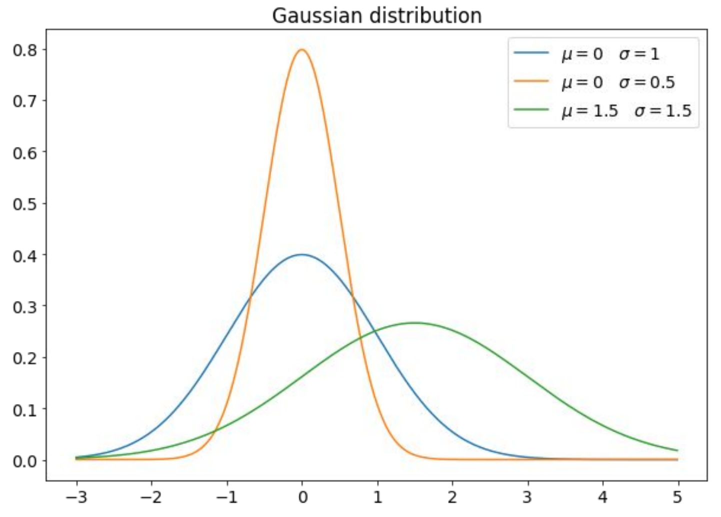


Common Distributions Cheat Sheet

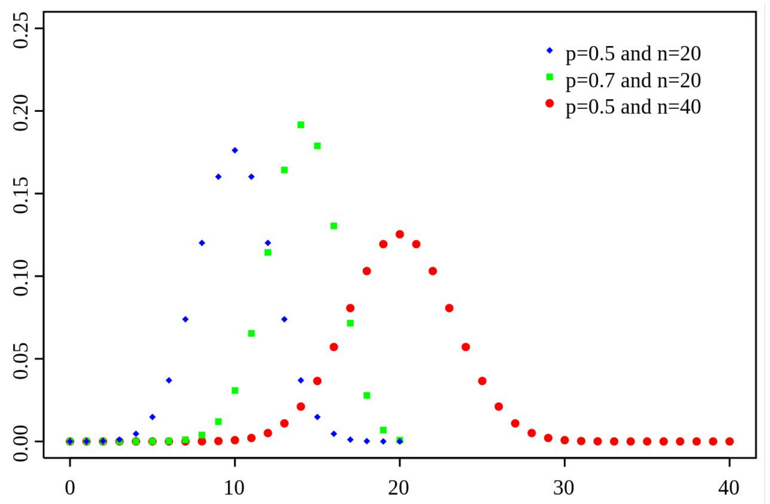
V2020.10.15

(Dr Yan Xu)

Discrete: Bernoulli (*p*)

- ➔ Variable space: {0, 1}
- ➔ Pr(***X***=1) = ***p*** = 1 - Pr(***X***=0)
- ➔ hyperparameters: ***p***
- ➔ Mean: ***p***
- ➔ Variance: ***p* * (1 - *p*)**

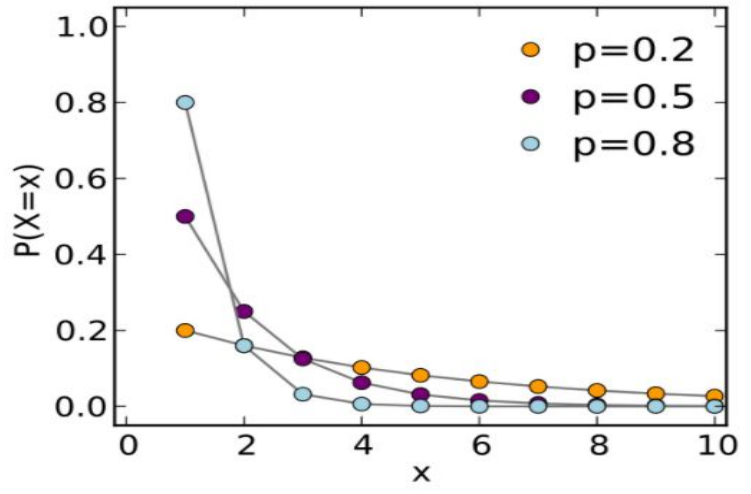
Discrete: Binomial *B(n, p)*



- ➔ Variable space: ***k*** = {1, 2, 3, … n}
- ➔ Pmf: $\binom{n}{k} p^k q^{n-k}$
- ➔ Hyperparameters: ***n, p***
- ➔ Mean: ***n*p***
- ➔ Variance: ***n* * *p* * (1 - *p*)**

