

UNIVERSITY OF ILORIN



**THE EIGHTIETH
INAUGURAL LECTURE**

**“REASONING IN THE REALM OF
UNCERTAINTY”**

By

PROFESSOR BENJAMIN AGBOOLA OYEJOLA
B.Sc (ABU); M.Sc.; Ph.D. (Reading)
Professor of Statistics

THURSDAY, 31st AUGUST, 2006

COURTESIES

The Vice-Chancellor,
Deputy Vice-Chancellors (Academic & Administration),
Other Principal Officers of the University,
Provost, College of Health Sciences,
Deans of Faculties, Post Graduate School and Student Affairs,
Professors and other members of Senate,
My Academic and Professional Colleagues,
Non-Teaching Staff of the University,
My Lords, Spiritual and Temporal,
Great Unilorin Students,
Distinguished Invited Guests,
Gentlemen of the Print and Electronic Media,
Ladies and Gentlemen.

INTRODUCTION

It is a great honour and privilege to stand before you this evening to deliver the 80th Inaugural Lecture of this great University. It is the 1st in 2006. It is also the 3rd from the Department of Statistics.

My problem in preparing for this lecture was how to present the materials to satisfy such a heterogeneous audience as are present here today. I will attempt to present it in a non-mathematical way as much as possible. I have chosen the topic "**Reasoning in the Realm of Uncertainty**" as the title of this lecture. My professional colleagues in the Nigerian Statistical Association will recollect that part of our anthem which says "We march on steadfastly as we reason in the realm of uncertainty."

Omniscience is an attribute of God. He is the only one who knows when, why and how things happen for certain. He is the only one who knows tomorrow for certain. The statistician does not play God and would not claim to know the value of anything for certain. Rather, he claims to know the value with some degree of confidence, and not for certain (Adegboye, 1997). He reasons in the realm of uncertainty.

Descartes said:

"When it is not in our power to follow what is true, we ought to follow what is most probable"

STATISTICS

To the layman, Statistics is synonymous with data where data refers simply to facts put in numerical form - figures. He is therefore a collector and compiler of data. To the statistician however, statistics is the science of collecting, analyzing and interpreting numerical information. The subject is concerned with data that

are subject to **uncertainty** and with the methods of drawing conclusions or inferences from uncertain data – **Reasoning in the realm of uncertainty.**

Torrie defined statistics as **the science, pure and applied, of creating, developing and applying techniques such that uncertainty of inductive inferences may be evaluated.** It may also be defined as a body of methods and theory that is applied to numerical evidence when making inferences in the face of uncertainty. The use of statistical methods allows one to put a level of confidence to conclusions made from available data. This implies that the statistician knows that the conclusions may be wrong even when the correct analysis has been carried out. It is only in the statistical practice that one is allowed to be wrong a certain percentage of the time! – The table below shows the type of decisions that can be made from data in the face of uncertainty.

		TRUTH (certainty)	
		YES	NO
DATA (uncertainty)	YES	Correct Decision	(Type 2 Error)
	NO	(Type I Error)	Correct Decision

Modern statistics can be traced to Galton’s invention of the method of correlation, which was built out of the law of heredity (Porter, 1986). It is not an abstract technique of numerical analysis. Fisher’s statistical works were centered on problems in plant breeding and development of analysis of variance (variability) and the theory of experimental design. These were influenced by the need to analyze and interpret plant breeding experiments.

Statistics is the key to technology. It is the language in which man reads the Universe. It is a language with numerical vocabulary, a mathematical grammar and, like any language, has its own view of shaping the speaker’s view of the world. Statistics employs the logic and methodology of mathematics but its domain is strictly in the realm of probability. That is why statisticians are said to **reason in the realm of uncertainty.**

USES OF STATISTICS

The uses of statistics can be summarised as:

- To evaluate the existing conditions.
- To provide information that can be useful in formulating plan for development programme.
- To measure progress

- To guide research and
- For decision making and forecasting.

A lot of data are generated from experiments and surveys in practically all areas of human endeavour. Results of political, economic and social surveys as well as increasing emphasis on drug and product testing are evidence of need for intelligent evaluation of data. Statistics is essential to develop a discerning sense of rational thought that will enable one to evaluate numerical data. It is used in making intelligent decisions, inferences and generalizations under uncertainty.

Statistics has been shown to be relevant to a very wide range of subjects including biology and agriculture (biometry), medicine and epidemiology (biostatistics), economics (econometrics), education and psychology (psychometrics), physical, chemical and engineering sciences (technometrics), and business (quality control). The basic principles and methodologies of each are the same. They only differ in details of application.

Statistics constitutes the only method of subjecting values liable to chance variation (uncertainties) to fixed and reproducible criteria based on logical mathematical considerations. This implies that no scientific investigation is capable of proving anything without the aid of statistics.

VARIABILITY IN DATA

Mr. Vice-Chancellor, Sir, permit me to narrate an experience I had at the beginning of my career. A friend (and classmate) and I were employed as Graduate Assistants at the Institute for Agriculture Research, Samaru in August 1976. As part of our orientation programme, we were taken on a visit to the Horticultural field. Tomatoes from each plot had been harvested and placed on the plots. The two of us were to count the tomatoes on each plot from one end while the horticulture assistants were to start from the other end. Fortunately or unfortunately there was an overlap in the plots counted by the two teams. Contrary to expectation, results from the overlap plots were different. On investigation, my team had the correct figures. The reason was that the other team believed that they had enough experience and can therefore correctly guess the number of tomatoes in each heap. Rather than count, they guessed. This was not only a humbling experience (for a graduate in 1976 to be asked to count tomatoes on the field) but also more importantly an experience that opened my eyes to a major source of unwanted variability and cause of **uncertainty** or error in data.

I shall come back to this later.

The following table was extracted from a publication of the National Population Commission/Federal Office of Statistics.

Table 1: Nigeria's Annual Population Growth.

Year	%Growth Rate
1991	2.83
1992	2.83
1993	2.83
1994	2.83
1995	2.83
1996	2.83
1997	2.83
1998	2.83

Source: National Population Commission/Federal Office of Statistics

What census data were used as the basis for the computation? Do we have a system of registering vital events (births, deaths, migration, emigration etc) or continuous population registers from which the rates could have been calculated? Vital registration in this country is at best voluntary. How were these growth rates then arrived at? Your guess is as good as mine!

Sir Josiah Stamp said:

“Public agencies are keen on amassing statistics – they collect them, add them... and prepare wonderful diagrams (and tables). But what you must never forget is that every one of those figures comes in the first instance from the village watchman, who just puts down what he well pleases”.

Let me discuss three sources of uncertainties or variability.

