Chris Piech
CS 109

Handout
Oct 12, 2018

All Discrete Distributions

Bernoulli
An indicator variable that takes on the value 1 (“success") or 0. Often the variable is defined to be 1 if an underlying event has occurred, 0 otherwise.

Notation	$X \sim \text{Bern}(p)$
Parameters:	p: The probability of the variable being 1
Range(X):	$\{0, 1\}$
pmf:	$\Pr(X = k) = \begin{cases} p & \text{if k = 1} \\ (1-p) & \text{if k = 0} \end{cases}$
E[X]:	p
Var(X):	$p(1-p)$

Note: Sometimes in machine learning algorithms a derivable version of the PMF is used: $f(X = k) = p^k(1-p)^{1-k}$. We will talk about that later.

Binomial
A variable which represents the number of successes in a fixed number of independent trials. The probability of success must be the same for each trial.

Notation	$X \sim \text{Bin}(n, p)$
Parameters:	n : the number of trials p: the probability of success in each trial
Range(X):	$\{0, 1, \ldots, n\}$
pmf:	$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
E[X]:	np
Var(X):	$np(1-p)$

Note: $\text{Bin}(1, p) = \text{Bern}(p)$

Poisson

The number of events occurring in a fixed interval of time or space if these events occur independently with a constant rate.

Notation	$X \sim \text{Poi}(\lambda)$
Parameters:	λ : the rate of events in one interval
Range(X):	$\{0, 1, \ldots, \infty\}$
pmf:	$\Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
E[X]:	λ
Var(X):	λ

Note: The Poisson is the number of events in an interval of time. The Exponential is a continuous distribution which is the time until the next event. They have the same parameter (λ).

Geometric

The number of independent Bernoulli trials until the first success.

Notation	$X \sim \text{Geo}(p)$
Parameters:	p: the probability of success of each trial
Range(X):	$\{1, 2, \ldots, \infty\}$
pmf:	$\Pr(X = k) = (1 - p)^{k-1} p$
E[X]:	$1/p$
Var(X):	$\frac{1-p}{p^2}$

Negative Binomial
The number of Bernoulli trials until the first r success.

Notation	$X \sim \text{NegBin}(r, p)$
Parameters:	p: the probability of success of each trial
Range(X):	$\{r, r+1, \ldots, \infty\}$
pmf:	$\Pr(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$
E[X]:	r/p
Var(X):	$\frac{r(1-p)}{p^2}$

Those are all the discrete random variable types we are going to teach you in CS109. Here is an example of another random variable, which is outside the scope of the class. Even though you won't have to memorize it, we would like you to be able to use a PMF when given one.

Zipf
The rank index of a chosen item (eg 1st most common, 2nd etc).

Notation	$X \sim \text{Zipf}(s)$
Parameters:	s: the value of the exponent characterizing the distribution
Range(X):	$\{1, 2, \ldots, \infty\}$
pmf:	$\Pr(X = k) = \frac{1}{k^s \cdot H}$

Note: H is the Nth harmonic number where N is the size of the language. It plays the role of a normalizing constant.