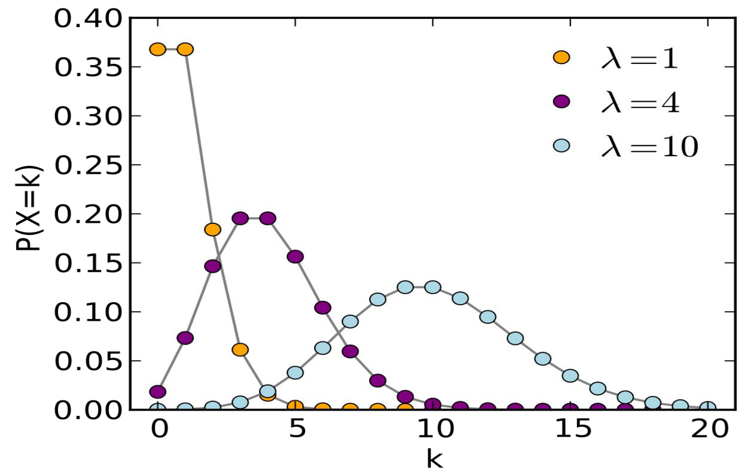
1: If n is large enough, => ***Normal(np, np(1 - p))***
2: If n is large enough and p small enough, => ***Poisson (λ = np)***, e.g. n ≥ 20 and p ≤ 0.05
3: The generalization to multiple (2+) outcomes is called a **Multinomial distribution**.
4: **Beta distribution** provides conjugate prior for **p**.

Discrete: Geometric (*p*)



- ➔ the number of Bernoulli trials needed to get one success
- ➔ Variable space: {1, 2, 3, … }
- ➔ Pmf: $(1-p)^{k-1}p$
- ➔ Hyperparameters: ***p***
- ➔ Mean: **1.0 / *p***
- ➔ Variance: **(*1* - *p*) / *p*²**

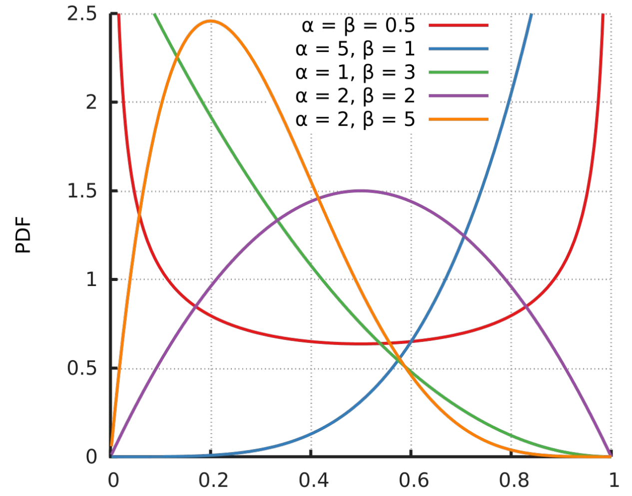
Discrete: Poisson (*λ*)



- ➔ the number of independent events occurring in a fixed interval
- ➔ Variable space: {0, 1, 2, … }
- ➔ Pmf: $\frac{\lambda^k e^{-\lambda}}{k!}$
- ➔ Hyperparameters: **λ**
- ➔ Mean: **λ**
- ➔ Variance: **λ**

1: If **λ** is large enough, => ***Normal(λ, λ)***
2: inter-arrival times **=> *Exponential*(1/*λ*)**
3: **Gamma distribution** provides conjugate prior for λ.

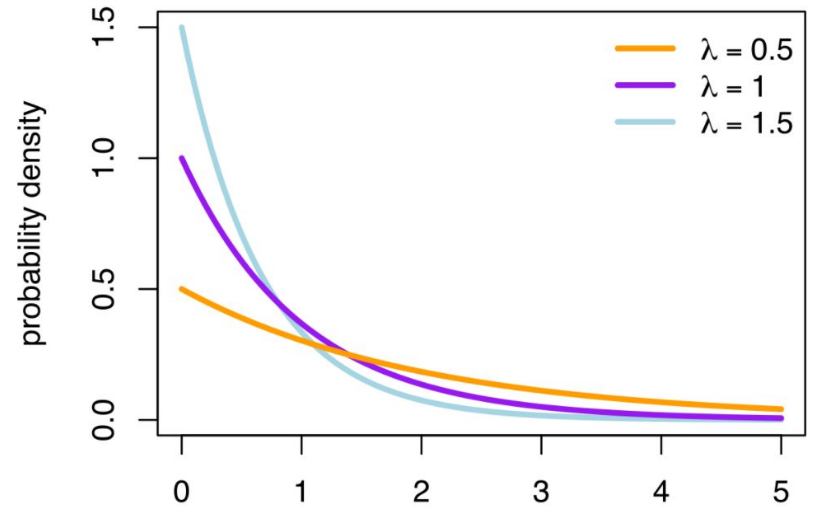
Continuous: Beta (*α, β*)



