Probability Distributions

1 Binomial (n, p), $(x = 0, 1, \ldots, n)$ R:$x$binom

$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$

$M(t) = (pe^t + 1 - p)^n$

$E(X) = np$, $\text{Var} X = np(1-p)$

NOTE: $n \to \infty$ with $\mu = np \Rightarrow \text{Poisson}(\mu)$
$F_{\text{Bin}(n,p)}(x) = F_{\text{Beta}(n-\lfloor x \rfloor, \lfloor x \rfloor + 1)}(1-p)$
$F_{\text{Bin}(n,p)}(r-1) = 1 - F_{\text{NegBin}(r,p)}(n-r)$
$X_i \sim \text{Bin}(n_i, p) \Rightarrow \sum X_i \sim \text{Bin}\left(\sum n_i, p\right)$

2 Geometric (p) R:$x$geom

$f(x) = p(1-p)^x$, $(x = 0, 1, \ldots)$

$F(t) = 1 - (1-p)^{\lfloor t \rfloor + 1}$, where $\lfloor t \rfloor = \max\{m \in \mathbb{Z} : m \le t\}$

$S(t) = P[X \ge t] = (1-p)^{\lceil t \rceil}$ with $\lceil t \rceil = \min\{m \in \mathbb{Z} : m \ge t\}$

$R(t) = P[X > t] = (1-p)^{\lfloor t \rfloor + 1}$

$Q(t) = \inf\{x : F(x) \ge y\} = \max\{\lceil \log(1-y)/\log(1-p) - 1 \rceil, 0\}$

$M_X(t) = \dfrac{p}{1 - (1-p)e^t}$, $(t < -\log(1-p))$

$E(X) = \dfrac{1-p}{p}$, $\text{Var} X = \dfrac{1-p}{p^2}$

$f(y) = p(1-p)^{y-a}$, $(y = a, a+1, a+2 \ldots)$

$F(t) = 1 - (1-p)^{\lfloor t \rfloor + 1 - a}$

$S(t) = P[Y \ge t] = (1-p)^{\lceil t \rceil - a}$

$R(t) = P[Y > t] = (1-p)^{\lfloor t \rfloor + 1 - a}$

$Q(t) = \max\{\lceil \log(1-y)/\log(1-p) - 1 + a \rceil, a\}$

$M_Y(t) = \dfrac{pe^{at}}{1 - (1-p)e^t}$, $(t < -\log(1-p))$

$E(Y) = \dfrac{1-p}{p} + a$, $\text{Var}(Y) = \dfrac{1-p}{p^2}$

NOTE: $\min\limits_{1 \le i \le n}(X_i) \sim \text{Geo}\left(1 - (1-p)^n\right)$: self-reproducing

3 Hypergeometric (N, M, n) R:$x$hyper

$f(x) = \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$,

$(\max(0, M - (N-n)) \le x \le \min(n, M))$

$E(X) = n\left(\dfrac{M}{N}\right)$, $\text{Var}(X) = n\dfrac{M}{N}\dfrac{(N-M)(N-n)}{N(N-1)}$

4 Negative Binomial (r, p) R:$x$nbinom

$f(x) = \binom{r+x-1}{x} p^r (1-p)^x$, $(x = 0, 1, \ldots)$

$M_X(t) = \left(\dfrac{p}{1 - (1-p)e^t}\right)^r$

$E(X) = r\dfrac{1-p}{p}$, $\text{Var}(X) = r\dfrac{1-p}{p^2}$

$f(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$, $(y = r, r+1, \ldots)$

