Cumulative Distribution Function	**Probability Density Function**
• $F_X(x) = P(X < x) = \int_{-\infty}^{x} f_X(t)dt$ • $F_X(x) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$	• $f_X(x) = \frac{d}{dx}F_X(x) \ge 0$, $f_X(\boldsymbol{x}) = \frac{\partial^n F_X(x)}{\partial x_1 \partial x_2 \dots \partial x_n}$ • $P(a \le x \le b) = \int_a^b f_X(t)dt$ • $\int_{-\infty}^{\infty} f_X(x)dx = 1$
Expected value	**Variance**

• $E\,X = \sum_i x_i p_i = \sum_{\omega \in \Omega} X(\omega)P(\omega)$
• $E\,h(X) = \int_{-\infty}^{\infty} h(x) f_X(x)dx$
• $E(E(X \mid Y)) = EX$
  ◦ $E(XY) = E\big(E(XY\|X)\big) = E(X\,E(Y\|X))$
  ◦ $E(XY) = E\big(E(XY\|Y)\big) = E(Y\,E(X\|Y))$
• $E(Y \mid X) = \sum_j y_j P\big(y = y_j \mid X\big)$
	• $Var\,X = E(X - E\,X)^2 = E\,X^2 - (E\,X)^2$
	• $Var\,c = 0$
	• $Var\,(cX) = c^2\,Var\,X$
	• $Var\,(X \pm Y) = Var\,X + Var\,Y \pm 2\,Cov(X,Y)$
	• $Var(a_1X_1 + \dots + a_nX_n) =$
	$= Cov(a_1X_1 + \dots + a_nX_n, a_1X_1 + \dots + a_nX_n)$
	$= (a_1 \dots a_n)\begin{pmatrix} Cov(X_1,X_1) & \dots & Cov(X_1,X_n) \\ \dots & \dots & \dots \\ Cov(X_n,X_1) & \dots & Cov(X_n,X_n) \end{pmatrix}\begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix}$
	• $Var\,(Y \mid X) = \sum_j \big(y_j - EY\big)^2 P\big(y = y_j \mid X\big)$
	• $Var\,Y = E\big(Var(Y \mid X)\big) + Var\big(E(Y \mid X)\big)$
	• $Var(XY) = Var(X)Var(Y) + Var(X)E(Y)^2 + Var(Y)E(X)^2$ For independent X. Y (proof)

Covariance	**Correlation Coefficient**
• $Cov(X,Y) = \sigma_{XY} = E\,(X - EX)(Y - EY) = E(XY) - E\,X * E\,Y$ • $Cov(X,X) = Var\,X$ • $Cov(X_1 + X_2, Y) = Cov(X_1,Y) + Cox(X_2,Y)$ • $Cov(aX + b, cY + d) = acCov(X,Y)$	• $Corr(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var\,X * Var\,Y}}$ • $-1 \le Corr(X,Y) \le 1$

Characteristic function
• $\phi_X(t) = E\big(e^{itX}\big), t \in \boldsymbol{R}$
  ◦ **Discrete:** $\phi_X(t) = \sum_k e^{itx_k} P(X = x_k)$
  ◦ **Continuous:** $\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)dx$
    ▪ $\phi_X(0) = 1, \|\phi_X(t)\| \le 1$
    ▪ X_1, X_2 are independent $\Longrightarrow$ $\phi_{X_1+X_2}(t) = \phi_{X_1}(t)\phi_{X_2}(t)$
    ▪ $\phi_{aX+b}(t) = e^{ibt}\phi_X(at)$
Inverse formula: $f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi_X(t)dt$
	Markov's inequality: $P(\|X\| \ge \varepsilon) \le \frac{E\|X\|^t}{\varepsilon^t}, \varepsilon > 0, t > 0$
	Chebyshev's inequality: $P(\|X - EX\| \ge \varepsilon) \le \frac{Var\,X}{\varepsilon^2}, \varepsilon > 0$
	Marginal PDF: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$
	Joint distribution: $f_{U,V}(u,v) = f_{X,Y}\big(x(u,v), y(u,v)\big)\left\|\det\begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix}\right\|$
	The Convolution formula: $f_{U+V}(t) = \int_{-\infty}^{\infty} f_{U,V}(t - y, y)dy$

