5 Research Methodology

5.1 INTRODUCTION

General tendency of a human being is to know things around him. Usually he asks interesting questions such as Why, How, and When, concerning a phenomenon. Such questions arise out of his observation of an event. It is this hunger for knowledge which paved way for searching information in a systematic way to get answer to the questions.

The purpose of such information may be for satisfying the thirst for knowledge or for solving problems or to test the validity of existing / established knowledge. The process of information search has given birth to a new discipline called research.

The word “research” is used to describe a number of activities involving a search information. Therefore, Research is a matter of gathering information from varying sources usually in relation to a specific topic and for a specific purpose.

The core concept underlying research is its methodology which controls the study, dictates the data acquisition by arranging & establishing logical relationships, sets up the ways of refining the raw data, identify an approach so that the meaning can be extracted from data and finally reveals a series of conclusions that leads to an expansion of knowledge.

Research Methodology, as term means, a systematic way to address or solve a problem. When a research is done scientifically, it might be understood as a study of science. While studying any research problem, a researcher adopt several process explaining the methodology for research.

According to J. W. Best (1999), “Research is considered to be formal, systematic, intensive process of carrying on the scientific method of analysis. It involves a more systematic structure of investigation usually resulting in some of formal record of procedures and report of result or conclusions.”
According to P. M. Cook, “Research is an honest, exhaustive, intelligent searching for facts and their meanings or implications with reference to a given problem. It is the process of arriving at dependable solution to problem through planned and systematic collection, analysis and interpretation of data. The best research is that which is reliable, verifiable and exhaustive so that it provides information in which the researcher has confidence.”

Research involves systematic structure. The keyword that distinguishes research from haphazard gathering of observation is systematic. The systematic conduct of research requires particularly two qualities; (1) Orderliness, in which the measurements are accurate and the cross section is fair, and (2) Impartiality in analysis and interpretation.

### 5.2 RESEARCH STATEMENT

The research statement under studied is entitled, “In – depth Study of Factors influencing Customer Behavior in the Purchase of Life Insurance Products”. The present study focuses on the factors influencing consumer behavior in the purchase of Life Insurance products in Gujarat state.

The study has been undertaken to address the following issues;

1. The reasons for selecting either public or private life insurance company.

2. The type of life insurance policy the customer has preferred and their total premium paying amount.

3. The extent of the knowledge of purchased life insurance policy and the main sources of information.

4. The current premium paying option and sources of buying life insurance policy.

5. The factors influencing while purchasing life insurance policy.

6. The extent of satisfaction with respect to the factor which has influenced.
5.3. RESEARCH DESIGN

Research design is an outline which directs the research process. The details of procedures required for achieving preferred information are identified by research design so that the research problems can be addressed. Hence, providing the structure to solve research problems. It also establishes the outline for assortment, measurement and investigation of data.

According to Fred N. Kerlinger, a research design is the plot and construction of examination so apprehended as to obtain answer to research question. The research design is the general scheme or program of the research. Research design comprises summary of hypotheses writing, investigation of data and their operational inferences made by the researcher.

Figure 5.1 Design in Research Process.

The study is Exploratory in nature, followed by Descriptive Study. I. M. Halman (2002) specified that, the foremost objective of exploratory research is to offer comprehensions into, and an understanding of, the problem challenging the researcher. Exploratory study is commonly used to explore the insight of the general research problem. While conducting research, the researcher faces a problem of not knowing anything about the problem.
The purpose of such study is to determine the fairly accurate area where the problem exists and also to identify some attractive courses of action to solve it. It may incidentally also enable a sharper definition of the problem to be made.

The descriptive research is undertaken to describe the business or market characteristics. Descriptive designs have that name because they describe phenomena without establishing association between factors. The data may be; (1) the behavioral variables of people (or other subjects) who are under study and (2) the situational variables that existed or are forthcoming. The descriptive research designs are intended to produce accurate descriptions of variables relevant to the problem being faced, without demonstrating that some relationship exists between variables.

This design serves a variety of research objectives;

- Description of phenomenon or characteristics associated with a subject population (who, what, when, where and how of a topic).
- Estimates of frequency of appearance and the proportion of the population that has these characteristics.
- Discovery of associations among different variables.

