Cont: x can be any value in range, uncountably infinite. If diff of neighbour vals similar, consider cont.
Need to consider intervals of x for a probability.
P(-infty < X < infty) = 1, < same as <=, > same as >=
Mode(X) = arg_max$_x$ $f_X(x)$

Entropy: $$H(X) = -\int_x f_X(x)\log[f_X(x)]\mathrm{d}x.$$ Mean: $$\mathbb{E}(X) = \int_x x\, f_X(x)\mathrm{d}x.$$
Entropy can be negative for cont.

Variance: $$\mathrm{Var}(X) = \mathbb{E}[(X-\mu_X)^2] = \int_x (x-\mu_X)^2 f_X(x)\mathrm{d}x.$$

Median: AuC = 0.5, its empirically-based definition is the "middle value" after sorting the outcomes from the smallest to the largest value. Probably best for single decision for random outcome.
Quantile: Q(p) where P(X <= Q(p)) = p
Prediction interval: lower = Q((1-p)/2), upper = Q((1+p)/2)

Skewness: $$\mathrm{Skewness}(X) = \mathbb{E}\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^3\right] = \int_x \left(\frac{x-\mu_X}{\sigma_X}\right)^3 \cdot f_X(x)\mathrm{d}x.$$

	Left (negative skew)	Symmetric (0 skew)	Right (positive skew)
Typical dist:	mean<med<mode	All same	Mode < med < mean

PDF: $f_X(x)$, CDF: $F_X(x)$ = P(X <= x) $$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t.$$
CDF has no units, PDF has units 1 / units of x
CDF strictly increasing, must be between 0 & 1, x -> -infty, CDF -> 0, vice-versa

Survival: $S_X(x)$ = P(X > x) = 1 - $F_X(x)$, prob of surviving after x. CDF upside down

Quantile: Q(p) = $F^{-1}(p)$ for 0 <= p <= 1. Inverse of CDF.

X ~ Unif(a,b), PDF = 1 / (b-a), mean = (a+b) / 2, var = (b - a)^2 / 12 a <= x <= b

X ~ Norm(mu, var), PDF = $$f_X(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$
Mean = mu, Var = var   -infty < x < infty

X ~ Log-Norm(mu, var), pdf same but log(x), mu = $e^{mu+var/2}$ x >= 0
Var = $$\mathrm{Var}(X) = \exp\left[2\left(\mu+\sigma^2\right)\right] - \exp\left(2\mu+\sigma^2\right).$$

X ~ Exp(lambda), PDF = lambda * $e^{-lambda*x}$ x >= 0, wait time for event, memoryless meaning regardless of how much time has passed, same dist
mean = 1/lambda, var = 1/lambda^2 NOTE: lambda = 1 / beta (rate vs. wait time)

X ~ Beta(alpha, beta), PDF = dbeta(x, alpha, beta) - support between 0 and 1.
Mean = alpha / (alpha + beta), Var = $$\mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

X ~ Weibull(lambda, k), PDF = $$f_X(x \mid \lambda, k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}\exp^{-(x/\lambda)^k}$$ x >=0
Event more likely the longer you wait, time until event.
Mean = $$\mathbb{E}(X) = \lambda\Gamma\left(1+\frac{1}{k}\right).$$, Var = $$\mathrm{Var}(X) = \lambda^2\left[\Gamma\left(1+\frac{2}{k}\right) - \Gamma^2\left(1+\frac{1}{k}\right)\right].$$

X ~ Gamma(k, theta), PDF = $$f_X(x \mid k, \theta) = \frac{1}{\Gamma(k)\theta^k}x^{k-1}\exp(-x/\theta)$$ x>=0

Mean = k*theta, Var = k*theta^2
Bivariate case: $f_{X,Y}(x,y)$ compute volume under density surface

Identify surface area, determine probability for event

For conditional, P(A) = 0 then use density instead of probability.
$f_{X|X>=2500}(x)$ = 0 for x < 2500
When P(B) = 0, can use general formula
$f_{Y|X}(y)$ = ($f_{Y,X}(y,x)$) / ($f_X(x)$)

If x and y indep, multiply indivi. Marginal probs.

RS: collection of RV's/outcomes, iid
MLE: estimate pop. parameters, relies on n obs from pop.
Given data and some family, finds values of params that fit data the best.
Since iid, multiply each indiv. Pdf which is also likelihood since equiv.
Use d*() to compute values of potential dist.
We take log since the likelihood can be extremely small. Empirical solution
Ex: exp_values <- tibble(
  possible_betas = seq(5, 50, 0.5),
  likelihood = map_dbl(1 / possible_betas, ~ prod(dexp(sample_n30$values, .))),
  log_likelihood = map_dbl(1 / possible_betas, ~ log(prod(dexp(sample_n30$values, .)))))

Deterministic: certain outcome, stochastic is opposite
set.seed(#), np.random.seed(#)
sample(outcomes, size = n, replace = TRUE, prob = probs)
np.random.choice(a = outcomes, size = n, p = probs)
r*() vs scipy.stats.dist.rvs()

Sample mean = sum (obs) / n, Sample var = sum(obs - mean) / (n - 1)
For PMF, use table() or tabyl()
Theoretical: use formula, Empirical: use built-in functions (mean, sd, var etc…)
LLN: as n go up, empirical mean approaches true mean we estimating.

iid if 1. Each pair of observations are independent, and
2. each observation comes from the same distribution.
Assuming no additional sampling assumptions.