PASSENGER DISTRIBUTION FUNCTIONS FOR SMALL AIRPORTS

by

ERICA GEDDES

B.Sc., Queen's University, 1979

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

FACULTY OF APPLIED SCIENCE

DEPARTMENT OF CIVIL ENGINEERING

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1984

Department of Civil Engineering

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date April 27/1989

ABSTRACT

This research looks at one input required for the design and planning of small airports. It investigates the number of passengers expected to use the terminal.

Data describing passenger volumes was gathered from airline flight records at eight airports in British Columbia. The volumes were formed into frequency distributions and a theoretical model was found that would best describe the data. The selection of the model was based on the overall fit of the curve (as measured by the Chi-Squared statistic and by visual inspection) and the ability of the model to predict the right hand tail of the observed curve (as measured by the 90th percentile values).

Three model distributions were studied: the Normal, the Poisson and the Lognormal. According to the selection criteria, the lognormal distribution was found to be the best model for use in air terminal design.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I wish to express my thanks to Professor G.R. Brown for his guidance throughout this thesis.

I would also like to thank the Terminal Services staff of Transport Canada's Pacific Region Airports Branch who made the study possible.

1. INTRODUCTION

1.1 Statement of the Problem

Expenditure for the design and construction of airport terminals is considerable. Even though regional airports are not as large as the international and national airports they connect to, the amount of capital and time involved can still be significant.

For example, the expansion of the air terminal complex at Castlegar, British Columbia is expected to cost approximately $6 million. Of this, $2.5 million will be spent to enlarge and renovate the terminal building with the remaining $3.5 million going to parking lot reconstruction, relocation of services and design fees. The planning of the project began in 1981 and completion is expected to be in 1987. The fact that the design and construction will take six years illustrates the magnitude of the effort involved.

The purpose of this study was to improve the input to the analytical processes of air terminal design. The particular input looked at was the number of passengers expected to occupy small terminals. Airline records of passengers enplaning and deplaning for each flight were used to determine the number of passengers expected.

The study looked at the frequency distribution curves of the flight volumes. Knowledge of the shape of these distributions will help the terminal design process. For example, a peak volume (such as the 90th percentile) can be calculated and used as a design criterion. Alternatively, the full distribution can be used for simulation models which randomly sample from the expected values. With this more accurate representation of passenger occupancies, the terminal design will be more efficient.

1.2 Approach

Records of the number of passengers getting on and off of aircraft were collected from small regional airports. They were compiled into frequency distributions. A common statistical distribution model was then found which would adequately describe the actual data so that it could be used for the design of terminals.

When passenger volume data is used for terminal design, it is typically in one of the following forms:

(1) design hour volume of passengers;

(2) design flight load;

(3) distribution of expected passengers;

(4) distribution of expected flight loads;

(5) design daily pattern of passenger volumes, or

(6) design daily flight schedule.

If the passenger volumes (and flight load volumes) are described by a model distribution, the values to be used for the terminal design can be better determined.

In this work, the data used to determine expected passenger volumes at the terminal was flight load data. The individual observations are the number of deplaned and enplaned passengers of one flight stop. In other words, each data point is the sum of all of the passengers getting off of the airplane when it arrives at the airport, and all those boarding the airplane as it departs. These two movements will be designated as one "flight event". All airports in the study have one arrival and one departure association with each event - that is, the flight routes do not originate or terminate at these particular sites.

These airports have only a few major flight events daily, and for each, the arrival and departure occur within the space of a half-hour. For these reasons, the passenger volumes of a flight event are equivalent to half-hourly volumes. This simplifies the analysis since flight event volumes can be measured to directly determine design volumes for planning.

The flight events are grouped together into years, such that a "flight" will be defined as the total of all of the flight events that occur at the same time of the day over the course of one year. This means there will be 366 or less flight events in one flight. Since the volumes of passengers involved in each flight event vary over the year, each flight will have a certain distribution of the frequency of occurrence of the flight volumes.

To derive hourly planning volumes, however, all hours with activity must be compiled for the year. Therefore, as a second step, all events of all flights at an airport will be combined to form another frequency distribution.

This, then, will be the data under study - individual flights and flights compiled at each airport. Each distribution will be formed into a histogram so that it can be compared to theoretical statistical models.

Originally, nine possible distribution models were considered:

(1) Binomial

(2) Poisson

(3) Normal

(4) Gamma or Erlang

(5) Weibull

(6) Lognormal

(7) Negative Binomial

(8) 5th Degree Polynomial

(9) Beta

Of these, three were selected for further study: the Normal, Poisson and Lognormal. The three are relatively simple to understand, to calibrate, and to apply. They also appeared to reasonably represent the shape of the observed distributions. Table I shows some of the distributions used by airlines and aircraft manufacturers.

Both quantitative and qualitative methods were used to select the distribution which would best replicate the actual data. A computer performed the most of quantitative work by doing two things. First, the Chi-squared statistic was calculated for each distribution model and compared to the theoretical Chi-squared values. This comparison determined if the model provided a statistically significant fit.

The second application of the computer was to measure the ability of each model to accurately predict the behaviour of the upper tail (the right hand end) of the distribution. This is particularly useful in the determination of peak design volumes. Actual and predicted 90th percentile volumes were calculated to measure the tail behaviour.

The third criterion used to evaluate the three models was more subjective. It involved visually inspecting each observed and expected histogram and ranking each model according to its ability to reproduce the observed data.

Finally, the selection of the best model was based on its ability to be understood and to be applied.

TABLE I

Distribution Models Used by Airlines and Aircraft Manufacturers

Distribution	Users
Binomial	Quantas (business and 1st classes)
Poisson	----
Normal	United Airlines, Boeing, Lockheed, KLM, Quantas (economy class) Pan American, Air Canada
Gamma/Erlang	Swiss Air
Weibull	American (now switched to Rayleigh)
Lognormal	McDonnell-Douglas
Negative Binomial	British Airways
5th Degree Polynomial	Lufthansa
Beta	----
(empirical model)	Cathay Pacific

Source: References - Lauchli[14]; Vella, et al[22]; Wang[24]; Soumis et al[16]

2. LITERATURE REVIEW

An airport terminal is a transfer point between ground and air transportation systems. By most definitions, the air terminal includes the building structure, the roadway curb, the station platform if the airport is served by transit, and the aircraft apron. The flow between ground and air is shown schematically in Figure 1.

The purpose of the air terminal is to aid this transfer between ground and air and also, in the case of connecting passengers, between air and air. Although the system of pedestrian movement is complex, the transfer must be done as quickly, as comfortably and as efficiently as possible.

Planning an air terminal is a complicated and usually lengthy process. Careful design will be even more critical as capital funds are reduced and a premium is placed on the space available. A typical framework for the planning process is given in Figure 2. There is, at present, no universal procedure for the generation of terminal designs nor for the evaluation of proposed terminal concepts. This is not to say, of course, that methodologies do not exist. There are numerous ways to size facilities and to model the movement of pedestrians between them. These will be discussed below.

2.1 Air Terminal Sizing

Planning of airport terminals incorporates the sizing of their facilities and the arrangement of these facilities within a building structure. Some of these elements are mandatory stops for passengers; others are optional. Essential for processing are the ticketing and bag

FIGURE 1

Airport Terminal Passenger Flows

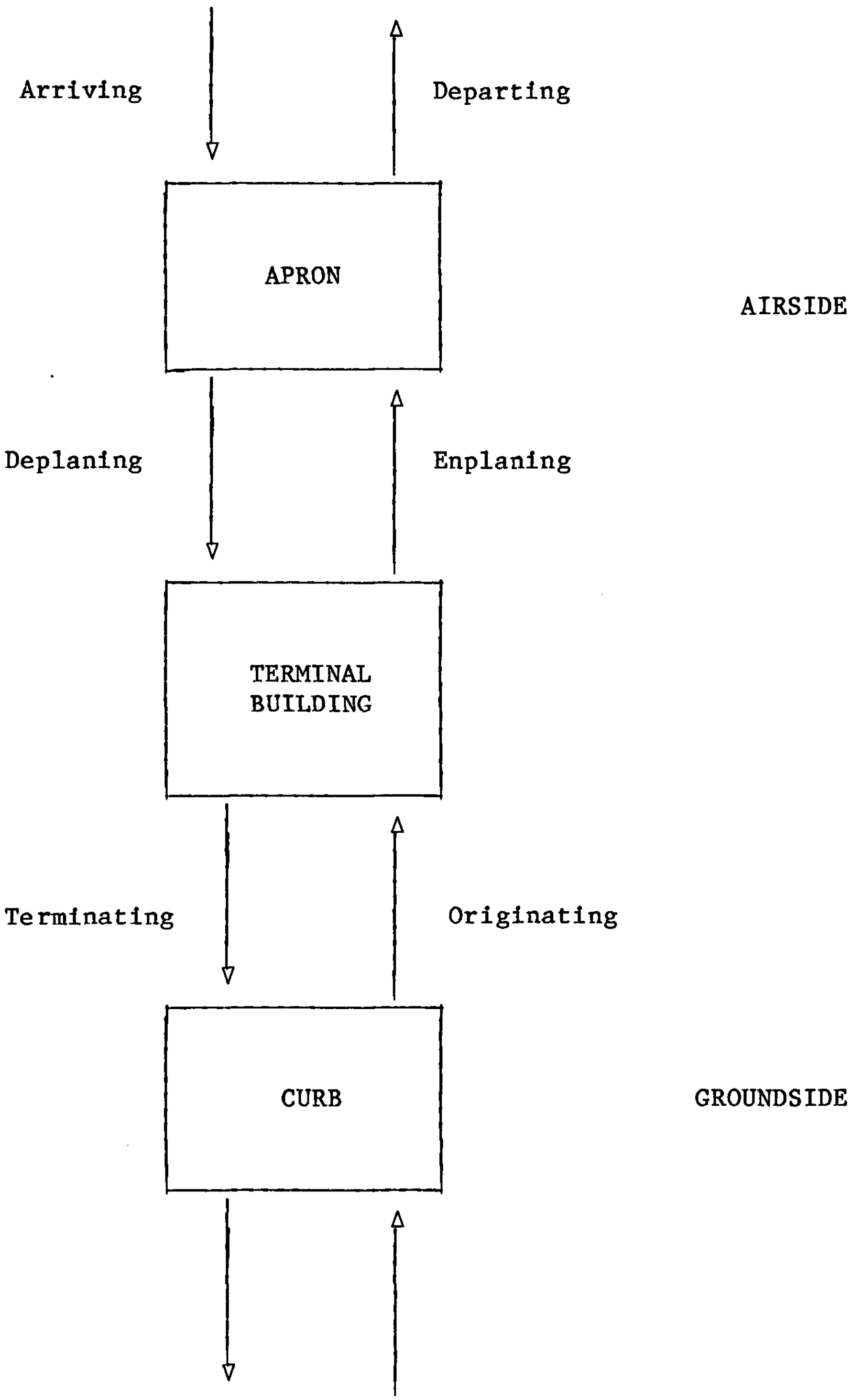
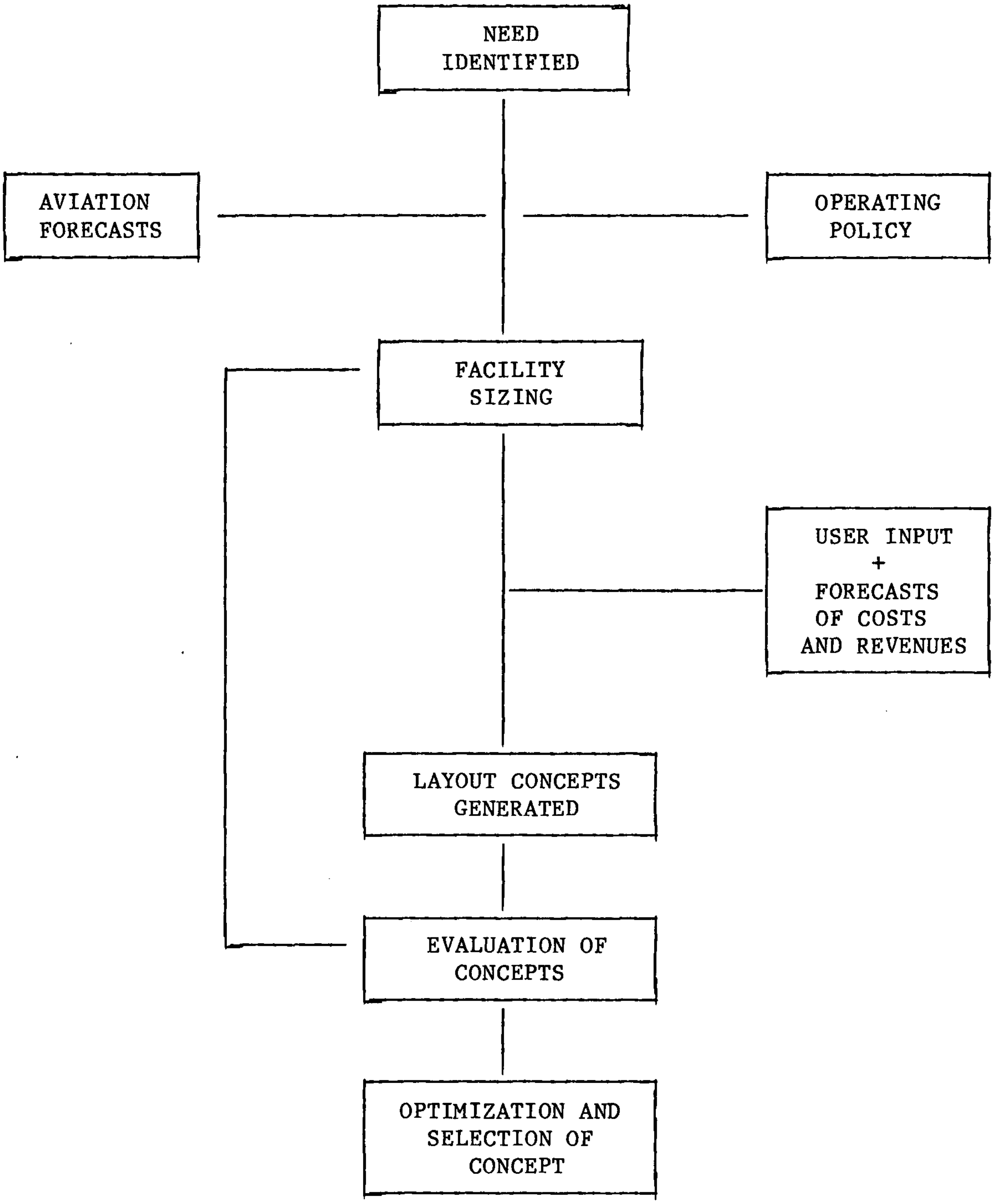


FIGURE 2

Terminal Planning Process



check-in counters, security checkpoints, holdrooms, gates and baggage claim devices. Occasionally, some of these may be bypassed if, for example, a passenger has no checked luggage or if ticketing is done on board the aircraft.

Optional components vary from airport to airport. Some examples are restaurants, washrooms, telephones, giftshops and banks. Space is also provided for the offices of airline and airport employees as well as for electrical and mechanical utilities.

Facilities can be further divided into those used by enplaners (ticketing, holdroom) and those used by deplaners (baggage claim). Figure 3 shows the basic passenger flow for a simple terminal layout.

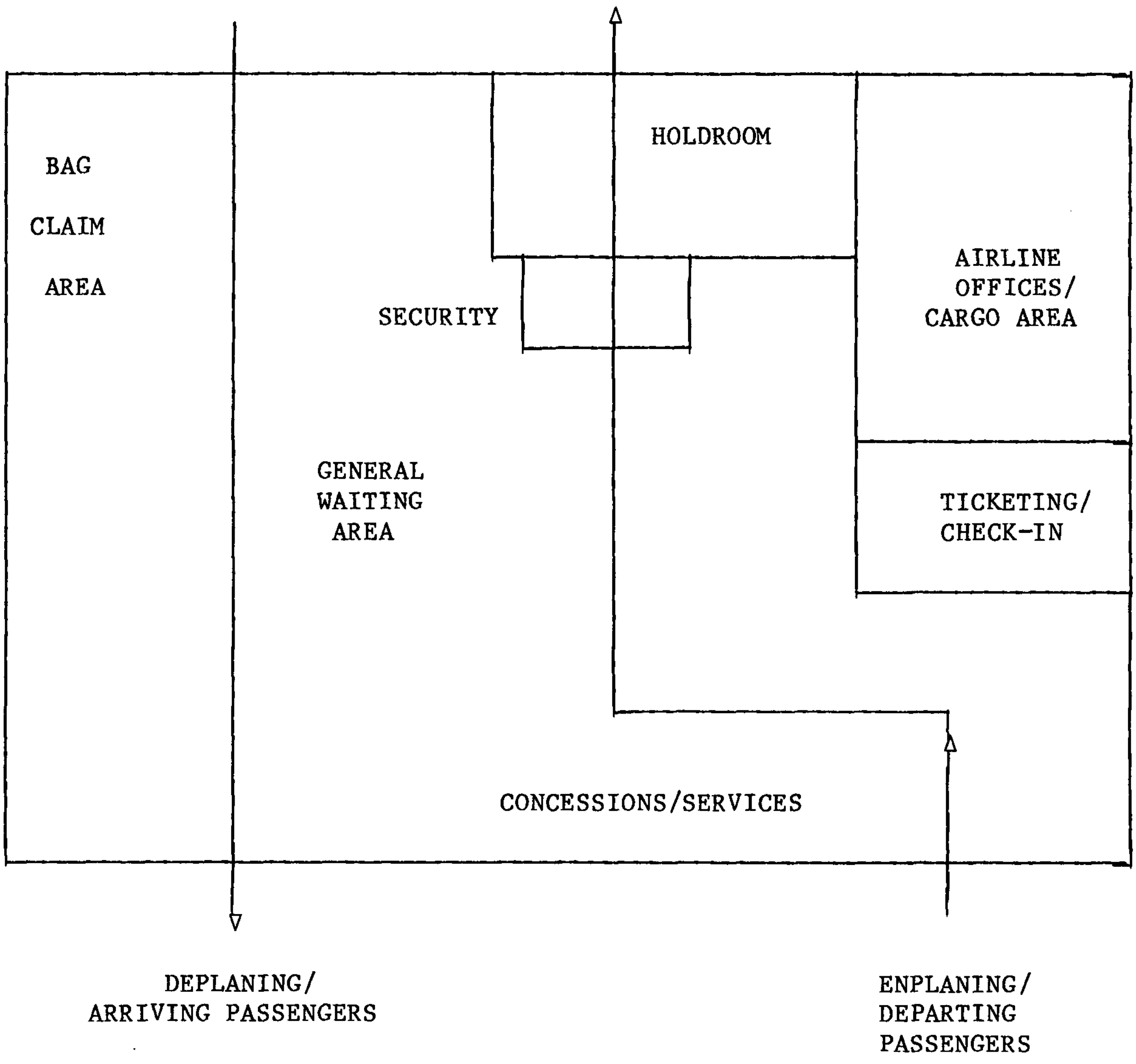
The function of terminal planners is to balance the demands of passengers, airline companies, government agencies, concessionaires and other airport users with the services to be supplied by the facilities. Obviously, the objectives of these parties will often conflict. There are, however, three tenets that are geneally accepted as being fundamental to good design - that the terminal be flexible, economic, and provide an acceptable level of service to the users. Some facets of each of these are:

(1) Flexibility to allow for:

- staged growth
- new technology
- unforeseen circumstances

FIGURE 3

Terminal Facilities



(2) Economic Optimization of:

- capital costs
- operations and maintenance costs
- revenues
- benefits to users (often intangible)

(3) Level of Service as Affected by:

- area per person
- waiting times
- walking distances (inside and outside; with and without bags)
- temperature and humidity
- lighting
- amenities (such as seating, no-smoking areas)
- concessions
- handling of disabled persons
- information systems

In most terminal design methods, the space required for each function is calculated on the basis of an expected peak occupancy. The occupants are each allotted a certain area, the amount of which is dependent upon the purpose of the area and upon some measure of personal comfort. The latter is quantified into discrete categories known as Levels of Service. For example, at a given Level of Service each person in the general waiting area may be given 1.5 square metres of space but the occupants of the holdroom would be deemed to need only 1.0 square metre. For a reduced Level of Service, these areas would be, say, 1.2 and 0.7 square metres respectively.

The design occupancy is either determined directly from a design volume or is calculated by a model of the terminal flows. In the former case, the number of passengers in the area is taken as a proportion of a design flight volume, or of the airport's peak hourly design volume. The proportion is based on historical patterns. If greeters and well-wishers are permitted in the facility, the ratio of non-passengers to passengers is multiplied by the number of passengers to find the total occupancy.

This method of fixed proportions was used extensively by Transport Canada until a few years ago. It is the simplest way of calculating facility space requirements, aside from using standard terminal layouts. For this reason it is still used, both in Canada and the United States, when more advanced tools are unavailable or for preliminary estimates.

There are difficulties, however, with the use of typical proportions of design volumes. De Neufville[8] explains this as being due to the method not incorporating the stochastic features of the movements through the terminal.

Models which incorporate pedestrian flows are better able to predict the dynamic nature. They can also point out critical areas of congestion (often the ticketing area and the bag claim area).

Flow models are typically used to evaluate proposed layouts. They do not generate layouts, which is a largely subjective process, although attempts have been made to quantify it. For example, Braaksma and Ramsey[5] developed two indices to catalogue terminal layouts. Braaksma also developed[3] a computerized method of creating preliminary layouts.

There are basically four categories of analytical methods for analyzing terminal flows (based on Horonjeff[12]):

(1) Network Models

Network models illustrate the airport functions and describe their interrelationship in the processing system. Once the processing times of each link and the passengers' paths through the network are known, the total trip time can be estimated. Analysis of the network can identify critical links that affect the entire system. Braaksma applied a CPM network model to evaluate passenger delay (Simulating the Turnaround Operation of Passenger Airfract using the Critical Path Method, University of Waterloo, 1970). This approach does not consider the volume of passengers travelling on any link or path. It does not assign passengers to paths, predict the effects of queue building nor model random behaviour.

(2) Queueing Models

Entrance and exit queueing models can be developed for each facility. Standard formulae can not be used because the demand is not steady, but builds up and dissipates with each flight. The facilities have to be analyzed in the order in which passengers go through them. For example, an analysis of the ticket counter could use a cumulative distribution curve of passenger arrival times and the average service rate of the ticketing agents. Both would be plotted. The queue length and waiting times would then be determined graphically from the differences in the curve. (Queueing models are well-explained by Horonjeff[12] and de Neufville[8]). Ashford and Wright[1] describe the difficulty with these models when several facilities are linked together in chains: the mathematics may become intactable when random arrivals and exponential service times are incorporated.

(3) Simulation Models

These are computer models which can provide very detailed information, but can be expensive to run. By definition their inner workings cannot be explained by simple equations and so can be difficult to validate. An example is the Vancouver Airport Simulation Model [21]. Simulation models operate on a projected flight schedule. Passenger arrival ratio, processing rates, walking speeds and passenger routes are also input the variables may be fixed, or the model may randomly select them from a distribution. By iteration, the movement of all persons throughout the day are found producing computations of delays, transit time, and occupancies.

(4) Hydraulic/Hydrologic Models

This is a relatively new type of model for terminals which assumes pedestrians behave in a manner similar to fluid flow. Ramsey and Hutchison[14] used a flood flow analogy and found it less expensive than the Vancouver Airport Simulation Model. Their model routes passengers through the system in the way in which a storm proceeds through the various reaches of a river. As input, a daily schedule is required which initiates the "storms" of passengers and determines the volume of passengers flowing through the terminal. Resistance characteristics of the processors and links as well as the desired level of service are also required model inputs.

There are obvious benefits to the use of these models. Once set-up, they can be repeated in order to evaluate proposed layouts and to examine their sensitivity to variations in input. They can model the entire system or only a part of it.

Like all analytical methods, terminal models are only as useful as the information input into them, which typically includes:

(1) characteristics of passengers, non-passengers and baggage;

(2) aircraft types and characteristics;

(3) activity levels of passengers and aircraft;

(4) rates of arrival, usually in relation to flight times;

(5) processing and flow rates; and

(6) the variations of all of the above factors over the course of the day.

This information may be difficult, if not impossible to obtain. Survey information such as found by the extensive surveys done in the Canadian Airports System Evaluations[4], can be used for model input. Alternatively, a survey can be used directly to determine facility occupancy but since a survey can only be done for a few days, the results may not be representative.

These are the usual methods of airport terminal sizing. As explained above, peak occupancies are determined from design volumes or from models. The areas are calculated by multiplying the number of occupants by a given unit area. Some iteration may be necessary since the size of an area will affect the travel times through it and, therefore, the flows.

In an effort to simplify design Transport Canada is now using standard layouts for all new terminals. The process is called the Systemized Terminal Expansion Program, or STEP[20]. The purpose of the program is to avoid repetition of the design process since the requirements of small terminals tend to be similar. It also speeds the selection, review, and approval processes, as well as the preparation of the contract drawings. Furthermore, by incorporating a pre-planned expansion capability, terminals are able to adjust to changing traffic conditions, which are often difficult to forecast at small sites.

The design year for STEP buildings is the year of opening, although the chosen size must suffice for three years. By minimizing the time to the design year, there is more certainty in forecasting the requirements. The lifespan of three years was chosen to balance the added cost of expansion with the savings made by delaying the construction.

Originally, the selection of a STEP terminal was based upon six criteria:

(1) Total Annual Passengers

This is an easily obtained statistic which gives a general indication of the airport size. However, it is too broad to be of use in facility sizing.

(2) Planning Volume

This hourly volume would more accurately reflect the demand made on the facilities. It is not yet officially defined for small airports, but the 90th percentile (of all hours with traffic) has been used. Complete data is difficult to collect, however, for small airports.

(3) Critical Aircraft

The largest scheduled aircraft also gives a reasonable idea of demand on the terminal. (In British Columbia, the critical aircraft is usually the Boeing 737).

(4) Daily Movements of Critical Aircraft

This also provides an effective demand measure.

(5) Involvement Ratio

The involvement ratios is defined as the ratio of the airport's passenger volume to the aircraft's available seats. It is also not well measured - especially for multi-stop flight routes.

(6) Maximum Passenger Loads

This is the largest load of either enplaning or deplaning passengers. It is neither commonly used nor measured.

The problem with the use of multiple criteria was that one criterion might indicate a different STEP size than the others did. This often made selection of a size a matter of judgement.

To simplify the process of selection, the latest draft (1983) of the STEP Planning and Design Manual[20] proposes that the entire selection be based on a half-hourly design volume of passengers. Again, there is no official definition of this volume, although the 90th percentile by passenger volume is often used. This approach places a great deal of emphasis on a single design value. An error in the measurement of the value or in its forecasting can lead to an erroneous STEP selection. For example, an error of the order of magnitude of ten passengers above the actual half-hourly volume would cause selection of a STEP 6 terminal, when a STEP 5 would have been sufficient. More research into the behaviour of small terminal flight loads should improve the accuracy of design volume calculations and, therefore, the selection of appropriate building sizes.

2.2 Design Volume Determination

There are a multitude of design volumes for passengers. Statistics can describe annual, daily or hourly volumes. They can be classified by origin and destination or by enplanement and deplanement. These can be further broken down into major carrier, charter, domestic, transborder or international categories.

Terminal design is usually based on a peak hourly design volume. Half-hourly and six-hourly periods are also used.

Definitions of what constitutes the peak hour abound. Horonjeff[11] suggests that a planner simply select a reasonable volume. The American Federal Aviation Authority[13] suggests the busiest hour of the busiest day of a typical week. Although American airports are not guided by a single body, many seem to favour the use of a percentage of either the annual total[8] or the average day of the busiest month[19]. Other definitions proposed include the peak hour of the average weekday in the busiest quarter and the nth highest hour of all hours of the year.

Until recently Transport Canada[17] has used an hour or half-hour percentile definition. For larger airports, the accepted planning volume was the 90th percentile of the annual distribution of passengers. This more statistical approach relies on the prediction of the upper tail of a distribution curve. In the absence of complete data, some assumptions must be made as to the form of this curve.

Some terminal design procedures rely heavily on the hourly (or half-hourly) design volume. For example Transport Canada's STEP method uses it. Simulation models may or may not make use of it. Most of them simulate activity over the course of eighteen or twenty-four hours, and so require daily input instead (such as a flight schedule or the passenger/non-passenger ratio for each hour of the day).

Some of the inherent features of basing terminal design on the hourly design volume are listed below. This compilation is based on the comments of Braaksma[3], de Neufville[8], Horonjeff[12], and Hamzawi[19].

(1) no commonly agreed upon definition;

(2) very dependent upon aircraft size, schedule and routing;

(3) statistic does not incorporate stochastic variability of the queueing process;

(4) does not reflect individual airport characteristics such as type of traveller (commuter, vacationer) or catchment area size; and

(5) not directly useful for many computer simulations.

2.3 Passenger Distribution Functions

Transport Canada has historically assumed that flight loads at small airports are normally distributed. This selection has been made for convenience only, since it has also determined that the Normal is not the best model for all cases. Transport Canada is of the opinion that each airport follows a different distribution.

Airlines (and the manufacturers who sell aircraft to them) have a different approach to the study of passenger loads. They are more interested in the number of occupied seats in the aircraft than in the passenger volumes at the airports.

Airlines use a different combination of the volumes. For example, in Canadian Pacific's Vancouver-Terrace-Prince Rupert-Vancouver flight, the airline might be interested in knowing the probability of filling the seats on the second leg - between Terrace and Prince Rupert. Since there is virtually no Terrace to Prince Rupert traffic, this would be the total of those going from Vancouver to Prince Rupert (Prince Rupert's deplaners) and those going from Terrace to Vancouver (Terrace's enplaners). Therefore, the number of seats occupied is the result of th summation of two independent randomly distributed variables. It is desirable, therefore for an airline to use a distribution form which is additive - that is, the summation takes the same distribution form as the parts.

Of course, this would be only one reason for an airline to select a particular distribution, since it is only one use for the distribution. Lauchli[14] selected the Erlang function during research for Swiss Air to determine optimal seating configurations of aircraft. Vella, Martin and Whale[22] continued this work for Quantas Airlines but decided that the normal and binomial distributions produced better results. Wang[24] used an empirical distribution function to determine booking levels for Cathay Pacific's long haul flights.

The behaviour aboard the aircraft, which interests the airlines, is obviously related to activity at the airports, which is of interest in this work. For example, the availability of seats limits the number of passengers that may board the aircraft. Also, a second flight may be warranted at a certain point, even though the increase in demand is occurring at another stop in the flight route.

Because of this close interaction, the frequency distributions that the airlines and the aircraft manufacturers have selected for use are of interest for airport terminal sizing.

3. METHODOLOGY

3.1 Data Description

The flight load volumes came from eight airports in British Columbia. They were originally released by the airline carriers to Transport Canada in order to assist in the planning of airport terminal buildings. Although the carriers are not obliged to release this information, they did so to ensure reasonable sizing of the facilities which they will be leasing and to promote co-operation with the government. The data was not, however, meant to be used publicly so the airports have been designated by letters (A through H).

All of the flights occurred between 1978 and 1982 at airports with three jet stops or less each day. All flights were served with Boeing 737 jets.

The list of airports in Table II illustrates the years and flights of the available data.

As described earlier, the term "flight event" will be used for the sum of the deplaned and enplaned passengers during a single visit of an aircraft. A "flight" will be the total of all flight events that occur over the course of one year at the same time of day. Therefore, each flight will contain 366 flight events or less.

Flight events cancelled due to poor weather (a very common occurrence) were excluded. Also not considered were flights which ran for only a portion of the year. This meant that a distribution of all airport activity could not be assembled for Airport H since it had several flights which ran in the summer only.

The characteristics of flight events at small airports have simplified the analysis. Volumes of deplaned and enplaned passengers are

TABLE II

Quantity of Available Data

Airport Designation	Years of Data	Number of Flights per Year	Total Number of Flights
A	1981	2	2
B	1978 to 1982	3	15
C	1980, 1981	2	4
D	1979, 1980	1	2
E	1979, 1980, 1982	1	3
F	1980, 1981	1	2
G	1981, 1982	3	6
H	1980 to 1982	1	3
			37 Flights

equivalent to half-hourly volumes because the flight turnarounds are less than thirty minutes and because the flights are separated from each other. Furthermore, there are no connecting or transiting passengers to account for.

The prevalence of triangular routing has been mentioned. All of the airports in this study are part of such routes - most of which originate or terminate in Vancouver. The routes with both stops included in this study are:

Vancouver - A - C - Vancouver;

Vancouver - F - G - Vancouver (or reverse);

and Vancouver - H - D - Vancouver.

The raw data was assembled and entered into APL computer language, such that each flight was a vector. The flight vectors have from 137 to 361 elements. Each element is a flight event.

To analyze distributions for airports with more than one flight, the flight vectors for that year were concatenated.

The descriptive details of these variables are given in Tables III and IV.

When arranged into frequency classes, the histograms had a right skew. A good model should reproduce this tendency.

3.2 Features of the Distributions

Three model distributions will be compared to the observed data. A more informed selection can be made if the characteristics of each one are understood.

The Normal or Gaussian is the most widely used of all frequency distributions. Its formula is:

TABLE III

Data Description for Flights

Airport	Flight	Number of Flight Events	Total Number of Passengers	Year
A	A1	292	18,957	1981
	A2	358	32,694	1981
B	B1	137	12,629	1978
	B2	247	28,167	1979
	B3	258	32,201	1980
	B4	287	30,809	1981
	B5	222	21,059	1982
	B6	259	17,813	1978
	B7	291	24,884	1979
	B8	274	26,074	1980
	B9	279	25,726	1981
	B10	167	16,154	1982
	B11	339	39,195	1978
	B12	291	34,316	1979
	B13	273	34,359	1980
	B14	294	16,519	1981
	B15	291	30,682	1982
C	C1	332	15,388	1980
	C2	196	12,010	1980
	C3	317	13,225	1981
	C4	357	19,816	1981
D	D1	328		1979
	D2	352	23,857	1980
E	E1	361	36,607	1979
	E2	360	42,031	1980
	E3	354	29,899	1982
F	F1	344	36,941	1980
	F2	358	40,427	1981
G	G1	344	13,995	1981
	G2	352	28,758	1981
	G3	313	44,647	1981
	G4	328	13,964	1982
	G5	343	27,324	1982
	G6	329	43,684	1982
H	H1	355	30,790	1980
	H2	328	28,664	1981
	H3	331	24,705	1982

TABLE IV

Data Description for Airports

Airport	Year	Combination of Flights	Number of Flight Events	Total Number of Passengers
A	1981	A1 + A2	650	51,651
B	1978	B1 + B6 + B11	735	69,637
	1979	B2 + B7 + B12	829	87,367
	1980	B3 + B8 + B13	548	60,455
	1981	B4 + B9 + B14	860	73,054
	1982	B5 + B10 + B15	680	67,895
C	1980	C1 + C2	528	27,398
	1981	C3 + C4	674	33,041
D	1979	D1	328	
	1980	D2	352	23,857
E	1979	E1	361	36,607
	1980	E2	360	42,031
	1982	E3	354	29,899
F	1980	F1	344	36,941
	1981	F2	358	40,427
G	1981	G1 + G2 + G3	1009	87,400
	1982	G4 + G5 + G6	1000	84,972
H				

$$f(x) = \frac{1}{\sigma \sqrt{2 \pi}} \exp \left[- \frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2\right]$$

μ = mean - parameter approximated by $\Sigma \frac{x_i}{n}$ where n = number of data points.

σ = variance - parameter approximated by $\Sigma \frac{(x_i - x)^2}{n}$

x = variable value.

The normal is a continuous distribution function that is symmetrical about its mean. It is commonly used to describe variations in physical measurements. The sum of two normally distributed variables is also normally distributed. This term for this feature is additive regenerative.

The formula for the Poisson distribution is:

$$P(r) = \frac{e^{-\mu}\mu^r}{r!}$$

r = discrete variable value

μ = mean = variance - parameter approximated by $\Sigma \frac{x_i}{n}$ approximated.

It is used for such things as determining the number of accidents in a given time interval. It has a right skew, but this decreases as the parameter (μ) increases. The Poisson distribution is also additive regenerative.

The lognormal distribution has the logarithm of its variable values distributed normally. Natural logarithms are usually used but another base is possible. The function for a base ten lognormal distribution is:

$$f(x) = \frac{.4343}{x\sigma_x \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log x - \mu}{\sigma_x}\right)^2\right]$$

$$f(x) = \frac{.4343}{x\sigma_{\log} \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log_{10}(x/\mu_x)}{\sigma_{\log x}}\right)^2\right]$$

The lognormal is a highly flexible distribution which skews to the right. It is not regenerative by addition, however, but by multiplication. That is, the multiplication of lognormally distributed variables is also lognormal but the addition of them is not.

3.3 Procedure

After the data was assembled, an APL computer program (owned by I.P. Sharp Associates) formed each vector into a frequency distribution. The parameters of each of the three model distributions were calculated from the data, and the program compared the expected curves to those of the observed data. The three quantitative methods used to make the comparison were:

i) the "goodness-of-fit", as measured by the Chi-Squared statistics;

ii) the ability of the model to predict the tail of the distribution, as measured by the 90th percentile;

iii) ranking of the models by visual inspection.

The quantitative procedures used are outlined in this chapter. Discussion of the two qualitative selection criteria - the ability of the models to be easily understood and their applicability - has been deferred until the next chapter.

Three sources of uncertainty inherent in any curve-fitting are:

i) the natural variation of the data due to its randomness (due to unknown factors);

ii) the statistical failure to effectively estimate the parameters from the data;

iii) the fact that a given model is poor for describing the curve.

Only the latter two sources can be minimized with a larger sample.

The data had to be grouped into intervals and the expected and observed frequencies of each interval studied. A cumulative distribution form would have eliminated the need to use intervals, but the APL program used was not able to construct it.

One inherent feature of histograms is that each individual interval has a certain probability of matching the frequency that the model has predicted for it. Even if the model is a good one, a perfect fit over all intervals, while being the most likely event, is still not very likely. As the number of intervals increases, a perfect fit becomes more rare. If fewer intervals are used, fitting the data to the model is more likely. However, if several models are under consideration, more of them will fit. This makes a selection difficult. Therefore, some sort of trade-off is needed.

Secondly, histogram class divisions should theoretically be made so that the number of data points is the same in each. For example, the region of higher frequency will have narrower intervals. Even though this is statistically preferable, variable interval widths are not commonly used.

In any case, the APL program used had limited flexibility. It was a standard statistical program and could only accept equal band widths. A band width of ten was selected for all distributions. This meant that the number of divisions ranged from 10 to 27, according to the spread in data values.

The APL program used for the Normal and Poisson distributions is shown in Appendix A. It is an interactive program which requests end points and the class width from the user. It then requires a selection as to whether the Normal or Poisson distribution is to be fit to the data.

Originally, the lognormal curve fitting was done by taking the logarithm of each data point and then running the standard program with the Normal option. This approach proved to be unsatisfactory since the scale was the data logarithm. The histogram intervals could not be compared to those of the Normal and Poisson.

A new program was written for the lognormal in order to permit direct comparison. It is similar to the standard program (although less refined) and is listed in Appendix A. The body was written by the author (LN program) but the histogram plotting function (HISTO and CLASSIFY) were written by I.P. Sharp Associates.

The three quantitative criteria used were: the statistical fit of the model; the ability of the model to predict a design volume; and the overall fit of the model as judged by a visual inspection.

3.3.1 Goodness-of-Fit Criterion

The Chi-Squared statistic is produced by the program as a measure of the "goodness-of-fit" of the model. It is used to decide whether or not a distribution should be retained or rejected. The Chi-Squared statistic is not meant to be used to choose among models.

The definition is:

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

where o_i = observed frequency

e_i = expected frequency

i = index of interval

The calculation should only be performed when the expected frequency of each interval is at least five, otherwise distortions can occur. For frequencies less than five, intervals should be combined. The programs did not do this, so the Chi-Squared value was hand-corrected by the author. Original and corrected values are in Appendix B.

A few of the flights could not have their Chi-Squared values corrected because the program did not display enough significant figures. These five flights were omitted from further calculations.

Use of the Chi-Squared values will be discussed in the next chapter. It should be noted, however, that several pieces of information have been obtained from the data. As described in Section 3.2, the Normal and lognormal have two parameters (the mean and variance) and the Poisson has one (the mean). These have been estimated by the average or the standard deviation of the data (or its logarithm). In addition, the total number of flight events has been used to determine the expected frequencies. The degrees of freedom of each distribution will depend upon this information.

3.3.2 Design Volume Criterion

As previously mentioned, there are many definitions of the planning design volume. Knowledge of this distribution will allow for a more informed decision as to which definition should be used.

This study will look at the 90th percentile of the flight events as a representation of the upper tail of the distribition. The actual 90th percentiles were calculated from the observed data. Short programs were written to derive the expected figures. The Normal and lognormal were written in APL but the Poisson was written in FORTRAN. They are listed in Appendix A.

The other percentile definition used for large Canadian airports. It is the 90th percentile by passenger volume - that is, 10 percent of the passengers will experience congestion. This is in contrast to the above percentile definition which would allow 10 percent of the flights to be above it and thus experience congestion. The second definition could have been used to measure the upper tail predictability, but it is difficult to calculate and is usually only slightly higher from the 90th percentile by passenger event. Figure 3 illustrates the 90th percentiles by passenger volume load were calculated for comparison purposes.