It attempts to reveals who should be surveyed (Life Insurance Policy Holders), What (Factors Influencing their Behavior), at what time (at the time of purchasing Life Insurance policy), from where (Gujarat State) and how (method of data collection).

However, descriptive designs are not proficient of addressing any of the why questions accompanying with a given research problem (Hair et al., 2002).

An effective research design ought to have at least following particulars.

- A statement of objectives of the study or the research output.
- A statement of the data inputs required on the basis of which the research problem is to be solved.
- The methods of analysis which shall be used to treat and analyze the data inputs.
5.3.1. Objectives of Study

Main Objective:

- “To Study Factors influencing Consumer Behavior in the Purchase of Life Insurance Products in Gujarat state”.

Sub Objectives:

- To identify the most preferred life insurance product and the most admired company in insurance sector.
- To identify the purpose of investments, attitude and personality with respect to their investment in insurance sector.
- To evaluate the knowledge of the Consumer with respect to the purchased Life insurance products and its sources of information.

5.3.2. Nature of Data

In this study the data are categorized in general terms of their meaning. There are distinct differences among the meaning of facts, knowledge, opinions, intensions and motives.

**Facts** include the measurements of anything that actually exists or has existed. Usually, facts describe tangible things, although they can intangible as long as they can really be determined. Facts are ideal in the sense of possible measurement accuracy. The interpretation placed on real facts, though, may be inaccurate. The researcher must be responsive of the risk that what he is dealing with are ‘quasi-facts’. Many ‘facts’ too are based on estimates or on sample that have unreliability and should not be treated as absolute truth. Here we will mention four of them;

*Demographic:* These are the facts that describe the population to which the data refer and much of our description is economic.

*Sociological:* These are the facts that describe how people are organized in and relate to the society.
Psychological: These are the facts that describe the life – style of an individual pertinent to study.

Behavioral: These are the facts that describe what people do, how they actually behave and these facts are of high importance in this study.

Knowledge means what people know, may be desire data, since that information (be it true or false) may be determinant of what they do. Here consumers’ knowledge or awareness of Life Insurance product is taken into consideration.

Opinion are how people perceive something – what they believe about Life Insurance products and what those beliefs signify. The most potent form of opinions tends to be attitude, which is a mental set or dispositions to act in some manner. Opinions are significant, of course, as they affect behavior, and attitude exerts a general and consistent influence.

Intensions are the acts that people have in mind to do, expectations of their behavior. The extent to which people intend to commit a particular behavior and changes in these intentions may be key information.

Motives are internal forces that cause people to behave as they do. This study would dearly like to have accurate data on the motives that impel consumers’ action toward the purchase of Life Insurance products. Man motives are quite obvious or are subject about which people will speak freely.

5.3.3. Sources of Data

There are numerous possible sources of data, and again we cannot list them in detail. Anyway, this is a determination that, it is special to this study. A step towards this determination is having, first, general classification of sources, which is now offer in several dimension.

Primary as well as secondary sources data are used for the study.

Secondary sources should be first be considered, which refer to those for already gathered and available data. There may be internal sources within the client’s firm.
Externally, these sources may include books or periodicals, published reports, data services and computer data banks.

*Primary* data are obtained from individual who are under study and mainly collected by a researcher to address the research problem. In other words, these data are not readily available from various sources; rather researcher has to systematically collect it for pre – defined research problem.

### 5.3.4. Method of Data Collection

Primary data was collected by adopting survey method in which a questionnaire was used as a tool. In a *survey method*, acquiring the information is based on the questioning the respondents. A “*questionnaire*” is a formal segmentation of structured questions which are framed in order to attain suggestion from the respondents under study. Respondents, who have purchased Life Insurance Products, were questioned with a several questions concerning their behavior, intentions, assertiveness, approaches, awareness, motivations, demographic characteristics. Characteristically, the questions were *structured* which denotes the degree of standardization imposed on procedure to collect the information. Finally, a formal questionnaire was primed where questions were queried with preset instruction and making this process direct collection of data.

Almost all the questions of questionnaire were *fixed – alternative* questions which necessitate the respondents had to choose from a prearranged set of replies. It was also consists of multiple choice, close ended questions having likert scale in which 1 stands for lowest importance to 5 for higher importance.

