Formulae

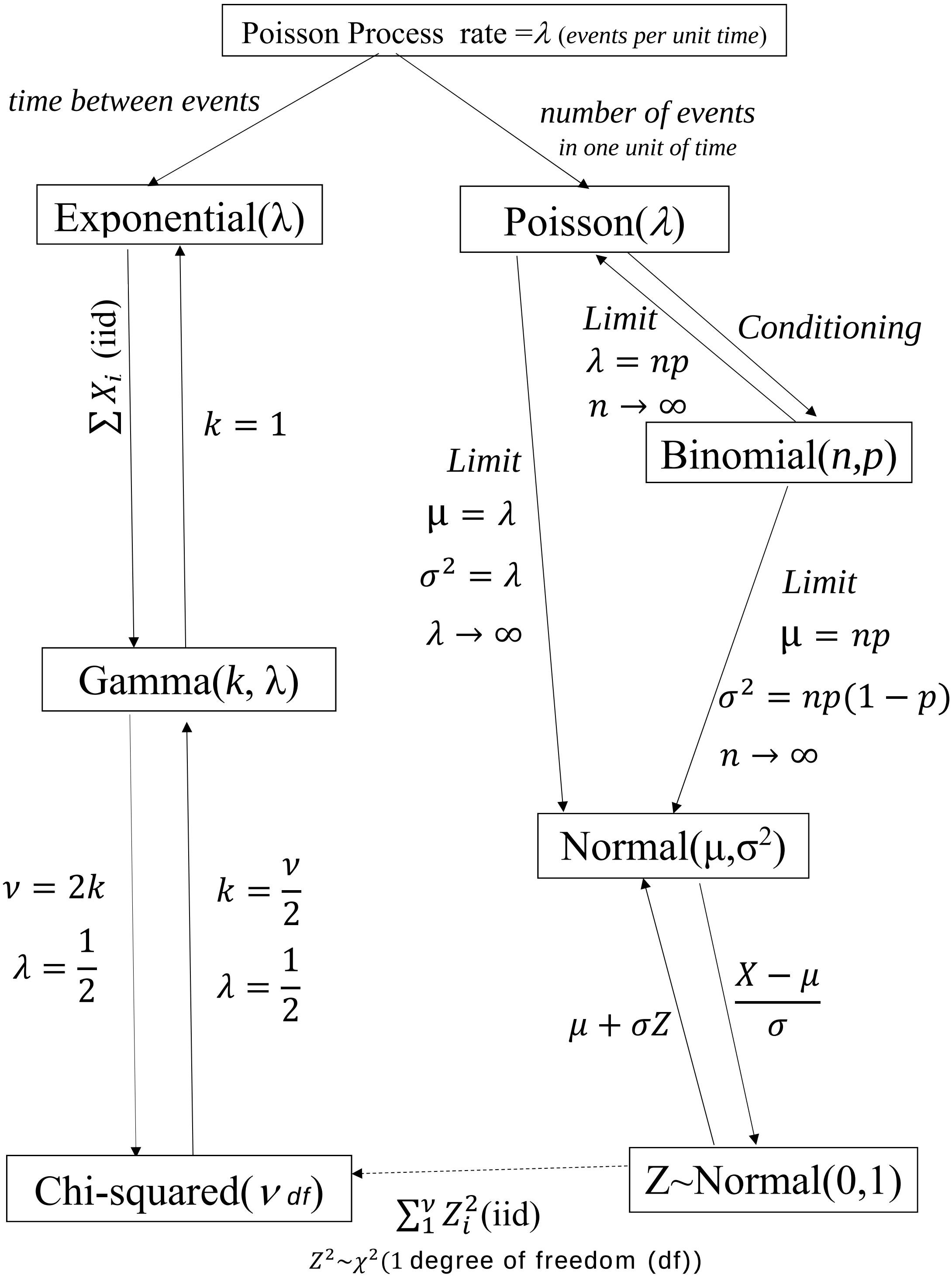
Discrete Distributions

Notation and Parameters	pmf $f_X(x)$	Mean $\mathbb{E}(X)$	Variance $\mathrm{Var}(X)$
Binomial			
$X \sim \mathrm{Bin}(n, p)$	$\binom{n}{x} p^x (1-p)^{n-x}$	np	$np(1-p)$
$0 < p < 1$	$x = 0, 1, \ldots, n$		
	`dbinom`(x, n, p)		
Hypergeometric			
$X \sim \mathrm{Hyp}(n, M, N)$	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$	$\frac{nM}{N}$	$n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}$
$n = 1, 2, 3, \ldots, N$	$x = \max(0, n - N + M) \ldots,$		
$M = 0, 1, 2, \ldots, N$	$\ldots, \min(n, M)$		
	`dhyper`$(x, M, N - M, n)$		
Poisson			
$X \sim \mathrm{Poi}(\lambda)$	$\frac{\lambda^x}{x!} e^{-\lambda}$	λ	λ
$\lambda > 0$	$x = 0, 1, 2, \ldots$		
	`dpois`(x, λ)		
Negative binomial			
$X \sim \mathrm{NegBin}(k, p)$	$\binom{k+x-1}{x} p^k (1-p)^x$	$\frac{k(1-p)}{p}$	$\frac{k(1-p)}{p^2}$
$0 < p < 1$	$x = 0, 1, 2, \ldots$		
$k = 1, 2, \ldots$	`dnbinom`(x, k, p)		
Geometric			
$X \sim \mathrm{Geo}(p)$	$(1-p)^x p$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
	$x = 0, 1, 2, 3 \ldots$		
	`dgeom`(x, p)		

Continuous Distributions

Notation and Parameters	**pdf** $f_X(x)$	**Mean** $\mathbb{E}(X)$	**Variance** $\mathrm{Var}(X)$
Uniform			
$X \sim$Uniform(a,b)	$\frac{1}{b-a}$	$\frac{b-a}{2}$	$\frac{(b-a)^2}{12}$
$a \neq b$	$a < x < b$		
$F_X(x) = \frac{x-a}{b-a}$	$F_X(x) = \texttt{punif}(x,a,b)$		
Exponential			