Conditional Expectation for Joint probability distribution

Discrete case

$$E(X \mid Y = y_0) = \frac{1}{P(Y = y_0)} \sum_{x_i} x_i P(X = x_i, Y = y_0)$$

$$E(X \mid Y) \sim \begin{pmatrix} \sum_{x_i} x_i P(X = x_i, Y = y_1) & \dots & \sum_{x_i} x_i P(X = x_i, Y = y_n) \\ P(Y = y_1) & \dots & P(Y = y_n) \end{pmatrix}$$

Continuous case

$$E(X \mid Y = y_0) = \int_{-\infty}^{\infty} x\, f_{X|Y=y_0}(x) dx\,,\ f_{X|Y=y_0}(x) = \frac{f_{X,Y}(x, y_0)}{f_Y(y_0)}$$

$$E(X \mid Y) = E(X \mid Y = y) = H(y)$$

Binomial distribution $X \sim \mathrm{Bin}(n, p)$ <u>"number of successes in n Bernoulli trials"</u>

PMF	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
Expected value & Variance	$E\,X = np, E\,X^2 = np(q - np), \mathrm{Var}\,X = npq$
$c * \mathrm{Bin}(n, p) \sim N(cnp, c^2)$ $X_1, X_2, \dots, X_k \sim \mathrm{Bin}(n, p) \Longrightarrow X_1 + X_2 + \dots + X_k \sim \mathrm{Bin}(nk, p)$	

Continuous Uniform distribution $X \sim U[a, b]$

CDF	$F_X(x) = P(X < x) = \frac{x - a}{b - a} I_{x \geq a}$
PDF	$f_X(x) = P(X = x) = \frac{1}{b - a} I_{a \leq x \leq b}$
Expected value & Variance	$E\,X = \frac{a + b}{2}, E\,X^2 = \frac{a^2 + ab + b^2}{3}, \mathrm{Var}\,X = \frac{(b - a)^2}{12}$

Geometric distribution $X \sim G(p)$ <u>"number of Bernoulli trials till the first success"</u>

CDF	$F_X(x) = P(X < x) = P(X \leq x - 1) = 1 - q^{x-1}$
PMF	$P(X = k) = G(k) = pq^{k-1}$
Expected value & Variance	$E\,X = \frac{1}{p}, E\,X^2 = \frac{q + 1}{p^2}, \mathrm{Var}\,X = \frac{q}{p^2}$
Lack of memory	$P(X > a + b \mid X > a) = P(X > b) = q^b$

Poisson distribution $X \sim \mathrm{Po}(\lambda), \lambda > 0$ <u>"number of times an event occurs in some interval"</u>

CDF	$X \sim \mathrm{Po}(\lambda), \lambda > 0$
PMF	$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$
Expected value & Variance	$E\,X = \mathrm{Var}\,X = \lambda$
$X \sim \mathrm{Po}(\lambda), Y \sim \mathrm{Po}(\theta) \Rightarrow X + Y \sim \mathrm{Po}(\lambda + \theta)$	

Exponential distribution $X \sim Exp(\lambda), \lambda > 0$

CDF	$F_X(x) = 1 - e^{-\lambda x} I_{x>0}$
PDF	$f_X(x) = \lambda\, e^{-\lambda x} I_{x>0}$
Expected value & Variance	$E\,X = \frac{1}{\lambda}, E\,X^2 = \frac{2}{\lambda^2}, Var\,X = \frac{1}{\lambda^2}$
Lack of memory	$E\,X = \frac{1}{\lambda}, E\,X^2 = \frac{2}{\lambda^2}, Var\,X = \frac{1}{\lambda^2}$
$X \sim U[a,b] \Longrightarrow -\ln X \sim Exp(1)$ $X \sim Exp(\lambda) \Longrightarrow E\,X^k = \frac{k!}{\lambda^k}$	

