THE PLACE OF REGRESSION IN HYPOTHESES TESTING AND DECISION MAKING

Waribugo Sylva
Faculty of Management Sciences, University of Port Harcourt, Nigeria

Selina Dan-Albert
Garden City Premier Business School, Port Harcourt, Nigeria

ABSTRACT

On a daily basis, managers are faced with problems that require the selection of alternative choices among several sets of possible solutions. In order to achieve this, decision makers are required to develop the most appropriate models that will give optimal decision within some parameters and assumptions. This study investigates the place of regression as Hypothesis Testing tool for decision making. The paper extensively reviews the concept of decision making, hypothesis testing and the applications of regression analysis as a hypothesis testing tool for managerial decision. Furthermore, traditional manual method and spread sheet modelling were deployed to illustrate the process of hypothesis testing in an organisational setting.

KEYWORDS: Decision making, Regression Analysis, Hypothesis Testing.

1.0: INTRODUCTION

The great task before organisations is to make decisions in order to create value for stakeholders (Lovallo, Koller, Uhlaner & Kahneman, 2020). Decision making is one of the key roles performed by managers (Mintzberg, 1980; Robins, & Coulter, 2012). It is the process of selecting a solution from alternative courses of action (Williams & Ward, 2007) in order to derive optimal results. A good decision, when made, promotes the bottom-line and leads to competitive advantage (Clemen, 1996). Moreover, decision-making is a critical component of leadership success and strategic planning in organisations (Sandras & Nelson, 1996; Akdere, 2011), which cuts across diverse processes such as negotiations, demand forecasting, new product lunch, risk analysis and facility location, among others (Clemen, 1996).

Furthermore, decisions are made in a bid to solve problems or correct errors (Anderson, Sweeney, Williams, Camm & Martin, 2012). According to Render, Stair and Hanna (2012), “a good decision is one that is based on logic, considers all available data and possible alternatives, and applies the quantitative approach” (p.70).

One of such quantitative approaches is regression analysis – an approach that estimates and predicts the relationship between two or more variable(s). Regression analysis is useful in cost estimation, demand forecast, advertising expenditure versus sales analysis, amongst other countless subjects. Moreover, experts commonly agree that the utility of regression analysis is expressed by hypothesis test for significance (Render, Stair & Hanna, 2012).

According to Groebner, Shannon, Fry and Smith (2011), hypothesis testing is a rational and scientific method used by managers for making decisions in many organisations. It involves the use of sample data to estimate population parameters (Gravetter & Wallnau, 2017). Specifically, the null hypothesis test allows managers make decisions, while controlling for decision errors. Whereas, the hypothesis testing is ineffective in times of environmental uncertainty, its application enhances the ability of decision makers to identify and mitigate unexplained external factors (Groebner, Shannon, Fry & Smith, 2011).

A core problem in organisational studies concerns the development of valid models for decision making. Regression models are suitable tools for solving decision problem in several repetitive scenarios. Despite the popularity of regression analysis and hypothesis testing in decision-making, not much has been practically demonstrated in details by simultaneously using manual and spreadsheet models in referred journals. This paper seeks to fill this gap in literature.

The rest of this article comprehensively discusses the concepts of decision making, hypothesis testing, regression analysis, and their strengths, limitations as well as their interdependence. The final part of paper deploys manual and spreadsheet modelling to demonstrate hypothesis testing in an organisational setting, using the parametric regression tool.

2.0: LITERATURE REVIEW

2.1: Decision Making

The word “decision” is derived from Latin decisio, which stands for provision, settlement, and resolution. Jones (2013)
views decision making as the “process of responding to a problem by searching for and selecting a solution or course of action that will create the most value for organizational stakeholders” (p. 334). Decision is made at all levels and functional areas of management. That is, managers make choices at top, middle and lower levels, as well as in finance, production, marketing, and personnel departments (Robins & Coulter, 2012).

According to Verma (2014), decision making facilitates implementation of managerial functions and the evaluation of managerial performance. The effectiveness, health and survival of a firm reflect the robustness of decisions that are made. Also, decision-making could involve the participation of stakeholders, which leads to higher levels of creativity and innovation (Rossiter & Lilien, 1994). Furthermore, collective decision making promotes organizational agility and fosters belongingness among organizational members (Levesque & Walker, 2007).

2.1.1: Decision Making Perspectives

Decision making comprises three perspectives which include rationality, bounded rationality and intuition. Classical rational decision making is an eight-step process that involves identification of a problem, and decision criteria, allocation of weights to the Criteria, development and analysis of alternatives, making a choice among the alternatives, and implementing the choice. In addition, decision makers evaluate the effectiveness of the decision (Robins & Coulter, 2012) through feedback.

The neoclassical paradigm of rational decision making (Savage, 1954) assumes that the decision maker has a sufficient and stable utility function, knows all the available decision variables, has perfect computational acuity to create maximum utility. Essentially, the rational model is an analytic, logical, step-by-step, iterative, and clearly specified methodology for decision making.