Natural Variation

Variability is always present in measurements and it is universal. No two individuals or objects are exactly alike. God made each object different from another. There is no one exactly like me! Two individuals subjected to the same experience would not react in exactly the same manner. No matter how many factors we control, we shall always find variation in measurements. No matter how identical a set of twins are, they will have some differences. Variability is inherent in units or individuals, managements or environments. **Variability is the source of uncertainty.**

Variability is greater in biological measurements due to the ways in which the various genetic systems of inheritance act to maintain variability. This variation, which is part of the process, is usually referred to as **natural or random variation.**

Examples of natural variation include differences in:

- Weights, heights or intelligence of identical twins.
- Heights of plants of the same variety of a crop on the same heap.
- Life length of bulbs from the same batch.

This natural variation is often said to be due to *unassignable or unpredictable causes*. In a strict sense however, unassignable variation could be more than the natural variation since it is hardly possible to control all sources of variability. When the cause of the variation is not known we attribute it to natural variation. Estimation of the magnitude of this random or natural variation is important to the statistician in hypothesis testing. It is the standard with which he compares other variations.

Bias

Another cause of variation is bias. This includes *systematic errors, personal errors and mistakes*. Sometimes, when we take a type of measurement on the same object several times, we get slightly different results. This is an example of bias or systematic error. My story on counting tomatoes is a good illustration of bias in measurements. Bias may also be in the form of using substandard measuring devices or not setting a scale to zero before the commencement of weighing. Non-uniform management of individuals, units, plots or materials used for an experiment introduces bias into an experiment. Multiple voting or multiple counting in elections is also an example of bias in elections and census figures respectively.

The quality of data would usually depend on how much of the bias has been eliminated during data collection. Unfortunately the statistician can not correct for the bias after the data have been collected. Both bias and natural variation are considered to be natural variation. It is therefore imperative that bias is eliminated as much as possible in data collection so that natural variation can be efficiently isolated.

Assignable Causes

Variation may also be caused by known factors introduced in the experiment or present in the sample. Such variation is said to be due to ***assignable causes***. For example, the use of different varieties or feeds or different nitrogen levels will produce variations in measurements. Different drugs will induce different reactions in patients. Level of education may affect the access to loan. Such are said to be assignable causes.

Both the researchers and statisticians are usually interested in variability produced by assignable causes. It is used to assess the effects or changes produced by these assignable causes. When the variability in response, produced by a factor, is higher than natural variation we conclude that the effect of the factor is important (or significant) on the response. Therefore, when bias is considered as part of natural variation an otherwise important effect may not be detected.

Total variation in any set of measurements can therefore be written as:

$$\textbf{Total Variation = Assignable causes + Bias + Random Error (Variation)}$$

“Error” in measurement consists of *bias + random error*.

As pointed out earlier, statistics cannot take care of uncertainties caused by bias or systematic error. As much as possible the bias can be reduced or eliminated in experiments or measurements by taking necessary precautions. The precision of an experiment depends largely on how much of the bias/systematic errors are eliminated.

Nothing can be done to remove the random error from measurements (as explained earlier) although it should be taken into consideration in planning experiments. Indeed, this random variation is used as the standard or yardstick in assessing the contribution of the assignable causes to total variation.

For the sake of statistical analysis therefore we have:

$$\textbf{Total Variation = Variation due to Assignable causes + Random Error (Variation)}$$

BASIC PRINCIPLES OF EXPERIMENTAL DESIGN

Let me quickly illustrate how statistics manages unwanted variations through what is today referred to as the three basic principles of experimental design.

Replication

Replication is the application of a treatment to at least two similar units. Randomness of nature can be observed from observations from similar plots, objects or units with the same treatment (under the same management). Estimate of this “natural variation”, usually referred to as Residual or Error Variation, is used as a yardstick for comparing variation due to treatments. If variation due to treatments is higher than natural variation, then we are confident that real differences exist between the effects of the treatments. We need to estimate this natural variation. Estimate of variation can only be obtained from at least two observations. Statistics also tells us that the higher the number of units in the experiment the higher the precision (chances of a correct decision) of the estimates and the greater the confidence in the results. This is the idea of replication. In surveys, replication is referred to as sample size.

Here variability is used to assess effects or impact of treatments. Its reduction is used to increase precision of or confidence in results.

Blocking

We may not always be able to have as many similar units as are required in an experiment. However it may be possible to group the units such that the units in each group (or blocks) are as similar as possible. This is the idea of blocking. Variables used in grouping are those that are likely to influence the response

variables. Treatments can then be applied to the similar units and comparisons of treatment effects compared within the group. Information from different groups can then be pooled together. This is a method through which systematic error or bias can be reduced. Blocking reduces experimental error by identifying and correcting for assignable causes of variation which would otherwise have been taken as part of random variation and therefore increasing the precision of the results.

This concept forms the basis of experimental designs like Randomized Complete Block, Latin Square and Incomplete block designs. It is also the basis of stratification and many other procedures in surveys.

*The more numerous the observations (**replication**) and the less the units vary among themselves (**blocking**), the more their results approach the truth.*

Randomization

We have earlier seen that no matter how similar the units are the responses from these units are bound to differ. Fisher, who developed the concept of randomization, was a gambler and saw randomization as the experimenter playing a game with the devil who has an unknown strategy. His solution to the problem was to allocate treatments to similar units such that each has an equal chance of being allocated in any position thereby counteracting any strategy the devil may choose. *Randomization is not synonymous with haphazardness. It is a procedure!*

Randomization is the basis of many statistical analyses. It justifies our inference from the experiment or sample to the target population. If an experimenter does not carry out randomization and analyses his data, the validity of the conclusion will largely depend on personal judgement rather than on statistical theory. Randomization ensures that comparisons are not biased.

MY CONTRIBUTION TO KNOWLEDGE

During my National Youth Service year (1975/76), I was invited to apply for the position of Graduate Assistant at the Data Processing Unit of the Institute for Agriculture, Ahmadu Bello University, Zaria. The main functions of the Unit were to advise the experimenters on the design and analysis of their experiments, to carry out analysis of data from experiments and develop programs for data analysis.

This meant that I had to concentrate on Biometry which is the area of statistics that relates to the use and development of statistical theory for application in biological sciences. My postgraduate studies were therefore in Biometry. Most of my contribution is also in the area of biometrics.

I would however like to discuss my contribution under four headings.

- (a) Statistical Consulting
- (b) Design and Analysis of Mixed Cropping (Intercropping) Experiments
- (c) On-Farm Research
- (d) Statistics Education

STATISTICAL CONSULTING

Between 1976 and 1986, my major assignment was statistical consulting particularly to researchers in Agriculture. The main role of the statistical consultant includes:

- (i) What relevant data or information to collect
- (ii) How to collect the relevant data
- (iii) What analysis to be carried out, and
- (iv) How to correctly interpret the outcome of the analysis

Categories of People serviced by the Statistician

The statistical consultant is faced with broadly three categories of customers classified by their knowledge of statistics as follows (Oyejola, 1988):

1. **None** – It is extremely difficult to discuss statistical concepts with this group. The statistician therefore needs to be versed in the basic principles of statistics. He also needs to be able to speak the language of his client. It is however easy to convince this group of the need for the involvement of the statistician at every stage of experimentation as well as in making inferences from the data collected.

