CSE312 Final Notes

**Counting:**

* , where A is desired event and is the sample space.
* Product rule: if 1st step has choices and 2nd step has choices, 2 steps together have choices.
* Permutation (order matters): pick objects from and permutes
* Combination (order doesn’t matter): choose objects from / choose

Identity:

Binomial Theorem: Corollary:

* Complementing: P(contains at least 1) = 1 – P(contains 0)
* Inclusion-Exclusion: + single – pairs + triples – quads
* Pigeonhole Principle: If you have pigeons and holes, then some hole has 1 pigeon.

**Probability:**

* 2 events and are *mutually exclusive* if and only if
* Axioms of Probability:
* If and are mutually exclusive, then

Implications of Axioms:

* If , then
* Equally likely outcomes: for every
* Conditional Probability: suppose conditional probability of given

Chain rule:

* Law of Total Probability:
* Conditional Independence: and are conditionally independent if and only if
* Bayes’ Theorem: Corollary:
* Naive Bayes Classifier:
* Assumption: words in the email are conditionally independent given we know the email is spam/ham.
* and are fractions of spam/ham emails in training data.
* Laplace Smoothing for each
* Independent Events: and are independent if and only if

If , then and are independent if and only if .

**Discrete random variables and expectation:**

* Random Variable: numerical function of the outcome

(Discrete: countable number of possible values.)

* Independent Random Variables: Random Variables and are independent if and only if
* Probability Mass Function (pmf): Let be the set of outcomes and be an outcome:
* Expectation of a random variable:
* Linearity of Expectation: ,
* If and are independent,
* Indicator Random Variable: for
* Variance () and Standard deviation ():
* Let , then
* Theorem:
* If and are independent,
* Distributions:
* Uniform: if is equally likely to be any integer in .
* Bernoulli: is a random indicator variable with and .
* Binomial: is the sum of independent Bernoulli random variables such that for .
* Geometric: is independent Bernoulli trials with parameter until and including 1st success.
* for
* Poisson: when evets happen independently with average rate of per unit time.
* Summations:

**Continuous Random Variables**

* Continuous Random Variables: takes values from an uncountable set.
* Probability Density Function:
* Cumulative Distribution Function: , thus
* Distributions:
* Uniform: indicates each real number from to be equally likely.
* ,
* Exponential: represents the waiting time to the first success where is the average number of events per unit time.
* ,
* ,
* Memoryless: for any ,
* Normal: if has the probability density function of

,

* Standard normal:
* Closure of Normal Distribution: linear transformation of normal is still normal

Suppose , then .

* Reproductive property of Normal: Sum of normal distributions is still normal.

**Central Limit Theorem**

* Suppose are identical, independent distributed random variables with and , so we have the sample mean:

with and

Thus, as ,

* Same as let with and , in this case:
* Continuity Correction: when being estimated is discrete

**Tail Bounds**

* Markov’s Inequality:
* Chebyshev’s Inequality: Suppose and , then

for any

* Chernoff’s Bound: Suppose . Then for any ,

**Law of Large Numbers**

* Let be identical, independent distributed variables with common mean and variance . Let be the sample mean for a sample size . Then:
* Weak Law of Large Numbers:

for any ,

* Strong Law of Large Numbers:
* The strong law implies the weak law but not vice versa.

**Likelihood**

* Realization/sample of a random variable: the actual values observed.
* Let be realizations of random variable , we define the likelihood function to be the probability of seeing these data:
* If is discrete with mass function :
* If is continuous with density :
* Maximum Likelihood Estimator (MLE): maximizes the likelihood function, denote as .

Steps of finding MLE:

* Find likelihood and log-likelihood of data
* Take derivative of log-likelihood and find critical points
* Use second derivative test to show is a maximizer, that at , also check points of non-differentiability and boundary points.
* Bias: the bias of an estimator for the true parameter is defined as

.

An estimator is unbiased if and only if the bias of the estimator is .

**Confidence Intervals**

* is a confidence interval for if and only if

.