Calabretta, Gemser and Wijnberg (2017) itemize the characteristics of rational decision making as (i) availability of relevant information (ii) formal and systematic analysis of the available information (iii) comprehensive information (iv) step-by-step process (v) logical approach where choices are based on rules and causality (vi) the presence of cognitive capacity intentionality, whereby the decision maker intentionally commits time and cognitive capacity to making the choice.

In real business situations, decisions are made under some limitations. Simon (1954) submits that decision making is bounded by the cognitive ability to process information, skill and competencies of the individual, the availability, cost and overload of information, and the complexity of the environment.

Also, Wall (1993) interrogated the rational model, claiming that the agent’s information processing is parsimonious, solutions are simple-minded, search is local and sequential, and search for a new and better solution is undertaken only when it is deemed necessary. Moreover, individuals’ decision making acuity is also limited by their beliefs, values, idiosyncrasies, perception, and organisational experiences (McKelvie, Haynie & Gustavsson, 2011). To this end, Schiliro (2018) defines bounded rationality as “the limitations and difficulties of the decision maker to behave in the way the traditional rational choice theory assumes, due to his insufficient cognitive and computational capacities to process all the relevant information” (p. 64).

Under conditions of bounded rationality, agents adopt a less ambitious decision model that is “good enough” instead of achieving the desired maximum — a term known as “satisficing” behaviour (Cyert & March, 1962; Jones, 1999). For a satisficing individual, “as soon as he discovered an alternative for choice that meets his level of aspiration, he would terminate the search and choose that alternative” (Simon, 1979, p. 503). If the choice falls below the agent’s aspiration, a search for another bundle of alternatives is repeated until a new, satisfactory outcome is achieved.

Although the rational model is the popularly adopted paradigm, some agents deploy a combination of rationality and intuition to optimize value (Elbanna & Child, 2007). Intuition (also called hunch or gut-feeling) is a decision-making approach in which the agent rapidly and non-consciously recognizes patterns and associations and deploy experiences, imagination, feelings and creative intelligence to arrive at value-creating judgments. It is that inner experience of knowing a thing without knowing any reasons why it is, and a “quick appraisal[s] based on integrating information in a sketchy way” (Segalowitz, 2007). Intuition has been nuanced as expertise (Burke & Miller, 1999), heuristics (Gigerenzer, 2007), deep sub-conscious learning and memory (Lowenstein, 2000), as well as non-scientific perceptual information evaluation (Volz & von Cramon, 2006).

Moreover, individuals deploy intuitive process rather than rational decision making because the systematic and mechanical nature of the latter makes it slow, time-consuming, and tiresome, thereby limiting the ability of the agent in coping with time pressure in acomplex, dynamic and uncertain business environment. Furthermore, intuition not only assists individuals to navigate uncertainty but also stimulates creative cognitions necessary for problem solving (Hodgkinson et al., 2009). As the experience and expertise of the decision maker increases in a specific area of specialization, the degree of certitude in intuitive judgement becomes less random (Gore & Sadler-Smith, 2011).

Some scholars (e.g. Kahneman, Slovic, & Tversky, 1982) are of the view that intuition is the primary window through which an alternative is chosen, and rational process is merely a way of evaluating the outcome of intuitive decision. However, it is commonly held in management theory and practice that the rational model remains the most scientific and desirable, even if contingent variables jeopardize its efficacy (Cabantous & Gond, 2011).

2.1.2: Types of Decisions

Managers in various organisations face various problems and make various types of decisions to solve the problems. Simon (1977) identified programmed and non-programmed decisions as the two classes of decisions. Depending on the nature of the problem or task, a manager can deploy any of them.

Programmed Decisions

Programmed decisions are those carried out under structured, straightforward, familiar or clear problem situations, whereby the information about the problem is easily defined and complete. They are tied to the procedures, rules and policies of the organisation. According to Simon (1977), “decisions are programmed to the extent that they are repetitive and routine, to the extent that a definite procedure has been worked out for handling them so that they don’t have to be treated from scratch each time they occur” (p. 46). The daily, standard operating procedures of most
organisations predominantly require programmed decisions, whereby routine rules are observed in the management of inventories, job scheduling, allocation of machines, cost estimation, salary determination, pricing decision, etc. In most cases, lower-level managers generally make programmed decisions.

**Non-programmed Decisions**

There are some problems that a manager cannot solve using programmed decisions. In fact, most organisational scenarios are inundated with unstructured problems that have novel, complex, ill-defined and insufficient information. Problems which are new and laden with ambiguity are classified as unstructured problems (Ivancevich & Matteson, 1990). According to Simon (1977), decisions are non-programmed “to the extent that they are novel, unstructured and unusually consequential” (p. 46). Non-programmed decisions deploy unique, broad, innovative and intuitive problem solving approach that do not have standard or pre-specified procedure. Thus, non-programmed decisions call for application of the agent’s individual opinion and conceptual skills to solve unstructured problems (Kreitner & Kinicki, 1992).

