STAT 400 Study Guide

Cliff Sun

February 17, 2025

Definition 0.1. *An ordered range of r distinct things is called a permutation. The number of ways to number r objects in n positions is*

$$_nP_r \tag{1}$$

Definition 0.2. *The number of methods for partiioning n objects in r groups is*

$$\binom{n}{n_1\ n_2\ \dots} = \frac{n!}{n_1!n_2!\dots} \tag{2}$$

Definition 0.3. *The number of ways to take n objects out of r total is*

$$C_r^n \tag{3}$$

Definition 0.4. *The conditional probability of event A given event B is*

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{4}$$

Bayes theorem looks into partitions of an event space S. For example, a person can drive. The person can also drink. The event space of driving can be divided up by probabilities of drinking. The partitions must be mutually exclusive and encompassing over S. Bayes rule states

$$P[A] = \sum_{i=1}^{m} P(B_i \cup A) \tag{5}$$

This is just summing over the event space. Using equation (4),

$$P[A] = \sum_{i=1}^{m} P(B)P(A|B) \tag{6}$$

Definition 0.5. *Given Y is a random variable, then*

1. *The domain of Y is the outcome space*
2. *The range of Y is the space*

Definition 0.6. *The support of X is the domain that doesn't map to 0.*

Definition 0.7. *A <u>Bernoulli Experiment</u> is when an outcome is binary. If a random variable distributes like a Bernoulli distribution, then we say*

$$X \sim \text{Bernoulli(p)} \tag{7}$$

It has the following equation

$$f(x) = p^x(1-p)^{1-x}\ x = \{0, 1\} \tag{8}$$

Where x dictates how many times an event of probability p occurred. Note,

$$E[X] = p \tag{9}$$

$$Var[X] = p(1-p) \tag{10}$$

Definition 0.8. *A Binomial Distribution is repeated Bernoulli experiment. That is, we calculate the probability of an event of probability p occurring r out of n tries. Note,*

$$E[X] = np \tag{11}$$

$$Var[X] = np(1-p) \tag{12}$$

We say

$$X \sim \text{Binomial(n,p)} \tag{13}$$

Where n is the total number of trials and p is the probability of this event occurring. Implicitly, x is the number of times the event has occurred.

Definition 0.9. *A Geometric Distribution is the number of trials that must occur before the first success occurs. This is analagous to the inverse of the Bernoulli Experiment. We say that*

$$X \sim \text{Geom(p)} \tag{14}$$

Note,

$$E[X] = \frac{1}{p} \tag{15}$$

$$Var[X] = \frac{1-p}{p^2} \tag{16}$$

There are two types of equations,

$$f(x) = p(1-p)^{x-1} \tag{17}$$

means a total of x trials. In R studio, $x-1$ would be replaced with x' where x' is the number of failures before the first success is observed.

Definition 0.10. *A Negative Binomial Distribution is a generalization of the Geometric Distribution. That is, the number of trials needed until we observe r successes. Note, the pmf of this distribution is a function of the number of trials. We say that*

$$X \sim \text{NB(r,p)} \tag{18}$$

Where r is the number of successes. Note,

$$E[X] = \frac{r}{p} \tag{19}$$

$$Var[X] = \frac{r(1-p)}{p^2} \tag{20}$$

Note that there are a number of x trials, with $r-1$ failures distributed over $x-1$ trials. In R studio, we say

dnbinom(x, r, p)

Where x is the number of failures, and r is the number of successes. Note $x+r$ is the total number of trials.

Definition 0.11. *In R studio, the prefixes mean*

1. *d: Calculate pmf (NOT CDF)*
2. *p: Calculate cdf (NOT PMF, this is confusing)*

Definition 0.12. *A Hypergeometric distribution describes the number of successes in a sample of size n, where the sampling is done without replacement. Note*

$$E[X] = n\frac{r}{N} \tag{21}$$

Where r is the maximum number of successes, n is size of the sample, N is the size of the population. Note x is the number of actual observed successes.

Definition 0.13. *A Poisson Distribution is used in scenarios when we count the number of occurrences given a certain rate λ. Note,*

$$E[X] = \lambda \tag{22}$$

$$Var[X] = \lambda \tag{23}$$

We say that

$$X \sim \text{Pois}(\lambda) \tag{24}$$

Where x is the number of occurrences of an event.