$M_Y(t) = \left(\dfrac{pe^t}{1 - (1-p)e^t}\right)^r$

$E(Y) = r\dfrac{1}{p}$, $\text{Var}(Y) = r\dfrac{1-p}{p^2}$

NOTE: $X = V_1 + \cdots + V_r$, $(V_i \sim \text{Geometric}(p))$
$(1-x)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} x^k$ $(|x| < 1)$
$\mu = r(1-p), r \to \infty \Rightarrow \text{Poisson}(\mu)$
$X_i \sim \text{NB}(r_i, p) \Rightarrow \sum X_i \sim \text{NB}\left(\sum r_i, p\right)$
$F_{\text{NB}(r,p)}(x) = F_{\text{Beta}(r, \lfloor x \rfloor + 1)}(p)$
$F_{\text{NB}(r,p)}(n-r) = 1 - F_{\text{Bin}(n,p)}(r-1)$

5 Poisson (μ) R:$x$pois

$f(x) = \dfrac{e^{-\mu}\mu^x}{x!}$, $(x = 0, 1, \ldots)$

$M(t) = \exp\left(\mu(e^t - 1)\right)$

$E(X) = \mu$, $\text{Var}(X) = \mu$

NOTE: $X_i \sim \text{Poi}(\mu_i) \Rightarrow \sum X_i \sim \text{Poi}(\sum \mu_i)$

$P_{\text{Poi}(\mu)}[X \ge n] = F_{\Gamma(n,1)}(\mu) = F_{\Gamma(n,\mu)}(1)$
$= F_{\chi^2_{2n}}(2\mu)$

1 Beta (α, β), $(\alpha > 0, \beta > 0)$ R:$x$beta

$f(x) = \dfrac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, $(0 < x < 1)$

$M(t) = 1 + \sum_{k=1}^{\infty}\left(\prod_{r=1}^{k-1}\dfrac{\alpha+r}{\alpha+\beta+r}\right)\dfrac{t^k}{k!}$

$E(X) = \dfrac{\alpha}{\alpha+\beta}$, $\text{Var}(X) = \dfrac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$

NOTE: $f_{\text{Beta}(\alpha,\beta)}(x) = f_{\text{Beta}(\beta,\alpha)}(1-x)$
$F_{\text{Beta}(\alpha,\beta)}(x) = 1 - F_{\text{Beta}(\beta,\alpha)}(1-x)$

2 BS (α, β), $(\alpha > 0, \beta > 0)$: Birnbaum-Saunders

$f(t) = \dfrac{1}{2\alpha\beta}\sqrt{\dfrac{\beta}{t}}\left(1 + \dfrac{\beta}{t}\right)\phi\left[\dfrac{1}{\alpha}\left(\sqrt{\dfrac{t}{\beta}} - \sqrt{\dfrac{\beta}{t}}\right)\right]$, $(t > 0)$

$F(t) = \Phi\left[\dfrac{1}{\alpha}\left(\sqrt{\dfrac{t}{\beta}} - \sqrt{\dfrac{\beta}{t}}\right)\right]$

$F^{-1}(p) = \dfrac{1}{4}\left[\alpha\sqrt{\beta}\Phi^{-1}(p) + \sqrt{\alpha^2\beta\{\Phi^{-1}(p)\}^2 + 4\beta}\right]^2$

$= \beta\left\{1 + \gamma(p)^2 + \gamma(p)\sqrt{\gamma(p)^2 + 2}\right\}$,

where $\gamma(p) = \alpha\Phi^{-1}(p)/\sqrt{2}$

$E(T) = \beta(1 + \tfrac{1}{2}\alpha^2)$, $\text{Var}(T) = (\alpha\beta)^2(1 + \tfrac{5}{4}\alpha^2)$

NOTE: $\text{median}(T) = \beta$, $cT \sim \text{BS}(\alpha, c\beta)$, $T^{-1} \sim \text{BS}(\alpha, \beta^{-1})$
$X = \left(\sqrt{\dfrac{T}{\beta}} - \sqrt{\dfrac{\beta}{T}}\right) \sim N(0, \alpha^2)$
$\log T \sim \text{sinh-Normal}(\log \beta, \alpha)$