Instances of non-programmed decisions are: the acquisition of a company, decision to make blue-chip investment or build a new manufacturing plant, divestment decision, going global, product diversification, and introducing a new business model to have strategic fit with the changing business context. Since non-programmed decisions require the use of conceptual skill, they are mostly carried out by top managers.

**2.1.3: Decision-Making Conditions**

When making decisions, managers may face different conditions. Thus, decision making is also classified according to the situation in which a decision is made. Literature suggests that agents make decisions under conditions of certainty, risk and uncertainty (Zimmewicz, 2008).

**Decision Making Under Certainty**

Certainty is the perfect condition under which decisions are made. Under certainty, agents have well-defined information about the decision criteria and can easily predict the outcome (Zimmewicz, 2008). Although, it is commonly agreed that most business decisions are not made under conditions of certainty (Alvarez & Barney, 2005), managers can take some steps to increase the certainty of the outcomes while making decisions. Such steps include collection of market data, analysis of competitors’ strengths and weaknesses, and the application of functional strategies and effective forecasting models. Even at that, the decision outcomes in the dynamic business environment will scarcely be certain as desired. This lack of certainty is presented in economic literature as risk and uncertainty (e.g., Shane, 2003).

**Decision Making Under Risk and Uncertainty**

In a strict sense, risk is the probability of having a disadvantageous or less desirable outcome. Broadly, risk is the variance (or standard deviation) reflected in potential loss or gain of a decision outcome (Park & Shapira, 2017). Under risky situation, the agent is able to estimate the probability of certain outcomes. This is achieved by assigning probability figures to various states of nature, based on a company’s past data, secondary sources or experience. A state of risk occurs when the probability of desired outcome is greater than 0 but less than 1, subject to the vicissitudes of the business environment (Wald, 1950).

Most economic, financial and statistical models involve the evaluation of risk in business decision (Brealey & Myers, 1988). A popular model is the expected value estimation (Raiffa, 1968) which entails the sum-product of the value of the outcomes and their corresponding probabilities of occurrence. The decision will then fall on the highest sum-product of the Expected Monetary Value (Ragsdale, 2012). Another model of decision making under risk is the Capital Asset Pricing Model which uses the β value as a measure of the systematic risk of an investment.

On the contrary, decision making under uncertainty takes place when the possible outcomes of the decision as well as the probability of occurrence are not known at the time the decision is made (Knight, 1921). Under such condition, a set of explanatory or target variables are unidentifiable, unquantifiable and fuzzy in terms of occurrence and evolution, thereby having no known probability estimates.

According to Ragsdale (2012), non-probabilistic models are deployed to solve problems under uncertainty via maximax, maximin and minimax (regret or opportunity loss) modes. Maximax is an optimistic decision making process under uncertainty whereby the manager considers the maximum possible payoff; in maximin, the decision maker pessimistically assumes that state of nature will only enable the organisation to maximize the minimum possible payoff regardless of the quality of decision; while a manager who desires to minimize his maximum “regret” will opt for a minimax or opportunity loss choice (Ragsdale, 2012). In minimax regret decision rule, the payoff matrix is first converted into “a regret matrix that summarizes the possible opportunity losses that could result from each decision alternative under each state of nature” (Ragsdale, 2012, p. 693).

In sum, decision making under risk and uncertainty differ in the sense that risk is the situation under which the outcomes and probabilities of occurrences are known to the agent, whereas uncertainty connotes lack of information on the outcomes and probabilities of occurrences (Knight, 1921). Nevertheless, decision-maker’s perception of risk and uncertainty is a function of the decision context and the psychological characteristics of the agent (Park & Shapira, 2017).

**2.1.4: Cognitive Styles, Biases and Risk Attitude in Decision Making**

Two situations that affect decision making process and outcomes are the agent’s cognitive styles (which have biases) and attitude towards risk (Phillips-Wren, Power & Mora, 2019).

According to Leonard, Scholl and Kowalski (1999), cognitive styles are “the way in which people process and organize information and arrive at judgments or conclusions based on their observations” (p. 407). Managers adopt either a linear cognitive style or a nonlinear cognitive style, or both. For Robbins and Coulter (2012), a linear cognitive style is characterized by a manger’s “preference for using external data and facts and processing this information through rational, logical thinking to guide decisions and actions”, while a nonlinear cognitive style is “characterized by a preference for internal sources of information (feelings and intuition) and processing this information with internal insights, feelings, and hunches to guide decisions and actions” (p.190). Moreover, nonlinear cognitive style falls under heuristics, a class of shortcut cognitive techniques that enhance problem solving. 
solving and simplify decision making amid uncertainty.

