*Quest Journals*
*Journal of Software Engineering and Simulation*
*Volume 2 ~ Issue 7(2014) pp: 08-15*
*ISSN(Online) :2321-3795 ISSN (Print):2321-3809*
www.questjournals.org

**Research Paper**

# Test of Hypothesis for Small and Large Samples and Test of Goodness of Fit

*[1]Dr. Mcchester Odoh and [2]Dr. Ihedigbo Chinedum E.
*[1]Director, ICT Michael Okpara University Of Agriculture Umudike, Umuahia, Abia State Of Nigeria.*
*[2]Department of computer science, Michael okpara university of agriculture umudike, Umuahia, abia state of Nigeria.*

## I. INTRODUCTION

We Are Always Faced With Decision Making Situations In Management. In An Attempt To Reach Decision, It Is Useful To Make Assumptions (Or Guess) About The Population Of Study. Such Assumption Which May Be True Or False Are Called Statistical Hypothesis. To Find Where It Is True Or False. The First Chapter Of This Work Cleared Us On Types Of Hypothesis, Errors In Testing, Level In Significant Steps In Testing The Hypothesis. And All We Are Doing Is To See Whether It Is True Or False (Ewurum U. F, 2000).

In Many Circumstances Decision Must Be Made Based Only On Sample Information. A Quality Control Manager Must Determine Whether A New Marketing Strategy Will Increase Sales. An Accountant Must Be Able To Ascertain The Appropriate Book That Gives A True And Fair View Balance. In Drawing This Type Of Conclusion, The Decision Maker Must Be Sure No Error Is Committed. Another Way Of Looking At That Is To Ask If The Frequency Distribution Fits A Specific Pattern. Two Values Are Involved, An Observed Value Which Is The Frequency Of A Category From A Sample And The Expected Frequency, Which Is Calculated Based Upon The Claimed Distribution. The Derivation Of The Formular Is Very Similar To That Of The Variance Frequency Is Really Close To The Claimed Distributions. (Ewurum U.J. F, 2000) (Chukwu B. I. 2006).Furthermore, The Following Are The Test Of Hypothesis.

A. The Data Are The Observed Frequencies And One Data Value For Each Category.
B. The Degree Of Freedom Is One Less Than The Number Of Categories, Not One Less Than The Sample Price.
C. It Is Always A Right Tail Test But Something Wrong.
D. It Has A Chi -Square Distributions.
E. The Value Of Distribution Doesnt Change If The Order Of The Categories Is Switched. The Test Statistic Is Interpreting The Claim. The Chapter Two And Three After Term Paper Looked Into The:
A. Population Distribution.
B. Calculated 2
C. Procedure Of Testing
D. Testing For One Normal Distribution
E. Testing For The Normal Distribution. All The Above Mentioned Are Go The Chapter Two. The Chapter Three Talked About The Following:
A. Test Of One Population Distribution.
B. Test Of Two Population Distribution. (Nwandu Ezima C (2003)(Chukwu B. I.2007) (Www.Goafle.Com), Ewurum Jonathan (2003).

**Types Of Hypothesis**

1. Null Hypothesis: This Is The Original Position Prior To The Experiment. It Is Denoted Mathematically As Ho. It Is The Hypothesis That States That There Is No Contradiction Between The Arrested And The Sample Statistic And That The Different Can Therefore Be Ascribed To Chance.
   In Many Instances We Formulate A Statistic Hypothesis For Sole Purpose Of Rejecting Or Nullifying It. For Instance We Want To Decide Whether A Given Coin Is Loaded, We Formulate The Hypothesis That The Coin Is Fair (I.E P -0.5 Where P Is Probability Of Heads). Also If We Want To Decide Whether One Procedure Is Better Than Another, We Formulate The Hypothesis That There Is On Difference Between The Procedures, That Any Observed Differences Are Due Merely To Fluctuations In Sampling From The Same Population.