3 Cauchy (α, β) R:$x$cauchy

$f(x) = \dfrac{\beta}{\pi}\dfrac{1}{\beta^2 + (x - \alpha)^2}$

$F(x) = \dfrac{1}{\pi}\left[\arctan\left(\dfrac{x-\alpha}{\beta}\right) + \dfrac{\pi}{2}\right]$

$\phi(t) = \exp(it\alpha - \beta|t|)$

NOTE: $\text{Cauchy}(0, 1) = t(1)$
$X, Y \sim N(0, 1) \Rightarrow X/Y \sim \text{Cauchy}(0, 1)$
$cX \sim \text{Cauchy}(c\alpha, c\beta)$
$X_i \sim \text{Cauchy}(\alpha, \beta) \Rightarrow \dfrac{1}{n}\sum X_i \sim \text{Cauchy}(\alpha, \beta)$

4 Chi-Square (n) R:$x$chisq

$f(x) = \dfrac{1}{\Gamma(n/2)2^{n/2}} x^{n/2-1} e^{-x/2}$, $(x \ge 0)$

$E(X) = n$, $\text{Var}(X) = 2n$

$M(t) = \left(\dfrac{1}{1-2t}\right)^{n/2}$, $(t < \tfrac{1}{2})$

$E(X^m) = 2^m\Gamma(m + n/2)/\Gamma(n/2)$

NOTE: $\chi^2(n) = \text{Gamma}(n/2, 2)$
$\chi^2(2) = \text{Exponential}(\beta = 2)$
$X_i \sim \chi^2(n_i) \Rightarrow \sum X_i \sim \chi^2(\sum n_i)$
$X_i \sim N(0, 1) \Rightarrow \sum_{i=1}^n X_i^2 \sim \chi^2(n)$
$F_{\chi^2(p+2)}(x) = F_{\chi^2(p)}(x) - 2xf_{\chi^2(p)}(x)/p$
$F_{\chi^2(1)}(x) = 2\Phi(\sqrt{x}) - 1$
$F_{\chi^2(2)}(x) = 1 - \sqrt{2\pi}\phi(\sqrt{x})$
$F_{\chi^2(3)}(x) = 2\Phi(\sqrt{x}) - 1 - 2\sqrt{x}\phi(\sqrt{x})$

5 Inverse-Chi-Square (n)

$f(x) = \dfrac{1}{\Gamma(n/2)2^{n/2}} x^{-n/2-1} e^{-1/(2x)}$, $(x \ge 0)$

$E(X) = 1/(n-2)$ for $n > 2$

$\text{Var}(X) = 2/[(n-2)^2(n-4)]$ for $n > 4$

NOTE: If $X \sim \chi^2(n)$, then $1/X \sim \text{Inverse-}\chi^2(n)$.
$F_{\text{Inv-}\chi^2(n)}(x) = 1 - F_{\chi^2(n)}(1/x)$

6 Exponential (θ) R:$x$exp

$f(x) = \dfrac{1}{\theta}e^{-x/\theta}$ $(x \ge 0,\ \theta > 0)$

$F(x) = 1 - e^{-x/\theta}$

$M(t) = \dfrac{1}{1 - \theta t}$, $(t < \tfrac{1}{\theta})$

$E(X) = \theta$, $\text{Var}(X) = \theta^2$

NOTE: Memoryless property
$cX \sim \text{Exponential}(c\theta)$
$\sum X_i \sim \text{Gamma}(n, \theta)$
$\min\limits_{1 \le i \le n}(X_i) \sim \text{Exponential}(\theta/n)$: self-reproducing

$$Y = X^{1/\alpha} \sim \text{Weibull}(\alpha, \theta)$$
$$Y = \sqrt{2X/\theta} \sim \text{Rayleigh}(1)$$
$$Y = \alpha - \gamma \log(X/\theta) \sim \text{Gumbel}(\alpha, \gamma)$$

7 $F(m, n)$ R:`xf`

$$F(m, n) = \frac{\chi_m^2/m}{\chi_n^2/n}$$
$$E(X) = \frac{n}{n-2} \quad (n > 2),$$
$$\text{Var}(X) = 2\left(\frac{n}{n-2}\right)^2 \frac{m+n-2}{m(n-4)} \quad (n > 4)$$

NOTE: $[F(m,n)]^{-1} = F(n,m)$
$F_{1-\alpha}(m,n) = [F_\alpha(n,m)]^{-1}$
$F(1,k) = t^2(k)$
If $X \sim F(m,n)$, $mX/(mX+n) \sim \text{Beta}(\frac{m}{2}, \frac{n}{2})$.