Although heuristic cognitive styles sometimes improve the ease and efficiency of decision making, there are times they attract biases (Dvorsky, 2013). When cognitive biases creep in, suboptimal decisions become inevitable. There are empirical studies which show that agents often make biased decisions, especially in uncertain contexts (Tversky & Kahneman, 1974). Some scholars have identified various cognitive biases exhibited by decision makers, which yield systematic errors. These biases include: overconfidence, illusion of control, immediate gratification, anchoring effect (Skład & Diekstra, 2013), selective perception, confirmation (McKenzie et al, 2011), framing (Entman, 2007), availability (Tversky & Kahneman, 1973; Chatfield, 2016), representation, randomness, sunk costs (Arkes & Blumer, 1985; Staw, 1976; Heath 1995), self-serving (Heider, 1958; Campbell & Sedikides, 1999), hindsight (Fischhoff, 1975; Mazursky & Ofir, 1990), false consensus (Mullen et al., 1985; Engelmann & Strobel, 2000; Hammond et al, 2006) and status quo (Samuelson & Zeckenhauser, 1988; Henman, 2008; Nicolich, 2018).

Aside cognitive bias, the decision maker’s attitudes toward risk - expressed as risk-averse, risk-neutral and risk-seeking (or risk-prone)- also affect decision choice (Schiebener & Brand, 2015). A risk-averse agent is a conservative decision maker that is comfortable with small positive (or marginal positive) returns on investments. They are bent on avoiding losses, however infinitesimal, because of the grave agony such outcomes could inflict. Usually, risk-avoiding managers place more emphasis on the possible economic losses that could emanate from their decisions than on the potential equivalent benefits (Lovallo, Koller, Uhlner & Kahneman, 2020). However, they take risk only when they can estimate that the reward far exceeds the aversion of the risk.

Risk neutral decision makers focus on the value of the investment and are ambivalent to risk. Investments in assets only appeal to them when the value of the asset increases linearly based upon its net present value (NPV). Thus, risk-neutral decision makers are objective and deploy quantitative models such as Expected Monetary Value (EVM) and the Decision Tree approach to maximize the utility of their choice (Perloff, 2014).

Risk-seeking decision makers (such as gamblers) attach less value to their assets, and feel less agony for any loss or decline in assets. They cope with losses by readily giving justification for losses in their asset investments, believing that it is all part of life (Swalm, 1966).

On the whole, the cognitive styles and risk attitude of managers shape decision making. However, the performance outcomes from the choices are influenced by external variables such as activities of competitors, regulators, and consumers. Moreover, other imponderables like natural disasters, price volatility, and the state of the economic also affect the decision outcomes.

2.2: Hypothesis Testing in Decision Making

According to Gravetter and Wallnau, (2017), hypothesis testing is an inferential “statistical method that uses sample data to evaluate a hypothesis about a population” (p. 225). Specifically, the Null hypothesis statistical testing (Fisher, 1925, 1935, 1956,1973; Neyman & Pearson, 1933) is arguably the most commonly deployed method of analysing data of organisational phenomena (Nickerson, 2000). Groebner, Shannon, Fry and Smith (2011) submit that “statistical hypothesis testing provides managers with a structured analytical method for making decisions” (p.347) in diverse sectors such as the pharmaceutical industry where hypothesis tests are performed to determine the efficacy and safety of new drugs before they are administered. Moreover, hypothesis testing is deployed as a decision making tool among judges when weighing evidence in jurisprudence.

Again, the null hypothesis test is a method that is used to ascertain the meaningfulness or statistical significance of regression models (Render, Stair & Hanna, 2012). This means that it lets managers make decisions in such a way that the probability of decision errors can be controlled, or at least measured (Groebner, Shannon, Fry & Smith, 2011). According to Lind, Marchaland Wathen (2012, p.483), hypothesis testing enables the decision analyst to evaluate if the slope of the regression line in the population is different from zero, so that conclusion can be drawn regarding the predictive strength of a regression model. Although the null hypothesis test of significance does not eliminate environmental uncertainty, the methodology often empowers decision makers to identify and mitigate unexplained factors in the dynamic environment (Groebner, Shannon, Fry, & Smith, 2011).

2.2.1: Procedure for Testing Hypothesis

Hypothesis testing comprises seven steps as follows:

Step1: Specify the population parameter of interest

A parameter (θ) is a numerical value that describes a property of a population. Examples of population parameters are: population mean (μ) and population standard deviation (σ). In management sciences, the corresponding sample statistics (e.g., X, S) are used as substitutes for population parameters, since it is more difficult to observe entire populations.

Step2: Formulate the null hypothesis (H₀) and the alternate hypothesis (H₁) in terms of the population parameter

The Null Hypothesis (H₀) is a declarative statement which specifies no-difference, association or effect between two variables (Nickerson, 2000). It is assumed to be true unless there is contrary statistical evidence. Once the sample data negates the nil-null hypothesis, its alternative (H₁) is accepted. Alternative Hypothesis is hypothesis that includes all population values not included in the null hypothesis.

Step 3: Specify the desired significance level (α)

The researcher states the level of significance (α), which is the probability of rejecting the null hypothesis when it is true. Sometimes called level of risk or probability of having Type I error, the alpha is arbitrarily set a priori, usually at 0.05. A researcher commits Type I (false positive) error when a true null hypothesis is rejected. Instead of accepting the null hypothesis, the researcher erroneously accepts the alternative hypothesis (Neyman & Pearson, 1933). Simply put, it is the act of concluding that there is a significant difference, whereas there is no significant difference in real setting. In an organisational setting, for instance, a manager commits Type I error if an employee is sacked based on some evidence of theft, whereas the employee did not really steal. This is similar to jailing an innocent suspect.

