

Probability Distributions

1 **Binomial** (n, p) , $(x = 0, 1, \dots, n)$

R: **xbinom**

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

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

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

NOTE: $n \rightarrow \infty$ with $\mu = np \Rightarrow \text{Poisson}(\mu)$

$$F_{\text{Bin}(n,p)}(x) = F_{\text{Beta}(n-\lfloor x \rfloor, \lfloor x \rfloor + 1)}(1-p)$$

$$F_{\text{Bin}(n,p)}(r-1) = 1 - F_{\text{NegBin}(r,p)}(n-r)$$

$$X_i \sim \text{Bin}(n_i, p) \Rightarrow \sum X_i \sim \text{Bin}(\sum n_i, p)$$

2 **Geometric** (p)

R: **xgeom**

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

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

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

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

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

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

$$f(y) = p(1-p)^{y-1}, \quad (y = 1, 2, \dots)$$

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

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

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

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

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

NOTE: $\min_{1 \leq i \leq n} (X_i) \sim \text{Geo}(1 - (1-p)^n)$: self-reproducing

3 **Hypergeometric** (N, M, n)

R: **xhyper**

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

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

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

4 **Negative Binomial** (r, p)

R: **xnbinom**

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

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

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

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

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

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

NOTE: $X = V_1 + \dots + V_r$, $(V_i \sim \text{Geometric}(p))$

$$(1-x)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} x^k \quad (|x| < 1)$$

$$\mu = r(1-p), r \rightarrow \infty \Rightarrow \text{Poisson}(\mu)$$

$$X_i \sim \text{NB}(r_i, p) \Rightarrow \sum X_i \sim \text{NB}(\sum r_i, p)$$

$$F_{\text{NB}(r,p)}(x) = F_{\text{Beta}(r, \lfloor x \rfloor + 1)}(p)$$

$$F_{\text{NB}(r,p)}(n-r) = 1 - F_{\text{Bin}(n,p)}(r-1)$$

5 **Poisson** (μ)

R: **xpois**

$$f(x) = \frac{e^{-\mu} \mu^x}{x!}, \quad (x = 0, 1, \dots)$$

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

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

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

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

1 **Beta** (α, β) , $(\alpha > 0, \beta > 0)$

R: **xbeta**

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

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

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

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

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

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

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

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

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

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

NOTE: median(T) = β , $cT \sim \text{BS}(\alpha, c\beta)$, $T^{-1} \sim \text{BS}(\alpha, \beta^{-1})$

$$X = \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim N(0, \alpha^2)$$

$$\log T \sim \text{sinh-Normal}(\log \beta, \alpha)$$

3 **Cauchy** (α, β)

R: **xcauchy**

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

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

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

NOTE: Cauchy(0, 1) = $t(1)$

$$X, Y \sim N(0, 1) \Rightarrow X/Y \sim \text{Cauchy}(0, 1)$$

$$cX \sim \text{Cauchy}(c\alpha, c\beta)$$

$$X_i \sim \text{Cauchy}(\alpha, \beta) \Rightarrow \frac{1}{n} \sum X_i \sim \text{Cauchy}(\alpha, \beta)$$

4 **Chi Squared** (n)

R: **xchisq**

$$f(x) = \frac{1}{\Gamma(n/2) 2^{n/2}} x^{n/2-1} e^{-x/2}, \quad (x \geq 0)$$

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

$$M(t) = \left(\frac{1}{1-2it} \right)^{n/2}, \quad (t < \frac{1}{2})$$

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

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 **Exponential** (θ)

R: **xexp**

$$f(x) = \frac{1}{\theta} e^{-x/\theta} \quad (x \geq 0, \theta > 0)$$

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

$$M(t) = \frac{1}{1 - \theta t}, \quad (t < \frac{1}{\theta})$$

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

NOTE: Memoryless property

$$cX \sim \text{Exponential}(c\theta)$$

$$\sum X_i \sim \text{Gamma}(n, \theta)$$

$\min_{1 \leq i \leq n} (X_i) \sim \text{Exponential}(\theta/n)$: self-reproducing

$$Y = X^{1/\alpha} \sim \text{Weibull}(\alpha, \theta)$$

$$Y = \sqrt{2X/\theta} \sim \text{Rayleigh}(1)$$

$$Y = \alpha - \gamma \log(X/\theta) \sim \text{Gumbel}(\alpha, \gamma)$$

6 **F** (m, n) R:**x**f

$$F(m, n) = \frac{\chi_m^2/m}{\chi_n^2/n}$$

$$E(X) = \frac{n}{n-2} \quad (n > 2),$$

$$\text{Var}(X) = 2 \left(\frac{n}{n-2} \right)^2 \frac{m+n-2}{m(n-4)} \quad (n > 4)$$

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 $X \sim F(m, n)$, $mX/(mX + n) \sim \text{Beta}(\frac{m}{2}, \frac{n}{2})$.

