Bernoulli Distribution. $X \sim Bern(p)\ on\ \{0, 1\}$

$P(X = x) = p^x(1-p)^{1-x}$ $\quad E(X) = E(X^2) = p$ $\quad Var(X) = p(1-p)$

Binomial Distribution. $X \sim Bin(n, p)\ on\ \{0, 1, \dots, n\}$ (X = # of successes in *n* independent trials, with replacement)

$P(X = x) = \binom{n}{x} p^x(1-p)^{n-x}$ $\quad P(X \geq x) = (1-p)^x$ $\quad E(X) = np$ $\quad Var(X) = np(1-p)$

$Mode(X) = m = \lfloor np + p \rfloor = \begin{cases} m & np + p \notin \mathbb{Z} \\ m, m-1 & np + p \in \mathbb{Z} \end{cases}$ $\quad Multinomial: P(X = x) = \binom{n}{x_1, x_2, \dots, x_k} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$

Geometric Distribution. $X \sim Geom(p)\ on\ \{1, 2, \dots\}$ (X = # of trials until the first success)

$P(X = x) = pq^{x-1}$ $\quad P(X > x) = q^x$ $\quad E(X) = \frac{1}{p}$ $\quad E(X^2) = \frac{2-p}{p^2}$ $\quad Var(X) = \frac{q}{p^2}$

$X \sim Geom(p)\ on\ \{0, 1, \dots\}$ (X = # of failures before the first success)

$P(X = x) = pq^x$ $\quad P(X \geq x) = q^x$ $\quad E(X) = \frac{q}{p}$ $\quad E(X^2) = \frac{q(2-p)}{p^2}$ $\quad Var(X) = \frac{q}{p^2}$

Poisson Distribution. $X \sim Pois(\mu)\ on\ \{0,1, \dots\}$

$P(X = x) = e^{-\mu}\frac{\mu^x}{x!}$ $\quad P(X = 0) = e^{-\mu}$ $\quad E(X) = Var(X) = \mu$ $\quad Mode(X) = m = \lfloor np \rfloor = \begin{cases} m & m \notin \mathbb{Z} \\ m, m-1 & m \in \mathbb{Z} \end{cases}$

Relationship: $X \sim Pois(\mu_X), Y \sim Pois(\mu_Y), S = X + Y \sim Pois(\mu_X + \mu_Y)$.

Discrete Uniform Distribution. $X \sim Unif(1, \dots, n)\ on\ \{0,1, \dots, n\}$

$P(X = x) = \frac{1}{n}$ $\quad P(X \leq x) = \frac{x}{n}$ $\quad E(X) = \frac{n+1}{2}$ $\quad E(X^2) = \frac{(2n+1)(n+1)}{6}$ $\quad Var(X) = \frac{n^2-1}{12}$

Standard Uniform Distribution. $T \sim Unif(0,1)\ on\ the\ range\ (0,1)$

$f_T(t) = 1$ $\quad F_T(t) = P(T < t) = t$ $\quad E(T) = \frac{1}{2}$ $\quad E(T^2) = \frac{1}{3}$ $\quad Var(T) = \frac{1}{12}$

Relationship: $X \sim Unif(0,1), Y = 1 - X \sim Unif(0, 1), Z = -logX \sim Exp(1)$.

General Uniform Distribution. $T \sim Unif(a, b)\ on\ the\ range\ (a, b)$

$f_T(t) = \frac{1}{b-a}$ $\quad F_T(t) = P(T < t) = \frac{t-a}{b-a}$ $\quad E(T) = \frac{a+b}{2}$ $\quad E(T^2) = \frac{a^2+ab+b^2}{3}$ $\quad Var(T) = \frac{(b-a)^2}{12}$

Standard Normal Distribution. $T \sim Norm(0,1)\ on\ the\ range\ (-\infty, \infty)$

$f_T(t) = \phi(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}$ $\quad F_T(t) = P(T < t) = \Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx$ $\quad E(X) = 0$ $\quad E(X^2) = Var(X) = 1$

$P(-1 < T < 1) \approx 68\%$ $\quad P(-2 < T < 2) \approx 95.4\%$ $\quad P(-3 < T < 3) \approx 99.7\%$ $\quad P(-1.96 < T < 1.96) \approx 95\%$

General Normal Distribution. $T \sim Norm(\mu, \sigma^2)\ on\ the\ range\ (-\infty, \infty)$

$$f_T(t) = \frac{1}{\sigma}\phi\left(\frac{t-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \qquad P(a \le T \le b) \approx \Phi\left(\frac{b + \frac{1}{2} - \mu}{\sigma}\right) - \Phi\left(\frac{a - \frac{1}{2} - \mu}{\sigma}\right)$$

Relationship: $X \sim Norm(0,1), Z = \mu + \sigma X \sim Norm(\mu, \sigma^2)$.

$X \sim Norm(\mu_X, \sigma_X^2), Y \sim Norm(\mu_Y, \sigma_Y^2), Z = aX + bY \sim Norm(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2)$.

Exponential Distribution. $T \sim Exp(\lambda)$ $on\ the\ range$ $(0, \infty)$

$$f_T(t) = \lambda e^{-\lambda t} \qquad F_T(t) = P(T < t) = 1 - e^{-\lambda t} \qquad P(T > t) = e^{-\lambda t} \qquad E(T) = \frac{1}{\lambda} \qquad Var(T) = \frac{1}{\lambda^2}$$

$$Median(T) = \frac{\log 2}{\lambda} \qquad P(a < T < b) = F(b) - F(a) = P(T > a) - P(T > b) \qquad P(X < Y) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

Relationship: $X_1 \sim Exp(\lambda_1),\ X_2 \sim Exp(\lambda_2), Y = min(X_1, X_2) \sim Exp(\lambda_1 + \lambda_2)$. $X \sim Exp(\lambda), Y = cX \sim Exp(\lambda/c)$.