The survey as a method and especially, using questionnaire as tool for data collection has numerous benefits. First of all, it can be easily managed and administered. And second, the acquired statistics are dependable since the responses are limited to specified alternatives. The inconsistency in the result caused by dissimilarities of questioners can be reduced by with use of limited responses of a questions. Finally, making the process of coding, analysis and interpretation very simple and modest.
5.3.5 Sample Design

Usually, research conclusions which describes the individualities of a large number of items of a phenomenon generally spoken as population, are constructed by investigating the limited number of items from that phenomenon stated as sample. Research usually reaches its conclusions on the basis of sampling, but its methods must adhere to certain scientific rules.¹

Sampling Design Process includes five steps that are shown in sequentially in figure 5.2 and these steps are diligently interconnected and applicable to all most all the phases of the research study.²

**Figure 5.2 Sampling Design Process.**

```
Define the Target Population
---
Determine the Sample Frame
---
Select the Sampling Technique
---
Determine Sample Size
---
Execute Sampling Process
```


**Step 1. Define the Target Population**

The sampling design initiates with specification of the targeted population. The target population comprises of gathering the elements or objects possessing the information which researcher pursues in order to prepare certain interpretations. A sampling design must exactly define the targeted population. It must be ascertained with respect to elements, sampling units, extent and time. An element is an object about which or from which the information is anticipated. Here in this study the element is respondent. An element which is available for selection at the phase of sampling procedure is termed as sampling unit.
- **Population**: Population includes consumers possessing Life Insurance Policies.
- **Element**: Respondents Possessing Life Insurance Policy.
- **Sampling Unit**: An individual consumer possessing Life Insurance Policy who can spare time and provide the needed information through questionnaire.
- **Extent**: Metropolitan cities of Gujarat State.
- **Time**: Year 2011-12.

**Step 2. Determining the Sampling Frame.**

A sampling frame presents the elements in the target population and comprises a list or set of course to recognize the target population. It also serves as a boundary which circumscribe the targeted population. Often, these boundaries are geographical. Here, in this study the sampling frame is the individual who has purchased or possessed life insurance policy within the Gujarat State.

**Step 3. Select Sampling Techniques.**

Principally, sampling techniques are distributed into Probability and non-probability sampling techniques. Probability sampling techniques offers an equal opportunity for each and every element of the population being selected as sample. This method utilizes some form of *random selection*. But non-probability sampling does not involve *random selection*. Briefly this concept has been explained in figure 5.3.

**Figure 5.3 Sampling Techniques Overview.**

**Sampling Technique:** *Non-probability sampling Technique*, where each element does not have equal opportunity to select as sample. Further, *Snow-ball Sampling* method will be used for the determination of data assembly. The process of snow-ball sampling is much like asking subjects to nominate another person with the same trait. Then nominated subjects were observed and continued the same way until sufficient number of respondent obtained. This type of sampling technique works like chain referral.

**Step 4: Determine Sample Size**

Selection of the sample size is considered to be the major part of all statistical analysis. Defining the number of sample units to be carefully chosen for examination is based on various factors. These factors are;

- **Size of Universe:** If the universe is large, the sample size should be large.
- **Nature of Universe:** If the universe is homogenous, a small sample may also provide accurate results. If the universe is heterogeneous, it is advisable to select a large number of sample units.
- **Nature of Respondents:** It is assumed that all the respondents may not respond, a large sample needs to be addressed.
- **Availability Resources:** If the researcher has considerable amount of time and money at his disposal, large sample can be selected.
- **Sampling Method Adopted:** Sampling method selected will also influence the size of sample.
- **Degree of Accuracy Required:** If the researcher expect greater degree of accuracy, it is advisable to go for a bigger sample size. It should be noted that the larger sample size need not mean the greater accuracy.

- **Sample Size:** The total sample size of 500 respondents was taken from Gujarat State.