One other definition of the planning design volume is also included. This is based on the average load factor and is calculated by adding fifteen percent to the mean load factor and multiplying this by the total number of arriving and departing seats available:

$$[\text{mean load factor} + 15\%] \times [\#\text{arrival and departure seats}]$$

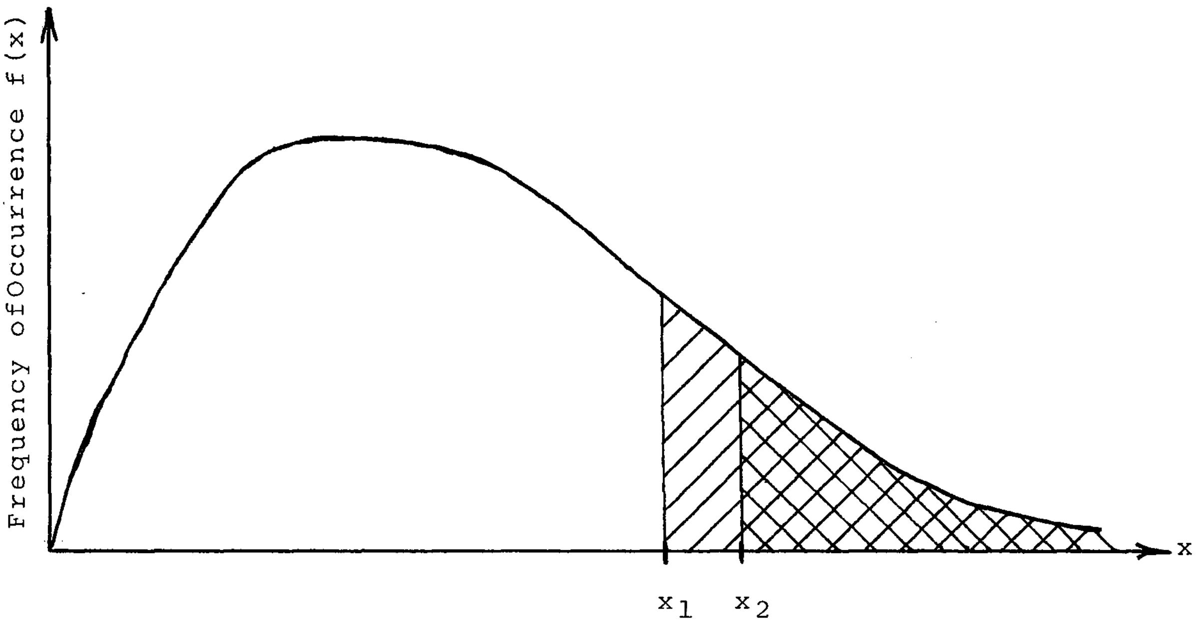
Because the load factor is the proportion of available seats that are used, for a Boeing 737 with 117 seats this expression reduces to:

$$\left[\frac{(\text{mean deplaned + enplaned passengers})}{2 \times 117 \text{ seats}} + .15\right] \times [2 \times 117 \text{ seats}]$$

$$= (\text{mean deplaned + enplaned passengers}) + 35.$$

FIGURE 3

<u>90th Percentile by Flight Event and by Passenger Volume</u>



Deplaned + Enplaned Passenger Loads

X_1 = 90th percentile by flight event

X_2 = '90th percentile' by passenger volume

// = 10 percent of flights $\left[\int_{x_1}^{\infty} f(x)dx = .10 \times \int_{-\infty}^{\infty} f(x)dx\right]$

\\ = 10 percent of passengers $\left[\int_{x_2}^{\infty} x\, f(x)dx = .10 \times \int_{-\infty}^{\infty} xf(x)dx\right]$

3.3.3 Visual Inspection Criterion

This somewhat judgemental method of histogram selection was included to ensure that there was some measure of the reasonableness of each model. It also allowed for detection of any unexpected deviations in the data or any trends in the curve-fitting.

Since all histograms have equal interval width, a visual comparison was given a ranking from best to worst and the results totalled. The rating was made on the basis of how the model matched the overall shape of the curve without trying to duplicate the Chi-Squared or 90th percentile measurements.

4. ANALYSIS

4.1 Criteria for Acceptance or Rejection

4.1.1 Goodness-of-Fit

The Chi-Squared value was used to accept or reject a model distribution for each set of observed data. The number of acceptances among all of the flights or all of the airports was then calculated as a percentage. The statistic was not used directly to decide which of the three model distributions best fit one particular flight or airport, since such comparisons are not its purpose.

Acceptance of a model is the "null hypothesis". This hypothesis states that there is no difference between the expected and observed curves that cannot be attributed to randomness. It is assumed that the null hypothesis is true until it has been proven otherwise. The onus is, therefore, to prove that a model should be rejected.

To generate the proof, the critical value of Chi-Squared is found from the theoretical Chi-Squared distribution:

$$\chi^2_{\alpha,\nu}$$

where ν = the degree of freedom (number of intervals less the number of parameters estimated from the data)

α = the level of significance (area under the Chi-Squared curve, above critical value).

If the Chi-Squared value calculated from the observed and expected frequencies is less than the critical Chi-Squared, the fit is a good one and the null hypothesis is true. The specified level of significance (α) is equal to the probability that the calculated Chi-Squared (χ^2) will exceed the critical value ($\chi^2_{\alpha,\nu}$) even though the fit is a good one. Therefore, there is a probability, α, of rejecting a model that was, in fact, a good fit.

In this study, the comparison was made at three different significance levels: .05, .01, and .001. As the level of significance decreases, there is more probability of the model being accepted. This is due to the reduction of the probability of rejecting the model, even though it fits the data. This is called a Type I error. But by lowering this probability, the chances of accepting a model that is actually a poor one are increased - a Type II error. Therefore, a balance is needed since minimizing one type of error increases the probability of the other. Both types can be reduced, however, by increasing the sample size.

In this analysis, the Chi-Squared is not the only criterion for selection. Therefore, the total number of acceptances can be compared at the three significance levels without forcing a conclusive decision on this criterion alone.

The calculated and critical Chi-Squared values are shown in Table V for flights and Table VI for airports.

Table VII and VIII show the percentage of acceptance for each model at each significance level. The results of Table VII for the thirty-seven individual flights show that the lognormal model was a statistically good fit to the data more often than the Normal or Poisson.

When the flights were combined to get yearly distributions by airport, all three models fit less often, although the lognormal was still slightly more successful.

TABLE V (a)

Comparison of Calculated and Critical Chi-Squared Values for Flights

Airport	Flight	NORMAL			
		χ^2(d.o.f)	$\chi^2_{.05}$	$\chi^2_{.01}$	$\chi^2_{.005}$
A	A1	19.6 (7)	14.1	18.5	20.3*
	A2	29.5 (10)	18.3	23.2	25.2
B	B1	48.6 (11)	19.7	24.7	26.8
	B2	23.0 (11)	19.7	24.7*	26.8*
	B3	22.3 (12)	21.0	26.2*	28.3*
	B4	28.1 (9)	16.9	21.7	23.6
	B5	12.9 (8)	15.5*	20.1*	22.0*
	B6	64.0 (11)	19.7	24.7	26.8
	B7	68.9 (12)	21.0	26.2	28.3
	B8	49.5 (13)	22.4	27.7	29.8
	B9	56.4 (11)	19.7	24.7	26.8
	B10	15.6 (11)	19.7*	24.7*	26.8*
	B11	38.7 (14)	23.7	29.1	31.3
	B12	46.1 (13)	22.4	27.7	29.8
	B13	45.6 (15)	25.0	30.6	32.8
	B14	213.3 (9)	16.9	21.7	23.6
	B15	32.8 (12)	21.0	26.2	28.3
C	C1	21.8 (4)	9.5	13.3	14.9
	C2	12.7 (5)	11.1	15.1*	16.8*
	C3	10.2 (3)	7.8	11.3*	12.8*
	C4	53.0 (5)	11.1	15.1	16.8
D	D1	5.1 (6)	12.6*	16.8*	18.5*
	D2	13.3 (7)	14.1*	18.5*	20.3*
E	E1	21.8 (8)	15.5	20.1	22.0*
	E2	16.4 (13)	22.4*	27.7*	29.8*
	E3	†			
F	F1	4.7 (11)	19.7*	24.7*	26.8*
	F2	41.0 (12)	21.0	26.2	28.3
G	G1	†			
	G2	†			
	G3	20.1 (14)	23.7*	26.1*	31.3*
	G4	12.7 (5)	11.1	15.1*	16.8*
	G5	9.9 (9)	21.0*	26.2*	28.3*
	G6	21.8 (14)	23.7*	29.1*	31.3*
H	H1	7.0 (9)	16.9*	21.7*	23.6*
	H2	9.1 (8)	15.5*	20.1*	22.0*
	H3	21.5 (8)	15.5	20.1	22.0*
Number of Acceptances			11	16	19

*Acceptance at this Level of Significance ($\chi^2 < \chi^2_{\alpha,\nu}$)

†Significant figures of program do not allow Ch-Squared calculations.

Source: Freund and Williams[4].

TABLE V (b)

Comparison of Calculated and Critical Chi-Squared Values for Flights

		POISSON			
Airport	Flight	χ^2(d.o.f)	$\chi^2_{.05}$	$\chi^2_{.01}$	$\chi^2_{.005}$
A	A1	39.2 (9)	16.9	21.7	23.6
	A2	29.3 (12)	21.0	26.2	28.3
B	B1	43.3 (12)	21.0	26.2	28.3
	B2	22.2 (13)	22.4*	27.7*	29.8*
	B3	22.0 (13)	22.4*	27.7*	29.8*
	B4	40.8 (12)	21.0	26.2	28.3
	B5	29.6 (11)	19.7	24.7	26.8
	B6	57.1 (10)	18.3	23.2	25.2
	B7	97.7 (11)	19.7	24.7	26.8
	B8	72.0 (11)	19.7	24.7	26.8
	B9	89.5 (12)	21.0	26.2	28.3
	B10	19.6 (11)	19.7*	24.7*	26.8*
	B11	48.7 (14)	23.7	29.1	31.3
	B12	43.5 (13)	22.4	27.7	29.8
	B13	69.6 (14)	23.7	29.1	31.3
	B14	†			
	B15	30.4 (12)	21.0	26.2	28.3
C	C1	76.9 (8)	15.5	20.1	22.0
	C2	50.4 (8)	15.5	20.1	22.0
	C3	88.8 (6)	12.6	16.8	18.6
	C4	92.1 (7)	14.1	18.5	20.3
D	D1	27.3 (9)	16.9	21.7	23.6
	D2	34.1 (9)	16.9	21.7	23.6
E	E1	37.1 (11)	19.7	24.7	26.8
	E2	14.7 (14)	23.7*	29.1*	31.3*
	E3	24.6 (11)	19.7	24.7*	26.8*
F	F1	20.3 (13)	22.4*	27.7*	29.8*
	F2	27.2 (13)	22.4	27.7*	29.8*
G	G1	†			
	G2	55.2 (11)	19.7	24.7	26.8
	G3	18.1 (14)	23.7*	29.1*	31.3*
	G4	24.1 (7)	14.1	18.5	20.3
	G5	29.0 (10)	21.0	26.2	28.3*
	G6	36.6 (15)	25.0	30.6	32.8
H	H1	27.7 (12)	21.0	26.2	28.3
	H2	39.2 (12)	21.0	26.2	28.3
	H3	45.9 (10)	18.3	23.2	25.2
Number of Acceptances			6	8	9

*Acceptance of this Level of Significance ($\chi^2 < \chi^2_{\alpha,\nu}$)
†Significant figures of program do not allow Chi-squared calculation.
Source: Freund and Williams[4]

TABLE V (c)

Comparison of Calculated and Critical Chi-Squared Values for Flights

Airport	Flight	LOGNORMAL χ^2(d.o.f)	$\chi^2_{.05}$	$\chi^2_{.01}$	$\chi^2_{.005}$
A	A1	14.5 (7)	14.1	18.5*	20.3*
	A2	19.7 (9)	16.9	21.7*	23.6*
B	B1	22.3 (9)	16.9	21.7	23.6*
	B2	3.4 (11)	19.7*	24.7*	26.8*
	B3	7.7 (12)	21.0*	26.2*	28.3*
	B4	19.6 (9)	16.9	21.7*	23.6*
	B5	4.1 (8)	15.5*	20.1*	22.0*
	B6	16.5 (11)	19.7*	24.7*	26.8*
	B7	21.5 (12)	21.0	26.2*	28.3*
	B8	16.7 (12)	21.0*	26.2*	28.3*
	B9	18.0 (12)	21.0*	26.2*	28.3*
	B10	9.2 (9)	16.9*	21.7*	23.6*
	B11	52.9 (13)	22.4	27.7	29.8
	B12	34.0 (12)	21.0	26.2	28.3
	B13	43.1 (14)	23.7	29.1	31.3
	B14	53.4 (9)	16.9	21.7	23.6
	B15	23.9 (9)	16.9	21.7	23.6
C	C1	6.9 (6)	12.6*	16.8*	18.5*
	C2	14.1 (5)	11.1	15.1*	16.8*
	C3	3.7 (3)	7.8*	11.3*	12.8*
	C4	9.1 (6)	12.6*	16.8*	18.5*
D	D1	11.2 (8)	15.5*	20.1*	22.0*
	D2	36.8 (8)	15.5	20.1	22.0
E	E1	18.2 (13)	22.4*	27.7*	29.8*
	E2	14.9 (12)	21.0*	26.2*	28.3*
	E3	2.8 (10)	18.3*	23.2*	25.2*
F	F1	34.6 (11)	19.7	24.7	26.8
	F2	41.1 (12)	21.0	26.2	28.3
G	G1	14.4 (6)	12.6	16.8*	18.5*
	G2	31.4 (8)	15.5	20.1	22.0
	G3	16.1 (13)	22.4*	27.7*	29.8*
	G4	16.4 (6)	12.6	16.8*	18.5*
	G5	64.3 (10)	21.0	26.2	28.3
	G6	35.8 (14)	23.7	29.1	31.3
H	H1	26.0 (9)	16.9	21.7	23.6
	H2	24.2 (8)	15.5	20.1	22.0
	H3	28.1 (8)	15.5	20.1	22.0
Number of Acceptances			15	22	23

Accepted at this Level of Significance ($\chi^2 < \chi^2_{\alpha,\nu}$)

Source: Freund and Williams[4]

TABLE VI (a)

Comparison of Calculated and Critical Chi-Squared Values for Airports

Airport	Year	No. of Flights	NORMAL χ^2(d.o.f)	NORMAL $\chi^2_{.05}$	NORMAL $\chi^2_{.01}$	NORMAL $\chi^2_{.005}$
A	1981	2	75.8 (10)	18.3	23.2	25.2
B	1978	3	92.6 (17)	27.6	33.4	35.7
	1979	3	53.9 (16)	26.3	32.0	34.3
	1980	3	82.0 (17)	27.6	33.4	35.7
	1981	3	132.6 (15)	25.0	30.6	32.8
	1982	3	55.4 (13)	22.4	27.7	29.8
C	1980	2	34.7 (6)	12.6	16.8	18.5
	1981	2	52.8 (5)	11.1	15.1	16.8
D	1979	1	5.1 (6)	12.6*	16.8*	18.5*
	1980	1	13.3 (7)	14.1*	18.5*	20.3*
E	1979	1	21.8 (8)	15.5	20.1	22.0*
	1980	1	16.4 (13)	22.4*	27.7*	29.8*
	1982	1	†			
F	1980	1	4.7 (11)	19.7*	24.7*	26.8*
	1981	1	41.0 (12)	21.0	26.2	28.3
G	1981	3	270.7 (19)	30.1	36.2	38.6
	1982	3	194.1 (18)	28.9	34.8	37.2
H						
Number of Acceptances				4	4	5

*Acceptance at this Level of Significance ($\chi^2 < \chi^2_{\alpha,\nu}$)
†Significant figures of program do not allow Chi-Squared calculations
Source: Freund and Williams (Reference 3)

TABLE VI (b)

Comparison of Calculated and Critical Chi-Squared Values for Airports

Airport	Year	No. of Flights	POISSON			
			χ^2(d.o.f)	$\chi^2_{.05}$	$\chi^2_{.01}$	$\chi^2_{.005}$
A	1981	2	58.5 (13)	22.4	27.7	29.8
B	1978	3	240.4 (14)	23.7	29.1	31.3
	1979	3	120.3 (15)	25.0	30.6	32.8
	1980	3	210.2 (15)	25.0	30.6	32.8
	1981	3	528.5 (14)	23.7	29.1	31.3
	1982	3	46.6 (14)	23.7	29.1	31.3
C	1980	2	78.2 (9)	16.9	21.7	23.6
	1981	2	131.3 (8)	16.9	20.1	22.0
D	1979	1	27.3 (9)	16.9	21.7	23.6
	1980	1	34.1 (9)	16.9	21.7	23.6
E	1979	1	37.1 (11)	19.7	24.7	26.8
	1980	1	14.7 (14)	23.7*	29.1*	31.3*
	1982	1	24.6 (11)	19.7	24.7*	26.8*
F	1980	1	20.3 (13)	22.4*	27.7*	29.8*
	1981	1	27.2 (13)	22.4	27.7*	29.8*
G	1981	3	2230.0 (14)	23.7	29.1	31.3
	1982	3	1490.0 (14)	23.7	29.1	31.3
H						
Number of Acceptances				2	4	4

*Acceptance at this Level of Significance ($\chi^2 < \chi^2_{\alpha,\nu}$)

Source: Freund and Williams[4]

TABLE VI (c)

Comparison of Calculated and Critical Chi-Squared Values for Airports

Airport	Year	No. of Flights	LOGNORMAL χ^2(d.o.f)	$\chi^2_{.05}$	$\chi^2_{.01}$	$\chi^2_{.005}$
A	1981	2	28.5 (11)	19.7	24.7	26.8
B	1978	3	71.5 (16)	26.3	32.0	34.3
	1979	3	59.1 (17)	27.6	33.4	35.7
	1980	3	52.7 (17)	27.6	33.4	35.7
	1981	3	109.4 (16)	26.3	32.0	34.3
	1982	3	28.2 (13)	22.4	27.7	29.8*
C	1980	2	9.2 (7)	14.1*	18.5*	20.3*
	1981	2	9.3 (6)	12.6*	16.8*	18.5*
D	1979	1	11.2 (8)	15.5*	20.1*	22.0*
	1980	1	36.8 (8)	15.5	20.1	22.0
E	1979	1	18.2 (13)	22.4*	27.7*	29.8*
	1980	1	14.9 (12)	21.0*	26.2*	28.3*
	1982	1	2.8 (10)	18.3*	23.2*	25.2*
F	1980	1	34.6 (11)	19.7	24.7	26.8
	1981	1	41.1 (12)	21.0	26.2	28.3
G	1981	3	99.0 (19)	30.1	36.2	38.6
	1982	3	75.9 (20)	31.4	37.6	40.0
H						
Number of Acceptances				6	6	7

*Acceptance at this Level of Significance ($\chi^2 < \chi^2_{\alpha,\nu}$)

Source: Freund and Williams[4]

TABLE VII

Acceptance Rate By Flight

Level of Significance (α)	Normal	Poisson	Lognormal
.05	33%	17%	41%
.01	47%	23%	59%
.005	56%	26%	62%

TABLE VIII

Acceptance Rate By Airport (All Airports)

Level of Significance (α)	Normal	Poisson	Lognormal
.05	25%	12%	35%
.01	25%	24%	35%
.005	31%	24%	41%

TABLE IX

Acceptance Rate By Airport

(Airports with Multiple Flights Only)

Level of Significance (α)	Normal	Poisson	Lognormal
.05	0%	0%	20%
.01	0%	0%	20%
.005	0%	0%	30%

These results are somewhat misleading, however, since they include airports with only one flight, which are also included as single flights in Table VII. The airports which had multiple flights had a much poorer acceptance rate. In fact, the Normal and Poisson did not provide a good fit even once. These results are shown by Table IX.

It is significant that the theoretical distributions fit more poorly as more flights were included. This was to be expected as the flights at one airport can be distinct in their characteristics (average loads, days of the week, et cetera). The final outcome may be that these theoretical models should be used only to describe individual flights.

4.1.2 Design Volumes

Tables X and XI list the 90th percentiles by flight event for flights and airports respectively. The predicted values for Normal, Poisson and lognormal functions were calculated from the estimated parameters. Short APL programs were used for the Normal and lognormal. The Poisson distribution required a separate program because of the rounding errors involved. The APL functions could not handle such high means, so a FORTRAN program was written. All are listed in Appendix A.

The 90th percentile by passenger volume and the "mean load factor plus 15 percent" volume are attached for interest. Both are alternative definitions of the peak design volume and have been calculated directly from the data.

Inspection of Tables X and XI reveals that both the Normal and lognormal provide reasonable predictions of the 90th percentile values. The Poisson distribution is consistently low in its estimate. The average

TABLE X

Comparison of Actual and Predicted 90th Percentiles by Flight

		90th Percentile by Event					
Airport	Flight	Actual	Predicted by Normal	Predicted by Poisson	Predicted by Lognormal	90th Percentile by Pax Volume	Mean Load Factor + 15%
A	A1	89	91	76	93	98	100
	A2	132	127	104	130	138	126
B	B1	141	133	106	132	161	127
	B2	157	155	128	156	170	149
	B3	168	166	140	169	180	160
	B4	141	141	121	144	147	142
	B5	129	126	108	128	135	130
	B6	119	107	80	110	130	104
	B7	143	133	98	135	159	121
	B8	151	144	108	148	165	130
	B9	146	140	106	143	162	127
	B10	144	142	110	147	156	132
	B11	163	162	130	170	169	151
	B12	169	162	132	166	177	153
	B13	188	179	141	186	201	161
	B14	106	101	66	94	150	91
	B15	151	145	119	149	159	141
C	C1	66	65	56	67	74	81
	C2	83	83	72	85	90	96
	C3	58	58	51	59	61	77
	C4	78	76	66	77	86	91
D	D1	91	89	73	92	101	97
	D2	95	94	79	99	103	103
E	E1	136	137	115	142	145	137
	E2	163	159	131	164	170	152
	E3	122	121	97	122	132	120
F	F1	144	144	121	152	150	143
	F2	154	155	127	164	168	148
G	G1	61	63	50	66	75	76
	G2	109	111	94	115	115	117
	G3	198	192	158	197	206	178
	G4	65	64	52	69	72	78
	G5	110	112	92	123	114	115
	G6	190	183	148	191	199	168
H	H1	116	118	99	122	125	122
	H2	115	116	100	120	124	123
	H3	102	102	86	105	112	110

TABLE XI

Comparison of Actual and Predicted 90th Percentiles by Airports

Airport	Year	No. of Flights	90th Percentile by Event: Actual	90th Percentile by Event: Predicted by Normal	90th Percentile by Event: Predicted by Poisson	90th Percentile by Event: Predicted by Lognormal	90th Percentile by Pax Volume	Mean Load Factor + 15%
A	1981	2	119	115	92	115	132	145
B	1978	3	155	145	108	157	165	130
	1979	3	160	154	119	161	171	141
	1980	3	174	165	124	174	190	145
	1981	3	139	136	97	149	155	120
	1982	3	141	139	113	144	153	135
C	1980	2	74	74	62	76	82	87
	1981	2	70	70	58	71	79	84
D	1979	1	91	89	73	92	101	97
		1	95	94	79	99	103	103
E	1979	1	136	137	115	142	145	137
	1980	1	163	159	131	164	170	152
	1982	1	122	121	97	122	132	120
F	1980	1	144	144	121	152	150	143
	1981	1	154	155	127	164	168	148
G	1981	3	162	150	99	165	189	122
	1982	3	156	145	97	161	179	120
H								

TABLE XII

Average Differences Between Actual and Predicted 90th Percentiles

	Normal	Poisson	Lognormal
For Flights $= \frac{\Sigma \lvert \text{predicted-actual} \rvert}{37} \times 100\%$	3.3%	24.9%	4.1%
For Airports $= \frac{\Sigma \lvert \text{predicted-actual} \rvert}{17} \times 100\%$	4.4%	31.9%	4.0%

difference between actual and predicted values are shown in Table XII for both flights and airports. The Poisson is clearly unacceptable.

A good model should predict the design volume within ten passengers if it is to be used in the Canadian STEP method. For individual flights the Normal failed to do this twice (Flights B6 and B7) while the lognormal exceeded ten three times (Flights B14, F2 and G5). The Poisson, however, could only predict the 90th percentile within ten three times (Flights C3, G4 and G6). Similarly, for the airport values the Normal missed four times and the lognormal twice, but the Poisson was never within ten passengers of the actual value.

The other two planning values show that the 90th percentile by passenger volume is slightly higher, but very close to the 90th percentile by event. The second definition (mean plus 15 percent) is much more variable.

4.1.3 Visual Inspection

The subjective ranking of the overall curve-fit is shown in Tables XIII and XIV. The lognormal was the most effective, followed by the Normal.

The inspection also pointed out some of the trends in the data and in the models. The high peak and the right skew common in most of the histograms were not well reproduced by the Poisson and Normal distributions. Also, many histograms had at least one other secondary peak to the right of the highest peak. This second mode may be due to the upper limit of the aircraft capacity, but proof of this conjecture would be beyond the scope of this study.

TABLE XIII
Visual Inspection of Histograms for Flights

Airport	Flight	RANKING (1=BEST FIT, 3=WORST FIT) Normal	Poisson	Lognormal
A	A1	3	2	1
	A2	2	3	1
B	B1	3	2	1
	B2	2	3	1
	B3	2	3	1
	B4	2	3	1
	B5	2	3	1
	B6	3	2	1
	B7	2	3	1
	B8	2	3	1
	B9	3	2	1
	B10	3	2	1
	B11	2	1	3
	B12	3	2	1
	B13	2	3	1
	B14	3	2	1
	B15	2	3	1
C	C1	1	3	2
	C2	1	3	2
	C3	2	3	1
	C4	2	3	1
D	D1	1	3	2
	D2	1	3	2
E	E1	1	2	3
	E2	1	2	3
	E3	3	2	1
F	F1	1	3	2
	F2	1	2	3
G	G1	2	3	1
	G2	2	3	1
	G3	2	3	1
	G4	1	3	2
	G5	1	2	3
	G6	2	1	3
H	H1	1	3	2
	H2	1	3	2
	H3	1	3	2
TOTAL		69	95	58

TABLE XIV

Visual Inspection of Histograms for Airports

Airport	Year	No. of Flights	RANKING (1=BEST FIT, 3=WORST FIT) Normal	Poisson	Lognormal
A	1981	2	2	3	1
B	1978	3	2	3	1
	1979	3	3	1	2
	1980	3	2	3	1
	1981	3	2	3	1
	1982	3	2	3	1
C	1980	2	2	3	1
	1981	2	2	3	1
D	1980	1	1	3	2
	1981	1	1	3	2
E	1979	1	1	2	3
	1980	1	1	2	3
	1982	1	3	2	1
F	1980	1	1	3	2
	1981	1	1	2	3
G	1981	3	3	2	1
	1982	3	3	2	1
H	(incomplete)				
		TOTAL	32	43	27

One other note should be made about the models. The Normal distribution is able to handle negative values, which cannot occur in the real data. This causes a distortion of the expected frequencies in the first left-hand interval. Since the Poisson and lognormal can handle positive values only, they are more representative of the lower frequency classes.

Visual inspection can only indicate preferences among the models and any gross tendencies of the data. It cannot be used as an independent criterion for acceptance or rejection.

4.1.4 Ease of Use

The Normal distribution is the most prevalent of the three distributions. It is a common assumption made by analysts that the data they are dealing with fits a Normal pattern. Because of the familiarity and general understanding, the Normal is easy to use.

The Poisson and lognormal are less common but are still known and understood by most engineers. All three distributions have parameters which are simple to determine and are tabulated in most texts, although the Poisson is not usually calculated for parameters greater than twenty.

If standard statistics computer programs are to be used, the Normal distribution is easier to find curve-fitting routines for. In this study, the data had to be scaled down by a factor of ten to use the Poisson program and a separate program had to be written for the lognormal. If an extensive program library is available, of course, this problem will not arise.

4.1.5 Applicability

In terms of applicability to small airport design, the prediction of the planning volume is the most important feature of any distribution. The normal and lognormal did this acceptably well for simple distributions.

More complicated calculations may need to be done to determine the design volume under changing conditions. For example, as traffic increases, the upper tail of the distribution will be limited by the aircraft capacity. A truncated curve will then have to be used. If the flight route has several stops, the passenger volumes of all the aircrafts will have to be added, and then this distribution truncated. A distribution that was additive regenerative would simplify this calculation.

Also, if the function were additive (the Normal and Poisson are), it could be assumed that if the total of the deplaned and enplaned passengers followed the distribution, that both each separately would be distributed according to the same function. Deplaned passenger distributions and enplaned passenger distributions could be determined.

4.2 Selection of a Model

Of the three distributions considered, none describes the data in all situations. The lognormal is, however, the preferred model - followed by the Normal.

As measured by the Chi-Squared statistic, the lognormal provided a good fit more often. None of the models were good at describing airports with multiple flights. The percentages of acceptance do not indicate whether a model should be taken for use in all cases. The lognormal was a good fit for 41 percent of the flights and 33 percent of the airports (at a

.005 level of significance), but no deduction can be made as to whether or not these percentages are sufficient to unconditionally use the lognormal model. The decision remains judgemental.

With reference to tail predictability, both the Normal and lognormal perform reasonably well. Use of the Poisson would lead to serious errors.

Visual inspection suggests that the lognormal is the better model, followed again by the Normal.

The fourth criterion - that the model be easy to use - would lead to the selection of the Normal. There are no serious complications, however, with the use of the other two.

Finally, the criterion that the model be applicable would indicate that it be additive regenerative. Only the Normal and Poisson are.

The original hypothesis of this work was that a statistical model could be found that would approximate the data well enough for use in the sizing of small airports. If a model is to be selected, it would be the lognormal, although if an additive quality was required, the normal would have to be used. The Poisson distribution can be discarded according to most of the criteria.

Although a selection was made, there is some doubt as to whether or not any of the three distributions is satisfactory. If only planning design volumes are to be determined from collected data, either the lognormal or the normal is adequate. However, any analysis that requires use of the entire distribution should consider other models - perhaps one of the six listed in the beginning of this study. This is especially true if airports with multiple flights are under analysis.

5. CONCLUSIONS

5.1 Assessment

The final decision to use a statistical model has to ultimately depend upon professional judgement. The study illustrates that using distributions for different purposes can result in the selection of different models for each purpose.

The lognormal provided the best model overall, although it had drawbacks. Some flights were better described by other distributions. Also, the lognormal did not have the additive feature of the Normal and Poisson. Nonetheless, the study had revealed some of the strengths and weaknesses of the three.

The scope of the study can be categorized in three areas: the number of models considered; the data itself; and the computer programs used.

Three distribution models were studied. The Chi-squared test measured overall "goodness-of-fit" and the 90th percentile test measured tail predictability.

The inherent assumption in the entire approach was that the statistical models assumed that the data was random, when actually the number of passengers choosing a particular flight depends on the complex interaction of many variables. The models incorporate these unknown forces as randomness.

The data itself has particular features which simplify the study. First all of the airports were in British Columbia. Air transportation in this province has certain unique and consistent features. For example, there are relatively few towns and these are typically separated by mountain highways or waterways. Therefore, air travel is more common than in other provinces. Also, most air traffic funnels through Vancouver or Calgary/Edmonton. The majority of flight routes originate and terminate at these cities. For example, a Canadian Pacific flight follows a triangular Vancouver-Terrace-Prince Rupert-Vancouver route since neither Terrace nor Prince Rupert can generate sufficient demand to warrant a single stop. Since there is virtually no demand between Prince Rupert and Terrace, it can be safely assumed that all enplaners at Terrace are bound for Vancouver, and that all of the deplaners at Prince Rupert came from Vancouver.

Another feature of these flights is that they are all served by Boeing 737 aircraft which have a capacity of 117 seats. This situation has evolved because the carriers have found the Boeing 737 to be the most suitable aircraft for the region, although this may change in the future. Furthermore, at these airports, flight events are isolated throughout the day. This means that there is no overlapping in the use of the facilities.

Lastly, the scope of the research was defined by the computer programs used. A statistical packaged program was used to calculate the Chi-squared values and to plot the histograms for the Normal annd Poisson distributions. It was found to be too restrictive for the Lognormal, however, and a separate program had to be written. Budgetary considerations limited the extent of the analysis in this regard.

The scope - as defined by the models, the nature of the data, and the computer programs - did not seriously hinder the process. There is no evidence that the use of a cumulative distribution form or a different statistical measure (such as the Kolmogorov-Smirnov) would have significantly changed the results.

The results of the design volume analysis were consistent enough to allow the conclusion that the Normal and Lognormal are satisfactory models.

The purpose of this study was to further the planning of small airport facilities. Once the correct distribution is known, it can be used directly in, say, a Monte Carlo simulation where passenger loads are randomly sampled from the distribution. The simulation would then produce expected occupancies for facility sizing. The distribution can also produce specific planning volumes (hourly or half-hourly). Facilities are then sized from a method of proportions or from a selection process such as Transport Canada's STEP.

A frequency model would also be needed to determine more complicated effects on the airport passenger volumes. It can quantify the effects of route changes and aircraft capacity.

5.2 Further Research

The possibilities for further work are numerous. Little research has been done in the field of smalll airports for several reasons:

(1) Carriers are not required to submit data by flight or by day to the government;

(2) detailed study was never considered necessary since aircraft were small, and facilities could be incrementally adjusted;

(3) air carriers are reluctant to release detailed information to competitors; and

(4) small airports have not been deemed as important as larger ones when research was to be done.

Increased reliance on a single design volume, as well as less the reduced availability of construction capital, may change this situation. Further research might include:

(1) Consideration of other statistical models (Gamma, Weibull, Rayleigh, Negative Binomial, Beta, et cetera);

(2) calculation and comparison of the 90th percentile by passenger volume as a design volume;

(3) categorization of airports by parameters or by distribution type;

(4) calculations of the effects of multi-stop flight routes with respect to aircraft capacity;

(5) calculation and comparison of other percentiles (75th, 80th, 85th);

(6) using a cumulative distribution form and a Kolmogorov-Smirnov goodness-of-fit measure;

(7) consideration of the effects of a trend to smaller aircraft, especially the Dash 7 in British Columbia;

(8) derivation of demand distributions from the measured load distributions;

(9) analysis of the costs of errors in forecasting the design volumes on all facilities; and

(10) a network analysis for British Columbia air traffic.

REFERENCES

1. Ashford, Norman and Wright, Paul H. Airport Engineering. John Wiley & Sons, U.S.A., 1979.

2. Benjamin, Jack, R. and Cornell, C. Allin. Probability, Statistics and Decision for Civil Engineers. McGraw-Hill Book Co., U.S.A. 1970.

3. Braaksma, John P. A Computerized Design Method for Preliminary Airport Space Planning. Technical Report, Department of Civil Engineering, University of Waterloo, Waterloo, Ontario. October 1973.

4. Braaksma, John P. Time Stamping: a new way to survey pedestrian traffic in airports. Traffic Engineering and Control, Volume 17, Number 5, May 1976, pp. 204-206.

5. Braaksma, John P, and Ramsay, W. Alex. Air Terminal Design: Decentralization and Shape. Transportation Engineering Journal, ASCE, Volume 105, TE6, November 1979, pp. 669-714.

6. Bury, Karl V. Statistical Models in Applied Science. John Wiley and Sons, Inc., U.S.A., 1975.

7. Chatfield, Christopher. Statistics for Technology. Chapman and Hall Limited, London, 1976.

8. de Neufville, Richard and Grillot, Michel. Design of Pedestrian Space in Airport Terminals. Transportation Engineering Journal, ASCE, Volume 108, TE1, January 1982, pp. 87-101.

9. de Neufville, Richard and Rusconi-Clerici Ignazio. Designing Airport Terminals for Transfer Passengers. Transportation Engineering Journal, ASCE, Volume 104, TE6, November 1978, pp. 775-787.

10. Dunlay, William, J., Jr. Simulation Model Validation: airport applications. Transportation Engineering Journal, ASCE, Volume 107, TE4, July 1981, pp. 401-412.

11. Freund, John E. and Williams, Frank, J. Elementary Business Statistics: the modern approach. Prentice-Hall, Inc., New Jersey, U.S.A., 1972.

12. Horonjeff, Robert. Planning and Design of Airports, Second Edition. McGraw-Hill Book Company, U.S.A., 1975.

13. International Air Transport Association. Airport Terminals Reference Manual, Sixth Edition.

14. Lauchli, U. Matching Transportation Capacity to a Stochastic Transportation Demand. Proceedings of Sixteenth Annual AGIFORS Symposium, Florida, 1976.

15. Ramsey, G.R. Stuart and Hutchinson, B.G. An Airport Terminal Pasenger Flow Model. Paper Presented at 63rd Annual Meeting of the Transportation Research Board, Washington, D.C., January 1984.

16. Soumis, Francois, Ferland, Jacques-A. and Rousseau, Jean-Marc. MAPUM A Model for Assigning Passengers to a Flight Schedule. Université de Montréal, Centre de recherche sur les transports - #142, August 1979.

17. Transport Canada, Air. Airport Traffic Allocation Model, Manual. Document AK 41-02-300, Airports Branch.

18. Transport Canada, Air. CATA's National Aviation Forecasting Models and Other Forecasting Methods. TP #2046. Policy, Planning and Programming, March 1979.

19. Transport Canada, Air. Study of the Air Terminal Building Planning Standard. TP 4771E. (Salah G. Hamzawi). Policy, Planning and Programming, July 1983.

20. Transport Canada, Air. Systemized Terminal Expansion Program (STEP), Draft. Document AK 62-08-000, Airports Branch, September 1983.

21. Transport Canada, Air. Vancouver Airport Simulation Model User's Manual. (E.J.M. Greco and G.R. Haverty), Airports Branch, September 1974.

22. Vella John, Jaul Martin and Whale Les. Determining the Best Aircraft Seat Configuration. Proceedings of Twenty-First Annual AGIFORS Symposium, California, 1981.

23. Walpole, Ronald E. and Myers, Raymond H. Probability and Statistics for Engineers and Scientists, 2nd Edition. MacMillan Publishing Co., Inc., New York, 1978.

24. Wang, Ken. Modelling the Interaction between Payload Restriction, Passenger Demand and Reservation Booking Levels. Proceedings of Twenty-Second Annual AGIFORS Symposium, Greece, 1982.

APPENDIX A

COMPUTER PROGRAMS

APL PACKAGE PROGRAM TO COMPARE DATA TO NORMAL AND POISSON DISTRIBUTIONS (I.P. SHARP & ASSOCIATES LTD.)

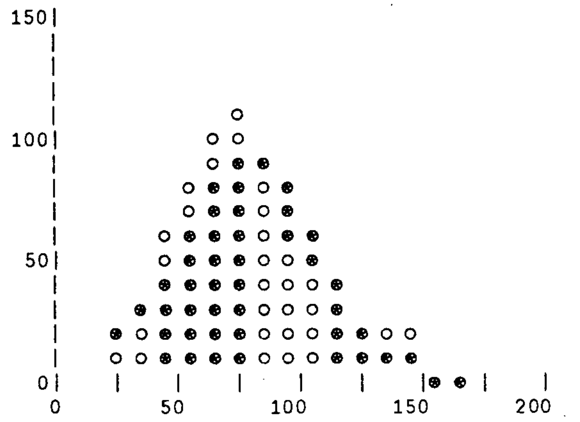
```
      FREQ A1981
ENTER THE FOLLOWING DATA.
LEFT HAND END OF THE FIRST FREQUENCY CLASS;YOUR DATA MIN=26
⎕:
      30
CLASS  WIDTH  AND  THE  NUMBER  OF CLASSES;YOUR DATA MAX=164
⎕:
      10 13


DO YOU WISH A FIT DONE ON YOUR DATA? Y OR N
Y
NORMAL OR POISSON ? N OR P
N

(1)  DATA MEAN = 79.46307692 AND STANDARD DEV. = 27.90987073

DO YOU WISH A HISTOGRAM ? Y OR N
Y
DO YOU WISH TABULAR OUTPUT  ? Y OR N
Y
```

-ENDPOINTS-

L	R	MID	OBS	°/°	EXP	°/°
20	30	25.0	6	0.9	24.666	3.8
30	40	35.0	18	2.8	26.485	4.1
40	50	45.0	63	9.7	43.617	6.7
50	60	55.0	76	11.7	62.998	9.7
60	70	65.0	101	15.5	80.890	12.4
70	80	75.0	106	16.3	91.276	14.0
80	90	85.0	84	12.9	90.802	14.0
90	100	95.0	51	7.8	79.188	12.2
100	110	105.0	39	6.0	60.912	9.4
110	120	115.0	42	6.5	41.614	6.4
120	130	125.0	20	3.1	24.895	3.8
130	140	135.0	23	3.5	12.938	2.0
140	150	145.0	15	2.3	5.893	0.9
150	160	155.0	4	0.6	2.408	0.4
160	170	165.0	2	0.3	1.420	0.2

```
TOTAL OBSERVATIONS          650
CHI-SQUARE              78.0978
```

OBSERVED- ○ : EXPECTED- ⊛ : VARIABLE (COLUMN NO.) 1



APL PROGRAM TO COMPARE DATA TO LOGNORMAL DISTRIBUTION

```
      ∇LN[⎕]∇
    ∇ LN ω;LOGMEAN;LOGSD;LH;RH;RHO;START;END;FRQ;PROB
[1]   LOGMEAN←MEAN(10⍟ω)
[2]   LOGSD←STDDEV(10⍟ω)
[3]   MS←LOGMEAN,LOGSD
[4]   ' LEFT HAND END '
[5]   LH←⎕
[6]   ' RIGHT HAND END '
[7]   RH←⎕
[8]   VECTOR←LH,LH+(10×⍳((RH-LH)÷10))
[9]   LOGVECTOR←10⍟VECTOR
[10]  PROB←(⌊(0.5+(MS NORMALPROB LOGVECTOR)×1000))÷1000
[11]  RHO←⍴PROB
[12]  START←(LH-10),VECTOR
[13]  END←VECTOR,(RH+10)
[14]  FRQ←PROB×⍴ω
[15]  EXP←LOGVECTOR CLASSIFY(10⍟ω)
[16]  DEL←((FRQ-EXP)*2)÷FRQ
[17]  DEL←(⌊(0.5+DEL×1000))÷1000
[18]  ' '
[19]  ' LH   RH   PROB    EXP   OBS     X'
[20]  ' '
[21]  SC←⍕((RHO,1)⍴START)
[22]  EC←⍕((RHO,1)⍴END)
[23]  PC←⍕((RHO,1)⍴PROB)
[24]  FC←⍕((RHO,1)⍴FRQ)
[25]  ECHAR←⍕((RHO,1)⍴EXP)
[26]  DC←⍕((RHO,1)⍴DEL)
[27]  SC,' ',EC,' ',PC,' ',FC,' ',ECHAR,' ',DC
[28]  ' ' ◇ ' '
[29]  CS←+/DEL
[30]  'CHI-SQUARED  =  ',⍕(CS)
[31]  ' ' ◇ 'TOTAL OBSERVATIONS =  ',⍕(⍴ω)
[32]  ' ' ◇ ' '
[33]  EXP HISTO FRQ
[34]  ' '
    ∇
```

... Subroutines HISTO and CLASSIFY on next page.

