Probability and Statistics

Cookbook

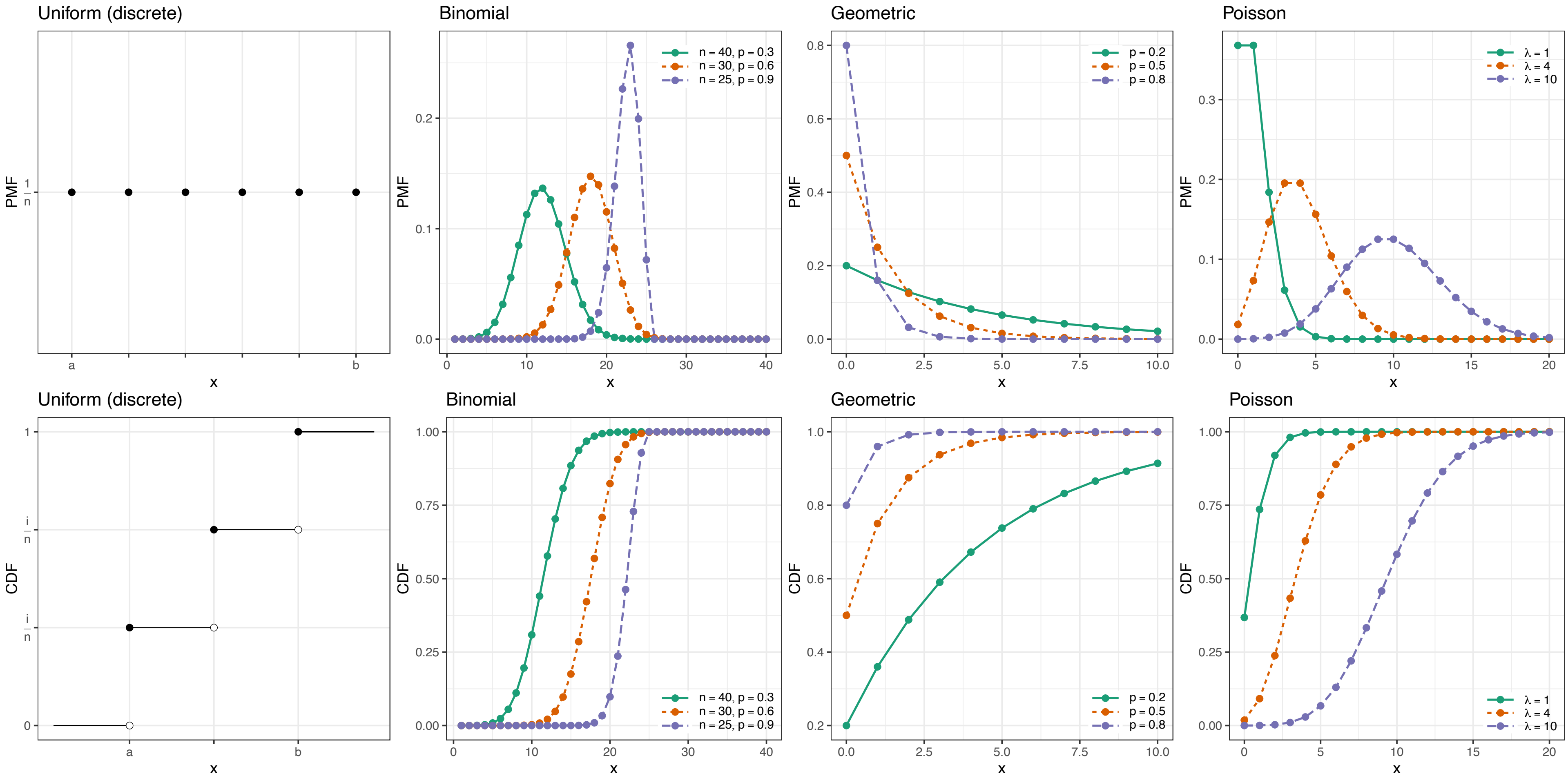
Version 0.2.7
27th July, 2021
http://statistics.zone/

