Definition **1.1**  **The *mean* of a sample**

Definition **1.2**

***Variance* of a sample**

Definition **1.3**

**The *Standard Deviation* of a sample**

Definition **2.1**

**An *experiment***

is the process by which an observation is made

Definition **2.2**

**A *Simple Event***

is an event that cannot be decomposed, Each simple event corresponds to one and only one *SAMPLE POINT*

Definition **2.3**

**The *Sample Space***

Associated with an experiment is the set consisting of all sample points

**A diagram of a circle with black dots and letters

Description automatically generated**

Definition **2.4**

**A *Discrete Sample***

space is one that contains either a finite or a countable number of distinct sample points.

Definition **2.5**

**An** ***Event***

In a **Discrete Sample** space **S** is a collection of sample points that is any subset of S

Definition **2.6**

***Probability of A***

Suppose S is a sample space associated with an experiment. To every event A

in S ( A is a subset of S), we assign a number, P( A), called the probability of

A, so that the following axioms hold:

**AXIOM 1: P(A) 0**

**AXIOM 2: P(S) = 1**

**AXIOM 3: If … form a sequence of pairwise mutually exclusive events in S (that is, ∩ = if i**

The Sample Point Method

1. Define the experiment and clearly determine how to describe one simple event.
2. List the simple events associated with the experiment and test each to make certain that it cannot be decomposed. This defines the sample space s.
3. Assign reasonable probabilities to the sample points in S, making certain that P AND
4. Define the event of interest, A as a specific collection of sample points. (A sample point is in A if A occurs when the sample point occurs. Test *all* sample points in S to identify those in A.)
5. Find **P(A)**

Theorem ***2.1***

With *m* elements …, elements it is possible to form mn = m \* n pairs containing one element from each group.

Verification of the theorem can be seen by observing the rectangular table in Figure 2.9. There is one square in the table for each  *pair and hence a total of m\*n squares.*

Definition **2.7**

An ordered arrangement of r distinct objects is called a ***permutation***. The number of ways of ordering n distinct objects taken r at a time will be designated

by the symbol

Theorem **2.2**

***= n(n-1)(n-2)…(n-r+1)=***

We are concerened with the number of the ways of filling r positions with n distinct objects. Applying the extension of the mn rules, we see that the first object can be chosen in one of n ways. After the first is chosen, the second can be chosen in (n-1) ways, the third in (n-2), and the rth in (n-r+1) ways. Hence, the total number of distinct arrangements is

Expressed in terms of factorials,

Where n! = n(n-1)….(2)(1) and O! = 1.

Theorem ***2.3***

The number of ways of partitioning n distinct objects into ***k*** distinct groups containing objects, respectively, where each object appears in exactly one group and  
is

***N*** is the number of distinct arrangements of ***n*** objects in a row for a case in which rearrangement of the objects within a group does not count. For example, the letters a to l are arranged in three groups, where **= 3,**

Is one such arrangement.

The number of distinct arrangements of the ***n*** objects assuming all obkects are distinct is (Theorem 2.2) then equals the number of ways of partitioning the ***n*** objects into ***k*** groups(ignoring order within groups) multiplied by the number of ways of ordering the elements within each group. This application of the extended ***mn*** rule gives

Whereis the number of distinct arrangements of the objects in the group ***i***.

Solving for ***N***, we have

Definition ***2.8***

***Combinations***

The number of ***Combinations*** of ***N*** objects taken ***R***  at a time is the number of subsets each of size ***R*** that can be formed from ***N*** objects. This is denoted by  **or**

Theorem ***2.4***

The number of unordered subsets of size ***r*** chosen (without replacement) from ***n*** available objects is

The selection of ***r*** objects from a total of ***n*** is equivalent to partitioning the ***n*** objects into ***k = 2*** groups, the ***r***  selected, and the ***(n-r)*** remaining. This is a special case of the general partitioning problem dealt with in Theorem 2.3. In the presend case, ***k = 2, = r, and*** ***= (n – r)***and therefore,

Definition ***2.9***

The ***Conditional Probability***

Of an event ***A, given that*** event ***B*** has occurred is equal to

Provided that P(B) > 0. [The symbol P(A|B) is read “probability of A given B.”]