SUBROUTINES 'HISTO' AND 'CLASSIFY' USED IN APL PROGRAM 'LN'

```
      ∇HISTO[⎕]∇
    ∇ R←EXP HISTO OBS;⎕IO;H;HH;TEST;B001;B002;LBL
[1]  ⍝FOR EXPECTED AND OBSERVED VALUES
[2]   ⎕IO←0
[3]   H←⌈/EXP,OBS ◇ SCL←10*⌊10⍟H ◇ HH←SCL× 1 1.5 2 3 4 5 6 8 10
[4]   H←⌊/(HH≥H)/HH
[5]   TEST←(⍳20)×H÷20 ◇ B001←TEST∘.≤EXP ◇ B002←TEST∘.<OBS
[6]   R←⊖' ○*⊛'[B001+2×B002]
[7]   R←((4×1↓⍴R)⍴ 0 0 0 1)\R ⍝SPACE IT OUT A BIT
[8]   R←R,[0] '-'
[9]   LBL←⍕ 3 1 ⍴H× 1 0.5 0 ◇ LBL←(21⍴10↑1)⍀LBL ◇ R←LBL,R
[10]  LBL←'EXPECTED: *  OBSERVED: ○'
[11]  ' ' ◇ ' ' ◇ ' ' ◇ ' '
[12]  W←(⍴LBL)⌈1↓⍴R ◇ R←(W↑LBL),[0]((1↑⍴R),W)↑R
    ∇
```

```
      ∇CLASSIFY[⎕]
    ∇ R←α CLASSIFY ω
[1]   R←α∘.≤ω
[2]   R←(1,[1] R)-R,[1] 0
[3]   R←+/R
    ∇
[4]   ∇
```

APL PROGRAM TO CALCULATE EXPECTED DECILES FOR NORMAL AND LOGNORMAL DISTRIBUTIONS

```
    ∇TEN[⎕]∇
    ∇ R←TEN ω;M;S;Z
[1]   M←MEAN ω
[2]   S←STDDEV ω
[3]   Z←GAUSS 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
[4]   R←⌊(0.5+(10*(M+(S×Z))))
    ∇
```

(Line 4 without '10*' for Normal Distribution)

FORTRAN PROGRAM TO CALCULATE EXPECTED POISSON 90TH PERCENTILE

```
Listing of POI at 13:34:50 on MAR  2, 1984 for CCid=FCE6

C
C234567
      REAL M,PROBRR,RR
      DIMENSION R(20),P(20)
      REAL Z,ZZ,F,D,DD
      RR=100.0
      M=64.9
C
      R(1)=RR
C
      DO 100 I=2,20,1
         R(I)=R(I-1)+1.0
100   CONTINUE
C
      F=0.0
      PROBRR=1.0
C
      DO 200 J=1,200,1
       IF (J.GE.RR) GO TO 210
       F=F+ALOG10(FLOAT(J))
       PROBRR=PROBRR+10**(FLOAT(J)*ALOG10(M)-F)
200   CONTINUE
210    CONTINUE
      PROBRR=PROBRR*EXP(-1.0*M)
C
      P(1)=PROBRR
      DO 300 K=2,20
       F=F+ALOG10((RR+FLOAT(K)-2.0))
       P(K)=P(K-1)+EXP(-1.*M)*10**(R(K-1)*ALOG10(M)-F)
300   CONTINUE
C
      WRITE(6,15)J
15    FORMAT('J WAS :',I4)
C
      WRITE(6,18)PROBRR
18    FORMAT('PROBRR WAS :',F9.7)
      WRITE(6,13)M
13    FORMAT('PARAMETER :',F6.2)
      WRITE (6,10)
10    FORMAT('  X       +/PROB')
      WRITE(6,12)
12    FORMAT('---------------')
      DO 400 I=1,20,1
         WRITE(6,11)R(I),P(I)
11       FORMAT(1X,F4.0,3X,F5.3)
400   CONTINUE
C
      STOP
      END
```

OUTPUT
(Sample)

```
PROBRR WAS :0.5653268
PARAMETER :117.90
  X       +/PROB
---------------
 120.   0.565
 121.   0.601
 122.   0.636
 123.   0.670
 124.   0.702
 125.   0.733
 126.   0.762
 127.   0.789
 128.   0.814
 129.   0.837
 130.   0.858
 131.   0.877
 132.   0.895
 133.   0.910
 134.   0.924
 135.   0.936
 136.   0.946
 137.   0.956
 138.   0.963
 139.   0.970
```

APPENDIX B

DETAILED CHI-SQUARED CALCULATIONS

APPENDIX B

TABLE BI

χ^2 CALCULATON FOR THE NORMAL DISTRIBUTION BY FLIGHT

Airport	Flight	PARAMETERS		INITIAL		CORRECTED	
		μ	σ	NUMBER OF DIVISIONS	χ^2	NUMBER OF DIVISIONS	χ^2
A	A1	64.9	20.4	13	78.8	10	19.6
	A2	91.3	27.6	14	29.6	13	29.5
B	B1	92.2	32.0	20	80.0	14	48.6
	B2	114.0	32.1	20	47.5	14	23.0
	B3	124.8	32.3	20	26.9	15	22.3
	B4	107.3	26.6	14	30.3	12	28.1
	B5	94.9	24.5	16	25.5	11	12.9
	B6	68.8	30.0	18	70.1	14	64.0
	B7	85.5	36.7	18	83.5	15	68.9
	B8	95.2	37.9	21	61.7	16	49.5
	B9	92.2	37.0	18	59.7	14	56.4
	B10	96.7	35.0	20	24.7	14	15.6
	B11	115.6	36.1	19	41.8	17	38.7
	B12	117.9	34.5	16	46.1	16	46.1
	B13	125.9	41.5	19	46.8	18	45.6
	B14	56.2	34.6	19	363.8	12	213.3
	B15	105.4	31.1	19	38.7	15	32.8
C	C1	46.3	14.9	10	46.7	7	21.8
	C2	61.3	17.0	11	29.4	8	12.7
	C3	41.7	12.4	8	17.7	6	10.2
	C4	55.5	16.1	9	53.3	8	53.0
D	D1	62.3	21.1	11	29.7	9	5.1
	D2	67.8	20.7	12	15.8	10	13.3
E	E1	101.4	27.7	20	51.4	11	21.8
	E2	116.8	33.3	18	17.4	16	16.4
	E3	84.8	28.7	20	305.7		*
F	F1	107.4	28.8	20	42.5	14	4.9
	F2	112.9	33.0	27	310.7	15	41.0
G	G1	40.7	17.6	12	658.8		*
	G2	81.7	22.5	18	378.8		*
	G3	142.6	38.3	17	20.1	17	20.1
	G4	42.6	17.0	10	21.6	8	12.7
	G5	79.9	25.1	20	121.6	12	9.9
	G6	132.8	39.4	22	25.4	17	21.8
H	H1	86.7	24.3	17	21.2	12	7.0
	H2	87.4	22.7	14	10.7	11	9.1
	H3	74.6	21.1	12	21.5	11	21.5

* Significant figures of the program do not allow calculatons of Chi-squared.

APPENDIX B

TABLE BII

χ^2 CALCULATON FOR THE POISSON DISTRIBUTION BY FLIGHT

Airport	Flight	PARAMETER μ	INITIAL		CORRECTED	
			NUMBER OF DIVISIONS	χ^2	NUMBER OF DIVISIONS	χ^2
A	A1	64.9	13	40.1	11	39.2
	A2	91.3	14	29.3	14	29.3
B	B1	92.2	20	56.0	14	43.3
	B2	114.0	20	25.1	15	22.2
	B3	124.8	20	21.1	15	22.0
	B4	107.3	14	40.8	14	40.8
	B5	94.9	16	29.8	13	29.6
	B6	68.8	18	62.8	12	57.1
	B7	85.5	18	124.0	13	97.7
	B8	95.2	21	93.0	13	72.0
	B9	92.2	18	93.9	14	89.5
	B10	96.7	20	26.2	13	19.6
	B11	115.6	19	61.6	16	48.7
	B12	117.9	16	43.5	15	43.5
	B13	125.9	19	73.1	16	69.6
	B14	56.2	19	1636.6		*
	B15	105.4	19	31.3	14	30.4
C	C1	46.3	10	76.9	10	76.9
	C2	61.3	11	50.5	10	50.4
	C3	41.7	8	88.8	8	88.8
	C4	55.5	9	92.0	9	92.1
D	D1	62.3	11	27.3	11	27.3
	D2	67.8	12	35.5	11	34.1
E	E1	101.4	20	39.5	13	37.1
	E2	116.8	18	15.6	16	14.7
	E3	84.8	20	72.7	13	24.6
F	F1	107.4	20	29.7	15	20.3
	F2	112.9	27	117.6	15	27.2
G	G1	40.7	12	61.9	13	*
	G2	81.7	18	60.6	16	55.2
	G3	142.6	17	19.0		18.1
	G4	42.6	10	24.4	9	24.1
	G5	79.9	20	43.6	12	29.0
	G6	132.8	22	56.8	17	36.6
H	H1	86.7	17	28.3	14	27.7
	H2	87.4	14	39.2	14	39.2
	H3	74.6	12	45.9	12	45.9

* Significant figures of the program do not allow calculatons of Chi-squared.

APPENDIX B

TABLE BIII

χ^2 CALCULATON FOR THE LOGNORMAL DISTRIBUTION BY FLIGHT

Airport	Flight	PARAMETERS		INITIAL		CORRECTED	
		μ LOG	σ LOG	NUMBER OF DIVISIONS	χ^2	NUMBER OF DIVISIONS	χ^2
A	A1	1.79	.14	13	15.6	10	14.5
	A2	1.94	.14	14	20.7	12	19.7
B	B1	1.94	.14	18	31.8	12	22.3
	B2	2.04	.12	20	10.1	14	3.4
	B3	2.08	.11	20	9.4	15	7.7
	B4	2.02	.11	14	20.4	12	19.6
	B5	1.96	.11	16	9.1	11	4.1
	B6	1.80	.19	17	18.3	14	16.5
	B7	1.89	.19	18	23.6	15	21.5
	B8	1.94	.18	21	24.1	15	16.7
	B9	1.93	.18	18	23.4	15	18.0
	B10	1.96	.16	20	14.7	12	9.2
	B11	2.04	.15	18	56.0	16	52.9
	B12	2.05	.13	16	34.3	15	34.0
	B13	2.07	.15	19	45.9	17	43.1
	B14	1.69	.22	19	77.7	12	53.4
	B15	2.00	.13	18	26.4	12	23.8
C	C1	1.64	.14	10	6.9	9	6.9
	C2	1.77	.12	11	15.2	8	14.1
	C3	1.60	.13	8	4.1	6	3.7
	C4	1.73	.12	9	9.1	9	9.1
D	D1	1.77	.15	11	11.2	11	11.2
	D2	1.81	.15	12	43.9	11	36.8
E	E1	1.99	.13	20	67900.	16	18.2
	E2	2.05	.13	18	19.0	15	14.9
	E3	1.90	.14	20	11.3	13	2.8
F	F1	2.01	.13	20	486.	14	34.6
	F2	2.03	.14	27	21300.	15	41.1
G	G1	1.57	.20	14	26.5	9	14.4
	G2	1.89	.13	18	1110.	11	31.4
	G3	2.14	.12	17	16.1	16	16.1
	G4	1.59	.20	10	32.7	9	16.4
	G5	1.88	.15	20	34300.	13	64.3
	G6	2.10	.14	22	809.	17	35.8
H	H1	1.92	.13	17	32.6	12	26.0
	H2	1.93	.12	14	24.6	11	24.2
	H3	1.85	.13	12	28.1	11	28.1

APPENDIX B

TABLE BIV

χ^2 CALCULATON FOR THE NORMAL DISTRIBUTION BY AIRPORT (AIRPORTS WITH MULTIPLE FLIGHTS)

Airport	Year	PARAMETERS		INITIAL		CORRECTED	
		μ	σ	NUMBER OF DIVISIONS	χ^2	NUMBER OF DIVISIONS	χ^2
A	1981	79.5	27.9	15	78.1	13	75.8
B	1978	94.7	39.3	21	94.2	20	92.6
	1979	105.4	37.6	22	55.1	19	53.9
	1980	110.3	42.7	21	82.4	20	82.0
	1981	84.9	39.5	20	133.9	18	132.6
	1982	99.8	30.5	20	73.1	16	55.4
C	1980	51.9	17.2	12	60.1	9	34.7
	1981	49.0	16.0	10	94.7	8	52.8
G	1981	86.6	49.4	23	276.2	22	270.7
	1982	85.0	46.6	24	199.0	21	194.1

APPENDIX B

TABLE BV

χ^2 CALCULATON FOR THE POISSON DISTRIBUTION BY AIRPORT (AIRPORTS WITH MULTIPLE FLIGHTS)

Airport	Year	PARAMETER μ	INITIAL NUMBER OF DIVISIONS	INITIAL χ^2	CORRECTED NUMBER OF DIVISIONS	CORRECTED χ^2
A	1981	79.5	15	58.5	15	58.5
B	1978	94.7	21	255.4	16	240.4
	1979	105.4	22	122.6	17	120.3
	1980	110.3	21	226.1	17	210.2
	1981	84.9	20	535.7	16	528.5
	1982	99.8	20	49.3	16	46.6
C	1980	51.9	12	78.2	11	78.2
	1981	49.0	10	131.3	10	131.3
G	1981	86.6	23	3411.9	16	2228.7
	1982	85.0	24	2161.3	16	1488.9

APPENDIX B

TABLE BVI

χ^2 CALCULATON FOR THE LOGNORMAL DISTRIBUTION BY AIRPORT

(AIRPORTS WITH MULTIPLE FLIGHTS)

Airport	Year	PARAMETERS		INITIAL		CORRECTED	
		LOG	LOG	NUMBER OF DIVISIONS	χ^2	NUMBER OF DIVISIONS	χ^2
A	1981	1.87	.15	15	31.2	14	28.5
B	1978	1.94	.20	21	1670.	19	71.5
	1979	1.99	.17	22	62.1	20	59.1
	1980	2.01	.18	21	58.0	20	52.7
	1981	1.88	.23	20	109.4	19	109.4
	1982	1.99	.14	20	43.2	16	28.2
C	1980	1.69	.15	12	9.4	10	9.2
	1981	1.67	.14	10	9.6	9	9.3
G	1981	1.86	.28	23	103.1	22	99.0
	1982	1.86	.27	24	91.0	23	75.9

APPENDIX C

ACTUAL DECILES BY FLIGHT EVENT

APPENDIX C

TABLE CI

ACTUAL DECILES BY FLIGHT EVENT

AIRPORT	FLIGHT	10th	20th	30th	40th	50th	60th	70th	80th	90th
A	A1	42	47	53	58	63	69	74	80	89
	A2	58	66	74	80	88	96	105	118	132
B	B1	59	67	75	80	84	88	97	113	141
	B2	76	88	95	101	109	118	126	141	157
	B3	86	99	106	115	123	128	136	149	168
	B4	76	83	89	96	105	113	123	131	141
	B5	67	74	81	86	92	98	106	114	129
	B6	38	44	49	56	63	70	77	90	119
	B7	46	54	61	67	77	87	101	115	143
	B8	52	62	71	77	89	103	111	128	151
	B9	49	59	66	75	86	98	110	124	146
	B10	56	68	74	84	93	101	111	127	144
	B11	69	81	92	102	115	127	138	150	163
	B12	76	87	94	103	114	124	136	152	169
	B13	75	94	103	109	120	131	147	164	188
	B14	26	32	36	41	44	48	55	74	106
	B15	70	77	84	93	100	111	122	135	151
C	C1	29	35	38	41	44	48	52	58	66
	C2	40	46	51	57	61	64	68	74	83
	C3	28	31	34	37	41	44	47	51	58
	C4	37	42	46	49	53	57	61	68	78
D	D1	38	44	50	54	60	65	71	81	91
	D2	42	51	56	61	67	73	78	84	95
E	E1	69	77	87	92	98	106	115	122	136
	E2	73	85	97	105	115	124	134	146	163
	E3	51	61	67	74	81	89	94	105	122
F	F1	72	85	93	100	108	116	121	131	144
	F2	78	89	97	102	109	116	125	136	154
G	G1	22	28	31	34	37	40	46	54	61
	G2	56	65	72	76	81	86	91	99	109
	G3	95	106	119	129	140	151	166	178	198
	G4	22	27	33	37	41	45	50	57	65
	G5	45	61	68	74	80	75	93	100	110
	G6	79	99	111	121	130	141	156	165	190
H	H1	54	65	73	80	86	93	100	108	116
	H2	57	67	75	81	86	93	100	107	115
	H3	48	59	64	68	73	79	84	91	102

APPENDIX C

TABLE CII

ACTUAL DECILES BY FLIGHT EVENT

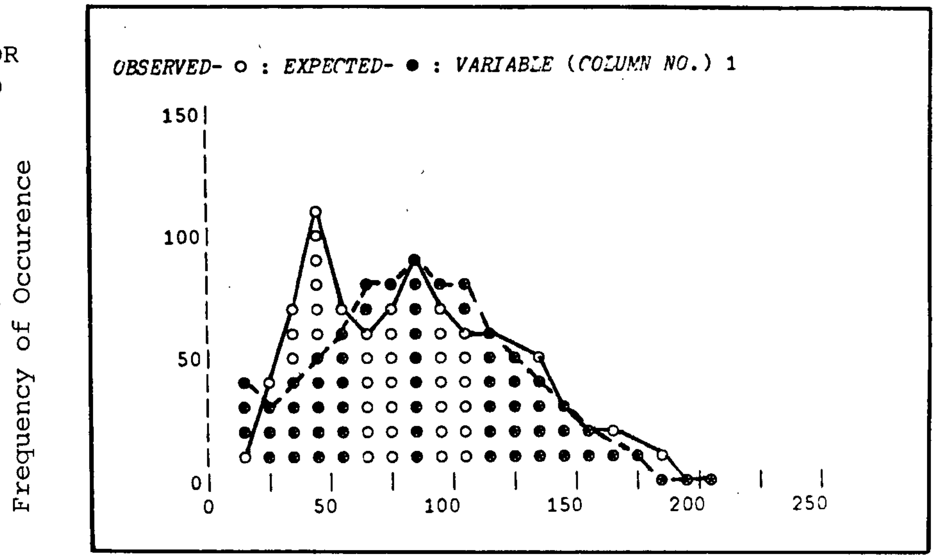
AIRPORT	YEAR	PERCENTILES BY EVENT 10th	20th	30th	40th	50th	60th	70th	80th	90th
A	1981	47	55	62	69	75	81	90	104	119
B	1978	47	60	70	79	87	98	115	132	155
	1979	60	71	84	92	101	111	122	139	160
	1980	58	71	81	97	106	117	130	147	174
	1981	36	46	56	70	82	93	106	121	139
	1982	67	74	81	88	95	103	113	126	141
C	1980	32	37	41	45	50	54	60	66	74
	1981	30	36	40	43	47	51	55	60	70
D	1979	38	44	50	54	60	65	71	81	91
	1980	42	51	56	61	67	73	78	84	95
E	1979	69	77	87	92	98	106	115	122	136
	1980	73	85	97	105	115	124	134	146	163
	1981	51	61	67	74	81	89	94	105	122
F	1981	72	85	93	100	108	116	121	131	144
	1982	78	89	97	102	109	116	125	136	154
G	1981	31	39	53	67	79	90	105	128	162
	1982	31	42	55	66	77	90	104	122	156

APPENDIX D

HISTOGRAMS

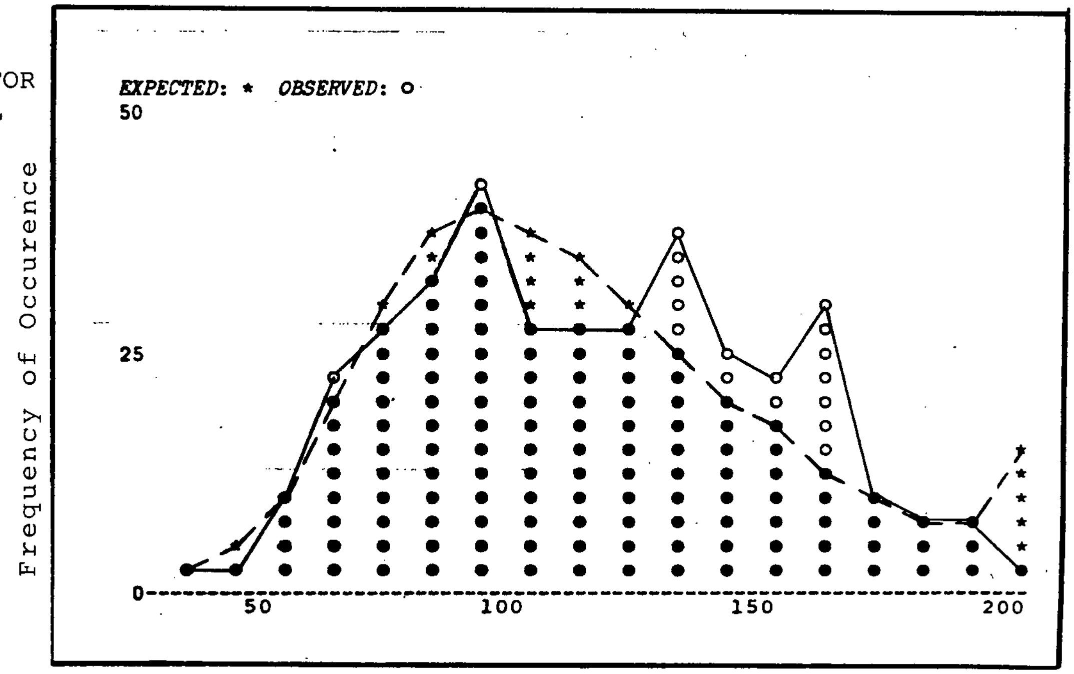
KEY TO HISTOGRAMS

EXAMPLE FOR
NORMAL AND
POISSON



NUMBER OF ENPLANED + DEPLANED PASSENGERS

EXAMPLE FOR
LOGNORMAL



NUMBER OF ENPLANED+DEPLANED PASSENGERS

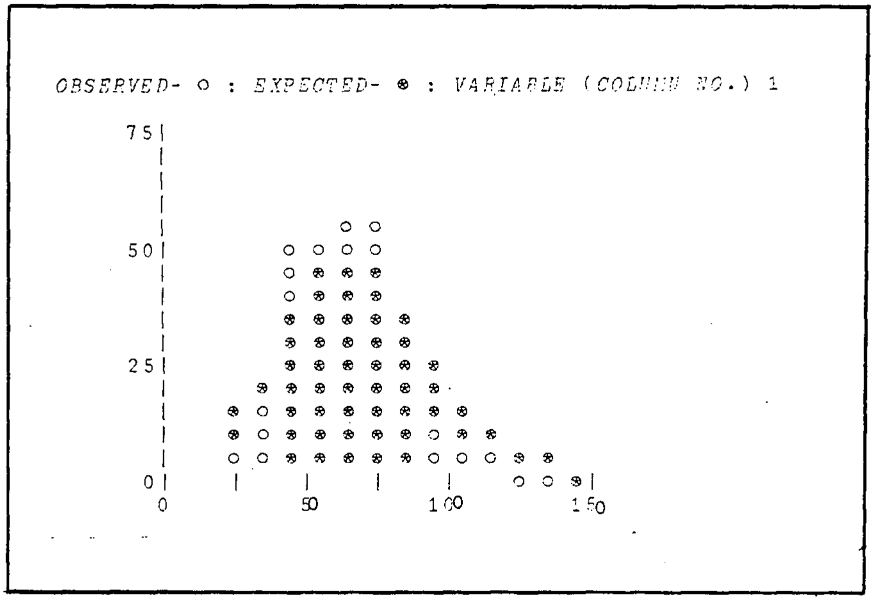
— — Expected Frequency Distribution

—— Observed Frequency Distribution

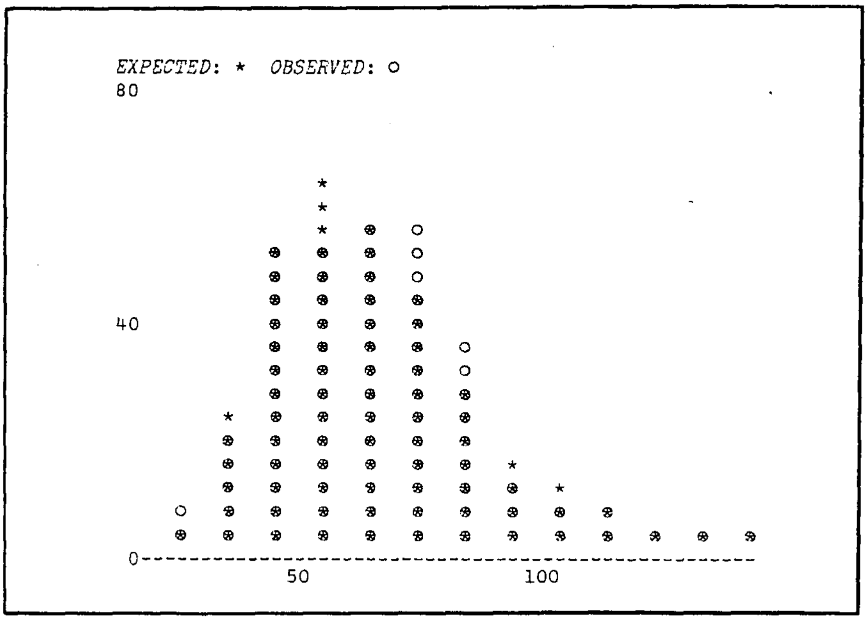
AIRPORT A, FLIGHT A1

(Normal missing)