1 Distribution Overview

1.1 Discrete Distributions

	Notation[1]	$F_X(x)$	$f_X(x)$	$\mathbb{E}[X]$	$\mathbb{V}[X]$	$M_X(s)$
Uniform	$\text{Unif}\{a,\ldots,b\}$	$\begin{cases} 0 & x < a \\ \frac{\lfloor x \rfloor - a + 1}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$	$\dfrac{I(a \le x \le b)}{b-a+1}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a+1)^2-1}{12}$	$\dfrac{e^{as}-e^{-(b+1)s}}{s(b-a)}$
Bernoulli	$\text{Bern}(p)$	$(1-p)^{1-x}$	$p^x(1-p)^{1-x}$	p	$p(1-p)$	$1-p+pe^s$
Binomial	$\text{Bin}(n,p)$	$I_{1-p}(n-x,x+1)$	$\binom{n}{x}p^x(1-p)^{n-x}$	np	$np(1-p)$	$(1-p+pe^s)^n$
Multinomial	$\text{Mult}(n,p)$		$\dfrac{n!}{x_1!\ldots x_k!}p_1^{x_1}\cdots p_k^{x_k} \quad \sum_{i=1}^k x_i = n$	$\begin{pmatrix} np_1 \\ \vdots \\ np_k \end{pmatrix}$	$\begin{pmatrix} np_1(1-p_1) & -np_1p_2 \\ -np_2p_1 & \ddots \end{pmatrix}$	$\left(\sum_{i=0}^k p_i e^{s_i}\right)^n$
Hypergeometric	$\text{Hyp}(N,m,n)$	$\approx \Phi\left(\dfrac{x-np}{\sqrt{np(1-p)}}\right)$	$\dfrac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$	$\dfrac{nm}{N}$	$\dfrac{nm(N-n)(N-m)}{N^2(N-1)}$	
Negative Binomial	$\text{NBin}(r,p)$	$I_p(r,x+1)$	$\binom{x+r-1}{r-1}p^r(1-p)^x$	$r\dfrac{1-p}{p}$	$r\dfrac{1-p}{p^2}$	$\left(\dfrac{pe^s}{1-(1-p)e^s}\right)^r$
Geometric	$\text{Geo}(p)$	$1-(1-p)^x \quad x \in \mathbb{N}^+$	$p(1-p)^{x-1} \quad x \in \mathbb{N}^+$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$\dfrac{pe^s}{1-(1-p)e^s}$
Poisson	$\text{Po}(\lambda)$	$e^{-\lambda}\sum_{i=0}^x \dfrac{\lambda^i}{i!}$	$\dfrac{\lambda^x e^{-\lambda}}{x!}$	λ	λ	$e^{\lambda(e^s-1)}$