2. Alternative Hypothesis: Any Hypothesis That Differs From A Given Hypothesis States That P=0.5, Alternative Hypothesis Might Be P=0.7, P=0.5 Or P> 0.5 Alternative Hypothesis Is Demonstrated By Hi.
3. Error In Testing: We Have Two Main Error Testing
   A. Type One Error: He We Reject A Hypothesis When It Should Be Accepted, We Say That A Type 1 Error Has Been Committed.
   B. Type Two Error: If We Accept A Hypothesis When It Should Be Selected, A Type 2 Error Has Been Committed.

| | **Ho True** | **Hi True** |
|---|---|---|
| **Ho Accepted** | Corrected Decision | Type Ii Error |
| **Ha Accepted** | Type I Error | Corrected Decision |

In Order For Decision Rules To Be Good, They Must Be Designed To Minimize Error Of Decision. This Is Not A Simple Matter Because As We Decrease The Risk Of One Type Of Error We Increase The Risk Of The Other Type Of Error.

In Testing A Given Hypothesis, The Maximum Probability With Which We Would Be Willing To Risk A Type 1 Error Is Called Level Of Significance. It Is Denoted Mathematically As A. The Rejection Of Significance Level Depends On Which Error We Consider More Serious. The Selected Of A True Hypothesis Or To Fell To Select One That Is False.

In Practice A Significance Level Of 0.05 Or 0.01 Is Simple, The 0. 05 (Or 5%) Significance Level Is Closer In Designing A Decision Rule, Them There Are About 5 Chases In 100 That We Would Select The Hypothesis Yen It Should Be Accepted. That Is We Are 95% Confident That The Right Decision Would Be Taken And 5% Chance Of Taken A Wrong Decision. However It Is Very Important That Significance Level Is Selected Before The Testing Of Hypothesis So That The Result Obtained Will Not Influence Our Choice.

**1d. Level Of Significance:**

In Level Of Significance, We Mean The Region At Acceptance Or Rejection At The Hypothetical Statement. In Its Table, It Is Mathematically Called The Tabulated Z And Denoted By A.

The Table Below

| **Level Of Significance** | **0.10** | **0.05** | **0.01** | **0.005** | **0.002** |
|---|---|---|---|---|---|
| **Critical Value Of Z For One Fail Test** | -1.2800 | -645 | -2.33 | -2.5800 | -2.88 |
| **Critical Value Of Z For One Fail Test** | -1.645 Or 1.645 | -1.96 Or 1.96 | -2.58 Or 2.58 | -2.81 Or 281 | -3.08 Or 3.08 |

**1e. Steps of Testing Hypothesis**

Many Authors Have Done Lots Of Research Work On The Steps Of Testing The Hypothesis; Some Came With The Conclusion That There Are Seven Steps. Some Say Nine, And Others Concluded That They Are Twelve (12) Steps. It Is As Following:

i. Statement Of Null Hypothesis.
ii. Statement Of The Alternative Hypothesis
iii. Statement Of The Significant Level Or Tabulated Z.
iv. Determination Of The Sample Size (N).
v. Determination Of Test Statistics.
vi. Setting Up The Critical Values That Divide The Region Into Rejection And Non Rejection Region.
vii. Statement Of The Decision Criterion.
viii. Collection Of The Date, Computing The Sample Value Of The Appropriate Test Statistics.
ix. Determine Whether The Test Statistic Has Fallen Into The Rejection Or The Non Rejection Region.
x. Determine The Statistical Decision.
xi. Expression Of The Statistical Decision Or Making Comparism.
xii. Giving The Administration Decision.

