ON SEQUENTIAL BINOMIAL ESTIMATION

ON SEQUENTIAL BINOMIAL ESTIMATION

By

BOON-CHONG CHANG

A Thesis

submitted to the Faculty of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

February, 1968

MASTER OF SCIENCE (Mathematical Statistics)

McMASTER UNIVERSITY Hamilton, Ontario

TITLE: On Sequential Binomial Estimation

AUTHOR: Boon-Chong Chang

SUPERVISOR: Dr. I. Z. Chorneyko

NUMBER OF PAGES: v, 104

SCOPE AND CONTENTS:

This thesis is concerned with the properties of sequential binomial estimation. It illustrates the construction of optimal sequential binomial sampling plans for point estimation problems in which, according to custom, each loss function is taken to be a constant times the square of the error. The way such a constant affects the sizes of the constructed sampling plans is also within the scope of this thesis.

PREFACE

The first part of this thesis deals with the uniqueness of unbiased estimate of p, proportion of defectives in the (binomial) population; the efficient estimators and sampling plans; the characterizations of simple binomial sampling plans.

In the second part, we summarize some properties of the truncated-sequential game and illustrate the construction of optimal sequential binomial sampling plans for point estimation problems.

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to his supervisor Dr. I. Z. Chorneyko for his willing assistance and guidance. Appreciation is also expressed to McMaster University for financial support.

TABLE OF CONTENTS

Part I BINOMIAL SAMPLING PLANS

1.1 Introduction

By a sample of size N drawn from a binomial distribution we mean a vector $\Upsilon = (u_1, u_2, \ldots, u_N)$ in which the independent observations $u_1, u_2, \ldots, u_N$ are to be taken on the random variable $\mathcal{U}$ so distributed that

$$\text{Prob}(\mathcal{U} = 1/p) = p, \ \text{Prob}(\mathcal{U} = 0/p) = q = 1 - p$$

where $0 < p < 1$. By letting

$$X(\Upsilon) = \text{number of zeroes in } \Upsilon$$

$$Y(\Upsilon) = \text{number of ones in } \Upsilon,$$

samples of this type may be referred to points in the xy-plane with nonnegative integral coordinates. The index $N(\Upsilon) = X(\Upsilon) + Y(\Upsilon)$ of the point Υ will be the size of that sample Υ. We shall interpret a sampling plan as a rule that specifies at each stage of a sequential sampling process whether sampling is to cease or another observation is to be taken and our main attention will be devoted to the development of certain criteria for the selection of an appropriate sampling plan for the family of binomial distributions.

1.2 Definitions and fundamental facts

Consider the plane lattice $\mathcal{L}$ of points γ whose coordinates $X(\gamma)$ and $Y(\gamma)$ are nonnegative integers.

Definition 1.2.1. A sampling plan is a function S defined on $\mathcal{L}$ taking only the values 0 and 1, and such that

$$S(\Theta) = S((0,0)) = 1.$$

Definition 1.2.2. A path to γ is a sequence of points $\Theta = \gamma_0, \gamma_1, \ldots, \gamma_n = \gamma$ such that $S(\gamma_k) = 1$ for $k = 0, 1, \ldots, n - 1$ and either $X(\gamma_{k+1}) = X(\gamma_k) + 1$, $Y(\gamma_{k+1}) = Y(\gamma_k)$ or $X(\gamma_{k+1}) = X(\gamma_k)$, $Y(\gamma_{k+1}) = Y(\gamma_k) + 1$.

Under a given sampling plan S, points of $\mathcal{L}$ are decomposed into the following three mutually exclusive classes:

Definition 1.2.3. γ is a boundary point if there exists a path to γ and $S(\gamma) = 0$.

Definition 1.2.4. γ is a continuation point if there exists a path to γ and $S(\gamma) = 1$ so that at least one path exists "through" γ.

Definition 1.2.5. γ is an inaccessible point if no path exists to γ and $S(\gamma) = 1$.

It should be noted that the origin Θ is always a continuation point. It is seen that if $\gamma_k = (x_k, y_k)$ is a continuation point then $(x_k + 1, y_k)$ and $(x_k, y_k + 1)$ are accessible. The values of S at inaccessible points are irrelevant to the sampling process. However, it is useful to take $S(\gamma) = 1$ for inaccessible points γ to facilitate the phrasing of certain definitions.

Definition 1.2.6. The boundary B of S is the set of all boundary points of S.

Let C denote the set of all continuation points of S.

Paths may be regarded as arising by a random process such that a path reaching $\Upsilon_i = (x_i, y_i)$, a continuation point, will be extended to $(x_i, y_i + 1)$ with probability p or to $(x_i + 1, y_i)$ with probability $q = 1 - p$. When a path is extended to a boundary point, the process ceases.

A sampling plan is completely determined by its boundary, so that any reasonable estimator depends on the observed sample sequence only through the boundary point reached by the sequence.

Definition 1.2.7. A sampling plan is said to be bounded if there is an integer n such that all points Υ with

$$N(\Upsilon) = X(\Upsilon) + Y(\Upsilon) > n$$

are inaccessible points.

The smallest such n is called the size of the plan.

The probability of reaching a particular point Υ is

$$P(p, \Upsilon) = K(\Upsilon)p^{Y(\Upsilon)}q^{X(\Upsilon)}$$

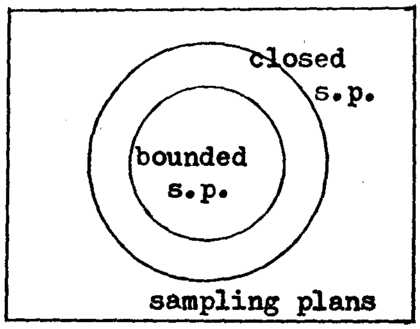
where $K(\Upsilon)$ is the number of distinct paths from the origin to Υ.

Definition 1.2.8. A sampling plan with boundary B is said to be closed if

$$\sum_{\Upsilon \in B} P(p, \Upsilon) = 1$$

for all p, $0 < p < 1$.

For a bounded sampling plan of size n, it is clear from the definitions that paths from the origin cannot include more than n + 1 points. This implies that a path from the origin strikes some boundary points with probability one. Thus, bounded sampling plans are closed. Their connection may be expressed in the following Venn diagram:



Wolfowitz (5) assured the closure of some infinite sampling plans by proving the assertion that a sampling plan is closed if

$$\lim_{n \to \infty} \inf \frac{A(n)}{\sqrt{n}} < \infty$$

where A(n) is the number of continuation points of index n.

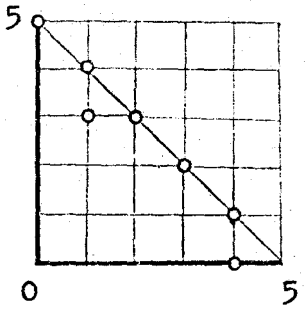
Definition 1.2.9. A boundary point γ_0 of a bounded sampling plan S is said to be essential if the sampling plan S_{γ_0} where

$$S_{\gamma_0}(\gamma) = S(\gamma) \text{ for } \gamma \neq \gamma_0$$

$$S_{\gamma_0}(\gamma_0) = 1,$$

with γ_0 a continuation point, is not bounded.

Example 1.2.1. Consider the following sampling plan S where dots denote boundary points:



The boundary point (3,2) is essential. Indeed, (3,2) is a continuation point of the sampling plan $S_{(3,2)}$ so that there is at least one path in $S_{(3,2)}$ from the origin to (3,2) which can be extended to (4,2) or (3,3). Similarly, (4,0), (4,1), (2,3), (1,4), and (0,5) are essential boundary points. On the other hand, (1,3) is a non-essential boundary point, for removal of the boundary point does not destroy the boundedness of the sampling plan.

Definition 1.2.10. A boundary B is said to be essential if all its points are essential.

Clearly, if (x_0, y_0) is an essential boundary point, then $(x_0 + 1, y_0)$ and $(x_0, y_0 + 1)$ cannot both be accessible.

Definition 1.2.11. An estimator f is a real-valued function defined on B. The only estimators to be considered are those for which

$$g(p) = E(f/p) = \sum_{\gamma \in B} f(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)}$$

is absolutely convergent.

Conventions

(i) For every sampling plan to be considered,

$$E(N^2/p) = \sum_{\gamma \in B} N^2(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)}$$

is uniformly convergent on every closed interval of values of p.

(ii) For every estimator f to be considered, E(f/p) is differentiable termwise in the open interval, $0 < p < 1$, and the derived series is absolutely convergent.

A well-known and useful sufficient condition for the termwise differentiability of the series E(f/p) is that the formal termwise derivative be absolutely uniformly convergent on every closed subinterval.

The functions defined on B and taking the values $X(\gamma)$, $Y(\gamma)$, and $N(\gamma)$ are denoted by X, Y, and N, respectively. Then we have the following:

Lemma 1.2.1. $E(Y^2/p)$, $E(X^2/p)$, and $E(XY/p)$ all exist and are at most $E(N^2/p)$.

Proof. Since $0 \leq X \leq N$ and $0 \leq Y \leq N$, the results follow from (i).

Lemma 1.2.2. N, X, and Y, considered as estimators, satisfy (ii).

Proof. $E(N/p) = \sum_{\gamma \in B} N(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)}$.

$$\left| \frac{d}{dp} \left(N(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)}\right) \right| = \left| N(\gamma)K(\gamma)p^{Y(\gamma)-1}q^{X(\gamma)-1}(qY(\gamma) - pX(\gamma)) \right|$$

$$= \frac{1}{pq} N(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)} \left| qY(\gamma) - pX(\gamma) \right|$$

$$< \frac{1}{pq} N^2(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)}$$

for $0 < p < 1$ and $\gamma \in B$. Then, by Weierstrass' M-test for uniform convergence and by (i), the series

$$\sum_{\gamma \in B} \left(\frac{d}{dp} N(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)}\right)$$

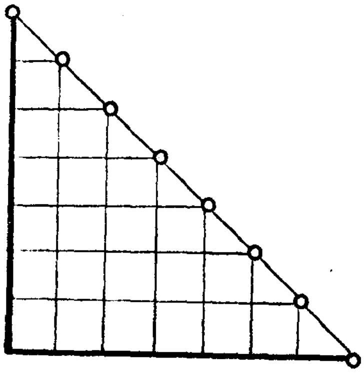
converges uniformly on every closed interval and, hence, according to the remark following (ii), E(N/p) is termwise differentiable. The proofs for X and Y are similar.

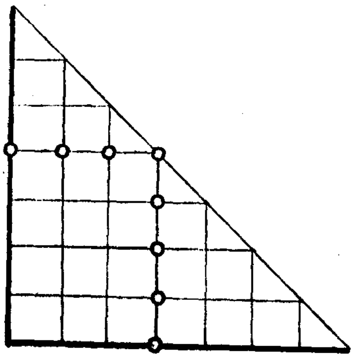
Definition 1.2.12. A sampling plan is said to be simple if every set

$$C_k = \left\{ (x,y):\ x + y = k \right\}$$

meets the class C of continuation points in an interval (possibly trivial or empty).

The sampling plan given in Example 1.2.1 is not simple, since the intersection of the sets C_4 and C is not an interval. However, the following sampling plans are simple.





1.3 The estimate $\hat{p}$

Girshick, Mosteller, and Savage (1) have shown how to construct an unbiased estimator of the parameter p, the fraction defective, from samples drawn from a binomial distribution. The estimator constructed is applicable to samples whose items are drawn and classified one at a time until the number of defectives and the number of nondefectives simultaneously agree with one of a set of preassigned number pairs. When this agreement takes place, the sampling process ceases and an unbiased estimate of the proportion p of defectives in the population may be made. Several results concerning this construction have also been discussed.

Construction: For a closed sampling plan S, let $\hat{p}(\Upsilon) = K^*(\Upsilon)/K(\Upsilon)$, where Υ is a boundary point and

$$K(\Upsilon) = \text{number of paths in S from } \Theta \text{ to } \Upsilon.$$

$$K^*(\Upsilon) = \text{number of paths in S from } (0,1) \text{ to } \Upsilon.$$

Theorem 1.3.1: $\hat{p}(\Upsilon)$ is an unbiased estimate of p.

Proof.

$$E(\hat{p}/p) = \sum_{\Upsilon \in B} \hat{p}(\Upsilon)P(p,\Upsilon)$$

$$= \sum_{\Upsilon \in B} K^*(\Upsilon)p^{Y(\Upsilon)}q^{X(\Upsilon)}$$

If $(0,1) \in B$, then $K^*((0,1)) = 1$ and $K^*(\Upsilon) = 0$ for $\Upsilon \neq (0,1)$, in which case

$$E(\hat{p}/p) = p.$$

If $(0,1) \notin B$, consider the plan S' obtained from S by taking the point (0,1) as a boundary point and $K'(\gamma)$, the number of paths in S' from the origin to the boundary points γ of S. We see that except for (0,1) every boundary point of S' is a boundary point of S and that $K'(\gamma) = 0$ except for the boundary points of S'. Now

$$K^*(\gamma) = K(\gamma) - K'(\gamma)$$

$$\therefore \quad E(\hat{p}/p) = \sum_{\gamma \in B} K(\gamma) p^{Y(\gamma)} q^{X(\gamma)} - \sum_{\gamma \in B} K'(\gamma) p^{Y(\gamma)} q^{X(\gamma)}$$

$$= 1 - \sum_{\gamma \in B} K'(\gamma) p^{Y(\gamma)} q^{X(\gamma)}.$$

But S' is closed since S contains S'. So

$$p + \sum_{\gamma \in B} K'(\gamma) p^{Y(\gamma)} q^{X(\gamma)} = 1$$

and the proof is complete.

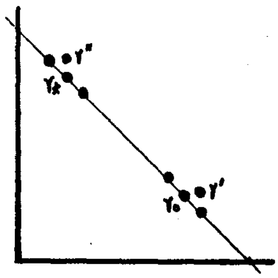
Theorem 1.3.2. A necessary condition that $\hat{p}(\gamma)$ be the unique unbiased estimate of p is that S be simple.

Proof. For a non-simple sampling plan we shall construct a function $m(\gamma)$, not identically zero, such that

$$E(m/p) = \sum_{\gamma \in B} m(\gamma) P(p,\gamma) = 0. \tag{1.3.1}$$

However, $\hat{p}(\gamma) + m(\gamma)$ will be an unbiased estimate of p different from $\hat{p}(\gamma)$.

Suppose we have a closed sampling plan S which is not simple. We consider the lowest index n where the continuation points are separated. There will be at least one uninterrupted sequence of points between some pair of continuation points that are not continuation points. By assumption of n, the points immediately below the points in this sequence are continuation points and hence all the points of this sequence are boundary points. Let this sequence be the points $\gamma_i = (x_0 - i, y_0 + i)$, $i = 0,1,\ldots, t$, $x_0 + y_0 = n$.



For boundary points in this sequence, let

$$m(\gamma_j) = \frac{(-1)^j}{K(\gamma_j)} \qquad 0 \leq j \leq t.$$

While, for those not in this sequence we proceed as follows:

Take

$$\gamma'' = (x_o - t, y_0 + t + 1) \text{ and } \gamma' = (x_0 + 1, y_0)$$

which are accessible.

Let

$$\ell''(\gamma) = \text{number of paths from } \gamma'' \text{ to } \gamma \in B$$

$$\ell'(\gamma) = \text{number of paths from } \gamma' \text{ to } \gamma \in B,$$

with

$$\ell''(\gamma'') = 1 = \ell'(\gamma')$$

if γ'' and γ' are boundary points.

To complete the construction, let

$$m(\Upsilon) = \frac{-[\ell'(\Upsilon) + (-1)^t \ell''(\Upsilon)]}{K(\Upsilon)}$$

for all boundary points Υ not in the mentioned sequence.

Note that $\sum_{\Upsilon \in B} \ell'(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)} = p^{y_0} q^{x_0+1}$

and $\sum_{\Upsilon \in B} \ell''(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)} = p^{y_0+t+1} q^{x_0-t}$.