It is as a result of my experience with this group that I have written two books - Design and Analysis of Experiments for Biology and Agriculture Students (Oyejola, 2003) and Basic Statistics for Biology and Agriculture Students (Oyejola & Adebayo, 2004). These are texts that explain basic statistical principles as they apply to Biology and Agriculture.

2. **Little** – This group could be problematic and rigid. They are the “this is the way we have always done it” type. They would collect data (without ascertaining that the design is correct) and expect you to obtain the results they desire even when the design is wrong. The relevant data might not even have been collected. They will resist the use of new and appropriate methods of analysis. They are the Amateur Statisticians. The only reason they come to the statistician is that they don't have the time to do the

analysis. Unfortunately, many researchers (both in Universities and Research Institutes) belong to this group. They have taken one course or the other in statistics and feel competent to handle all the relevant statistical issues involved in their projects.

Unfortunately statisticians have been made to succumb to providing non-informative or probably wrong solutions to problems. For example, a supervisor may insist that a particular statistical test is to be used otherwise he would not approve the write-up of the student. Lack of adequate input of statisticians to research and development projects has strong implications for the quality of project results and the poor rate of acceptance of publications in international journals.

R.A. Fisher said:

“To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may only be able to say what the experiment died of”.

3. **Adequate** – It is easy to work with this group but unfortunately few belong to the group. They consult before carrying out an experiment or survey because they appreciate the value of good designs and correct method of data collection.

Most researches are multi-disciplinary and interdisciplinary in nature and require participatory and collaborative approaches. Biometry as an aspect of scientific methodology involves fundamental instruments that are based on probability theory and stochastics which are subjects that are hardly attractive to collaborators and experts in other disciplines. Statistical theory alone however cannot provide explanations for phenomena in these disciplines. For useful results to come out of our research efforts there must be collaboration between scholars from different disciplines.

Demands on the Statistician

In a research institute or university, a consulting statistician is expected to advice on design and analysis of experiments and/or surveys. He is also often expected to assist in processing research or survey data and advice on the interpretation of the results of the analysis. He may or may not be a co-author in the resulting publications depending on the category the other author(s) is in the categories above.

For the sake of this audience however, I will like to present three ways in which the client views the statistical consultant.

1. **“Magician”** - Many researchers expect a certain pattern of results from their experiments. If the data collected does not follow the pattern, they

will expect the statistician to find a way of obtaining the expected pattern of results. He is to bring order to chaos and clarity to confusion. They believe that if you “torture the data enough, it will confess”.

Quite often the reason for not getting the expected result lies in wrong method of data collection. One wonders what results would have been obtained if the error in the data collected in my story on Counting Tomatoes had gone unnoticed! Could this be the situation with our census figures-both past and present?

2. **“Fire Brigade Officer”** – A researcher who had spent between six to twelve months in collecting data would come with a load of logbooks. His request can be summarized by “There in this mountain of data are *mine* precious jewels randomly deposited. I want them dug out. The job must be finished within 24 hours to meet the deadline”. “Is it not going to be done by the Computer?” he queries. The researcher may never have discussed the research with the statistician before and does not think the statistician can understand the details of the experiment. Even if he can, the researcher does not have time to explain to him.
3. **“The Technician”** – The statistician is often used as a data analyst, programmer, report writer, referee for papers etc. He is consulted only when a referee has faulted the design of an experiment or the analysis of the data contained in a paper submitted for publication. *After the “technician” has dug out the diamond, the researcher sells it and pockets all the cash. If only rusty iron is found, the statistician is the culprit.*

From my experience in consulting with students, particularly postgraduate research and final year undergraduate students, I believe that service courses in statistics for students in non-statistics disciplines should be taught by statisticians with a strong interest in the students discipline, rather than specialist in the students’ discipline who have some knowledge of statistics. Students taught by non-statisticians are usually not aware of the principles behind the statistical methods they have been taught.

Departments of statistics in our universities should have a unit which provides a service of statistical advice to staff and students undergoing research. The implication of this is that the staff strength in these departments will have to increase.

One advantage of this is that contacts arising from the service will improve the understanding of statistics among students and staff in the university. It will also expose the statisticians to a wider variety of examples that may be used to illustrate statistical principles in the various courses being taught by them.

DESIGN AND ANALYSIS OF MIXED CROPPING (INTERCROPPING) EXPERIMENTS

Mixed cropping is defined as growing two or more crops simultaneously intermingled on the same land (Kassam & Andrews, 1975). Statistically, mixed cropping can be considered to include intercropping, mixed cropping and relay cropping since the statistical problems are the same.

Before 1970, most agricultural research in the developing tropics was directed towards increasing production under sole cropping (a predominantly temperate system), instead of asking how to increase production under mixed cropping (which is the dominant system of tropical subsistence farmers). No wonder, farmers were not adopting technologies developed by researchers.

An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate or wrong problem.

In 1976 when I started work in an Agricultural Research Institute, the argument among agricultural researchers was whether intercropping (planting two or more crops simultaneously on the same plot of land) should be encouraged among farmers or whether research should continue on sole cropping. About this time, a survey by researchers at the Institute for Agricultural Research, Samaru, showed that less than 17% of the cropping systems consisted of sole crops. Although the advocates of sole cropping believed that intercropping could be more beneficial, especially to the peasant farmers, researchers were however uncertain about the types of designs and analyses that could be suitable for mixed cropping experiments. There was also uncertainty about the validity of methods of analyzing mixed cropping data and of comparing sole and mixed cropping systems.

Researchers found it difficult to publish results from intercropping experiments. Designs for intercropping experiments and analysis of data from such experiments became an obvious area of research for me.

Some reasons that made mixed cropping attractive to farmers include:

- Economy of labour – possible reductions in amount of input compared to output.
- Need to intensify production per unit area on the existing and potential cultivable area
- Need to grow two types of crops- e.g. cash crops for monetary values and food crops for family consumption.

- Increase in utilization of environmental factors and reduction of adverse factors. Crop mixtures have been found to be less likely to succumb to adversities. Crops are not equally affected by adverse conditions. Fig 1 shows the probability of failure of the different cropping systems to attain a certain percentage (disaster level) of the expected yield in a survey of different environments over a number of years.