7 **Gamma** (α, θ) R:**x**gamma

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} \exp(-x/\theta), \quad (0 < x < \infty)$$

$$M(t) = (1 - \theta t)^{-\alpha} \quad (t < 1/\theta)$$

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

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

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

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

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

8 **Laplace** (μ, σ): Double Exponential

$$f(x) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}$$

$$F(x) = \begin{cases} \frac{1}{2} e^{-|x-\mu|/\sigma} & (x < \mu) \\ 1 - \frac{1}{2} e^{-|x-\mu|/\sigma} & (x \geq \mu) \end{cases}$$

$$M(t) = \frac{e^{\mu t}}{1 - (\sigma t)^2}, \quad (|t| < \frac{1}{\sigma})$$

$$E(X) = \text{median}(X) = \mu, \quad \text{Var}(X) = 2\sigma^2$$

9 **Logistic** (μ, β) R:**x**logis

$$f(x) = \frac{e^{-(x-\mu)/\beta}}{\beta[1 + e^{-(x-\mu)/\beta}]^2} = \frac{1}{2\beta \left[1 + \cosh\left(\frac{x-\mu}{\beta}\right) \right]}$$

$$F(x) = \frac{1}{1 + e^{-(x-\mu)/\beta}}$$

$$M(t) = e^{\mu t} \Gamma(1 - \beta t) \Gamma(1 + \beta t), \quad |t| < \frac{1}{\beta}$$

$$E(X) = \text{median}(X) = \mu, \quad \text{Var}(X) = \frac{\pi^2 \beta^2}{3}$$

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

10 **Lognormal** (μ, σ^2) R:**x**lnorm

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}x} \exp\left(-\frac{1}{2}\left(\frac{\log(x)-\mu}{\sigma}\right)^2\right)$$

$$E(X^k) = e^{k\mu + k^2\sigma^2/2}, \quad \text{Var}(X) = e^{2(\mu + \sigma^2)} - e^{2\mu + \sigma^2}$$

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

11 **Normal** (μ, σ^2) R:**x**norm

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

$$M(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$$

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

$$F_{N(0,1)}(x) = \frac{1}{2} + \frac{1}{2} \text{sign}(x) F_{\Gamma(1/2,1)}\left(\frac{1}{2}x^2\right)$$

$$= \frac{1}{2} + \frac{1}{2} \text{sign}(x) F_{\chi^2(1)}(x)$$

$$\phi'(z) = -z\phi(z), \quad \phi''(z) = (z^2 - 1)\phi(z)$$

$$E(X^3) = \mu^3 + 3\mu\sigma^2, \quad E(X^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

12 **Rayleigh** (β)

$$f(x) = \frac{x}{\beta^2} \exp\left(-\frac{x^2}{2\beta^2}\right)$$

$$F(x) = 1 - \exp\left(-\frac{x^2}{2\beta^2}\right)$$

$$E(X) = \beta\sqrt{\pi/2}, \quad E(X^k) = (\sqrt{2}\beta)^k \frac{k}{2} \Gamma\left(\frac{k}{2}\right)$$

$$\text{median}(X) = \beta\sqrt{2 \ln 2}$$

$$\text{Var}(X) = (2 - \pi/2)\beta^2$$

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

13 **Slash** (α, β)

$$f(x) = \frac{1}{\sqrt{2\pi}(x-\alpha)^2/\beta^2} \left(1 - \exp\left(-\frac{1}{2}\left(\frac{x-\alpha}{\beta}\right)^2\right)\right)$$

$$F(x) = \Phi\left(\frac{x-\alpha}{\beta}\right) - \left(\frac{x-\alpha}{\beta}\right) f_{\text{Slash}(0,1)}\left(\frac{x-\alpha}{\beta}\right)$$

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

14 **Student** $t(k)$ R:**x**t

$$f(x) = \frac{\Gamma((k+1)/2)}{\Gamma(k/2)} \frac{1}{\sqrt{k\pi}} \left(1 + \frac{x^2}{k}\right)^{-(k+1)/2}, \quad (k \geq 1)$$

$$E(X) = 0 \quad (k > 1), \quad \text{Var}(X) = k/(k-2) \quad (k > 2)$$

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

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

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

15 **Uniform** (a, b) R:**x**unif

$$f(x) = \frac{1}{b-a}$$

$$M(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}, \quad (t \neq 0)$$

$$E(X) = \text{median}(X) = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

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

16 **Wald** (μ, λ): Inverse Gaussian (IG) R:**x**invgauss{statmod}

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\lambda \frac{(x-\mu)^2}{2\mu^2 x}\right] = \sqrt{\frac{\lambda}{x^3}} \phi\left(\sqrt{\frac{\lambda}{x}} \frac{x-\mu}{\mu}\right)$$

$$E(X) = \mu, \quad \text{Var}(X) = \mu^3/\lambda$$

$$F(x) = \Phi\left(\sqrt{\frac{\lambda}{x}} \frac{x-\mu}{\mu}\right) + \exp\left(\frac{2\lambda}{\mu}\right) \Phi\left(-\sqrt{\frac{\lambda}{x}} \frac{x+\mu}{\mu}\right)$$

$$M(t) = \exp\left[\frac{\lambda}{\mu} \left(1 - \sqrt{1 - 2\mu^2 t/\lambda}\right)\right]$$

$$\hat{\mu}_{\text{mle}} = \bar{X}, \quad \hat{\lambda}_{\text{mle}} = \left[\frac{1}{n} \sum \{X_i^{-1} - \bar{X}^{-1}\}\right]^{-1}.$$

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

17 **Weibull** (α, θ) R:**x**weibull

$$f(x) = \frac{\alpha}{\theta} \left(\frac{x}{\theta}\right)^{\alpha-1} \exp\left(-\left(x/\theta\right)^\alpha\right), \quad (x \geq 0, \alpha > 0, \theta > 0)$$

$$F(x) = 1 - \exp\left(-\left(x/\theta\right)^\alpha\right)$$

$$E(X^k) = \theta^k \Gamma(1 + \frac{k}{\alpha}), \quad \text{median}(X) = \theta (\ln 2)^{1/\alpha}$$

$$\text{Var}(X) = \theta^2 \left[\Gamma(1 + \frac{2}{\alpha}) - \Gamma^2(1 + \frac{1}{\alpha})\right]$$

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