By symmetry, it is enough to show the first equality. Indeed, if $\Upsilon' \in B$, $\ell'(\Upsilon') = 1$ and $\ell'(\Upsilon) = 0$ for $\Upsilon \neq \Upsilon'$, and the sum is the single term $p^{y_0} q^{x_0+1}$. If $\Upsilon' \notin B$, consider the sampling plan S' obtained from S by taking Υ' as a boundary point and $K'(\Upsilon)$, the number of paths in S' from the origin to the boundary points Υ of S. Every boundary point of S' except Υ' is a boundary point of S and $K'(\Upsilon) = 0$ for Υ not in the boundary of S'. Then it is easy to see that

$$K(\Upsilon) = K'(\Upsilon') \ell'(\Upsilon) + K'(\Upsilon),$$

and

$$1 = \sum_{\Upsilon \in B} K(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)}$$

$$= K'(\Upsilon') \sum_{\Upsilon \in B} \ell'(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)} + \sum_{\Upsilon \in B} K'(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)}.$$

Thus

$$K'(\Upsilon') \sum_{\Upsilon \in B} \ell'(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)} = 1 - \sum_{\Upsilon \in B} K'(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)}$$

$$= K'(\Upsilon') p^{y_0} q^{x_0+1}$$

Hence
$$\sum_{\Upsilon \in B} \ell'(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)} = p^{y_0} q^{x_0 + 1}.$$

We now check that $m(\Upsilon)$ satisfies equation (1.3.1):

$$\sum_{\Upsilon \in B} m(\Upsilon) K(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)} = \sum_{j=0}^{t} (-1)^j p^{y_0+j} q^{x_0-j} - \sum_{\Upsilon \in B} \ell'(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)} - \sum_{\Upsilon \in B} (-1)^t \ell''(\Upsilon) p^{Y(\Upsilon)} q^{X(\Upsilon)}$$

$$= \sum_{j=0}^{t} (-1)^j p^{y_0+j} q^{x_0-j} - p^{y_0} q^{x_0+1} - (-1)^t p^{y_0+t+1} q^{x_0-t}$$

$$= p^{y_0} q^{x_0-t} \left[\sum_{j=0}^{t} (-1)^j p^j q^{t-j} - q^{t+1} - (-1)^t p^{t+1} \right]$$

$$= 0.$$

Savage [6] found that simplicity is also a sufficient condition that ensures $\hat{p}$ to be the unique unbiased estimate of p for a closed sampling plan S.

<u>Theorem 1.3.3</u>. If S is simple there is at most one bounded unbiased estimate of any given function of p.

Proof. If the lemma were false, the difference of two unbiased estimates would yield a non-trivial bounded unbiased estimate of zero, i.e. $m(\Upsilon)$ such that $|m(\Upsilon)|$ is bounded by a constant m^*, $m(\Upsilon)$ not identically zero and

$$E(m/p) = \sum_{\Upsilon \in B} m(\Upsilon)K(\Upsilon)p^{Y(\Upsilon)}q^{X(\Upsilon)} = 0. \tag{1.3.2}$$

But this will be shown to be impossible. If $m(\Upsilon)$ were not identically zero, there would be an $\Upsilon_0 = (x_0, y_0) \in B$ such that $m(\Upsilon_0) \neq 0$ and

(a) $m(\Upsilon) = 0$ for all boundary points Υ of index less than that of Υ_0.

and (b) one of the coordinates of Υ_0 is less than the corresponding coordinate of any other boundary point Υ for which $m(\Upsilon) \neq 0$.

This follows from the simplicity requirement which implies that the boundary points of index $n = x_0 + y_0$ are broken into two sets,

(c) those whose y coordinates are less than the y coordinates of the continuation points of index n.

and (d) those whose x coordinates are less than the x coordinates of the continuation points of index n.

Since the situations (c) and (d) are symmetrical, we may suppose that Υ_0 is a boundary point such that $m(\Upsilon_0) \neq 0$, Υ_0 is below all continuation points of its own index and also below every other Υ for which $m(\Upsilon) \neq 0$. Equation (1.3.2) may be written

$$m(\Upsilon_0)K(\Upsilon_0)p^{y_0}q^{x_0} + \sum_{\substack{\Upsilon \in B \\ \Upsilon \neq \Upsilon_0}} m(\Upsilon)K(\Upsilon)p^{Y(\Upsilon)}q^{X(\Upsilon)} = 0.$$

Thus

$$|m(\Upsilon_0)|\,K(\Upsilon_0)p^{y_0}q^{x_0} = \left| \sum_{\substack{\Upsilon \in B \\ Y(\Upsilon) > y_0}} m(\Upsilon)K(\Upsilon)p^{Y(\Upsilon)}q^{X(\Upsilon)} \right| \tag{1.3.3}$$

$$\leq m^* \sum_{\substack{\Upsilon \in B \\ Y(\Upsilon) > y_0}} K(\Upsilon)p^{Y(\Upsilon)}q^{X(\Upsilon)}.$$

Let M denote the set of all accessible points at which $X(\gamma) < x_0$ and $Y(\gamma) = y_0 + 1$. There are at most x_0 points in M, say $\beta_1, \beta_2, \dots, \beta_n$. Considering the way in which γ_0 has been chosen, every path from the origin to a γ for which $Y(\gamma) > y_0$ passes through or to at least one point of M. Therefore when $Y(\gamma) > y_0$

$$P(p,\gamma) = K(\gamma)p^{Y(\gamma)}q^{X(\gamma)}$$

$$(1.3.4) \qquad \begin{aligned} &= P(p,\gamma/M)P(M) \\ &\leq P(p,\gamma/M)\sum_{1}^{n} K(\beta_j)p^{y_0+1}q^{X(\beta_j)} \\ &\leq p^{y_0+1}\sum_{1}^{n} K(\beta_j)P(p,\gamma/M). \end{aligned}$$

From inequalities (1.3.3) and (1.3.4),

$$(1.3.5) \qquad \begin{aligned} |m(\gamma_0)|\, K(\gamma_0)p^{y_0}q^{x_0} &\leq m^* p^{y_0+1}\left[\sum_{1}^{n} K(\beta_j)\right] \sum_{\substack{\gamma\in B \\ Y(\gamma)>y_0}} P(p,\gamma/M) \\ &\leq m^* p^{y_0+1}\sum_{1}^{n} K(\beta_j). \end{aligned}$$

But this is impossible that (1.3.5) should be satisfied for small p.

Combining Theorems 1.3.1, 1.3.2, and 1.3.3 we have the

Theorem 1.3.4. A necessary and sufficient condition that $\hat{p}(\gamma)$ be the unique (bounded) unbiased estimate of p for a closed sampling plan S is that S be simple.

1.4 Efficient Estimators and Sampling Plans

The following inequality serves as a basic tool of determining the efficiency of an estimator. It provides a lower bound for the variance of an estimator in terms of its expected value and the average sample size of the sampling plan. If, at p_0, this lower bound is attained for a particular estimator and sampling plan, we say that they are **efficient** of p_0. An estimator is efficient at p_0 if it is unbiased and if it possesses minimum variance among all unbiased estimators at p_0.

Lemma 1.4.1. For any estimator f,

$$\mathrm{Var}(f/p) \geq \frac{pq\left(g'(p)\right)^2}{E(N/P)} \quad , \; g(p) = E(f/p). \tag{1.4.1}$$

Equality holds at a particular value of p, say p_0, if and only if there exist constants a and b such that

$$f(\gamma) = a\left[q_0 Y(\gamma) - p_0 X(\gamma)\right] + b$$

for all boundary points γ, where $q_0 = 1 - p_0$.

A brief history of this inequality with references is given by Savage in [8] , Pg. 238. It was first proved for sequential plans by Wolfowitz [7]. Following Savage, (1.4.1) will be called the **information inequality**.

De Groot (2) has shown that the only efficient sampling plans are the single sample plans and the inverse binomial sampling plans. In a single sample plan, $B = \{\gamma: N(\gamma) = n\}$, n being a positive integer, and any non-constant function of the form $a + bY$ is an efficient estimator of $a + b\,np$, and these are the only efficient estimators. In an inverse binomial sampling plan, either $B = \{\gamma: Y(\gamma) = c\}$ or $B = \{\gamma: X(\gamma) = c\}$, c being a positive integer. When $B = \{\gamma: Y(\gamma) = c\}$, any non-constant function of the form $a + bN$ is an efficient estimator of $a + bc(1/p)$, and these are the only efficient estimators. When $B = \{\gamma: X(\gamma) = c\}$, any non-constant function of the form $a + bN$ is an efficient estimator of $a + bc(1/q)$, and these are the only efficient estimators.

The name "inverse binomial sampling" was suggested by Tweedie in (9) and this type of plan was first treated formally by Haldane in (10) and (11). We see that these plans are closed.

For the single sample plan with boundary $B = \{\gamma: N(\gamma) = n\}$, Y/n is an efficient estimator of p with $\mathrm{var}\left(\frac{Y}{n}/p\right) = pq/n$. For any estimator f

$$E(f/p) = \sum_{\gamma \in B} f(\gamma) \binom{n}{Y(\gamma)} p^{Y(\gamma)} q^{X(\gamma)},$$

which is a polynomial in p of degree at most n, say

$$E(f/p) = a_0 + a_1 p + a_2 p^2 + \dots + a_\gamma p^\gamma + \dots + a_n p^n.$$

Clearly, polynomials in p of degree at most n are the only functions which may be estimated unbiasedly. In addition,

$$E\left[\frac{Y(Y-1)\ \dots\ (Y-\gamma+1)}{n(n-1)\ \dots\ (n-\gamma+1)} \Big/ p\right] = p^{\gamma},\ \gamma = 1, 2, \dots, n.$$

Thus every such polynomial is estimable unbiasedly.

The analogous properties of an inverse binomial sampling plan are less familiar. De Groot (2) proved the following theorem for inverse binomial sampling plans which provides a rule for finding unbiased estimator of a given function h(q). For convenience, the functions are written as functions of q rather than of p.

Consider now the plan with boundary $B = \{\gamma: Y(\gamma) = c\}$. For each non-negative integer k, there exists a unique point γ_k of B such that $N(\gamma_k) = c + k$. Then we have:

<u>Theorem 1.4.1.</u> A function h(q) is estimable <u>unbiasedly</u> if and only if it can be expressed in a Taylor's series in the interval $|q| < 1$. If h(q) is estimable unbiasedly, then its unique estimator is given by

$$f(\gamma_k) = \frac{(c-1)!}{(k+c-1)!} \frac{d^k}{dq^k} \left[\frac{h(q)}{(1-q)^c}\right]_{q=0} \qquad k = 0, 1, 2, \dots .$$

Proof. We note that if h(q) can be expanded in Taylor's series in the given interval, then so also can $\frac{h(q)}{(1-q)^c}$, and conversely.

$$\text{Let } \frac{h(q)}{(1-q)^c} = \sum_{k=0}^{\infty} b_k q^k \text{ where } b_k = \frac{d^k}{dq^k} \left[\frac{h(q)}{(1-q)^c}\right]_{q=0} \Big/ k!$$

Then $h(q) = p^c \sum_{k=0}^{\infty} b_k q^k$.

Taking $$f(\gamma_k) = b_k \Big/ \binom{k+c-1}{k}$$

we have $$E(f/p) = \sum_{k=0}^{\infty} f(\gamma_k) \binom{k+c-1}{k} p^c q^k$$

$$= p^c \sum_{k=0}^{\infty} b_k q^k$$

$$= h(q).$$

Conversely, suppose that h(q) is estimable unbiasedly. Then there exists an estimator f such that

$$h(q) = E(f/p)$$

$$= p^c \sum_{k=0}^{\infty} f(\gamma_k) \binom{k+c-1}{k} q^k$$

or $$\frac{h(q)}{(1-q)^c} = \sum_{k=0}^{\infty} f(\gamma_k) \binom{k+c-1}{k} q^k.$$

Then $$\binom{k+c-1}{k} f(\gamma_k) = \frac{1}{k!} \frac{d^k}{dq^k} \left[\frac{h(q)}{(1-q)^c} \right]_{q=0}$$

or $$f(\gamma_k) = \frac{(c-1)!}{(k+c-1)!} \frac{d^k}{dq^k} \left[\frac{h(q)}{(1-q)^c} \right]_{q=0}.$$

Thus $$\frac{h(q)}{(1-q)^c} = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{d^k}{dq^k} \left[\frac{h(q)}{(1-q)^c} \right]_{q=0} \cdot q^k$$

which is the required expansion. The uniqueness of f follows from the uniqueness of the Taylor's series.

We have just discussed unbiased estimators for the efficient sampling plans. The rest of the present section will be devoted to the relationships between sampling plans and efficiently estimable functions. We have stressed before that, for a given sampling plan, the only estimators that are efficient at a given value p_0 are the non-constant functions f* of the form

$$f^*(\gamma) = a\left[q_0 Y(\gamma) - p_0 X(\gamma)\right] + b,$$

for some constants a and b and for all boundary points γ. The next theorem, developed by De Groot [2] determines the class of functions that are estimable efficiently at a given point, simply by evaluating $E(f^*/p)$.

Theorem 1.4.2. For a given sampling plan, a non-constant function $g(p)$ is estimable efficiently at p_0 if and only if there exists a constant $k \neq 0$, such that

$$(1.4.2) \qquad E(N/p) = k\left[g(p) - g(p_0)\right]/(p - p_0) \text{ for } p \neq p_0$$

$$E(N/p_0) = k\, g'(p_0).$$

Proof. Suppose that $g(p)$ is estimable efficiently at p_0, then its estimator is of the form

$$f^*(\gamma) = a\left[q_0 Y(\gamma) - p_0\, X(\gamma)\right] + b$$

for some constants a and b, $a \neq 0$, and all boundary points γ.

$$
\begin{aligned}
g(p) &= E(f^*/p) \\
&= a\, E(q_0 Y - p_0 X/p) + b \\
&= a\, E(Y - p_0\, Y - p_0 X/p) + b \\
&= a\, E(qY + pY - p_0 N/p) + b \\
&= a\, E(qY - pX + pN - p_0 N/p) + b \\
&= a\, E(qY - pX/p) + a(p - p_0)E(N/p) + b.
\end{aligned}
$$

By symmetry, $qE(Y/p) = pE(X/p)$, so that

$$g(p) = a(p - p_0)E(N/p) + b \tag{1.4.3}$$

and $g(p_0) = b$.

Thus
$$
\begin{aligned}
E(N/p) &= \left[g(p) - g(p_0)\right] / a(p - p_0) \\
&= k\left[g(p) - g(p_0)\right] / (p - p_0),\ p \neq p_0,\ k = \frac{1}{a}\,.
\end{aligned}
$$

Differentiating both sides of (1.4.3) at p_0, gives

$$g'(p_0) = a\, E(N/p_0)$$

and the proof of necessary part is complete.

The reverse steps prove the sufficient part.

From theorem 1.4.2 we see that there does not always exist a sampling plan that admits estimation of a given function efficiently at a given value of p. In order to estimate a function g(p) efficiently at p_0, a sampling plan must be selected such that (1.4.2) hold. Then the efficient estimator will be of the form

(1.4.4) $$f(\gamma) = a\left[q_0 Y(\gamma) - p_0\, X(\gamma)\right] + g(p_0),\ a = \frac{1}{k}.$$

For a given g(p), p_0, and k, there does exist more than one sampling plan satisfying (1.4.2). Let R denote the class of such plans. Since every plan of R yields the same E(N/p), and since, for every plan of R, the estimator f given by (1.4.4) is efficient at p_0, it follows from the information inequality that Var(f/p_0) is the same under each plan of R. In general, however, for values of p other then p_0, Var(f/p) will be different under the various plans of R. We are now interested in determining the plan of R for which Var(f/p) is smallest at some values of p other than p_0. De Groot [2] claimed in the following theorem that this is equivalent to determining the plan for which Var(N/p) is minimized at the relevant values of p.

<u>Theorem 1.4.3</u>. Let $f = a(q_0Y - p_0X) + b$.

Let p* be a value of p other than p_0.

Let S_1 and S_2 be two sampling plans such that

$$E(N/p,\ S_1) = E(N/p,\ S_2) \text{ for all p.}$$

Then $\mathrm{Var}(f/p^*,\ S_1) \lesseqgtr \mathrm{Var}(f/p^*,\ S_2)$ if and only if

$$\mathrm{Var}(N/p^*,\ S_1) \lesseqgtr \mathrm{Var}(N/p^*,\ S_2).$$

Proof. $f = a(q_0Y - p_0X) + b = a(qY - pX) + a(p - p_0)N + b$

$$f^2 = a^2(qY - pX)^2 + a^2(p - p_0)^2N^2 + b^2 + 2a^2(p - p_0)(qY - pX)N + 2ab(qY - pX) + 2ab(p - p_0)N.$$

We note that (a) $E\left[(qY - pX)f/p\right] = pq\ g'(p)$

(b) $E\left[(qY - pX)^2/p\right] = pq\ E(N/p)$.