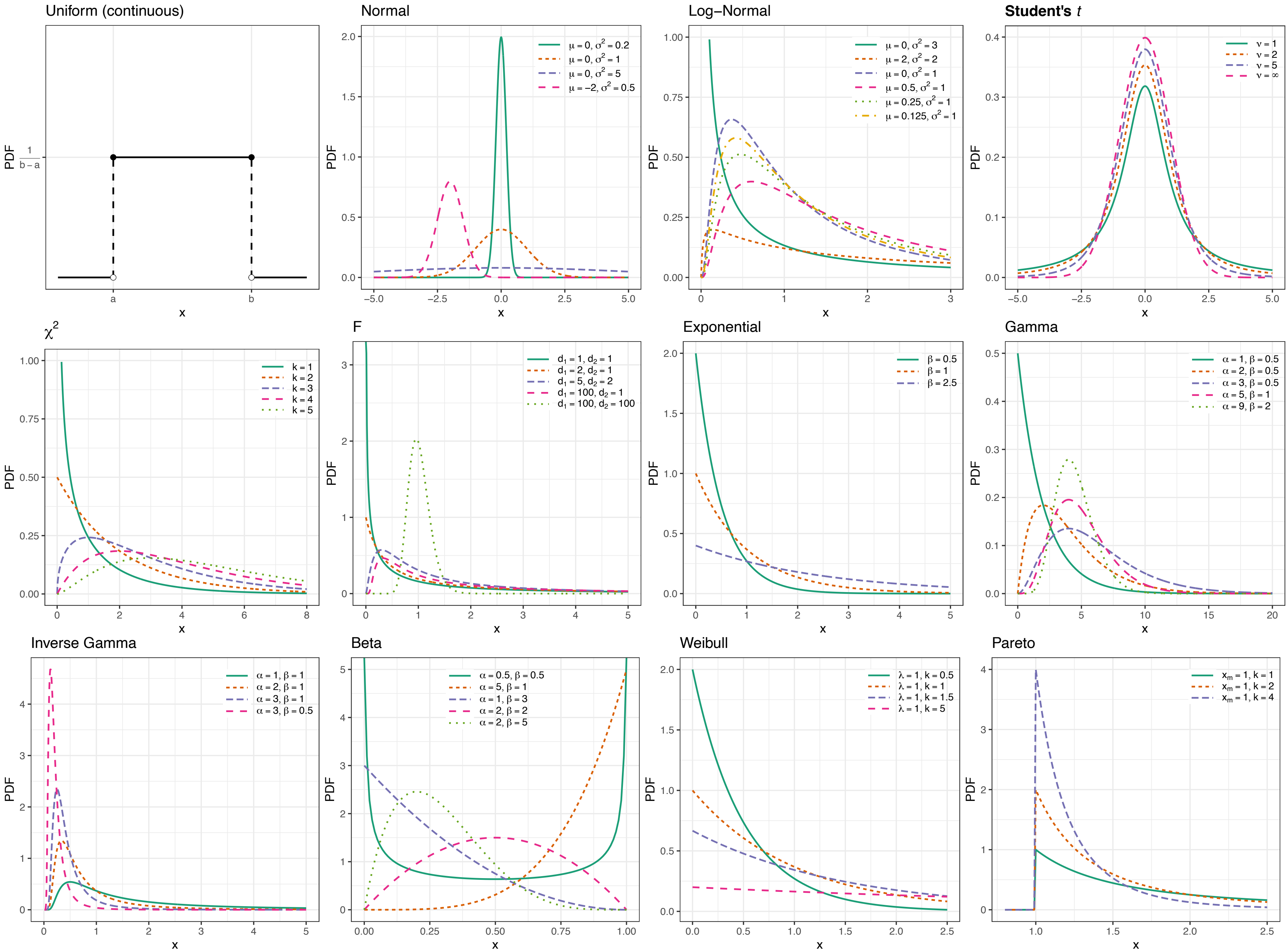
[1] We use the notation $\gamma(s,x)$ and $\Gamma(x)$ to refer to the Gamma functions (see §22.1), and use $\mathrm{B}(x,y)$ and I_x to refer to the Beta functions (see §22.2).

