Chains of Distributions

Classic Random Distributions:

Normal(fixed mean, fixed s.d.)

Normal(constant mean, constant s.d.)

The parameter is a fixed number

Unorthodox Random Distributions:

Normal(varied mean, varied s.d.)

Normal(random mean, random s.d.)

The parameter itself is a random number!

Uniform(a, b)

Parameter a = 0

parameter b = Uniform(5, 27)

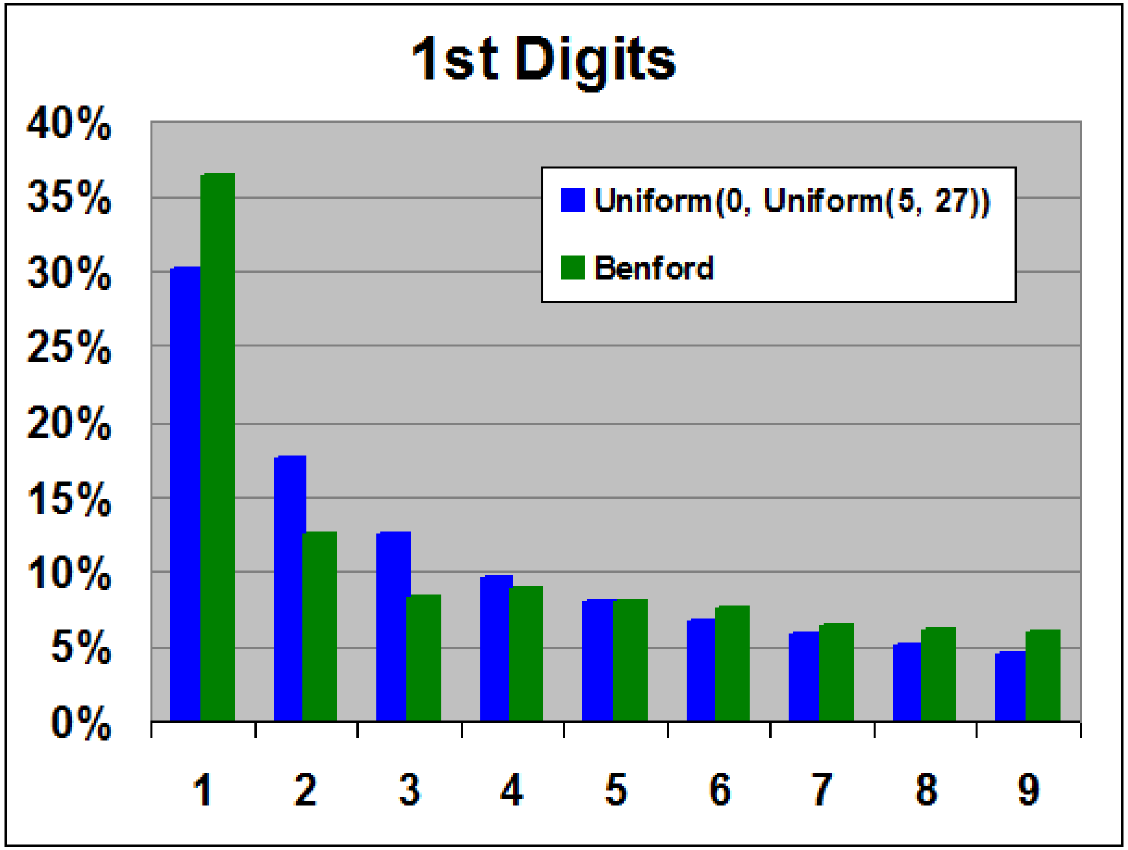
Uniform(0, Uniform(5, 27))

Uniform(0,)

↑

Uniform(5, 27)