- ➔ Variable space: [0, 1]
- ➔ Pdf: $\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}$
- ➔ Hyperparameters: ***α, β***
- ➔ Mean: **α / (*α*+ *β*)**
- ➔ Variance: **α * *β* / {(*α*+ *β*)²*(*α*+ *β*+1)}**

The generalization of Beta to multiple variables is called a **Dirichlet distribution**.

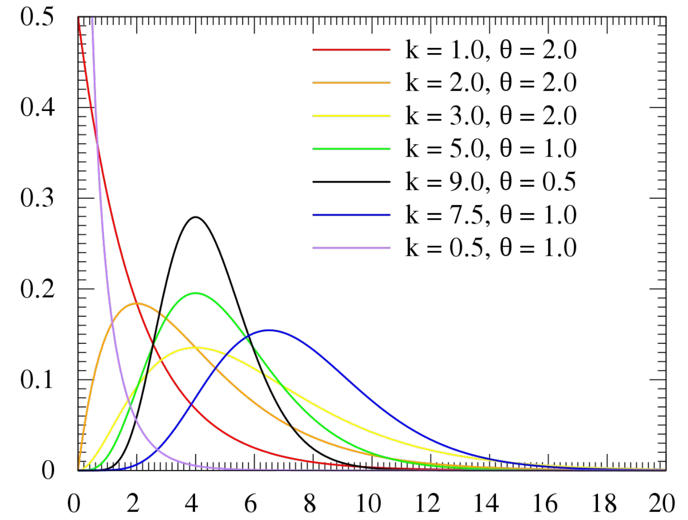
Continuous: Exponential (*λ*)



- ➔ Variable space: [0, ∞)
- ➔ Pdf: $\lambda e^{-\lambda x}$
- ➔ Hyperparameters: ***λ***
- ➔ Mean: **1.0 / *λ***
- ➔ Variance: **1.0 / *λ*²**
- ➔ Memorylessness: $\Pr(T>s+t \mid T>s) = \Pr(T>t)$

1: Exponential family: Exponential, Gamma, Normal, Chi-square, Beta, Dirichlet, Bernoulli, Possion, Geometric, etc.
2: Exponential families have conjugate priors, an important property in Bayesian statistics.

Continuous: Gamma (*α, β*)

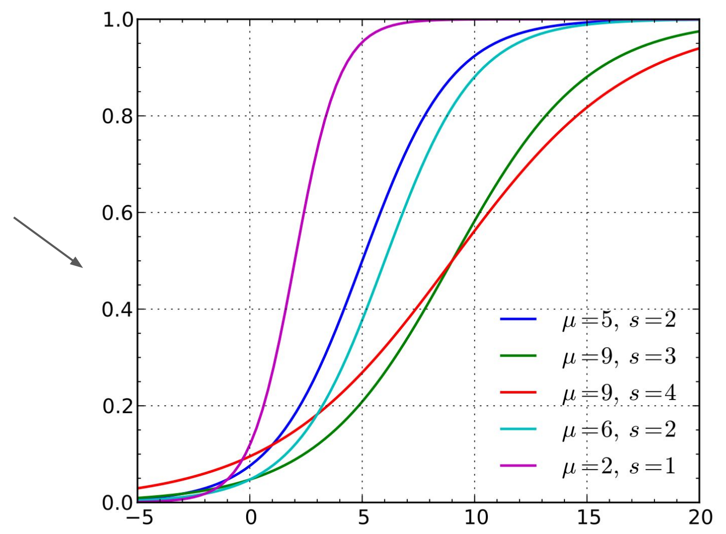


- ➔ Variable space: [0, ∞)
- ➔ Pdf: $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$
- ➔ Hyperparameters: **α** (shape), **β** (inverse scale)
- ➔ Mean: **α / β**
- ➔ Variance: **α / β²**