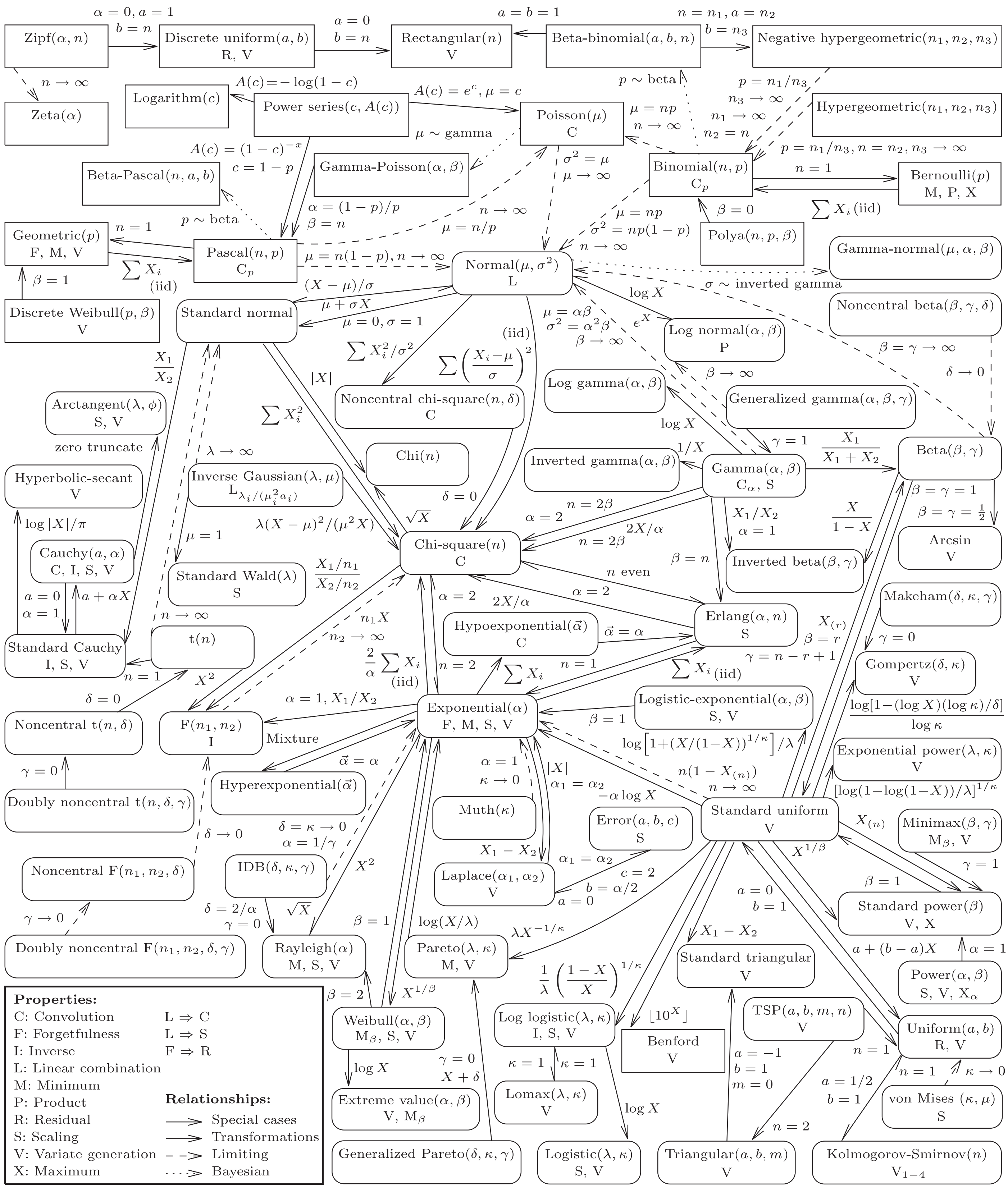

Uniform (discrete)
PMF
1/n
a
b
x
Binomial
PMF
n = 40, p = 0.3
n = 30, p = 0.6
n = 25, p = 0.9
x
Geometric
PMF
p = 0.2
p = 0.5
p = 0.8
x
Poisson
PMF
λ = 1
λ = 4
λ = 10
x
Uniform (discrete)
CDF
1
i/n
i/n
0
a
b
x
Binomial
CDF
n = 40, p = 0.3
n = 30, p = 0.6
n = 25, p = 0.9
x
Geometric
CDF
p = 0.2
p = 0.5
p = 0.8
x
Poisson
CDF
λ = 1
λ = 4
λ = 10
x

