Discrete Distributions

Bernoulli (p)

Models the probability of an event succeeding or failing, with $p = 1$ as a success and $p = 0$ as fail. There is only one trial.

Given $x = 0, 1;$ $0 \le p \le 1$

- p is the probability of getting selected trial
- Has x probability of being 1 and $1 - p$ probability of being 0

$$P(Y = x) = p^x(1-p)^{1-x}$$
$$CDF = \begin{cases} 0 & x < 0 \\ 1-p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$
$$\mu = p$$
$$\sigma^2 = p(1-p)$$
$$M(t) = (1-p) + pe^t$$

Binomial (n, p)

Models the probability of getting x successes out of n trials with p probability of each trial succeeding.

Given $x = 0, 1, 2, \ldots, n;$ $0 \le p \le 1$

- p is probability of entering a particular trial
- n is number of samples in a round of sampling
- Predicts probability of getting certain number of chosen trial in sample set

$$P(Y = x) = \binom{n}{x} p^x (1-p)^{n-x}$$
$$\mu = np$$
$$\sigma^2 = np(1-p)$$
$$M(t) = (pe^t + (1-p))^n$$

Discrete Uniform (N)

Models the chance of x as a result when all numbers from 1 to N are equally likely.

Given $x = 1, 2, \ldots, N;$ $N = 1, 2, \ldots$

- N is the largest possible sample
- All numbers from 1 to N are equally likely

$$P(Y = x) = 1/N$$
$$CDF = \frac{x}{N}$$
$$\mu = \frac{N+1}{2}$$
$$M(t) = \frac{1}{N} \sum e^{tx}$$

Geometric (p)

Models chance of taking x trials to get a success, with each success having p chance of succeeding.

Given $x = 1, 2, \ldots;$ $0 \le p \le 1$

- p is probability of getting certain trial
- Predicts number of samples needed to get a sample of particular trial

$$P(Y = x) = p(1-p)^{x-1}$$
$$CDF = \begin{cases} 0 & x < 0 \\ 1 - (1-p)^x & x \ge 0 \end{cases}$$
$$\mu = 1/p$$
$$\sigma^2 = \frac{1-p}{p^2}$$
$$M(t) = \frac{pe^t}{1-(1-p)e^t}$$

Hypergeometric (N, K, M)

Predicts likelihood of selecting x successes after K trials of selecting from a population of N with M samples within that are successes without replacement.

Given $x = 0, 1, 2, \ldots, K;$ $M - (N - K) \le x \le M;$ $N, M, K = 0, 1, 2, \ldots$

- N is the population size
- K is the number of samples in the population with a certain trait
- M number of samples taken in a round of sampling
- Predicts the likelihood of selecting x samples of type M after selecting K samples from population N

$$P(Y = x) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}$$
$$\mu = KM/N$$
$$\sigma^2 = \frac{KM(N-M)(N-K)}{N^2(N-1)}$$

Negative Binomial (r, p)

Predicts likelihood of sampling r successes of p probability after $x + r$ samples

Given $x = 0, 1, 2, \ldots;$ $0 \le p \le 1$

- p is the probability of getting a particular trait in one sample
- r is the desired number of samples with a particular trait
- Predicts number of likelihood of getting r samples of trait after $x + r$ samples

$$P(Y = x) = \binom{x + r - 1}{r - 1} p^r (1-p)^x$$
$$\mu = \frac{r(1-p)}{p}$$
$$\sigma^2 = \frac{r(1-p)}{p^2}$$
$$M(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r$$

Poisson Distribution (λ)

Assumes chances of an event happening in a short time is proportional to a large time. Predicts likelihood of r events happening in the next unit if λ events happen on average per unit where the unit does not need to be time.

- $X_1 = 0$
- For $\lambda > 0$, N and $N_1, \ldots, N$ are independent random variables
- Y_1 and $N_1, \ldots, N_1$ are the same distribution
- $\frac{Y}{N} = \lambda$ (Average rate of events will be λ everywhere)
- $\frac{Y}{N} = Y$ (Can't have two events at the same time)

Given $x = 0, 1, 2, \ldots;$ $0 \le \lambda$

- λ is the number of times on average an event will happen within an interval
- Predicts number of times an event will happen within an interval
- Approximates the Binomial Distribution

$$P(Y = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$
$$\mu = \lambda$$
$$\sigma^2 = \lambda$$
$$M(t) = e^{\lambda(e^t - 1)}$$

Continuous Distributions

Beta (α, β)

Often used to model proportions that lie on $(1, 0)$

Given $0 \le x \le 1;$ $\alpha > 0;$ $\beta > 0$

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$
$$\mu = \frac{\alpha}{\alpha+\beta}$$
$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$
$$M(t) = 1 + \sum_{k=1}^{\infty} \left(\prod \frac{\alpha+r}{\alpha+\beta+r}\right) \frac{t^k}{k!}$$

Cauchy (θ, σ)

Symmetric, bell-shaped, and has no mean, often used as an extreme edge case for statistical theory, but also represents the ratio between two normal distributions among other unsuspecting things.