1: Exponential distribution (***α=1***), Erlang distribution (***α=INT***), and Chi-squared distribution (only for hypothesis testing) are special cases of the gamma distribution.
2: If k is large enough, => ***Normal(α/β, α/β²)***
3: gamma distribution is widely used as a conjugate prior in Bayesian statistics

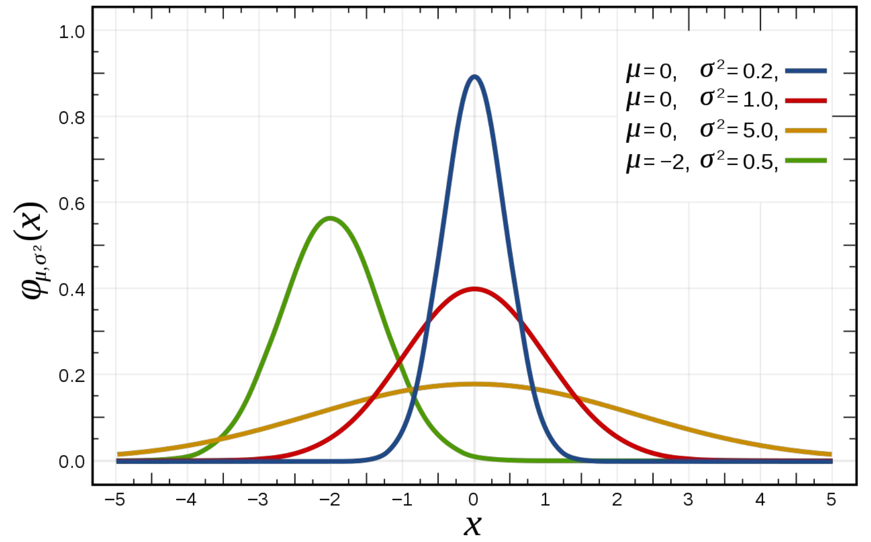
Continuous: Logistic (*μ, s*)

CDF



- ➔ Variable space: (-∞, ∞)
- ➔ Pdf: $\frac{e^{-(x-\mu)/s}}{s\left(1+e^{-(x-\mu)/s}\right)^2}$
- ➔ Cdf: $\frac{1}{1+e^{-(x-\mu)/s}}$
- ➔ Applications: Logistic regression

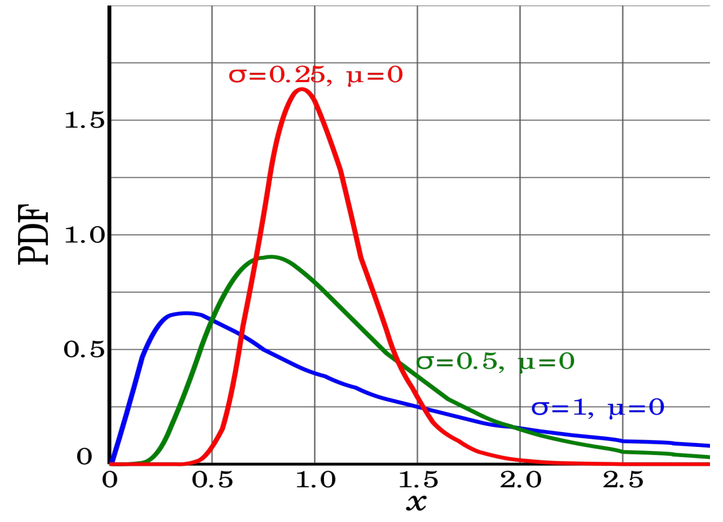
Continuous: Normal (*μ,σ²*) (or Gaussian)



- ➔ Variable space: (-∞, ∞)
- ➔ Pdf: $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
- ➔ Mean: **μ**
- ➔ Variance: ***σ*2**
- ➔ Sample variance: **n / (n - 1) * *σ*2**

1: Central limit theorem: the sum of many random variables will have an approximately normal distribution. It is the key reason why Normal is widely used in various observations.
2: If X, Y are independent Normal with means u_1, u_2 and standard deviation σ_1, σ_2, then X + Y will be normally distributed, with mean (u_1 + u_2) and variance ($\sigma_1^2 + \sigma_2^2$)
3: Bayesian analysis of normally distributed data is complicated

Continuous: Log-normal (*μ,σ²*)



- ➔ Variable space: (0, ∞)
- ➔ Pdf: $\frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right)$
- ➔ if X has a normal distribution, then the exponential function of exp(X), has a log-normal distribution.