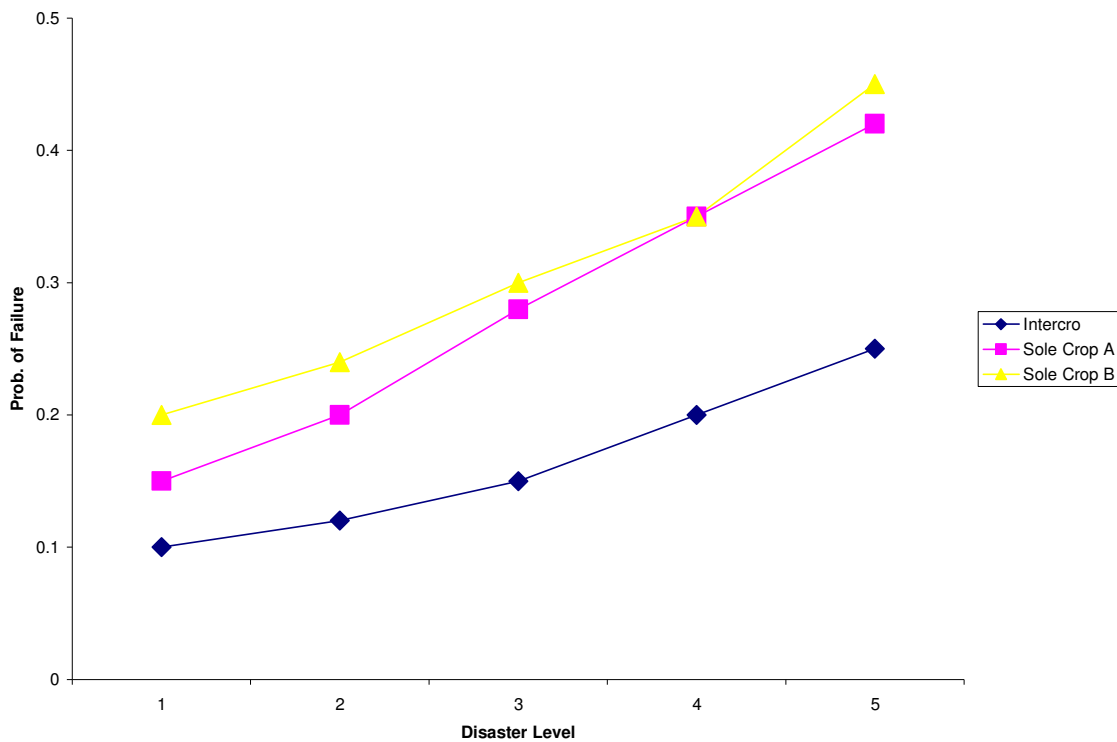


Fig. 1: Relative Risk in intercropping

In 1978, I carried out a review of some mixed cropping experiments focusing particularly on the validity of results from such experiments and suggesting alternative methods of design and analysis (Oyejola, 1978). The main findings are:

- Experiments should be designed so as to provide a good estimate of sole crop yields, to use in standardizing the intercrop yields.
- Many of the designs (particularly factorial and systematic designs) developed for sole cropping systems could easily be adapted for use in intercropping experiments.

- Incomplete blocks are useful. They can accommodate the large number of treatments that may be involved in intercropping experiments.

The Land Equivalent Ratio (LER) is an index found to be very useful by agronomists in assessing yields from mixed cropping. LER is defined as *the ratio of the area needed under sole cropping to one unit area of intercropping at the same management level to give an equal amount of produce*. Alternatively, it can be defined as *the relative area of sole crops required to produce the component yields obtained in intercropping*.

For example if 1 ha of intercrop of maize and cowpea produces 50kg of maize and 30 kg of cowpea, how much land will be required to produce 50kg of maize and 30kg of cowpea under sole cropping systems.

This can be expressed as:

$$\begin{aligned} \text{LER} &= M_A/S_A + M_B/S_B \\ &= L_A + L_B \end{aligned} \quad (1)$$

where M_A and M_B are yields of crops A and B from mixtures and S_A and S_B are corresponding yields from sole crops. L_A and L_B are the component LERs.

A value of LER greater than 1 indicates an advantage for intercropping.

Example 1

Consider the following results from 2 intercrops in a sorghum/cowpea experiment.

	Sole Crop Yield/(kg/ha)	Intercrop 1 Yield/(kg/ha)	LER	Intercrop 2 Yield (kg/ha)	LER
Sorghum	4313	3028	0.702	3886	0.901
Cowpea	2039	1185	0.581	579	0.284
Total			1.283		1.185

Both intercrops have yield advantages over the sole cropping system.

Univariate Analysis

Oyejola & Mead (1982) identified six choices of sole crop yields (standardization methods) in (1) as follows:

- Mean of all treatments from each block
- Mean of each treatment from each block
- Mean of the best treatment from each block
- Mean of all treatments from all blocks
- Mean of each treatment from all blocks

- Mean of the best treatment from all blocks

The main results from comparisons of these methods and the appropriateness of using the Analysis of Variance technique to analyze LER values and similar indices show:

- No statistical (and indeed no practical) advantages in the use of values from different blocks for standardization. (Oyejola & Mead, 1982)
- Calculation of LERs for comparative purposes should use a single sole crop yield for each crop and that the same sole crop yield should be used for all blocks. (Oyejola & Mead, 1982)
- The coefficient of variation (CV) of the divisors (S_A and S_B) should be less than 15% but preferably less than 10%. This implies that the values to be used as standards must be estimated as precisely as possible. This high precision can be obtained from good management of experiments and/or by increasing the number of observations or plots for sole crops (Oyejola & Mead, 1989)
- The distribution of the sum of two ratios of normal variables is at least not more non-normal than the distribution of the component ratios. (Oyejola, 1989)
- The distribution of the sum of two ratios is approximately normal provided the CV of the divisors are not more than 15% and can therefore be used in analysis of variance procedure. (Oyejola, 1989).
- Standard errors to be used for comparing intercrops through LER require modification. The modification was given by Oyejola and Mead (1989).
- In experiments where the effectiveness of treatments such as herbicides or insecticides) are to be compared, controls are often included to provide a standard with which to assess the effectiveness of the treatments. If the results are presented in form of percentages of the control mean, the standard error, obtained from the analysis of variance of the percentages require adjustment. The form of adjustment is given in Oyejola & Mead (1989). The adjustment applies to all measurements obtained as ratios.
- The Competitive Ratio Index is defined as $\lambda = (L_A/L_B)(P_A/P_B)$ where P_A and P_B are the proportions of intercropped area allocated to the two crops A and B. Oyejola & Mead (1989) proposed an alternative index defined as $\lambda' = (L_A/L)(1/P_B)$ whose distributional properties are more acceptable for statistical tests. The interpretations of the two indices are similar.

Bivariate Analysis

Since at least two crops are usually involved in intercropping experiments, analysis of data from such experiments should be at least bivariate.

A graphical (Bivariate-LER) method was developed for assessing the yield advantages when the LER is used (Oyejola, 1993).

The method is illustrated in Fig.2. A treatment whose circle of Confidence Region is above the line joining the points for the two sole crops gives significant yield advantage.