1.2 Continuous Distributions

	Notation	$F_X(x)$	$f_X(x)$	$\mathbb{E}[X]$	$\mathbb{V}[X]$	$M_X(s)$
Uniform	$\text{Unif}(a,b)$	$\begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & x > b \end{cases}$	$\dfrac{I(a<x<b)}{b-a}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	$\dfrac{e^{sb}-e^{sa}}{s(b-a)}$
Normal	$\mathcal{N}(\mu,\sigma^2)$	$\Phi(x)=\int_{-\infty}^{x}\phi(t)\,dt$	$\phi(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\}$	μ	σ^2	$\exp\left\{\mu s+\dfrac{\sigma^2 s^2}{2}\right\}$
Log-Normal	$\ln\mathcal{N}(\mu,\sigma^2)$	$\dfrac{1}{2}+\dfrac{1}{2}\operatorname{erf}\left[\dfrac{\ln x-\mu}{\sqrt{2\sigma^2}}\right]$	$\dfrac{1}{x\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right\}$	$e^{\mu+\sigma^2/2}$	$(e^{\sigma^2}-1)e^{2\mu+\sigma^2}$	
Multivariate Normal	$\text{MVN}(\mu,\Sigma)$		$(2\pi)^{-k/2}\lvert\Sigma\rvert^{-1/2}e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$	μ	Σ	$\exp\left\{\mu^T s+\dfrac{1}{2}s^T\Sigma s\right\}$
Student's t	$\text{Student}(\nu)$	$I_x\left(\dfrac{\nu}{2},\dfrac{\nu}{2}\right)$	$\dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\dfrac{x^2}{\nu}\right)^{-(\nu+1)/2}$	$0 \quad \nu>1$	$\begin{cases}\dfrac{\nu}{\nu-2} & \nu>2 \\ \infty & 1<\nu\le 2\end{cases}$	
Chi-square	χ^2_k	$\dfrac{1}{\Gamma(k/2)}\gamma\left(\dfrac{k}{2},\dfrac{x}{2}\right)$	$\dfrac{1}{2^{k/2}\Gamma(k/2)}x^{k/2-1}e^{-x/2}$	k	$2k$	$(1-2s)^{-k/2}\ s<1/2$
F	$\text{F}(d_1,d_2)$	$I_{\frac{d_1 x}{d_1 x+d_2}}\left(\dfrac{d_1}{2},\dfrac{d_2}{2}\right)$	$\dfrac{\sqrt{\frac{(d_1x)^{d_1}d_2^{d_2}}{(d_1x+d_2)^{d_1+d_2}}}}{x\mathrm{B}\left(\frac{d_1}{2},\frac{d_1}{2}\right)}$	$\dfrac{d_2}{d_2-2}$	$\dfrac{2d_2^2(d_1+d_2-2)}{d_1(d_2-2)^2(d_2-4)}$	
Exponential*	$\text{Exp}(\beta)$	$1-e^{-x/\beta}$	$\dfrac{1}{\beta}e^{-x/\beta}$	β	β^2	$\dfrac{1}{1-\frac{s}{\beta}}\ (s<\beta)$
Gamma*	$\text{Gamma}(\alpha,\beta)$	$\dfrac{\gamma(\alpha,\beta x)}{\Gamma(\alpha)}$	$\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$	$\dfrac{\alpha}{\beta}$	$\dfrac{\alpha}{\beta^2}$	$\left(\dfrac{1}{1-\frac{s}{\beta}}\right)^{\alpha}\ (s<\beta)$
Inverse Gamma	$\text{InvGamma}(\alpha,\beta)$	$\dfrac{\Gamma\left(\alpha,\frac{\beta}{x}\right)}{\Gamma(\alpha)}$	$\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{-\alpha-1}e^{-\beta/x}$	$\dfrac{\beta}{\alpha-1}\ \alpha>1$	$\dfrac{\beta^2}{(\alpha-1)^2(\alpha-2)}\ \alpha>2$	$\dfrac{2(-\beta s)^{\alpha/2}}{\Gamma(\alpha)}K_\alpha\left(\sqrt{-4\beta s}\right)$
Dirichlet	$\text{Dir}(\alpha)$		$\dfrac{\Gamma\left(\sum_{i=1}^k\alpha_i\right)}{\prod_{i=1}^k\Gamma(\alpha_i)}\prod_{i=1}^k x_i^{\alpha_i-1}$	$\dfrac{\alpha_i}{\sum_{i=1}^k\alpha_i}$	$\dfrac{\mathbb{E}[X_i](1-\mathbb{E}[X_i])}{\sum_{i=1}^k\alpha_i+1}$	
Beta	$\text{Beta}(\alpha,\beta)$	$I_x(\alpha,\beta)$	$\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\dfrac{\alpha+r}{\alpha+\beta+r}\right)\dfrac{s^k}{k!}$
Weibull	$\text{Weibull}(\lambda,k)$	$1-e^{-(x/\lambda)^k}$	$\dfrac{k}{\lambda}\left(\dfrac{x}{\lambda}\right)^{k-1}e^{-(x/\lambda)^k}$	$\lambda\Gamma\left(1+\dfrac{1}{k}\right)$	$\lambda^2\Gamma\left(1+\dfrac{2}{k}\right)-\mu^2$	$\sum_{n=0}^{\infty}\dfrac{s^n\lambda^n}{n!}\Gamma\left(1+\dfrac{n}{k}\right)$
Pareto	$\text{Pareto}(x_m,\alpha)$	$1-\left(\dfrac{x_m}{x}\right)^{\alpha}\ x\ge x_m$	$\alpha\dfrac{x_m^\alpha}{x^{\alpha+1}}\quad x\ge x_m$	$\dfrac{\alpha x_m}{\alpha-1}\ \alpha>1$	$\dfrac{x_m^2\alpha}{(\alpha-1)^2(\alpha-2)}\ \alpha>2$	$\alpha(-x_m s)^\alpha\Gamma(-\alpha,-x_m s)\ s<0$