POISSON

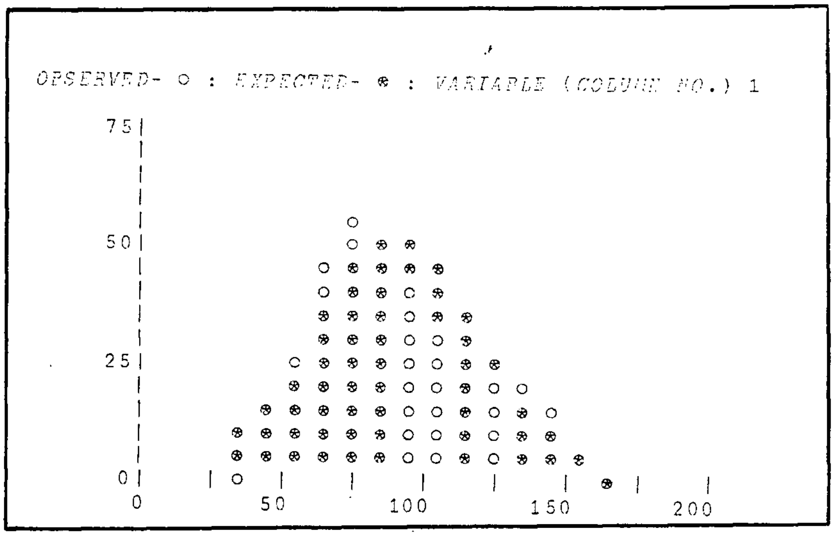


LOGNORMAL

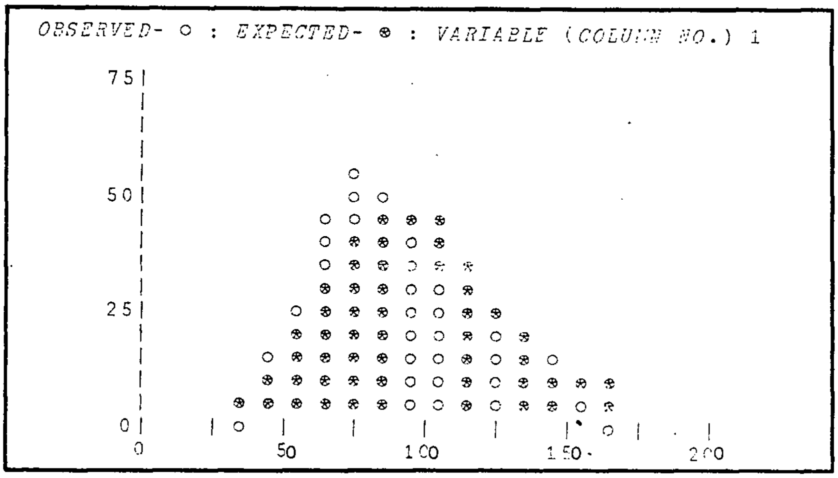


AIRPORT A, FLIGHT A2

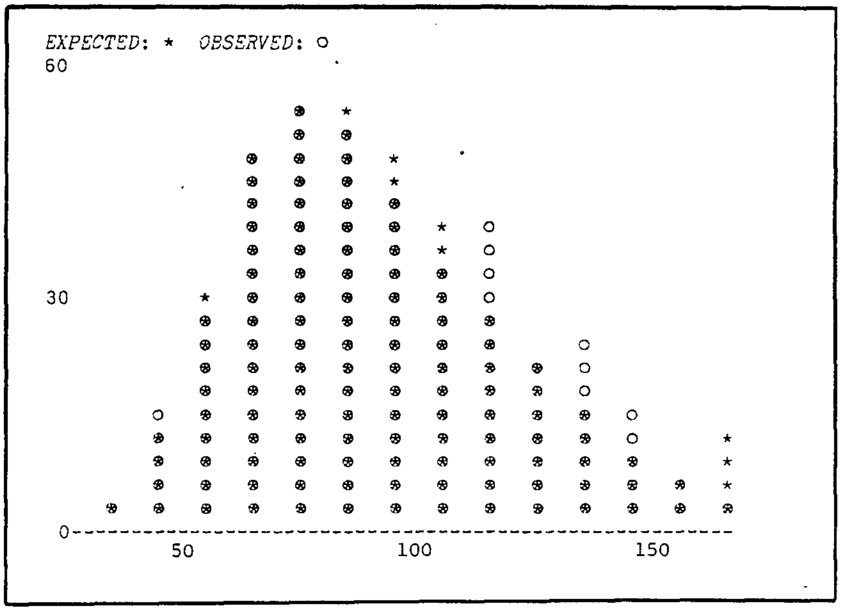
NORMAL



POISSON

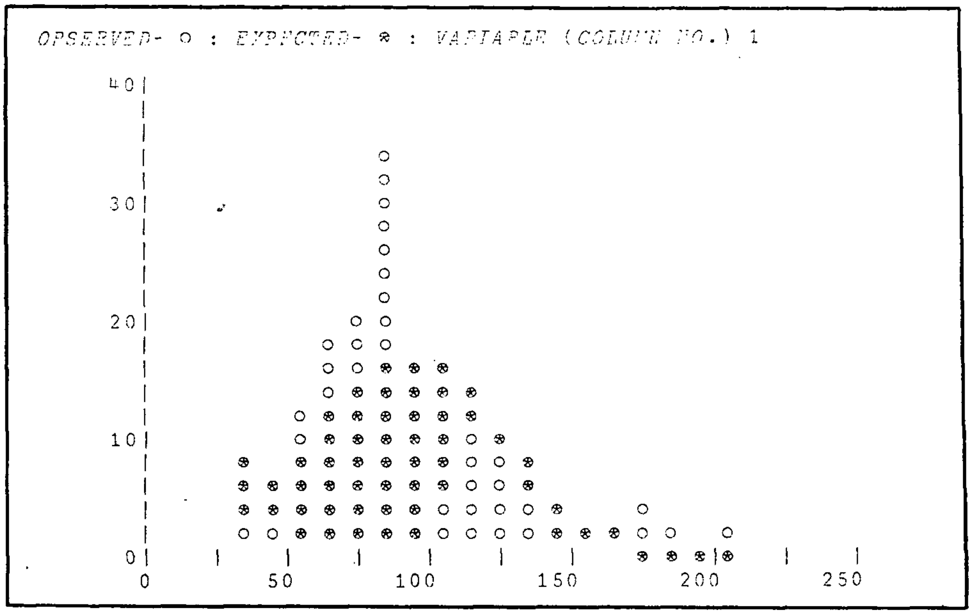


LOGNORMAL

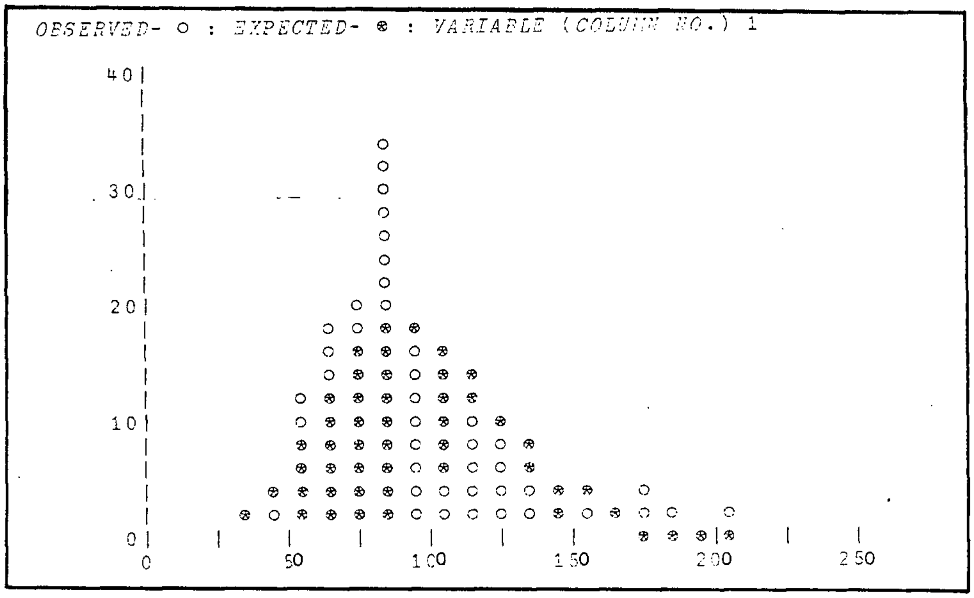


AIRPORT B, FLIGHT B1

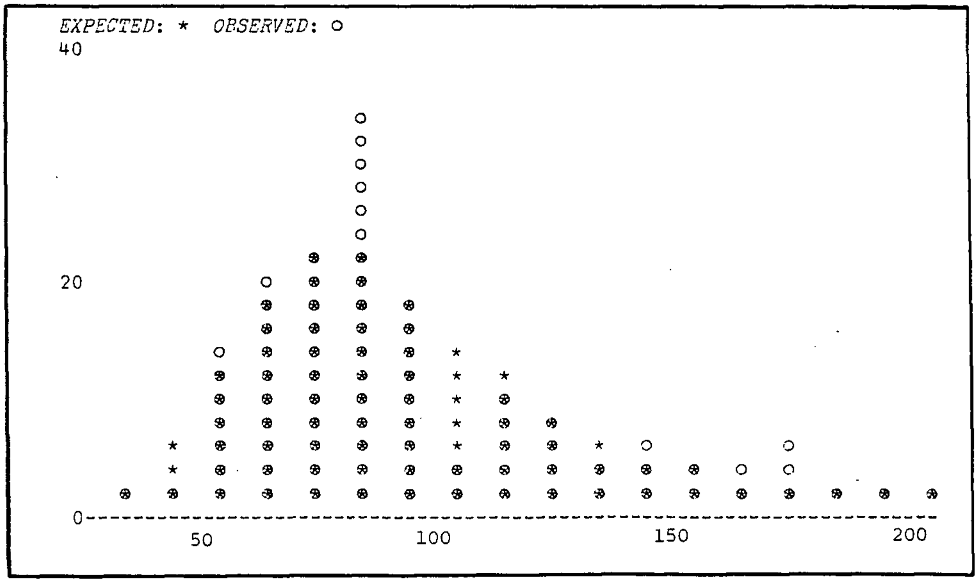
NORMAL



POISSON

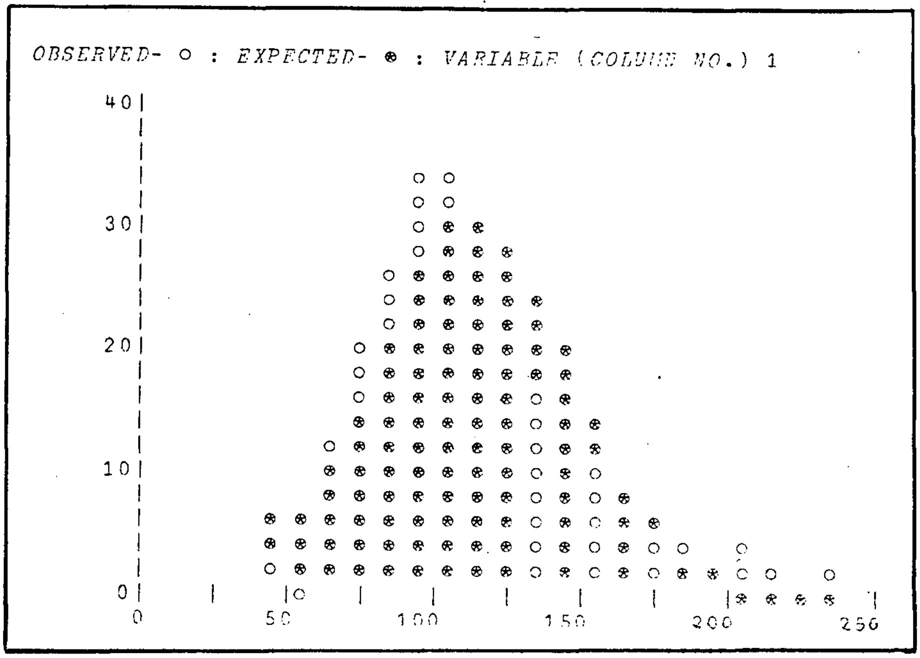


LOGNORMAL

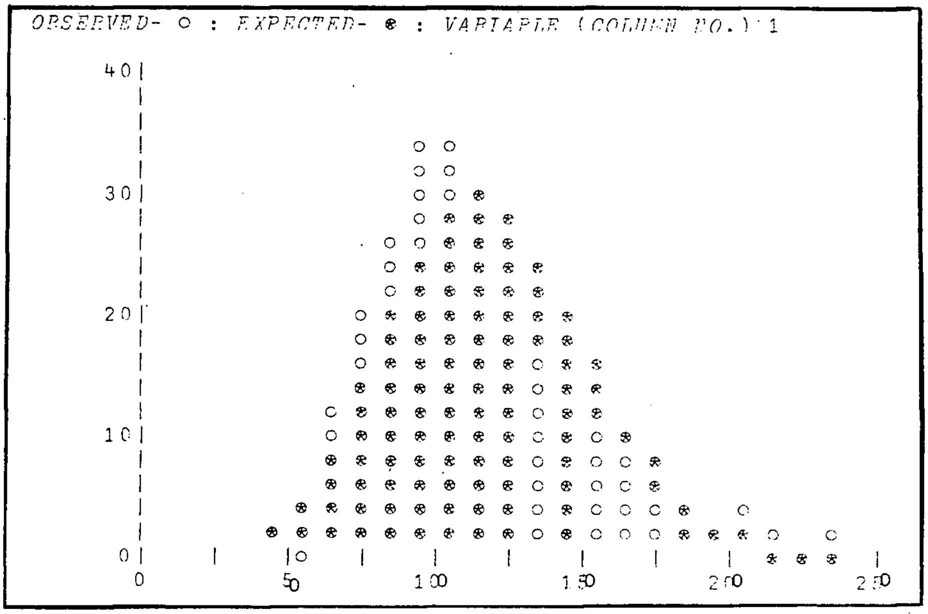


AIRPORT B, FLIGHT B2

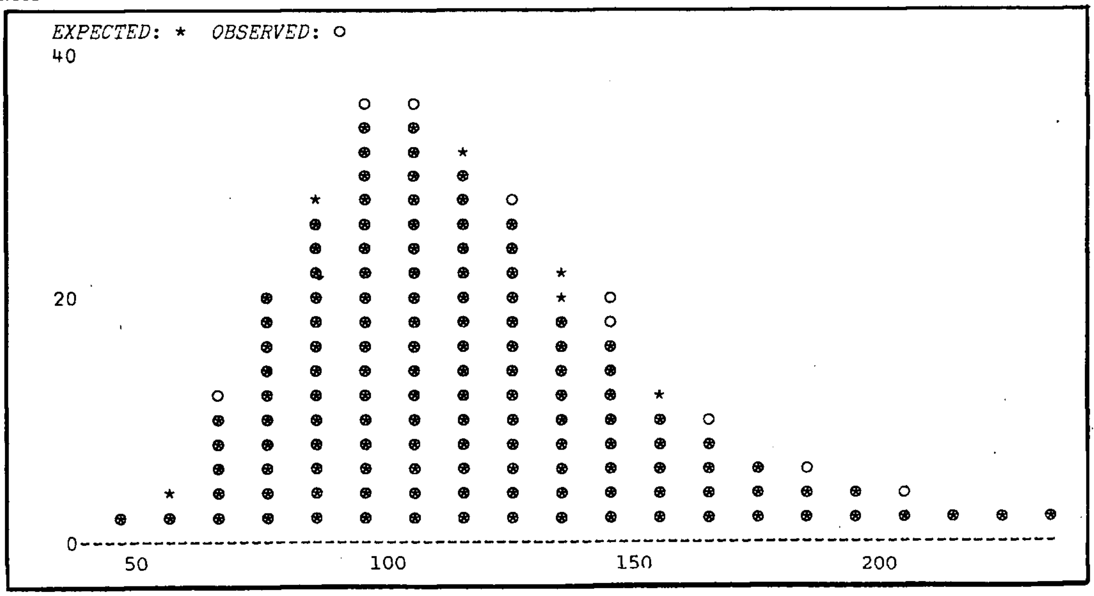
NORMAL



POISSON

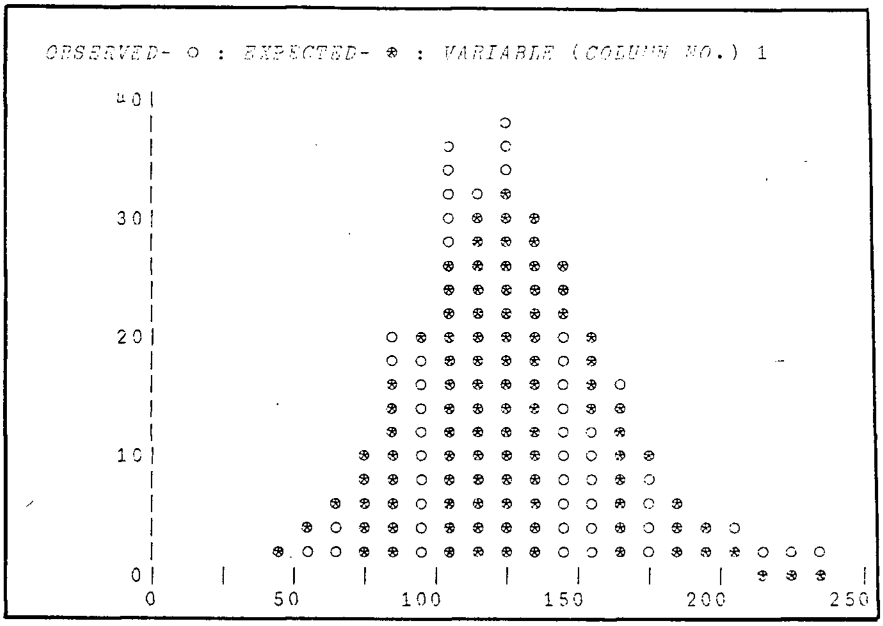


LOGNORMAL

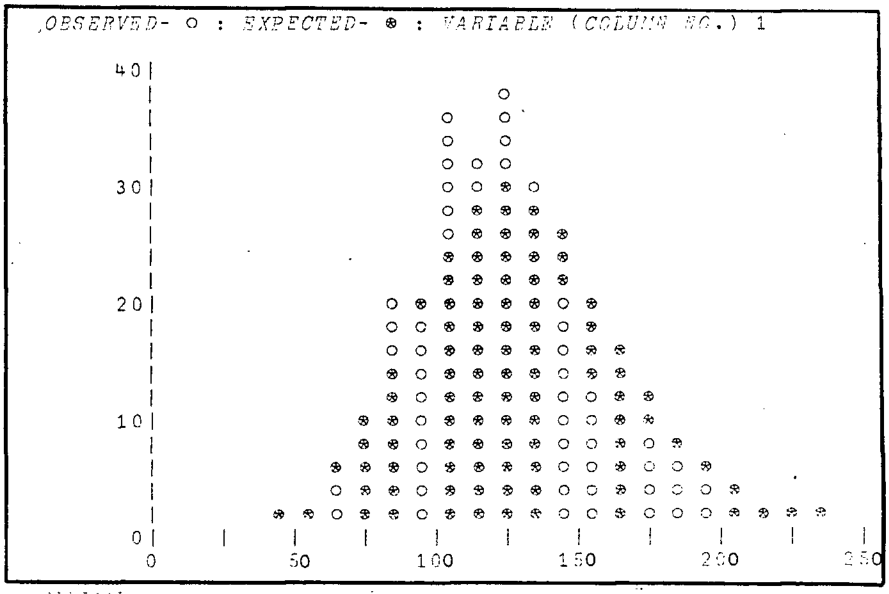


AIRPORT B, FLIGHT B3

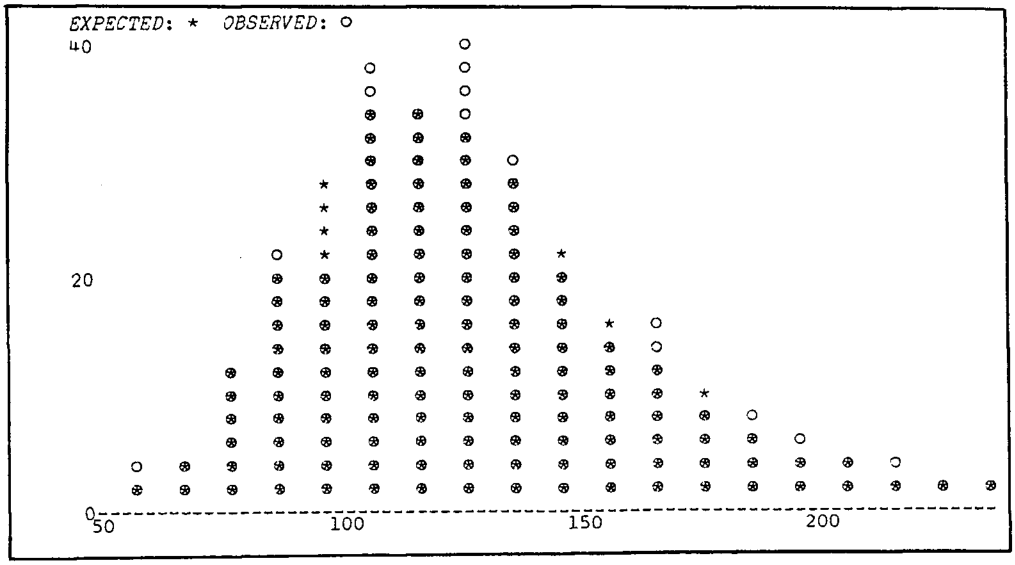
NORMAL



POISSON

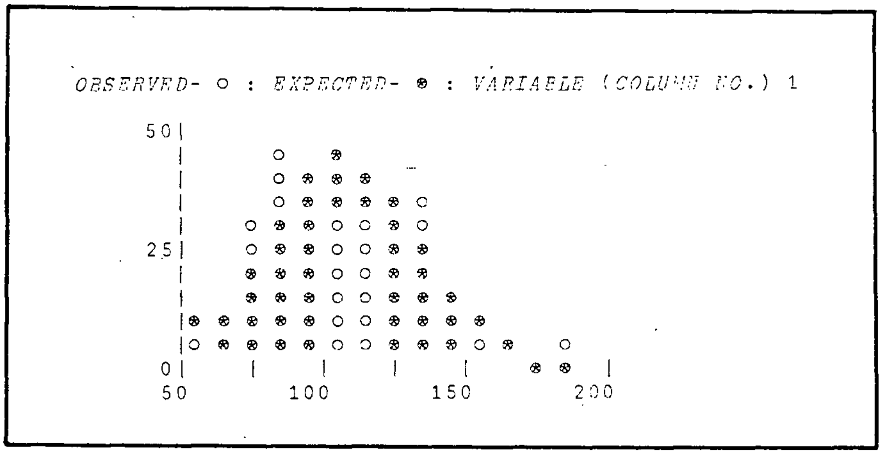


LOGNORMAL

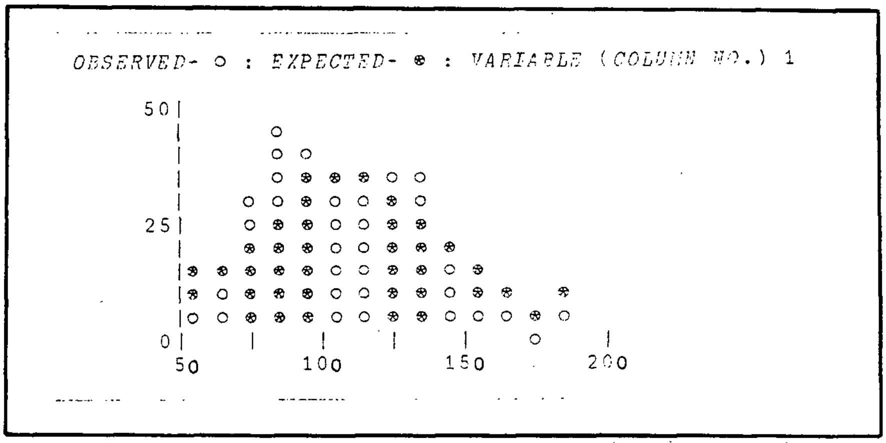


AIRPORT B, FLIGHT B4

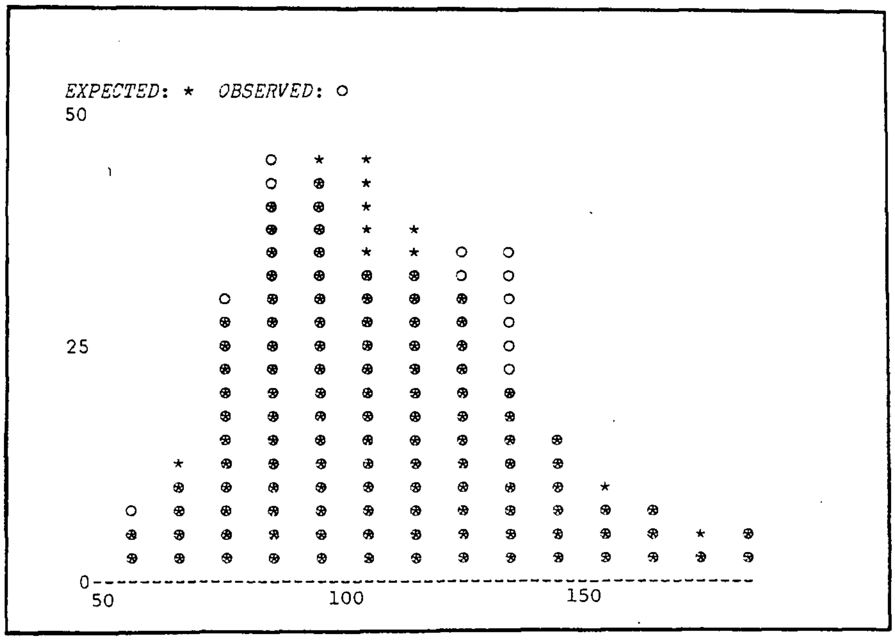
NORMAL



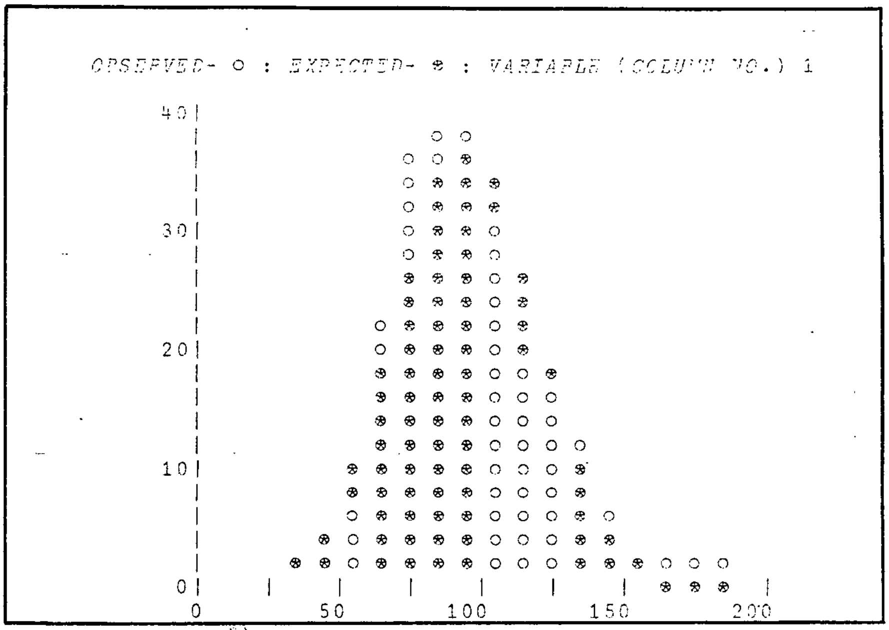
POISSON



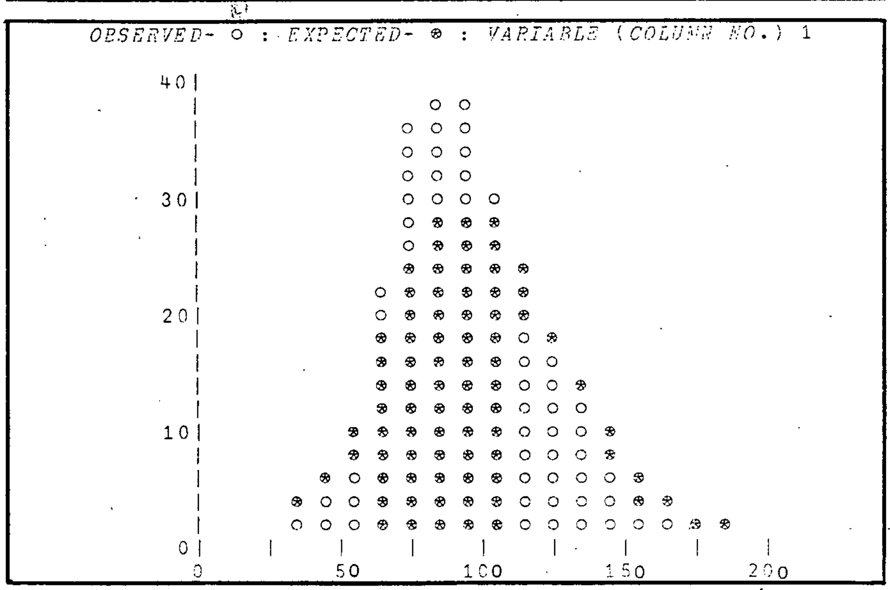
LOGNORMAL



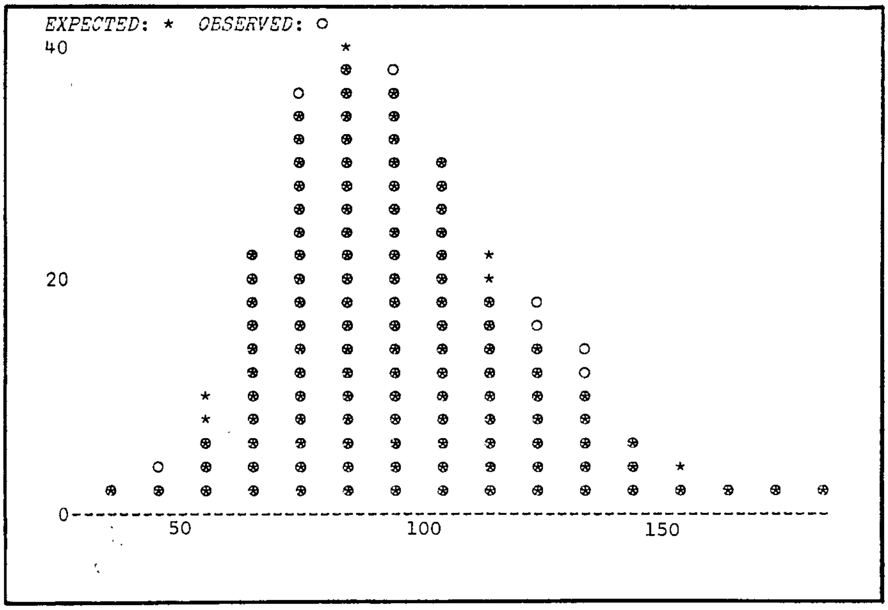
NORMAL



POISSON

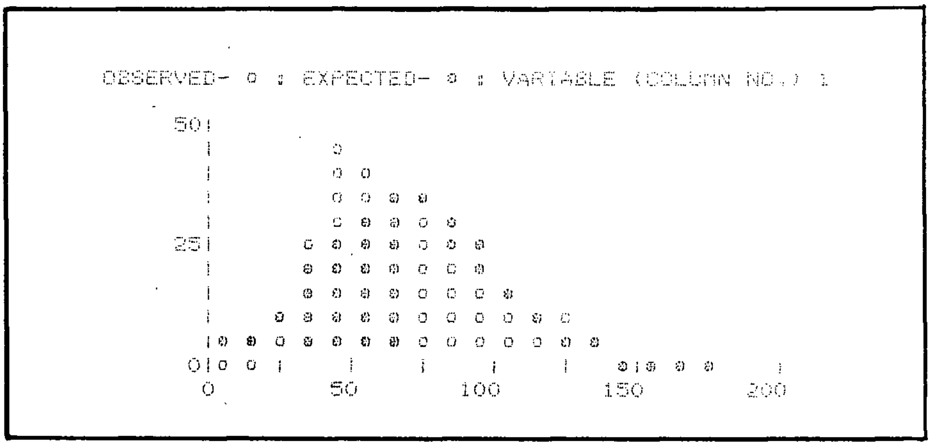


LOGNORMAL

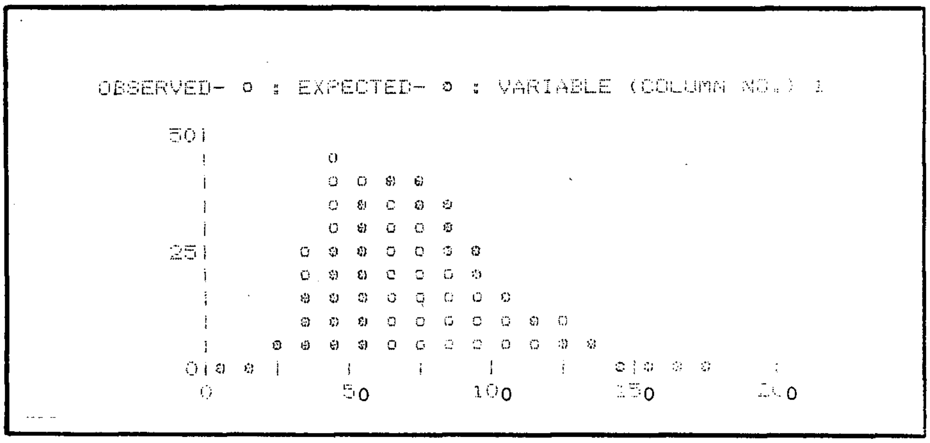


AIRPORT B, FLIGHT B6

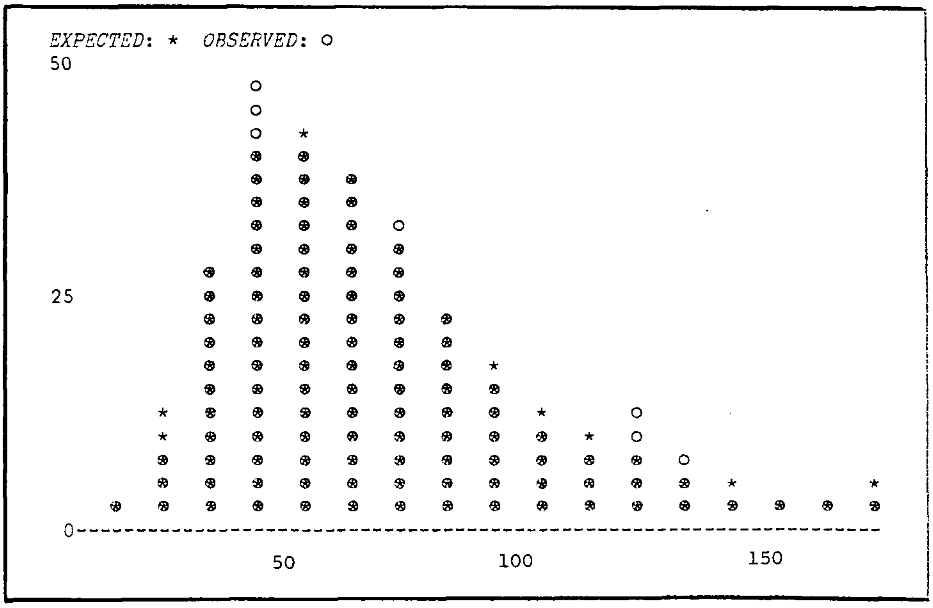
NORMAL



POISSON

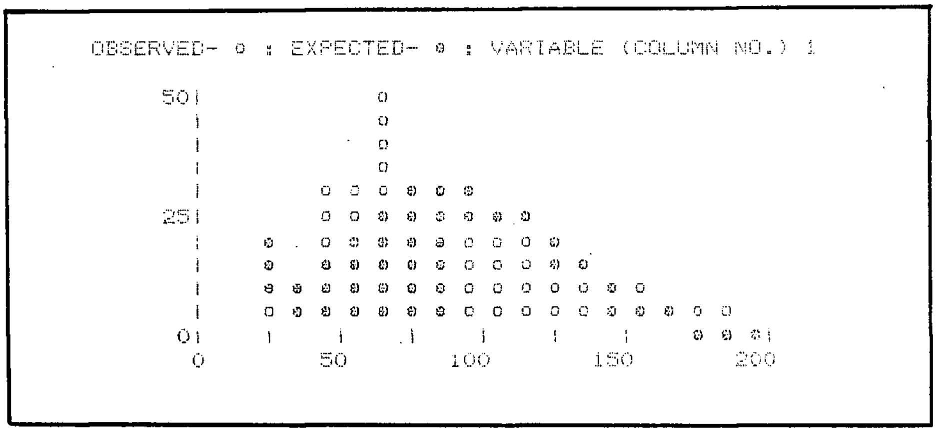


LOGNORMAL

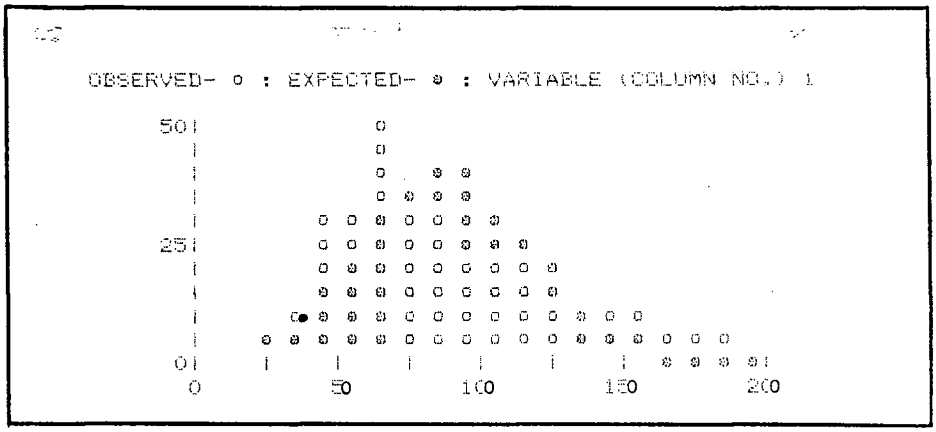


AIRPORT B, FLIGHT B7

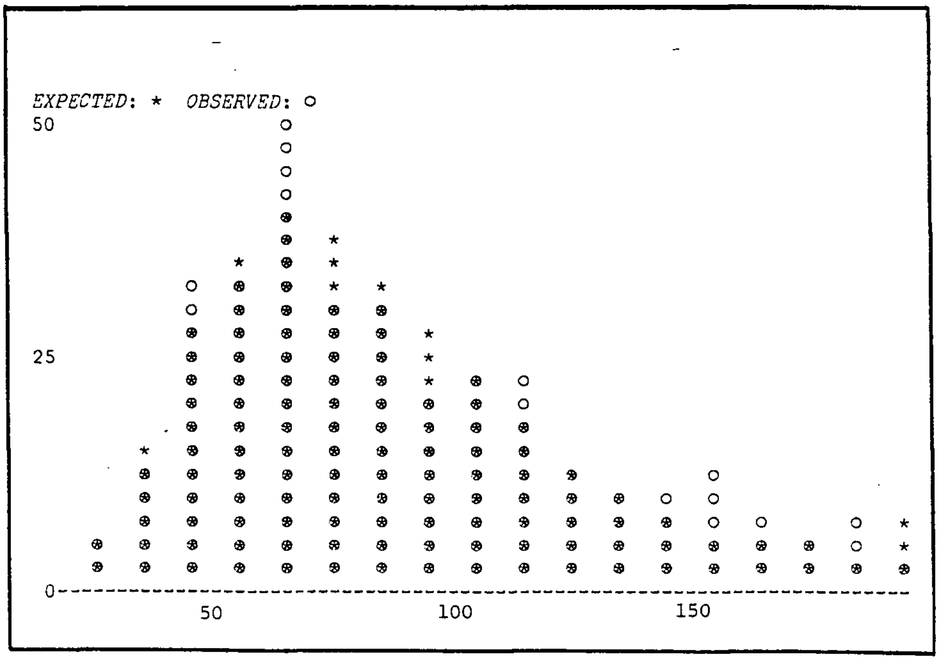
NORMAL



POISSON

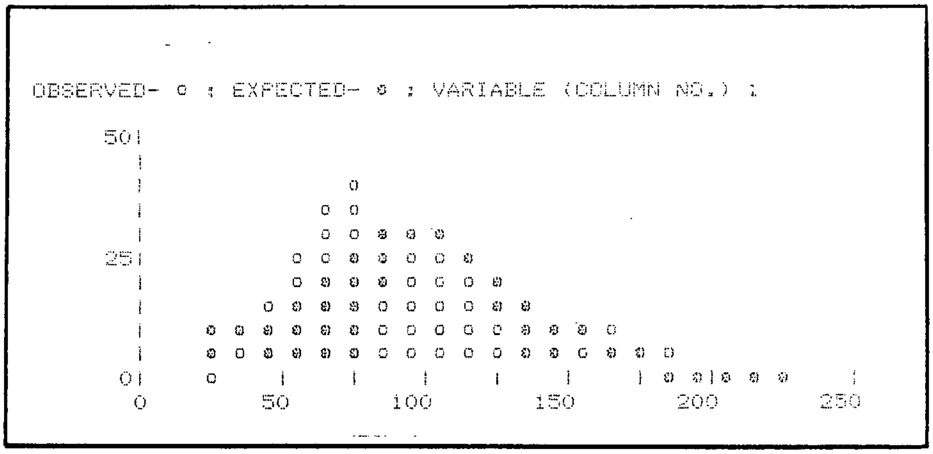


LOGNORMAL

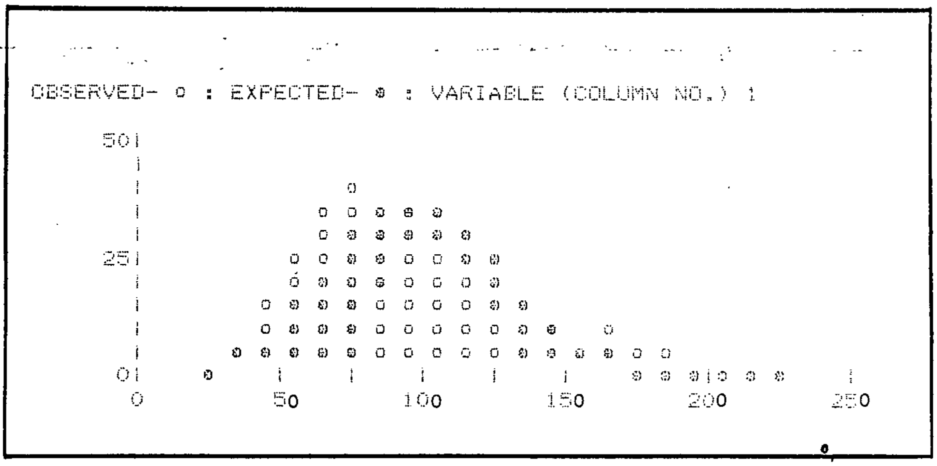


AIRPORT B, FLIGHT B8

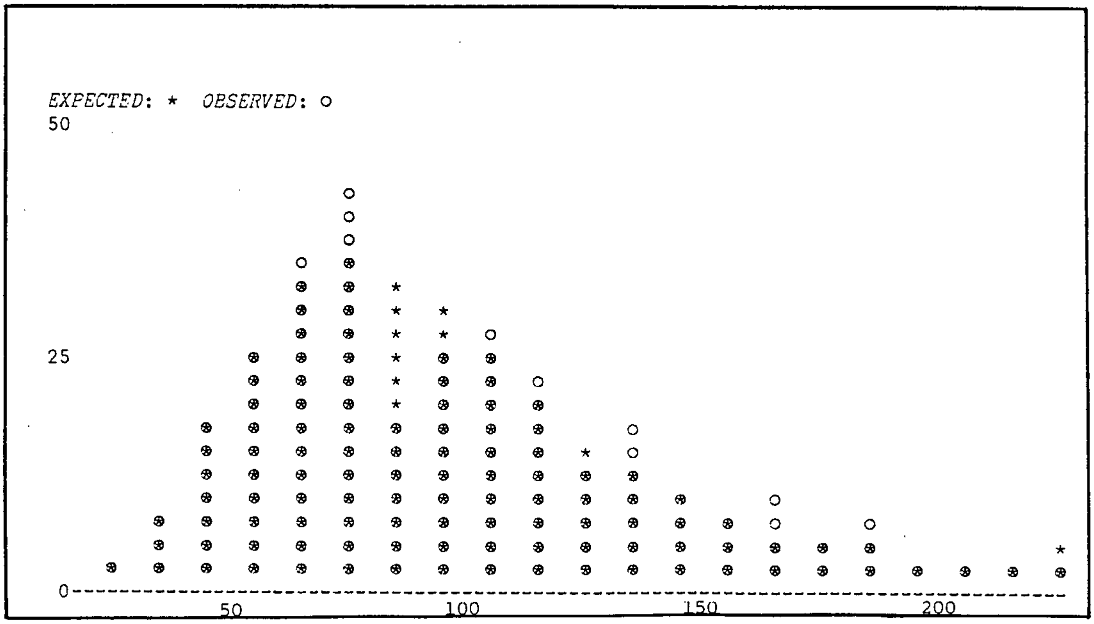
NORMAL



POISSON