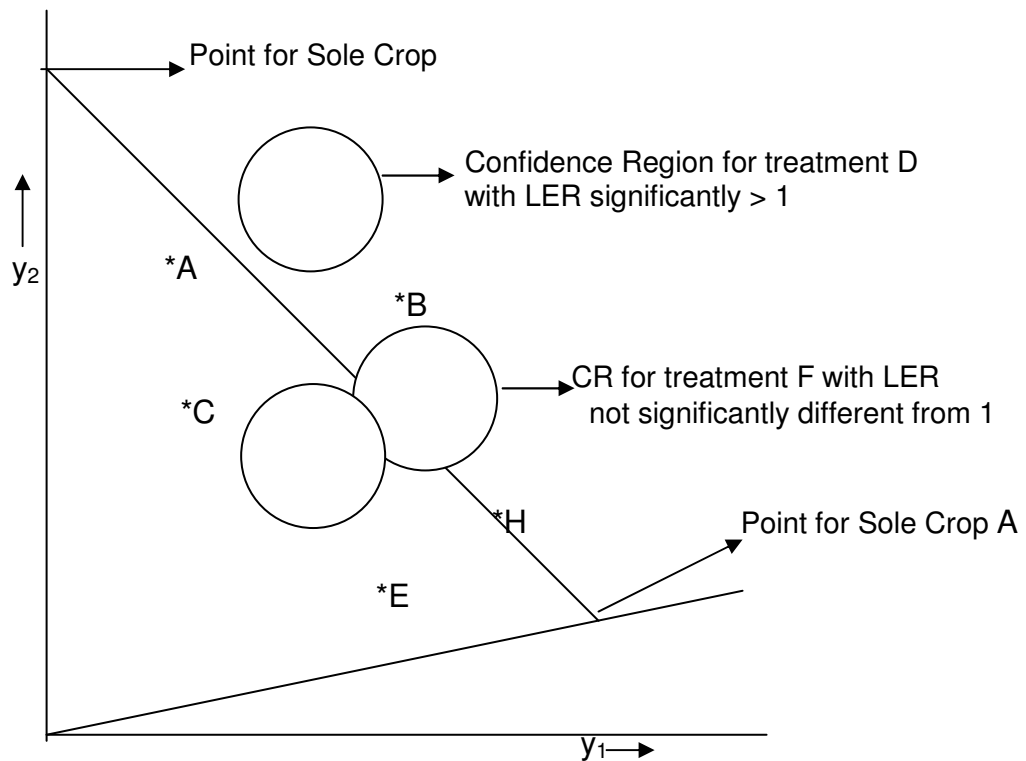


Fig. 2 Graphical assessment of yield advantages in intercropping

Treatment D, whose circle of confidence region lies entirely above the line shows that this intercrop treatment has yield advantage over the sole cropping system. An overlap of confidence regions of two intercrop treatments indicate lack of significant differences in their yield advantages

It could be misleading to conclude that intercrop 1 in Example 1 is better than intercrop 2 based on the total LER values. Assessment of yield advantage

through the use of LER assumes that the proportions of the component yields observed are those required by the farmer. Mead and Wiley (1980) suggested the concept of Effective LER (ELER) and Staple LER (SLER) for comparing yield advantages.

Oyejola (1990, 1994) obtained confidence intervals which are modifications of the graphical method described above for comparing yield advantages using ELER and SLER concepts.

ON-FARM RESEARCH

Results from research stations have consistently been found to differ from those of farmers. On-Farm experiments or trials as a means of exposing new technologies to real farm situations are therefore increasingly receiving attention at the different stages of the agricultural research cycle. On-station experiments concentrate more on basic research while On-farm experiments concentrate on applied research.

An on-farm participatory trial is that trial where the farmer, who is the beneficiary of a proposed technology, is an active participant in the development of the technology. The challenge in such experiments for the statistician is that higher variability is introduced by the involvement of the farmer and his farm conditions. The statistical approach to design and analysis of on-farm trials is challenged by practical considerations due to limited farm resources and their allocation.

Information is needed about all farms and this is a dynamic population. It is impossible to do research in all farms. There is need to take a few farms (sample) to represent all farms. Statistical problems in on-farm include:

- ❖ Mapping out zones and subzones and stratifying environments and regions and selection of farms/farmers typical of each subdivision. This is done prior to field trials although the zoning may be modified in the course of research when more information is obtained. Similar farmers' environments require similar technologies to solve similar problems.
- ❖ Selection of technologies appropriate for different ecological zones.
- ❖ It may also not be possible to use the same farms or farmers for several years. Grouping of farms and farmers allows similar farms/farmers to be chosen to replace existing ones.
- ❖ Deciding the optimum number of farmers or locations to be selected in each subdivision or checking if the number of subdivisions is optimal.

- ❖ Extrapolation of results from selected farms (sample) to all farms (population) – Analysis of on-farm data.

Given relevant data, the Additive Main effects and Multiplicative Interaction (AMMI) procedure was found to be effective in grouping locations. (Oyejola and Danbaba, 1999). The procedure produces a map of locations grouped according to their similarities based on relevant information. For example, in variety trials, relevant information would include soil characteristics, weather parameters and so on.

Fig. 3 shows the result of the AMMI procedure adapted to group locations for a multilocational trial.

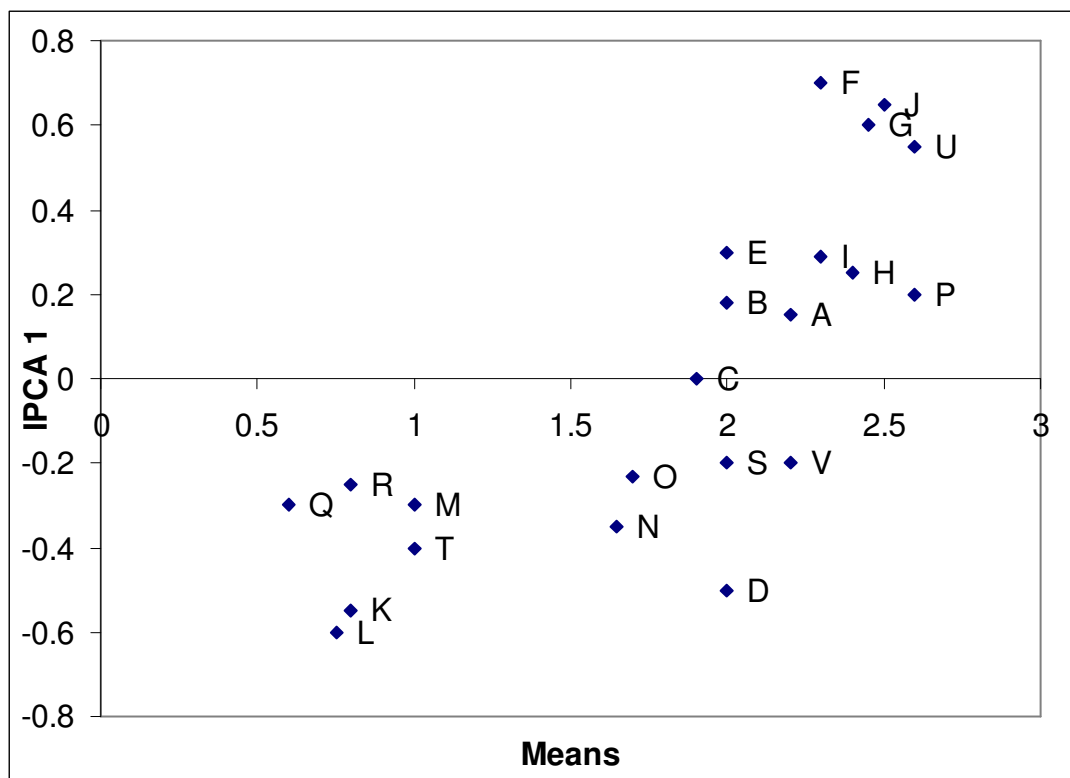


Fig. 3: AMMI BIPLLOT for Grouping Locations

The approach adapted here can be extended to grouping of sites (towns or Local Government areas or even states) from which representative sites may be selected for pilot projects).