* We use the *rate* parameterization where $\beta=\frac{1}{\lambda}$. Some textbooks use β as *scale* parameter instead [6].


Uniform (continuous)
PDF
1/(b − a)
a
b
x
Normal
μ = 0, σ² = 0.2
μ = 0, σ² = 1
μ = 0, σ² = 5
μ = −2, σ² = 0.5
Log−Normal
μ = 0, σ² = 3
μ = 2, σ² = 2
μ = 0, σ² = 1
μ = 0.5, σ² = 1
μ = 0.25, σ² = 1
μ = 0.125, σ² = 1
Student's t
ν = 1
ν = 2
ν = 5
ν = ∞
χ²
k = 1
k = 2
k = 3
k = 4
k = 5
F
d1 = 1, d2 = 1
d1 = 2, d2 = 1
d1 = 5, d2 = 2
d1 = 100, d2 = 1
d1 = 100, d2 = 100
Exponential
β = 0.5
β = 1
β = 2.5
Gamma
α = 1, β = 0.5
α = 2, β = 0.5
α = 3, β = 0.5
α = 5, β = 1
α = 9, β = 2
Inverse Gamma
α = 1, β = 1
α = 2, β = 1
α = 3, β = 1
α = 3, β = 0.5
Beta
α = 0.5, β = 0.5
α = 5, β = 1
α = 1, β = 3
α = 2, β = 2
α = 2, β = 5
Weibull
λ = 1, k = 0.5
λ = 1, k = 1
λ = 1, k = 1.5
λ = 1, k = 5
Pareto
xm = 1, k = 1
xm = 1, k = 2
xm = 1, k = 4



Univariate distribution relationships, courtesy Leemis and McQueston [2].