**Explanation**

Step 1 And 2: Statement Of Null And Alternate Hypothesis. The Null And Alternate Hypothesis Must Be Stated In Statistical Terms. That Is He We Are Testing The Number Of Building In A Given Community Or Number Of Tyre Life

Or Anything. Example

Ho=Θ
Ho≠Θ
Step 3: The Level Of Significance: This Has A Table Stated Above For It. Here We Make Use Of The Table And It Is Specified According To The Importance Of A And Θ In The Problem.
Step 4: The Sample Size Determined By Taking Into Account The Importance Of A And Θ, And By Considering Budgeting Constraints In Carrying Out The Study.
Step 5: Taking A Number Of Factors Into Consideration The Appropriate Test Statistics Must Be Determined.
Step 6: Once The Null And Alternate Hypothesis Are Known And The Level Of Significance And Sample Size Decided Upon, The Critical Values Of The Appropriate Distribution Can Be Set Up To Indicate The Rejection And Non Rejection Region.
Step 7: The Decision Rule Must Be Stated. In This Example, The Rule Would Indicate As Follows: Rejection Ho He The Computed Value Of The Appropriate Test Statistic Is Greater Than The Critical Value Otherwise Do Not Reject Ho.

The Method To Be Used To Determine Whether The Sample Statistic Has Fallen Into The Rejection Or Non Rejection Region Must Be Determined. The Appropriate Test Statistic Must Be Indicated Along With How The Sample Statistic Is To Be Compared To The Hypothesized Parameter. For Example In The Case Of Means, We Find How Far The Statistic Deviates From The Hypothesis Parameter In Standard Deviation Units.
Formula =Ho/Standard Error
Step 8: Here We Collect The Data And Make The Appropriate Computation.
Step 9: The Value Of The Test Statistic Is Compared To The Critical Value Of The Appropriate Distribution To Determine Whether It Falls In The Rejection Or Non Rejection Region.
Step 10: The Hypothesis Testing Decision Is Determined He The Test Statistic Falls Into The Non Rejection Region, The Null Hypothesis Ho Cannot Be Rejected. If The Test Statistic Falls Into The Rejection Region The Null Hypothesis Ho Is Rejected.
Step 11: Once The Decision Is Made, Its Consequences Must Be Expressed In Terms Of The Particular Problem. Example, Using The Number Of Building In A Community Or Number Of Tyre Produced By A Company E.G Machline.
Step 12: The Conclusion May Be Taken Based On The Correctiveness Action Determined By The Process Not Working Properly.
Conclusion: In This Chapter, We Have Stated The Test Of Hypothesis, The Steps On How Hypothetical Statement Can Be Derived And Computed. We Also Talked About The Level Of Significance And Confidence Level.
Lastly, We Discussed About The Error In Testing, We Found Out Two Types Of Error I.E Error I And Ii.

## II. POPULATION DISTRIBUTION

**2.1 We Have Two Distributions**
A. One Tail Test
B. Two Tail Test.
A. **One Tail Test:** This Is Only When One Particular Hypothetical Statement Is Being Tested.
B. **Two Tail Test**: This Is When We Are Testing Or Carrying Out A Test On Two Population Hypothetical Distribution At The Same Time.
**2.2 Calculated Z**: The Calculated Z Is The Result Gotten From The Test After All The Collections And Computation Of Data. While Tabulated Z Is The One We Discussed Above. It Is Given In A Table, That Means That After So Many Computations, We Compare The Calculated Z And The Tabulated Z To How To Draw The Conclusion. That Is Whether To Reject Or Accept The Hypothetical Statement.
Formula For Computation
X-U/ Σ /√N
X= The Given Distribution In Testing
U = The Mean Of The Given Distribution
Σ = The Standard Deviation Of The Normal Null Or Alternate Distribution.
N= The Number Of Sample Or The Sample Size Given.

**2.3 Procedure Of Testing**

Testing Of Hypothesis For The Mean Of One Normal Distribution. Let Us Consider The Following Problem. It Is Believed That A Particular Plant Has Two Varieties. Whenever Researchers Finds A New Plant Of This Kind, They Try To Classify The Plant Into One Of The Two Varieties By Its Top Leafs Vertical Diameter. But They Are Not Hundred Percent Certain About This. It Was Also Observed That Some Of The Varieties Have A Slight Different Top Leaf Vertical Diameter. The Researchers Has Found That Variety A Has

A Mean Variety Diameter Of 19.0 Millimeters Whereas The Same For Variety B Is 19.6mm. The Standard Deviation For Both Varieties Is 0.8mm A New Search Produced 12 Plants. The Researchers Feet That There Is Reasons To Believe That These Belong To Variety A. The Mean Diameter Of There 12 Plants Is 19.4m. The Problem Is To Test The Hypothesis That These Plants Belong To A Rather Than B. That Is, Which Of These Hypotheses Is True.