In Fig. 4, the procedure is used to assess adaptability of 16 varieties to 5 ecological zones (A, B, C, D and E). See also Ajibade *et al* (2003).

Varieties whose points are close to a point for an environment are more suitable for that environment. For example varieties 1, 12, 13 and 16 are suitable for environment A.

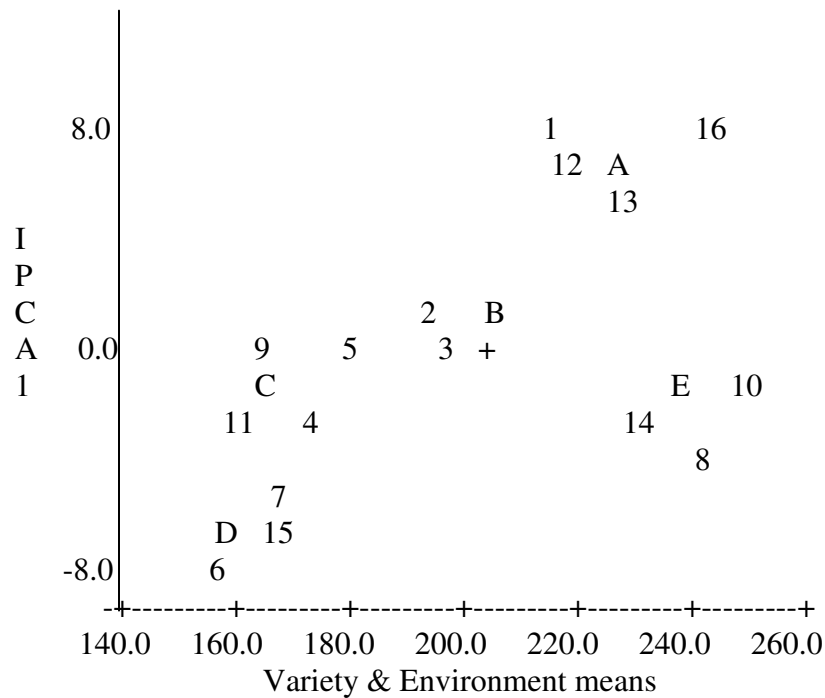


Fig. 4 Adaptation of Varieties to Different Environments

Oyejola *et al* (1998) suggested and compared three methods that can be used to estimate sample sizes in zones, sub-zones and farming environments. The method of variance components was found to be most efficient. Sample sizes were found to depend on levels of variability in zones, subzones and farming environments as well as the magnitude of differences to be detected.

Other methods of analysis that are useful for data collected from on-farm trials include stability and risk analysis, correspondence analysis, regression modeling with covariates, Mixed model combined ANOVA with unstructured variance-covariance matrices, meta analysis and non-parametric analysis (Nokoe, 1999 ; Oyejola, 1999).

STATISTICS EDUCATION

A lecturer is expected to teach and do research. My involvement in teaching statistics started in 1983 after my Ph.D. I have since been involved in teaching statistics particularly to Statistics and Agriculture students. Rather than assess my contribution in statistics education, I would like to make comments on some deficiencies in the teaching of statistics to these two groups of students.

1. Statistics curricula for both B. Agric. and M.Sc. in various fields of Agriculture can best be described as conventional. Emphasis should be on statistical reasoning applied to real life problems. Concentration is on formal manipulation and mechanical application of statistical methods at the expense of training them on how to plan and manage their investigations. One reason for this is the short time (low number of credits) assigned to the courses. For some years, I introduced some experiment games to illustrate the basic principles of experimental design. This involved having extra sessions outside the allocated periods. Increase in class sizes made this difficult in later years.
2. Lack of computing facilities also limits how comprehensive the curriculum can be. Availability of computers will facilitate practical sessions. Students would not have to spend much time on calculations using hand calculators. Departments in Faculties of Agriculture have not seen the need to have computing facilities for use by their students.
3. Even in Departments of Statistics, computing facilities are at best very inadequate. Some university authorities do not even appreciate the fact that statistics students do laboratory (computing) work and therefore need both equipment and consumables. This situation makes statistical courses theoretical and does not give suitable or necessary practical training. Students are not exposed to realistic large data sets and find it difficult to appreciate the relevance of statistical principles to practical situations.
4. Statistics is currently being taught in secondary schools by mathematics teachers as part of mathematics curriculum. One problem with this is that there is shortage of mathematics teachers at this level. Those who are mathematically qualified typically have inadequate statistical background and no training in teaching statistics to teach statistics effectively and interestingly. Whatever statistics they must have learnt is of a theoretical nature. Quite often, the statistics component of the curriculum is left uncovered or at best the coverage is scanty. Students therefore come into tertiary institutions ill-prepared.
5. Statistics is in a sense more difficult than many other branches of mathematics. Statistics is not merely a collection of techniques but is a

practical subject devoted to obtaining and processing data with a view to making statements that extend beyond the data to the real situation generating them. Students fail to appreciate this and depend on memorizing the techniques.

CONCLUDING REMARKS

Mr. Vice Chancellor, Sir, let me conclude this lecture with some remarks:

- The value of statistics in enhancing the validity and usefulness of results of research cannot be over-emphasized. Inclusion of statisticians in research and development project teams where data collection and data analysis are involved should therefore be made mandatory. I hasten to add, however, that a statistician who is knowledgeable in the particular discipline should be involved.
- It is a waste of resources to use a statistician as a technician. He should be made a collaborator who is involved from the planning stage to the report dissemination stage.
- Statistical resources are available for grouping environments, selecting “sites” that are representative of each environment and the number of such “sites” for pilot studies and projects. Government and researchers should use such objective resources rather than politics unless such projects are politically oriented.
- A functional vital registration system should be established. Registration of vital events should then be made compulsory. This would provide some essential demographical data which can be used for planning purposes. It can also be used to augment census figures.
- A Statistical Advisory or Consultancy Centers should be established, preferably in Statistics Departments, to cater for statistical needs of researchers and students. This will enhance efficiency of research efforts. It will serve as training ground for statistics students. Relevant and useful research areas will also result from interaction sessions.
- There is need to **restore** statistics as a subject in secondary education. Virtually every discipline requires statistics. At this stage, studies of real life problems, collection of relevant data, drawing of sound conclusions from those data and presentation of results in simple ways should be emphasized. Statistical ideas needed in other disciplines will also be provided. This will also require adequate statistical training for teachers at that level.