**Step 5: Execute the Sampling Process**

Implementation of sampling process involves a detail description of the sampling design decisions regarding determination of population, sampling frame, sample size, sampling unit and sampling technique.
Execute: Entire state of Gujarat was divided into zones namely North Gujarat, South Gujarat, Central Gujarat and Shaurashtra and Kutch. Approximately 100 consumers were selected from each zone. The respondents having Life insurance policies were designated to fulfill the objectives of the study. Further, through referrals, other respondent were surveyed until the required number was obtained.

5.3.6 Data Scaling and Measurements

Measurement refers to allocating a number or any symbol for certain characteristics of objects as per definite predefined rules\(^3\). Here, the objects are not measured but some of its characteristics. Thus, respondents are not measured but only the measurement of their perceptions, attitudes, preferences, or any related individualities. In this research, there were mainly two reasons for allocating the numbers to certain individualities. First of all, it provide easiness for statistical analysis of the collected information. And secondly, the communication of measurement rules and result can easily be facilitated through numbering.

The explanation of guidelines in order to convey some numbers to certain characteristics is one of the supreme characteristic of measurement. The process of assigning numbers should be isomorphic. In other words, one-to-one correspondence must exist amongst the characteristics which is being measured and numbers assigned. Moreover, while assigning the numbers to these characteristics, the adopted rule must be standard and uniform. The rule must not change with the change in object or time.

An extension of measurements might be considered as Scaling where a continuum is created upon the location of measured objects. Generally, the scales of measurements are; nominal, ordinal, interval or ratio. The summarized properties of these measurement scales are mentioned in Table 5.1. Here, in this study, the primary scales like Nominal, Ordinal and Interval scales have been used to measure the characteristics of respondents.

Non – Comparative Scale: Under this scaling techniques, scaling of each object is autonomous of the other objects of set. It consists of one of the scale known as itemized scale, which is described in above Table 5.2.
### Table 5.1 Primary Scales of Measurements.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Basic Characteristics</th>
<th>Common Examples</th>
<th>Marketing Examples</th>
<th>Permissible Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>Numbers identify and classify objects.</td>
<td>Social Security Numbers, Numbering of football players.</td>
<td>Brand Numbers, Store types, Sex classification.</td>
<td>Percentage, Mode</td>
</tr>
<tr>
<td>Ordinal</td>
<td>Number indicates the relative positions of the objects but not the magnitude of difference between them.</td>
<td>Quality ranking, Ranking of teams in tournaments.</td>
<td>Preference ranking, Market position, Social class.</td>
<td>Percentile, Median</td>
</tr>
<tr>
<td>Interval</td>
<td>Difference between objects can be compared; zero point is arbitrary.</td>
<td>Temperature (Fahrenheit, centigrade)</td>
<td>Attitudes, Opinions, Index number.</td>
<td>Range, Mean, Standard Deviation.</td>
</tr>
<tr>
<td>Ratio</td>
<td>Zero point is fixed; ratios of scale values can be computed.</td>
<td>Length, Weight.</td>
<td>Age, Income, Costs, Sales, Market shares.</td>
<td>Geometric mean, Harmonic mean.</td>
</tr>
</tbody>
</table>


### Table 5.2 Basic Non-comparative Scales.

<table>
<thead>
<tr>
<th>Itemized Rating Scale</th>
<th>Basic Characteristics</th>
<th>Examples</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likert Scale</td>
<td>Degree of Agreement on a 1 (strongly disagree) to 5 (strongly agree) scale</td>
<td>Measurement of Attitude</td>
<td>Easy to conduct, Administer and understand</td>
<td>Time consuming</td>
</tr>
<tr>
<td>Semantic Differential</td>
<td>Seven-point scale with bipolar labels</td>
<td>Brand, Product and Company image</td>
<td>Versatile</td>
<td>Controversy as to whether the data are interval</td>
</tr>
<tr>
<td>Stepel Scale</td>
<td>Unipolare ten point scale, -5 to +5, without a neutral point (zero)</td>
<td>Measurement of attitude and image</td>
<td>Easy to construct</td>
<td>Confusing and Difficult to apply</td>
</tr>
</tbody>
</table>

As per Itemized Scale Rating, a scale having a specific number or brief explanation linked to respective category were distributed amongst the respondents and these categories were well-organized with respect to position of scale. Finally, respondents were has to choose the specified category pronouncing the best of the object which under assessment. In this study, an itemized rating scales such as Likert Scale has been used to measure the extent of various factors which has influences their behavior in purchase of Life Insurance Products. Titled after its developer ‘Rensis Likert’, it is one of the most common which is being used by researchers.