Normal distribution $X \sim N(\mu, \sigma^2), \sigma > 0$

$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$	**PDF**
$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-0.5t^2} dt = 1$	**Area under the curve is always = 1**
$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-0.5t^2} = 0.5 + \Phi_0(x)$	**Definition of Φ** **Relation between Φ and Φ_0**
$\Phi_0(x) = \int_{0}^{x} \frac{1}{\sqrt{2\pi}} e^{-0.5t^2} = 0.5 erf\,(\frac{x}{\sqrt{2}})$	**Definition of Φ_0** **Relation between Φ_0 and erf**
$\Phi(-x) = 1 - \Phi(x), \Phi_0(-x) = -\Phi_0(x)$	**Negative arguments**
$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < T < \frac{b-\mu}{\sigma}\right)$ $= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$	**Probability using Φ**
$E\,X = \mu, E\,X^2 = \sigma^2 + \mu^2, Var\,X = \sigma^2$	**Expected value and Variance**
$\Phi(x) = c \Longrightarrow x = \Phi^{-1}(c) = \sqrt{2}\, erf^{-1}(2c - 1)$	**Inverse Φ function** $\boldsymbol{erf^{-1}(x) = inverf(x)}$ **in Wolfram**
$f_X(\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\det \Sigma}} \exp(-0.5(\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}))$ $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \cdots \\ \mu_n \end{pmatrix}, \mu_i = E\,X_i$ $\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \cdots & \cdots & \cdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}, \sigma_{ij} = Cov(X_i, X_j)$	**Multivariate normal distribution**

$X \sim Bin(n,p) \Longrightarrow \frac{X-np}{\sqrt{npq}} = Y, Y \sim N(0,1)$	**Relation between Binomial and Normal distribution (for large n)**
X, Y are independent $\Leftrightarrow X, Y$ are uncorrelated $\Leftrightarrow Cov(X,Y) = 0, \rho_{X,Y} = 0$ $Z = aX + bY + c \Longrightarrow Z \sim N(E\,Z, Var\,Z)$ $X \sim N(0,1) \Longrightarrow E\,X^{2n-1} = 0, E\,X^{2n} = (2n-1)!! = (1)(3)(5)\dots(2n-1)$ $\binom{\xi}{\eta} \sim N(\mu, \sigma^2) \Longrightarrow (\xi \mid \eta = k) \sim N(\mu_\xi + \rho_{\xi,\eta}\left(\frac{\sigma_\xi}{\sigma_\eta}\right)(k-\mu_\eta), \sigma_\xi^2(1-\rho_{\xi,\eta}^2))$	

Gamma distribution $X \sim Gam(\alpha, \lambda)$

PDF	$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\lambda x}\, I_{x\geq 0}, \alpha > 0, \lambda > 0$ $\Gamma(x) = \int_0^{+\infty} t^{\alpha-1} e^{-t}$
Expected value & Variance	$E\,X = \frac{\alpha}{\lambda}, E\,X^2 = \frac{\alpha(\alpha+1)}{\lambda^2}, Var\,X = \frac{\alpha}{\lambda^2}$
Characteristic function	$\phi_X(t) = \frac{\lambda^\alpha}{(\lambda - it)^\alpha}$
$X_1 \sim Gam(\alpha_1, \lambda), X_2 \sim Gam(\alpha_2, \lambda), X_1, X_2$ are independent $\Longrightarrow X_1 + X_2 \sim Gam(\alpha_1 + \alpha_2, \lambda)$	

Chi-squared distribution $X \sim \chi_n^2$

$\chi_n^2 = X_1^2 + X_2^2 + \cdots + X_n^2, X_i \sim N(0,1), X_i$ are independent	
PDF	$f_{\chi_n^2}(x) = \frac{1}{2^{0.5n}\,\Gamma(0.5n)} x^{0.5n-1} e^{-0.5x} \sim Gam(0.5n, 0.5)\, I_{x>0}$
Expected value & Variance	$E\,X = n, Var\,X = 2n$

Students t-distribution $X \sim t_n$

$t_n = \frac{X}{\sqrt{\frac{1}{n}\sum_{i=1}^n X_i^2}} = \frac{X}{\sqrt{\frac{Y}{n}}}, X \sim N(0,1), Y \sim \chi_n^2$	
PDF	$f_{t_n}(x) = \frac{\Gamma\left(\frac{x+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$
Expected value & Variance	$E\,X = 0$ for n > 1 $Var\,X = \frac{n}{n-2}$ for n > 2