Solution

Ho = 19.0

Hi = 19.6 The Mean = 12

The Plant Level Is 19.6

Note: Now, The Problem Is Km Decide (Test) Which One Of These Hypothesis Should Be Accepted. Many People May Use Good Sense In This Particular Problem And Decide In Favour Of Ho He The Samsue Mean Is Closer To 19 Vice 19.6 And In Favour Of H1 It Is Otherwise. Ant The Researchers Have Other Reasons To Believe Vice Ho Is True So They Would Insist On Ho Being True, Except If The Sample Mean (X) Is Fairly Close To The Mean Corresponding To H1, Before They Would Be Willing To Give Up The Hypothesis Ho In Favour Of Hi.

In Taking These Decision One Must Be Conscious Of The Probability Of Making Either A Type 1 Or Type Ii Of Calculating These Probabilities, Let Up Assume That X Is Normally Distributed. Then X Is Normally Distributed With Standard Deviation.

**Illustration I**

Many Years Ago Experience Has Shown That Waec Exams In Mathematics Show An Average Of 64 And A Standard Deviation Of 10. A Sample Of 49 Students From A Certain Town Obtain A Score Mean Of 60 Can We Conclude That Students From This Town Are Inferior In Mathematics. Using 5% Significance Level.

Solution : Using The Steps Above.

Step 1: Statement Of Hypothesis.

Ho=64

Hi≠64

Step 2: Statement Of Level Of Significance Is Equal To 5%

Step 3: Statement Of Critical Value A (Tabulated Z)=±1.645

Step 4: Statement Of Sample Statistics

X=60

U=64

Σ =10

N=49

Step 5: Test Of Main Hypothesis. Here We Calculate Out Our Calculated Z

Formular

$$\frac{X - U}{\frac{\Sigma}{\sqrt{N}}}$$

X=Sample Mean

U=Population Mean

Σ=Standard Diviation

N=Sample Size

Substitution Where

X=60, U=64, Σ =10, N=49

$$Z = \frac{60-64}{\frac{10}{\sqrt{49}}}$$

$$\text{Calculated } Z = \frac{-4}{\frac{10}{7}}$$

$$\text{Calculated } Z = \frac{-4}{10} \times \frac{7}{1}$$

= -2.8

Calculated Z = -2.80

Step 6: Comparism Tabulated Z = -1.645

Calculated Z = -2.80

Step 7: Decision From The Above Analysis Rejects The Null Hypothesis Because The Tabulated Z Is -1.645 While Calculated Z Is -2.80
Step 8: Conclusion: We Are Making Our Conclusion That, The Past Experience Does Not Show The Average Performance Of The Students In That Community On Waec Is Not 64, Therefore, We Unanimously Reject It.