Under this scale the respondents are required to designate the extent of agreement or disagreement with each of the declarations regarding Life insurance policy. Characteristically, respective scale item is having five response types, oscillating between “strongly disagree” and “strongly agree”.

### 5.3.7 Scale Evaluation

According to Walter C. Borman, correctness and applicability of a multi – item scale must be assessed. In other word, this constitutes an assessment of scale with respect to its reliability, validity and Generalizability, as mentioned in Figure 5.4. Assessment of reliability of the scale comprises the approaches of Test – retest, Alternative – forms and Internal Consistency. Whereas an assessment of validity is performed through investigating Content, Criterion and Construct Validity.

---

**Figure 5.4 Evaluation of a Multi-item Scale.**

![Diagram of Scale Evaluation]


---

13
The most fundamental to evaluation of scale is *Measurement Accuracy*. A measurement is nothing but a number which reveals about characteristics of an item or an object which is under study. A number which is used under measurement do not reflects the characteristic’s value, but rather reflects an observation of that characteristic. In this measurement, there are several factors which leads to *measurement error*[^4], ultimately causing difference between the measurement or observed score and the true score of the measured characteristic. The *true score model*[^5] which conveys a background to understand the correctness of measurement can be stated as;

\[ X_O = X_T + X_S + X_R \]

Where,
- \( X_O \), the observed score or measurement.
- \( X_T \), the true score of the characteristic being measured.
- \( X_S \), Systematic error.
- \( X_R \), Random error.

The entire measurement error encompasses both the systematic error (\( X_S \)) as well as Random error (\( X_R \)). The *Systematic error* explains the constant measurement error representing the stable factors which touch the observed score in the similar manner at each time of measurement. Whereas in case of Random error, it is not consistence and arising out of random fluctuations or alterations in respondents or measurement circumstances.

### 5.3.7.1 Reliability

According to *Pratibhusan Sinha*, “*Reliability denotes to the degree to which a scale produces consistent results if repeated measurements are made on the characteristics*”. Reliability do not get influenced by Systematic sources of error. As these errors influence the measurement in a constant manner and hence, not leading to inconsistency. On other hand, random error leads to inconsistency and ultimately lower reliability. The magnitude to which measures are free of random errors can be referred as Reliability, \( X_R \). The measures are faultlessly reliable, if \( X_R = 0 \).

Here, in this study, the reliability was assessed by using *Internal Consistency Reliability* method with popular approach of *Cronbach’s Alpha* or *Coefficient Alpha*. In order to evaluate the reliability of the scale, *Internal Consistency Reliability* was adopted. Here, each item consistently measures certain characteristic of the construct.
which is measured by means of the complete scale and designate the specific characteristics. The internal consistency of the items forming the scale was focused during this measure of reliability.

**Cronbach’s Alpha or Coefficient Alpha:** It is an internal consistency reliability measurement, where the average of all possible split – half coefficients arising out of dissimilar splitting of scale items. The value of this coefficient fluctuates between 0 and 1. If the value is 0.6 or below normally specifies that internal consistency reliability is unsatisfactory. A value of coefficient alpha upsurges as number of scale items increases and this tends to be a vital characteristic. For that reason, coefficient alpha might be artificially exaggerated by adding a number of no longer needed scale items (Tom J. Brown et al., 2002).

### 5.3.7.2 Validity

A scale validity is defined as the magnitude to which difference in scale of observed scores make a replica of right variances amongst the characteristics of objects which under measurement. In order to have desirable validity, there must be no errors in measurements, i.e. \( X_O = X_T, X_S = 0, X_R = 0 \). Researcher might take content validity, criterion validity or construct validity into consideration (Gilad Chen, 2001).

**Content Validity**

It sometimes also known as face validity consisting an independent but organized assessment which represents the content of the scale in order to measure the task on hand. The study scrutinizes whether the entire sphere of the constructs, which are being measured, are satisfactorily roofed by the scale items or not. As content validity is subjective in nature, single-handedly, it do not reflects an adequate measure of the scale validity. A supplementary recognized assessment might be acquired by investigating criterion validity.