Indeed,

$$pq\ g'(p) = pq \frac{d}{dp}\left(\sum_{\gamma \in B} f(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)} \right)$$

$$= pq \sum_{\gamma \in B} f(\gamma)K(\gamma)\left[\frac{d}{dp}\left(p^{Y(\gamma)}q^{X(\gamma)}\right)\right]$$

$$= pq \sum_{\gamma \in B} f(\gamma)K(\gamma)\left[Y(\gamma)p^{Y(\gamma)-1}q^{X(\gamma)} - p^{Y(\gamma)}X(\gamma)q^{X(\gamma)-1}\right], \quad q = 1-p$$

$$= \sum_{\gamma \in B} qY(\gamma)f(\gamma)K(\gamma)p^{Y(\gamma)}q^{X(\gamma)} - \sum_{\gamma \in B} pX(\gamma)f(\gamma)K(\gamma)\cdot p^{Y(\gamma)}q^{X(\gamma)}$$

$$= E\left[(qY - pX)f/p\right]$$

and

$$E\left[(qY - pX)^2/p\right] = pq\ E'(qY - pX/p), \text{ by (a)}$$

$$= pq \frac{d}{dp} E(qY - pX/p).$$

Since, by symmetry, $qE(Y/p) = p\ E(X/p)$

$$E\left[(qY - pX)^2/p\right] = pq \frac{d}{dp} E(pX - qY/p)$$

$$= pq\ E\left(\frac{d}{dp}(pX - qY/p)\right)$$

$$= pq\ E\left(\frac{d}{dp}(pX - Y + pY/p)\right)$$

$$= pq\ E(X + Y/p)$$

$$= pq\ E(N/p).$$

Now, in virtual of the above two equalities,

$$E(f/p) = a(p - p_0)E(N/p) + b$$

and

$$E(f^2/p) = a^2pq\ E(N/p) + a^2(p - p_0)^2E(N^2/p) + b^2 + 2a^2(p - p_0)pq\ E'(N/p)$$
$$+ 2ab(p - p_0)E(N/p).$$

Thus

$$Var(f/p) = a^2pq\ E(N/p) + a^2(p - p_0)^2Var(N/p) + 2a^2(p - p_0)pq\ E'(N/p) +$$
$$2ab(p - p_0)E(N/p).$$

This expression completes the proof.

1.5 Simple Sampling Plans

In this section, a new characterization of simplicity is given for bounded sampling plans and it is shown that the dimension of the linear space of unbiased estimators of 0 can be determined simply by counting the number of boundary points. We further determine the number of simple sampling plans of size n. It will be helpful to keep in mind that the boundedness of a sampling plan implies its closure. The following theorem plays an important role in determining a sampling plan that is simple.

Theorem 1.5.1. A sampling plan of size n is said to be simple if and only if it contains exactly n + 1 boundary points.

This theorem is implied by results of De Groot [2] and Girshick, Mosteller, and Savage [1] . However, the proof given by Brainerd and Narayana [3] gives insight into the structure of closed bounded sampling plans. Theorem 1.5.1 will be proved by a series of lemmas.

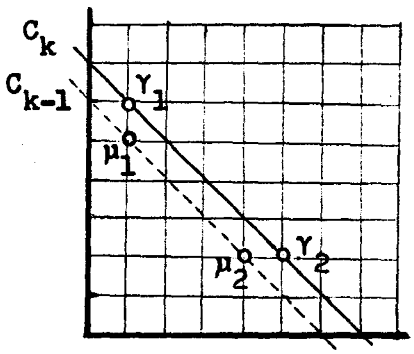
Lemma 1.5.1. The boundary of a sampling plan of size n contains at least n + 1 points.

Proof. If n = 1, then (0,1) and (1,0) must both be boundary points. Suppose the lemma is true for n = m and consider a plan S of size m + 1 with boundary B. Since $m + 1 \geq 2$, the points (0,1) and (1,0) cannot both be boundary points. We may assume that (1,0) is not a boundary point. By shifting the origin to the point (1,0), we obtain a plan S^* of size m and hence its boundary B^* involves at least m + 1 points. But S is bounded, so there must also exist a boundary point of the form (0,y). Hence B contains at least m + 2 points.

Induction completes the proof.

Lemma 1.5.2. If S is a bounded simple sampling plan, then for each k and for each pair of boundary points γ_1 and γ_2 on $C_k = \{(x,y): x + y = k\}$, there are no inaccessible points between γ_1 and γ_2 on C_k.

Proof. Let S be simple, bounded, and assume that γ_1 is above γ_2. Figure 1.5.1 illustrates the construction used in the proof.



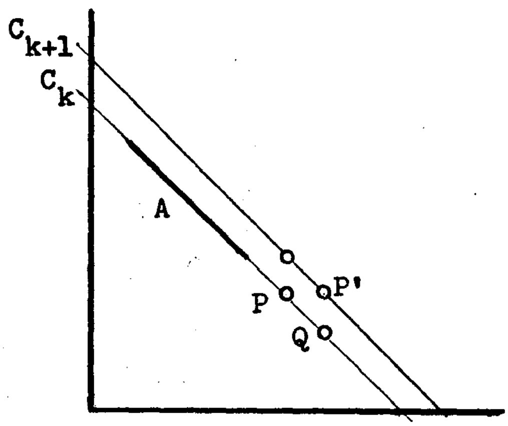
(figure 1.5.1)

Now $C_{k-1} \cap C$ where C is the set of all continuation points of S is a non-void interval. In fact, $\gamma_i = (x_i, y_i)$, i = 1, 2, can be reached by a path, implying that either $(x_i - 1, y_i)$ or $(x_i, y_i - 1)$ is a continuation point. However, the simplicity condition demands that all points on C_{k-1} between any two continuation points on C_{k-1} are continuation points. We may let $\mu_1 = (x_1, y_1-1) \in C$ and the $\mu_2 = (x_2-1, y_2) \in C$, and hence all points on C_{k-1} between μ_1 and μ_2 are in C. Therefore, no points on C_k between γ_1 and γ_2 can be an inaccessible point.

Corollary. Either of the boundary points γ_1, γ_2 in lemma 1.5.2 could be a continuation point and the lemma is still valid.

<u>Lemma 1.5.3</u>. If S is a bounded simple sampling plan, then its boundary is essential. In other words, if S contains a non-essential boundary point, then S cannot be simple.

Proof. Let S be simple, bounded, and contain a non-essential boundary point. Let k be the smallest integer such that C_k contains a non-essential boundary point γ_o. Clearly $k > 1$. Let $A = C_k \cap C$. C_k contains at least one continuation point. In fact, if all points on C_k are boundary points or inaccessible points, then all points of index > k are inaccessible which in turn implies that γ_o is essential. Thus A is a non-void interval. Lemma 1.5.2 assures that the point contiguous to at least one end of A must be a boundary point. Call this point P = (x,y) and let P be a lower boundary point. (see figure 1.5.2)



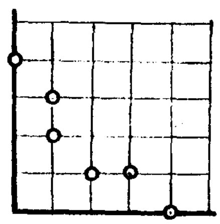
(fig. 1.5.2)

Suppose P is essential. Because A is non-void, the point (x, y + 1) is accessible. Then the essentiality of P demands that the point P' = (x + 1, y) must be inaccessible and hence all points below P' on C_{k+1} are inaccessible points. It follows that Q = (x + 1, y - 1) is either an essential boundary point or an inaccessible point. If Q is an inaccessible point, then all points on C_k below Q are inaccessible. If Q is an essential boundary point, repeat the argument until an inaccessible point is reached. The modifications required when p is an essential upper boundary point are obvious. Hence if P is an essential upper (lower) boundary point, C_k cannot contain a non-essential upper (lower) boundary point. We are now left to consider the case where the non-essential boundary point Υ_0 is a lower (upper) boundary point contiguous with A. We may assume Υ_0 to be a lower boundary point. Then all points on C_k below Υ_0 are either boundary points or inaccessible points. Thus all points $(x_0 + \ell, y_0), \ell = 1, 2, \ldots,$ are inaccessible points. In this situation, Υ_0 becomes essential. This is a contradiction.

Remark: The converse of lemma 1.5.3 is not true.

Consider the following sampling plan with boundary

$B = \{(0,4), (1,3), (1,2), (2,1), (3,1), (4,0)\}$.



B is essential but the plan is not simple.

Let S be a bounded sampling plan with essential boundary B. Let $\Upsilon_0 = (x_0, y_0)$ be a point of B. We note that $(x_0 + 1, y_0)$ and $(x_0, y_0 + 1)$ cannot both be accessible points. Thus three cases arise:

(a) $(x_0 + 1, y_0)$ and $(x_0, y_0 + 1)$ both inaccessible.

(b) One of these points inaccessible and the other a boundary point.

(c) One of these points inaccessible and the other a continuation point.

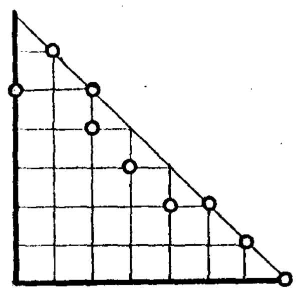
Definition 1.5.1. S' is said to be a deformation of S at Υ_0 if $S'(\Upsilon) = S(\Upsilon)$ for all $\Upsilon \neq \Upsilon_0$, $(x_0 + 1, y_0), (x_0, y_0 + 1)$ and in case (a), either $S'(x_0 + 1, y_0) = 0$ or $S'(x_0, y_0 + 1) = 0$ but not both,

in case (b), if γ_1 is the inaccessible point and γ_2 the boundary point, then $S'(\gamma_1) = 0$ and $S'(\gamma_2) = 0$,

in case (c), if γ_1 is the inaccessible point and γ_2 the continuation point, then $S'(\gamma_1) = 0$ and $S'(\gamma_2) = 1$,

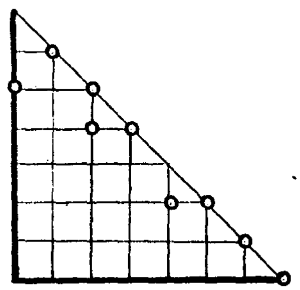
and in all cases $S'(\gamma_0) = 1$.

It is obvious that a deformation of S at a particular boundary point is nothing but simply "shift" the boundary point to an inaccessible point immediately beyond this boundary point. As an example, consider the following sampling plan S.



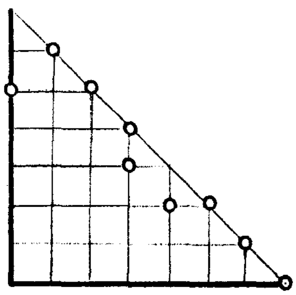
A deformation of S at (3,3) is

S'_a



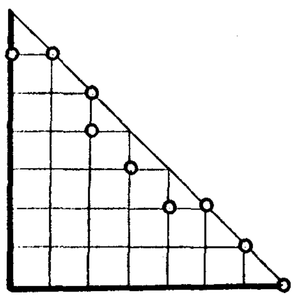
The deformation of S at (2,4) is

S_b'



The deformation of S at (0,5) is

S_c'



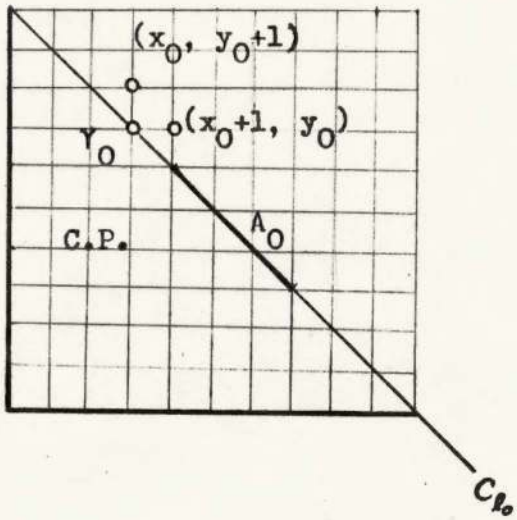
<u>Definition 1.5.2.</u> S' is said to be an <u>admissible</u> deformation of S if S' is bounded and its boundary is essential.

In the above example, S_b' and S_c' are admissible deformations. However, S_a' is an inadmissible deformation.

<u>Lemma 1.5.4.</u> If S is a simple sampling plan of size n, then there exists a sequence of sampling plans $S = S_0, S_1, \ldots, S_k$ such that S_{i+1} is an admissible deformation of S_i, $i = 0, 1, \ldots, k-1$, S_k has exactly the points of index n as boundary points and each S_i has n + 1 boundary points.

Proof. Let B_i be the boundary of S_i. Let ℓ_0 be the smallest integer such that $C_{\ell_0} \cap B_0 \neq \emptyset$. (We assume $\ell_0 \neq n$). Observe that

B_0 is essential and the continuation points on C_{ℓ_0} form an interval $A_0 \neq \emptyset$. If there is a contiguous boundary point $\gamma_0 = (x_0, y_0)$ on C_{ℓ_0} above A_0, then $(x_0 + 1, y_0)$ is an accessible point and the point $(x_0, y_0 + 1)$ is inaccessible. (Refer figure 1.5.3). Hence the



(Fig. 1.5.3)

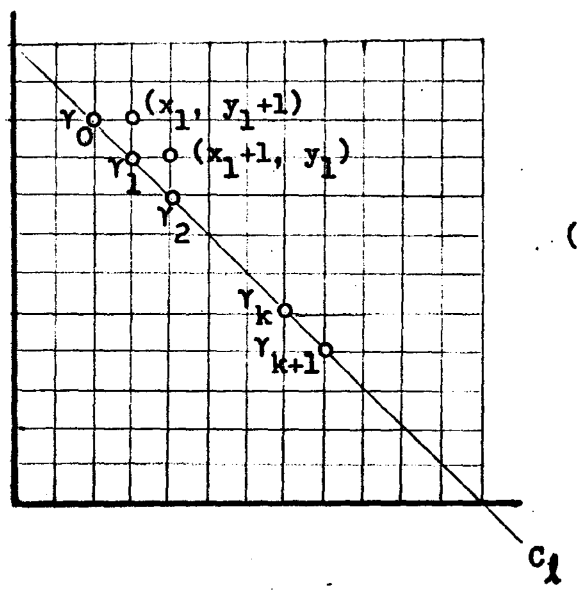
deformation S_1 of S_0 at γ_0 has $S_1(\gamma) = S(\gamma)$ for all $\gamma \neq \gamma_0$, $(x_0, y_0 + 1)$ and $S_1(\gamma_0) = 1$, $S_1(x_0, y_0 + 1) = 0$. Since there are no boundary points on C_ℓ ($\ell < \ell_0$), there are no inaccessible points on C_{ℓ_0} above γ_0. Thus above γ_0 on C_{ℓ_0} there are only boundary points. Since the only change in S_1 from S is to shift the boundary point γ_0 to the inaccessible point $(x_0, y_0 + 1)$ making γ_0 a continuation point, the admissibility of S_1 is obvious. Repeat shifting the boundary points one by one, whether from above or below A_0 on C_{ℓ_0}, to the line $C_{\ell_0 + 1}$. If C_{ℓ_0+1} has no continuation points, then we are finished. Otherwise continue the process. Induction guarantees that we finally reach the region S_k where boundary points are exactly the points of index n. Clearly, S_k and S_0 (in fact, every S_i) contain the same number of boundary points, namely n + 1.

So far, we have proved the necessary part of Theorem 1.5.1. To the sufficient part, we proceed as follows:

Lemma 1.5.5. If S is a non-simple sampling plan of size n, then S contains more than n + 1 boundary points.

Proof. If S is non-simple and contains a non-essential boundary point γ_0, then, by definition, the sampling plan S_{γ_0} obtained from S by taking γ_0 as a continuation point is of size n and hence, by lemma 1.5.1, the boundary of S_{γ_0} contains at least n + 1 points. Therefore the boundary of S contains more than n + 1 points.

Now we restrict ourselves to consider the case where S has essential boundary. Let ℓ be the smallest integer such that C_ℓ intersect C, the set of all continuation points, in a configuration which is not an interval. The following configuration occurs in C_ℓ , (Refer figure 1.5.4)



(figure 1.5.4)

where $\Upsilon_0 = (x_0, y_0)$ is a continuation point,

$\Upsilon_i = (x_0 + i, y_0 - i) = (x_i, y_i)$, $i = 1, 2, \ldots, k$, are boundary points, and

$\Upsilon_{k+1} = (x_0 + k + 1, y_0 - k - 1)$ is a continuation point.