Definition ***2.10***

***Independent Events***

Two events A and B are said to be ***independent*** if any one of the following holds:

Otherwise the events are said to be ***Dependent.***

Theorem ***2.5***

**The Multiplicative Law of Probability**

The probability of the intersection of two events A and B is

If ***A and B*** are independent, then

***.***

The multiplicative law follow directly from Definition ***2.9,*** the definition of conditional probability.

Theorem ***2.6***

**The Additive Law of Probability** The probability of the union of two event ***A and B*** is

If ***A and B*** are Mutually exclusive***,*** and

The proof of the additive law can be followed by inspecting the Venn diagram in Figure 2.10

Notice that are mutually exclusive events. Further, B = ***(A***  whereare mutually exclusive events. Then, by Axiom 3,

and

The equality given on the right implies that ***P*** ***= P(B) – P(A***

*Substituting this expression for P(A****P(A)*** given in the left-hand equation of the preceding pair, we obtain the desired result:

Theorem ***2.7***

If A is an event, then

Observe that ***S = A*** because ***A*** and are mutually exclusive events, it follows that Therefore, and the result follows.

Definition ***2.11***

***Partition of S***

For some positive integer k, let the sets

**be such that**



Then the collection of sets  
 is said to be a partition of ***S***

Theorem ***2.8***

Assume that is a partition of ***S*** (see definition 2.11) such that  
Then for any event A

Any subset of ***A*** and ***S*** can be written as

Notice that, because ***{}*** is a partition of ***S*** if ***i j,***

And that (are mutually exclusive events. Thus,

Theorem ***2.9***

***Bayes’ rule***

Assume that ***{}*** is a partition of ***S*** (see definition 2.11) such that ***P() > 0, for i = 1,2,…,k. Then***

The proof follows directly from the definition of conditional probability and the law of total probability. Note that

Definition ***2.12***

A ***Random Variable*** is a real-valued function for which the domain is a sample space.

Definition **2.13**

***Random Sample***

Let ***N*** and ***n*** represent the numbers of elements in the population and sample, respectively. If the sampling is conducted in such a way that each of the samples has an **equal probability** of being selected, the sampling is said to be random, and the result is said to be a ***Random Sample***.

Definition **3.1**

***Random Variable Discrete***

A random variable ***Y*** is said to be ***discrete*** if it can assume only a finite or countably infinite number of distinct values.

Definition **3.2**

***sum of the probabilities of all sample points in S***

The probability that ***Y*** takes on the value y, ***,*** is defined as the ***sum of the probabilities of all sample points in S*** that are assigned the value y. We will sometimes denote  ***by***

Definition **3.3**

The ***Probability Distribution***

For a discrete variable Y can be represented by a formula, a table, or a graph that provides  
for all ***y***

Theorem **3.1**

For any discrete probability distribution, the following must be true:

1. ***0 p(y) 1 for all y***
2. where the summation is over all values of ***y*** with nonzero probability.

Definition ***3.4***

Let ***Y*** be a discrete random variable with the probability function ***p(y).*** The *expected value* of Y, E(Y), is defined to

Theorem **3.2**

Let ***Y*** be a discrete random variable with the probability function p(y) and g(Y) be a real-valued function of Y. Then the expected value of g(Y) is given by

We prove the result in the case where the random variable Y takes on the finite number of values ***,…, .*** because the function g(y)may not be one-to-one, supper that g(Y) takes on values ***,…, . ( where m*** It follows that g(Y) is a random variable such that I = 1, 2, …, m,

