

# Discrete Distributions

## Bernoulli ( $p$ )

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

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

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

$$P(X = x) = p^x(1 - p)^{1-x}$$

$$\text{CDF} = \begin{cases} 0 & x < 0 \\ 1 - p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

$$\mu = p$$

$$\sigma^2 = p(1 - p)$$

$$M(t) = (1 - p) + pe^t$$

## Binomial ( $n, p$ )

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

Given  $x = 0, 1, 2, \dots, n$ ;  $0 \leq p \leq 1$

- $p$  is probability of selecting a particular trait
- $n$  is number of samples in a round of sampling
- Predicts probability of getting certain number of chosen trait in sample set

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

$$\mu = np$$

$$\sigma^2 = np(1 - p)$$

$$M(t) = (pe^t + (1 - p))^n$$

## Discrete Uniform ( $N$ )

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

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

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

$$P(X = x) = 1/N$$

$$\text{CDF} = \frac{\lfloor x \rfloor + 1}{N}$$

$$\mu = \frac{N+1}{2}$$

$$\sigma^2 = \frac{(N+1)(N-1)}{12}$$

$$M(t) = \frac{1}{N} \sum_{i=1}^N e^{it}$$

## Geometric ( $p$ )

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

Given  $x = 1, 2, \dots$ ;  $0 \leq p \leq 1$

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

$$P(X = x) = p(1 - p)^{x-1}$$

$$\text{CDF} = \begin{cases} 1 - (1 - p)^{\lfloor x \rfloor} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\mu = 1/p$$

$$\sigma^2 = \frac{1-p}{p^2}$$

$$M(t) = \frac{pe^t}{1 - (1-p)e^t}$$

## 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  $x = 0, 1, 2, \dots, K$ ;  $M - (N - K) \leq x \leq M$ ;  $N, M, K = 0, 1, 2, \dots$

- $N$  is the population size

- $M$  is the number of samples in the population with a certain trait
- $K$  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  $N$

$$P(X = x) = \frac{\binom{M}{x} \binom{N-M}{K-x}}{\binom{N}{K}}$$

$$\mu = KM/N$$

$$\sigma^2 = \frac{KM(N-M)(N-K)}{N^2(N-1)}$$

## Negative Binomial ( $r, p$ )

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

Given  $x = 0, 1, 2, \dots$ ;  $0 \leq p \leq 1$

- $p$  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

$$P(X = x) = \binom{r+x-1}{x} p^r (1-p)^x$$

$$\mu = \frac{r(1-p)}{p}$$

$$\sigma^2 = \frac{r(1-p)}{p^2}$$

$$M(t) = \left( \frac{p}{1-(1-p)e^t} \right)^r$$

## Poisson Distribution ( $\lambda$ )

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

- $N_0 = 0$
- For  $s > 0$ ,  $N_t$  and  $N_{t+s} - N_t$  are independent random variables
- $N_s$  and  $N_{t+s} - N_t$  are the same distribution
- $\lim_{t \downarrow 0} \frac{P(N_t=1)}{t} = \lambda$  (Average rate of events will be  $\lambda$  everywhere)

- $\lim_{t \downarrow 0} \frac{P(N_t > 1)}{t} = 0$  (Can't have two events at the same time)

Given  $x = 0, 1, 2, \dots$ ;  $0 \leq \lambda$

- $\lambda$  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

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

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

## Continuous Distributions

### Beta ( $\alpha, \beta$ )

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

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

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

$$\mu = \frac{\alpha}{\alpha + \beta}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

$$M(t) = 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \prod_{r=0}^{k-1} \frac{\alpha + r}{\alpha + \beta + r}$$

### Cauchy ( $\theta, \sigma$ )

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

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

### Chi squared ( $p$ )

Given  $0 \leq x < \infty$ ;  $p = 1, 2, 3, \dots$

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

$$\mu = p$$

$$\sigma^2 = 2p$$

$$M(t) = \left(\frac{1}{1-2t}\right)^{p/2}$$

## Double Exponential $(\mu, \sigma)$

Given by reflecting the exponential around its mean. Does not have a bell shape.