Conversely, Type II (false negative) error is made when a researcher fails to reject the null hypothesis when it is false. If a researcher fails to observe a difference when in reality there is one, a Type II (β) error is made. Suppose a car owner concludes that it was not yet time to service the car because
the engine was not noisome, whereas the engine pump was failing, the car owner commits Type II error. Organisations endeavour not to commit Type II error because it is costly and regrettable. Imagine not sacking the employee that actually stole company property, or declaring the suspect innocent, whereas he actually committed the crime.

Step 4: Specify the test statistic

Test statistic is a value, determined from sample information, used to decide whether to reject the null hypothesis or not. Test statistics (e.g., $t$, $p$, $F$) and a corresponding $p$ value are calculated, most often by computer, where:

$$ t = \frac{b_1 - \beta_1}{S_b} = \frac{b_1 - 0}{S_b} = \frac{\hat{b}_1}{S_b} = \frac{SSE}{n-1} = \frac{\sum x^2 - (\sum x)^2}{SSE} $$

$b_1$ = Sample regression slope coefficient
$\beta_1$ = Hypothesized Slope (usually, $\beta_1 = 0$); degree of freedom (df) = $n - 2$
$S_b$ = Estimator of the standard error of the slope
$S_x$ = Standard error of the estimate; $SSE$ = Sum of Squares Error; $n$ = Sample size

The $p$ value indicates the probability of obtaining a value of the test statistic that deviates as extremely (or more extremely) as it does from the null hypothesis prediction if the null hypothesis were true for the population from which the data were sampled.

Test to see whether $p$ is significantly different from $\zeta$, the test statistic is:

$$ t = \frac{r}{\sqrt{1 - r^2}} $$

Test to see whether $p^2$ is significantly greater than 0, the test statistic is:

$$ F = \frac{SSR}{SSE} $$(n-2)

Where SSR = Sum of Squares Regression; SSE = Sum of Squares Error

Step 5: Formulate the Decision Rule

A decision rule is a statement of the specific conditions under which the null hypothesis is rejected and the conditions under which it is not rejected. The region or area of rejection defines the location of all those values that are so large or so small that the probability of their occurrence under a true null hypothesis is rather remote.

For $t - test$, if the calculated $t - value$ falls outside $ \pm $ the tabulated (critical), it means there is a significance difference. Hence, the null hypothesis will be rejected. Also, if a significance test yields a value of equal to or less than, the null hypothesis is rejected and the result is said to be statistically significant at that level.

Step 6: Compute the test statistic and compare it to the critical value

At this stage, the researcher computes the test statistic and compares it to the critical value. Hence, the calculated is compared with the table (critical), or the critical $p$-value (0.05) criterion is deployed.

At this stage, it is advisable to draw or construct a diagram that shows the rejection region of the sample distribution.

Step 7: Draw a conclusion regarding the null hypothesis

The researcher makes a conclusion regarding the null hypothesis based on the sample information. The decision is to reject or not to reject the null hypothesis. The researcher then interprets the results of the test for managerial decision making.

2.3: Regression Analysis

Regression analysis is a modelling technique that specifies and quantifies the relationship between an endogenous variable and one or more exogenous variables(s). Specifically, the regression approach is targeted at identifying a function that explains the relationship between the endogenous and exogenous variables in order to predict the amount of variation in the endogenous variable accounted by a unit change in the exogenous variables(s) (Ziegel, & Ragsdale, 1998). The endogenous variable (typically denoted as Y) is also called dependent-, criterion-, target-, observed-, output-, prognostic- response- or outcome variable or regressand, while the exogenous variable (often denoted as X) is also known as independent-, predictor-, selected-, carrier-, input-, manipulated-, or explanatory variable or regressor.

Render, Stair and Hanna (2012) submit that regression analysis is useful in predicting future business outcomes by using historical data. Managers could deploy regression analysis to forecast demand, prices and profits, while governments and econometricians use it to project population, Gross National Product, revenue and expenditure for efficient economic planning. Other areas in which regression analysis is used include facility location problems, level of education and income, the price of a house and the square footage, and the sales volume for a company relative advertising budget. Moreover, regression models are used in cost estimation (Elfahham, 2019), quality control (Mandel, 1969; Hayati, 2017), capital asset pricing (Sharpe, 1964; Lintner, 1965), employee turnover forecasting (Wong, Chan & Chiang, 2011; Zhu, Seaver, Sawhney, Ji, Holt, Sanil & Upreti, 2016) and countless quantitative domains.
Regression models could be simple (bivariate), multiple (multivariate) or non-linear. A bivariate model involves only one independent - and one dependent variable. A multivariate model comprises one dependent variable and two or more predictors. In multiple nonlinear regression (MNLR) analysis, the relationship between the exogenous and the endogenous variables are assumed to be nonlinear. Nonlinear regression can estimate a model using random relationship between dependent and independent variables. This paper focuses on simple linear regression.