**Criterion Validity**

This type of validity inspects whether the measured scale achieves as predicted with respect to additional variables nominated as expressive criteria or not. Criterion variable take into account the demographics and psychographic, attitudinal and
behavioral characteristics. A criterion validity can be of two kinds, concurrent and predictive validity. *Concurrent validity* is evaluated once collection of the data regarding the scale under evaluation and the data regarding the criterion variables are at similar time. *Predictive validity* is assessed by collecting data with respect to the scale under evaluation and data on the criterion variable are not in the same period of time.

**Construct Validity**

This type of validity focuses on problems as what characteristic is being measured by the scale. An effort is also made in order to respond the theoretic questions such as; why a scale performs and what deductions are to be prepared in relation to scale theory. Construct validity consist of convergent, discriminant and nomological validity. *Convergent validity* is the extent of positive correlation of the scale with other measurements of the similar construct. Discriminant validity is the extent of difference in correlation of scale with other measurements of the similar construct. Nomological validity is the degree of theoretically predicted correlation of scale with measures of dissimilar but correlated constructs.

### 5.4 DATA PREPARATION

#### 5.4.1 Tabulation and Classification of Data

The data collected through questionnaire were *edited* by reviewing the questionnaire with an objective to increase the accurateness and exactness. This step contains screening of questionnaire in order to recognize unintelligible, inadequate, inconsistence or indefinite responses. Then after, the data were coded through assigning a code to signify a specific response to a precise question.

At last, the data were tabulated and classified on the basis of Age, Education, Occupation, Household Income and other responses given by the respondents. The cross – tabulation has been done according to different variables. *Cross – tabulation* is a statistical procedure which simultaneously defines two or more variables. Consequently, a table reflecting the dual distribution of two or more variables having limited number of distinct values.⁶
5.4.2 Framework of Data Analysis

Statistical Package for Social Science (SPSS 17 Version) has been used to investigate the data. SPSS is the one of the most commonly statistical software package is being used for data analysis. It offers a broad range of statistical procedure to summarize data and determines whether the differences between the groups are statistically significant or not. SPSS software also includes several tools for studying data containing functions for recording data and computing new variables as well as merging files. Reliability and validity tests were conducted by using SPSS software. The statistical procedures like; Cross –tabulation, Factor Analysis, ANOVA Test and other necessary analysis were performed with SPSS.

5.5 STATISTICAL METHODS OF DATA ANALYSIS

5.5.1 Chi – Square Test

The \textbf{Chi – Square Statistic} ($\chi^2$) is used to check the statistical significance of the observed relationship in a cross – tabulation. It supports us in determine existence of systematic association between the variables. Karl Pearson in 1900 developed a non – parametric test for testing the significance of the discrepancy between experimental i.e. observed frequencies and the theoretical frequencies i.e. expected obtained under some theory or hypothesis. This test is also known as Test of Goodness of fit and it is used to test whether the discrepancy between expected and observed values may be attributed the chance (fluctuation of sampling) or whether the deviation is really because of the inadequacy of the theory to fit the observed data.

The null hypothesis, $H_0$, is that there is no association between the variables. The test is accompanied by computing the cell frequencies that would be expected if no association were present between the variables, given the existing row and column totals. For the calculation of Chi-square statistics, the expected cell frequencies ($f_e$) are matched with actual observed frequencies ($f_o$) established under cross – tabulation. Higher values of the statistics are revealed by larger differences between these expected and actual frequencies of cross – tabulation.
Undertaking the cross – tabulation having $r$ rows and $c$ columns and a random samples of $n$ observations, then the expected frequency for each cell can be calculated by using a simple formula;

$$f_e = \frac{n_r n_c}{n}$$

Where, $n_r = \text{total number in the row}$  
$n_c = \text{total number in column}$  
$n = \text{total sample size}$

Then, the value of $\chi^2$ is calculated as follows;

$$\chi^2 = \sum_{\text{all cells}} \frac{(f_o - f_e)^2}{f_e}$$

Where, $f_o = \text{frequency of observed values}$  
$f_e = \text{frequency of expected values}$

There is a different distribution for $\chi^2$ for each number of degree of freedom ($d.f.$), defined as;

$$d.f. = (k - 1) \quad \text{where, } k = \text{the number of categories.}$$

With Chi – square contingency tables of the two – samples variety, we have both rows and column in the cross – tabulation.