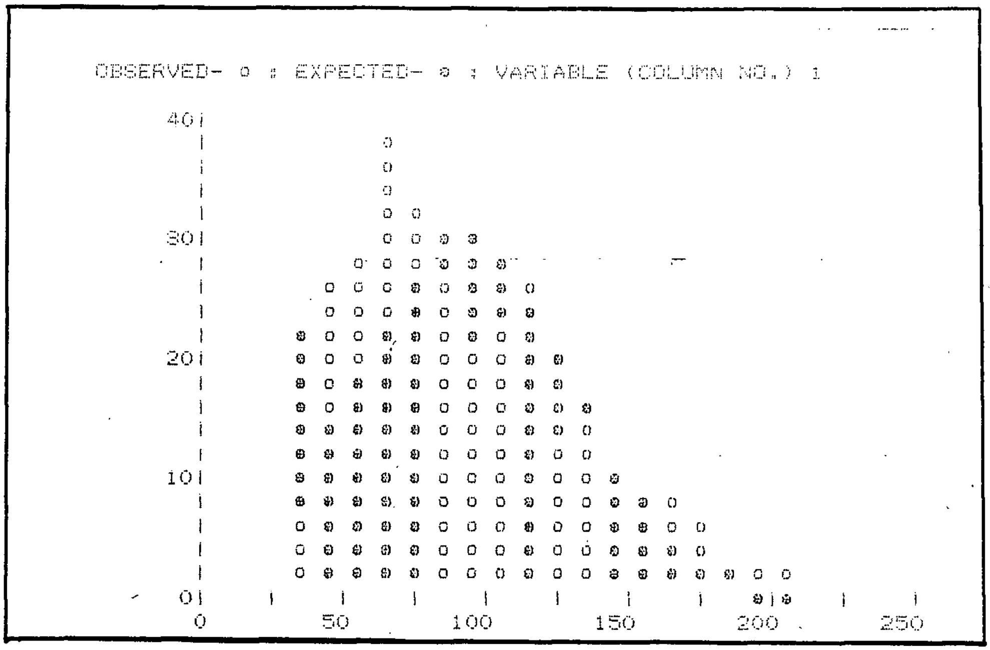


LOGNORMAL

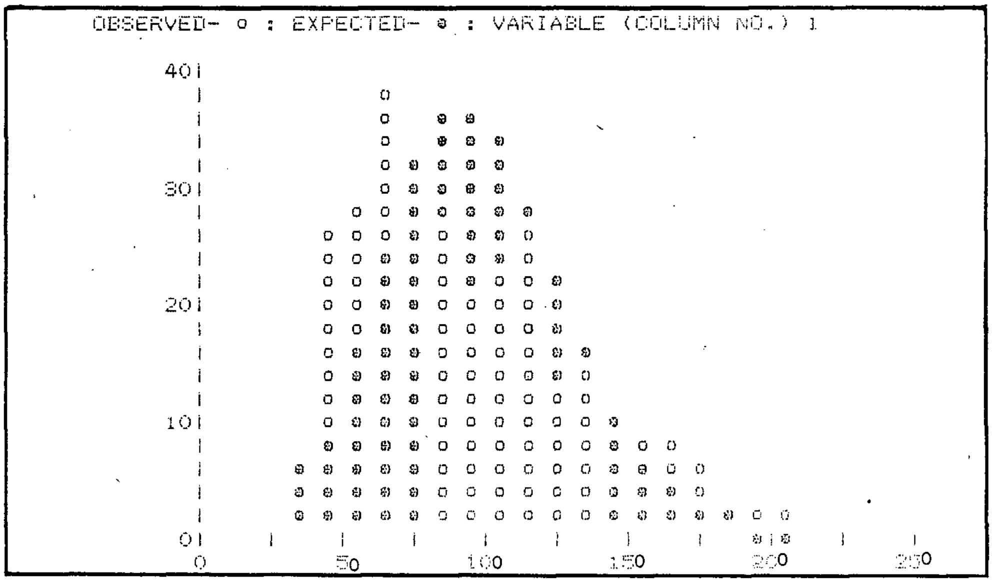


AIRPORT B, FLIGHT B9

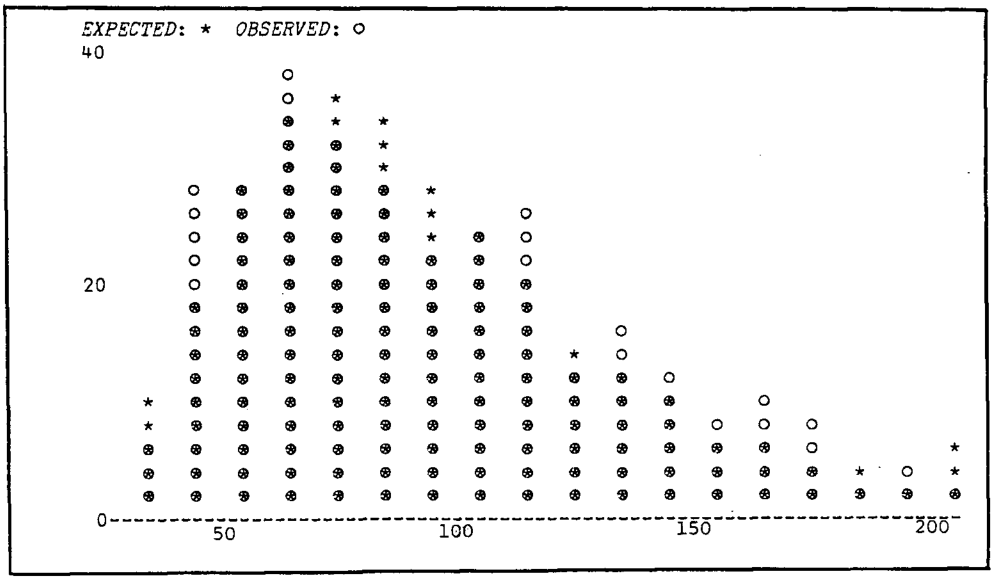
NORMAL



POISSON

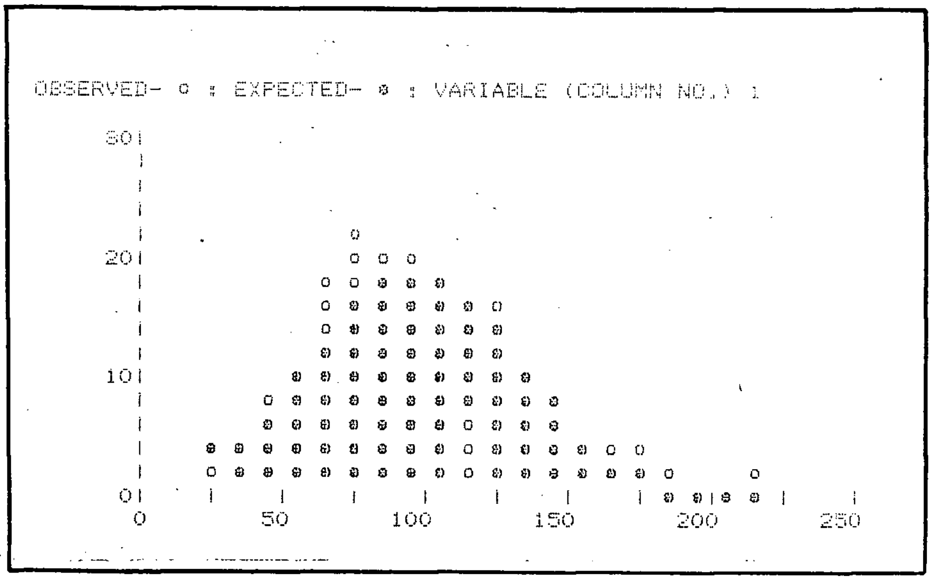


LOGNORMAL

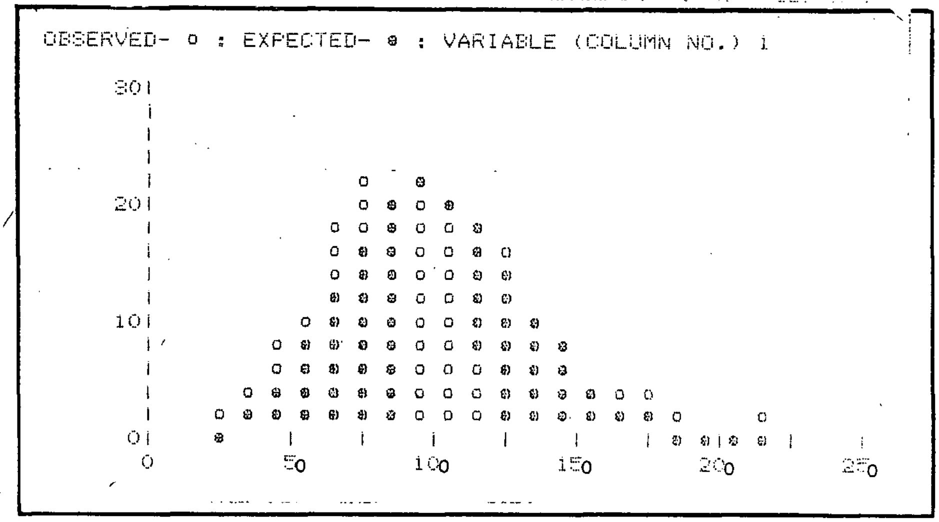


AIRPORT B, FLIGHT B10

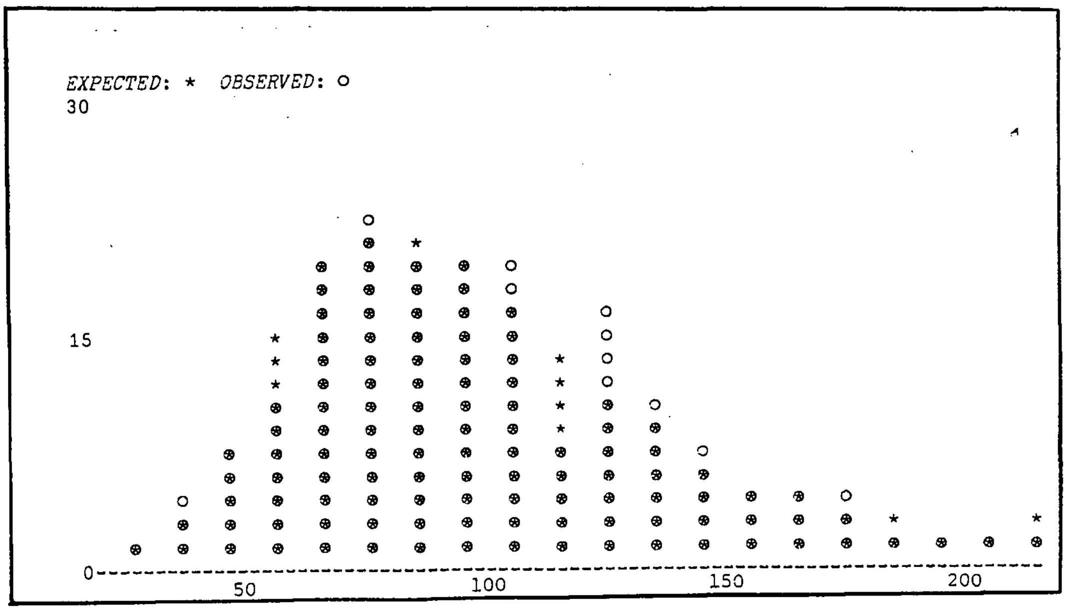
NORMAL



POISSON

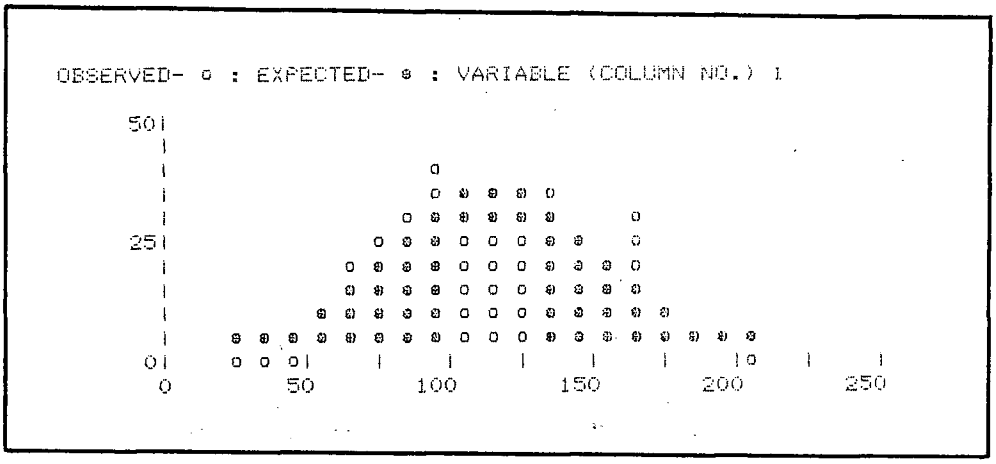


LOGNORMAL

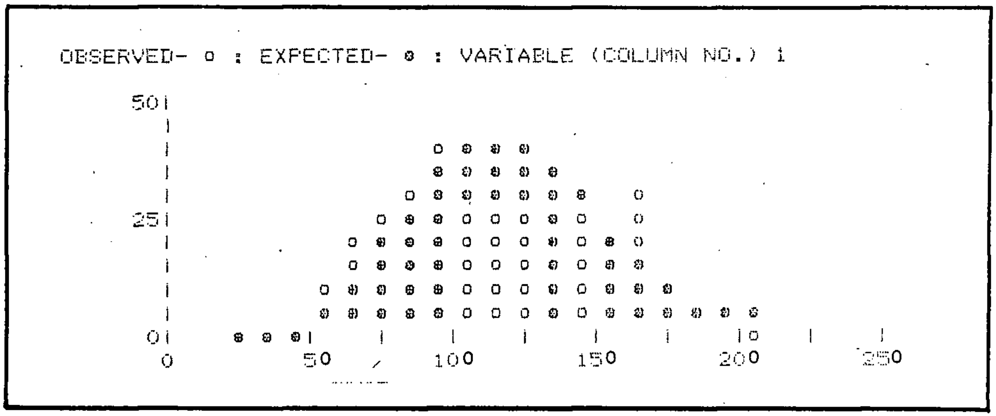


AIRPORT B, FLIGHT B11

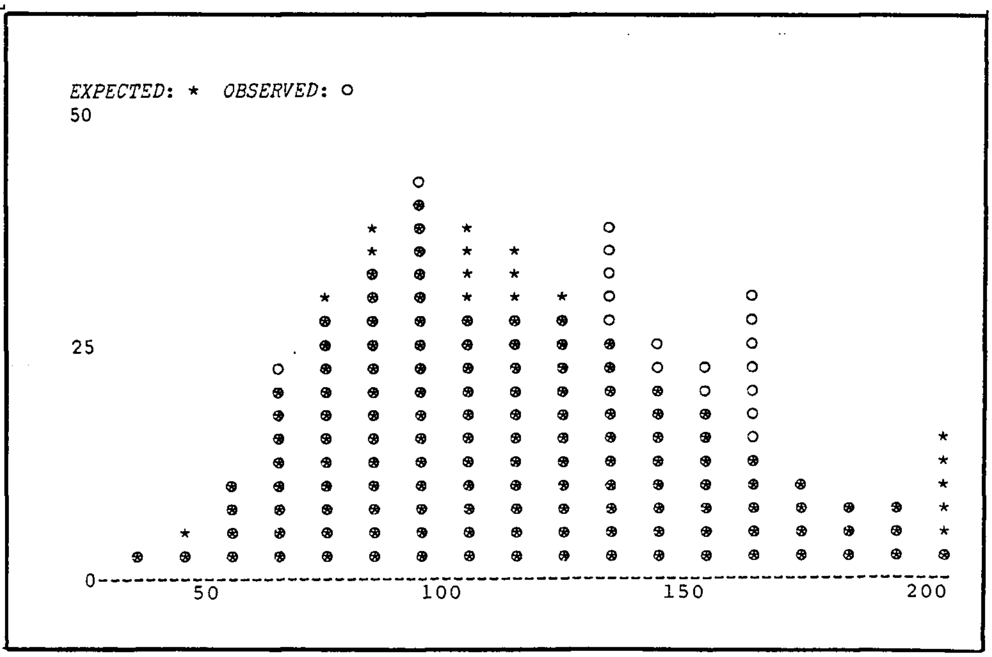
NORMAL



POISSON

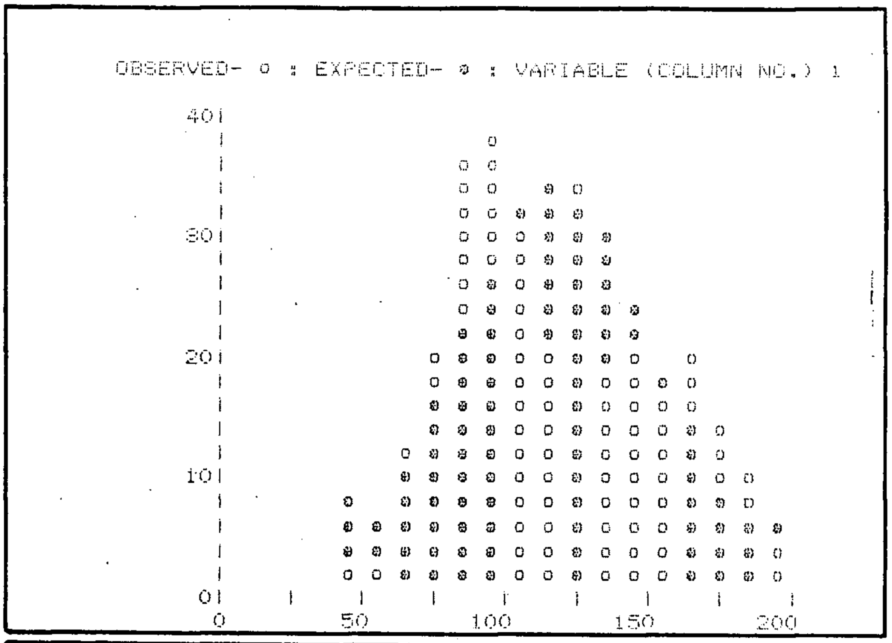


LOGNORMAL

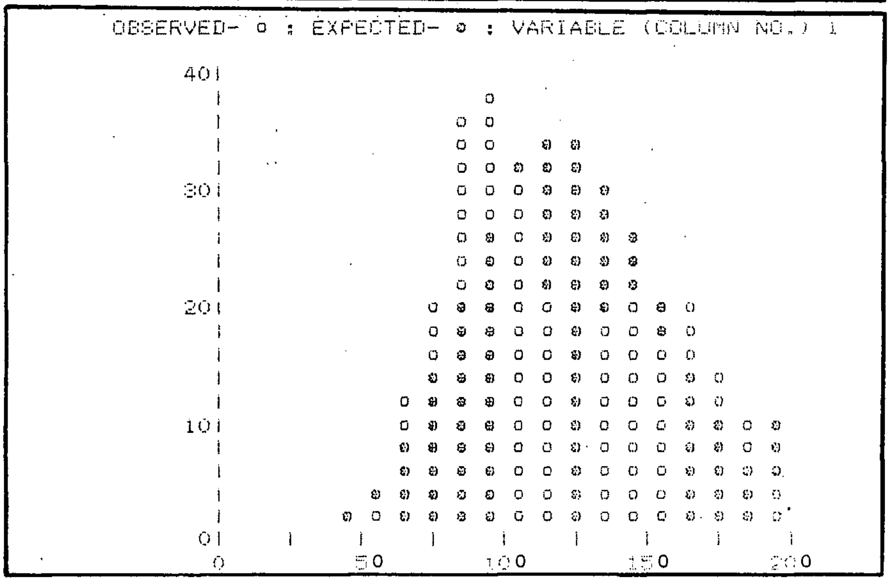


AIRPORT B, FLIGHT B12

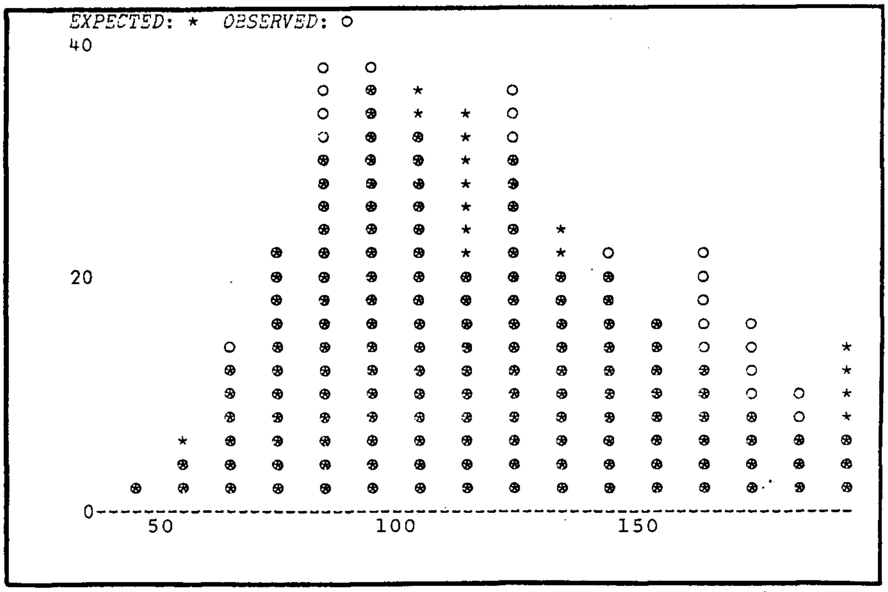
NORMAL



POISSON

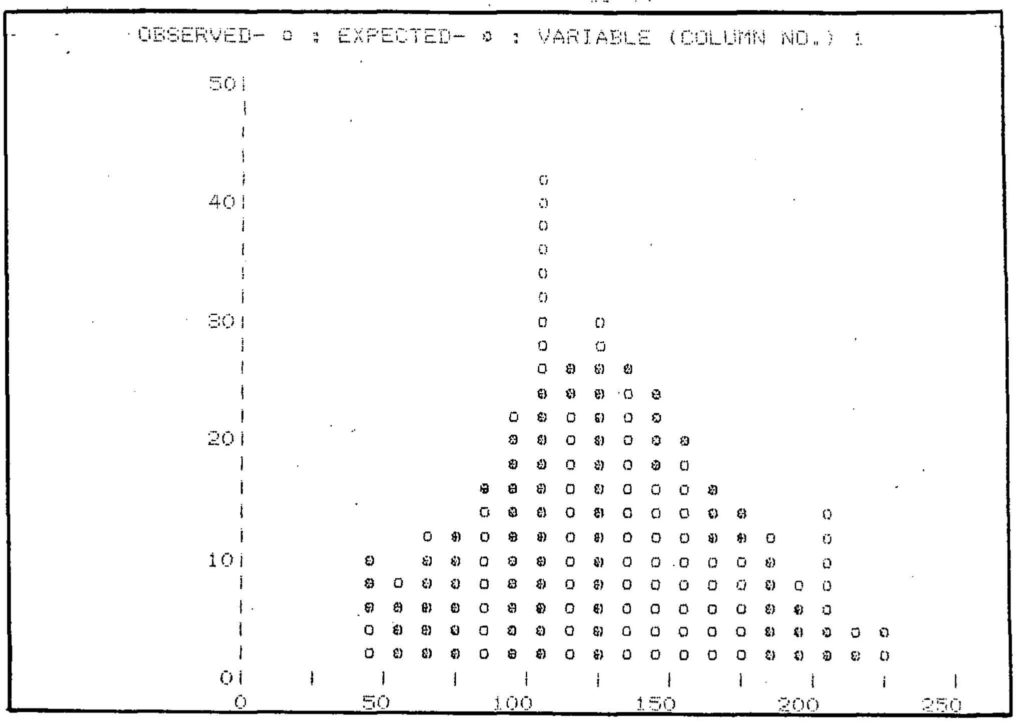


LOGNORMAL

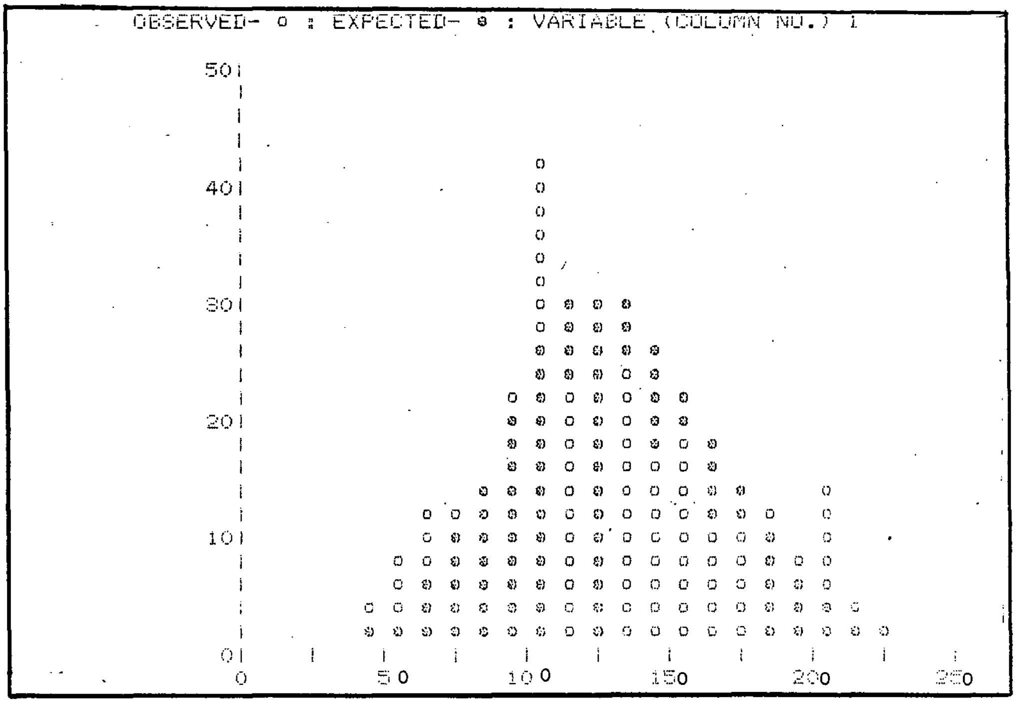


AIRPORT B,
FLIGHT
B13

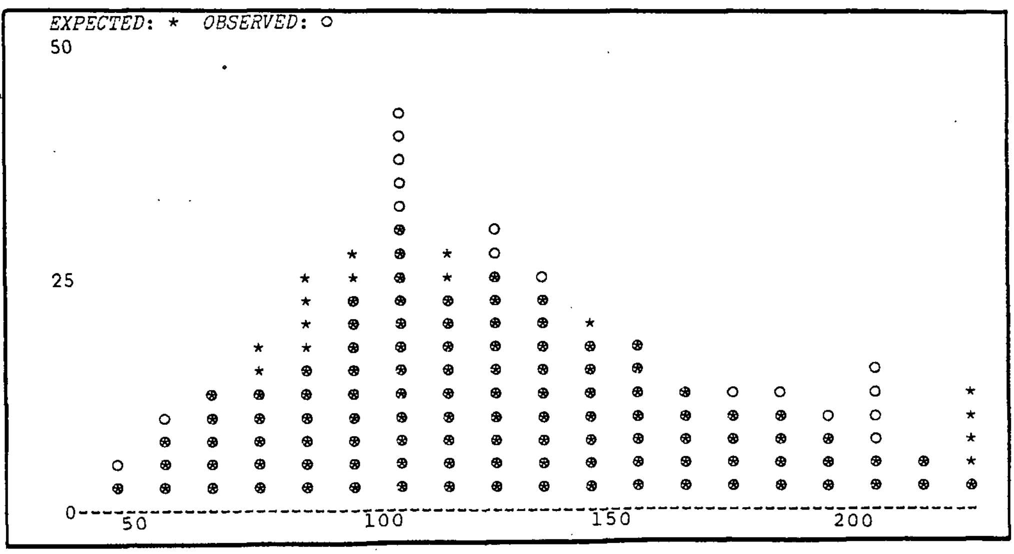
NORMAL



POISSON

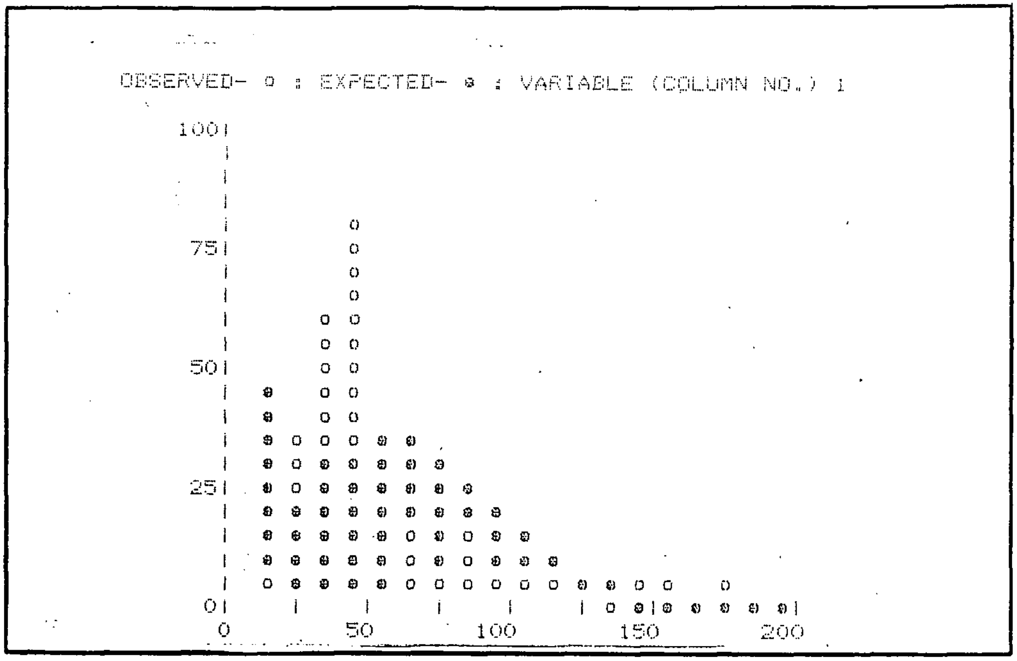


LOGNORMAL

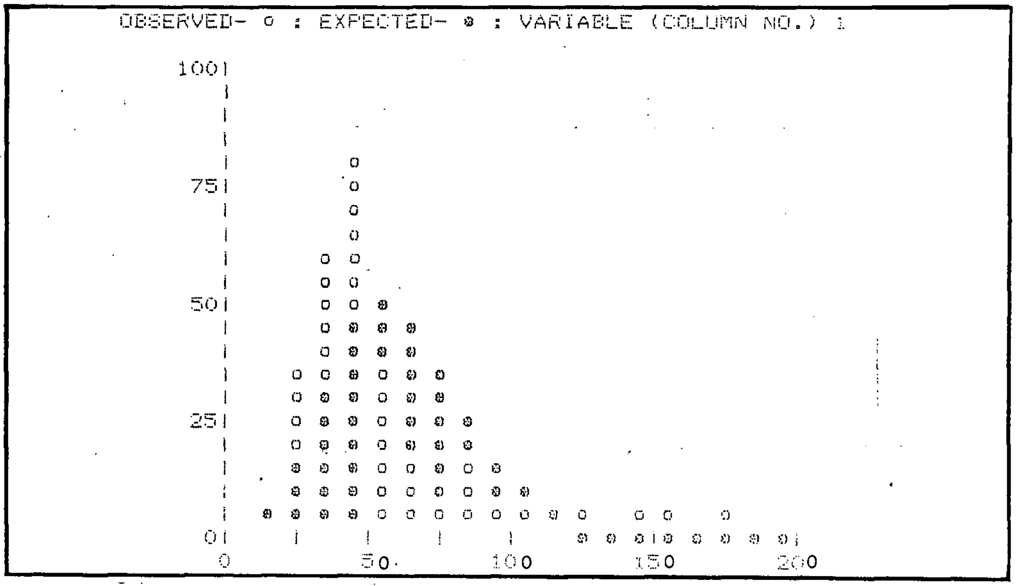


AIRPORT B, FLIGHT B14

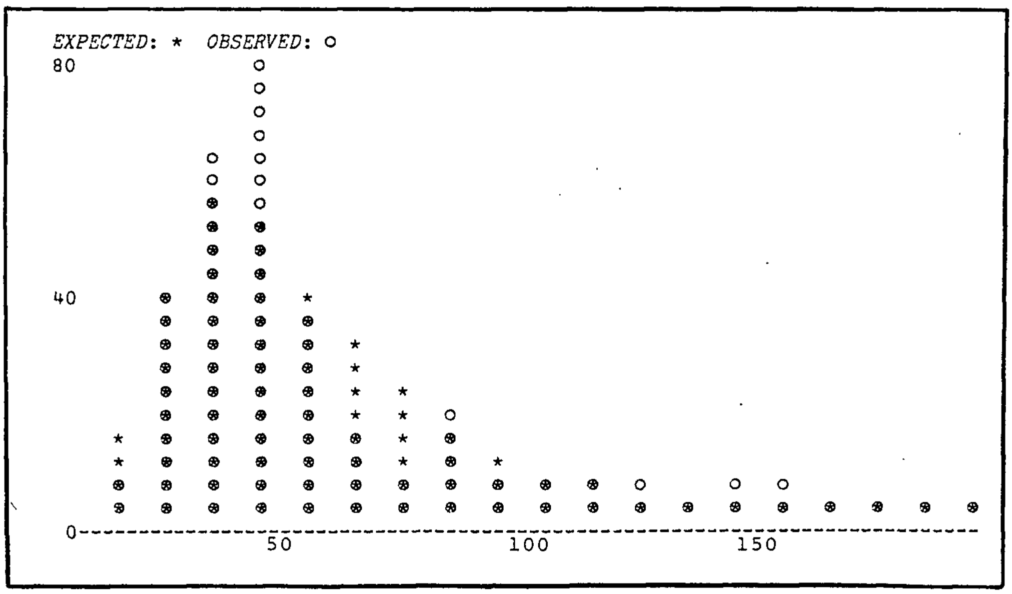
NORMAL



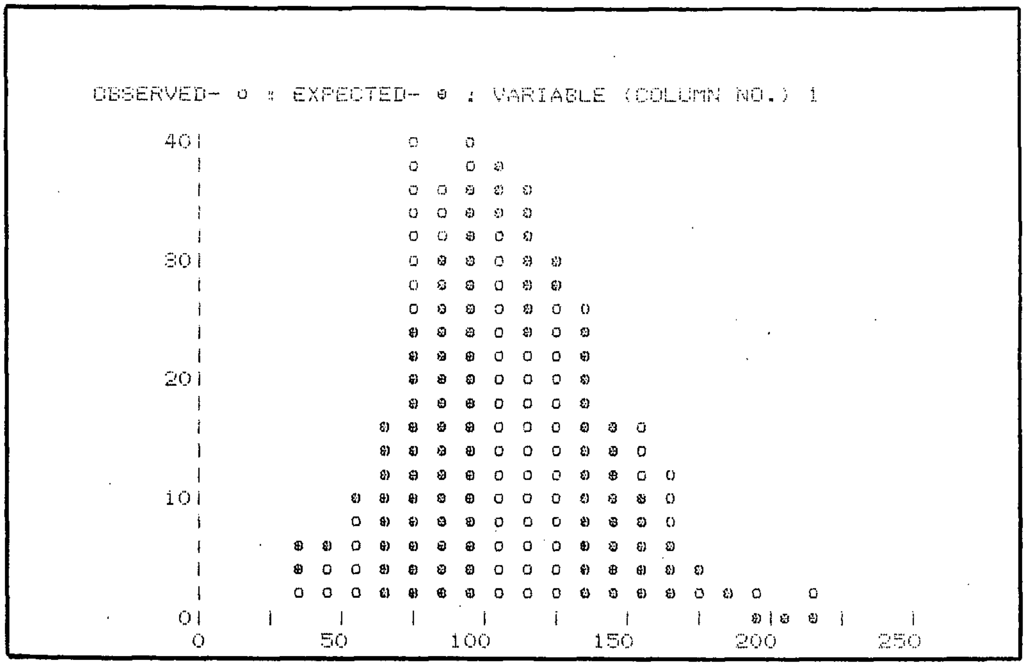
POISSON



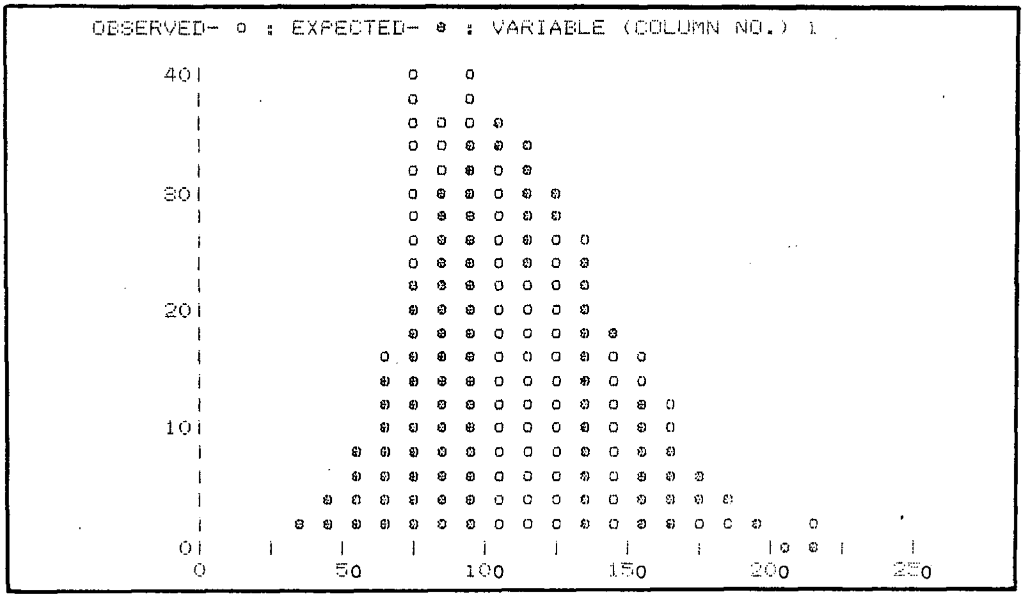
LOGNORMAL



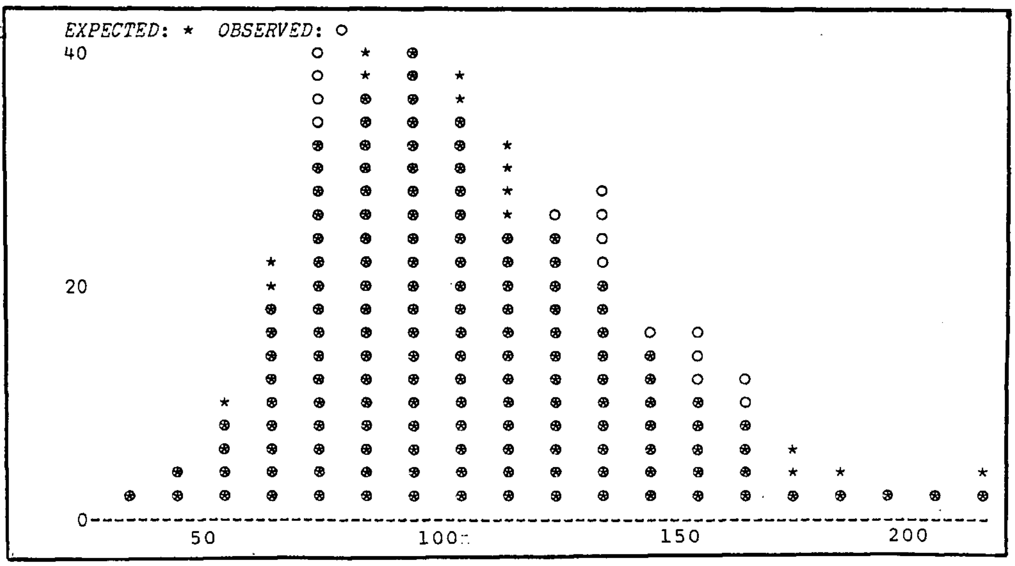
NORMAL



POISSON

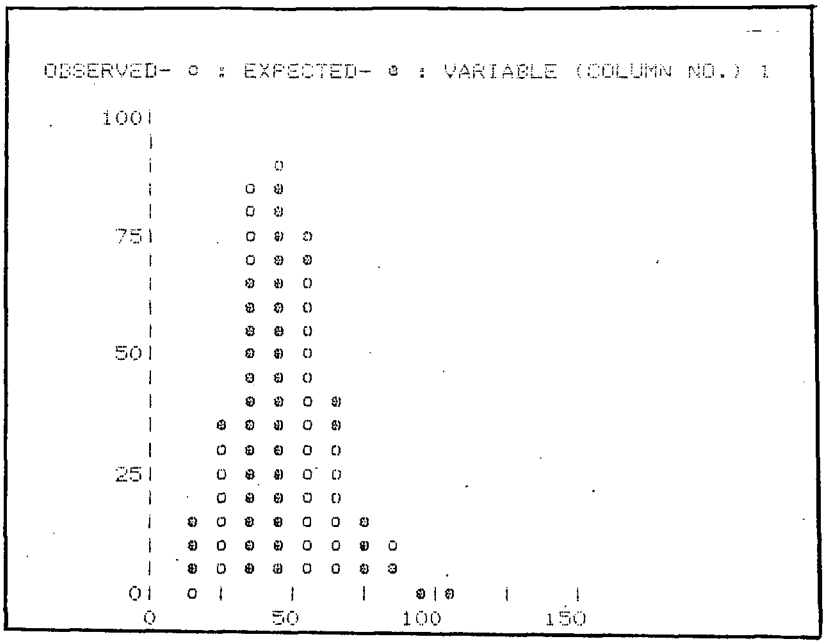


LOGNORMAL

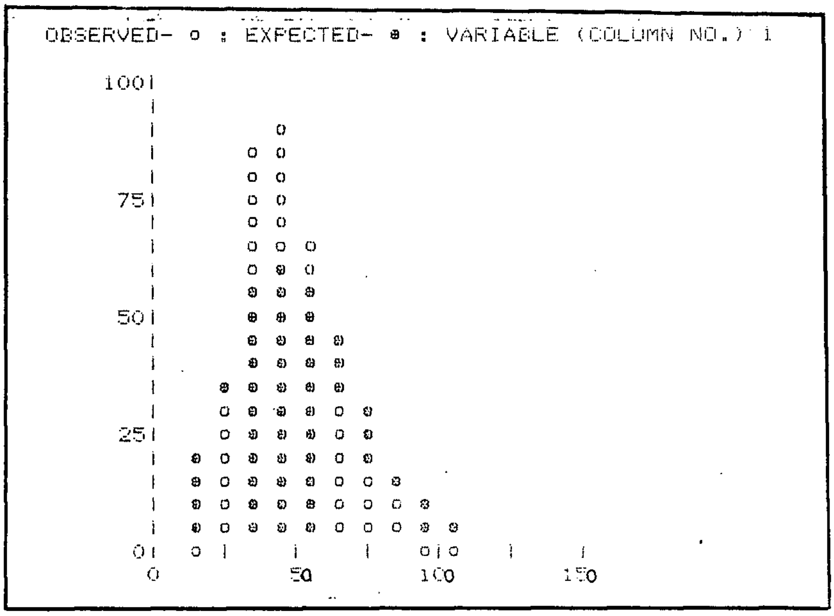


AIRPORT C, FLIGHT C1

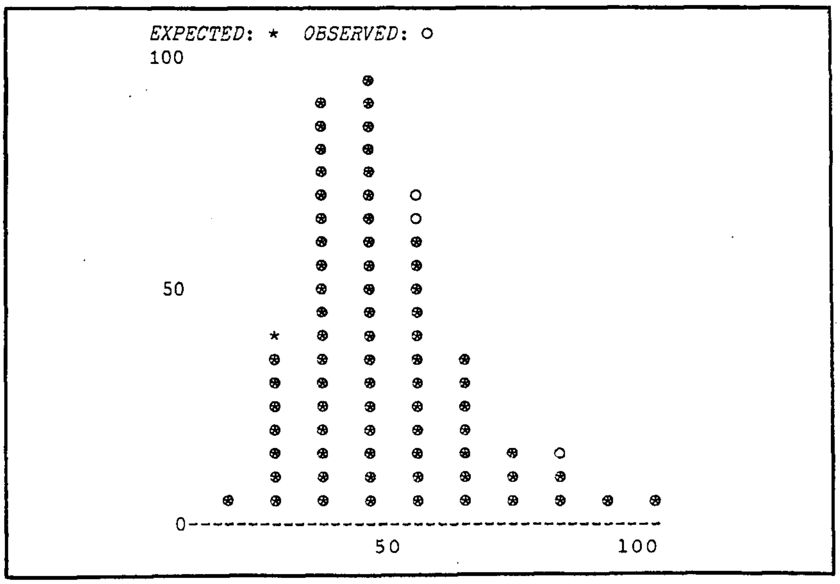
NORMAL



POISSON

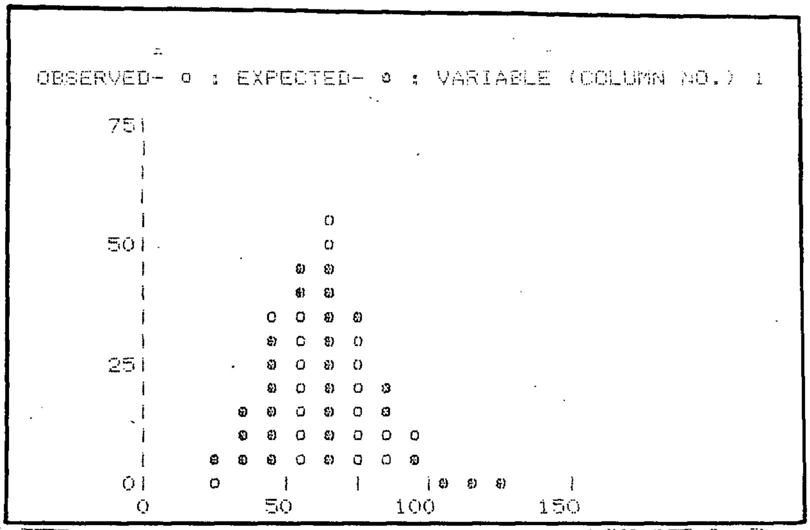


LOGNORMAL

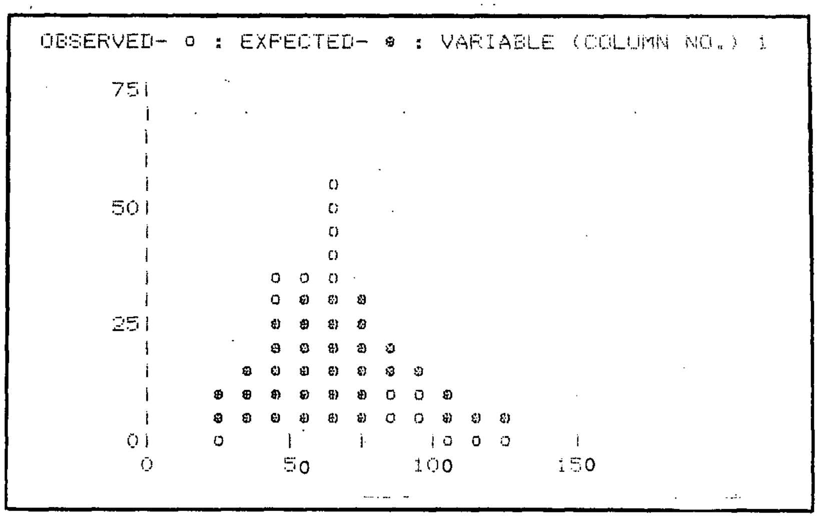


AIRPORT C, FLIGHT C2

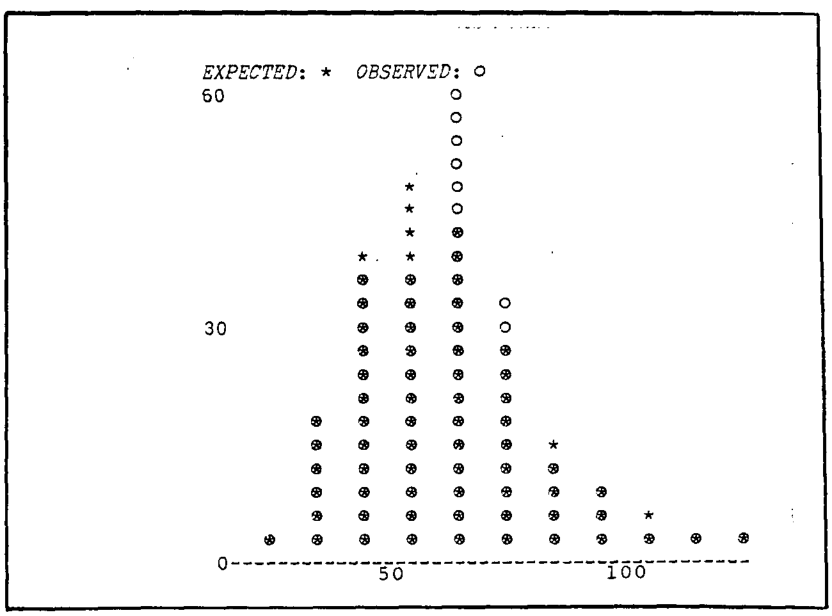
NORMAL



POISSON

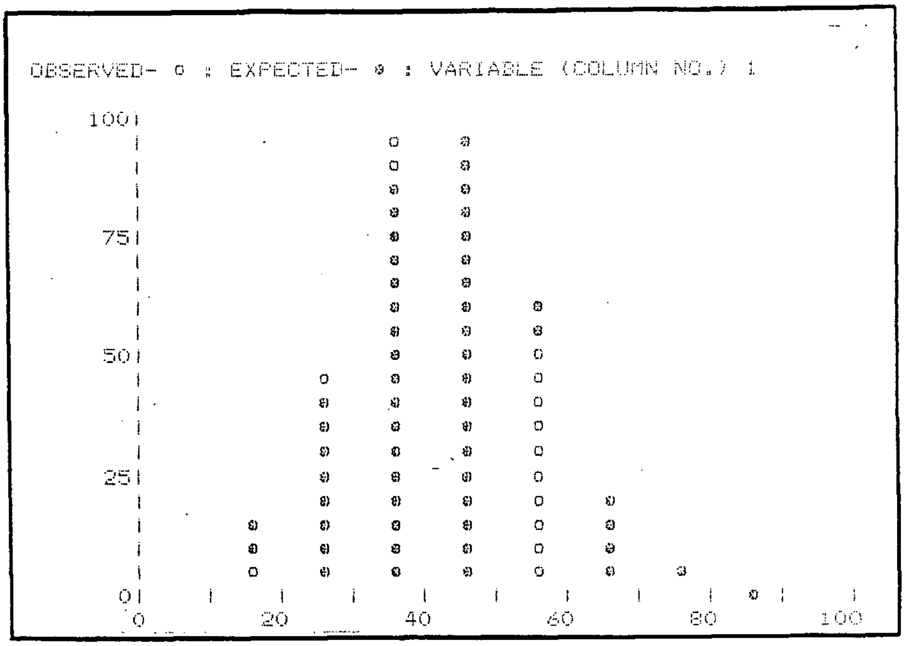


