Distribution Tool Box

N. T. Hobbs, ESS 575, Systems Ecology, Spring 2013

Discrete distributions	**Data** (y_i)	**Shape parameters**	**Moments**	**R functions**	**JAGS functions**	Conjugate relationship

Poisson
$P\left(y_i \mid \lambda\right) =$
$\frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$
	Counts of things that occur randomly over time or space, e.g., the number of birds in a forest stand, the number of fish in a km of river, the number of prey captured per minute.	λ, the mean number of occurrences per time or space	$\mu = \lambda$
			$\sigma^2 = \lambda$
				`dpois(y, lambda, log = FALSE)`
				`ppois(q, lambda)`
				`qpois(p, lambda),`
				`rpois(n, lambda)`
					`y[i] ~ dpois(lambda)`	$P(\lambda\vert\mathbf{y}) =$
						gamma $\left(\alpha + \sum\limits_{i=1}^{n} y_i, \beta + n\right)$

Binomial
$P\left(y_i \mid n, p\right) =$
$\binom{n}{p} p^{y_i}\left(1-p\right)^{n-y_i}$
$\binom{n}{p} = \frac{n!}{y_i!(n-y_i)!} p^{y_i}$
Because $\binom{n}{p}$ is a normalizing constant
$P\left(y_i \mid n, p\right) \propto$
$p^{y_i}\left(1-p\right)^{n-y_i}$
	Number of "success" on a given number of trials, e.g., number of survivors in a sample of individuals, number of plots containing an exotic species from a sample, number of terrestrial pixels that are vegetated in an image.	n, the number of trials
		p, the probability of a success
		$p = 1 - \sigma^2/\mu$
		$n = \mu^2/\left(\mu - \sigma^2\right)$
			$\mu = np$
			$\sigma^2 = np\left(1-p\right)$
				`dbinom(x, size, prob, log = FALSE)`
				`pbinom(q, size, prob)`
				`qbinom(p, size, prob)`
				`rbinom(n, size, prob)`
					`y[i] ~ dbin(p,n)`	$P(p\vert\mathbf{y}) =$
						beta $\left(\alpha + y, \beta + n - y\right)$

Bernoulli
$P\left(y_i \mid p\right) =$
$p^{y_i}\left(1-p\right)^{1-y_i}$
for $y_i \in \{0, 1\}$
	A special case of the binomial where the number of trials = 1 and the y_i can take on values 0 or 1. Widely used in survival analysis, occupancy models.	p, the probability that the random variable = 1	$\mu = p$
			$\sigma^2 = p\left(1-p\right)$
				`dbinom(x, size=1, prob, log = FALSE)`
				`pbinom(q, size=1, prob)`
				`qbinom(p, size=1, prob)`
				`rbinom(n, size=1, prob)`
				Note that size *must* = 1.
					`y[i]~dbern(p)`	

Multinomial $P\left(\mathbf{y} \mid \mathbf{p}, n\right) =$ $n! \prod_{i=1,k} \frac{p_i^{y_i}}{y_i}$ $\mathbf{y}$ and $\mathbf{p}$ are vectors.	Counts that fall into > 2 categories, so that the y must be represented as a vector of counts. e.g., number individuals in age classes, number of pixels in different landscape categories, number of species in trophic categories in a sample from a food web.	$\mathbf{y}$ a vector giving the number of counts in each category, $\mathbf{p}$ a vector of the probabilities of occurrence in each category $\sum_{i=1,k} p_i = 1$ $\sum_{i=1,k} y_i = n$	$\mathrm{E}\left[y_i\right] = np_i$ $\mathrm{Var}[y_i] =$ $np_i\left(1-p_i\right)$	`rmultinom(n, size, prob)` `dmultinom(x, size, prob, log = FALSE)`	`y[i,]~dmulti(p[],n)`	

Continuous Distributions	**Data (y_i)**	**Shape parameters**	**Moments**	**R functions**	**JAGS function**	**Conjugate prior for**	**Vague Prior**

Normal
$P(y_i \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}$
	Continuously distributed quantities that can take on positive or negative values. Also applied to strictly positive values when tail of distribution has low probability of overlapping 0. Sums of things.	μ, σ	μ, σ^2	`dnorm(x, mean, sd, log = FALSE)`
				`pnorm(q, mean, sd)`
				`qnorm(p, mean, sd)`
				`rnorm(n, mean, sd)`
					`# tau = 1/sigma^2#`
					`#likelihood`
					`y[i]~dnorm(mu,tau)`
					`#prior`
					`theta ~ dnorm(mu,tau)`
						normal mean (with known variance)	`dnorm(0,1E-6)`
							#This is scale dependent. The larger the parameter value, the smaller tau must be to make the prior uninformative.