In that instance, $d.f.$ is defined as rows minus 1 $(r - 1)$ times column minus 1 $(c - 1)$.

$$d.f. = (r - 1) (c - 1)$$

**Conditions Characterizing Chi-square test**

The chi-square test can be validly applied if the following conditions are satisfied;

- The observations recorded are collected on random basis.
- The sample observation should be independent, i.e. no individual item should be included twice or more in the samples.
- The total number of observations should be rationally large, say $N > 50$. 
The data should be expressed in original units for convenience of comparison and
the given distribution should never be replaced by relative frequencies.

Small theoretical frequencies should be avoided while calculating Chi-square test.
Small is a relative term. Preferably, each theoretical frequency should be larger
than 10, but in any case not less than 5. Since, Chi-square distribution is a
continuous distribution, it cannot maintain its characteristics of continuity, if cell
frequency is below less than 5. In that case, we adopt pooling techniques, which
consists of adding the frequencies which are less than 5 with preceding or
subsequent frequencies to enable the resulting sum to exceed 5 and adjust
accordingly for the degree of freedom is adopted.

5.5.2 ANOVA (Analysis of Variance)

I. A. Fisher founded statistical methodology, while analyzing design of experiment,
called Analysis of variance abbreviate as ANOVA. The most universally used
procedures for testing the significance of the difference among several means
simultaneously are collectively called Analysis of Variance (ANOVA). The equality
tested is usually called the partition sum of squares, for the ANOVA test calls for the
experimentally different samples to be partitioned into two distinct parts. And the
values termed sum of squared deviation are calculated rather than the simple variance
themselves.

The meaning of this partition of the sum of squared deviation (also called Sum of
Squared) into two parts is the essence of the ANOVA test. An individual’s score in
any sample can differ from another’s. These observed differences can be attributed to
two sources. First, some individuals are in different treatment group, and their
differences could be due to different treatments.

The measure termed sum of squared between groups reflects the contribution of the
different treatment to intergroup differences. Second, individuals in the same
treatment groups can differ because of chance variation or individual differences, for
each individual within the group receives exactly the same experimental treatment.
The sum of squared within groups reflects the intragroup difference.
The above explanation suggests that the measure of the difference *between* the different groups relative to the differences that exist *within* each group would determine the equality or inequality of the means. If the null hypothesis were true, the study expects the variance among the sampling groups to be equal, except for possible sampling error. Whereas, if the null hypothesis were not true, then study expects the variance *between* the groups to be higher than the variance *within* the groups.

The relationship of the variance between the groups and the variance within the groups is expressed as an F ratio, computed by dividing the treatment mean square (i.e. the sum of squares between – group divided by its degree of freedom) by the within group mean square (i.e. the sum of squares within – group divided by its degree of freedom).

If the computed F ratio is greater than the theoretical F ratio from the F table, the null hypothesis is rejected, which means that the difference among the means are significant. The null hypothesis is accepted, if the computed value of F ratio is below the critical F value, resulting into the difference among the means is not significant.

Here in this study ANOVA is carried out Age wise, Income wise, Education wise and Occupation wise to that whether there is difference in perceptions towards factors influencing in the purchase of Life Insurance Products or not. If one test variable is used then it is called one way analysis of variance and if two test variables are used then it is called two way analysis of variance.

### 5.5.3 Factor Analysis

*Factor Analysis* is a common name symbolizing a course of techniques mainly used to reduce the data for summarization. A research may comprises of variables which are large in number (in this study 29 variables) and utmost are interrelated. These variables must be reduced so as to explain relationship of association. This relationship amongst many interconnected variables are scrutinized and represented in a few fundamental factors.

Here in this study, factors influencing consumers’ behavior was measured by asking respondents to evaluate series of items on a semantic differential scale. These items
evaluations were analyzed to determine the dimension underlying the factors influencing consumer behavior while purchasing life insurance products.