In simple linear regression, values of the independent - and dependent variables are computed, and a model (known as the regression line or regression equation, or line of best fit) is constructed to establish a relationship between the two variables. Hence, more probable values of a dependent variable can be predicted based on selected quantities of the independent variable.

For instance, Alpha Star, a company in the Niger Delta Region of Nigeria that manufactures paints, wishes to predict sales volume of its product with respect to advertisement budget. After collecting data from the past 12 months, the operations research department hypothesizes that higher levels of advertising expenditure will significantly promote sales performance. Here, the simple regression model would specify advertising expenditure as "x" and sales performance as "y". Hence, sales performance (y) will be regressed on advertising expenditure (x).

The simple (bivariate) regression line is modelled as:

\[ Y = \alpha + \beta x + \epsilon \]

Where:
- \( x \) = independent variable (the regressor or variable to be manipulated, advertising expenditure)
- \( Y \) = dependent variable (the outcome variable, sales performance)
- \( \alpha \) = constant which indicates the point of intercept between the regression line and the \( Y = \alpha \) axis (the value of sales performance when advertising expenditure is zero)
- \( \beta \) = slope or regression coefficient, which estimates how many units of sales performance (Y) will change when advertising expenditure (x) changes by one unit.
- \( \epsilon \) = residual (Gaussian) error, i.e., variation due to random error from poor estimation of parameters or unsystematic component of the model.

Apart from modelling the regression line, the amount of variation in the dependent explained by the model is also determined. This proportion of variability in the endogenous variable that is explained by the regression model is called the coefficient of determination \( r^2 \). Although \( r^2 \) indicates the extent of prediction in the regression model, sometimes it scores high values on the relationship between the variables, even when the sample size is very small.

Hence, \( r^2 \) does not guarantee predictive accuracy (Gravetter, & Wallnau, 2017), which may lead to spurious conclusions. Moreover, adding more predictors to model can artificially inflate and cause overestimation. The introduction of - adjusted does some remediation in this aspect, but that is not enough to prove statistical significance. Gatekeepers of management science suggest that a good method of determining the meaningfulness of the regression model is the hypothesis test for significance (Render, Stair & Hanna, 2012).

Mathematically, \( r^2 = 1 - \frac{SSE}{S_{yy}} = 1 - \frac{SSE}{\sum y^2 - \frac{(\sum y)^2}{n}} \)

where \( SSE = \text{Sum of Squares Error}; 0 \leq r^2 \leq 1 \)

3.1: REGRESSION ANALYSIS AS PREDICTION AND HYPOTHESIS TESTING TOOL

Managers of yore adopted qualitative approaches in decision making, but contemporary managers have embraced quantitative methods in decision making. Chiefly among the methods is regression analysis. This approach performs the dual purposes of enabling the decision maker understand the relationship between variables, and to predict the value of one variable due to a unit change in the other. A simple regression line for a population is modelled as \( Y = \alpha + \beta x + \epsilon \)

Suppose all points are on the regression line, the model becomes deterministic and error free. In such case, the regression line becomes \( Y = \alpha + \beta x \)

However, since most business data utilize sample statistics and not population parameters, the regression formula can be re-specified as:

for a given number of observations, \( \hat{y} = b_0 + b_1 x \), where:

\[ b_1 = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} \]

The expression in the numerator of the slope \( b_1 \) formula is denoted as \( SS_{xy} \), while the denominator is denoted \( SS_{xx} \).

Thus,

\[ b_1 = \frac{SS_{xy}}{SS_{xx}} \]

Also, the y intercept of the regression line \( b_0 = \bar{y} - b_1 \bar{x} \)

To determine the equation of the regression line for a sample of data, the researcher must determine the values for \( b_0 \) and \( b_1 \). This process is sometimes referred to as least squares analysis. Least squares analysis is a process whereby a regression model is developed by producing the minimum sum of the squared error values (Black, 2010).

The remaining part of this paper demonstrates the estimation and predictive properties of linear regression. A statistical exercise is also carried out on the use of regression in hypothesis testing of significance. We collected data on advertising expenditure and sales of drums of paints in Alpha Star Paints Industry Limited in Port Harcourt, Nigeria. Monthly sales (in drums of paints) and monthly advertisement expenditure (in 0,000 naira) for the past 12 months are shown in table 1:

<table>
<thead>
<tr>
<th>Month</th>
<th>Advertising Expenditure (0,000 naira)</th>
<th>Sales Performance (in drums of paints)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>Feb</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>Mar</td>
<td>30</td>
<td>70</td>
</tr>
<tr>
<td>Apr</td>
<td>35</td>
<td>65</td>
</tr>
<tr>
<td>May</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Jun</td>
<td>45</td>
<td>55</td>
</tr>
<tr>
<td>Jul</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Aug</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>Sep</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>Oct</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td>Nov</td>
<td>70</td>
<td>30</td>
</tr>
<tr>
<td>Dec</td>
<td>75</td>
<td>25</td>
</tr>
</tbody>
</table>