The point $(x_1, y_1 + 1)$ is accessible. If $k = 1$, $(x_1 + 1, y_1)$ is also an accessible point implying that Υ_1 is non-essential. Thus the essentiality of boundary demands that $k \geqslant 2$ and hence $(x_1 + 1, y_1)$ is an inaccessible point. A deformation S_1 of S at Υ_1 is of the form

$$S_1(\Upsilon) = S(\Upsilon) \text{ for all } \Upsilon \neq \Upsilon_1, (x_1 + 1, y_1)$$
$$S_1(\Upsilon_1) = 1, \; S_1(x_1 + 1, y_1) = 0.$$

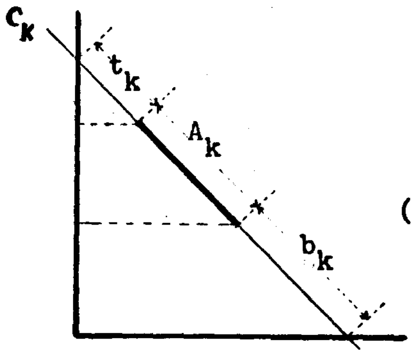
S_1 is bounded because any path through Υ_1 either coincides with an S-path from $(x_1, y_1 + 1)$ onwards or stops at $(x_1 + 1, y_1)$. In the same manner, S_i can be constructed from S_{i-1}, $i = 1, 2, \ldots, k - 1$. In S_{k-1}, $\Upsilon_0, \Upsilon_1, \ldots, \Upsilon_{k-1}$ are continuation points, Υ_k is a boundary point, and Υ_{k+1} is a continuation point. Clearly, Υ_k is a non-essential boundary point of S_{k-1} which is of size n. Thus S_{k-1} contains more than n + 1 boundary points and hence does S.

Theorem 1.5.1 reflects that for a given n there exists more than one sampling plan of size n that are simple. How many such simple sampling plans can we have corresponding to a particular value of n? In answering this question we would like to refer to a paper written by Brainerd and Narayana [12].

<u>Theorem 1.5.2</u>. The number of simple sampling plans of size n is

$$n^{-1} \binom{3n}{n-1}.$$

Proof. Let C be the class of continuation points and let $A_k = C \cap C_k$, $C_k = \{(x, y): x + y = k\}$. Clearly, A_k is non-empty if and only if $k < n$. For every simple sampling plan of size n, each non-empty A_k is characterized by the distance t_k between its top and (0, k) and the distance b_k between its bottom and (k, 0). (Refer figure 1.5.5)



(figure 1.5.5)

The only restrictions on $\{t_k, b_k\}$ are

(1) $t_k + b_k \leq k$, $k = 0, 1, \ldots, n - 1$

$0 \leq t_k \leq t_{k+1}$, $0 \leq b_k \leq b_{k+1}$, $k = 0, 1, \ldots, n - 2$.

The number of different solutions of the above set of inequalities is the number of different simple sampling plans of size n. Now let $(x, y)_n$ denotes the number of simple sampling plans of size n with $t_{n-1} = x$ and $b_{n-1} = y$, then (1) implies that

(2) $$(x, y)_n = \sum_{a=0}^{x} \sum_{b=0}^{y} (a, b)_{n-1} \quad \text{for } x + y < n$$

$$= 0 \quad \text{for } x + y \geq n.$$

We note that $(a, b)_n = (b, a)_n$. The condition $(0,0)_1 = 1$ together with (2) determines $(x, y)_n$ recursively for all non-negative integers x, y and positive integers n. Then the number of different simple sampling plans of size n with $t_{n-1} + b_{n-1} = k$ is $k_{(n)}$, where

(3)

$$k_{(n)} = \sum_{x+y=k} (x,y)_n$$

$$= \sum_{x+y=k} \sum_{a=0}^{x} \sum_{b=0}^{y} (a,b)_{n-1}$$

$$= \sum_{a=0}^{0} \sum_{b=0}^{k} (a,b)_{n-1} + \sum_{a=0}^{1} \sum_{b=0}^{k-1} (a,b)_{n-1} + \dots + \sum_{a=0}^{k-1} \sum_{b=0}^{1} (a,b)_{n-1} + \sum_{a=0}^{k} \sum_{b=0}^{0} (a,b)_{n-1}$$

$$= (k+1) \sum_{a+b=0} (a,b)_{n-1} + (k) \sum_{a+b=1} (a,b)_{n-1} + \dots + 2 \sum_{a+b=k-1} (a,b)_{n-1} + 1 \sum_{a+b=k} (a,b)_{n-1}$$

$$= \sum_{c=0}^{k} (k - c + 1) \sum_{a+b=c} (a,b)_{n-1}$$

$$= \sum_{c=0}^{k} (k - c + 1) c_{(n-1)} \qquad \text{for } k < n$$

$$k_{(n)} = 0 \qquad \text{for } k \geq n$$

and $0_{(1)} = 1$.

These conditions determine $k_{(n)}$ recursively. Experiment leads to the conjectured solution

(4)
$$k_{(n)} = \frac{2n-2k}{2n+k}\binom{2n+k}{2n}$$

$$= \frac{2n-2k}{2n+k}\binom{2n+k}{k}$$

$$= \binom{2n+k}{k} - 3\binom{2n+k-1}{k-1} \quad \text{for } k \leqslant n$$

$$k_{(n)} = 0 \text{ for } k \geqslant n.$$

We shall make use the following formula which will be proved in the accompanied remark.

(5)
$$\sum_{b=0}^{c} \binom{a+b}{b} = \binom{a+c+1}{c}.$$

Thus the total number of simple sampling plans of size n is

$$\sum_{k=0}^{n-1} k_{(n)} = \sum_{k=0}^{n-1} \binom{2n+k}{k} - 3\sum_{k=0}^{n-1}\binom{2n+k-1}{k-1}$$

$$= \binom{3n}{n-1} - 3\binom{3n-1}{n-2}$$

$$= \frac{1}{n}\binom{3n}{n-1}.$$

<u>Remark</u>.
$$\sum_{b=0}^{c}\binom{a+b}{b} = \binom{a+c+1}{c}.$$

Proof. By induction.

When $c = 0$, it is true

$$\text{suppose} \sum_{b=0}^{k} \binom{a+b}{b} = \binom{a+k+1}{k}$$

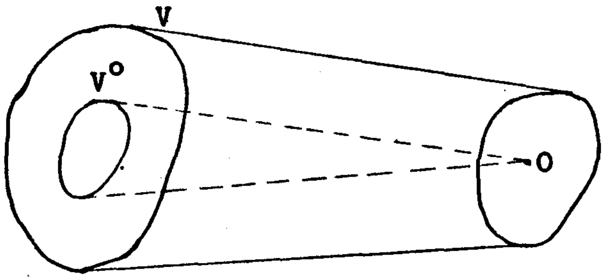
$$\text{then} \sum_{b=0}^{k+1} \binom{a+b}{b} = \sum_{b=0}^{k} \binom{a+b}{b} + \binom{a+k+1}{k+1}$$

$$= \binom{a+k+1}{k} + \binom{a+k+1}{k+1}$$

$$= \binom{a+k+2}{k+1} .$$

<u>Theorem 1.5.3</u>. If the boundary of a sampling plan of size n contains $n + 1 + k$ points, $k > 0$, then there exist exactly k linearly independent unbiased estimators of 0.

Proof. Let $\gamma_1, \gamma_2, \ldots, \gamma_{n+1+k}$ be the boundary points. Each estimator f can be regarded as a vector $(f_1, \ldots, f_{n+1+k})$ where $f(\gamma_j) = f_j$, $j = 1, \ldots, n + 1 + k$. Thus, the space of estimators can be considered as an $(n + 1 + k)$- dimensional vector space V. Theorem 1.3.1 assures that p^m is estimable unbiasedly for all non-negative integers $m \leq n$. Thus $p^{Y(\gamma)}q^{X(\gamma)}$, a polynomial in p of degree $N(\gamma)$ is estimable unbiasedly and so is any linear combination of such forms. It follows that all polynomials in p of degree at most n are estimable unbiasedly. Now, since the expectation of every estimator is a linear combination of the polynomials $P(p; \gamma)$, $\gamma \in B$, the expectation operator E is a linear mapping from V ONTO the (n + 1)- dimensional linear space of polynomials in p of degree at most n. The subspace $V^0 = \{ f: E(f/p) = 0 \}$ is the null space of this mapping and it follows from the standard theorems concerning rank and nullity that V^0 has dimension k. (See figure 1.5.6)

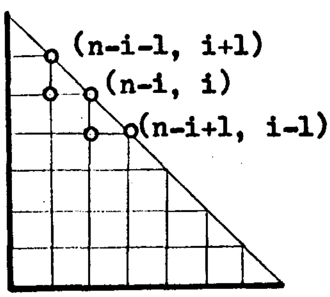


(figure 1.5.6)

Theorem 1.5.4. The boundary B of a sampling plan of size n contains at least two contiguous points of index n.

Proof. Suppose the contrary, i.e. suppose that there is only one boundary point on $\{ (x, y): x + y = n \}$.

If (n, 0) is that boundary point, then (n - 1, 0) must be a continuation point and hence (n - 1, 1) is a boundary point. This is a contradiction. Similarly for (0, n).



Now if (n - i, i) is in B, i = 1, 2, ..., n - 1, then, by hypothesis, (n - i - 1, i + 1) and (n - i + 1, i - 1) are inaccessible points. It follows that (n - i - 1, i) and (n - i, i - 1) are not continuation points and hence (n - i, i) is inaccessible. Again, this is a contradiction.

Hence there are at least 2 contiguous boundary points of index n.

Part II SEQUENTIAL SAMPLING PLANS

2.1 Introduction

In part I a description of optimality of binomial sampling plans has been given and we claimed that single binomial sample plans are efficient sampling plans. At present we shall be concerned with the study of games which enables us to discuss the optimality problem in a more general situation. Our main aim will be to show that it is always possible to construct an optimal sequential-sampling plan. The single-experiment (fixed sample size experiment) game though often encountered in statistical-decision theory and practice, is a very special type of game. In most real situations, observations are costly, and instead of exhausting all N (sample size) observations, the experimenter might greatly improve his situation if at each stage of experimentation he balances the cost of taking future observations against the expected gain in information from such observations. We shall thus restrict ourselves to the consideration of truncated sequential games.

2.2 Basic Notions of Truncated Sequential Games

For the sake of completeness we shall summarize a few definitions. The materials we shall encounter can be found in [4].

A sample space is a triple $\mathcal{Z} = (Z, \Omega, p')$ where Z, the outcome space, and Ω, the parameter space, are non-empty sets, and p' is a function defined on $Z \times \Omega$ such that, for a fixed $w \in \Omega$, p'_w is a probability distribution on Z.

Let $g_1, g_2, \ldots, g_N$ be random variables defined on Z. Let $\mathcal{X} = (X, \Omega, p)$ be the new sample space in which $X = X_1 \times X_2 \times \ldots \times X_N$, where X_i is the range of g_i, $i = 1, 2, \ldots, N$, and for each $w \in \Omega$, p_w is the joint probability distribution of $g_1, g_2, \ldots, g_N$.

Let A be a space of terminal actions. A may be any arbitrary set.

Let D be a set of functions which map $J \times X$ into A, where J is the set $\{0, 1, \ldots, N\}$ and such that, if $x, y \in X$ and $x_i = y_i$, $i = 1, 2, \ldots, j$, then $d(j,x) = d(j, y)$ for all $d \in D$.

Let $\mathfrak{S}$ be a class of partitions of X such that if $S \in \mathfrak{S}$ then $S = (S_0, S_1, \ldots, S_N)$ where S_j is a cylinder set over $K = \{\gamma \in J: 0 < \gamma \leq j\}$, i.e., if $x, y \in X$ and $x_i = y_i$ for $i = 1, 2, \ldots, j$ then $x \in S_j$ if and only if $y \in S_j$.

The product space $\mathfrak{S} \times D$ is the space of sequential-decision functions. A partition S in $\mathfrak{S}$ determines a sequential-sampling plan. The sets S_j of S are sometimes referred to as "stopping regions". Since S_j is a cylinder set, it is always known in a sequence of experiments whether or not the observations belong to a stopping region, i.e. whether the experiment is to be continued or terminated. Thus, a sequential-decision function $(S, d) \in \mathfrak{S} \times D$ is in fact a procedure that tells the experimenter at each state whether to take another observation or to stop experimenting and make a decision.

Perhaps, the following example will illustrate some of the concepts mentioned above. An inspector at an Army Proving Ground has to decide whether to accept or reject a lot of rocket-propellent powder on the basis of the performance of 5 randomly selected rockets which he is to fire. A propellent is called defective if the pressure developed in the rocket chamber is 3,000 pounds or more per square inch. The acceptability of the lot depends on the proportion of defective items in the lot, and the decision is to be based on the number of defective rockets found in the sample of size 5. For each rocket fired, let y be the random variable which has value 1 if the propellent is defective and 0 otherwise. Then, if all 5 rockets are fired, the space of outcomes X consists of the 32 points $x^1, x^2, \ldots, x^{32}$ given below:

$x^1=(0,0,0,0,0)$	$x^9=(0,0,0,1,0)$	$x^{17}=(0,0,0,0,1)$	$x^{25}=(0,0,0,1,1)$
$x^2=(0,1,0,0,0)$	$x^{10}=(0,1,0,1,0)$	$x^{18}=(0,1,0,0,1)$	$x^{26}=(0,1,0,1,1)$
$x^3=(1,0,0,0,0)$	$x^{11}=(1,0,0,1,0)$	$x^{19}=(1,0,0,0,1)$	$x^{27}=(1,0,0,1,1)$
$x^4=(1,1,0,0,0)$	$x^{12}=(1,1,0,1,0)$	$x^{20}=(1,1,0,0,1)$	$x^{28}=(1,1,0,1,1)$
$x^5=(0,0,1,0,0)$	$x^{13}=(0,0,1,1,0)$	$x^{21}=(0,0,1,0,1)$	$x^{29}=(0,0,1,1,1)$
$x^6=(0,1,1,0,0)$	$x^{14}=(0,1,1,1,0)$	$x^{22}=(0,1,1,0,1)$	$x^{30}=(0,1,1,1,1)$
$x^7=(1,0,1,0,0)$	$x^{15}=(1,0,1,1,0)$	$x^{23}=(1,0,1,0,1)$	$x^{31}=(1,0,1,1,1)$
$x^8=(1,1,1,0,0)$	$x^{16}=(1,1,1,1,0)$	$x^{24}=(1,1,1,0,1)$	$x^{32}=(1,1,1,1,1)$

The space A of actions consists of only two points a_1, a_2 where a_1 stands for the acceptance and a_2 the rejection of the lot. A possible sequential rule is as follows: fire the rockets 1 at a time and stop as soon as 2 defectives are found; in any case stop when 5 rockets have been fired. Let n be the number of rockets fired by this rule, n = 2, 3, 4, 5. In order to decide whether to accept or reject the lot, a possible criterion may be given as: take action a_1 if $n \geq 4$, and action a_2 otherwise. We shall now consider a procedure δ that to each point x of X assigns two numbers, an integer j = 1, 2, ..., 5 which specifies the number of coordinates of x to observe before terminating experimentation, and an element a of A which specifies what action is to be taken once experimentation is terminated. The function δ then has the following values:

$\delta(x^1)=(5,a_1)$	$\delta(x^9)=(5,a_1)$	$\delta(x^{17})=(5,a_1)$	$\delta(x^{25})=(5,a_1)$
$\delta(x^2)=(5,a_1)$	$\delta(x^{10})=(4,a_1)$	$\delta(x^{18})=(5,a_1)$	$\delta(x^{26})=(4,a_1)$
$\delta(x^3)=(5,a_1)$	$\delta(x^{11})=(4,a_1)$	$\delta(x^{19})=(5,a_1)$	$\delta(x^{27})=(4,a_1)$
$\delta(x^4)=(2,a_2)$	$\delta(x^{12})=(2,a_2)$	$\delta(x^{20})=(2,a_2)$	$\delta(x^{28})=(2,a_2)$
$\delta(x^5)=(5,a_1)$	$\delta(x^{13})=(4,a_1)$	$\delta(x^{21})=(5,a_1)$	$\delta(x^{29})=(4,a_1)$
$\delta(x^6)=(3,a_2)$	$\delta(x^{14})=(3,a_2)$	$\delta(x^{22})=(3,a_2)$	$\delta(x^{30})=(3,a_2)$
$\delta(x^7)=(3,a_2)$	$\delta(x^{15})=(3,a_2)$	$\delta(x^{23})=(3,a_2)$	$\delta(x^{31})=(3,a_2)$
$\delta(x^8)=(2,a_2)$	$\delta(x^{16})=(2,a_2)$	$\delta(x^{24})=(2,a_2)$	$\delta(x^{32})=(2,a_2)$.