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

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

$$\mu = \mu$$

$$\sigma^2 = 2\sigma^2$$

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

## Exponential $\beta$

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

$$f(x) = \frac{e^{-x/\beta}}{\beta}$$

$$\mu = \beta$$

$$\sigma^2 = \beta^2$$

$$M(t) = \frac{1}{1-\beta t}$$

## F $(v_1, v_2)$

Given  $0 \leq \infty$ ;  $v_1, v_2 = 1, 2, 3, \dots$

$$f(x) = \frac{\Gamma(\frac{v_1+v_2}{2})}{\Gamma(\frac{v_1}{2})\Gamma(\frac{v_2}{2})} \left(\frac{v_1}{v_2}\right)^{v_1/2} \frac{x^{(v_1-2)/2}}{(1+\frac{x}{v_2})^{(v_1+v_2)/2}}$$

$$\mu = \frac{v_2}{v_2-2}$$

$$\sigma^2 = 2\left(\frac{v_2}{v_2-2}\right)^2 \frac{v_1+v_2-2}{v_1(v_2-4)}$$

$$EX^n = \frac{\Gamma(\frac{v_1+2n}{2})\Gamma(\frac{v_2-2n}{2})}{\Gamma(v_1/2)\Gamma(v_2/2)} \left(\frac{v_2}{v_1}\right)^n \quad ; n < \frac{v_2}{2}$$

## Gamma $(\alpha, \beta)$

$\alpha$  being the shape parameter and  $\beta$  being the scale parameter, it is also a generalization of the exponential, erlang, weibull, and chi-squared distributions.

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

$$f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha) \beta^\alpha}$$

$$\mu = \alpha\beta$$

$$\sigma^2 = \alpha\beta^2$$

$$M(t) = \left(\frac{1}{1-\beta t}\right)^\alpha$$

## Logistic $(\mu, \beta)$

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

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

$$\mu = \mu$$

$$\sigma^2 = \frac{\pi^2 \beta^2}{3}$$

$$M(t) = e^{\mu t} \Gamma(1 + \beta t)$$

## Lognormal $(\mu, \sigma^2)$

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

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

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

$$\mu = e^{\mu + (\sigma^2/2)}$$

$$\sigma^2 = e^{2(\mu + \sigma^2)} - e^{2\mu + \sigma^2}$$

$$EX^n = e^{n\mu + n^2\sigma^2/2}$$

## Normal $(\mu, \sigma^2)$

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

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

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

$$\mu = \mu$$

$$\sigma^2 = \sigma^2$$

## Pareto ( $\alpha, \beta$ )

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

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}}$$

$$\mu = \frac{\beta\alpha}{\beta-1} \quad ; \beta > 1$$

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

## t ( $v$ )

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

$$f(x) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})} \frac{1}{\sqrt{v\pi}} \frac{1}{(1+(\frac{x^2}{v}))^{(v+1)/2}}$$

$$\mu = 0 \quad ; v > 1$$

$$\sigma^2 = \frac{v}{v-2} \quad ; v > 2$$

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

## Uniform ( $a, b$ )

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

Given  $a \leq x \leq b$

- $a$  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 $(\gamma, \beta)$

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

$$f(x) = \frac{\gamma}{\beta} x^{\gamma-1} e^{-x^\gamma/\beta}$$

$$\mu = \beta^{1/\gamma} \Gamma(1 + \frac{1}{\gamma})$$

$$\sigma^2 = \beta^{2/\gamma} (\Gamma(1 + \frac{2}{\gamma}) - \Gamma^2(1 + \frac{1}{\gamma}))$$

$$EX^n = \beta^{n/\gamma} \Gamma(1 + \frac{n}{\gamma})$$

## 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

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

### 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:

$$g(x|\mu, \sigma) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$$