The data collected were used to determine the regression line for the relationship between advertising expenditure and sales performance. The slope of the regression line is calculated as:

\[ b_1 = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = \frac{1500 - \frac{1800}{12}}{300 - \frac{900}{12}} = \frac{1500 - 150}{300 - 75} = \frac{1350}{225} = 6 \]

The y intercept of the regression line is calculated as:

\[ b_0 = \bar{y} - b_1 \bar{x} = 71.54 - 6 \times 33.33 = 7.50 \]

Hence, the regression line is given by:

\[ \hat{y} = 7.50 + 6x \]

This equation can be used to predict the sales performance for any given level of advertising expenditure. For example, if the advertising expenditure is 50,000 naira, the predicted sales performance would be:

\[ \hat{y} = 7.50 + 6 \times 50 = 327.50 \]

The statistical significance of the regression model can be tested using the Student’s t-test. The t-statistic for the slope is calculated as:

\[ t = \frac{b_1}{SE(b_1)} \]

where \( SE(b_1) \) is the standard error of the slope. The t-statistic is compared to the critical t-value from the t-distribution table for the given degrees of freedom and level of significance.

In this case, the t-statistic is calculated as:

\[ t = \frac{6}{0.5} = 12 \]

The critical t-value for a two-tailed test with 10 degrees of freedom at a 0.05 level of significance is 2.228. Since the calculated t-statistic (12) is greater than the critical t-value, we can reject the null hypothesis and conclude that there is a significant linear relationship between advertising expenditure and sales performance.
4.0: ANALYSIS

The manual method to estimate the regression model is as follows:

Next is the determination of residuals and sum of squares error. The residual is the difference between the observed sales ($y'$) and the predicted sales ($\hat{y}$). Residual = $y - \hat{y}$

The sum of all the squared residual errors is the Sum of Squares Error = $\sum (y - \hat{y})^2$

Table 3 shows the predicted values based on the regression model, the residuals and SSE.
Next is the calculation of $r^2$ which shows the proportion of variation in sales that is explained by advertisement expenditure.

\[
\text{Recall that } r^2 = 1 - \frac{\text{SSE}}{\sum y^2} = 1 - \frac{\text{SSE}}{\sum y^2 - (\overline{y})^2} \]

We have $\sum y = 2.636$. Also recall that, in table 2, $\sum y^2 = 741.022$

Hence, $r^2 = 1 - \frac{89.416.71}{741.022 - \frac{(2.636)^2}{N}} = 1 - \frac{89.416.71}{741.022 - \frac{(2.636)^2}{N}} = 0.5552 = 0.448$

The $r^2$ tells a story that a unit increase in advertisement expenditure will determine 44.8% increase in sales. The balance of 55.2% is due to unexplained factors not captured by the model.

However, the computation of does not reveal statistical significance. The next section of this paper demonstrates a test of the hypothesis which states that “advertisement expenditure does not significantly predict sales in Alpha Star Paint Industries Limited”.

**Step 1: Specify the population parameter of interest**

The parameter of interest is the population slope ($\beta$). The test is carried out to know whether the slope of the regression line will be different from zero, if all the pairs of data points for the population were available. The question is: can the sample slope ($b_1$) be a good estimator of the population slope ($\beta$)?

**Step 2: Formulate the null hypothesis ($H_0$) and the alternate hypothesis ($H_1$) in terms of the population parameter**

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

**Step 3: Specify the desired Significance Level ($\alpha$)**

$\alpha = 0.05$, which is the probability of making Type I error

Since this test is two tailed, $\alpha/2 = 0.025$; while, $df = n - 2 = 12 - 2 = 10$

The table (critical) $t_c$ value at $t_{0.025,10}$ = ± 2.228

**Step 4: Specify the Test Statistic**

Test Statistic for Significance of the Slope is given as:

\[
t = \frac{b_1 - \beta_1}{s_{b_1}} \text{ where: } S_b = \sqrt{\frac{SSE}{n-2}}; S_x = \sqrt{\frac{\sum x^2}{n}} \sum x^2 - \frac{(\sum x)^2}{n}
\]
Step 5: Formulate the Decision Rule

For $t$-test, a calculated $t$-value that falls outside $±\tau_c$ the table (critical) $t_c$-value is an evidence of significance that warrants rejection of the null hypothesis. For $p$-value criterion, if $p$ is equal or less than .05, the null hypothesis is will be rejected, and the result will be regarded as statistically significant at that level.

Step 6: Compute the test statistic and compare it to the critical value

$$t = \frac{b_1 - 0}{S_b} = \frac{b_1}{S_b}$$

where: $S_b = \frac{S_{xx}}{\sqrt{SS_{xx}}}$; $S_p = \sqrt{\frac{SSE}{n-1}}$; $S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$

$SSE = 89,416.71$ (see table 3); $\sum x^2 = 386; \sum x = 64$ (see table 2); $b_1 = 40.30$

$$S_p = \sqrt{\frac{89,416.71}{12 - 1}} = 90.16$$

$$S_{xx} = 386 - \frac{(64)^2}{12} = 44.67$$

$$S_b = \frac{90.16}{\sqrt{44.67}} = 13.49$$

Hence, the observed $t$-value for this sample slope is:

$$t = \frac{40.30 - 0}{13.49} = 2.987$$

This stage entails a comparison between the calculated $t$-value and the table (critical) $t_c$ value, or the deployment of the $p$-value (0.05) criterion. But, before that is done, the excel software is utilized here to substantiate the manual method.