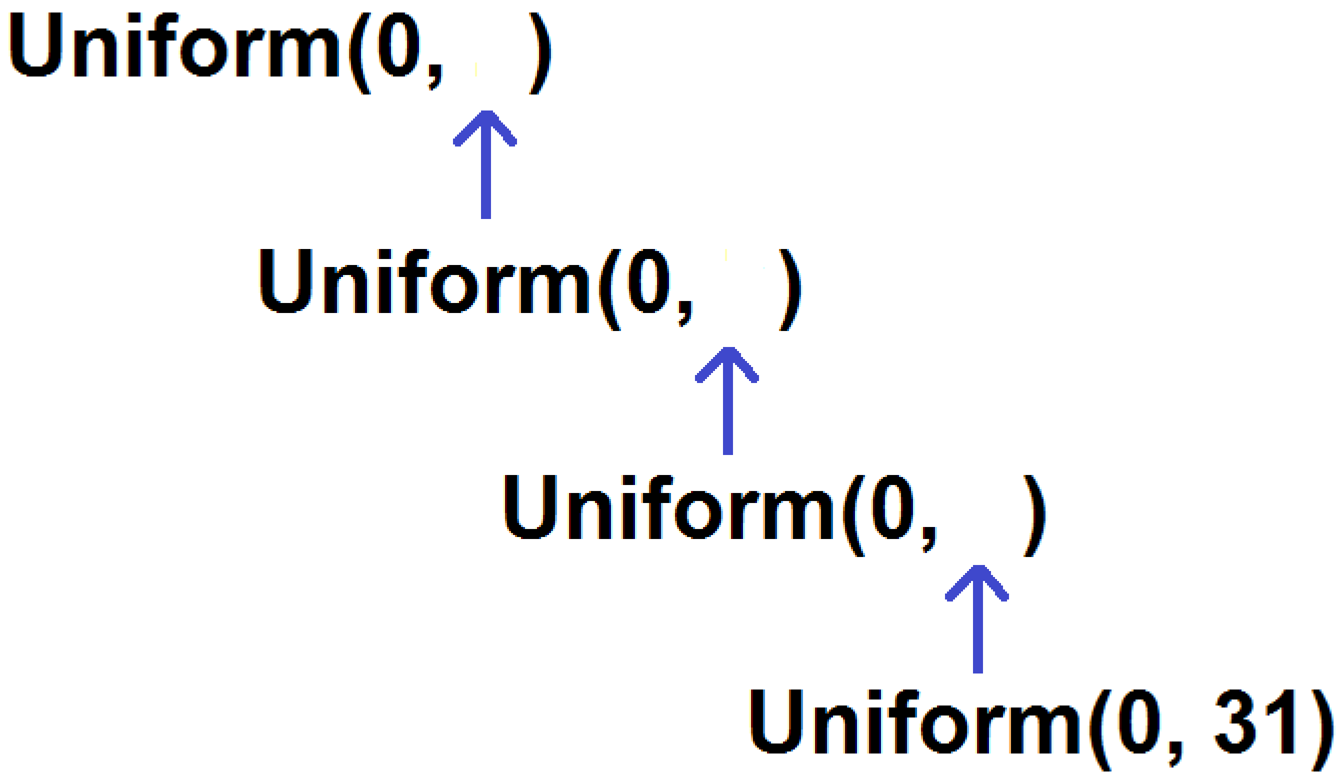
A Chain of 2 Uniform Distributions



More generally:

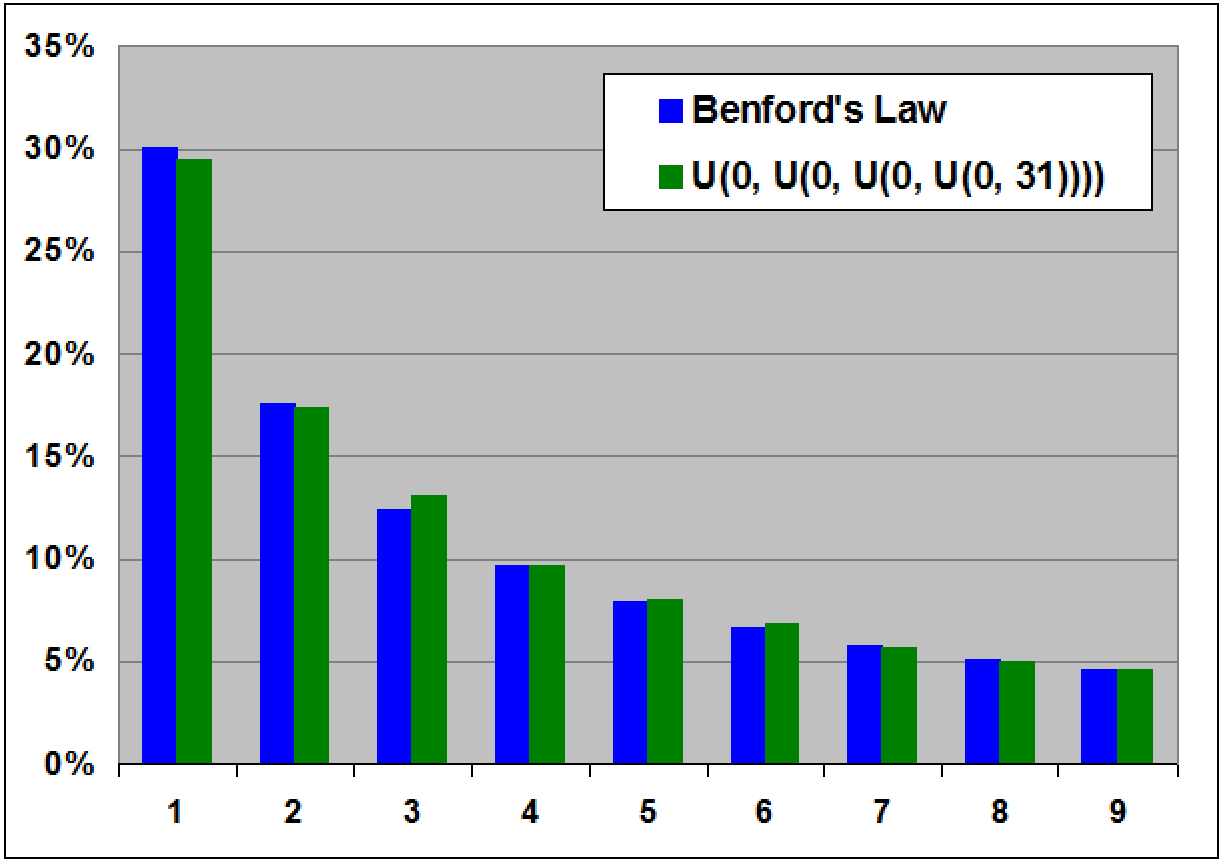
A Chain of 4 Uniform Distributions

Uniform(0, Uniform(0, Uniform(0, Uniform(0, 31))))