Nigeria cannot afford to continue in the 21st century without having a critical mass of its youths grounded in the art of statistical reasoning. This habit can only be cultivated through an early exposure to the studying of statistics at the secondary school level.

ACKNOWLEDGEMENTS

I give all the glory, honour and majesty to the Almighty, Unchanging God for making me what I am today. I thank Him for His blessings and protection. He is my joy, peace and salvation. It is by His grace that I am able to give an inaugural lecture today. He has used various people to achieve His will in my life and I will like to acknowledge some of them.

- I wish to appreciate the love and support of my parents, Late Rev. James Oyejola and Mrs. Abigail Oyejola. Your prayers have been a source of encouragement throughout my career. You brought me up in the way of the Lord. You gave all you could to see that I succeed. You trusted me so much and I thank God that I have not disappointed you.
- To my brothers and sisters, Wale, Bolanle, Biola, Adeoye and Nike, I appreciate your love and respect. You have always been there for me.
- To my uncle, Mr. J.A. Oyedepo, I thank God for your life. You have allowed God to use you tremendously in my life. The Lord will reward you abundantly.
- To my extended family, thank you for your support and encouragement.
- I thank all my teachers in primary and secondary schools. I wish to particularly mention Chief Adeseke. I still remember your words of encouragement, at that tender age, that I would reach the peak of the academic career. Here I am today by the grace of God.
- To my supervisors, Dr. Roger Stern and Prof. Roger Mead, I appreciate you. Your special ways of reasoning in the realms of uncertainty has gone a long way to bring me to this stage today.
- To my students, you have sharpened my understanding of my statistics as I try to impart same to you. Many of them are now professional colleagues. Thank you for being there.
- To those who have brought me in as collaborators in their research work and those who have joined me as collaborators in mine, I say a big thank you. The collaboration has enriched my understanding of some basic issues and problems in those disciplines particularly in Agriculture and lately in Medicine. Many of the publications resulting from such collaborations are listed in the references as a sign of my appreciation.
- To my colleagues, (both teaching and non-teaching, past and present) particularly in the unique Department of Statistics of this University, you have made the office a place I always long to be. You have created a wonderful environment for reasoning in the realms of uncertainty.

- To my fellow statisticians, particularly members of the International Biometric Society and Nigerian Statistical Association, I thank you for the many opportunities you have made available for me to serve the Association particularly as the Editor-In-Chief of JNSA for many years.
- Members of the UMCA Chapel, Tanke, Chapel of Redemption, Gaa-Akanbi and Calvary Baptist Church, Samaru, Zaria, you have provided a wonderful place of fellowship and worship. I have enjoyed your support and love. Reverends Gbenga Odebiri and Simeon Oladimeji, I thank God for you for spiritual leadership.
- I have enjoyed the friendship of many people and I will not attempt to mention their names here. You have been wonderful people.
- I however have to mention my special friends, Prof. J. A. Gbadeyan and “my twin brother”, Prof. Teju Jolayemi. It has been a relationship closer than that of a friend. I appreciate all that we have been to each other.
- God has blessed me with three lovely and God-fearing children. Tayo, Titi and Tolu (the 3-T’s), I appreciate your love, support and understanding. You make home so lively and pleasant that I always look forward to returning home whenever I am away.
- I wish to acknowledge the invaluable companionship of my wife, Grace Olufunmilayo Oyejola who passed on to be with the Lord on 24th May, 2005. You toiled with me for over 25 years and I wish you were here today to share in the joy of this day with me. I am however comforted in the certainty that the joy of where you are far exceeds this earthly joy.
- Finally, I thank every one of you here today who have honoured the invitation to attend this lecture. I appreciate your love and support. For those who have come from afar, I wish you journey mercies as you travel back to your destinations.

Mr. Vice Chancellor, Sir, distinguished ladies and gentlemen, let me end this lecture by saying:

IF IN DOUBT CONSULT A STATISTICIAN – HE MAKES VALID DECISIONS IN THE REALM OF UNCERTAINTY

Thank you and God bless you richly.

REFERENCES

Adegboye, O.S. (1997). The Magicians, the Prophets and the Statisticians. 50th Inaugural Lecture of the University of Ilorin.

Aderibigbe, A., Ologe, F.E. and **Oyejola, B.A.** (2005). Hearing Thresholds in Sickle Cell Anemia Patients: Emerging New Trends? Journal of the Medical Association. Vol. 97, No. 8: 1135-1142.

Aganga, A.A., **Oyejola, B.A.**, Aganga, A.O. & Yaakugh, I.D.I. (1986). "Reproductive Performance of White Fulani Cows". Thai J. Agric. Sci. 19: 225-229

Ajibade, S.R., Ogunbodede, B.A. and **Oyejola, B.A.** (2003). AMMI Analysis of Genotype x Environment Interaction of Open Pollinated Maize Varieties Evaluated in the Major Agro-ecologies of Nigeria.

Akinola, J.O. and **Oyejola, B.A.** (1994). "Planting Date and Density effects on six pigeonpea (*Cajanus cajan*) cultivars at three Nigerian Savannah Locations". Jour. of Agric. Sci., Camb. 123:233-246.

Buvanendran, U., Adu, I.F & **Oyejola, B.A.** (1981). "Breed and Environmental Effects on Lamb Production in Nigeria". J. Agric. Sci. Camb. 96: 9-15.

Buvanendran,U., Olayiwole,M.B., Piotroska, K.I. & **Oyejola, B.A.** (1981). "A comparison of Milk Production Traits in Friesian X Fulani cross Breed Cattle". Animal Prod. 32,165-170.

Elemo, K.A. & **Oyejola, B.A.** (1991). "Performance of Cowpea (*Vigna Unguiculata* (l) Walp) Cultivars in Maize/Cowpea Mixture". Appropriate Agricultural Technologies for Resource - Poor Farmers. Proceedings of the National Farming Systems Research Network Workshop held in Calabar, Nigeria. August 14-16, 1990. Edited by Olukosi, J.O., Ogungbile, A.O., & Kalu, B.A. Pgs 73-82.

Erinle, I.D., Quinn, J.G. & **Oyejola, B.A.** (1986). "Effect of a Fungicide Spray Programme on Performance of Tomato Cultivars in Nigeria". Tropical Pest Management. 32(2) 111-114.

Fielding, W.J., Riley, J and **Oyejola, B.A.** (1998). Ranks are Statistics: Some Advice for their Interpretation. PLA Notes . 33: 35-39.

Johnson, A.O., Buvanendran, V., & **Oyejola, B.A.** (1984). "Dairy Potential of Bunaji and Bokoloji Breeds". Tropical Agric. (Trinidad) 61: 267-268.

Kassam A.H and Andrews D.J. (1975) Importance of Multiple Cropping in Increasing World Food Supplies. Report, I.A.R., Samaru – Nigeria.