Factor analysis is considered as an *interdependent technique* where entire set of interdependent relationship is scrutinized. Unlike in analysis of variance where one variable is reflected as the dependent or criterion and the others as independent or predictor. However, such differences are not made under factor analysis.

**Interdependent Technique**: Multivariate statistical techniques in which the whole set of interdependent relationship is examined.

**Factors**: A principal dimension which explains the correlations between variables set.

**Principal Components Analysis** is one of the method used for factor analysis where total variance in the data is taken into consideration. In correlation matrix, the diagonal is unities and complete variance is carried into the factor matrix. This is method is recommended when the principal apprehension is to define minimum number of factors which are accounting maximum variance amongst the data. Ultimately, these extracted factors are known as *principal components*.

5.5.3.1 Statistics Linked with Factor Analysis

**Barlett's Test of Sphericity**: This test statistic examines the hypothesis that in the population the variables are uncorrelated. Simply, the population correlation matrix is an identity matrix where each variable has perfect correlation with itself ($r = 1$) and there do not correlate with other variables ($r = 0$).

**Correlation Matrix**: It is a lower triangle matrix indicating the simple correlations ($r$) amongst all promising pairs of variables in the analysis and the diagonal elements which are generally all 1 are commonly excluded.

**Communality**: Communality explains the extent of variance a variable shares with all other variables. It also explains the percentage of variance clarified by the common factors.

**Eigenvalue**: The total variance explained by each factor is represented by this value.

**Factor Loading**: These are modest correlations amongst the variables and factors.
**Factor Loading Plot:** It is a plot of the original variables using the factor loading as coordinates.

**Factor Matrix:** It encompasses the factor loadings of every variables on all the extracted factors.

**Kaiser-Meyer-Olkin (KMO) Measure of Sample Adequacy:** It is a measure of sample adequacy index which helps to determine the suitability for factor analysis. If the data are appropriate for factor analysis, its value ranges from 0.5 and 1.

**Percentage of Variance:** This is the fraction of total variance accredited to each factor.

**Residuals:** Residuals are the variances explained amongst the observed correlations which are given in the input correlation matrix and the reproduced correlations which are predictable from the factor matrix.

**Scree Plot:** It is a graphical presentation of the eigenvalues against the number of extracted factors in order.

### 5.4. SCOPE OF THE STUDY

The present study has made an effort to study the factor influencing consumers’ behavior in the purchase of life insurance products in Gujarat State. It involves understanding the basic concept of Life Insurance products, Development of Insurance industry in India with key legislations and its regulatory body, detail outline of insurance industry and Indian economy and the factor influencing consumers’ behavior in the purchase of life insurance products.

The study is conducted to track consumers’ behavior towards life insurance products in Gujarat and analyze the factors which are influencing in the purchase of life insurance products. The scope is limited to the extent of geographical region, time and the information collected during this study. The scope of this study is limited to the responses given by respondent during this study. The scope of this study is limited to the responses supplied by the 500 respondents only, which may not be a significant amount to represent the whole population. The results drawn from the responses of
500 respondents may or may not be biased. The study is confined to life insurance products only.

Similar studies on this line may be conducted in other cities also and for different insurance products in India. Further research can also be conducted for studying consumers’ perception with respect to public and private insurance players and also with respect to other insurance products of non – life like general insurance and health insurance.

5.5 LIMITATIONS OF THE STUDY

For the research work, data was collected and interpreted with utmost reliability and consistency but due to predisposition of a few respondents, certain limitations of the study are as follows:

- The study represents the present scenario in the selected cities of Gujarat and hence the result may not be applicable to another geographical area.
- Responses were collected during the year 2011 – 12, the result may vary from another period of time.
- The study is limited to 500 respondents of the selected cities of Gujarat.
- The data is collected from the respondent possessing life insurance policies only.
- Answer to the questionnaire depends upon the belief and prejudices of customers possessing life insurance policies.
- It is anticipated that respondents are true and honest in expressing their opinion and have filled the questionnaire without any bias.
- The present study is constrained to the information collected about the life insurance products and factors influencing in the purchase of life insurance products with the help of questionnaire.
Endnotes