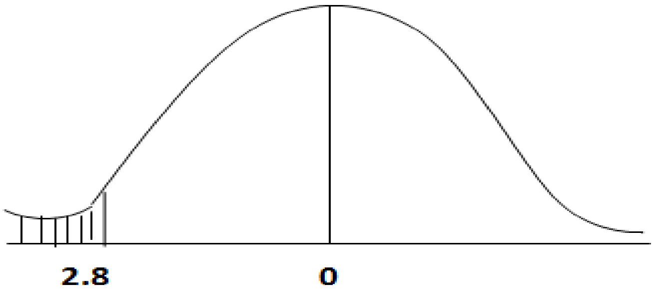


**Illustration 2**
A Large Retailer Wants To Determine Whether The Mean Monthly Income Of Families Living Within 2km Of A Proposed Building Site Exceeds N24.400. What Can Be Concluded At 0.05 Level Of Significance, He The Mean Income Of A Random Sample Of N =60 Families Living Within 2 Km Of The Proposed Site Is X =24524 And The Standard Deviation Is @=733?.
Solution
Step I: Statement Of Hypothesis  Ho =24,400
Hi #24,400
Step Ii: Statement Of Significance Level @=5%
Step Iii: Statement Of Tabulated Z =5% That Is + 1.645
Step Iv: Statement Of Sample Mean (Statistics)
X =24524
U =24400
Σ =763
N=60
Step V : Testing Of The Hypothesis Proper Formula
X -U / Σ√N
Substitution

(a) 24524-24400

$$\frac{762}{\sqrt{60}}$$

Calculated Z=   124
763
7.7
Calculated Z= 124
98.57
Calculated Z= 1.257,
=1.26
Step 6: Comparism
Calculated Z =1.26
Tabulated  Z =1.645
Step Vii: Decision: We Shall Reject The Null Hypothesis As The Calculated Z Is Greater Than The Tabulated Z I.E Calculated Z Is 1.26 While The Tabulated Z Is Greater Than Tabulated Z.
Graph Below
(X: X> X < 1.2)

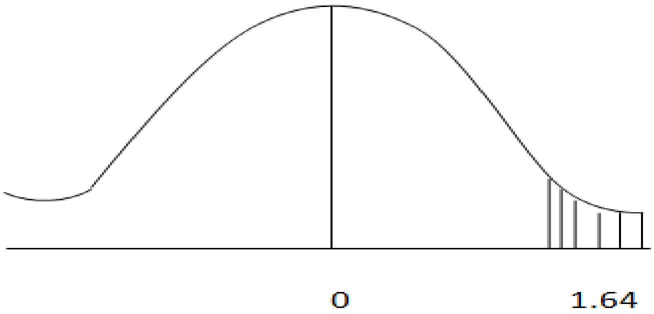


**Illustration 3:**
A Trucking From Suspect That The Average Life Time Of 25,000km Claim, The Firm Puts A Random Sample Of 40 Of These Tyres On Its Trucks And Later Finds That Mean Literature Is 24,421km And The Standard Deviation Of 1349km What Can Be Concluded At 0.01 Level Of Significance.

**Solution**

Step 1: Statement Of Hypothesis
Ho=25,000
Hi=25,000.
Step Ii: Statement Of Level Of Significance=1% (1 Percent).
Step Iii: Tabulated Z = + 2.326=2.33.
Step Iv: Statement Of Sample Statistic (Mean)
X=24421
U=25,000
Σ =1349
N=40.
Step V: Testing The Hypothesis.
Formula X-U/ Σ/√N
Substitution 24421-25000/1349/√40
Calculated Z= -579/1349/6.32
Calculated Z=-579/2.345
Calculated Z= -2.71
Step Vi: Comparism
Calculated Z=-2.71
Tabulated Z=-2.33.
Step Vii: Decision. We Reject The Null Hypothesis Because The Tabulated Z Is Greater Than Calculated Z.
Step Viii: Conclusion: We Are Concluding That The Average Determine Of 25,000km Claimed For Certain Types Of Trucking Is Not 25,000. Therefore We Reject It.

**Test Of Hypothesis For Two Normal Distribution**

We Talk Of Two Normal Distribution When We Have Two Of Sample Mean, Standard Deviation, Population Mean And Sample Size. Therefore We Use The Below Formulars.

(b) Calculated Z (X1-X2)-0
√S1 +S2
N1 N2

Where:
X1= Sample Mean Of 1st Population
X2= Sample Mean Of 2nd Population
S1= Standard Deviation Of The 1st Population
S2= Standard Deviation Of The 2nd Population
N1= Number Of Sample Of The 1st Population
N2= Number Of Sample Of The 2nd Population Or We Use The Formula.