LOGNORMAL

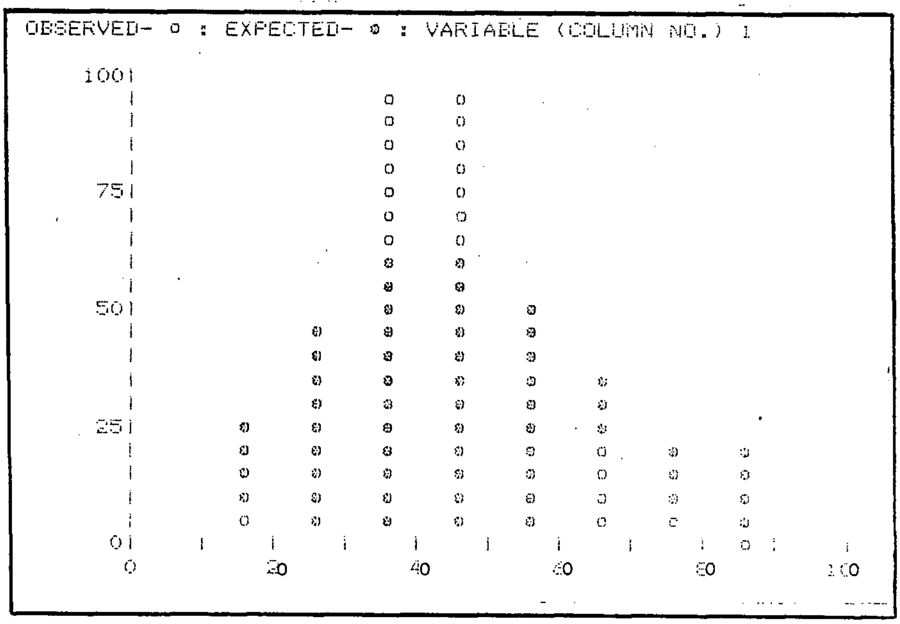


AIRPORT C, FLIGHT C3

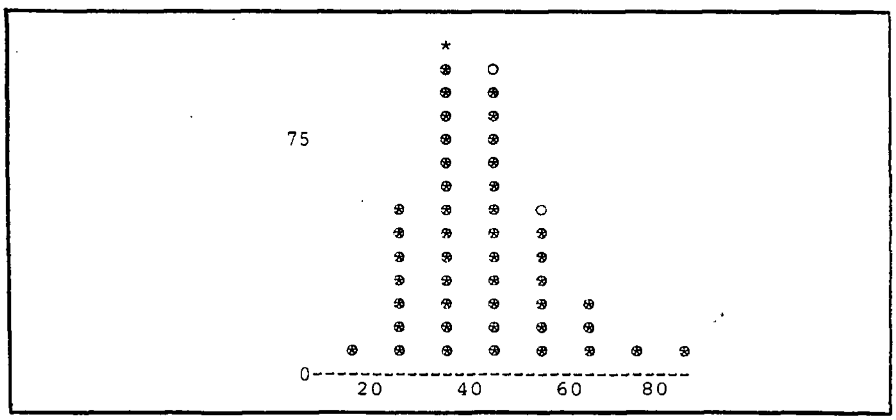
NORMAL



POISSON

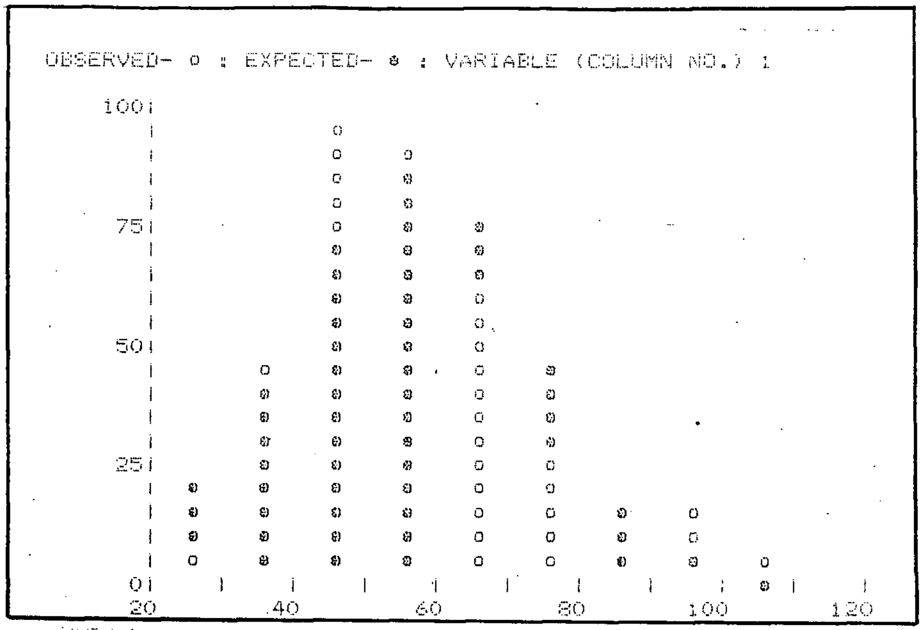


LOGNORMAL

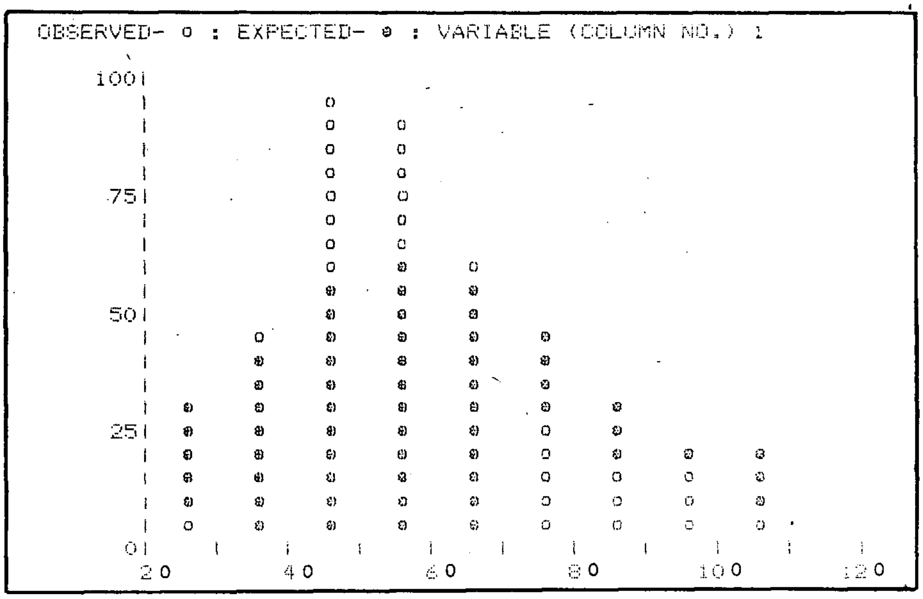


AIRPORT C, FLIGHT C4

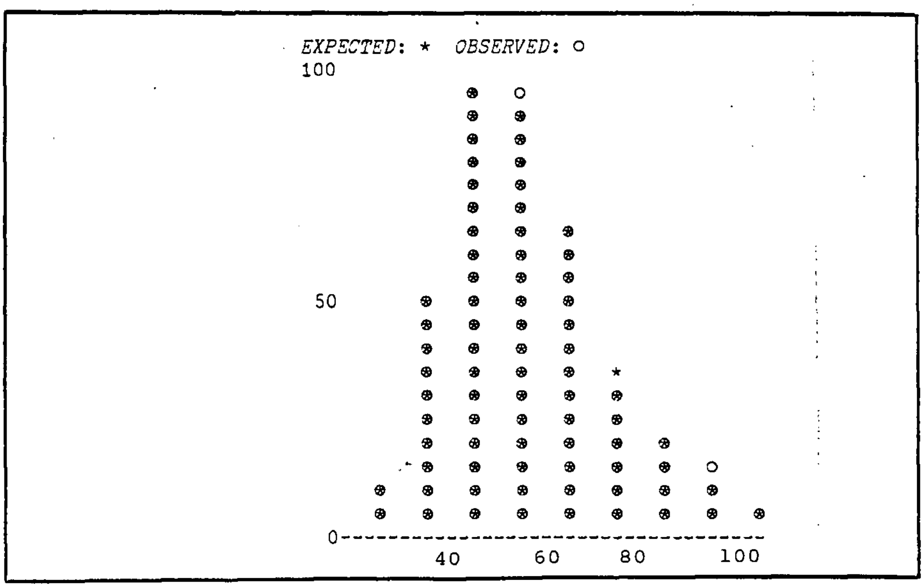
NORMAL



POISSON

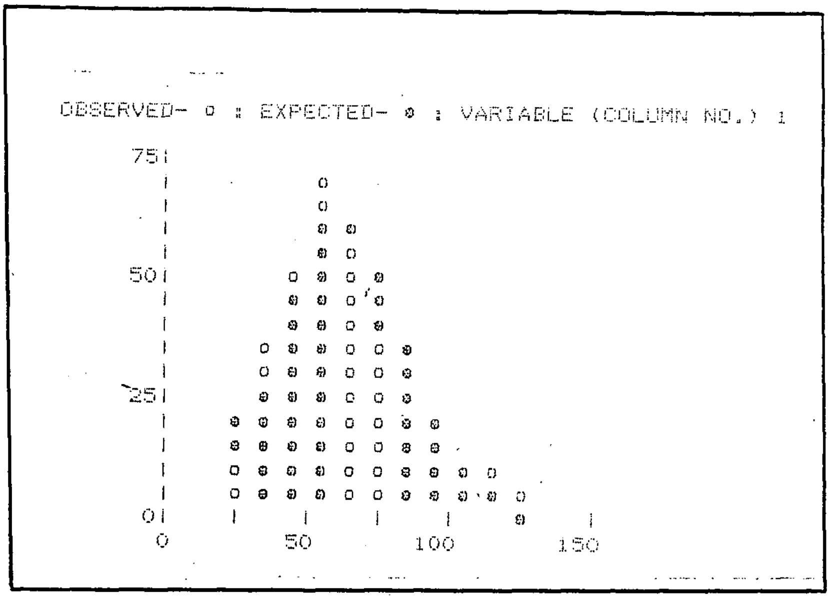


LOGNORMAL

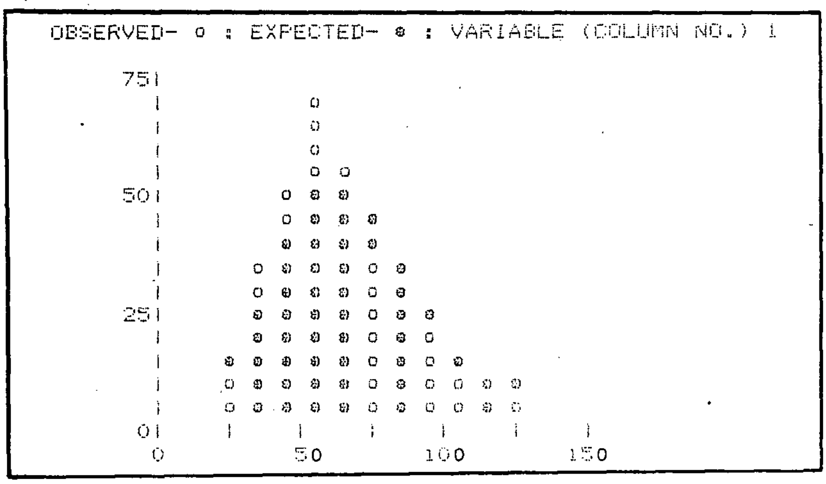


AIRPORT D, FLIGHT D1

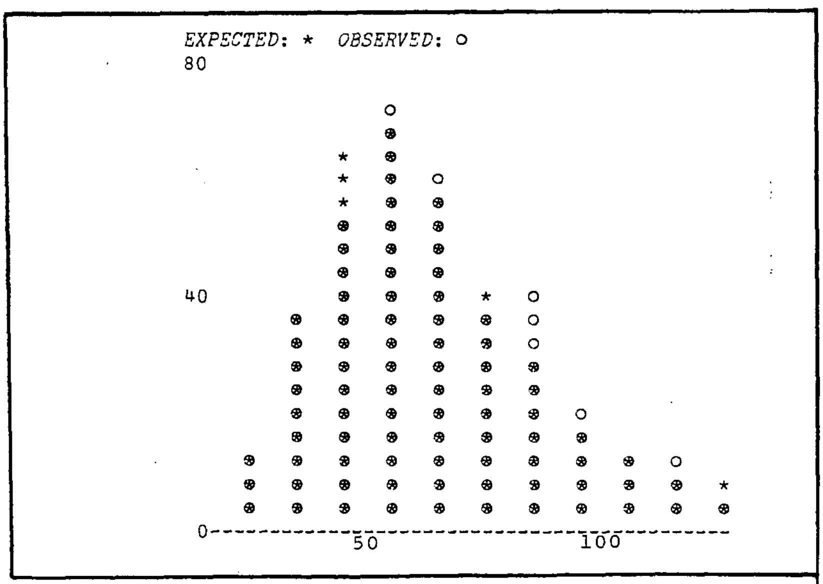
NORMAL



POISSON

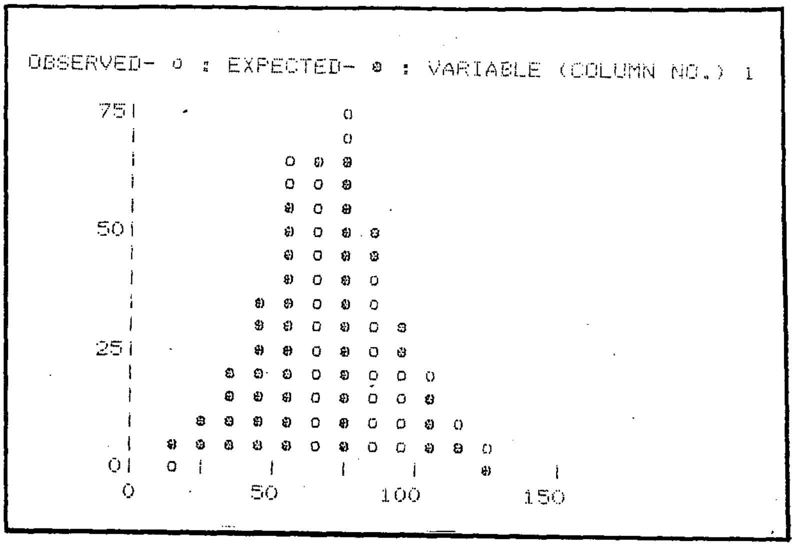


LOGNORMAL

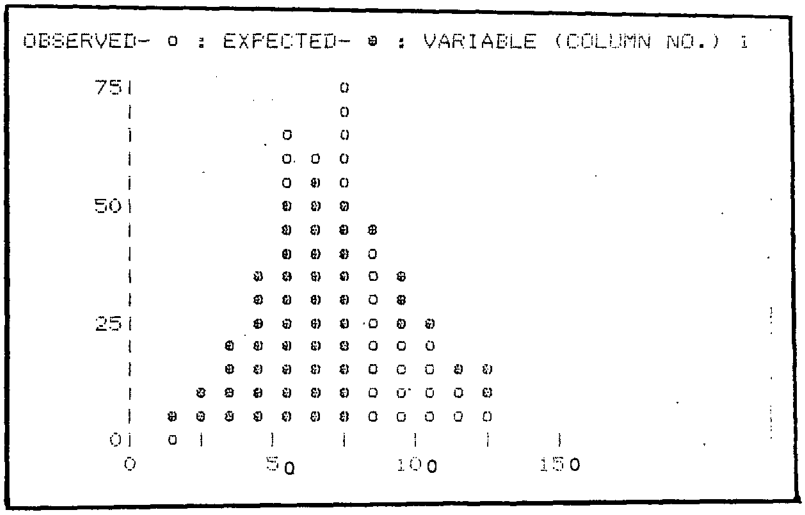


AIRPORT D, FLIGHT D2

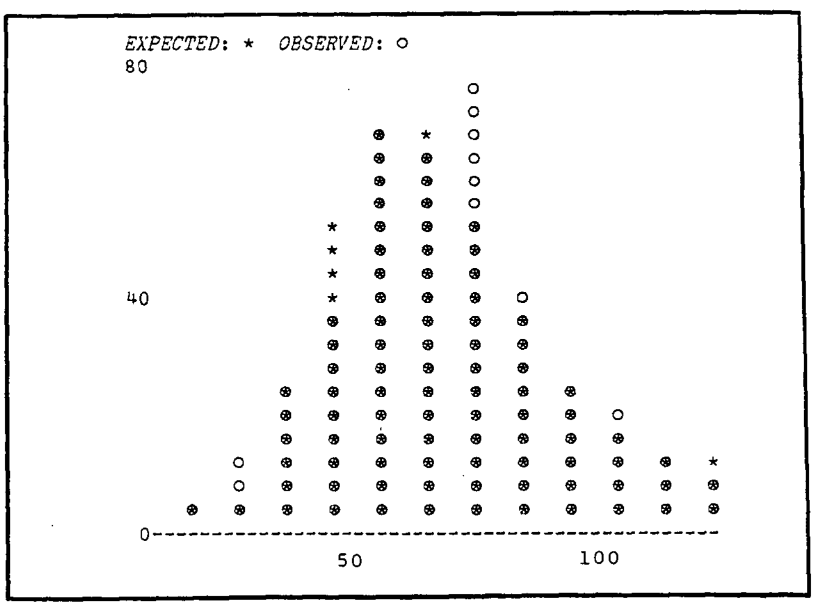
NORMAL



POISSON

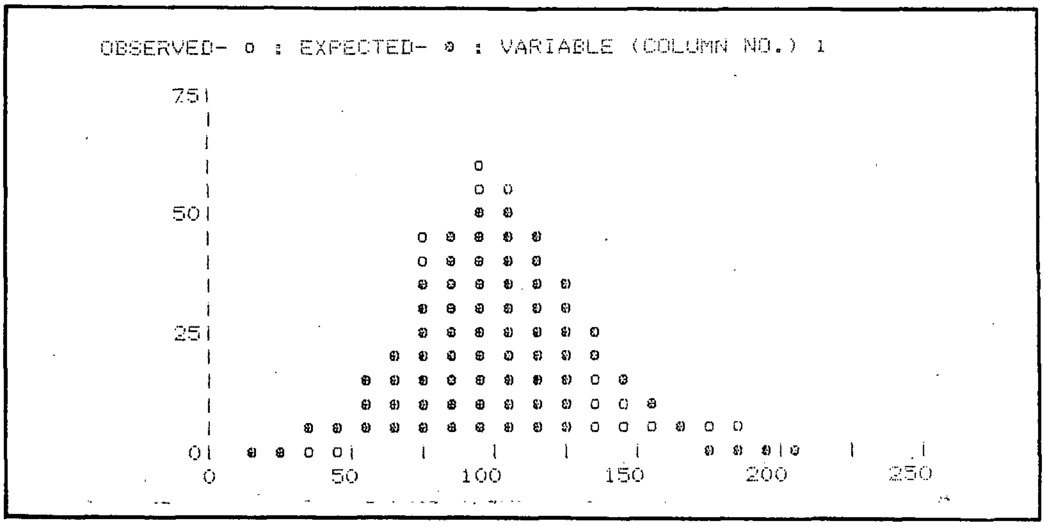


LOGNORMAL

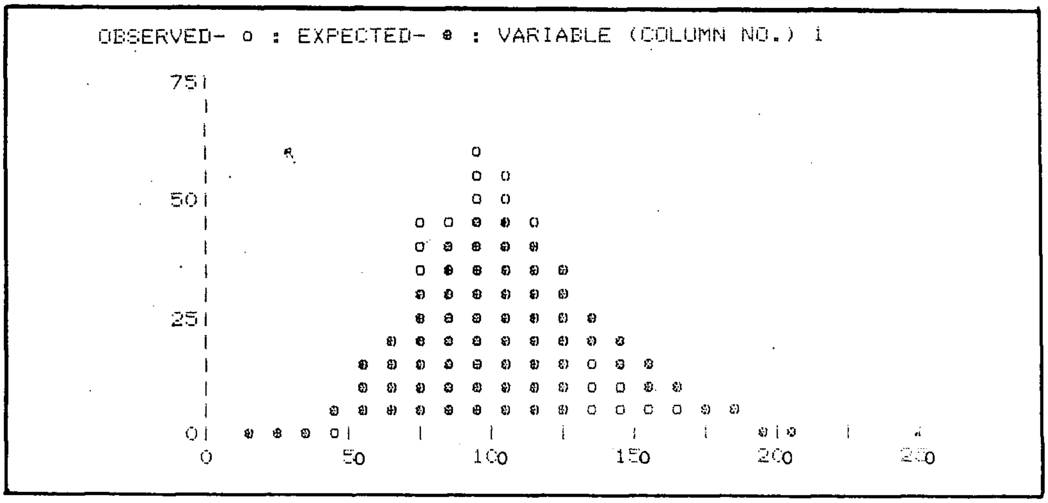


AIRPORT E, FLIGHT E1

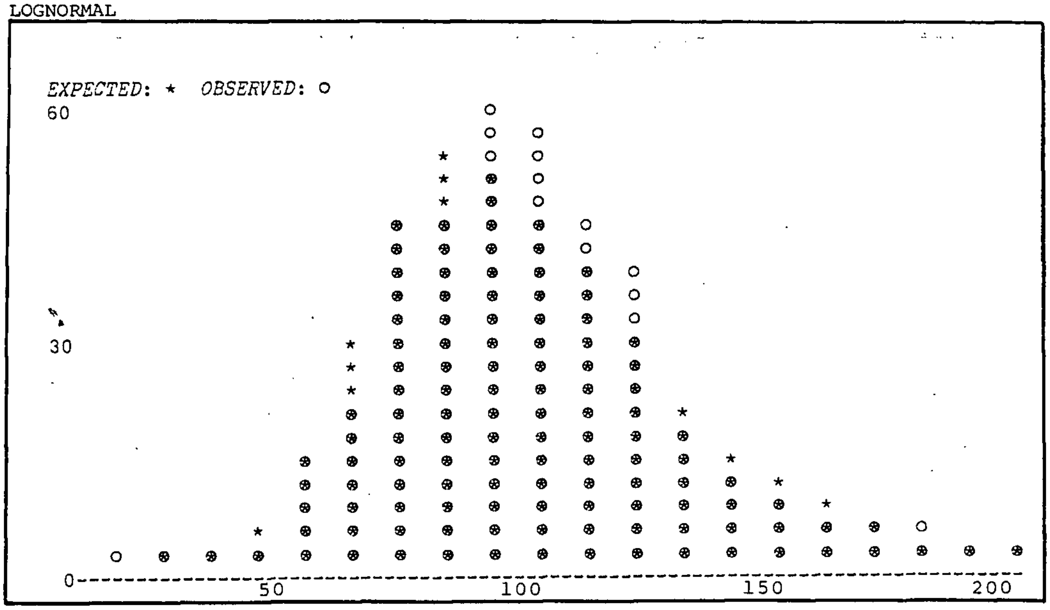
NORMAL



POISSON

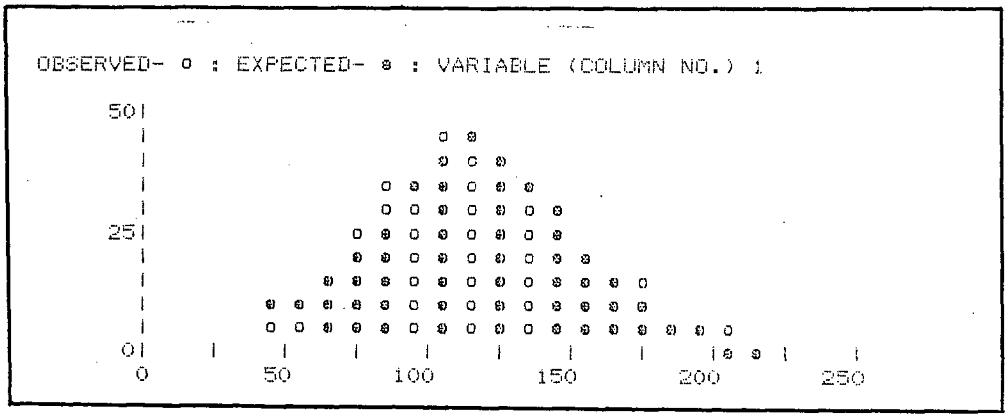


LOGNORMAL

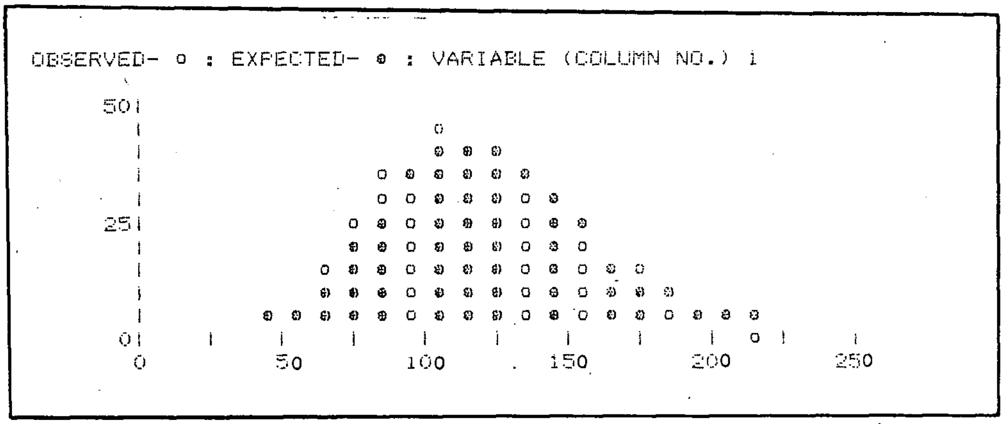


AIRPORT E, FLIGHT E2

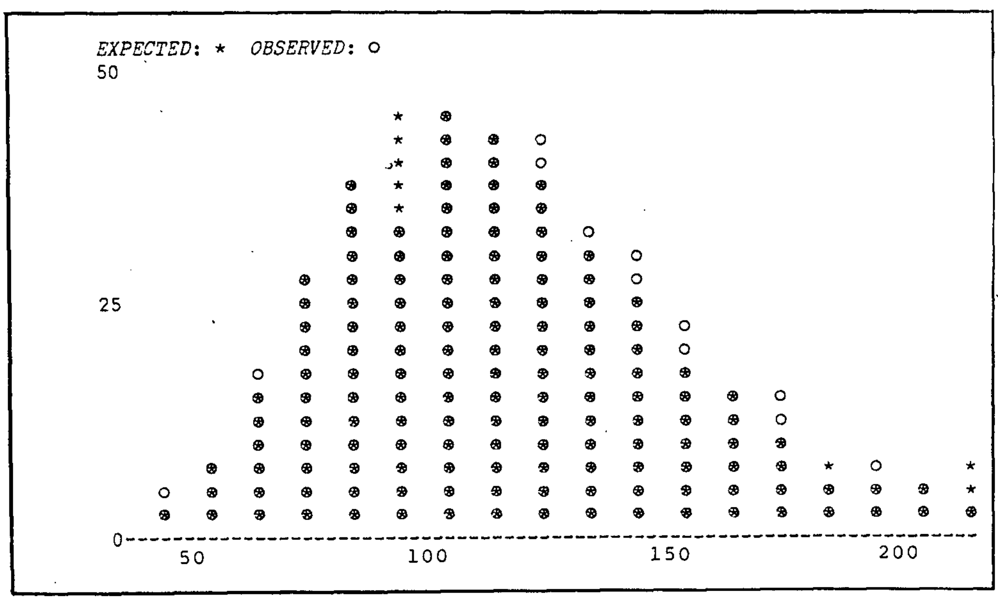
NORMAL



POISSON

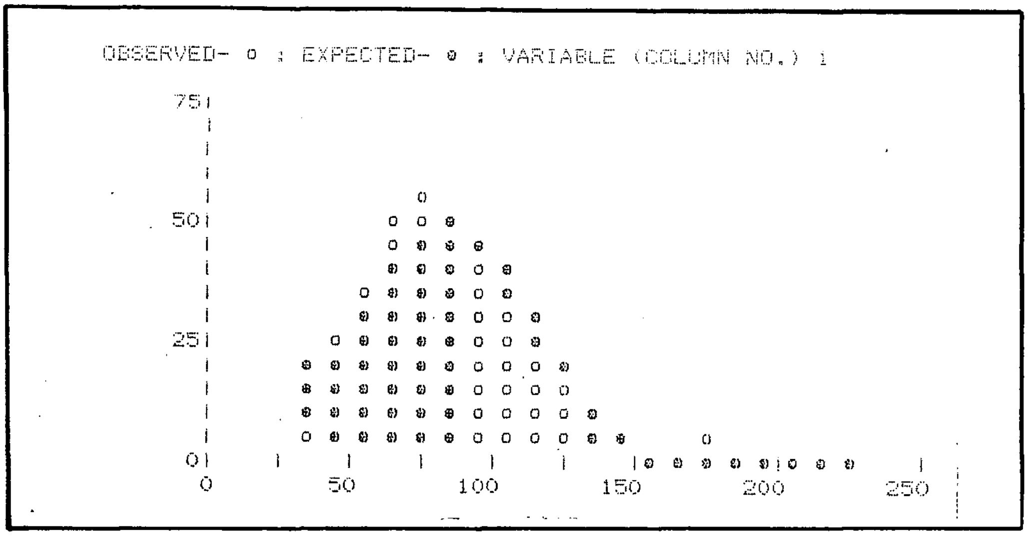


LOGNORMAL

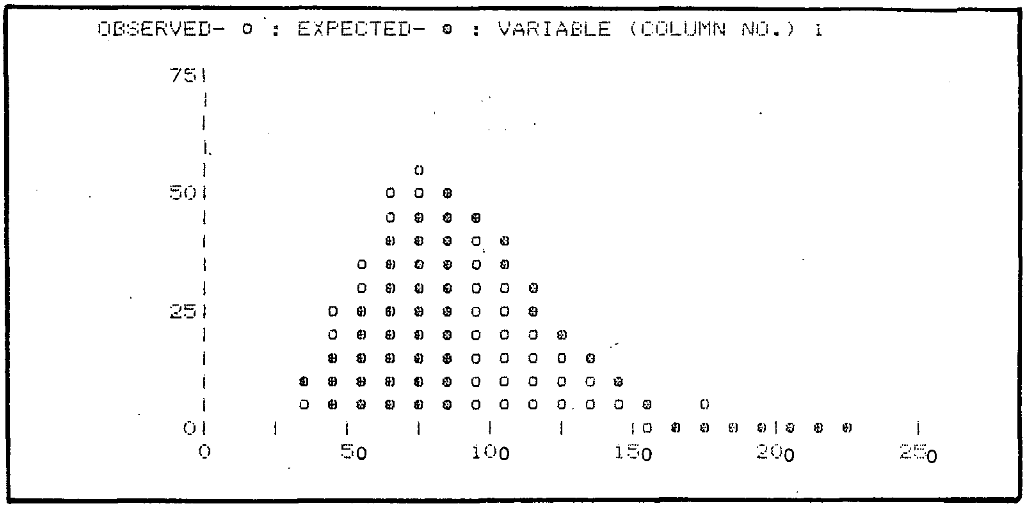


AIRPORT E, FLIGHT E3

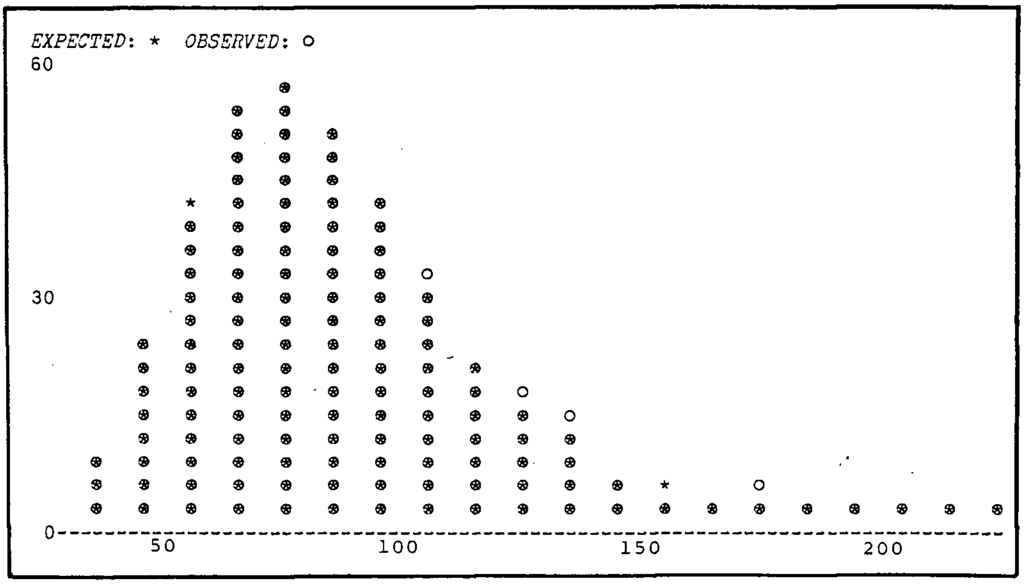
NORMAL



POISSON

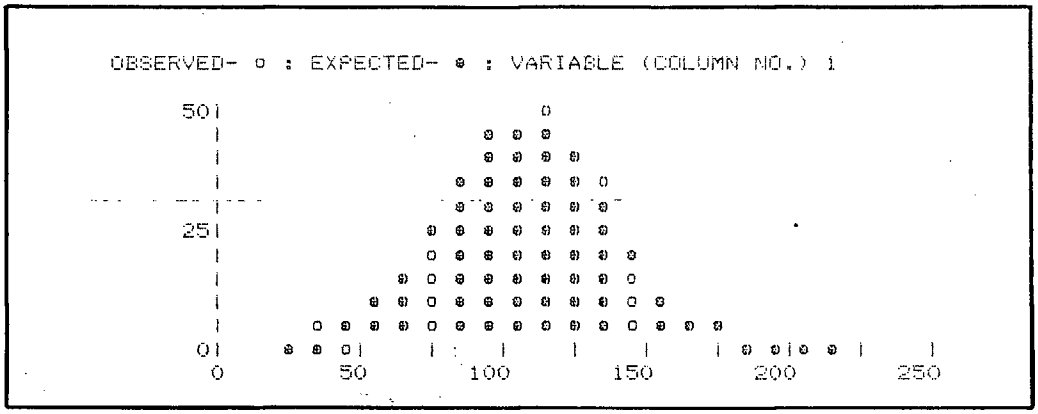


LOGNORMAL

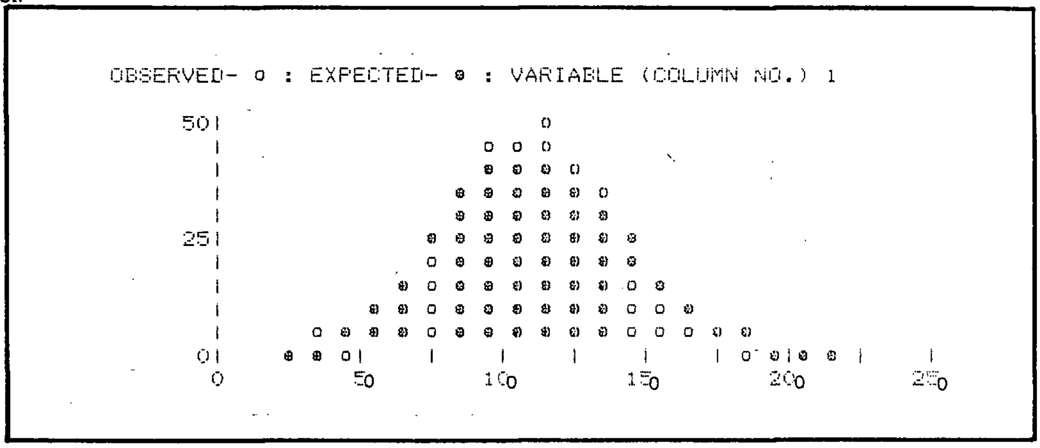


AIRPORT F, FLIGHT F1

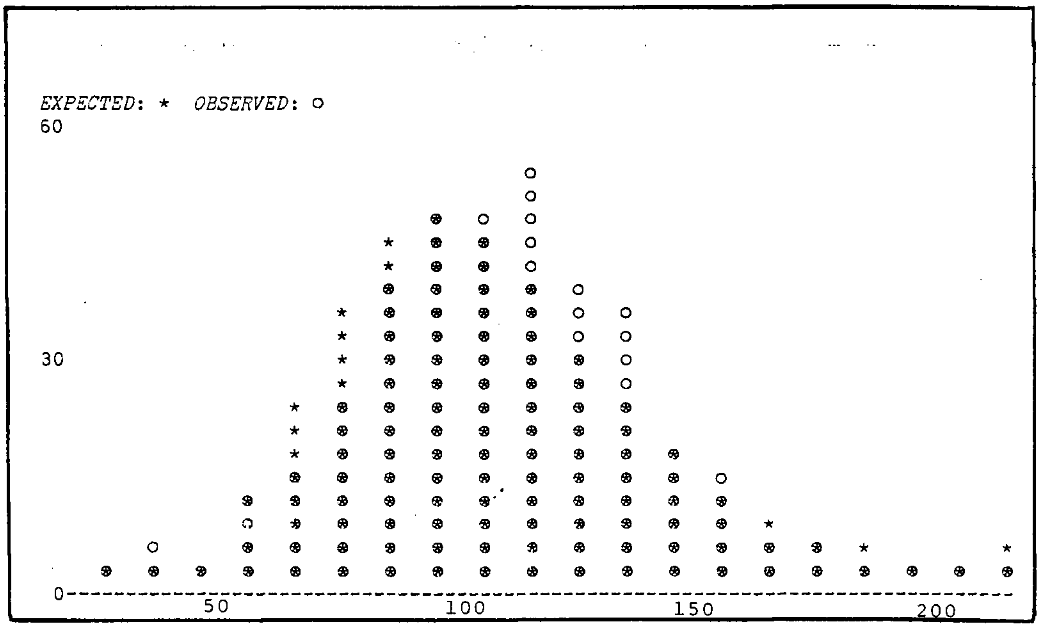
NORMAL



POISSON

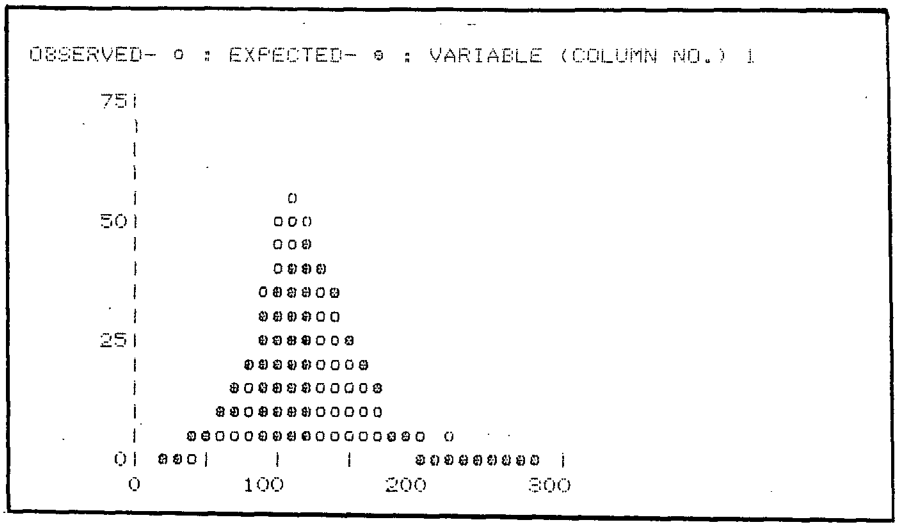


LOGNORMAL

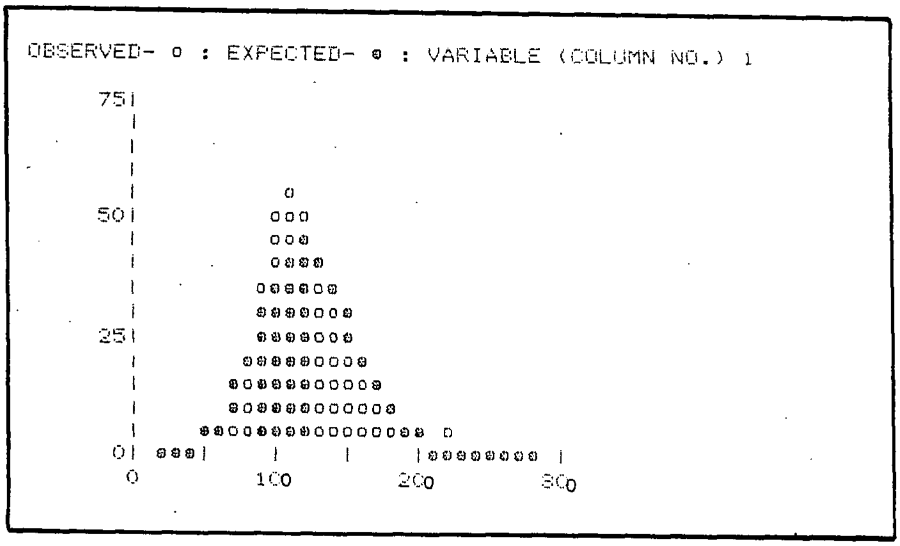


AIRPORT F, FLIGHT F2

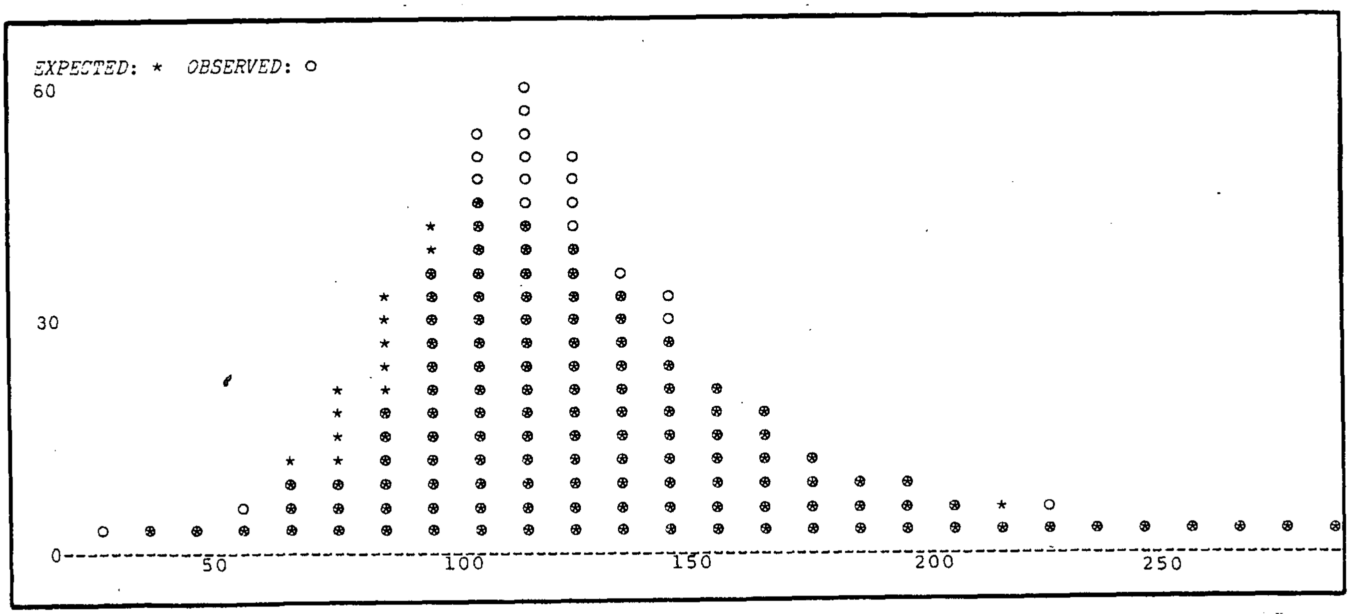
NORMAL



POISSON

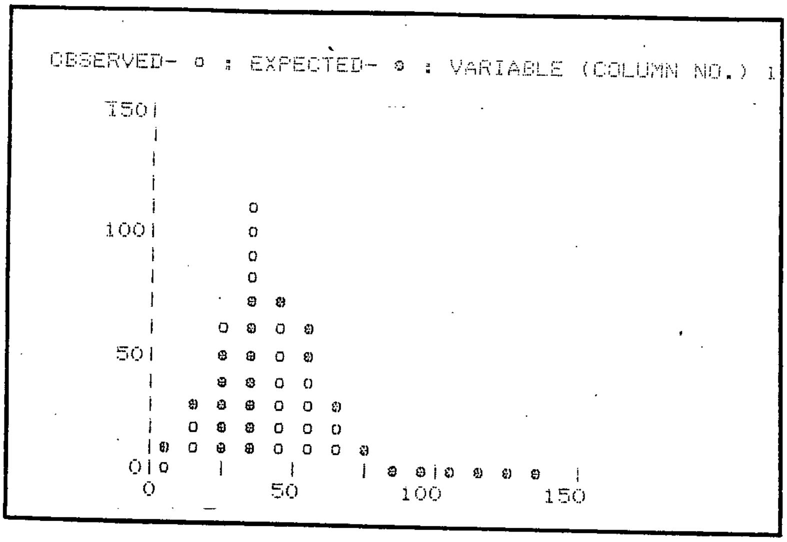