$X \sim$ Exponential(λ)	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\lambda > 0$	$x > 0$		
$F_X(x) = 1 - e^{-\lambda x}$	$F_X(x) = \texttt{pexp}(x,\lambda)$		
Gamma			
$X \sim$ Gamma(k,λ)	$\frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}$	$\frac{k}{\lambda}$	$\frac{k}{\lambda^2}$
$\lambda > 0$	$x > 0$		
	$F_X(x) = \texttt{pgamma}(x,k,\lambda)$		
Chi-squared			
$X \sim \chi^2_\nu$	$\frac{(1/2)^{\nu/2}}{\Gamma(k)} x^{\nu/2-1} e^{-x/2}$	ν	2ν
$\nu = 1, 2, \ldots$	$x > 0$		
	$F_X(x) = \texttt{pchisq}(x,\nu)$		
Normal			
$X \sim$ Normal(μ,σ^2)	$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\{-(x-\mu)^2/2\sigma^2\}}$	μ	σ^2
$-\infty < x < \infty$, $\sigma^2 > 0$	$-\infty < x < \infty$		
	$F_X(x)$=$\texttt{pnorm}(x,\mu,\sigma)$		

Connections between distributions



Joint distributions

Two continuous random variables

Joint cdf
$$F_{X,Y}(x,y) = \Pr(X \le x, Y \le y)$$

Joint pdf
$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

- $f_{X,Y}(x,y) \ge 0$
- $\int\limits_x \int\limits_y f_{X,Y}(x,y)\, dy dx = 1$

Marginal pdf
$$f_X(x) = \int\limits_y f_{X,Y}(x,y)$$
$$= \int\limits_y f_{X|Y}(x|y) f_Y(y)$$

Conditional pdf
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

Independence
$$\Pr(X \le x, Y \le y) = \Pr(X \le x) \Pr(Y \le y)$$
$$F_{X,Y}(x,y) = F_X(x) F_Y(y)$$
$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$
$$\text{for all } x \text{ and } y$$