Given $-\infty < x < \infty;$ $-\infty < \theta < \infty;$ $\sigma > 0$

$$f(x) = \frac{1}{\pi\sigma\left[1+\left(\frac{x-\theta}{\sigma}\right)^2\right]}$$

Chi-squared (p)

Given $0 \le x < \infty;$ $p = 1, 2, 3, \ldots$

$$f(x) = \frac{x^{p/2-1}e^{-x/2}}{\Gamma(p/2)2^{p/2}}$$
$$\mu = p$$
$$\sigma^2 = 2p$$
$$M(t) = \frac{1}{(1-2t)^{p/2}}$$

Double Exponential (μ, σ)

Given by reflecting the exponential around its mean. Gives rat have a left slope.

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

$$f(x) = \frac{1}{2\sigma}e^{-|x-\mu|/\sigma}$$
$$\mu = \mu$$
$$\sigma^2 = 2\sigma^2$$
$$M(t) = \frac{e^{\mu t}}{1-\sigma^2 t^2}$$

Exponential β

Given $0 \le x < \infty;$ $\beta > 0$

$$f(x) = \frac{1}{\beta}e^{-x/\beta}$$
$$\mu = \beta$$
$$\sigma^2 = \beta^2$$
$$M(t) = \frac{1}{1-\beta t}$$

F (ν_1, ν_2)

Given $0 < x < \infty;$ $\nu_1, \nu_2 = 1, 2, 3, \ldots$

$$f(x) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2}\frac{x^{\nu_1/2-1}}{\left(1+\frac{\nu_1}{\nu_2}x\right)^{(\nu_1+\nu_2)/2}}$$
$$\mu = \frac{\nu_2}{\nu_2-2}, \quad \nu_2 > 2$$
$$\sigma^2 = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}, \quad \nu_2 > 4$$

Gamma (α, β)

α being the shape parameter and β being the scale parameter, it is also a generalization of the exponential, erlang, weibull, and chi-squared distributions.

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

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}$$
$$\mu = \alpha\beta$$
$$\sigma^2 = \alpha\beta^2$$
$$M(t) = \frac{1}{(1-\beta t)^\alpha}$$

Logistic (μ, β)

Given $-\infty < x < \infty;$ $-\infty < \mu < \infty;$ $\beta > 0$

$$f(x) = \frac{e^{-(x-\mu)/\beta}}{\beta(1+e^{-(x-\mu)/\beta})^2}$$
$$\mu = \mu$$
$$\sigma^2 = \frac{\pi^2\beta^2}{3}$$
$$M(t) = e^{\mu t}\Gamma(1+\beta t)\Gamma(1-\beta t)$$

Lognormal (μ, σ^2)

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

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

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x}e^{-(\ln x - \mu)^2/(2\sigma^2)}$$
$$\mu = e^{\mu+\sigma^2/2}$$
$$\sigma^2 = e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$$
$$M(t) = e^{\mu t + \sigma^2 t^2/2}$$

Normal (μ, σ^2)

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

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

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

Pareto (α, β)

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

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

t (ν)

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

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\sqrt{\nu\pi}}\frac{1}{(1+x^2/\nu)^{(\nu+1)/2}}$$
$$\mu = 0, \quad \nu > 1$$
$$\sigma^2 = \frac{\nu}{\nu-2}, \quad \nu > 2$$
$$M(t) = \begin{cases} \frac{\nu}{\nu-2} & \nu > 2, \nu \text{ even} \\ 0 & \nu > 1, \nu \text{ odd} \end{cases}$$

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

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

Weibull (γ, β)

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

$$f(x) = \frac{\gamma}{\beta}x^{\gamma-1}e^{-x^\gamma/\beta}$$
$$\mu = \beta^{1/\gamma}\Gamma\left(1+\frac{1}{\gamma}\right)$$
$$\sigma^2 = \beta^{2/\gamma}\left[\Gamma\left(1+\frac{2}{\gamma}\right) - \Gamma^2\left(1+\frac{1}{\gamma}\right)\right]$$
$$M(t) = \beta^{t/\gamma}\Gamma\left(1+\frac{t}{\gamma}\right)$$

Families of Distributions

Exponential Family

Definition =

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

Definition =

Location Scale Family

Any statistical distribution or family of distributions that can fit into the form

$$\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right)$$

Probability Inequalities

Chebychev's Inequality

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

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

Normal Probability Inequality

With Z as a normal distribution,

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