LOGNORMAL

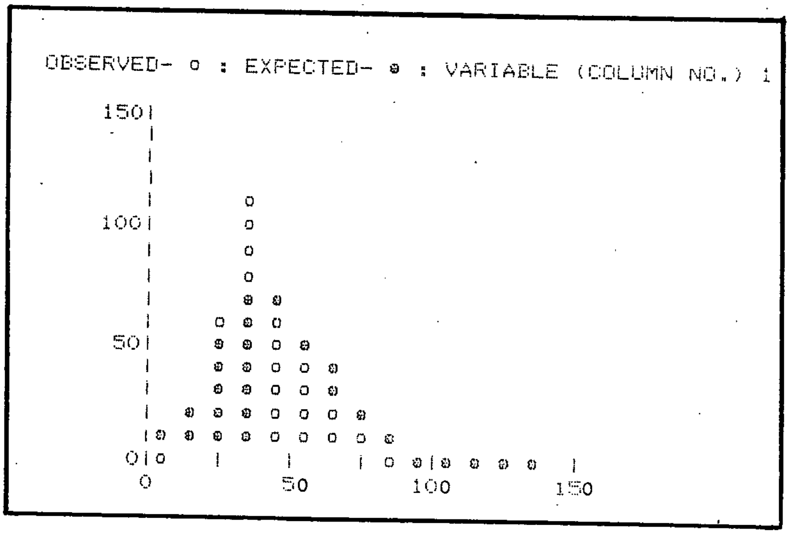


AIRPORT G, FLIGHT G1

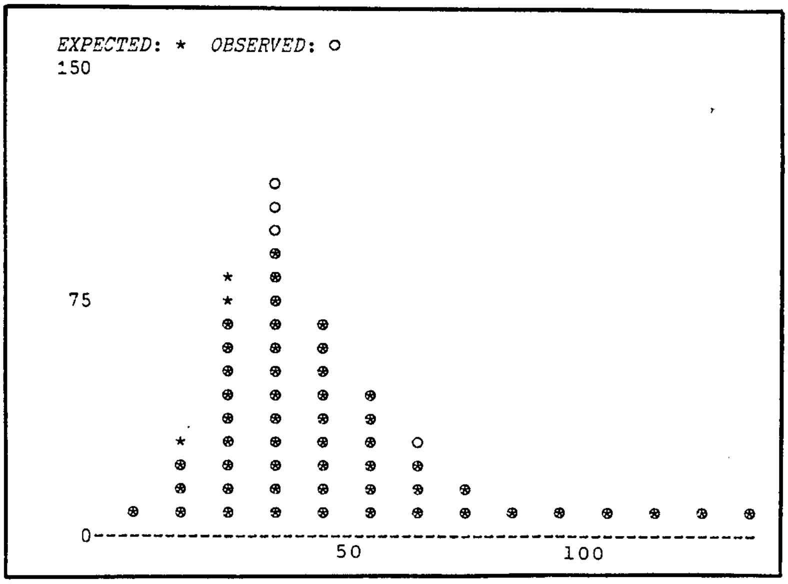
NORMAL



POISSON

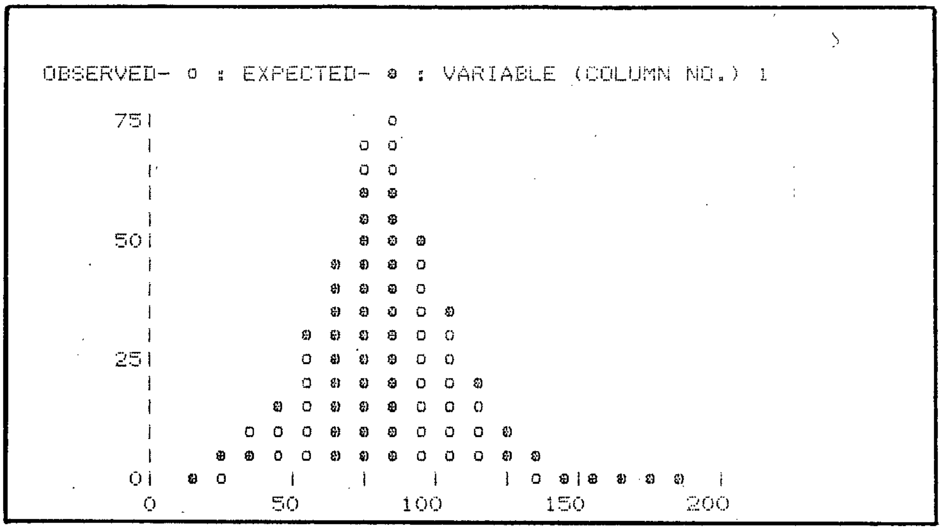


LOGNORMAL

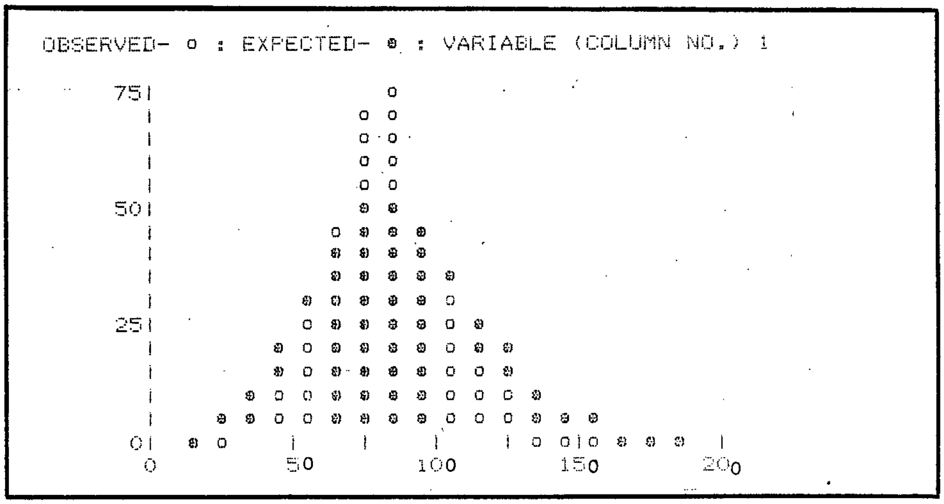


AIRPORT G, FLIGHT G2

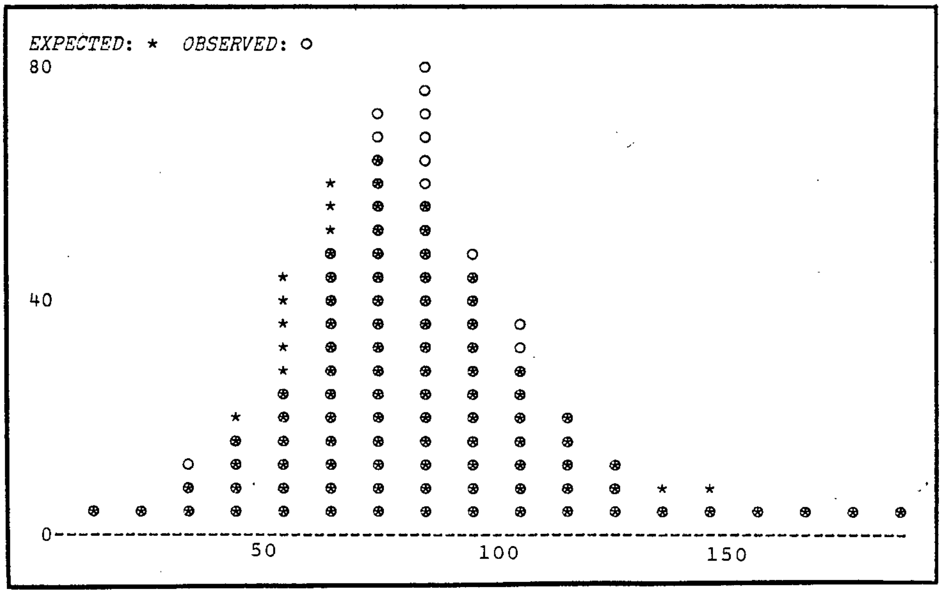
NORMAL



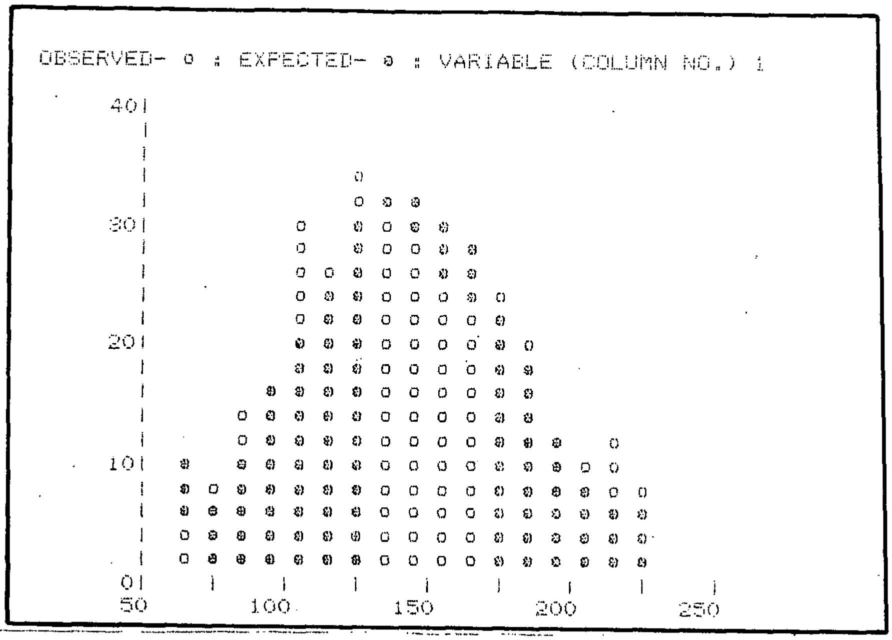
POISSON



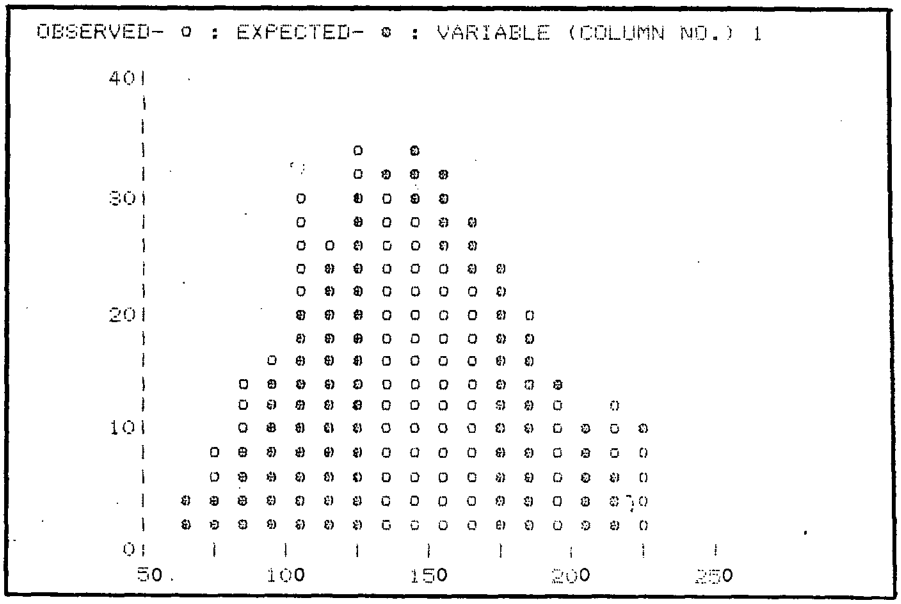
LOGNORMAL



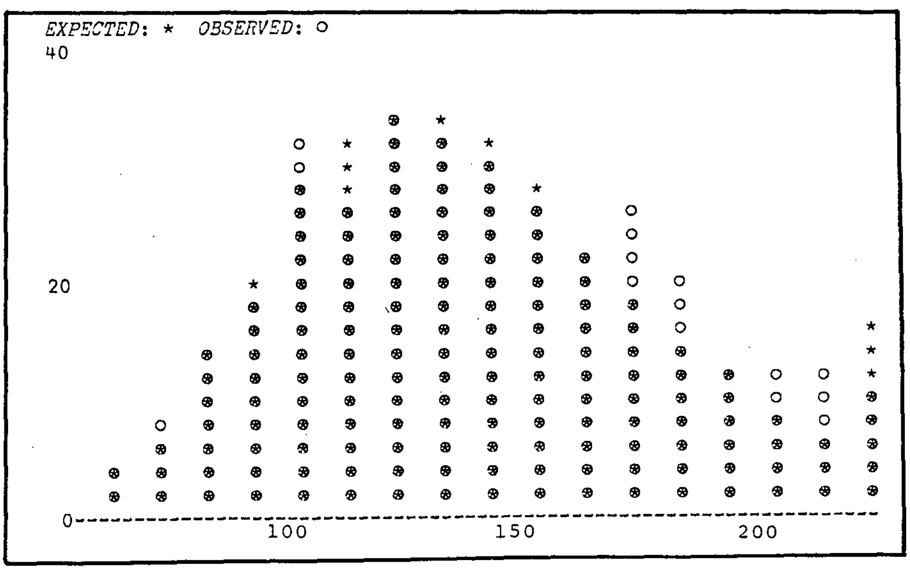
NORMAL



POISSON

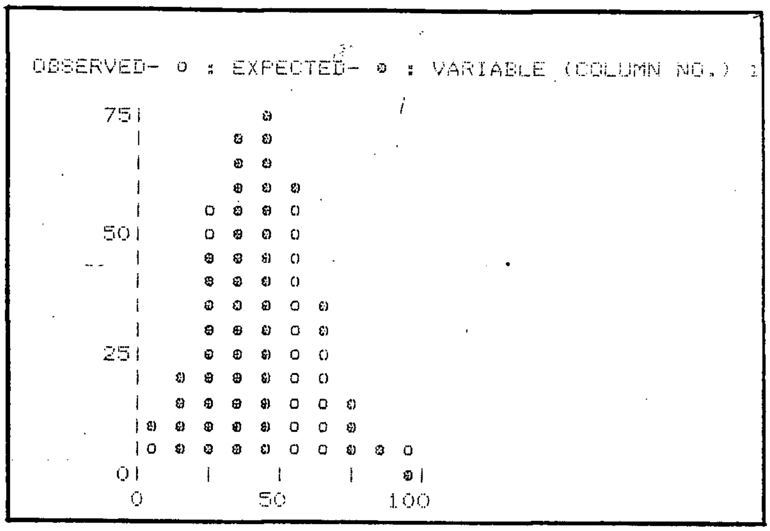


LOGNORMAL

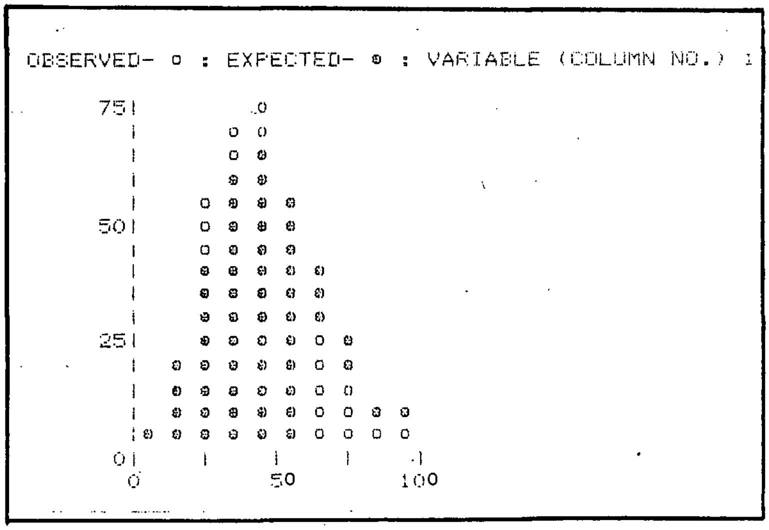


AIRPORT G, FLIGHT G4

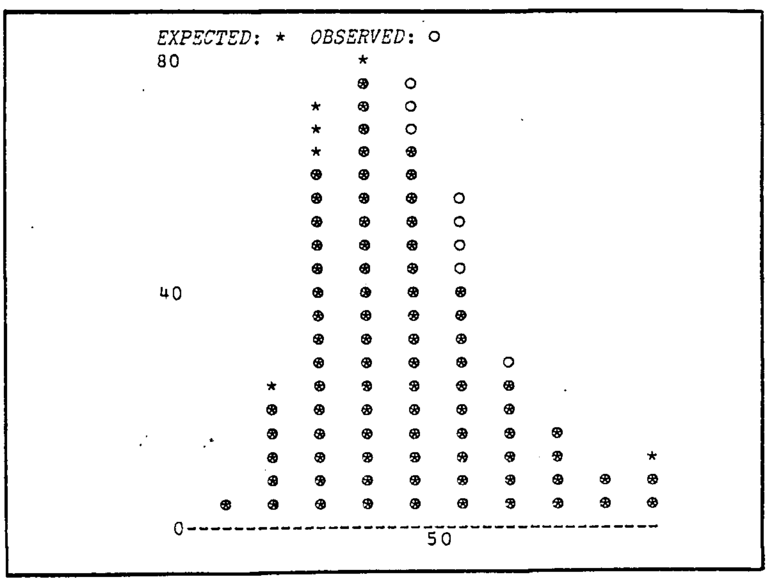
NORMAL



POISSON

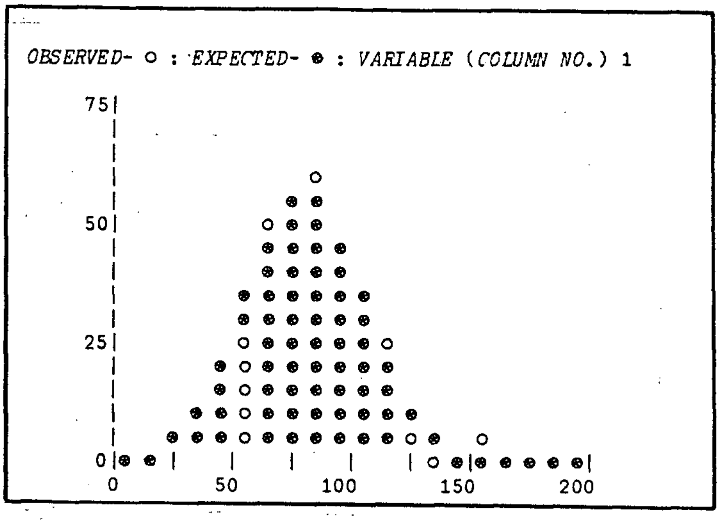


LOGNORMAL

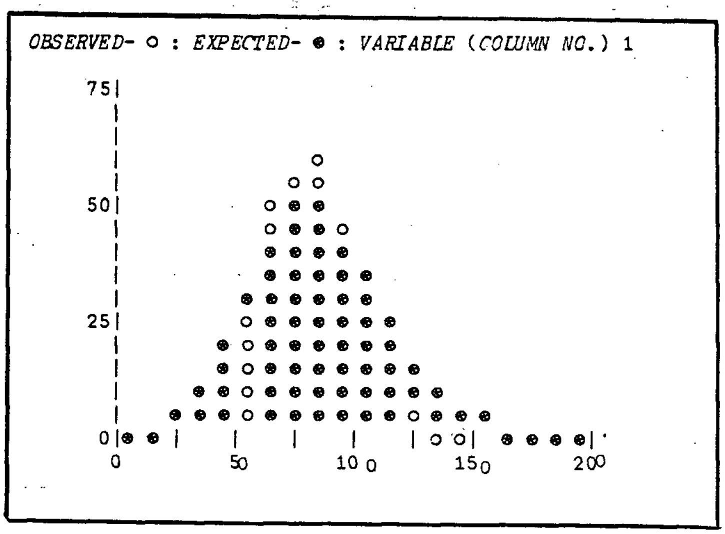


AIRPORT G, FLIGHT G5

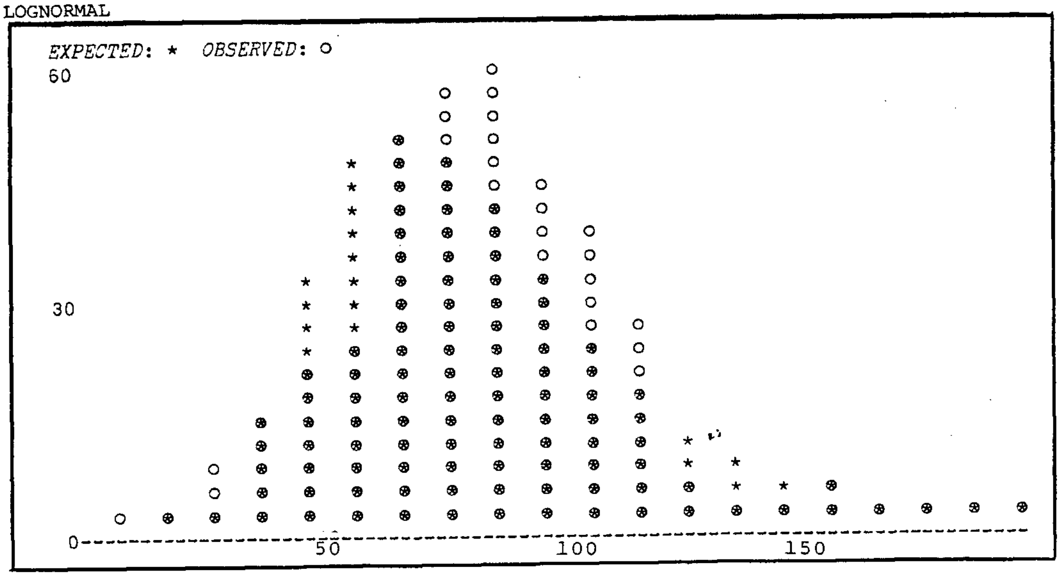
NORMAL



POISSON

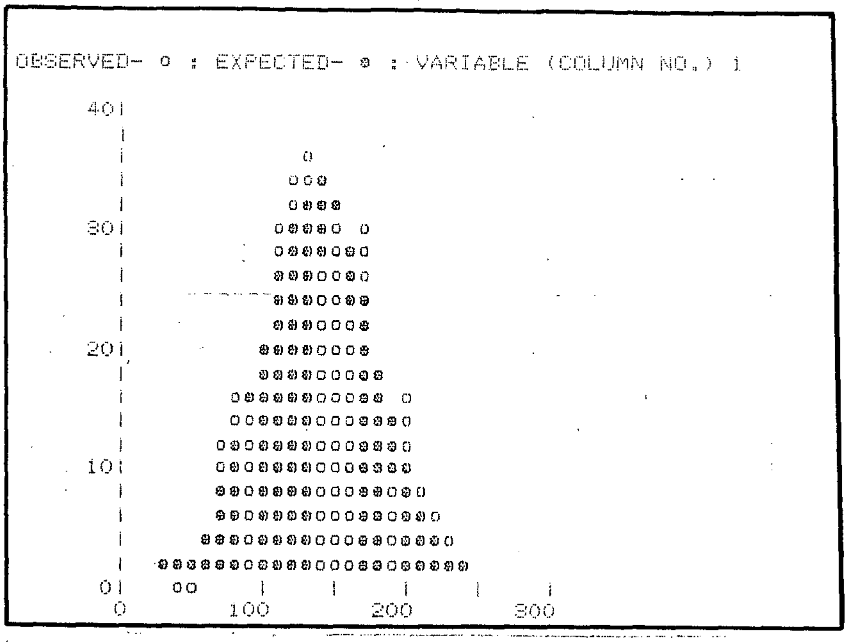


LOGNORMAL

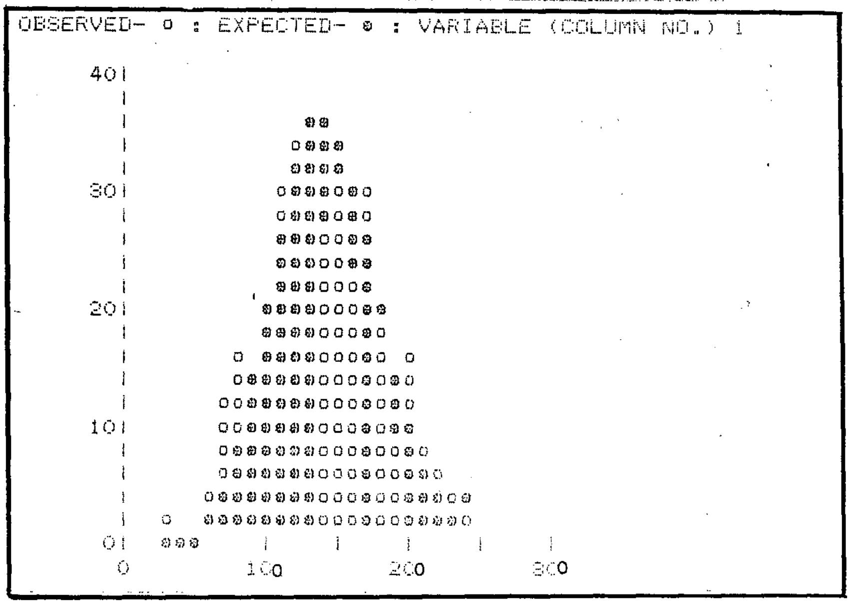


AIRPORT G, FLIGHT G6

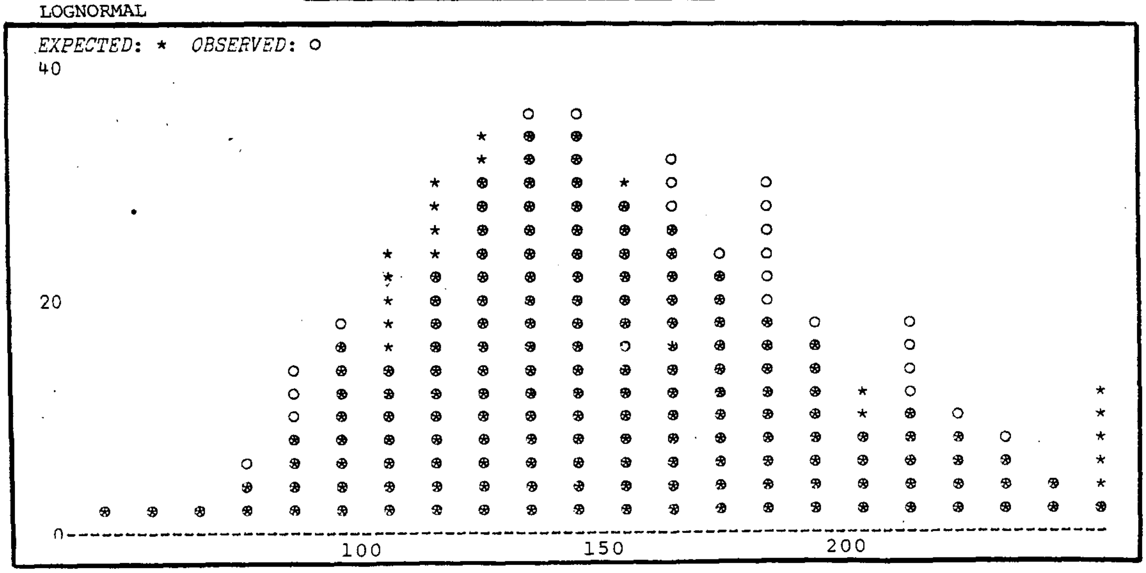
NORMAL



POISSON

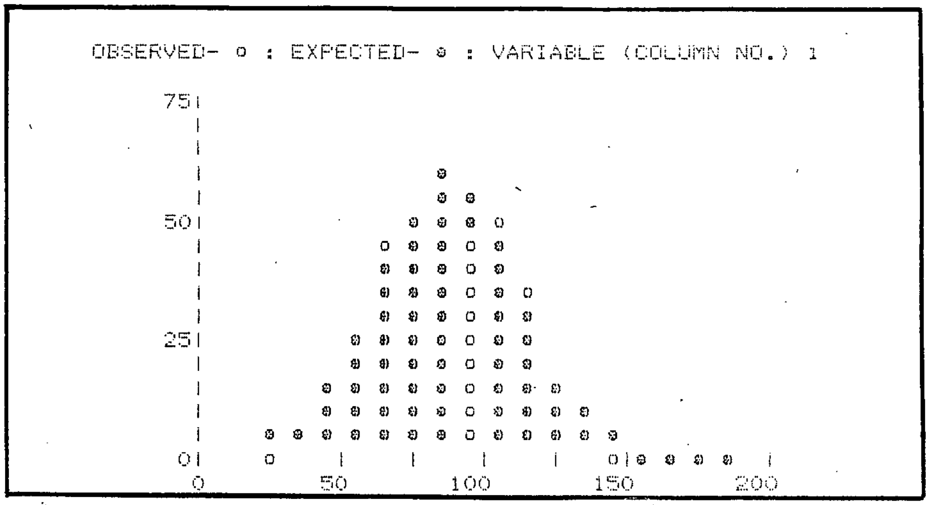


LOGNORMAL

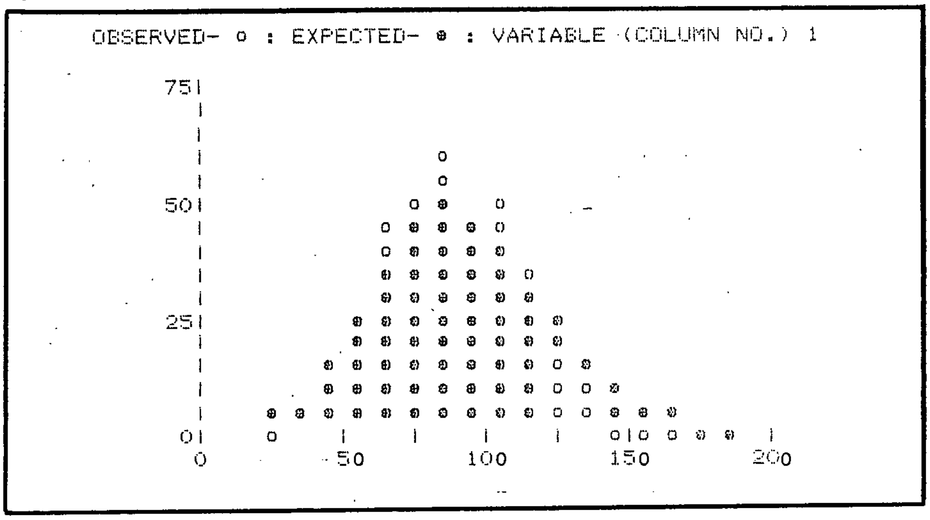


AIRPORT H, FLIGHT H1

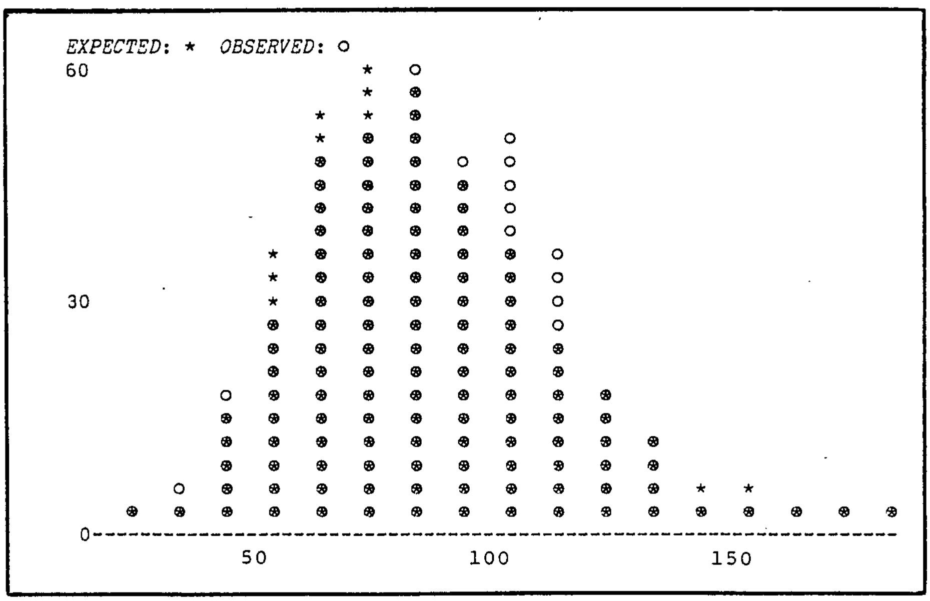
NORMAL



POISSON

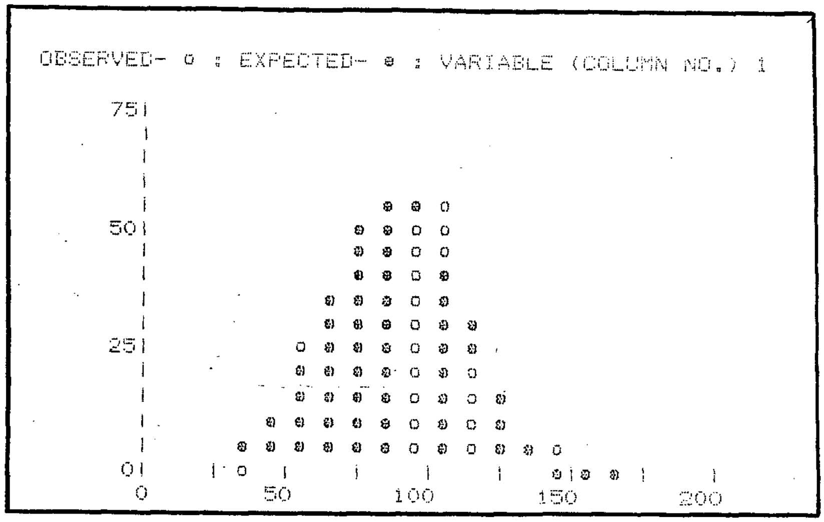


LOGNORMAL

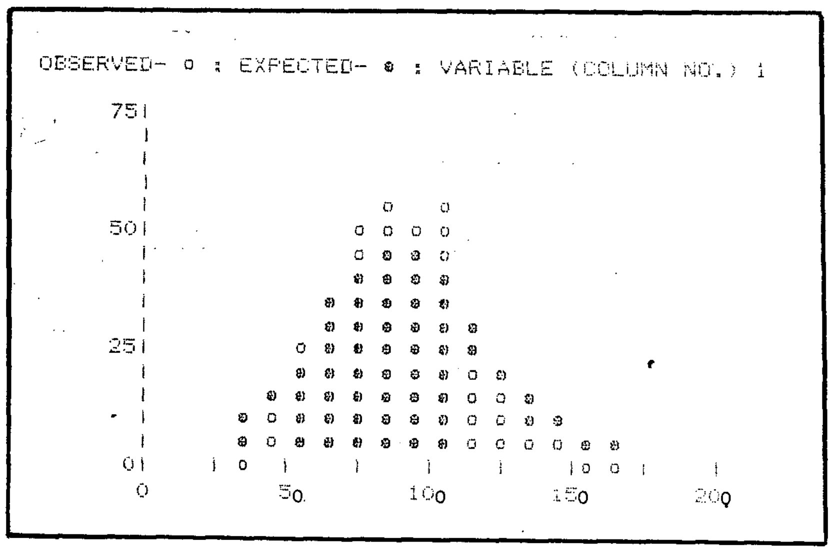


AIRPORT H, FLIGHT H2

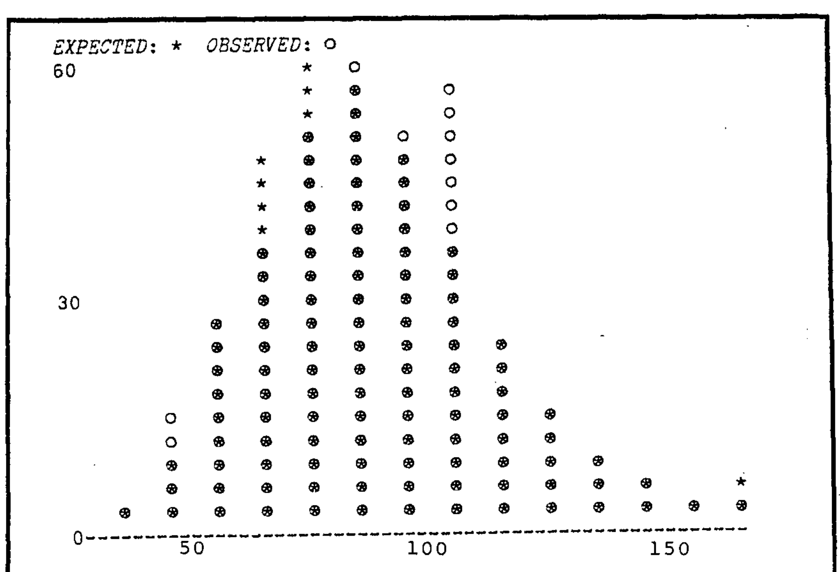
NORMAL



POISSON

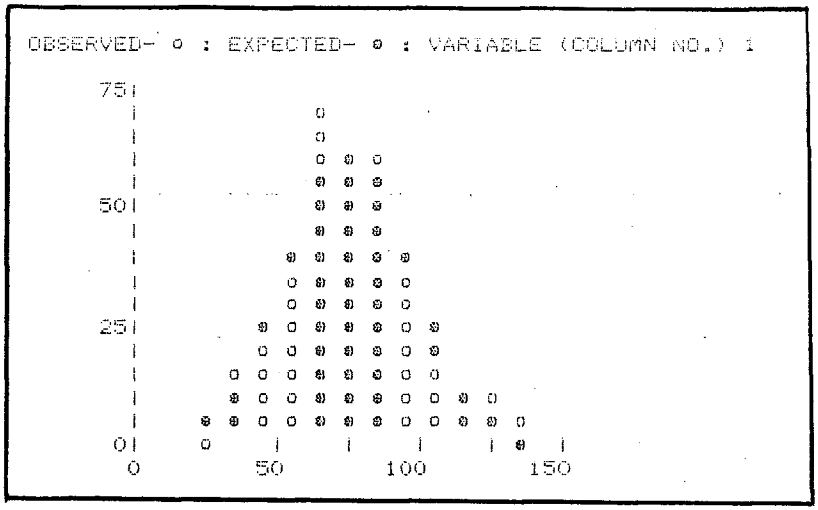


LOGNORMAL

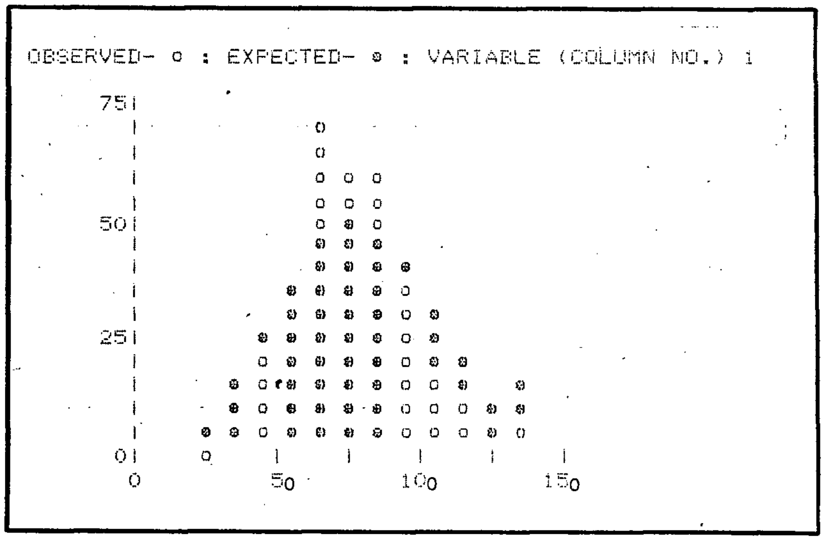


AIRPORT H, FLIGHT H3

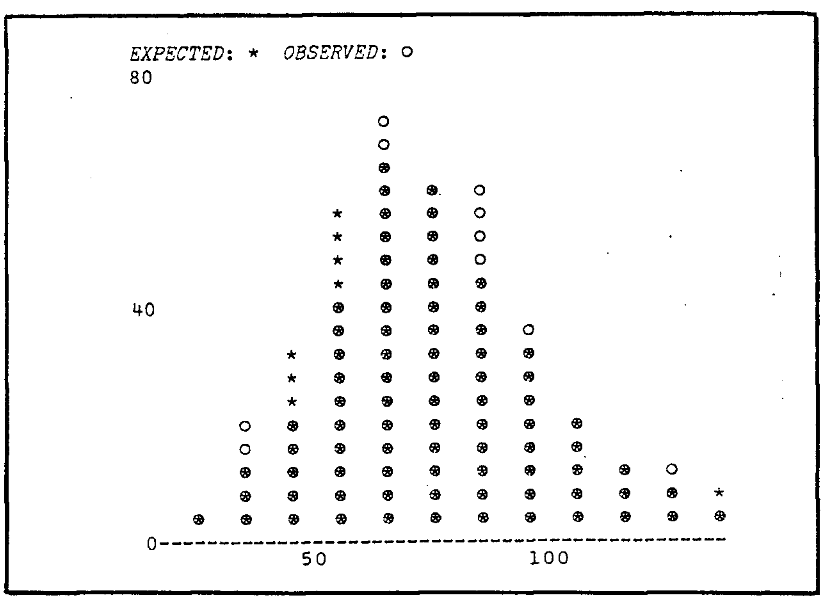
NORMAL



POISSON

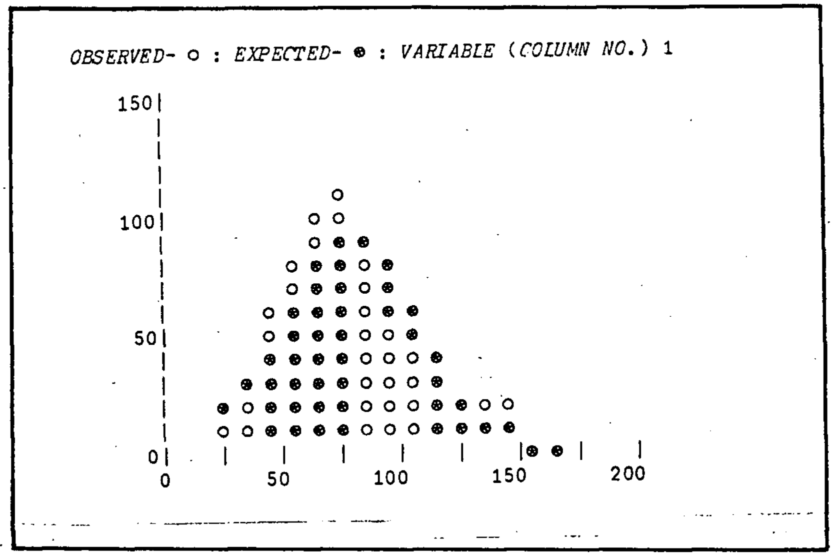


LOGNORMAL

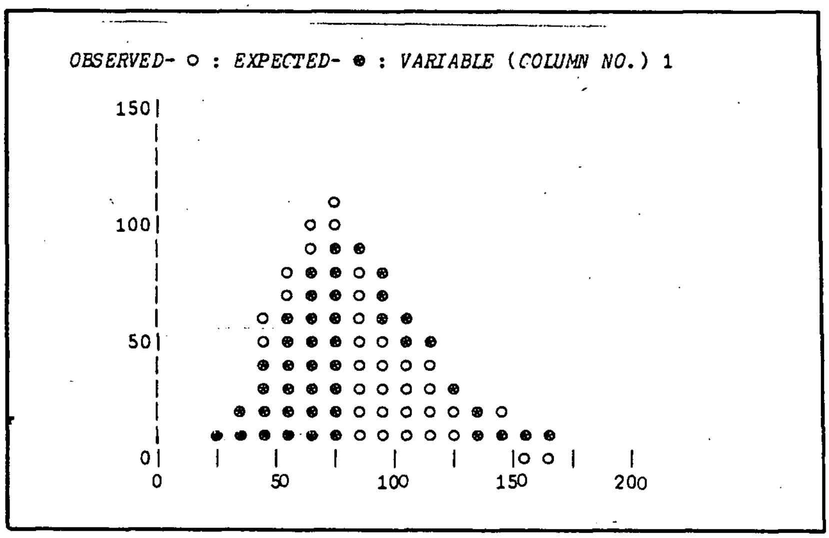