Calculated Z (X1-X2)-1
√S1 + √S2
N1 N2

We Use The Second Formular When Variance Is Given And Not Standard Deviation. The Formular Is Elaborated Below:

(X1-X2)-0
√S1 +√S2
N1 N2

Illustration I: A Test In Financial Management Was Given To Two Groups Of 2nd Year Students.
Consisting Of 50 Students From Banking And Finance And 60 Students From Accountancy. The Mean Score Of The Banking Class Was 65% With A Standard Deviation Of 8%.
While The Mean Score Of Accountancy Class Was 63% With Standard Deviation Of 5%, Is There Any Significant Difference Between The Performance Of The Two Class? Use 5% Level Of Significance.

**Solution.**

Step I: Statement Of Hypothesis
Ho=0
Hi≠0.
Step Ii: Statement Of Significant Level=5%
Step Iii: Statement Of Critical Value (Tabulated) + 1.96 (Using The Table Above). Step Iv: Statement Of Sample Statistics
X1=65%
X2=63%
S1=8%
S2=5%
N1=50
N2=60
Step V: Test Of Hypothesis
Using The Above Formular

(X1-X2)-0
Cal Z √S1 +√S2
N1 N2

Substitution

Cal Z= (60-63)-0
√8 +√5
50 60

Calculated Z= 2-0
√0.16+0.083
Calculated Z= 2
√0.243
Calculated Z= 2
0.4929
Calculated Z= 4.057
= 4.04
Step Vi
Calculated Z =4.06, Tab Z=1.96
I.E Cal Z>Tab Z
Step Vii: Is To Be Accepted Because Calculated Z Is Greater Than Tabulated Z.
Step Viii: Conclusion: We Conclude That There Are No Significant Difference In The Two Departmental Scores

**2.4 Test Of One Population Proportion**

Formular For Calculated Z
Cal Z = P-P
√P(1-P)
N
Where G = (1-P)
P = Sample Proportion Mean
P = Unknown Proportion Mean
N= Number In The Sample

**Illustration**

The Past Experience Shows That Mark Man Hit Target 60% Of The Time. If During The Next 100 Shut He Succeed 65 Times Would You Conclude Tat His Ability As A Mark Man Has Changed. Use 10% Level Of Significance.

**Solution**

Step I: Statement Of Hypothesis
Ho =60%
Hi ≠ 60%
Step Ii: Statement Of Significance Level =10%
Step Iii: Statement Of Critical Value A=±1.28
Step Iv: Statement Of Sample Statistics
P =65/100 I.E 65%
P = 60%
N= 100 (The Total Shuts)
Step V: Testing The Hypothesis
Formular Below

P-P
√P(1-P)
N

Cal Z = 0.65- 0.60
0.60(1-0.60)
100

Cal Z= 0.50
0.06(0.4)
100

Cal Z = 0.05
√0.24
100

Cal Z = 0.05
√0.0024

Cal Z= 0.05
0.04898
=1.02249
=1.022

Step Vi: Comparism Cal Z =M1.02
Tabulated Z = 1.28
Step Vii: Statement Of Significance Level: Tab Z> The Calculated Z
Step Viii: Decision
We Reject It Because The Calculated Z < The Tabulated Z.
Test About Two Different Population Proportions Formular To Calculated Z

Cal Z = (P1-P2)-0
√(P1-P2)+P2(1-P2)
N1 N2

## REFERENCES

[1]. Nwandu, E.C. (2003). Intermediate Statistic 1st Edition. (Nerce Publication Ogui Enugu, Enugu State Pp 59-89)
[2]. Ewurum, J.U. F. (2003), Diamond Press Ogui Enugu, Enugu State (Pp 123-164)
[3]. Www.Google.Com. (Test Of Hypothesis And Goodness Of Fit).
[4]. Chukwu,B.I. (2007). (Fundamental Business Statistics. A Horsthrone Concept At 2007 ).