Lognormal
$P(y_i \mid \alpha, \tau)$
$\frac{1}{y_i\sqrt{2\pi\beta^2}} e^{-\frac{(\ln y_i-\alpha)^2}{2\beta^2}}$
	Continuously distributed quantities with positive values. Data that have the property that their logs are normally distributed. Thus if z is normally distributed then $\exp(z)$ is lognormally distributed. Represents products of things. The variance increases with the mean squared.	α, the mean of y_i on the log scale
		β, the standard deviation of y_i on the log scale
		$\alpha = \log[\text{median}(y_i)]$
		$\alpha = \ln(\mu) - 1/2\ \ln\left(\frac{\sigma^2+\mu^2}{\mu^2}\right)$
		$\beta = \sqrt{\ln\left(\frac{\sigma^2+\mu^2}{\mu^2}\right)}$
			$\mu = e^{\alpha+\frac{\beta^2}{2}}$
			$\text{median}(y_i) = e^{\alpha}$
			$\sigma^2 = \left(e^{\beta^2}-1\right)e^{2\alpha+\beta^2}$
				`dlnorm(x, meanlog, sdlog)`
				`plnorm(q, meanlog, sdlog)`
				`qlnorm(p, meanlog, sdlog)`
				`rlnorm(n, meanlog, sdlog)`
					`#likelihood`
					`y[i]~dlnorm(alpha,tau)`
					`#prior`
					`theta~ dlnorm(alpha,tau)`		

Gamma
$P(y_i \mid \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} y_i^{\alpha-1} e^{-\beta y_i}$
$\Gamma(a) = \int_0^{\infty} t^{\alpha-1} e^{-t}\, dt\,.$
	Any continuous data that are strictly positive.	$\alpha =$ shape
		$\beta =$ rate
		$\alpha = \frac{\mu^2}{\sigma^2}$
		$\beta = \frac{\mu}{\sigma^2}$
		Note–be very careful about rate, defined as above, and scale $= \frac{1}{\beta}$.
			$\mu = \frac{\alpha}{\beta}$
			$\sigma^2 = \frac{\alpha}{\beta^2}$
				`dgamma(x, shape, rate, log = FALSE)`
				`pgamma(q, shape, rate)`
				`qgamma(p, shape, rate)`
				`rgamma(n, shape, rate)`
					`#likelihood`
					`y[i]~ dgamma(r,n)`
					`#prior`
					`theta~dgamma(r,n)`
						1) Poisson mean
						2) normal precision (1/variance)
						3) n parameter (rate) in the gamma distribution
							`dgamma(.001,.001)`

Beta
$P(y_i \mid \alpha, \beta) = B\, y_i^{\alpha-1}(1-y_i)^{\beta-1}$
$B = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$
Because B is a normalizing constant,
$P(y_i \mid \alpha, \beta) \propto y_i^{\alpha-1}(1-y_i)^{\beta-1}$
	Continuous data between 0 and 1–any data that can be expressed as a proportion; survival, proportion of landscape invaded by exotic, probabilities of transition from one state to another.	$\alpha = \frac{(\mu^2-\mu^3-\mu\sigma^2)}{\sigma^2}$
		$\beta = \frac{\mu-2\mu^2+\mu^3-\sigma^2+\mu\sigma^2}{\sigma^2}.$
			$\mu = \frac{\alpha}{\alpha+\beta}$
			$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
				`dbeta(x, shape1, shape2, log = FALSE)`
				`pbeta(q, shape1, shape2, )`
				`qbeta(p, shape1, shape2, )`
				`rbeta(n, shape1, shape2)`
					`#likelihood`
					`y[i] ~ dbeta(alpha, beta)`
					`#prior`
					`theta ~ dbeta(alpha, beta)`
						p in binomial distribution	`dbeta(1,1)`

Uniform
$P(y_i \mid a, b) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b, \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$
	Any real number.	$a =$ lower limit
		$b =$ upper limit
		$a = \mu - \sigma\sqrt{3}$
		$b = \mu + \sigma\sqrt{3}$
			$\mu = \frac{a+b}{2}$
			$\sigma^2 = \frac{(b-a)^2}{12}$
				`dunif(x, min, max,log = FALSE)`
				`punif(q, min, max)`
				`qunif(p, min max)`
				`runif(n, min, max)`
					`#prior`
					`theta~dunif(a,b)`
							a and b such that posterior is “more than entirely” between a and b