Expectation of a function
$$\mathbb{E}(g(X,Y)) = \int_x \int_y g(x,y) f_{X,Y}(x,y)\, dy dx$$

Covariance
$$\mathrm{Cov}(X,\, Y) {=} \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Correlation
$$\rho_{X,Y} = \frac{\mathrm{Cov}(X,\, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

Conditional Expectation
$$\psi_Y(x) = \mathbb{E}(Y|X = x) = \int\limits_{-\infty}^{\infty} y f_{Y|X}(y|x)\, dy.$$

All properties of expectation,variance and covariance are exactly the same for continuous and discrete random variables.

Properties of expectation and variance

For any random variables, X, Y, and for arbitrary constants a and b:

- $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$.
- $\mathbb{E}(ag(X) + b) = a\mathbb{E}(g(X)) + b$.
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.
- $\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X)$.
- $\operatorname{Var}(ag(X) + b) = a^2 \operatorname{Var}(g(X))$.

If $X_1, \ldots, X_n$ are *INDEPENDENT* random variables, and if $a_1, \ldots, a_n$ and b are arbitrary constants:

- $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.
- $\operatorname{Var}(a_1X_1 + \ldots a_nX_n + b) = a_1^2 \operatorname{Var}(X_1) + \ldots + a_n^2 \operatorname{Var}(X_n)$.

Covariance Properties

For any random variables X, Y and Z:

- $\operatorname{Cov}(X, X) = \operatorname{Var}(X)$
- $\operatorname{Cov}(aX + b, cY + d) = ac \operatorname{Cov}(X, Y)$
- $\operatorname{Cov}(X, Y) = \operatorname{Cov}(Y, X)$
- $\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2 \operatorname{Cov}(X, Y)$
- $\operatorname{Var}(X - Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) - 2 \operatorname{Cov}(X, Y)$
- $\operatorname{Cov}(X, Y + Z) = \operatorname{Cov}(X, Y) + \operatorname{Cov}(X, Z)$

Properties of the Bivariate Normal distribution

If the pdf of (X, Y) is Bivariate Normal then we have:

(BV$_1$) $\mathbb{E}(X) = \mu_X$, $\mathbb{E}(Y) = \mu_Y$, $\mathrm{Var}(X) = \sigma_X^2$, $\mathrm{Var}(Y) = \sigma_Y^2$ and $\rho_{X,Y} = \rho$.

(BV$_2$) The marginal distribution of X is Normal(μ_X, σ_X^2) and the marginal distribution of Y is Normal(μ_Y, σ_Y^2).

(BV$_3$) The conditional distribution of Y given that $X = x$ is the Normal distribution with mean

$$\mathbb{E}(Y|X = x) = \mu_Y + \rho\sigma_Y \frac{x - \mu_X}{\sigma_X}$$

and variance given by

$$\mathrm{Var}(Y|X = x) = (1 - \rho^2)\sigma_Y^2.$$

A similar result holds for the conditional distribution of X given that $Y = y$.

(BV$_4$) Let $U = aX + bY$, where $a, b \in \mathbb{R}$. Then the distribution of U is Normal, with mean $a\mu_X + b\mu_Y$ and variance $a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y$.

(BV$_5$) Let $U = aX + bY$ and $V = cX + dY$, where $a, b, c, d \in \mathbb{R}$ and $ad - bc \neq 0$. Then the joint distribution of U and V is Bivariate Normal and

$$\begin{aligned}
\mathbb{E}(U) =& a\mu_X + b\mu_Y, \\
\mathrm{Var}(U) =& a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y, \\
\mathbb{E}(V) =& c\mu_X + d\mu_Y, \\
\mathrm{Var}(V) =& c^2\sigma_X^2 + d^2\sigma_Y^2 + 2cd\rho\sigma_X\sigma_Y, \\
\mathrm{Cov}(U, V) =& ac\sigma_X^2 + bd\sigma_Y^2 + (ad + bc)\rho\sigma_X\sigma_Y.
\end{aligned}$$