Here a partition S is given by (S_2, S_3, S_4, S_5) where

$$S_2 = \{ x^4, x^8, x^{12}, x^{16}, x^{20}, x^{24}, x^{28}, x^{32} \}$$

$S_3 = \{ x^6, x^7, x^{14}, x^{15}, x^{22}, x^{23}, x^{30}, x^{31} \}$

$S_4 = \{ x^{10}, x^{11}, x^{13}, x^{26}, x^{27}, x^{29} \}$

$S_5 = \{ x^1, x^2, x^3, x^5, x^9, x^{17}, x^{18}, x^{19}, x^{21}, x^{25} \}$.

The function d maps the sets S_4 and S_5 into a_1; S_2 and S_3 into a_2.

In order to define a truncated sequential game, we also need a cost function and a loss function. A cost function is a non-negative bounded function c defined on J x X such that if x and y are in X and $x_i = y_i$ for i = 1, 2, ..., j, then c(j,x)=c(j,y). The sampling cost of experimentation is defined only on the subexperiments actually performed. A loss function is a bounded non-negative function L defined on Ω x A. L(w,a) represents the loss when p_w is the true distribution on X and the statistician takes action a. Using these two notions, the risk to the statistician is given by the function ρ which is defined as follows

$$\rho\,(w,S,d) = \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j,x) + L(w,d(j,x)) \right] p_w(x).$$

The triple (Ω, $\mathcal{S}$ x D, ρ) is a truncated-sequential game.

2.3 Bayes Procedures for Sequential Games

In this section we are going to prove the following theorems 2.3.1 and 2.3.2 which are direct consequence of the results of Blackwell and Girshick (4) .

Let Ξ be the class of mixed strategies for nature in a statistical game. By a mixed strategy for nature is meant a probability distribution over Ω . Then, for each ξ in Ξ, the expected risk is

$$(2.3.1)\quad \rho(\xi, S, d) = \sum_{j=0}^{N} \sum_{x \in S_j} \sum_{w} \left[c(j,x) + L(w, d(j,x)) \right] \xi(w) p_w(x).$$

Let $\mathcal{B}_j$ be the collection of all sets B_j such that, for some $x \in X$, $y \in B_j$ if and only if $y_i = x_i$ for all $i \leq j$. That is, B_j is a cylinder set over $K = \{ \gamma \in J: 0 < \gamma \leq j \}$. For any $x \in X$ we also define $F_j(x)$ as the set of all points having the same first j coordinates as x. Thus, for a fixed j, the sets $F_j(x)$ are equal for all $x \in B_j$.

We see that $\mathcal{B}_j$ so defined is a partition of X.

For any bounded function h on $\Omega \times X$, let $E_{j\xi}(h)$ be the conditional expectation of h given $x_1, \ldots, x_j$, when w has distribution ξ , and, for a fixed w, x has distribution p_w. For any x, the value of $E_{j\xi}(h)$ at x is

$$E_{j\mathfrak{Z}}(h) = \frac{\sum_{y \in F_j(x)} \sum_{w} \mathfrak{Z}(w)p_w(y)h(w,y)}{\sum_{y \in F_j(x)} \sum_{w} \mathfrak{Z}(w)p_w(y)} .$$

Denote $$P_{\mathfrak{Z}}(x) = \sum_{w} \mathfrak{Z}(w)p_w(x)$$

then

(2.3.2) $$E_{j\mathfrak{Z}}(h) = \frac{\sum_{y \in F_j(x)} \sum_{w} \mathfrak{Z}(w)p_w(y)h(w,y)}{\sum_{y \in F_j(x)} P_{\mathfrak{Z}}(y)} .$$

For $j = 0$, $\mathcal{B}_0 = \{X\}$, we write $E_{\mathfrak{Z}}(h)$ instead of $E_{0\mathfrak{Z}}(h)$.

The function $E_{j\mathfrak{Z}}(h) = V(x)$ is a function of $x_1, x_2, \ldots, x_j$ only and has the following property: For any bounded function f on X which depends only on $x_1, x_2, \ldots, x_j$,

(2.3.3) $$\sum_{w} \sum_{x} \mathfrak{Z}(w)p_w(x)f(x)h(w,x) = \sum_{w} \sum_{x} \mathfrak{Z}(w)p_w(x)f(x)v(x)$$

$$= \sum_{x} f(x)v(x)P_{\mathfrak{Z}}(x).$$

The second equality is immediate, we shall prove the first one.

In fact,

$$\sum_{w} \sum_{x} \mathfrak{Z}(w)p_w(x)f(x)v(x) = \sum_{w} \sum_{x} \mathfrak{Z}(w)p_w(x)f(x) \frac{\sum_{y \in F_j(x)} \sum_{\theta} \mathfrak{Z}(\theta)p_\theta(y)h(\theta,y)}{\sum_{y \in F_j(x)} P_{\mathfrak{Z}}(y)}$$

$$= \sum_{x} P_{\mathfrak{Z}}(x) \frac{\sum_{y \in F_j(x)} \sum_{\theta} f(y)\mathfrak{Z}(\theta)p_{\theta}(y)h(\theta,y)}{\sum_{y \in F_j(x)} P_{\mathfrak{Z}}(y)} \quad , \text{ f is constant}$$

over a given set $F_j(x)$

$$= \sum_{B_j \in \mathcal{B}_j} \sum_{x \in B_j} P_{\mathfrak{Z}}(x) \frac{\sum_{y \in F_j(x)} \sum_{\theta} f(y)\mathfrak{Z}(\theta)p_{\theta}(y)h(\theta,y)}{\sum_{y \in F_j(x)} P_{\mathfrak{Z}}(y)}$$

$$= \sum_{B_j \in \mathcal{B}_j} \left(\sum_{x \in B_j} P_{\mathfrak{Z}}(x) \right) \frac{\sum_{y \in B_j} \sum_{\theta} f(y)\mathfrak{Z}(\theta)p_{\theta}(y)h(\theta,y)}{\sum_{y \in B_j} P_{\mathfrak{Z}}(y)}$$

$$= \sum_{B_j \in \mathcal{B}_j} \sum_{y \in B_j} \sum_{\theta} f(y)\mathfrak{Z}(\theta)p_{\theta}(y)h(\theta,y)$$

$$= \sum_{x} \sum_{w} f(x)\mathfrak{Z}(w)p_w(x)h(w,x).$$

In case h is a function of x only, i.e. $h(w,x) = g(x)$, we have, from equation (2.3.3).

$$(2.3.4) \quad \sum_{x} f(x)g(x)P_{\mathfrak{Z}}(x) = \sum_{x} f(x)(E_{j\mathfrak{Z}}[g(x)])P_{\mathfrak{Z}}(x).$$

Now let $h(w, x) = h_j(w, x) = c(j, x) + L(w, d(j, x))$ and let $f(x) = f_j(x)$ be the indicator function of S_j, i.e. $f_j(x) = 1$ if $x \in S_j$ and zero otherwise; then (2.3.1) can be expressed as

(2.3.5)

$$\rho(\xi, S, d) = \sum_{j=0}^{N} \sum_{x \in S_j} \sum_{w} \left[c(j, x) + L(w, d(j, x)) \right] \xi(w) p_w(x)$$

$$= \sum_{j=0}^{N} \sum_{x \in S_j} E_{j\xi} \left[c(j, x) + L(w, d(j, x)) \right] P_\xi(x), \text{ by (2.3.3)}$$

$$= \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j, x) + E_{j\xi} \big(L(w, d(j, x)) \big) \right] P_\xi(x)$$

since $c(j, x)$ is constant for $x \in S_j$.

In most situations ξ will be fixed. For a given a in A, we thus define

(2.3.6) $$T_j(x, a) = E_{j\xi} \left[L(w, a) \right]$$

(2.3.7) $$T_j^*(x) = \inf_{a \in A} T_j(x, a).$$

Theorem 2.3.1. For a fixed ξ, there is a sequence of terminal-decision functions d_n such that

$$\lim_{n \to \infty} \rho(\xi, S, d_n) = \rho^*(\xi, S) = \inf_d \rho(\xi, S, d)$$

uniformly in S, where

$$\rho^*(\xi, S) = \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j, x) + T_j^*(x) \right] P_\xi (x).$$

Proof. Fix j and x in (2.3.7). Then, for any n we can find a point a(j, x) in A depending on j and x such that

(1) $$T_j(x, a(j, x)) \leq T_j^*(x) + \frac{1}{n}.$$

Since for a fixed n we can define d_n by

$$d_n(j, x) = a(j, x)$$

satisfying (1) for each j and x, then for this d_n,

(2) $$\rho(\xi, S, d_n) = \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j, x) + E_{j\xi}(L(w, d_n(j, x))) \right] P_\xi (x)$$

$$= \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j, x) + E_{j\xi}(L(w, a(j, x))) \right] P_\xi (x)$$

$$= \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j, x) + T_j(x, a(j, x)) \right] P_\xi (x)$$

$$\leq \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j, x) + T_j^*(x) + \frac{1}{n} \right] P_\xi (x)$$

$$= \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j, x) + T_j^*(x) \right] P_{\xi}(x) + \frac{1}{n} \sum_{j=0}^{N} \sum_{x \in S_j} P_{\xi}(x)$$

$$= \rho^*(\xi, S) + \frac{1}{n}$$

for all S. Thus

(3) $\inf_d \rho(\xi, S, d) \leq \rho(\xi, S, d_n) \leq \rho^*(\xi, S) + \frac{1}{n}.$

On the other hand, it follows from (2.3.5) that for all d

$$\rho(\xi, S, d) \geq \sum_{j=0}^{N} \sum_{x \in S_j} \left[c(j, x) + T^*(x) \right] P_{\xi}(x) = \rho^*(\xi, S)$$

so that

(4) $\inf_d \rho(\xi, S, d) \geq \rho^*(\xi, S).$

Now, since (3) and (4) hold for all n and S, the theorem is proved.

Theorem 2.3.1 says that, for any arbitrary truncated sequential-sampling plan S and for any a priori probability distribution ξ on Ω, there always exists, at least to within any $\varepsilon > 0$, an optimal terminal-decision rule. Moreover, the Bayes risk for a given ξ and arbitrary S may be taken to be $\rho^*(\xi, S)$, since this value can be approximated to arbitrary accuracy by an appropriate choice of d_n.

We shall next show that, for a given ξ, there exists an optimal sequential-sampling plan. This sampling plan will be briefly characterized as follows: at any stage of experimentation, if there exists a continuation that will reduce the risk below the present level, we perform an additional subexperiment. If, on the other hand, there exists no such continuation, we stop experimenting. A constructive proof is given as follows:

Let ξ be fixed, and for any function $h(w, x)$ we write $E_j(h)$ for the expression in (2.3.2). We also write

$$U_j(x) = c(j, x) + T_j^*(x).$$

We define $\quad \alpha_N(x) = U_N(x)$

and

$$\alpha_{N-1}(x) = \min\left[U_{N-1}(x),\ E_{N-1}(\alpha_N(x))\right]$$

$$\alpha_{N-2}(x) = \min\left[U_{N-2}(x),\ E_{N-2}(\alpha_{N-1}(x))\right]$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\alpha_1(x) = \min\left[U_1(x),\ E_1(\alpha_2(x))\right]$$

$$\alpha_0 = \min\left[U_0,\ E(\alpha_1(x))\right]$$

where

$$U_0 = \inf_{a \in A} E\left[L(w, a)\right] = \inf_{a \in A} \sum_w L(w, a)\,\xi(w).$$

Observe that if we had performed all N subexperiments and obtained $x = (x_1, \ldots, x_N)$ then $\alpha_N(x)$ would represent the best we can do with this x. Also, both $U_{N-1}(x)$ and $E_{N-1}(U_N(x))$

depend only on $x_1, \ldots, x_{N-1}$ and, moreover, $U_{N-1}(x)$ represents the best we can do with N-1 observations, and $E_{N-1}(U_N(x))$ represents the average of the best that we can do with an additional observation. Thus $\alpha_{N-1}(x)$ represents the smaller of two risks - the risk of stopping with N-1 observations and the average risk from an additional observation. Similarly, $\alpha_{N-2}(x)$ depends only on $x_1, \ldots, x_{N-2}$ and stands for the smaller of two risks - the risk of stopping with N-2 observations and the average risk of going on, provided that if we go on experimenting we do the best we can, i.e. we stop with one more observation if $\alpha_{N-1}(x) = U_{N-1}(x)$ and take a second observation if $\alpha_{N-1}(x) = E_{N-1}(U_N(x))$. The interpretation of $\alpha_{N-k}(x)$ for any k is now clear.

Let

(2.3.8) $$S_j^* = \left\{x: U_i(x) > \alpha_i(x) \text{ for } i < j,\ U_j(x) = \alpha_j(x)\right\}$$

$$S^* = (S_0^*, S_1^*, \ldots, S_N^*).$$

Then S^* forms a partition of the sample space X. Indeed, for any $m < n$,

$$S_m^* = \left\{x: U_i(x) > \alpha_i(x) \text{ for } i < m,\ U_m(x) = \alpha_m(x)\right\}$$

$$S_n^* = \left\{x: U_i(x) > \alpha_i(x) \text{ for } i < n,\ U_n(x) = \alpha_n(x)\right\}$$

it follows that

$$U_m(x) > \alpha_m(x) \text{ for all } x \in S_n^*$$

and

$$U_m(x) = \alpha_m(x) \text{ for all } x \in S^*_m$$

so that S^*_m and S^*_n are disjoint. Moreover, every point of X belongs to some S^*_j. We note that $\alpha_j(x) \leq U_j(x)$ for all j, with equality holding for j = N. Let $\gamma \leq N$ be the smallest non-negative integer for which the equality sign holds in this expression. Then $x \in S^*_\gamma$. Also the sets $S^*_j (j \neq 0)$ are clearly cylinder sets over $K = \left\{ \gamma \in J: 0 < \gamma \leq j \right\}$.

Thus S* is a possible sequential-sampling plan and is characterized as follows: at the j^{th} stage of experimentation j = 0, 1, ..., we compare the present risk $U_j(x)$ with the average risk $\alpha_j(x)$ resulting from a continuation if at each future stage we did the best we could with the resulting observations. We stop sampling if $U_j(x) = \alpha_j(x)$ and take another observation if $U_j(x) > \alpha_j(x)$. We shall now show that S* is in fact a Bayes sequential-sampling plan.