8 Gamma (α, θ) R:`xgamma`

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} \exp(-x/\theta), \ (0 < x < \infty)$$
$$M(t) = (1 - \theta t)^{-\alpha} \ (t < 1/\theta)$$
$$E(X) = \alpha\theta, \ \text{Var}(X) = \alpha\theta^2$$

NOTE: $X \sim \text{Gamma}(n, \theta) = \text{Erlang}(n, \theta)$
$X = V_1 + \cdots + V_n, \ (V_i \sim \text{Exponential}(\theta))$
$2X/\theta \sim \chi^2(2n)$

$$F_X(x) = 1 - \sum_{k=0}^{n-1} \left(\frac{x}{\theta}\right)^k \frac{e^{-x/\theta}}{k!} = \sum_{k=n}^{\infty} \left(\frac{x}{\theta}\right)^k \frac{e^{-x/\theta}}{k!}$$

$\text{Gamma}(n/2, 2) = \chi^2(n), \ \Gamma(1/2) = \sqrt{\pi}$
$X_i \sim \text{Gamma}(\alpha_i, \theta) \Rightarrow \sum_{i=1}^r X_i \sim \text{Gamma}(\sum_{i=1}^r \alpha_i, \theta)$
$X/d \sim \text{Gamma}(\alpha, \theta/d)$
$E[X^c] = \Gamma(\alpha+c)\theta^c/\Gamma(\alpha) \ (c > -\alpha)$

9 Gumbel (μ, β)

$$f(x) = \frac{1}{\beta} \exp\left(-\left(\frac{x-\mu}{\beta}\right) - e^{-(x-\mu)/\beta}\right)$$
$$M(t) = \Gamma(1 - \beta t)e^{\mu t}$$
$$E(X) = \mu + \beta\gamma \text{ with } \gamma \approx 0.5772157 \text{ (Euler–Mascheroni constant)}$$
$$\text{Var}(X) = \pi^2\beta^2/6 \approx 1.644934\beta^2.$$

10 Inverse-Gamma (α, θ)

$$f(x) = \frac{\theta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp(-\theta/x), \ (0 < x < \infty)$$
$$E(X) = \theta/(\alpha-1) \text{ for } \alpha > 1$$
$$\text{Var}(X) = \theta^2/[(\alpha-1)^2(\alpha-2)] \text{ for } \alpha > 2$$

NOTE: If $X \sim \text{Gamma}(\alpha, \theta)$, then $1/X \sim \text{Inv-Gamma}(\alpha, \theta)$.

11 Laplace (μ, σ): Double Exponential

$$f(x) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}$$
$$F(x) = \begin{cases} \frac{1}{2} e^{-|x-\mu|/\sigma} & (x < \mu) \\ 1 - \frac{1}{2} e^{-|x-\mu|/\sigma} & (x \geq \mu) \end{cases}$$
$$M(t) = \frac{e^{\mu t}}{1 - (\sigma t)^2}, \ (|t| < \frac{1}{\sigma})$$
$$E(X) = \text{median}(X) = \mu, \ \text{Var}(X) = 2\sigma^2$$

