Rules of **Probability**

$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$$

$$\begin{aligned} P(X \cap Y) &= P(X) \cdot P(Y \mid X) \\ &= P(Y) \cdot P(X \mid Y) \end{aligned}$$

$$\begin{aligned} P(Y) &= P(Y \cap X) + P(Y \cap\, !X) \\ &= P(X) \cdot P(Y \mid X) + P(!X) \cdot P(Y \mid !X) \\ &= P(Y) \cdot P(X \mid Y) + P(Y) \cdot P(!X \mid Y) \end{aligned}$$

$$\begin{aligned} P(X \mid Y) &= \frac{P(X \cap Y)}{P(Y)} \\ &= \frac{P(X) \cdot P(Y \mid X)}{P(Y)} = \frac{P(Y) \cdot P(X \mid Y)}{P(Y)} \\ &= \frac{P(X) \cdot P(Y \mid X)}{P(Y \mid X) \cdot P(X) + P(Y \mid !X) \cdot P(!X)} \\ &= \frac{P(Y) \cdot P(X \mid Y)}{P(Y \mid X) \cdot P(X) + P(Y \mid !X) \cdot P(!X)} \end{aligned}$$

$$P(X \cap Y \cap Z) = P(Y) \cdot P(X \mid Y) \cdot P(Z \mid X \cap Y)$$

$$P(X \cup Y \cup Z) = P(X) + P(Y) + P(Z) - P(X \cap Y) - P(X \cap Z) - P(Y \cap Z) + P(X \cap Y \cap Z)$$

Independent

$$\begin{aligned} P(X \cap Y) &= P(X) \cdot P(Y) \\ &= P(Y) \cdot P(X) \end{aligned}$$

$$P(X \mid Y) = P(X)$$

discrete

1. **probability mass function (PMF)**
   list all possible probabilities of x
   $p_X(x) = P(X = x)$
   (a) probability of x must be in [0, 1]
   (b) sum of all probabilities must equal to 1
   (c) R: hist(x)
   with x in x-axis and $p_X(x)$ in y-axis
2. **cumulative distribution function (CDF)**
   probability that a random variable is less than or equal to x
   $F_X(x) = P(X \leq x)$
   (a) $\lim_{x \to -\infty} F_X(x) = 0$
   (b) $\lim_{x \to \infty} F_X(x) = 1$
   (c) non-decreasing
3. $E_X(x)$ **= theoretical mean**
   For random variable x,
   $E_X(x) = \sum_{all\ x} x \cdot p_X(x)$
   For function g(x),
   $E_X(g(x)) = \sum_{all\ x} g(x) \cdot p_X(x)$

continuous

1. **probability density function (PDF)**
   list all possible probabilities of x
   $f_X(x) = \frac{\partial}{\partial x}\ F_X(x)$
   (a) $f_X(x) \geq 0$ for all x
   (b) area under the curve is equal to 1
   (c) R: plot(density(x))
   with x in x-axis and $f_X(x)$ in y-axis
2. **cumulative distribution function (CDF)**
   probability that a random variable is less than or equal to x
   $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)\ dx$
   $P(a \leq X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\ dx$
   (a) $\lim_{x \to -\infty} F_X(x) = 0$
   (b) $\lim_{x \to \infty} F_X(x) = 1$
   (c) non-decreasing
3. $E_X(x)$ **= theoretical mean**
   For random variable x,
   $E_X(x) = \int_{-\infty}^{\infty} x \cdot f_X(x)$
   For function g(x),
   $E_X(g(x)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)$

1. Rules of **Expectation**
   (a) $E(aX + b) = a \cdot E(X) + b$
   (b) $E(Y \mid X) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y \mid x)\ dy$
   (c) $E(XY) = E(E(XY \mid X)) = E(X \cdot E(Y \mid X))$
   (d) $E(XY) == E(X) \cdot E(Y)$ if X and Y are independent
   (e) $E(XY \mid X) = X \cdot E(Y \mid X)$
   (f) $E(Y) = E(E(Y \mid X))$
   (g) $E_X(x^2) = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)$
2. Rules of **Variance**
   (a) $var(X) = \sigma^2 = E(X^2) - E(X)^2$
   (b) $var(X + c) = var(X)$
   (c) $var(cX) = c^2 \cdot var(X)$
   (d) $var(X + Y) = var(X) + var(Y) + 2 \cdot cov(X, Y)$
   (e) $var(X + Y) = var(X) + var(Y)$ if X and Y are independent
3. Rules of **Standard Deviation**
   (a) $\sigma = \sqrt{var(X)}$
4. Rules of **Normal Distribution**
   (a) $E(X) = 0$
   (b) $E(X^2) = 1$
   (c) $E(X)^2 = 0$
   (d) $\sigma = 1$
   (e) $var(X) = 17$
5. Rules of **Joint Distribution**
   Remember the rules of probability above (for discrete) applies here.
   (a) **Marginal**
   $$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y)\ dy \\ &= \int_{-\infty}^{\infty} f_X(x) \cdot f_{Y|X}(y \mid x)\ dy \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y)\ dx \\ &= \int_{-\infty}^{\infty} f_Y(y) \cdot f_{X|Y}(x \mid y)\ dx \end{aligned}$$
   (b) $f_{X|Y} = \frac{f_{X,Y}(X,Y)}{f_Y(y)}$
   (c) $f_{X,Y} == f_X(x) \cdot f_Y(y)$ if X and Y are independent
6. Rules of **Covariance**
   Measure the strength of linear relationship b/w 2 variables
   (a) $cov(X, Y) = E((X - \mu_X)(Y - \mu_Y)) = E(XY) - E(X) \cdot E(Y)$
   (b) $cov(X, X) = var(X)$
   (c) $cov(aX, bY) = ab \cdot cov(XY)$
   (d) $cov(X, Y + Z) = cov(X, Y) + cov(X, Z)$
7. Rules of **Correlation**
   Normalize Covariance by scaling it b/w -1 and 1
   (a) $corr(X, Y) = \frac{cov(X,Y)}{\sigma_X \cdot \sigma_Y}$

Distribution	Type	Description	PMF or PDF	$E_X(x)$	$var(x)$
Bernoulli	Discrete	Binomial with n = 1	$p^x(1-p)^{1-x}$	p	$1-p$
Binomial	Discrete	n independent Bernoulli trials	$\binom{n}{x}p^x(1-p)^{n-x}$	np	$np(1-p)$
Geometric	Discrete	get first success in independent Bernoulli trials	$(1-p)^{x-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Negative Binomial	Discreite	get r^{th} success in independent Bernoulli trials	$\binom{x-1}{n-1}p^r(1-p)^{x-r}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$