<u>Theorem 2.3.2</u>: The sequential-sampling plan S* defined by (2.3.8) is Bayes against ξ, i.e.

$$\rho(\xi, S^*) = \rho^*(\xi) = \min_S \rho(\xi, S)$$

and furthermore,

$$\rho^*(\xi) = \sum_{j=0}^{N} \sum_{x \in S^*_j} \alpha_j(x) P_\xi(x) = \alpha_0$$

Proof. Let $S = (S_0, S_1, \ldots, S_N)$ be any arbitrary truncated sequential-sampling plan, and let

$$T_\gamma = S_\gamma \cup S_{\gamma+1} \cup \dots \cup S_N.$$

Define

$$g(\gamma) = \sum_{j=0}^{\gamma-1} \sum_{x \in S_j} \alpha_j(x) P_\mathfrak{J}(x) + \sum_{x \in T_\gamma} \alpha_\gamma(x) P_\mathfrak{J}(x)$$

$$= \sum_{j=0}^{\gamma} \sum_{x \in S_j} \alpha_j(x) P_\mathfrak{J}(x) + \sum_{x \in T_{\gamma+1}} \alpha_\gamma(x) P_\mathfrak{J}(x)$$

then

$$g(N) = \sum_{j=0}^{N} \sum_{x \in S_j} \alpha_j(x) P_\mathfrak{J}(x)$$

and

$$g(0) = \sum_{x \in T_0} \alpha_0(x) P_\mathfrak{J}(x)$$

$$= \alpha_0 \sum_{x} P_\mathfrak{J}(x)$$

$$= \alpha_0.$$

Now the set $T_{\gamma+1}$ is defined by $x \notin S_j$ for $j = 0, 1, \dots, \gamma$, so that $T_{\gamma+1}$ depends only on $x_1, x_2, \dots, x_\gamma$. Hence letting $\beta_{\gamma+1}$ represent the characteristic function of the set $T_{\gamma+1}$ and taking $\beta_{\gamma+1}$ for f and $\alpha_{\gamma+1}$ for g in (2.3.4), we obtain

$$\sum_{x \in T_{\gamma+1}} \alpha_{\gamma+1}(x) P_\mathfrak{J}(x) = \sum_{x \in T_{\gamma+1}} E_\gamma \left[\alpha_{\gamma+1}(x) \right] P_\mathfrak{J}(x).$$

Thus,

$$g(\gamma + 1) = \sum_{j=0}^{\gamma} \sum_{x \in S_j} \alpha_j(x) P_{\xi}(x) + \sum_{x \in T_{\gamma+1}} \alpha_{\gamma+1}(x) P_{\xi}(x)$$

$$= \sum_{j=0}^{\gamma} \sum_{x \in S_j} \alpha_j(x) P_{\xi}(x) + \sum_{x \in T_{\gamma+1}} E_{\gamma} \left[\alpha_{\gamma+1}(x) \right] P_{\xi}(x).$$

However,

$$\alpha_{\gamma}(x) \leq E_{\gamma} \left[\alpha_{\gamma+1}(x) \right].$$

Hence

(1) $$g(\gamma + 1) \geq \sum_{j=0}^{\gamma} \sum_{x \in S_j} \alpha_j(x) P_{\xi}(x) + \sum_{x \in T_{\gamma+1}} \alpha_{\gamma}(x) P_{\xi}(x) = g(\gamma)$$

for $\gamma = 0, 1, \ldots, N-1$. Thus $g(\gamma)$ is a non-decreasing function of γ. Again by Theorem 2.3.1, we have

(2) $$\rho(\xi, S) = \sum_{j=0}^{N} \sum_{x \in S_j} U_j(x) P_{\xi}(x)$$

$$\geq \sum_{j=0}^{N} \sum_{x \in S_j} \alpha_j(x) P_{\xi}(x) = g(N).$$

Hence, for all S we have

$$\rho(\xi, S) \geq g(N) \geq g(0) = \alpha_0.$$

However, if $S = S^*$, for $x \in T_{\gamma+1}$, $x \notin S_0^*, S_1^*, \ldots, S_{\gamma}^*$, it follows that $U_{\gamma}(x) > \alpha_{\gamma}(x)$ and hence $\alpha_{\gamma}(x) = E_{\gamma}\left[\alpha_{\gamma+1}(x)\right]$. Thus (1) becomes an equality for $\gamma = 0, 1, \ldots, N - 1$. Also inequality (2) becomes an equality since $U_j(x) = \alpha_j(x)$ for $x \in S_j^*$. This completes the proof of the theorem.

Remark. Since each set S_j^*, $j = 1, 2, \ldots, N$ of the optimal sequential sampling plan S^* is a cylinder set over $K = \{\gamma: \ 0 < \gamma \leq j\}$, we can justify whether or not a sequence of j observations is in a stopping region S_j^*. For any point in S_j^*, we shall cease sampling at j^{th} stage of experimentation.

2.4 Examples on Construction of optimal Sequential Binomial Sampling Plans for Point Estimation Problems.

Again, we shall concentrate our attention to the binomial sampling plans. Let us consider the following point estimation problems. Suppose

(a) y is a binomial random variable with $p_w(y = 1) = w$, $p_w(y = 0) = 1 - w$ for $0 \leq w \leq 1$.

(b) ξ is uniform on the interval $[0, 1]$, i.e. $\xi(w) = 1$ for all $0 \leq w \leq 1$.

(c) the cost per observation is 1 unit, and

(d) the loss $L(w, a) = k(w - a)^2$ for any estimate a, $0 \leq a \leq 1$.

We shall apply the construction developed in the last section to construct optimal sequential binomial sampling plans S* relative to certain numerical values of k.

For any point x in X, let

$$m_j = \sum_{i=1}^{j} x_i$$

= number of ones in the first j observations $x_1, \ldots, x_j$.

Then, by expressions (2.3.2) and (2.3.6),

$$T_j(x,\ a) = E_{j\xi}\left[L(w,\ a)\right]$$

$$= k\int_0^1 (w - a)^2 \binom{j}{m_j} w^{m_j}(1 - w)^{j-m_j}\,dw \Big/ \int_0^1 \binom{j}{m_j} w^{m_j}(1 - w)^{j-m_j}\,dw$$

$$= k\int_0^1 (w^2 - 2aw + a^2) w^{m_j}(1 - w)^{j-m_j}\,dw \Big/ \int_0^1 w^{m_j}(1 - w)^{j-m_j}\,dw.$$

The integrals are complete Beta-functions, so that

$$T_j(x,\ a) = k\,\frac{B_1(m_j+3,\ j-m_j+1) - 2aB_1(m_j+2,\ j-m_j+1) + a^2 B_1(m_j+1,\ j-m_j+1)}{B_1(m_j+1,\ j-m_j+1)}\,.$$

Recalling the expressions

$$B_1(m,n) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}, \qquad \Gamma(m+1) = m!$$

Then

$$T_j(x,a) = k\,\frac{\Gamma(j+2)}{\Gamma(m_j+1)\,\Gamma(j-m_j+1)}\left[\frac{\Gamma(m_j+3)\,\Gamma(j-m_j+1)}{\Gamma(j+4)} - 2a\,\frac{\Gamma(m_j+2)\,\Gamma(j-m_j+1)}{\Gamma(j+3)} + a^2\,\frac{\Gamma(m_j+1)\,\Gamma(j-m_j+1)}{\Gamma(j+2)}\right]$$

$$= k\,\frac{(j+1)!}{m_j!}\left[\frac{(m_j+2)!}{(j+3)!} - 2a\,\frac{(m_j+1)!}{(j+2)!} + a^2\,\frac{m_j!}{(j+1)!}\right]$$

$$= k \left[\frac{(m_j+2)(m_j+1)}{(j+3)(j+2)} - 2a \frac{m_j+1}{j+2} + a^2 \right]$$

$$= k\,\varphi(a), \text{ say.}$$

To find the infimum of $T_j(x, a)$ when a ranges over $A = [0,1]$, it suffices to minimize the value of $\varphi(a)$. We have

$$\varphi'(a) = 0 = -2\,\frac{m_j+1}{j+2} + 2a$$

$$= 2\left(a - \frac{m_j+1}{j+2}\right).$$

It follows that $a = \frac{m_j+1}{j+2}$ and clearly $\varphi\left(\frac{m_j+1}{j+2}\right)$ is a minimum value.

Thus $$T^*_j(x) = k\left[\frac{(m_j+2)(m_j+1)}{(j+3)(j+2)} - \left(\frac{m_j+1}{j+2}\right)^2\right]$$

$$= k\,\frac{m_j+1}{j+2}\left(\frac{m_j+2}{j+3} - \frac{m_j+1}{j+2}\right)$$

$$= k\,\frac{(m_j+1)(j-m_j+1)}{(j+2)^2(j+3)}$$

and

$$U_j(x) = c(j,\ x) + T^*_j(x)$$

or (2.4.1) $$U_j(x) = j + k\,\frac{(m_j+1)(j+1-m_j)}{(j+2)^2(j+3)}.$$

We shall next develop a formula for calculating average risk from taking additional observations. Let $p(x_{j+1} = 1/m_j)$ be the conditional probability that the (j+1)th observation will be 1 given the value of m_j. Then

$$p(x_{j+1}=1/m_j) = \int_0^1 p_w(x_{j+1}=1)p_w(m_j)\mathfrak{J}(w)dw \Big/ \int_0^1 p_w(m_j)\mathfrak{J}(w)dw$$

$$= \int_0^1 w\binom{j}{m_j} w^{m_j}(1-w)^{j-m_j}dw \Big/ \int_0^1 \binom{j}{m_j} w^{m_j}(1-w)^{j-m_j}dw$$

$$= \int_0^1 w^{m_j+1}(1-w)^{j-m_j}dw \Big/ \int_0^1 w^{m_j}(1-w)^{j-m_j}dw$$

$$= B_1(m_j+2,\ j-m_j+1) \Big/ B_1(m_j+1,\ j-m_j+1)$$

$$= \frac{\Gamma(m_j+2)\Gamma(j-m_j+1)}{\Gamma(j+3)} \times \frac{\Gamma(j+2)}{\Gamma(m_j+1)\Gamma(j-m_j+1)}$$

$$= \frac{(m_j+1)!(j+1)!}{(j+2)!(m_j)!}$$

$$= \frac{m_j+1}{j+2}.$$

Thus,

$$E_j\left[U_{j+1}(x)\right] = (j+1) + \frac{k}{(j+3)^2(j+4)} E_j\left[(1+m_j+x_{j+1})(j+2-m_j-x_{j+1})\right]$$

(2.4.2)

$$= (j+1) + \frac{k}{(j+3)^2(j+4)}\Big[(1+m_j)(j+2-m_j)-(1+m_j)E_j(x_{j+1})$$

$$+(j+2-m_j)E_j(x_{j+1})-E_j(x_{j+1}^2)\Big]$$

$$= (j+1) + \frac{k}{(j+3)^2(j+4)}\Big[(1+m_j)(j+2-m_j)-(1+m_j)\frac{m_j+1}{j+2} +$$

$$(j+2-m_j)\frac{m_j+1}{j+2} - \frac{m_j+1}{j+2}\Big]$$

$$= (j+1) + \frac{k(1+m_j)}{(j+3)^2(j+4)(j+2)}\left[(j+2)(j+2-m_j)-(1+m_j)+(j+2-m_j)-1\right]$$

$$= (j+1) + \frac{k(1+m_j)}{(j+2)(j+3)^2(j+4)}\,(j^2+5j-jm_j-4m_j+4)$$

$$= (j+1) + \frac{k(1+m_j)(j+1-m_j)}{(j+2)(j+3)^2} .$$

Again and again, we shall apply formulae (2.4.1) and (2.4.2) in our calculation. It should be pointed out that if $U_i(x) > \alpha_i(x)$, then $U_{i-1}(x) > \alpha_{i-1}(x)$. Indeed, the ith observation is an additional observation of the first (i-1) observations and hence the continuity of the point $(i-m_i,\ m_i)$ implies the continuity of $(i-1-m_{i-1},\ m_{i-1})$.

Once a value of k is fixed, we may construct its corresponding optimal sampling plan S*. Suppose k = 400, we have

$$S_8^* = \left\{ x:\ U_i(x) > \alpha_i(x) \text{ for } i < 8,\ U_8(x) = \alpha_8(x) \right\}$$
$$\alpha_7(x) = \min\left[U_7(x),\ E_7(\alpha_8(x))\right]$$

m_7	0	1	2	3	4	5	6	7
U_7	10.95	13.92	15.89	16.880	16.880	15.89	13.92	10.95
$E_7(\alpha_8)$	11.55	14.22	16.00	16.888	16.888	16.00	14.22	11.55

Thus S_8^* and hence S_n^* for $n > 8$ are empty.

$S_7^* = \{ x: U_i(x) > \alpha_i(x) \text{ for } i < 7,\ U_7(x) = \alpha_7(x) \}$

$\alpha_6(x) = \min \left[U_6(x),\ E_6(\alpha_7(x)) \right]$

m_6	0	1	2	3	4	5	6
U_6	10.87	14.34	16.43	17.12	16.43	14.34	10.87
$E_6(\alpha_7)$	11.33	14.42	16.27	16.89	16.27	14.42	11.33

Thus $S_7^* = \{ x: m_6 = 2, 3, 4 \}$

$S_6^* = \{ x: U_i(x) > \alpha_i(x) \text{ for } i < 6,\ U_6(x) = \alpha_6(x) \}$

$\alpha_5(x) = \min \left[U_5(x),\ E_5(\alpha_6(x)) \right]$

m_5	0	1	2	3	4	5
U_5	11.13	15.21	17.26	17.26	15.21	11.13
$E_5(\alpha_6)$	11.36	14.93	16.72	16.72	14.93	11.36

Thus $S_6^* = \{ x: m_5 = 1, 4,\ m_6 = 1, 5 \}$

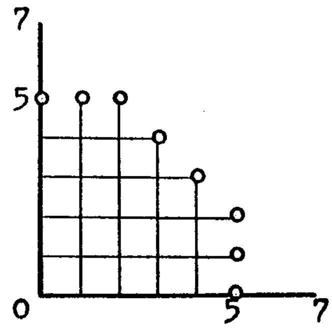
$S_5^* = \{ x: U_i(x) > \alpha_i(x) \text{ for } i < 5,\ U_5(x) = \alpha_5(x) \}$

$\alpha_4(x) = \min \left[U_4(x),\ E_4(\alpha_5(x)) \right]$

m_4	0	1	2	3	4
U_4	11.94	16.70	18.28	16.70	11.94
$E_4(\alpha_5)$	11.80	15.88	17.24	15.88	11.80

Thus $S_5^* = \{ x: m_5 = 0, 5 \}$ and $S_n^* = \emptyset$ for $n = 0, 1, 2, 3, 4$.

The partition $S^* = (S_5^*, S_6^*, S_7^*)$ determines an optimal sequential binomial sampling plan which can be illustrated graphically as follows: (dots stand for boundary points).



We are interested to know the relation between the values of k and the sizes of the plans. In other words, we shall try to find out the restriction on the values of k such that the optimal sampling plan so constructed is of certain fixed size. We have claimed before that if $U_j(x) = \alpha_j(x)$ we stop sampling while if $U_j(x) > \alpha_j(x)$ we take another observation. Hence $x = (j-m_j, m_j)$ is a boundary point or an inaccessible point if

$$U_j(x) - E_j\left[\alpha_{j+1}(x)\right] \leq 0.$$

Observe that if we stop sampling at the jth stage then the (j+1)th stage is inaccessible. Thus

$$U_j(x) = \alpha_j(x) \Longrightarrow U_{j+1}(x) = \alpha_{j+1}(x).$$

It follows that $(j-m_j, m_j)$ is a boundary or an inaccessible point if

$$U_j(x) - E_j\left[U_{j+1}(x)\right] \leq 0.$$

In virtual of formulae (2.4.1) and (2.4.2) we reduce the above condition to

$$k \leq \frac{(j+2)^2(j+3)^2}{(m_j+1)(j+1-m_j)} .$$

Set

$$D(j, m_j) = \frac{(j+2)^2(j+3)^2}{(m_j+1)(j+1-m_j)} .$$

Since $D(j, m_j) = D(j, j-m_j)$, we see that this kind of sampling plans are symmetric. A symmetric sampling plan is defined as one in which the boundary points are symmetric about the line $y = x$. Thus we eventually arrived the following result 2.4.1.

<u>Result 2.4.1</u>: If $k \leq D(j, m_j)$, then $(j-m_j, m_j)$ and $(m_j, j-m_j)$ are boundary or inaccessible points, $m_j = 0, 1, 2, \ldots, [j/2]$.

Table A in the Appendix gives the values of $D(j, m_j)$ for $j = 0, 1, \ldots, 50$ and $m_j = 0, 1, \ldots, [j/2]$. The reason zero is included in the values of j is to permit making a decision without taking any observation.

Since a sampling plan is completely determined by its boundary, Table A in the appendix as well as result 2.4.1 enable us to figure out the Table B in the appendix which gives a series of values of k showing how these values affect the sizes of the optimal sampling plans so constructed. For simplicity, we only figured out such plans of size 1 up to 16. We shall next investigate some properties of such plans in the rest of this section.