AIRPORT A, 1981

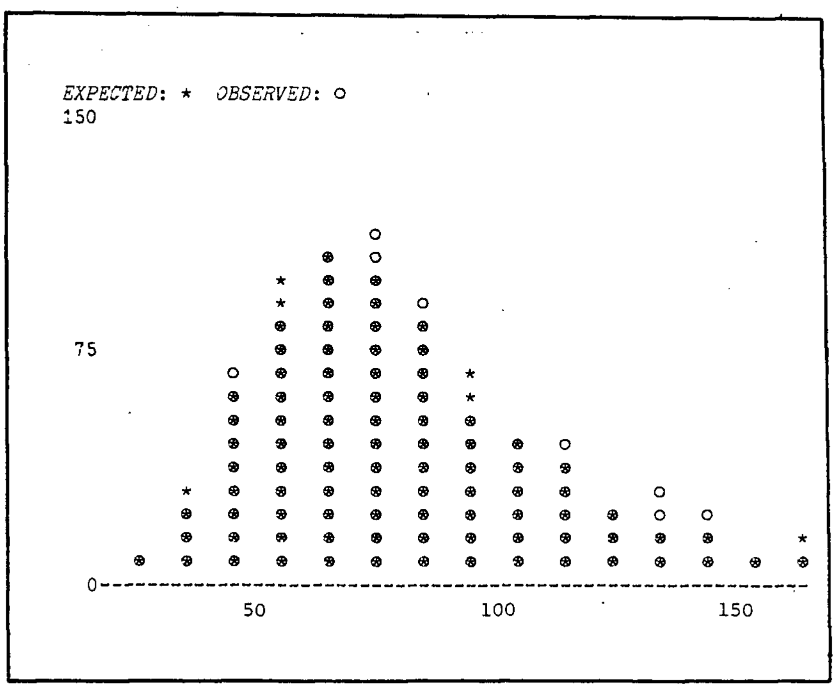
NORMAL



POISSON

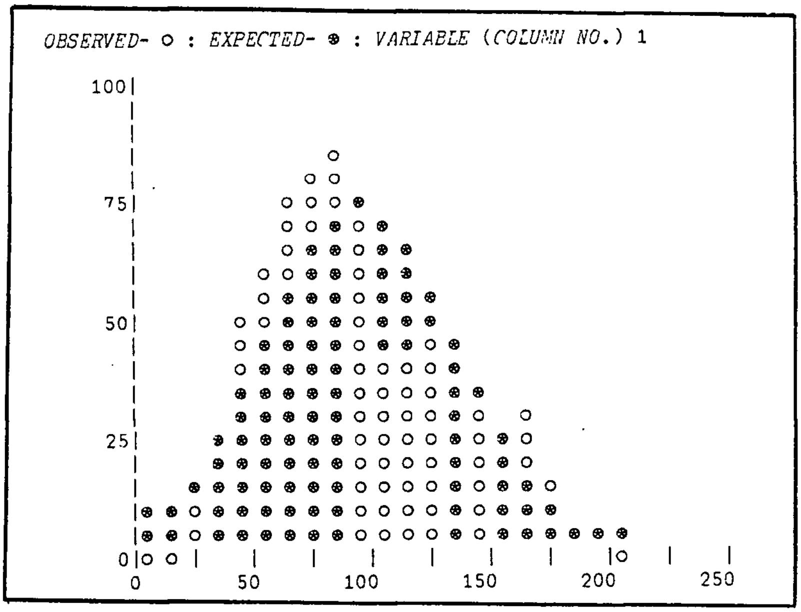


LOGNORMAL

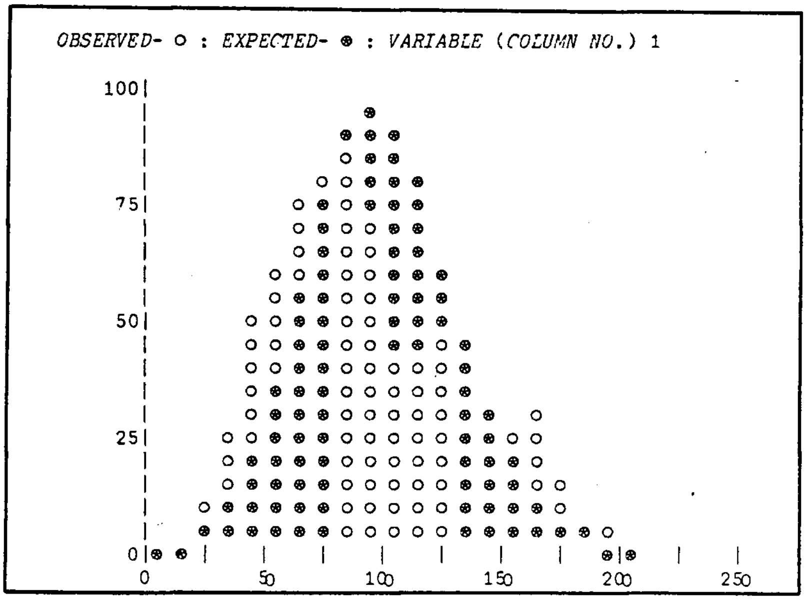


AIRPORT B, 1978

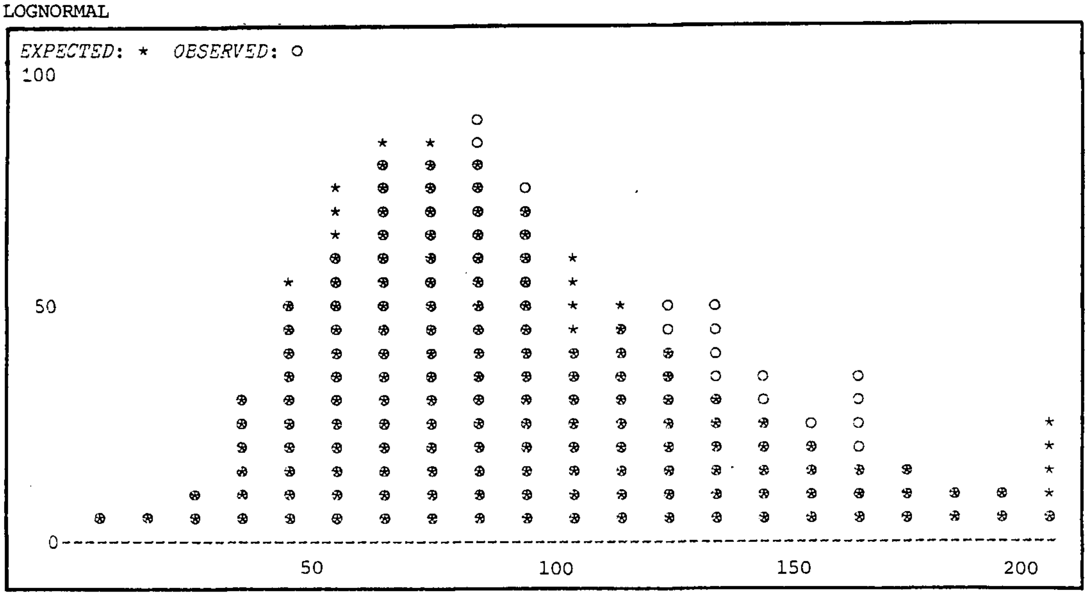
NORMAL



POISSON

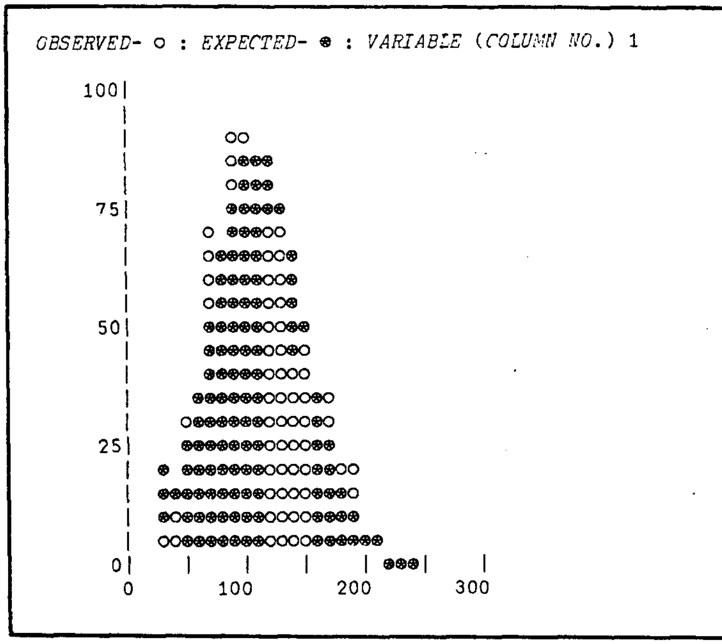


LOGNORMAL

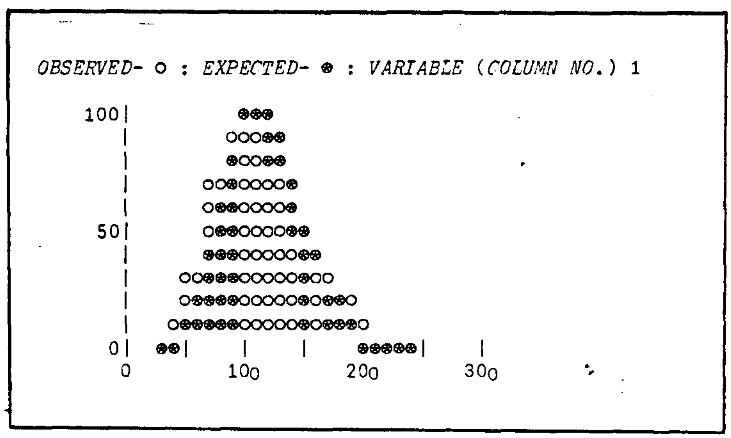


AIRPORT B, 1979

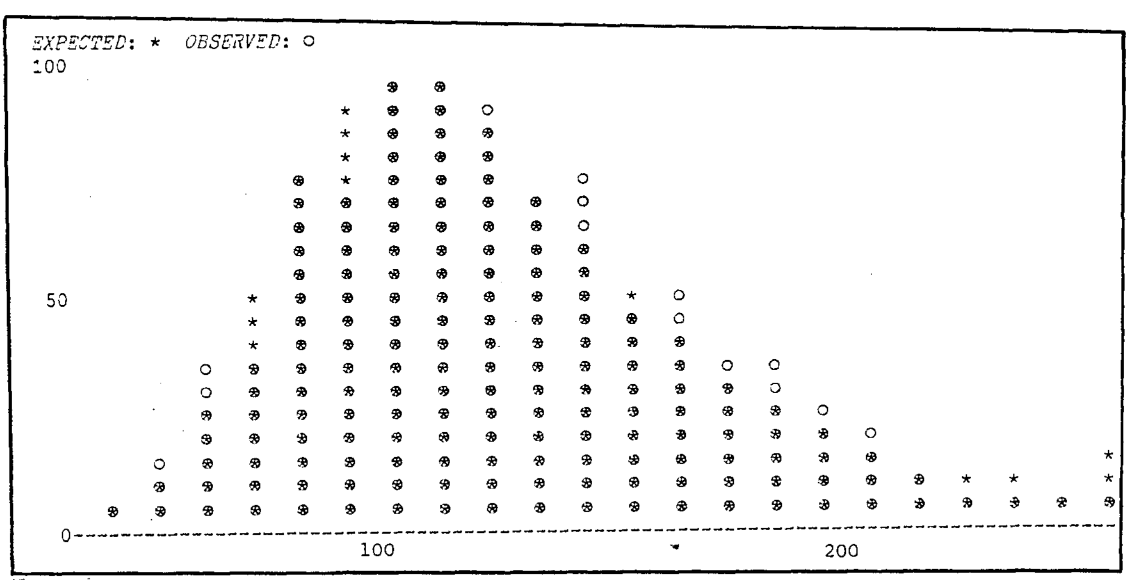
NORMAL



POISSON

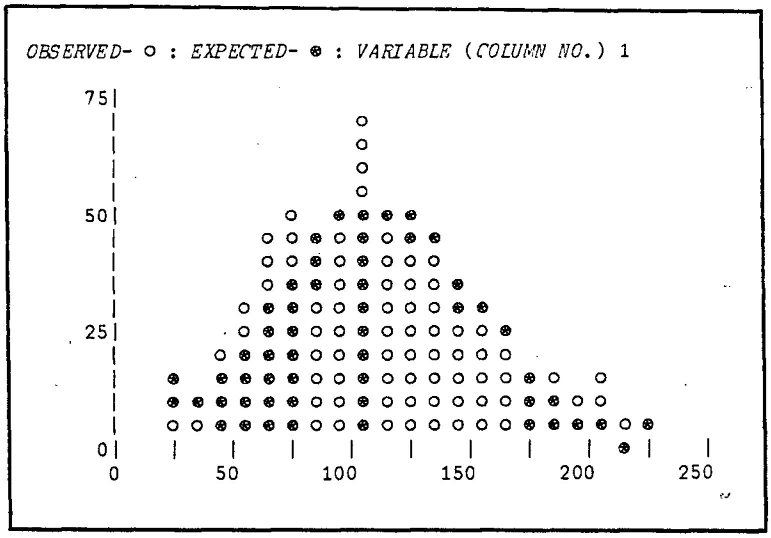


LOGNORMAL

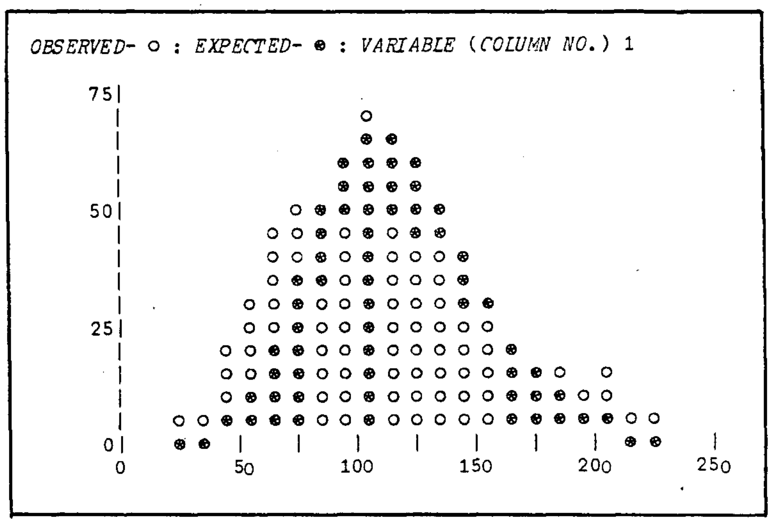


AIRPORT B, 1980

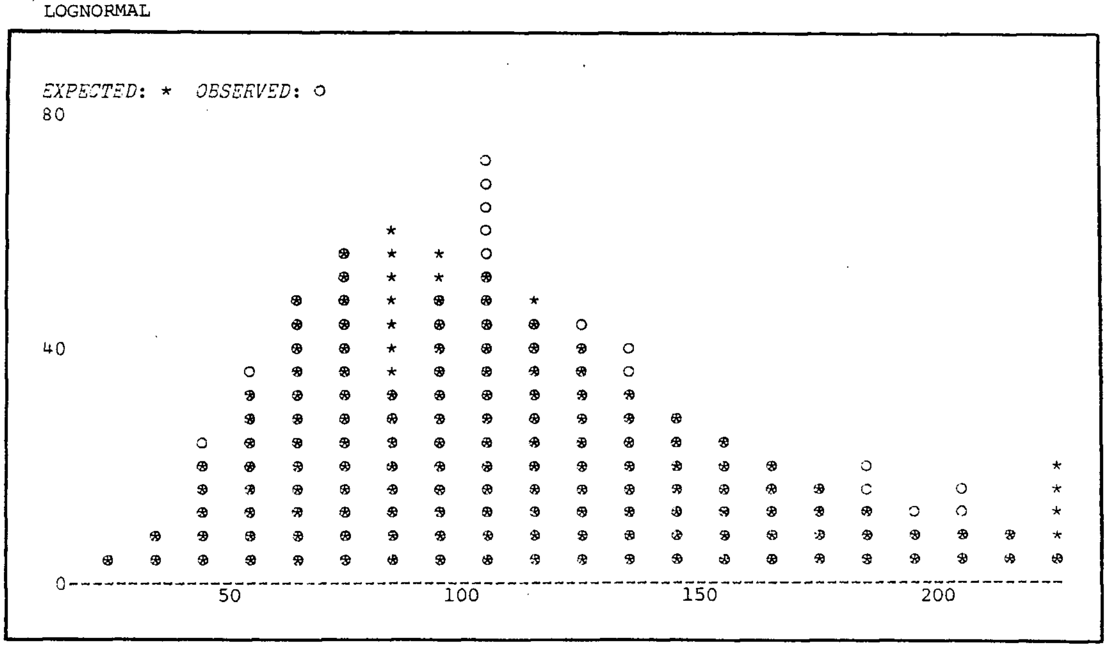
NORMAL



POISSON

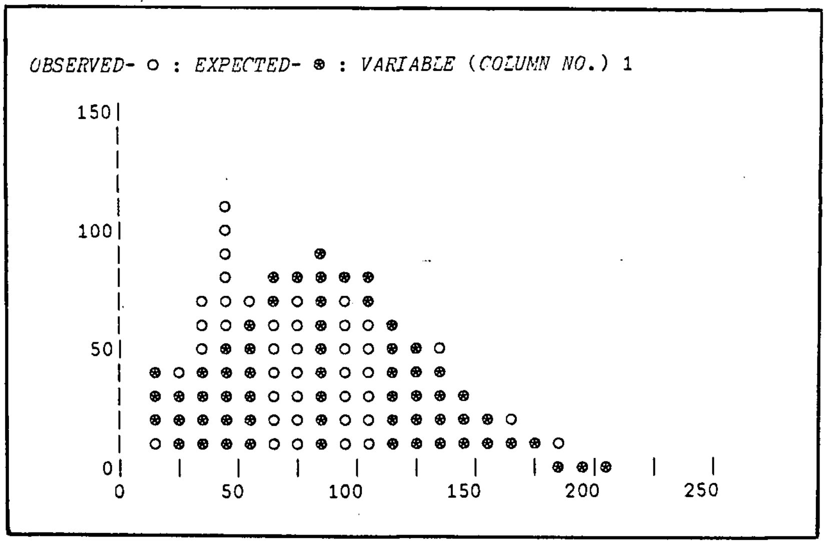


LOGNORMAL

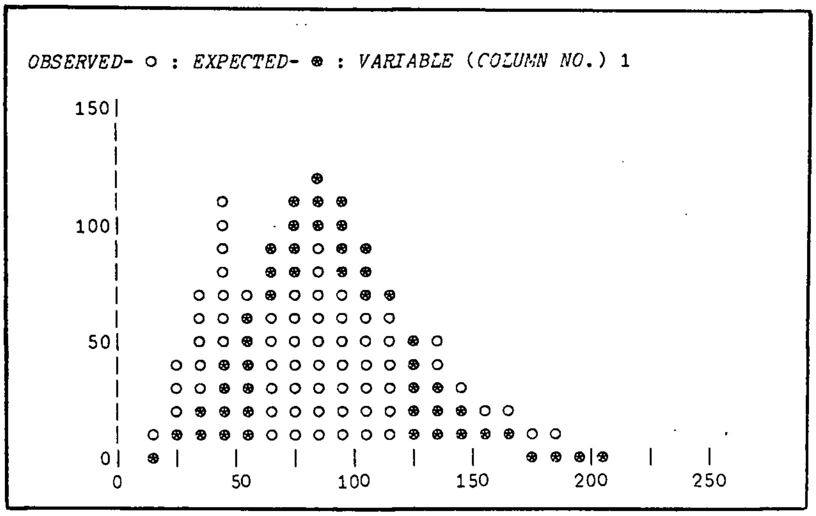


AIRPORT B, 1981

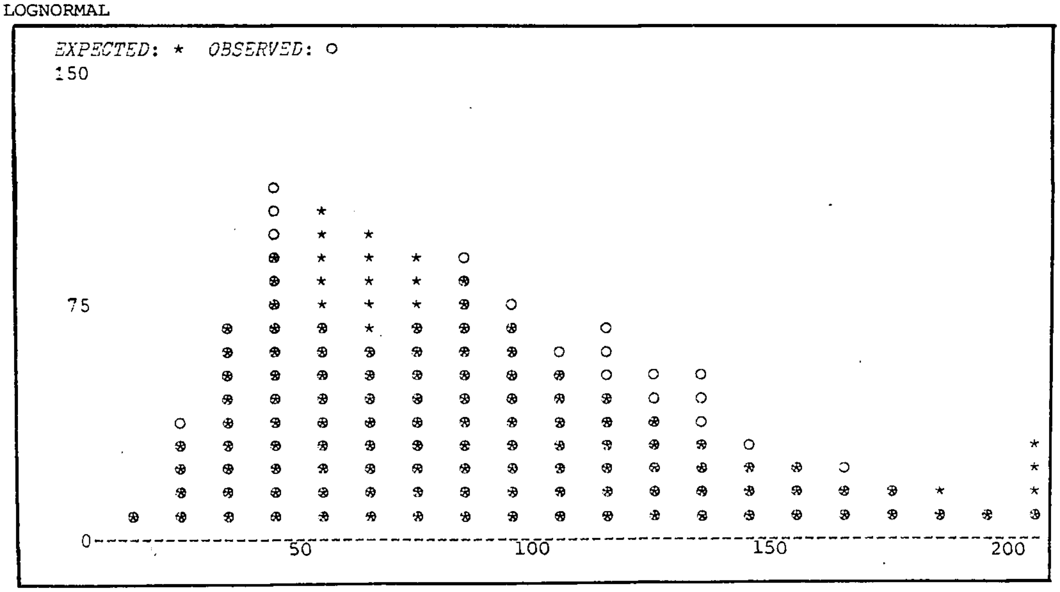
NORMAL



POISSON

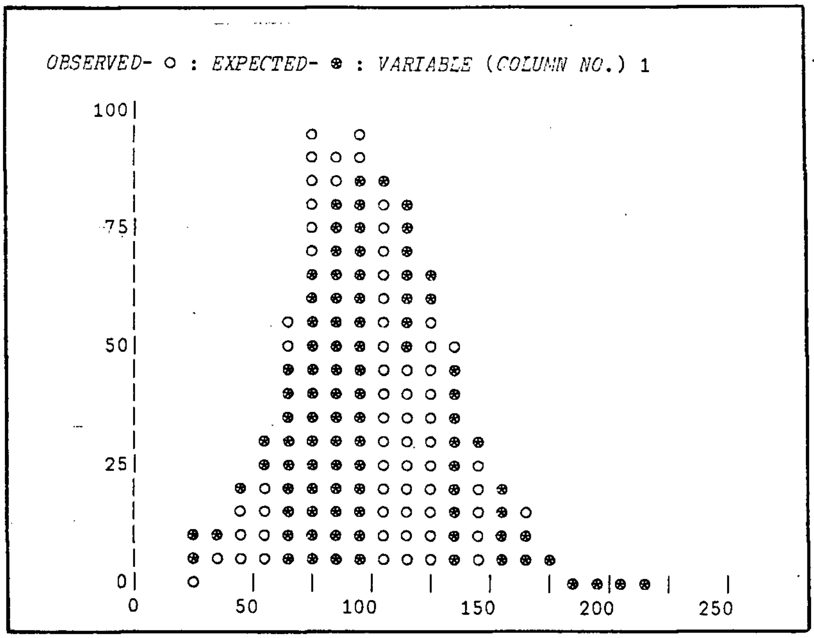


LOGNORMAL

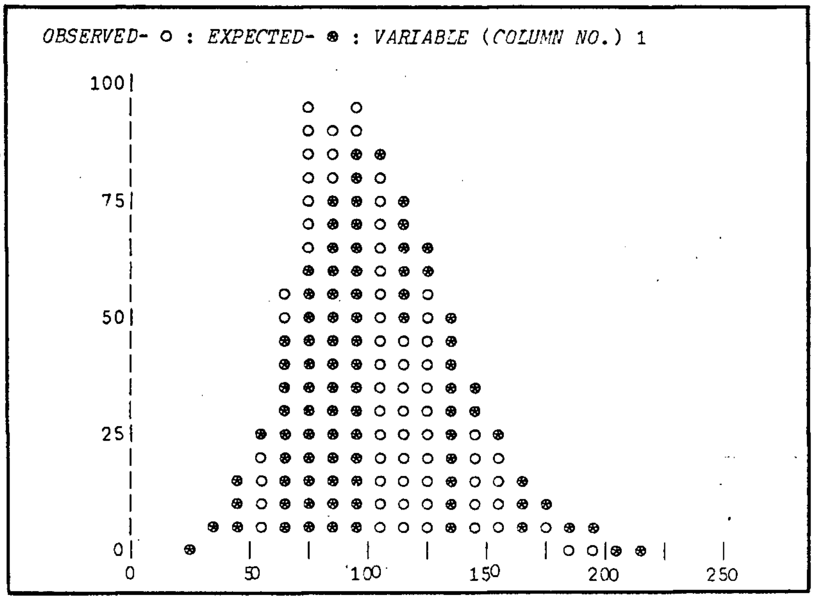


AIRPORT B, 1982

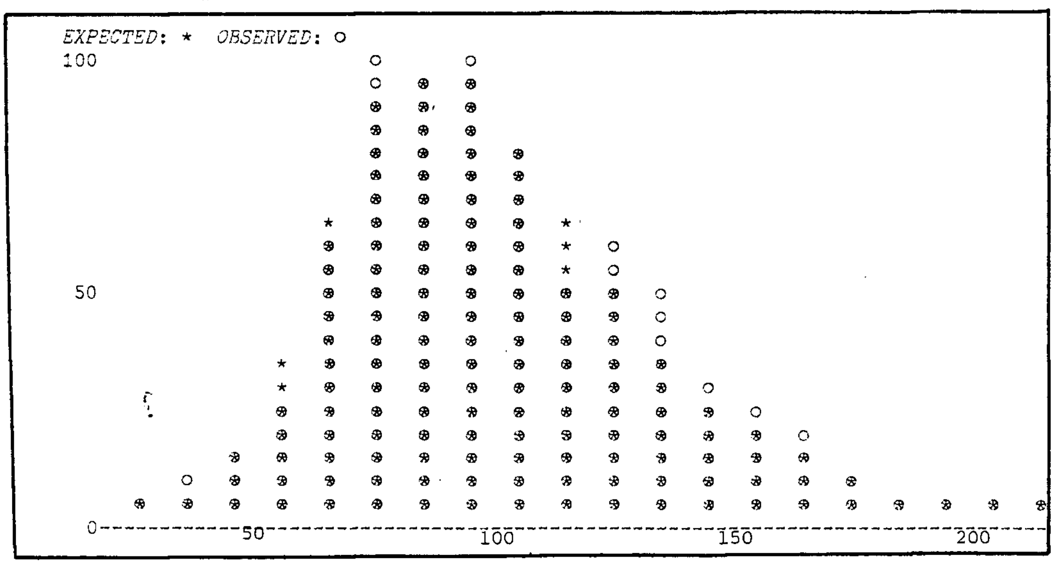
NORMAL



POISSON

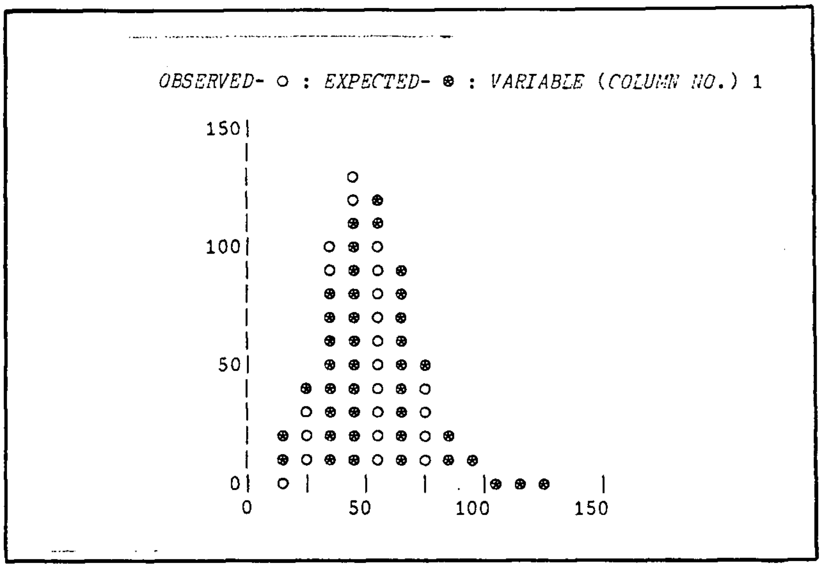


LOGNORMAL

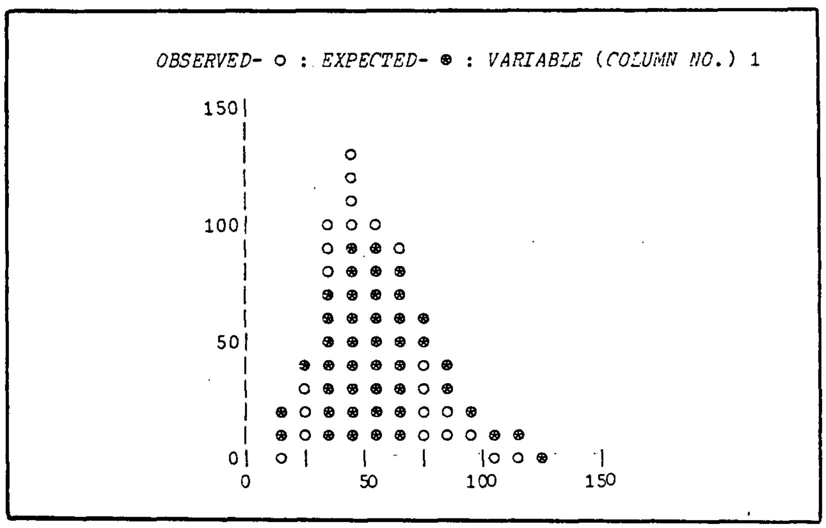


AIRPORT C, 1980

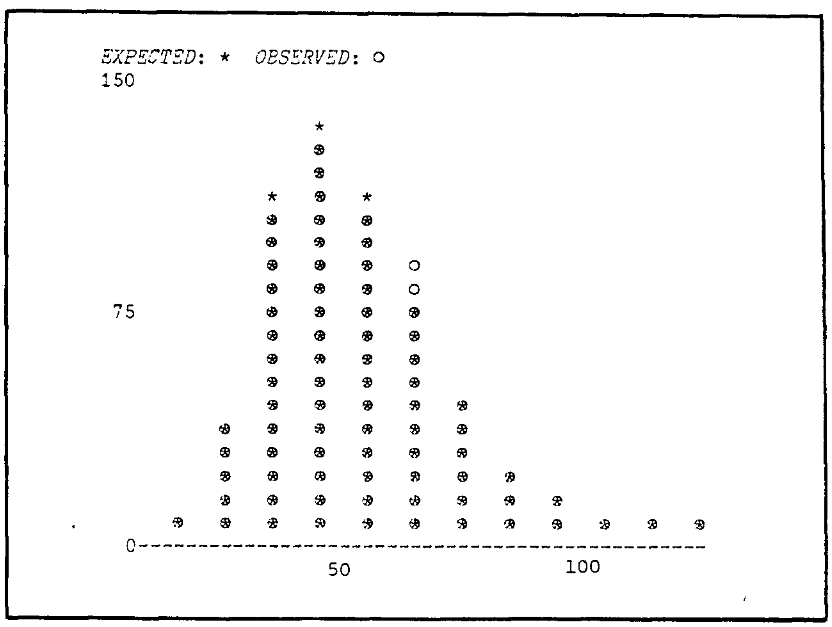
NORMAL



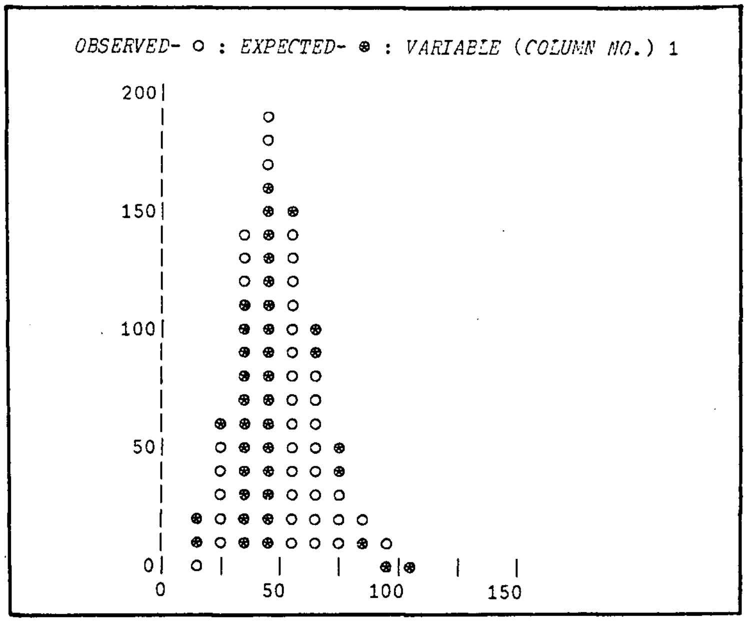
POISSON



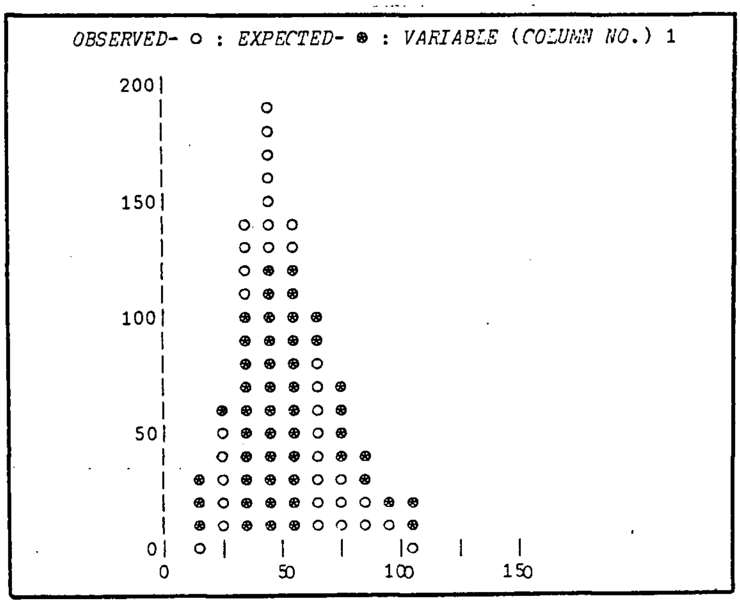
LOGNORMAL



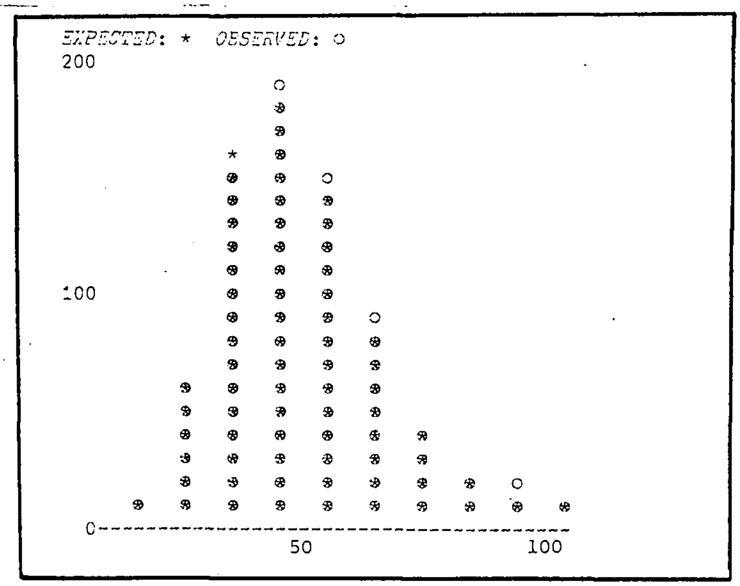
NORMAL



POISSON

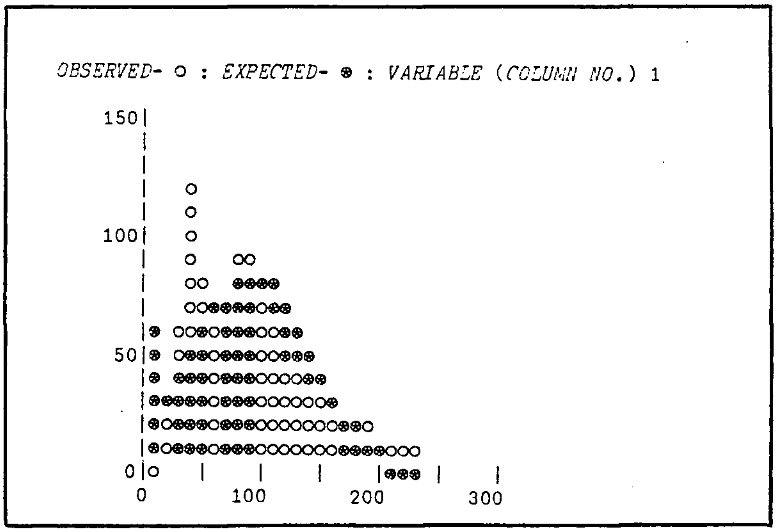


LOGNORMAL

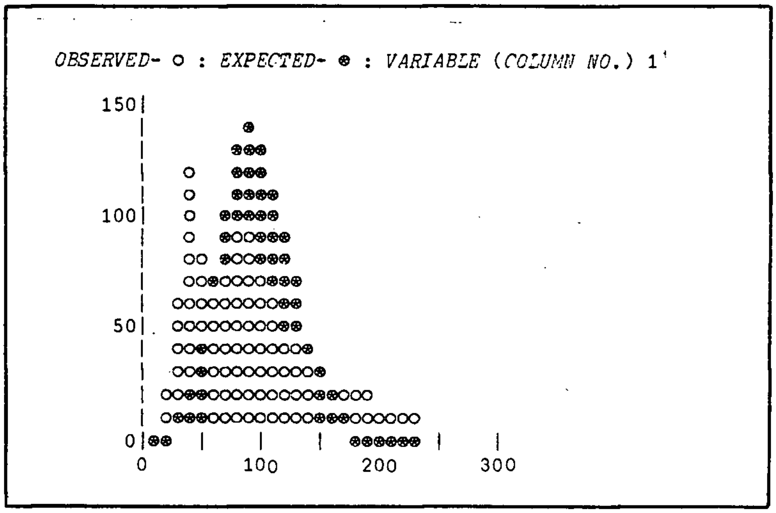


AIRPORT G, 1981

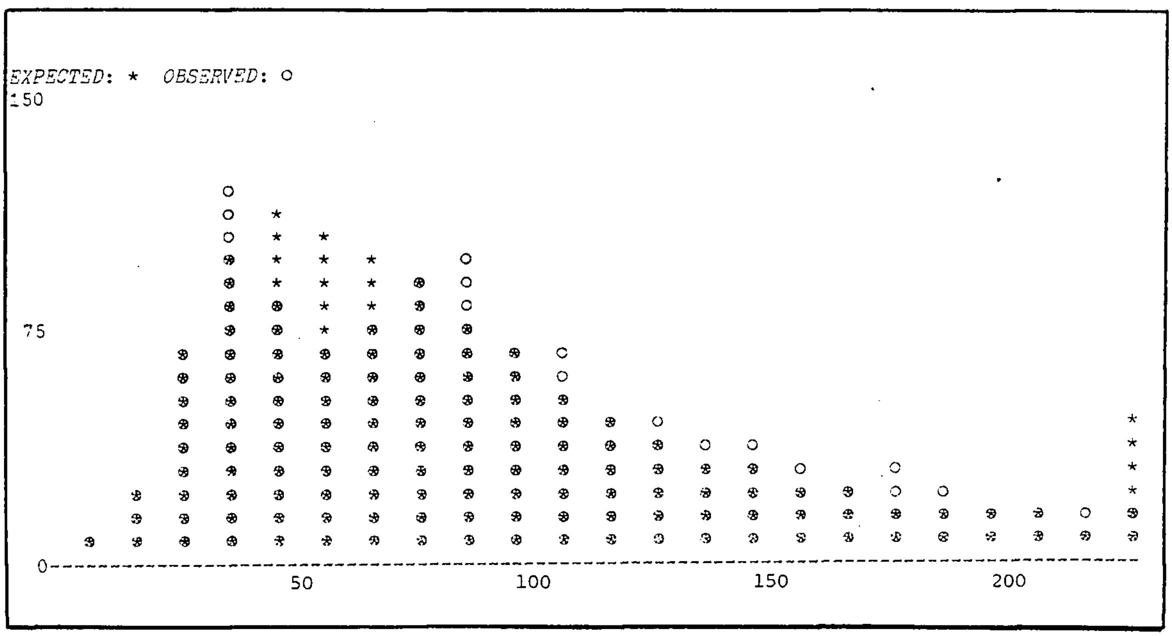
NORMAL



POISSON

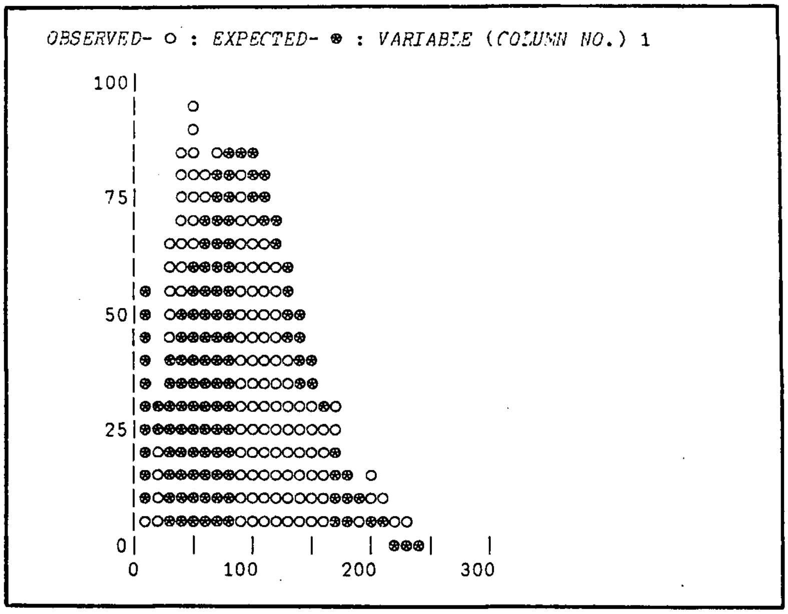


LOGNORMAL

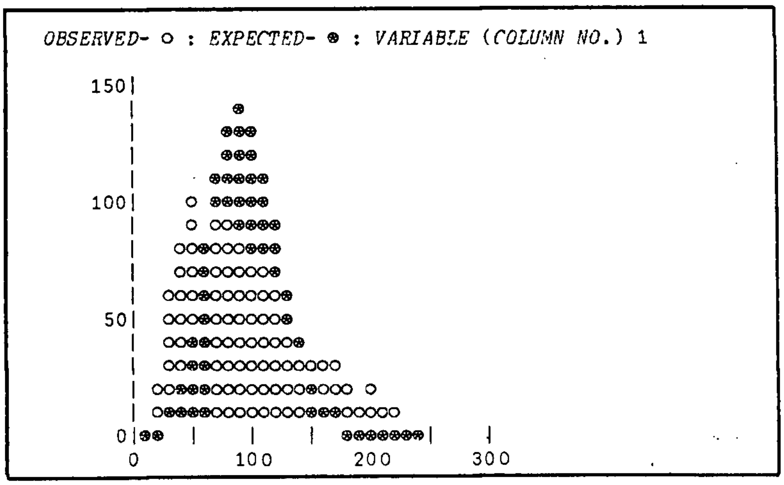


AIRPORT G, 1982

NORMAL



POISSON



LOGNORMAL