	Discrete	Continuous
Support	Countable set of values	Uncountable set of values
Probabilities	pmf: $P(X = x) = p(x)$	$P(X = x) = 0 \neq f(x)$ for all x $P(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a)$
Joint	$p_{XY}(x, y) = \mathbb{P}(X = x, Y = y)$	$\mathbb{P}(X \in [a, b], Y \in [c, d]) = \int_a^b \int_c^d f_{XY}(x, y)\, dx\, dy$
Marginal	$p_X(x) = \sum_y p_{XY}(x, y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy$
Conditioning	$p_{X\vert Y}(x\vert y) = p_{XY}(x, y)/p_Y(y)$	$f_{X\vert Y}(x\vert y) = f_{XY}(x, y)/f_Y(y)$
Independence	$p_{XY}(x, y) = p_X(x)p_Y(y)$	$f_{XY}(x, y) = f_X(x)f_Y(y)$
Expected Value	$\mu_X = \mathbb{E}[X] = \sum_x xp(x)$ $\mathbb{E}[g(X)] = \sum_x g(x)p(x)$ $\mathbb{E}[g(X, Y)] = \sum_x \sum_y g(x, y)p(x, y)$	$\mu_X = \mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)\, dx$ $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\, dx$ $\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f_{XY}(x, y)\, dx\, dy$

Table 1: Differences between Discrete and Continous Random Variables

Probability Mass Function $p(x)$	Probability Density Function $f(x)$
Discrete Random Variables	Continuous Random Variables
$p(x) = P(X = x)$	$f(x) \neq P(X = x) = 0$
$p(x) \geq 0$	$f(x) \geq 0$
$p(x) \leq 1$	$f(x)$ *can* be greater than one!
$\sum_x p(x) = 1$	$\int_{-\infty}^{\infty} f(x)\, dx = 1$
$F(x_0) = \sum_{x \leq x_0} p(x)$	$F(x) = \int_{-\infty}^{x} f(t)\, dt$

Table 2: Probability mass function (pmf) versus probability density function.

Definition of R.V.	$X: S \to \mathbb{R}$ (RV is a fixed function from sample space to reals)
Support	Set of all values the RV can take
CDF	$F(x_0) = P(X \leq x_0)$
Definition of Variance	$\sigma_X^2 = \text{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$
Shortcut for Variance	$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
Definition of Std. Dev.	$\sigma_X = \sqrt{\sigma_X^2}$
Covariance	$\sigma_{XY} = \text{Cov}(X,Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right]$
Cov. and Independence	X, Y indep. $\Rightarrow \text{Cov}(X,Y) = 0$ but $\text{Cov}(X,Y) = 0 \not\Rightarrow X, Y$ indep.
Functions and Independence	X, Y indep. $\Rightarrow g(X), h(Y)$ indep. where g, h are any functions
Shortcut for Covariance	$\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
Definition of Correlation	$\rho_{XY} = Corr(X,Y) = \sigma_{XY}/(\sigma_X \sigma_Y)$
Expectations of Functions	$\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$
Linear Functions of RVs	$\mathbb{E}[a + bX] = a + b\mathbb{E}[X]$ where a, b are constants and X is a RV
	$\text{Var}(a + bX) = b^2\text{Var}(X)$ where a, b are constants and X is a RV
	$\mathbb{E}[X_1 + \ldots + X_k] = \mathbb{E}[X_1] + \ldots \mathbb{E}[X_k]$ where $X_1, \ldots, X_k$ are *any* RVs
	$\text{Var}(X_1 + \ldots + X_k) = \text{Var}(X_1) + \ldots \text{Var}(X_k)$ if $X_1, \ldots, X_k$ are *independent* RVs
	$\text{Var}(aX + bY + c) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X,Y)$ for *any* RVs X, Y and constants a, b, c
	$\text{Cov}(a + bX, c + dY) = bd\text{Cov}(X,Y)$ for any constants a, b, c, d and RVs X, Y
	$\text{Cov}(X, Y + Z) = \text{Cov}(X,Y) + \text{Cov}(X,Z)$ for any RVs X, Y, Z

Table 3: Essential facts that hold for *all* random variables, continuous or discrete

	Sample Statistic	Population Parameter	Population Parameter
Setup	Sample from a population	Population viewed as list of objects	Population viewed as a RV
Mean	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$\mu_X = \frac{1}{N}\sum_{i=1}^{N} x_i$	Discrete $\mu_X = \sum_{x} xp(x)$ Continuous $\mu_X = \int_{-\infty}^{\infty} xf(x)\,dx$
Variance	$s_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$	$\sigma_X^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_X)^2$	$\sigma_X^2 = E\left[(X - E[X])^2\right]$
Std. Dev.	$s_X = \sqrt{s_X^2}$	$\sigma_X = \sqrt{\sigma_x^2}$	$\sigma_X = \sqrt{\sigma_x^2}$
Covariance	$s_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$	$\sigma_{XY} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_X)(y_i - \mu_Y)$	$\sigma_{XY} = E\left[(X - \mu_X)(Y - \mu_Y)\right]$
Correlation	$r_{XY} = s_{XY}/(s_X s_Y)$	$\rho_{XY} = \sigma_{XY}/(\sigma_X\sigma_Y)$	$\rho_{XY} = \sigma_{XY}/(\sigma_X\sigma_Y)$