Thus by Definition 3.4,

Definition ***3.5***

If ***Y*** *is a random variable with mean* ***E(Y) = ,*** the variance if a random variable ***Y*** is defined to be the expected value of That is,

The ***Standard deviation*** of ***Y***  is the positive square root of ***V(Y).***

Theorem ***3.3***

***Let Y*** be a discrete random variable with probability function ***p(y)*** and ***c*** be a constant. Then ***E(c)=c.***

Consider the function ***g(Y)***  By Theorem 3.2,

But  
(Theorem 3.1)and, hence,  
***.***

Theorem ***3.4***

***Let Y*** be a discrete random variable with probability function ***p(y), g(y)*** be a function of ***Y*** and ***c*** be a constant. Then

***E[cg(Y)] = c E[g(Y)].***

By theorem 3.2,

***E[cg(Y)] =***

Theorem ***3.5***

***Let Y*** be a discrete random variable with probability function ***,***  be k functions of ***Y***. Then

**E[**

we will demonstrate the proof only for the case k = 2, but analogous steps will hold for finite ***k.*** By Theorem 3.2,

**E[**

=

=.

Theorem ***3.6***

Let ***Y*** be a discrete random variable with probability function ***p(y)*** and mean ***E(Y)= ;*** then

***V(Y) = =***

***= =***

***=*** (By Theorem 3.5).

Noting that is a constant and applying Theorems 3.4 and 3.3 to the second and third terms, respectively, we have

But

Definition ***3.6***

A ***Binomial Experiment*** possesses the following properties:

1. The experiment consists of a fixed number, ***n,*** of identical trials.
2. Each trial results in one of two outcomes: success, ***S***, or failure, ***F.***
3. The probability of success on a single trial is equal to some value ***p*** and remains the same from trial to trail. The probability of a failure is equal to q = (1 – p).
4. The trials are independent.
5. The random variable of interest is ***Y***, the number of success observed during the ***n*** trial.

Definition ***3.7***

A random variable ***Y*** is said to have a ***Binomial Distribution*** based on ***n*** trials with success probability p if and only if

***y=0,1,2,…,n*** and ***0***

Let ***Y*** be a *random variable based on n* trials and success probability ***p. Then***

and

By definitions 3.4 and 3.7,

,

Notice that the first term in the sum is 0 and hence that

The summands in this last expression bear a striking resemblance to binomial probabilities. In fact, if we factor np out of each term in the sum and

Notice that is the binomial probability function based on (n-1) trials. Thus, and it follows that

From theorem 3.6, we know that . Thus, can be calculated if we find . Find directly is difficult because term in the sum is 0 and hence that

Abd the quantity does not appear as a factor of y!. where do we go from here? Notice that

And therefore,

In this case,

The first and second terms of this sum equal zero (when . Then

(Notice the cancellation that led to this last result. The anticipation of this cancellation is what actually motivated the consideration of

Again, the summands in the last expression look very much like binomial probabilities. Factor out of each term in the sum and let to obtain

Again note that is the binomial probability function based on trials. Then (again using the device illustrated in the derivation of the mean) and

Thus,

**And**

Definition 3.8

a random variable ***Y*** is said to have a ***geometric probability distribution*** if and only if

Theorem 3.8

If ***Y*** is a random variable with a geometric distribution,

This series might seem to be difficult to sum directly. Actually, it can be summed easily if we take into account that for

And hence,

(The interchanging of the derivative and the sum here can be justified.) Substituting, we obtain

The latter sum is the geometric series, which is equal to Therefore,

To summarize, our approach is to express a series that cannot be summed directly as the derivative of a series for which the sum van be readily obtained. Once we evaluate the more easily handled series, we differentiate to complete the process

A random variable Y is said to have a **negative binomial** probability distribution if and only if

If Y is a random variable with a negative binomial distribution,

A random variable Y is said to have a **hypergeometric probability** distribution if and only if

If Y is a random variable with a hypergeometric distribution,