Memoryless Property: $P(T > w + t \mid T > w) = P(T > t) = e^{-\lambda t}$.

Conditional Expectation Addition Rule. $E(X + Y \mid A) = E(X \mid A) + E(Y \mid A)$

Rule of Average Conditional Expectations. $E(Y) = E[E(Y \mid X)]$

Conditioning Formulae: Discrete Case. $P(X = x, Y = y) = P(X = x) \cdot P(Y = y \mid X = x)$

$$P(Y = y) = \sum_{all\ x} P(Y = y \mid X = x) \cdot P(X = x) \qquad E(Y \mid X = x) = \sum_{all\ y} y \cdot P(Y = y \mid X = x)$$

$$P(Y \in A \mid X = x) = \sum_{y \in A} P(Y = y \mid X = x) \qquad E(Y) = \sum_{all\ x} E(Y \mid X = x) \cdot P(X = x)$$

Conditioning Formulae: Density Case. $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y \mid X = x)$

$$f_Y(y) = \int f_Y(y \mid X = x) \cdot f_X(x)dx \qquad E(Y \mid X = x) = \int y \cdot f_Y(y \mid X = x)dy$$

$$P(Y \in A \mid X = x) = \int_A f_Y(y \mid X = x)dy \qquad E(Y) = \int E(Y \mid X = x) \cdot f_X(x)dx$$

Covariance. $Cov(X, Y) = E([X - E(X)][Y - E(Y)]) = E(XY) - E(X)E(Y)$

$Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$ $\qquad X \perp Y \nLeftarrow Cov(X, Y), X \perp Y \Rightarrow Cov(X, Y)$

Covariance of Indicators. I_A, I_B are indicators of Event A and B:

$$E[I_A] = P(A) \qquad E[I_B] = P(B) \qquad E[I_AI_B] = P(AB) \qquad Cov(I_A, I_B) = P(AB) - P(A)P(B)$$

Sum of Dependent and Identically-Distributed Indicators. Let $X = I_1 + \cdots + I_n$:

$$Var(X) = \sum_i Var(I_i) + \sum_{i \neq j} Cov(I_i, I_j) = nVar(I_1) + n(n-1)Cov(I_1, I_2)$$

$$= E(X^2) - [E(X)]^2 = nE(I_1) + n(n-1)E(I_1, I_2) - [nE(I_1)]^2$$

Bilinearity of Covariance. $Cov(X, Y + Z) = Cov(X, Y) + Cov(X, Z)$

Conditional Covariance. $Cov(X, Y) = E[Cov(X, Y \mid Z)] + Cov[(X \mid Z), E(Y \mid Z)]$

Conditional Independence. $P(X = x, Y = y \mid Z = z) = P(X = x \mid Z = z)\, P(Y = y \mid Z = z)$

$P(\, Y = y \mid X = x, Z = z) = P(Y = y \mid Z = z)$

Minimum and Maximum of Discrete Random Variables. Let $M = \min(X, Y)$, $N = \max(X, Y)$.

$P(M = m) = P(X = m, Y = m) + P(X > m, Y = m) + P(X = m, Y > m)$

$P(M > m) = P(X > m, Y > m)$ $\quad$ $P(M \leq m) = 1 - P(M > m)$ $\quad$ $P(M = m) = P(M > m - 1) - P(M > m)$

$P(N = n) = P(X = n, Y = n) + P(X < n, Y = n) + P(X = n, Y < n)$

$P(N \leq n) = P(X \leq n, Y \leq n)$ $\quad$ $P(N > n) = 1 - P(N \leq n)$ $\quad$ $P(N = n) = P(N \leq n) - P(N \leq n - 1)$

Relationship between Minimum and Maximum: $P(N = n) = P(X = n) + P(Y = n) - P(M = n)$.

Minimum and Maximum of Continuous Random Variables. Let $X = \min(T_1, \ldots, T_n)$, $Y = \max(T_1, \ldots, T_n)$.

$P(X > x) = [P(T > x)]^n$ $\quad$ $F_X(x) = 1 - [P(T > x)]^n$ $\quad$ $P(Y > y) = 1 - [P(T \leq y)]^n$ $\quad$ $F_Y(y) = [P(T \leq y)]^n$

Relationship between PDF, CDF. $f_X(x) = \frac{d}{dx}[F_X(x)] = -\frac{d}{dx}[P(X > t)]$ $\quad$ $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$

Markov's Inequality: If $X \geq 0$, then $P(X \geq a) \leq \dfrac{E(X)}{a}, \forall a > 0$. **Chebychev's Inequality:** $P(|X - \mu| \geq k\sigma) \leq \dfrac{1}{k^2}$.

De Morgan's Laws: $P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - P\left(\bigcap_{i=1}^{n} \overline{A_i}\right), P\left(\bigcap_{i=1}^{n} A_i\right) = 1 - P\left(\bigcup_{i=1}^{n} \overline{A_i}\right)$. **Boole's Inequality:** $P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$.

Multiplication Rule: $P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2 \mid A_1)\ldots P(A_n \mid A_1 A_2 \ldots A_{n-1})$.

Inclusion-Exclusion Formula: $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i} P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k) - \ldots + (-1)^{n+1} P(A_1 A_2 \ldots A_n)$.