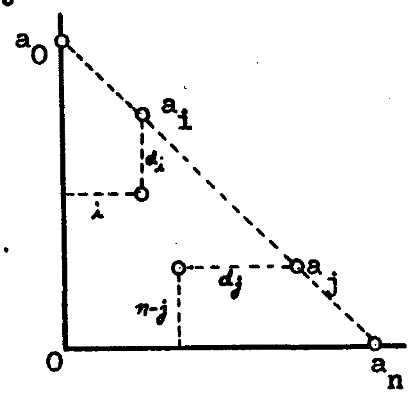
Let $a_i = (i, n-i)$, $i = 0, 1, \ldots, n$ be the n+1 points of index n of a sampling plan of size n. Then a_i is either a boundary or an inaccessible point. Now consider the following (n+1)-vectors

$$(d_0, \ldots, d_{\frac{n-1}{2}}, d_{\frac{n+1}{2}}, \ldots, d_n) \text{ when n is odd,}$$

and

$$(d_0, \ldots, d_{\frac{n}{2}-1}, d_{\frac{n}{2}}, d_{\frac{n}{2}+1}, \ldots, d_n) \text{ when n is even.}$$

For n being even (respective odd) let d_i, $i = 0, 1, \ldots, \frac{n}{2}-1$ (respective $i = 0, 1, \ldots, \frac{n-1}{2}$) be the least "distance" between a_i and a boundary point Υ with $X(\Upsilon) = i$, while d_j, $j = \frac{n}{2} + 1, \ldots, n$ (respective $j = \frac{n+1}{2}, \ldots, n$) be the least "distance" between a_j and a boundary point Υ with $Y(\Upsilon) = n-j$.



Note that when n is even, $X(a_{\frac{n}{2}}) = Y(a_{\frac{n}{2}}) = \frac{n}{2}$, it is immaterial which direction distance $d_{\frac{n}{2}}$ is measured. For example, with respect to sampling plan (39) (resp. (34)) we have the vector (6, 4, 3, 2, 1, 1, 0, 0, 0, 1, 1, 2, 3, 4, 6) (resp. (6, 4, 2, 2, 1, 0, 0, 0, 0, 1, 2, 2, 4, 6)).

For convenience, we call such a (n+1)-vector a B_n-vector in which the subscript n denote the size of the corresponding optimal binomial sampling plans. From now onwards, we shall simply write "a sampling plan" instead of "an optimal binomial sampling plan so constructed" and our B_n-vectors that follow will refer to those optimal plans. We have shown that these sampling plans in consideration are symmetric.

Proposition 1. If n is even, the $(\frac{n}{2} + 1)$th component of a B_n-vector is zero.

Proof. Suppose the contrary. Then $(n/2, n/2)$ is an inaccessible point and it follows that $(\frac{n}{2} -1, n/2)$ and $(n/2, \frac{n}{2} -1)$ are non-continuation points. Thus, by Result 2.4.1,

$$D(n-1, \tfrac{n}{2}) = D(n-1, \tfrac{n}{2}-1) \geqslant k.$$

But

$$D(n-1, \tfrac{n}{2}) \leqslant D(n-1, t) \text{ for all } t = 0, 1, \ldots, n-1.$$

Hence

$$D(n-1, t) \geqslant k \text{ for all } t = 0, 1, \ldots, n-1.$$

Therefore the size of the plan is less than n, which is a contradiction.

Proposition 2. If n is even (resp. odd), the components $d_{\frac{n}{2}-1}, d_{\frac{n}{2}}, d_{\frac{n}{2}+1}$ (resp. $d_{\frac{n-1}{2}}, d_{\frac{n+1}{2}}$) of each B_n-vector are zero.

Proof. When n is even, by Proposition 1, $(n/2, n/2)$ is a boundary point. If $(\frac{n}{2}-1, \frac{n}{2}+1)$ and $(\frac{n}{2}+1, \frac{n}{2}-1)$ were inaccessible points, it follows that $(\frac{n}{2}-1, \frac{n}{2})$ and $(\frac{n}{2}, \frac{n}{2}-1)$ are non-continuation points and in turn $(\frac{n}{2}, \frac{n}{2})$ is inaccessible. This is a contradiction. Thus $(\frac{n}{2}-1, \frac{n}{2}+1)$ and $(\frac{n}{2}+1, \frac{n}{2}-1)$ are boundary points and hence $d_{\frac{n}{2}-1} =$

$d_{\frac{n}{2}+1} = d_{\frac{n}{2}} = 0.$

When n is odd, we shall show that $(\frac{n-1}{2}, \frac{n+1}{2})$ and $(\frac{n+1}{2}, \frac{n-1}{2})$ are boundary points. Indeed, if they were inaccessible, then $(\frac{n-1}{2}, \frac{n-1}{2})$ is a non-continuation point and, by Result 2.4.1,

$$D(n-1, \tfrac{n-1}{2}) \geqslant k.$$

But

$$D(n-1, t) \geqslant D(n-1, \tfrac{n-1}{2}) \text{ for } t = 0, 1, \ldots, \tfrac{n-1}{2}.$$

So $D(n-1, t) \geqslant k$ for $t = 0, 1, \ldots, \frac{n-1}{2}$. It follows that the size of the plan is less than n which is a contradiction. Thus $(\frac{n-1}{2}, \frac{n+1}{2})$ and $(\frac{n+1}{2}, \frac{n-1}{2})$ are boundary points and hence $d_{\frac{n-1}{2}} = d_{\frac{n+1}{2}} = 0$.

Proposition 3. The components of a B_n-vector are non-decreasing from the middle to both ends.

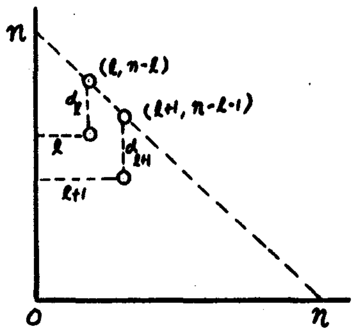
Proof. First let n be even and let

$$(d_0, \ldots, d_{\frac{n}{2}-2}, 0, 0, 0, d_{\frac{n}{2}+2}, \ldots, d_n)$$

be a B_n-vector. By symmetry, it is enough to prove that

$$d_i \geqslant d_{i+1} \text{ for } i = 0, \ldots, \tfrac{n}{2}-2.$$

Suppose that there exists an ℓ, $0 \leq \ell \leq \frac{n}{2}-2$, such that $d_\ell < d_{\ell+1}$.



The boundary point of $(\ell, n-\ell-d_\ell)$ implies that either $(\ell-1, n-\ell-d_\ell)$ or $(\ell, n-\ell-d_\ell-1)$ is a continuation point. If $(\ell-1, n-\ell-d_\ell)$ is a continuation point, then

$$k > D(n-d_\ell-1, \ell-1).$$

But

$$D(n-d_\ell-1, \ell-1) > D(n-d_\ell-1, \ell).$$

It follows that $k > D(n-d_\ell-1, \ell)$, i.e. $(\ell, n-\ell-d_\ell-1)$ is a continuation point. Hence in either case $(\ell, n-\ell-d_\ell-1)$ must be a continuation point. Repeating this kind of argument we finally show that $(\ell, n-\ell-d_{\ell+1})$ is a continuation point, i.e.

$$k > D(n-d_{\ell+1}, \ell) > D(n-d_{\ell+1}, \ell+1).$$

However, the boundary point of $(\ell+1, n-\ell-1-d_{\ell+1})$ demands that

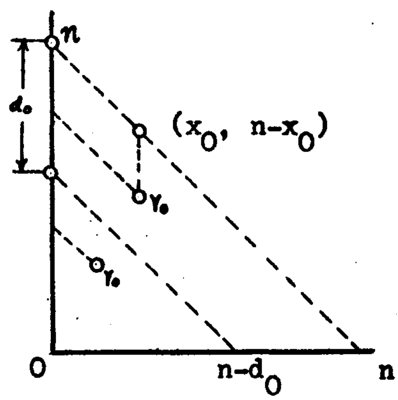
$k \leq D(n-d_{\ell+1}, \ell+1)$.

This is a contradiction.

Similarly, we can prove for odd integers n.

Proposition 4. There are no boundary points other than the n+1 boundary points defining the B_n-vector of a sampling plan of size n.

Proof. Let $\Upsilon_0 = (x_0, y_0)$ be an arbitrary boundary point. Without loss of generality, we may take $X(\Upsilon_0) < Y(\Upsilon_0)$.



If Υ_0 is of index less than $n-d_0$, then $(0, x_0 + y_0)$ is a non-continuation point. This follows from Result 2.4.1 and the inequalities

$$D(x_0 + y_0 , x_0) \geq k$$

and

$$D(x_0 + y_0, 0) \geq D(x_0 + y_0, x_0).$$

This is impossible for $(0, n-d_0)$ is a boundary point. Thus $N(\Upsilon_0) \geq n-d_0$. Since Υ_0 is a boundary point,

$$D(x_0 + y_0, x_0) \geq k.$$

Thus

$$D(x_0 + y_0, t) \geq D(x_0 + y_0, x_0) \geq k, \text{for } t = 0, 1, \ldots, x_0$$

This shows that $(x_0 - i, y_0 + i)$, $i = 1, \ldots, x_0$, are non-continuation points and hence $(x_0, y_0 + 1)$ is inaccessible. By similar reasoning, $(x_0, y_0 + j)$, $j = 1, \ldots, n - x_0 - y_0$, are inaccessible. It follows that the distance between Υ_0 and $(x_0, n - x_0)$ defines the $(x_0 + 1)$th component of the B_n-vector. Hence Υ_0 is a boundary point defining the B_n-vector of the sampling plan.

Propositions 3, 4 and Theorem 1.5.1 together imply the following:

Proposition 5. The optimal sampling plans so constructed are simple.

Let S_1 (resp. S_2) be a sampling plan with B_n-vector $(a_0, \ldots, a_n)$ (resp. B_m-vector $(b_0, \ldots, b_m)$). Then we say that S_1 is larger than S_2 if $n > m$. When $n = m$, S_1 is larger than S_2 if $a_i \leq b_i$ for all $i = 0, 1, \ldots, n$ with inequality holding for at least one value of i.

Consider the inequalities

(a) $D(n-1, \frac{n}{2}) < k \leq D(n, \frac{n}{2})$

(b) $D(n, \frac{n}{2}) < k \leq D(n+1, \frac{n}{2})$

where n being any positive even integer. Inequality (a) says that the sampling plan corresponding to each value of k satisfying (a) is of size n since points of index n are either boundary or inaccessible points and at the same time there are continuation points of index n-1. Similarly, inequality (b) determines sampling plans of size n+1.

For positive odd integers n, we have the following corresponding inequalities

(c) $D(n-1, \frac{n-1}{2}) < k \leq D(n, \frac{n-1}{2})$

(d) $D(n, \frac{n-1}{2}) < k \leq D(n+1, \frac{n+1}{2})$.

Between $D(n-1, \frac{n}{2})$ and $D(n, \frac{n}{2})$ there may exist certain values $D(m, j)$, $m \leq n-1$, $j \leq [m/2]$. If there exist $t+1$ such values, arrange them in order such that

$$D(n-1, \tfrac{n}{2}) < D(m_0, j_0) < D(m_1, j_1) < \dots < D(m_t, j_t) < D(n, \tfrac{n}{2}).$$

We see that every real k in the half open interval $I_0 = (D(n-1, n/2), D(m_0, j_0)]$ corresponds to a <u>fixed</u> sampling plan S_0 of size n (even) with B_n-vector say,

$$v_0 = (d_{0,0}, \dots, d_{0,j_0}, \dots, d_{0,\frac{n}{2}-2}, 0, 0, 0, d_{0,\frac{n}{2}+2}, \dots, d_{0,n-j_0}, \dots, d_{0,n}.)$$

As k increases and falls into $I_1 = (D(m_0, j_0), D(m_1, j_1)]$, again, every real k in I_1 defines another fixed sampling plan S_1 of size n with B_n-vector, say,

$$v_1 = (d_{1,0}, \dots, d_{1,j_0}, \dots, d_{1,\frac{n}{2}-2}, 0, 0, 0, d_{1,\frac{n}{2}+2}, \dots, d_{1,n-j_0}, \dots, d_{1,n}).$$

Observe that as k varies from I_0 to I_1, it turns only the non-continuation points (j_0, m_0-j_0) and (m_0-j_0, j_0) of S_0 to continuation points of S_1. Thus v_0 differs from v_1 only in the (j_0+1)th and $(n-j_0+1)$th components. By symmetry, we have

$$d_{0,j_0} = d_{0,n-j_0} \geqslant (n-j_0)-(m_0-j_0) = n-m_0.$$

However, (j_0, m_0-j_0) is a continuation point of S_1 which implies that

$$d_{1,j_0} = d_{1,n-j_0} < n-m_0.$$

Thus

$$d_{1,j_0} < d_{0,j_0}.$$

Hence S_1 is larger than S_0. Repeating this argument we shall eventually show that as k increases from D(n-1, n/2) to D(n, n/2) its corresponding sampling plan though preserves size n becomes larger and larger. If k goes beyond D(n, n/2), then its corresponding sampling plan is of size larger than n.

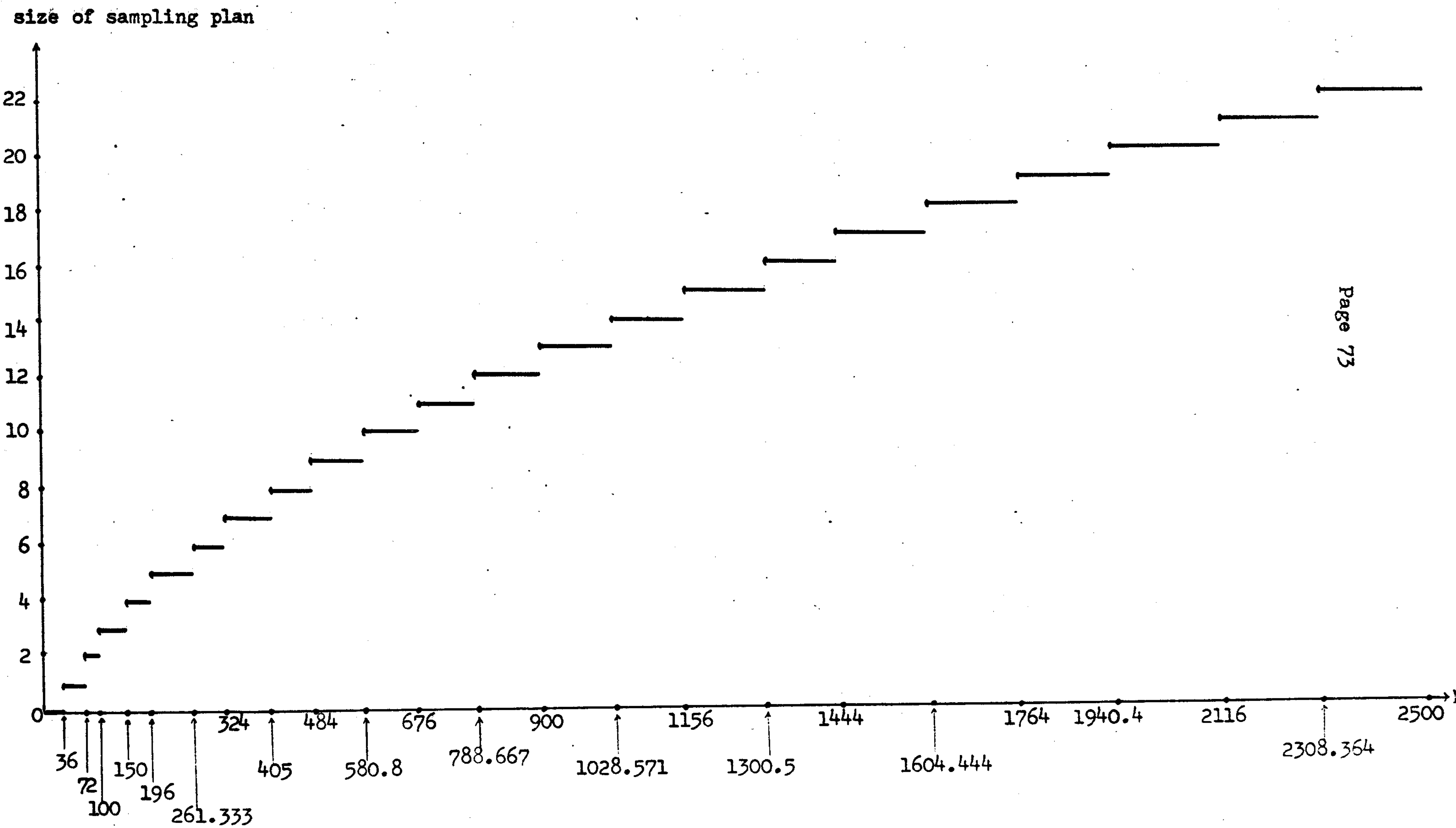
Similarly, we can show the previous result for odd integers n.

Hence we have proved the following:

<u>Proposition 6</u>. The increasing values of k <u>enlarge</u> the optimal sampling plans so constructed.

In fact, we have seen that the size of the optimal sampling plan does not increase strictly with the increasing k.