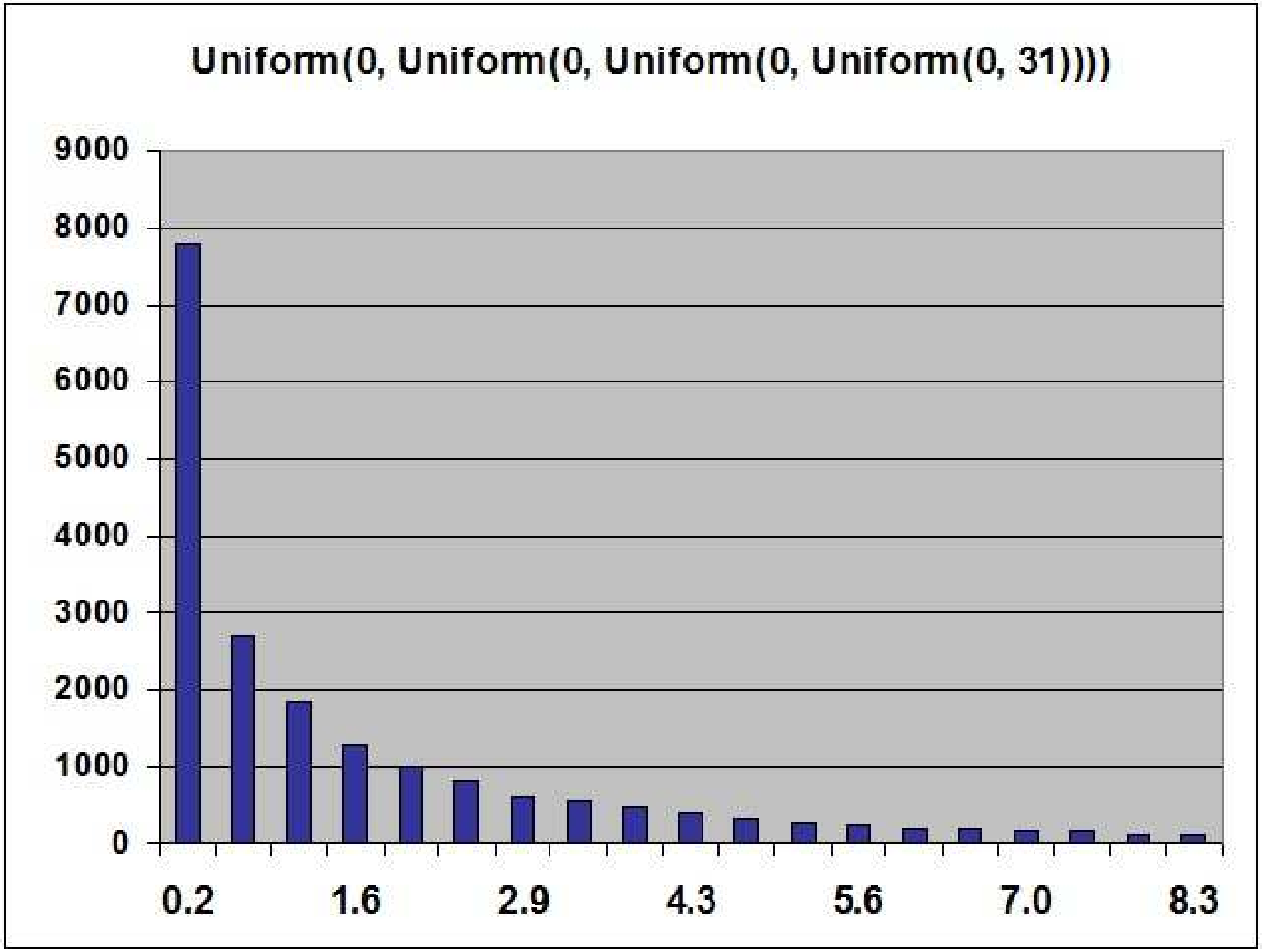

Uniform(0,)
Uniform(0,)
Uniform(0,)
Uniform(0, 31)