12 Logistic, (μ, β) R:`xlogis`

$$f(x) = \frac{e^{-(x-\mu)/\beta}}{\beta[1 + e^{-(x-\mu)/\beta}]^2} = \frac{1}{2\beta\left[1 + \cosh\left(\frac{x-\alpha}{\beta}\right)\right]}$$
$$F(x) = \frac{1}{1 + e^{-(x-\mu)/\beta}}$$
$$M(t) = e^{\mu t}\Gamma(1 - \beta t)\Gamma(1 + \beta t), \ |t| < \frac{1}{\beta}$$
$$E(X) = \text{median}(X) = \mu, \text{Var}(X) = \frac{\pi^2\beta^2}{3}$$

13 Lognormal (μ, σ^2) R:`xlnorm`

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}x} \exp(-\frac{1}{2}(\frac{\log(x)-\mu}{\sigma})^2)$$
$$E(X^k) = e^{k\mu + k^2\sigma^2/2}, \text{Var}(X) = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2}$$

NOTE: $F_{logN(\mu,\sigma^2)}(x) = F_{N(\mu,\sigma^2)}(\log x) = \Phi\left(\frac{\log x - \mu}{\sigma}\right)$
Self-reproducing under multiplication and division

14 Normal (μ, σ^2) R:`xnorm`

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)$$
$$M(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$$

NOTE: If $X \sim N(\mu, \sigma^2)$, $Y = e^X \sim logN(\mu, \sigma^2)$

$$F_{N(0,1)}(x) = \frac{1}{2} + \frac{1}{2}\text{sign}(x)F_{\Gamma(1/2,1)}(\frac{1}{2}x^2)$$
$$= \frac{1}{2} + \frac{1}{2}\text{sign}(x)F_{\chi^2(1)}(x)$$
$$\phi'(z) = -z\phi(z), \ \phi''(z) = (z^2 - 1)\phi(z)$$
$$E(X^3) = \mu^3 + 3\mu\sigma^2, \ E(X^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

15 Rayleigh (β)

$$f(x) = \frac{x}{\beta^2} \exp\left(-\frac{x^2}{2\beta^2}\right)$$
$$F(x) = 1 - \exp\left(-\frac{x^2}{2\beta^2}\right)$$
$$E(X) = \beta\sqrt{\pi/2}, \ E(X^k) = (\sqrt{2}\beta)^k \frac{k}{2}\Gamma(\frac{k}{2})$$
$$\text{median}(X) = \beta\sqrt{2\ln 2}$$
$$\text{Var}(X) = (2 - \pi/2)\beta^2$$

NOTE: $\text{Rayleigh}(\beta) = \text{Weibull}(2, 2\beta^2)$
$cX \sim \text{Rayleigh}(c\beta)$
$(X/\beta)^2 \sim \chi^2(2)$
$X^2 \sim \text{Exponential}(2\beta^2) = \text{Gamma}(1, 2\beta^2)$
$\min_{1 \leq i \leq n}(X_i) \sim \text{Rayleigh}(\beta/\sqrt{n})$: self-reproducing

16 Slash (α, β)

$$f(x) = \frac{1}{\sqrt{2\pi}(x-\alpha)^2/\beta^2}\left(1 - \exp(-\frac{1}{2}(\frac{x-\alpha}{\beta})^2)\right)$$
$$F(x) = \Phi\left(\frac{x-\alpha}{\beta}\right) - \left(\frac{x-\alpha}{\beta}\right)f_{\text{Slash}(0,1)}\left(\frac{x-\alpha}{\beta}\right)$$

NOTE: $X = \alpha + \beta\dfrac{Z}{U}$,
where $Z \sim N(0, 1)$ and $U \sim \text{Uniform}(0, 1)$.

