Named Probability Distributions

Binomial – $Y \sim \text{Binom}(n, p)$

$p(y) = \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}$, $y \in [0, n]$, $p \in [0, 1]$
$\mathbb{E}[Y] = np$
$\mathbb{V}[Y] = np(1-p)$

Geometric – $Y \sim \text{Geom}(p)$

$p(y) = (1-p)^{y-1} p$, $y \in [1, \infty)$, $p \in [0, 1]$
$\mathbb{E}[Y] = 1/p$
$\mathbb{V}[Y] = (1-p)/p^2$

Hypergeometric – $Y \sim \text{HG}(N, K, n)$

$p(y = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$, $k \in \{\max(0, n + K - N), \ldots, \min(n, K)\}$, $K \leq N; n \leq N$
$\mathbb{E}[Y] = \frac{nK}{N}$

Negative Binomial – $Y \sim \text{NBinom}(r, p)$

$P(Y = k) = \binom{r+k-1}{r-1} p^r (1-p)^k$, $k \in [r, \infty)$, $r \in \mathbb{Z}^+$, $p \in [0, 1]$
$\mathbb{E}[Y] = rq/p$
$\mathbb{V}[Y] = rq/p^2$

Poisson – $Y \sim \text{Poi}(\lambda)$

$p(y) = \frac{\lambda^y}{y!} e^{-\lambda}$, $y \in [0, \infty)$;
$\mathbb{E}[Y] = \mathbb{V}[Y] = \lambda$

Uniform – $Y \sim \text{Uniform}(a, b)$

$f(y) = (b-a)^{-1}$, $y \in [a, b]$
$\mathbb{E}[Y] = (a+b)/2$
$\mathbb{V}[Y] = (b-a)^2/12$

Normal – $Y \sim \text{N}(\mu, \sigma^2)$

$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}$ $y \in (-\infty, \infty)$, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$
$\mathbb{E}[Y] = \text{mode}[Y] = \mu$;
$\mathbb{V}[Y] = \sigma^2$

Exponential – $Y \sim \text{Exponential}(\lambda)$

$f(y) = \lambda e^{-\lambda y}$, $y \in [0, \infty)$, $\lambda \in \mathbb{R}^+$
$\mathbb{E}[Y] = 1/\lambda$
$\mathbb{V}[Y] = 1/\lambda^2$

Gamma – $Y \sim \text{Gamma}(s, r)$

$f(y) = \frac{r^s}{\Gamma(s)} y^{s-1} e^{-ry}$ for $y > 0$
$\Gamma(s) = \int_0^\infty z^{s-1} e^{-z} dz = (s-1)\Gamma(s-1)$
If s is a positive integer, $\Gamma(s) = (s-1)!$
$\mathbb{E}[Y] = \frac{s}{r}$
mode[Y] = $\frac{s-1}{r}$ for $s \geq 1$
$V[Y] = \frac{s}{r^2}$
$s = 1 \Rightarrow$ **exponential distribution**
$r = 1/2, s = \nu/2, \nu \in \mathbb{Z}^+ \Rightarrow$ **chi-square distribution**

Beta – $Y \sim \text{Beta}(\alpha, \beta)$

$f(y) = y^{\alpha-1}(1-y)^{\beta-1}/B(\alpha, \beta)$, $y \in [0, 1]$, $\alpha \in \mathbb{R}^+, \beta \in \mathbb{R}^+$
$B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$
$\mathbb{E}[Y] = \alpha/(\alpha+\beta)$
mode[Y] = $\frac{\alpha-1}{\alpha+\beta-2}$ for $\alpha, \beta > 1$
$V[Y] = \alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$

T – $Y \sim \text{T}(\nu)$

$f(y) = \Gamma(\frac{\nu+1}{2})/(\sqrt{\nu\pi}\Gamma(\nu/2))(1 + \frac{x^2}{\nu})^{-(\nu+1)/2}$, $y \in (-\infty, \infty)$, $\nu > 0$
$E[Y] = 0 (\nu > 1)$
$V[Y] = \frac{\nu}{\nu-2} (\nu > 2), \infty (1 < \nu \leq 2)$

Conjugate Models

Beta-Binomial

$\pi \sim \text{Beta}(\alpha, \beta)$
$Y|\pi \sim \text{Binomial}(n, \pi)$
$\pi|Y \sim \text{Beta}(\alpha + y, \beta + n - y)$

Gamma-Poisson

$\lambda \sim \text{Gamma}(s, r)$
$Y_i|\lambda \sim \text{Poisson}(\lambda)$
$\lambda|Y_1, Y_2, \ldots, Y_n \sim \text{Gamma}(s + \sum Y_i, r + n)$

Normal-Normal

$\mu \sim N(\theta, \tau^2)$
$Y_i|\mu \overset{iid}{\sim} N(\mu, \sigma^2)$
$\mu|\vec{y} \sim N\left(\frac{\theta\sigma^2/n + \bar{y}\tau^2}{\tau^2+\sigma^2/n}, \frac{\tau^2\sigma^2/n}{\tau^2+\sigma^2/n}\right)$

Using R

You can use R to compute probabilities for common discrete distributions. The basic commands for a distribution `xxx` with parameter values `...` are

`dxxx(x,...)` computes the probability mass function $p(x) = P(X = x)$ at the value x

`pxxx(x,...)` computes the cumulative distribution function $P(X \leq x)$ at the value x

`qxxx(p,...)` computes the pth quantile of distribution 'xxx'. (ie the value of x that makes $P(X \leq x) = p$ true)

`rxxx(n,...)` generates n (independent) values from the distribution 'xxx'

Discrete Distributions

Binomial(n,p) → `binom(x, n, p)`

Geometric(p) → `geom(x-1, p)` where x-1 is number of failures in x trials

Negative Binomial(r,p) → `nbinom(x-r, r, p)` where x-r is the number of failures in x trials

Hypergeometric(M, N, n) → `hyper(x, M, N, n)`

Poisson(λ) → `pois(x, lambda)`

Continuous Distributions

Beta(a, b) → `beta(x, a, b)`

Exponential(λ) → `exp(x, lambda)`

Gamma(s, r) → `gamma(x, s, r)`

Normal(μ, σ^2) → `norm(x, mean, sd)` (make sure to use sd instead of variance)

Uniform(a,b) → `unif(x, a, b)`