A Chain of 4 Uniform Distributions

20,000 simulated values

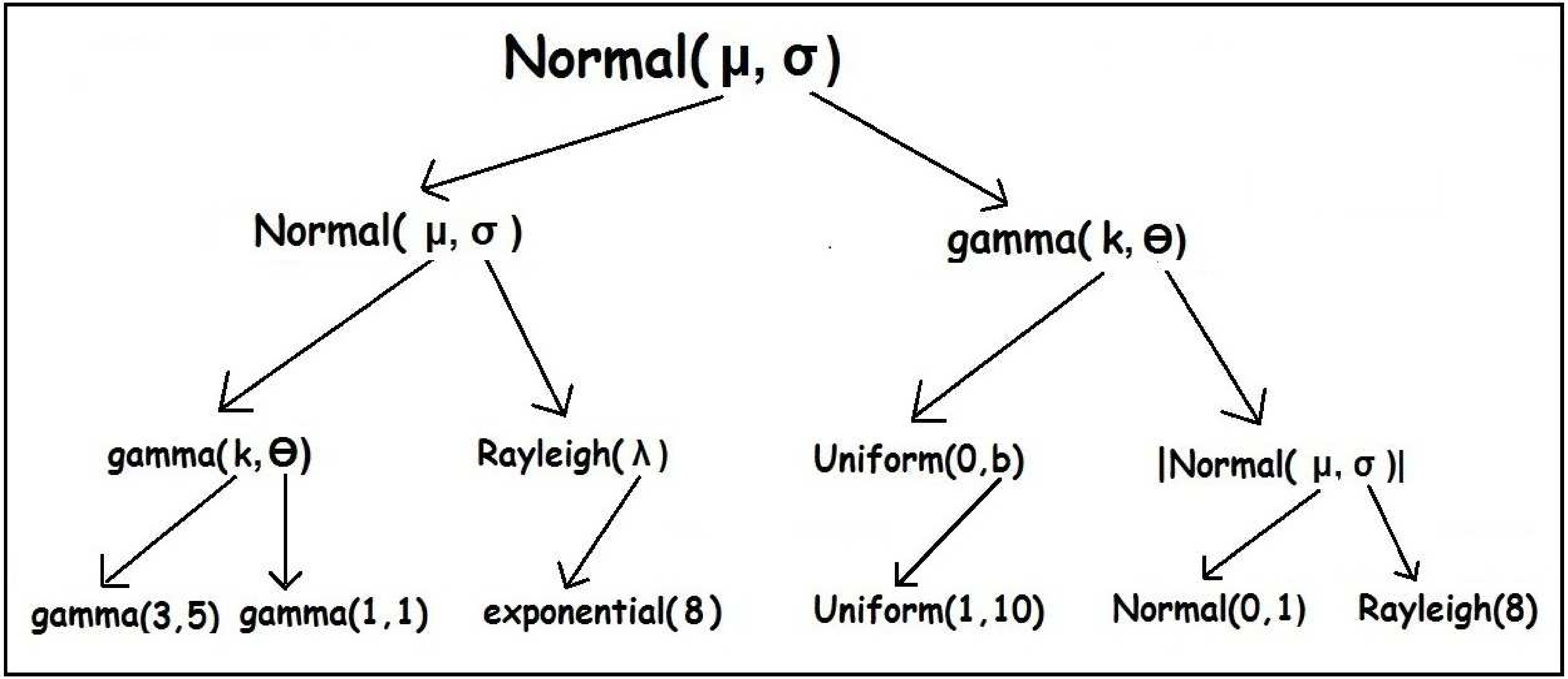


Quantitative Histogram of Chain of 4 Uniform Distributions



Small is beautiful !

The concept is by far more general, not limited to Uniform distributions. The figure below is but one manifestation of such possible constructions:



The following illustrative results demonstrate the ubiquity of the logarithmic distribution which turns up in almost all chains of distributions schemes.

===

Uniform**(*Normal*(**chi-sqr(die), Uniform(0, 3)**), *Normal*(**Uniform(77, 518), Uniform(0, 2)**))**

{33.1, 21.3, 14.0, 8.1, 5.1, 4.9, 5.0, 4.4, 4.3}
{30.1, 17.6, 12.5, 9.7, 7.9, 6.7, 5.8, 5.1, 4.6}

===

Normal(Normal(Normal(Normal(Normal(Normal(Uniform(0, 7), 13),13),13),13),13),13)

{25.0, 21.5, 16.1, 11.9, 8.9, 5.7, 4.3, 3.6, 2.6}
{30.1, 17.6, 12.5, 9.7, 7.9, 6.7, 5.8, 5.1, 4.6}

===

Normal(***Uniform***(50, 100), ***Uniform***(0, 133))

{27.0, 15.8, 12.7, 10.8, 8.5, 7.7, 6.4, 6.0, 4.8}
{30.1, 17.6, 12.5, 9.7, 7.9, 6.7, 5.8, 5.1, 4.6}

===

Normal**(*Uniform*(**0,Unif(0,Unif(0,Unif(0, 55)))**), *Uniform*(**0,Unif(0,Unif(0,Unif(0,Unif(0, 3)))**))**

{30.0, 17.5, 12.5, 10.0, 7.9, 6.6, 6.1, 4.8, 4.6}
{30.1, 17.6, 12.5, 9.7, 7.9, 6.7, 5.8, 5.1, 4.6}

===

==

Normal(0, **Uniform**(0, 3)). This is a very short chain, yet it rapidly converges!

{32.0, 19.4, 12.1, 9.5, 7.1, 5.3, 5.6, 4.7, 4.1}
{30.1, 17.6, 12.5, 9.7, 7.9, 6.7, 5.8, 5.1, 4.6}
==

Rayleigh(**Uniform**(0, **Exponential**(Rayleigh(Uniform(0, Exponential(Rayleigh… *etc.*
9 full such cycles of 3- sequence each, totaling 27sequences for the chain.

{30.1, 17.7, 12.7, 9.7, 8.0, 7.0, 5.4, 5.5, 3.9}
{30.1, 17.6, 12.5, 9.7, 7.9, 6.7, 5.8, 5.1, 4.6}

==

Weibul(***Normal(*** Wald(2, 25), Weibull(0.1, 1)**) ,**
Uniform(0, Rayleigh(Rayleigh(Weibull(Uniform(0, 65), Normal(87, 5)))**))**

{30.1, 17.2, 12.2, 9.5, 8.1, 7.3, 5.4, 5.5, 4.7}
{30.1, 17.6, 12.5, 9.7, 7.9, 6.7, 5.8, 5.1, 4.6}
==

1st Chain Conjecture:

An infinite chain of parametrical dependencies is Benford.

2nd Chain Conjecture:

AnyDensity(AnyBenford) is Benford

Namely, if the parameter is logarithmic in its own right, then the chain is also logarithmic, without any need to infinity chain more distributions.

Which distributions/parameters are chain-able, and which are not?

In extreme generality:

Scale parameters such as **λX** or **X/λ**
(divisions & multiplications)
YES!

Location parameters such as **X – μ**
(subtractions)
YES!