17 Student $t(k)$ R:`xt`

$$f(x) = \frac{\Gamma((k+1)/2)}{\Gamma(k/2)} \frac{1}{\sqrt{k\pi}}\left(1 + \frac{x^2}{k}\right)^{-(k+1)/2}, \ (k \geq 1)$$
$$E(X) = 0 \ (k > 1), \ \text{Var}(X) = k/(k-2) \ (k > 2)$$

NOTE: $X \sim \dfrac{N(0,1)}{\sqrt{\chi_k^2/k}}, \quad F_\alpha(1,k) = t_{\alpha/2}(k)^2$

$$F_{t(k)}(x) = 1 - \frac{1}{2}F_{\text{Beta}(k/2,1/2)}\left(\frac{k}{k+x^2}\right) \quad (x \geq 0)$$
$$= \frac{1}{2}F_{\text{Beta}(k/2,1/2)}\left(\frac{k}{k+x^2}\right) \quad (x < 0)$$

18 Uniform (a, b) R:`xunif`

$$f(x) = \frac{1}{b-a}$$
$$M(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}, \ (t \neq 0)$$
$$E(X) = \text{median}(X) = \frac{a+b}{2}, \ \text{Var}(X) = \frac{(b-a)^2}{12}$$

NOTE: $X \sim \text{Uniform}(0, 1), \ -\log X \sim \text{Exponential}(1)$

19 Wald (μ, λ): Inverse-Gaussian (IG)   R:`xinvgauss{statmod}`

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\lambda\frac{(x-\mu)^2}{2\mu^2 x}\right] = \sqrt{\frac{\lambda}{x^3}}\phi\left(\sqrt{\frac{\lambda}{x}}\frac{x-\mu}{\mu}\right)$$
$$E(X) = \mu, \ \text{Var}(X) = \mu^3/\lambda$$
$$F(x) = \Phi\left(\sqrt{\frac{\lambda}{x}}\frac{x-\mu}{\mu}\right) + \exp\left(\frac{2\lambda}{\mu}\right)\Phi\left(-\sqrt{\frac{\lambda}{x}}\frac{x+\mu}{\mu}\right)$$
$$M(t) = \exp\left[\frac{\lambda}{\mu}\left(1 - \sqrt{1 - 2\mu^2 t/\lambda}\right)\right]$$
$$\hat{\mu}_{\text{mle}} = \bar{X}, \ \hat{\lambda}_{\text{mle}} = \left[\frac{1}{n}\sum\left\{X_i^{-1} - \bar{X}^{-1}\right\}\right]^{-1}.$$

NOTE: $\lambda(X - \mu)^2/(\mu^2 X) \sim \chi^2(1)$
$X_i \sim \text{IG}(\mu, \lambda) \Rightarrow kX_i \sim \text{IG}(k\mu, k\lambda)$,
$\sum X_i \sim \text{IG}(n\mu, n^2\lambda), \ n\lambda/\hat{\lambda}_{\text{mle}} \sim \chi^2(n-1)$

20 Weibull (α, θ) R:`xweibull`

$$f(x) = \frac{\alpha}{\theta}\left(\frac{x}{\theta}\right)^{\alpha-1} \exp\left(-(x/\theta)^\alpha\right), \ (x \geq 0, \alpha > 0, \theta > 0)$$
$$F(x) = 1 - \exp\left(-(x/\theta)^\alpha\right)$$
$$E(X^k) = \theta^k\Gamma(1 + \frac{k}{\alpha}), \quad \text{median}(X) = \theta(\ln 2)^{1/\alpha}$$
$$\text{Var}(X) = \theta^2\left[\Gamma(1 + \frac{2}{\alpha}) - \Gamma^2(1 + \frac{1}{\alpha})\right]$$

NOTE: $\theta = 1 - e^{-1} \approx 63.2\%$ percentile
$Y = X^\alpha \sim \text{Exponential}(\theta^\alpha)$
$\text{Weibull}(1, \theta) = \text{Exponential}(\theta)$
$\text{Weibull}(2, \sqrt{2}\theta) = \text{Rayleigh}(\theta)$
$\min_{1 \leq i \leq n}(X_i) \sim \text{Weibull}(\alpha, \theta/n^{1/\alpha})$: self-reproducing
$Y = 1/X \sim \text{Inv. Wei}(\alpha, 1/\theta), \ F_Y(y) = \exp\left(-(\frac{1/\theta}{y})^\alpha\right)$