We migrated to adds-in on excel platform and added “data analysis”. After adding “data analysis”, we enter the data vertically on the excel platform. Values for the independent variable (advertisement expenditure) were entered under column $X$ and those of the dependent variable (sales) under column $Y$. Next, we clicked data analysis. We scrolled down to regression on the data analysis environment and clicked OK. For input range, we selected the values, and for input range, we selected the values. We then clicked OK. The output appeared on a different sheet, as shown in figure 1.

![Figure 1: excel output for Regression model](image-url)
Table 4 shows the comparison of relevant outputs from manual calculation and excel spreadsheet.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Manual Output</th>
<th>Excel Spreadsheet output</th>
<th>Critical $t_c$</th>
<th>Observed $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept ($b_0$)</td>
<td>4.871</td>
<td>4.702</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Slope ($b_1$)</td>
<td>40.303</td>
<td>40.306</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Coefficient of Determination ($r^2$)</td>
<td>0.448</td>
<td>0.448</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>SSE (Residual)</td>
<td>89416.711</td>
<td>89416.485</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>$t$-value</td>
<td>2.987</td>
<td>2.849</td>
<td>± 2.228</td>
<td>Either 2.987 or 2.849</td>
</tr>
<tr>
<td>$p$-value</td>
<td>Apriori value</td>
<td>0.017</td>
<td>0.050</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Furthermore, the $t$-value (2.987: manual method, 2.849: excel output) calculated from the sample slope falls in the rejection region since it is greater than the table (critical) $t_c = \pm 2.228$. Also, $p = \text{value} = 0.017 < 0.05 \Rightarrow \text{connotes statistical significance}$. Figure 2 is the diagram of the area of rejection.

Step 7: Draw a conclusion regarding the null hypothesis

Since the calculated $t$-value is greater than the table (critical) $t_c$, the null hypothesis that "advertisement expenditure does not significantly predict sales in Alpha Star Paint Industries Limited" is rejected.

Thus, there is statistical evidence that higher levels of advertisement expenditure will significantly magnify sales performance in Alpha Star Paint Industries Limited. Therefore, the management of Alpha Star Paint Industries Limited is advised to increase advertisement expenditure.

5.0: CONCLUSION

Based on the foregoing, this paper concludes that decision making is a complex process that comprises qualitative and quantitative approaches. Whereas employees and lower level managers often use rational methods in tandem with laid down policies and processes, top managers tend to adopt judgement or intuition as they encounter uncertainty in the business environment. The study also concludes that over-reliance on heuristics by managers in the decision making process jeopardizes the decision benefits due to inherent cognitive biases.

In order to mitigate the weaknesses of heuristic or cognitive styles of decision making, firms should be aware that the effectiveness of their decisions will improve through statistical analysis. Consequently, the implication is that there is need for managers to develop models in order to substantiate their claims, minimize the deleterious random effects of judgements, and provide new insights to optimize organisational outcomes. Furthermore, based on the empirical evidence from this study and literature, the paper concludes that regression analysis is a powerful tool that managers can deploy to understand relationship between two or more variables, and to predict future outcomes. Moreover, regression analysis is a scientific tool that helps managers to reduce large amounts of raw data to actionable information.

Also, the paper concludes that hypothesis testing is, indeed, one of the most effective statistical methods to determine if a regression model significantly estimates some factors of interest or whether a result is due to chance. Thus, firms utilize hypothesis test of significance to understand how strongly the results of analysis will influence managerial decisions.
Furthermore, despite the epistemological beititudes of regression analysis - due to its scientific credentials - it is not inoculated from certain flaws. The woes of regression analysis include its failure to establish causal relationships and its lack of identification of unmeasured variables outside models which contribute to variations in the dependent variable. Moreover, managers ought to be aware that testing the regression slope is not a silver bullet in decision making. This is because hypothesis testing based on significance tests is not reliable when the samplesize is small.

The paper recommends that managers should not make decisions based only on intuition. Doing so will be counterproductive and stiflorganisational performance. Managers should champion a synergistic combination of quantitative techniques - such as regression analysis and hypothesis testing- and informed judgment to birth quality decision. In addition, for more reliable results in testing regression models, decision analysts should use reasonably large samples.

Lastly, caution should be taken when managers are held accountable for decision outcomes. Organisations should encourage the empirical determination probabilities of occurrence of favourable outcomes, while desisting from making inconsistent risk choices.

REFERENCES


29. Gigerenzer, G., Todd, P. M., & the ABC Research Group 1999. Simple heuristics that ma


82. Strategic Decision-Making


84. The Influence of Decision, Environmental