Shape parameters such as X^k
(powers)
NO!

parameter always affects AVG => it's chainable

parameter does NOT affect AVG in the limit => it's NOT chainable

MORE PRECISELY:

A parameter that does **not** continuously involve itself in the expression of centrality [*mean, median, midpoint, Tukey's Biweight, etc.*] is **not** chain-able at all.

'*not continuously involved*' means that

$$\lim_{parameter \to \infty} \partial(\text{center}) \, / \, \partial(\text{parameter}) = 0$$

Chain-ability of Some Distributions Defined on (0, +∞)

Distribution	pdf expression	Param 1	Param 2	both 1,2	MEAN
Weibull $(0, +\infty)$	$(k/\lambda)(x/\lambda)^{(k-1)}e^{-(x/\lambda)^k}$	λ BEN	k NOT	λ & k BEN	$\lambda*\Gamma(1+1/k)$
Rayleigh $(0, +\infty)$	$\frac{x \exp\left(\frac{-x^2}{2\sigma^2}\right)}{\sigma^2}$	BEN			$\sigma\sqrt{\frac{\pi}{2}}$
Exp 1 $(0, +\infty)$	$\rho*\exp(-\rho x)$	BEN			$1/\rho$
Exp 2 $(0, +\infty)$	$(1/\rho)*\exp(-x/\rho)$	BEN			ρ
Exp 3 $(0, +\infty)$	$(1/\sqrt{\rho})*\exp(-x/\sqrt{\rho})$	some			$\sqrt{\rho}$
Exp 4 $(0, +\infty)$	$(1/\rho^{7.5})*\exp(-x/\rho^{7.5})$	BEN			$\rho^{7.5}$
Exp 5 $(0, +\infty)$	$(1/\rho^{8})*\exp(-x/\rho^{8})$	BEN			ρ^{8}
Exp 6 $(0, +\infty)$	$(1/\log(\rho))*\exp(-x/\log(\rho))$	some			$\log(\rho)$
Gamma $(0, +\infty)$	$x^{k-1}\frac{\exp(-x/\theta)}{\Gamma(k)\,\theta^k}$	θ BEN	k some	θ & k BEN	θk
Wald $(0, +\infty)$	$\left[\frac{\lambda}{2\pi x^3}\right]^{1/2} \exp\frac{-\lambda(x-\mu)^2}{2\mu^2 x}$	μ BEN	λ NOT	μ & λ BEN	μ
LogNormal $(0, +\infty)$	$\frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln(x)-\mu]^2}{2\sigma^2}\right)$	σ NOT	μ NOT	σ & μ NOT	$\exp(\mu+\sigma^2/2)$
Uniform a = 0 (0, b) or $(0, +\infty)$	$1/b$	BEN			$b/2$
Gompertz $(0, +\infty)$	$be^{-bx}e^{-\eta e^{-bx}}\left[1+\eta\left(1-e^{-bx}\right)\right]$	B BEN	η some	b & η BEN	long & complex expression
Nakagami $(0, +\infty)$	$\frac{2\mu^\mu}{\Gamma(\mu)\omega^\mu}x^{2\mu-1}\exp\left(-\frac{\mu}{\omega}x^2\right)$	ω some	μ NOT	ω & μ some	$\frac{\Gamma(\mu+\frac{1}{2})}{\Gamma(\mu)}\left(\frac{\omega}{\mu}\right)^{1/2}$
Gupta Kundu $(0, +\infty)$	$\alpha\lambda*\exp(-\lambda x)*(1-\exp(-\lambda x))^{(\alpha-1)}$	λ BEN	α NOT	λ & α BEN	$1/\lambda*[\psi(\alpha+1)-\psi(1)]$ ψ being the digamma

Chain-ability of Some Distributions Defined on $(-\infty, +\infty)$ & $(\mu, +\infty)$

