu = e^{\mu + (\sigma^2/2)}$$
$$\sigma^2 = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2}$$
$$EX^n = e^{n\mu + n^2\sigma^2/2}$$

Normal (μ, σ^2)

Can be used to approximate many different kinds of distributions as their population sizes increase.

Given $-\infty < x < \infty; \quad -\infty < \mu < \infty; \quad \sigma > 0$

$$f(x) = \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma}$$
$$\mu = \mu$$
$$\sigma^2 = \sigma^2$$
$$M(t) = e^{\mu t + \sigma^2 t^2/2}$$

Paretto (α, β)

Given $a < x < \infty; \quad \alpha, \beta > 0$

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}}$$
$$\mu = \frac{\beta\alpha}{\beta-1} \quad ;\beta > 1$$
$$\sigma^2 = \frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)} \quad ;\beta > 2$$

T (v)

Given $-\infty < x < \infty; \quad v = 1, 2, 3, \ldots$

$$f(x) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})}\frac{1}{\sqrt{v\pi}}\frac{1}{(1+(\frac{x^2}{v}))^{(v+1)/2}}$$
$$\mu = 0 \quad ;v > 1$$
$$\sigma^2 = \frac{v}{v-2} \quad ;v > 2$$
$$MX^n = \begin{cases} \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{v-n}{2})}{\sqrt{\pi}\Gamma(v/2)}v^{n/2} & n < v; n \text{ is even} \\ 0 & n < v; n \text{ is odd} \end{cases}$$

$$\mathrm{T}(v) = \frac{N(0,1)}{\sqrt{\frac{\chi_v^2}{v}}}$$

Uniform (a, b)

All values between a and b are evenly distributed and x has equal chance of landing anywhere on that range.

Given $a \le x \le b$

- a is the lower bound of the distribution
- b is the upper bound
- All values between a and b are equally distributed

$$f(x) = \frac{1}{b-a}$$
$$\mu = \frac{b+a}{2}$$
$$\sigma^2 = \frac{(b-a)^2}{12}$$
$$M(t) = \frac{e^{bt} - e^{at}}{t(b-a)}$$

Weibull (γ, β)

Given $0 \le x < \infty; \quad \gamma, \beta > 0$

$$f(x) = \frac{\gamma}{\beta}x^{\gamma-1}e^{-x^\gamma/\beta}$$
$$\mu = \beta^{1/\gamma}\Gamma(1 + \frac{1}{\gamma})$$
$$\sigma^2 = \beta^{2/\gamma}(\Gamma(1 + \frac{2}{\gamma}) - \Gamma^2(1 + \frac{1}{\gamma}))$$
$$EX^n = \beta^{n/\gamma}\Gamma(1 + \frac{n}{\gamma})$$

Multivariable Distributions

Covariance and Correlation

$$\mathrm{Cov}(X, Y) = E((X - \mu_X)(Y - \mu_Y))$$
$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

Multinomial Distribution

Very similar to a binomial, except there is more than one possible outcome per trial — as compared to success or failure in the binomial.

- Taking the marginal of any of the possible outcomes results in a regular binomial

Multivariable Normal $(\vec{\mu}, \mathbf{\Sigma})$

Given $\vec{\mu} \in \Re^k; \quad \mathbf{\Sigma} \in \Re^{k^2}; \quad k \in \mathbb{N}$

$$f(\vec{x}) = (2\pi)^{-k/2}\det(\mathbf{\Sigma})^{-1/2}\exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu})^T\mathbf{\Sigma}^{-1}(\vec{x} - \vec{\mu})\right)$$
$$\vec{\mu} = \vec{\mu}$$
$$\boldsymbol{\sigma^2} = \mathbf{\Sigma}$$
$$M(\vec{t}) = \exp\left(\vec{\mu}^T\vec{t} + \frac{1}{2}\vec{t}^T\mathbf{\Sigma}\vec{t}\right)$$

Bivariate case $(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\left(\left(\frac{\vec{x}-\vec{\mu}_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{\vec{x}-\vec{\mu}_x}{\sigma_X}\right)\left(\frac{\vec{y}-\vec{\mu}_Y}{\sigma_Y}\right) + \left(\frac{\vec{y}-\vec{\mu}_Y}{\sigma_Y}\right)^2\right)\right)$$

$$\vec{\mu} = \langle\mu_X, \mu_Y\rangle$$
$$\Sigma = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

Families of Distributions

Exponential Family

Defintion

Exponential Family
Any statistical distribution or family of distributions that can fit into the form:
$f(x|\theta) = h(x)c(\theta)\exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right)$

Binomial

$$\binom{n}{x}(1-p)^n\exp\left(\log(\frac{p}{1-p})x\right)$$

Normal

$$\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\mu^2}{2\sigma^2}\right)\exp\left(-\frac{x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2}\right)$$

Location Scale Family

Def

Location Scale Family
Any statistical distribution or family of distributions that can fit into the form:
$g(x|\mu, \sigma) = \frac{1}{\sigma}f(\frac{x-\mu}{\sigma})$

Probability Inequalities

Chebychev's Inequality

Let X be a random variable and let $g(x)$ be a non-negative function. Then, for any $r > 0$

$$P(g(X) \ge r) \le \frac{Eg(X)}{r}$$

Normal Probability Inequality

With Z as a normal distribution,

$P(|Z| \ge t) \le \sqrt{\frac{e}{\pi}}\frac{e^{-t^2/2}}{t}$ for all $t > 0$

Random Samples

Properties of the sample

Mean: $\bar{X} = \frac{X_1 + \ldots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$

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

$$E(\sum_{i=1}^{n} g(X_i)) = nEg(X_1)$$
$$\mathrm{Var}(\sum_{i=1}^{n} g(X_i)) = n\mathrm{Var}g(X_1)$$

Properties of properties of the sample of random variables

$$E(\bar{X}) = \mu$$
$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$$
$$E(S^2) = \sigma^2$$

Sample distributions of common distributions

Normal (μ, σ^2)

$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
S^2 is independent from $\bar{X}$

Gamma (α, β)

$$\bar{X} \sim \mathrm{Gamma}\left(n\alpha, \frac{\beta}{n}\right)$$

Cauchy

$$\mathrm{Cauchy}(0, \sigma_1) + \ldots + \mathrm{Cauchy}(0, \sigma_n) = \mathrm{Cauchy}(0, \sum_{i=1}^{n}\sigma_i)$$

Chi-squared

$$\mathrm{Chi}(a_1) + \ldots + \mathrm{Chi}(a_n) = \mathrm{Chi}(\sum_{i=1}^{n} a_i)$$