Mead, R. and Wiley, R.W. (1980). The Concept of a Land Equivalent Ratio and Advantage in Yield from Intercrop. *Expl. Agric.* 16, pp217-228.

Nokoe, S. (1999). On-farm trials: Surgical or Preventive Approach. *Journal of Tropical Forest Resources*. Vol 15: No 2: 76-83

Nwasike C.C., Thakare, R.B., **Oyejola, B.A.**, and Okiror, S.S. (1983). "Genotypic and Phenotypic Variances and Covariances in Pearl Millet". *Z. Pflanzensuchtg* 90: 259-264.

Nwasike, C.C., & **Oyejola, B.A.** (1989). "Combining Ability for yield". *Samaru J. Agric. Res.* 6:3-7

Ogunlela, V.B & **Oyejola, B.A.** (1991). "Comparative Responses of Photoperiod Sensitive and Insensitive Sorghums to Delayed Sowing in a Semi-Arid Tropical Environment". *Jour. of Agric. in the Tropics and Sub-tropics.* 62: 103-114.

Okoro, E.O., Jolayemi, E.T., and **Oyejola, B.A.** (2001). Observations on the use of low dose Hydrochlorothizode in the Treatment of Hypertension in Nigerians with Diabetics. *Heart/Drug*, 1: 83-88

Okoro, E.O., **Oyejola, B.A.**, and Jolayemi, E.T. (2002). Pattern of salt taste perception and blood pressure in normotensive offspring of hypertensive and diabetic patients. *Annals of Saudi Medicine*, 22(3/4), 249-251.

Okoro, E.O., Adejumo, A.O., and **Oyejola, B.A.** (2002) Diabetic Care in Nigeria; Report of a Self-Audit. *J. Diabetes & its Complications*. Vol. 16 pg. 159-164.

Okoro, E.O. and **Oyejola, B.A.** (2004). Inadequate Control of Blood Pressure in Nigerians with Diabetes. *Ethnicity & Disease*. 14, 82-85.

Okoro, E.O. and **Oyejola, B.A.** (2005). Long Term Effect of hydrochlorothiazide on Diabetic Control and Blood Pressure in Nigerians. *Kuwait Medical Journal*. 37(1), 18-21.

Ologe, F.E., Okoro, E.O. and **Oyejola, B.A.** (2004). Type 2 diabetes and hearing loss in black Africans. *Diabetic Medicine*. (In Press).

Ologe, F.E., Okoro, E.O. and **Oyejola, B.A.** (2005). Hearing function in Nigerian Children with a family history of type 2 diabetes. *Int. J. Paed. Otorhinolayngol.* 69(3): 387-91.

Ologe, F.E., Okoro, E.O. and **Oyejola, B.A.** (2006). Environmental Noise Levels in Nigeria: A Report. Journal of Occupational and Environmental Hygiene, 3: D19-D21.

Onukogu, I.B., **Oyejola, B.A.**, Ipinoyi, R.A. and Chigbu, P.E. (2002) Super Convergent Line Series in Optimal Design of Experiments and Mathematical Programming. AP Express Publishers.

Oyejola, B.A. (1978). On Design and Analysis of Mixed Cropping Experiments. Unp. M.Sc. Dissertation, Univ. of Reading, U.K.

Oyejola, B.A. & Mead, R. (1982) Statistical Assessment of Different Ways of Calculating Land Equivalent Ratios. Expl. Agric. 18:125-128.

Oyejola, B.A. (1988). Statistical Consultancy in Agricultural Research Institutes. Paper presented at the 12th Annual Conference of the Nigerian Statistical Association held at Abeokuta between 25th and 28th October, 1988.

Oyejola, B.A. (1989a). "Transformation of the Land Equivalent Ratios in Analysis of Variance Tests". Samaru J. Agric. Res. 6: 53-57.

Oyejola, B.A. and Mead, R. (1989). "On the Standard Errors and other Moments for Ratios of Biological Measurements". Expl. Agric. 25: 473-484.

Oyejola, B.A. (1989b). "The Distribution of the sum of Two Ratios of Normally Distributed variables". J.N.S.A. Vol 5 No. 1

Oyejola, B.A. (1990). "Confidence Intervals for staple and Effective LER values". Expl. Agric. 26: 213-220

Oyejola, B.A. (1993). "The Use of the Bivariate LER Method in the Analysis of Intercropping Data". Nig. Jour of Sci. 27: 231-237.

Oyejola, B.A. (1994). "Confidence Intervals for the Staple LER values from Intercrop Supplementation". Nig. Jour. of Sci. 28: 229-234.

Oyejola, B.A. (1995). "On the Moments and the Distribution of the Ratio of Two Normally (Restricted) Distributed Variables". ABACUS.

Oyejola, B.A. (1997). Impact of Computer on Statistical Training and Data Analysis. Paper presented at the 21st Annual Conference of the Nigerian Statistical Association held at Calabar between September 22 and 26, 1997.

Oyejola, B.A. and Jolayemi, E. Teju (1997). "A Comparison of Some Models for Studying Relationships of the Sigmoid Form". Nig. Jour. of Sci. 31:193-198.

Oyejola, B.A., Riley, J. and Bolton, S. (1998) A study of Alley-cropping data from Northern Brazil: II A Comparison of Methods to Estimate Sample Size. *Agroforestry Systems* 41: 167-179.

Oyejola, B.A. and Danbaba, A. (1999). Selection of Locations for Multilocational On-farm Trials. Paper presented at the 6th Scientific meeting of the Sub-Saharan African Network of the International Biometric Society, 23rd - 27th August, 1999.

Oyejola, B.A. (1999). An over-view of On-farm Participatory Trials. *Journal of Tropical Forest Resources*. Vol 15: No 2: 76-83

Oyejola, B.A. (2002). Data Analysis Using SPSS. Manual for the 2002 Pre-Conference Workshop of the Nigerian Statistical Association.

Oyejola, B.A. (2003) Design and Analysis of Experiments for Biology and Agriculture Students. Olad Publishers.

Oyejola, B.A. (2003). Linear Modeling Using SPSS. Manual for the 2003 Pre-Conference Workshop of the Nigerian Statistical Association.

Oyejola, B.A. (2004). Multivariate Data Analysis. Manual for the 2004 Pre-Conference Workshop of the Nigerian Statistical Association.

Oyejola, B.A. and Adebayo S.B. (2004). Basic Statistics for Biology and Agriculture Students. Olad Publishers.

Oyejola, B.A. (2005). Applications of SPSS. Manual for the 2005 Pre-Conference Workshop of the Nigerian Statistical Association.

Oyejola, B.A. (2005). Statistical Methods for Breeding Trials. Module 3 of Short Course for Graduate Research Fellows & Associates at IITA, Cotonou.

Sekoni, V.O., Njoku, C., Saror, D., Sanusi, A., **Oyejola, B.A.** (1990). "Effect of Chemotherapy on Elevated Ejaculation Time and Deteriorated Semen Characteristics Consequent to Bovine Trypanosomiasis". *Br. Vet. Journal* 146: 368-373.