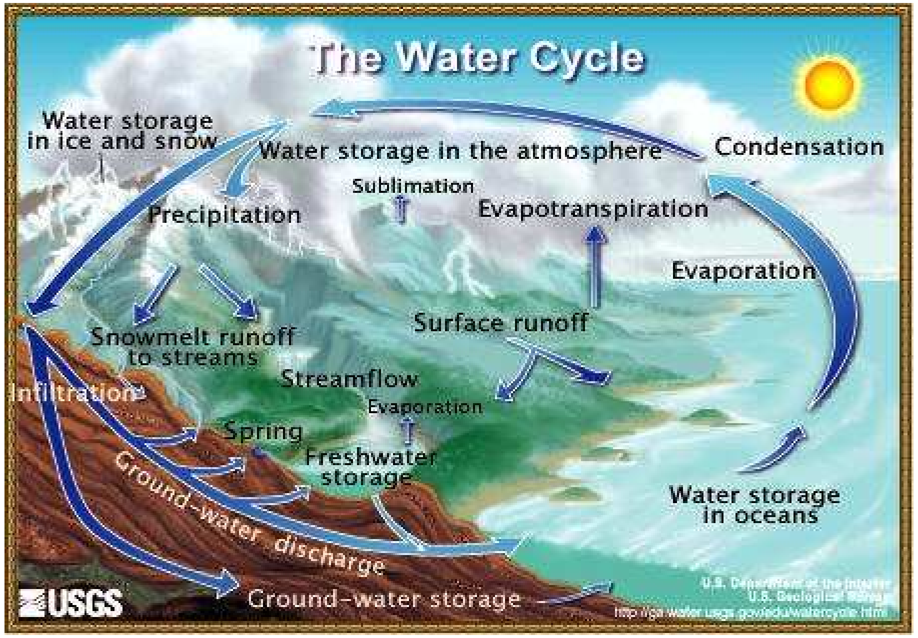
Distribution	pdf expression	Param 1	Param 2	both 1,2	MEAN
Normal $(-\infty, +\infty)$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	σ NOT	μ some	σ & μ BEN	μ
Origin-centered Normal $(-\infty, +\infty)$	$1/\sigma\sqrt{(2\pi)} * \exp(-x^2/2\sigma^2)$	BEN			0
Fisher-Tippett $(-\infty, +\infty)$	$(1/b)*e^{-(x-\mu)/\lambda}*\exp(-e^{-(x-\mu)/\lambda})$	λ NOT	μ some	λ & μ BEN	$\mu + 0.57721*\lambda$
Logistic $(-\infty, +\infty)$	$\frac{e^{-(x-\mu)/s}}{s\left(1+e^{-(x-\mu)/s}\right)^2}$	s NOT	μ some	s & μ BEN	μ
Cauchy-Lorentz $(-\infty, +\infty)$	$\frac{1}{\pi\gamma\left[1+\left(\frac{x-x_D}{\gamma}\right)^2\right]}$	γ NOT	X_D BEN (almost)	γ&X_D BEN	infinite and hence not part of our conjecture
Pareto $[a, +\infty)$	$(\theta/a)(x/a)^{-(\theta+1)}$	a BEN	θ NOT	a & θ BEN	$a*(\theta/(\theta-1))$ if $\theta>=1$ ∞ if $\theta < 1$
Generalized exp1 $[\mu, +\infty)$	$\rho*\exp(-\rho*(x-\mu))$	ρ NOT	μ some	ρ & μ BEN	$\mu + 1/\rho$
Generalized exp2 $[\mu, +\infty)$	$(1/\rho)*\exp(-(x-\mu)/\rho)$	ρ NOT	μ some	ρ & μ BEN	$\mu + \rho$

A compelling argument for the relevance of the chains of distributions to real life data could be made if we ponder the preponderance of the **causality in life, the multiple interconnectedness and dependencies** of entities that we normally measure and record; that so many measures are themselves parameters for other measures.

Exact **cause and effect relationships** in the deterministic realm such as in physics, chemistry, astronomy, and so forth, often lead to statistical and probabilistic results and scenarios especially when large aggregated phenomena are examined.

For example: **lengths and widths of rivers** depend on average rainfall (being the parameter); and rainfall in turns depends on sunspots, prevailing winds, and geographical location, all serving as parameters of rainfall.

For example: **weights of people** may depend on overall childhood nutrition; while nutrition in turns may depend on overall economic activity; which in turns depend of economic policy, war and peace, weather-related events such as droughts and flooding, and so forth.


The Water Cycle
Water storage in ice and snow
Water storage in the atmosphere
Condensation
Sublimation
Precipitation
Evapotranspiration
Evaporation
Surface runoff
Snowmelt runoff to streams
Streamflow
Infiltration
Evaporation
Spring
Freshwater storage
Water storage in oceans
Ground-water discharge
Ground-water storage
USGS
U.S. Department of the Interior
U.S. Geological Survey
http://ga.water.usgs.gov/edu/watercycle.html

END