The following figure illustrates the effect of k on the size of its corresponding sampling plan when k goes from 0 to 2500.


size of sampling plan
22
20
18
16
14
12
10
8
6
4
2
0
k
36
72
100
150
196
261.333
324
405
484
580.8
676
788.667
900
1028.571
1156
1300.5
1444
1604.444
1764
1940.4
2116
2308.364
2500

Appendix

TABLE A

Values of $D(j, m_j)$, $j = 0, 1, \ldots, 50$, $m_j = 0, 1, \ldots, [j/2]$.

D(0,0) = 36	D(9,0) = 1742.4
	D(9,1) = 968
D(1,0) = 72	D(9,2) = 726
	D(9,3) = 622.286
D(2,0) = 133.333	D(9,4) = 580.8
D(2,1) = 100	
	D(10,0) = 2212.364
D(3,0) = 225	D(10,1) = 1216.8
D(3,1) = 150	D(10,2) = 901.333
	D(10,3) = 760.5
D(4,0) = 352.8	D(10,4) = 695.314
D(4,1) = 220.5	D(10,5) = 676
D(4,2) = 196	
	D(11,0) = 2760.333
D(5,0) = 522.667	D(11,1) = 1505.636
D(5,1) = 313.6	D(11,2) = 1104.133
D(5,2) = 261.333	D(11,3) = 920.111
	D(11,4) = 828.1
D(6,0) = 740.571	D(11,5) = 788.667
D(6,1) = 432	
D(6,2) = 345.6	D(12,0) = 3392.308
D(6,3) = 324	D(12,1) = 1837.5
	D(12,2) = 1336.364
D(7,0) = 1012.5	D(12,3) = 1102.5
D(7,1) = 578.571	D(12,4) = 980
D(7,2) = 450	D(12,5) = 918.75
D(7,3) = 405	D(12,6) = 900
D(8,0) = 1344.444	D(13,0) = 4114.286
D(8,1) = 756.25	D(13,1) = 2215.385
D(8,2) = 576.191	D(13,2) = 1600
D(8,3) = 504.167	D(13,3) = 1309.091
D(8,4) = 484	D(13,4) = 1152
	D(13,5) = 1066.667
	D(13,6) = 1028.571

D(14,0) = 4932.267
D(14,1) = 2642.286
D(14,2) = 1897.026
D(14,3) = 1541.333
D(14,4) = 1345.164
D(14,5) = 1233.067
D(14,6) = 1174.349
D(14,7) = 1156

D(15,0) = 5852.25
D(15,1) = 3121.2
D(15,2) = 2229.429
D(15,3) = 1800.692
D(15,4) = 1560.6
D(15,5) = 1418.727
D(15,6) = 1337.657
D(15,7) = 1300.5

D(16,0) = 6880.235
D(16,1) = 3655.125
D(16,2) = 2599.2
D(16,3) = 2088.643
D(16,4) = 1799.446
D(16,5) = 1624.5
D(16,6) = 1519.013
D(16,7) = 1462.05
D(16,8) = 1444

D(17,0) = 8022.222
D(17,1) = 4247.059
D(17,2) = 3008.333
D(17,3) = 2406.667
D(17,4) = 2062.857
D(17,5) = 1815.282
D(17,6) = 1719.048
D(17,7) = 1640.909
D(17,8) = 1604.444

D(18,0) = 9284.21
D(18,1) = 4900
D(18,2) = 3458.824
D(18,3) = 2756.25
D(18,4) = 2352
D(18,5) = 2100
D(18,6) = 1938.462
D(18,7) = 1837.5
D(18,8) = 1781.818
D(18,9) = 1764

D(19,0) = 10672.20
D(19,1) = 5616.947
D(19,2) = 3952.667
D(19,3) = 3138.882
D(19,4) = 2668.05
D(19,5) = 2371.6
D(19,6) = 2178
D(19,7) = 2052.346
D(19,8) = 1976.333
D(19,9) = 1940.4

D(20,0) = 12192.19
D(20,1) = 6400.9
D(20,2) = 4491.86
D(20,3) = 3556.056
D(20,4) = 3012.188
D(20,5) = 2667.042
D(20,6) = 2438.438
D(20,7) = 2286.036
D(20,8) = 2188.342
D(20,9) = 2133.633
D(20,10) = 2116

D(21,0) = 13850.18
D(21,1) = 7254.857
D(21,2) = 5078.4
D(21,3) = 4009.263
D(21,4) = 3385.6
D(21,5) = 2987.294
D(21,6) = 2720.571
D(21,7) = 2539,2
D(21,8) = 2418.286
D(21,9) = 2343.877
D(21,10) = 2308.364

D(22,0) = 15652.17
D(22,1) = 8181.818
D(22,2) = 5714.286
D(22,3) = 4500
D(22,4) = 3789.474
D(22,5) = 3333.333
D(22,6) = 3025.21
D(22,7) = 2812.5
D(22,8) = 2666.667
D(22,9) = 2571.429
D(22,10) = 2517.483
D(22,11) = 2500

D(23,0) = 17604.17
D(23,1) = 9184.783
D(23,2) = 6401.515
D(23,3) = 5029.762
D(23,4) = 4225
D(23,5) = 3706.14
D(23,6) = 3353.175
D(23,7) = 3106.618
D(23,8) = 2934.028
D(23,9) = 2816.667
D(23,10) = 2743.507
D(23,11) = 2708.333

D(24,0) = 19712.16
D(24,1) = 10266.75
D(24,2) = 7142.087
D(24,3) = 5600.045
D(24,4) = 4693.371
D(24,5) = 4106.7
D(24,6) = 3705.293
D(24,7) = 3422.25
D(24,8) = 3220.941
D(24,9) = 3080.025
D(24,10) = 2986.691
D(24,11) = 2933.357
D(24,12) = 2916

D(25,0) = 21982.15
D(25,1) = 11430.72
D(25,2) = 7938
D(25,3) = 6212.348
D(25,4) = 5195.782
D(25,5) = 4536
D(25,6) = 4082.4
D(25,7) = 3760.105
D(25,8) = 3528
D(25,9) = 3361.976
D(25,10) =3247.364
D(25,11) =3175.2
D(25,12) =3140.308

D(26,0) = 24420.15
D(26,1) = 12679.69
D(26,2) = 8791.253
D(26,3) = 6868.167
D(26,4) = 5733.426
D(26,5) = 4995.03
D(26,6) = 4485.333
D(26,7) = 4120.9
D(26,8) = 3855.813
D(26,9) = 3663.022
D(26,10) = 3525.904
D(26,11) = 3434.083
D(26,12) = 3381.251
D(26,13) = 3364

D(27,0) = 27032.14
D(27,1) = 14016.67
D(27,2) = 9703.846
D(27,3) = 7569
D(27,4) = 6307.5
D(27,5) = 5484.783
D(27,6) = 4914.935
D(27,7) = 4505.357
D(27,8) = 4205
D(27,9) = 3983.684
D(27,10) =3822.727
D(27,11) =3710.294
D(27,12) =3638.942
D(27,13) =3604.286

D(28,0) = 29824.14
D(28,1) = 15444.64
D(28,2) = 10677.78
D(28,3) = 8316.346
D(28,4) = 6919.2
D(28,5) = 6006.25
D(28,6) = 5372.05
D(28,7) = 4914.205
D(28,8) = 4576.19
D(28,9) = 4324.5
D(28,10) = 4138.277
D(28,11) = 4004.167
D(28,12) = 3913.575
D(28,13) = 3861.161
D(28,14) = 3844

D(29,0) = 32802.13
D(29,1) = 16966.62
D(29,2) = 11715.05
D(29,3) = 9111.704
D(29,4) = 7569.723
D(29,5) = 6560.427
D(29,6) = 5857.524
D(29,7) = 5348.174
D(29,8) = 4970.02
D(29,9) = 4686.019
D(29,10) = 4473.018
D(29,11) = 4316.07
D(29,12) = 4205.402
D(29,13) = 4134.723
D(29,14) = 4100.267

D(30,0) = 35972.13
D(30,1) = 18585.60
D(30,2) = 12817.66
D(30,3) = 9956.571
D(30,4) = 8260.267
D(30,5) = 7148.308
D(30,6) = 6372.206
D(30,7) = 5808
D(30,8) = 5387.13
D(30,9) = 5068.8
D(30,10)= 4827.429
D(30,11)= 4646.4
D(30,12)= 4514.721
D(30,13)= 4425.143
D(30,14)= 4373.082
D(30,15)= 4356

D(31,0) = 39340.13
D(31,1) = 20304.58
D(31,2) = 13987.60
D(31,3) = 10852.45
D(31,4) = 8992.029
D(31,5) = 7770.889
D(31,6) = 6916.945
D(31,7) = 6294.42
D(31,8) = 5828.167
D(31,9) = 5473.409
D(31,10) = 5202
D(31,11) = 4995.571
D(31,12) = 4841.862
D(31,13) = 4732.647
D(31,14) = 4662.533
D(31,15) = 4628.25

D(32,0) = 42912.12
D(32,1) = 22126.56
D(32,2) = 15226.88
D(32,3) = 11800.83
D(32,4) = 9766.207
D(32,5) = 8429.167
D(32,6) = 7492.593
D(32,7) = 6808.173
D(32,8) = 6293.778
D(32,9) = 5900.417
D(32,10)= 5597.233
D(32,11)= 5364.015
D(32,12)= 5187.179
D(32,13)= 5057.5
D(32,14)= 4968.772
D(32,15)= 4917.014
D(32,16)= 4900

D(33,0) = 46694.12
D(33,1) = 24054.55
D(33,2) = 16537.50
D(33,3) = 12803.23
D(33,4) = 10584
D(33,5) = 9124.138
D(33,6) = 8100
D(33,7) = 7350
D(33,8) = 6784.615
D(33,9) = 6350.4
D(33,10)= 6013.636
D(33,11)= 5752.174
D(33,12)= 5551.049
D(33,13)= 5400
D(33,14)= 5292
D(33,15)= 5222.368
D(33,16)= 5188.235

D(34,0) = 50692.11
D(34,1) = 26091.53
D(34,2) = 1792.45
D(34,3) = 13861.12
D(34,4) = 11446.61
D(34,5) = 9856.8
D(34,6) = 8740.02
D(34,7) = 7920.643
D(34,8) = 7301.333
D(34,9) = 6823.938
D(34,10)= 6451.724
D(34,11)= 6160.5
D(34,12)= 5933.859
D(34,13)= 5760.467
D(34,14)= 5632.457
D(34,15)= 5544.45
D(34,16)= 5492.954
D(34,17)= 5476

D(35,0) = 54912.11
D(35,1) = 28240.51
D(35,2) = 19380.75
D(35,3) = 14976.03
D(35,4) = 12355.22
D(35,5) = 10628.15
D(35,6) = 9413.505
D(35,7) = 8520.845
D(35,8) = 7844.587
D(35,9) = 7321.615
D(35,10)= 6912.014
D(35,11)= 6589.453
D(35,12)= 6336.013
D(35,13)= 6139.242
D(35,14)= 5990.412
D(35,15)= 5883.44
D(35,16)= 5814.224
D(35,17)= 5780.222

D(36,0) = 59360.11
D(36,1) = 30504.50
D(36,2) = 20917.37
D(36,3) = 16149.44
D(36,4) = 13311.05
D(36,5) = 11439.19
D(36,6) = 10121.31
D(36,7) = 9151.35
D(36,8) = 8415.034
D(36,9) = 7844.014
D(36,10)= 7395.03
D(36,11)= 7039.5
D(36,12)= 6757.92
D(36,13)= 6536.679
D(36,14)= 6366.156
D(36,15)= 6239.557
D(36,16)= 6152.168
D(36,17)= 6100.9
D(36,18)= 6084

D(37,0) = 64042.10
D(37,1) = 32886.49
D(37,2) = 22533.33
D(37,3) = 17382.86
D(37,4) = 14315.29
D(37,5) = 12290.91
D(37,6) = 10864.29
D(37,7) = 9812.903
D(37,8) = 9013.333
D(37,9) = 8391.724
D(37,10)= 7901.299
D(37,11)= 7511.111
D(37,12)= 7200
D(37,13)= 6953.143
D(37,14)= 6760
D(37,15)= 6613.043
D(37,16)= 6506.952
D(37,17)= 6438.095
D(37,18)= 6404.21

D(38,0) = 68964.10
D(38,1) = 35389.47
D(38,2) = 24230.63
D(38,3) = 18677.78
D(38,4) = 15369.14
D(38,5) = 13184.31
D(38,6) = 11643.29
D(38,7) = 10506.25
D(38,8) = 9640.143
D(38,9) = 8965.333
D(38,10)= 8431.348
D(38,11)= 8004.762
D(38,12)= 7662.678
D(38,13)= 7389.011
D(38,14)= 7172.267
D(38,15)= 7004.167
D(38,16)= 6878.772
D(38,17)= 6791.919
D(38,18)= 6740.852
D(38,19)= 6724

D(39,0) = 74132.10
D(39,1) = 38016.46
D(39,2) = 26011.26
D(39,3) = 20035.70
D(39,4) = 16473.80
D(39,5) = 14120.40
D(39,6) = 12459.18
D(39,7) = 11232.14
D(39,8) = 10296.12
D(39,9) = 9565.432
D(39,10)= 8985.709
D(39,11)= 8520.931
D(39,12)= 8146.385
D(39,13)= 7844.667
D(39,14)= 7603.292
D(39,15)= 7413.21
D(39,16)= 7267.853
D(39,17)= 7162.522
D(39,18)= 7093.981
D(39,19)= 7060.2

D(40,0) = 79552.10
D(40,1) = 40770.45
D(40,2) = 27877.23
D(40,3) = 21458.13
D(40,4) = 17630.47
D(40,5) = 15100.17
D(40,6) = 13312.80
D(40,7) = 11991.31
D(40,8) = 10981.94
D(40,9) = 10192.61
D(40,10)= 9564.915
D(40,11)= 9060.1
D(40,12)= 8651.554
D(40,13)= 8320.5
D(40,14)= 8053.422
D(40,15)= 7840.471
D(40,16)= 7674.438
D(40,17)= 7550.083
D(40,18)= 7463.698
D(40,19)= 7412.809
D(40,20)= 7396

D(41,0) = 85230.09
D(41,1) = 43654.44
D(41,2) = 29830.53
D(41,3) = 22946.56
D(41,4) = 18840.34
D(41,5) = 16124.61
D(41,6) = 14205.02
D(41,7) = 12784.51
D(41,8) = 11698.25
D(41,9) = 10847.47
D(41,10)= 10169.50
D(41,11)= 9622.753
D(41,12)= 9178.626
D(41,13)= 8816.906
D(41,14)= 8523.009
D(41,15)= 8286.259
D(41,16)= 8098.787
D(41,17)= 7954.809
D(41,18)= 7850.14
D(41,19)= 7781.878
D(41,20)= 7748.19

D(42,0) = 91172.09
D(42,1) = 46671.43
D(42,2) = 31873.17
D(42,3) = 24502.50
D(42,4) = 20104.62
D(42,5) = 17194.74
D(42,6) = 15136.68
D(42,7) = 13612.50
D(42,8) = 12445.71
D(42,9) = 11530.59
D(42,10)= 10800
D(42,11)= 10209.38
D(42,12)= 9728.04
D(42,13)= 9334.286
D(42,14)= 9012.414
D(42,15)= 8750.893
D(42,16)= 8541.176
D(42,17)= 8376.923
D(42,18)= 8253.474
D(42,19)= 8167.5
D(42,20)= 8116.77
D(42,21)= 8100

D(43,0) = 97384.09
D(43,1) = 49824.42
D(43,2) = 34007.14
D(43,3) = 26127.44
D(43,4) = 21424.50
D(43,5) = 18311.54
D(43,6) = 16108.65
D(43,7) = 14476.01
D(43,8) = 13225
D(43,9) = 12242.57
D(43,10)= 11456.95
D(43,11)= 10820.45
D(43,12)= 10300.24
D(43,13)= 9873.041
D(43,14)= 9522
D(43,15)= 9234.698
D(43,16)= 9001.891
D(43,17)= 8816.667
D(43,18)= 8673.887
D(43,19)= 8569.8
D(43,20)= 8501.786
D(43,21)= 8468.182

D(44,0) = 103872.1
D(44,1) = 53116.41
D(44,2) = 36234.45
D(44,3) = 27822.88
D(44,4) = 22801.19
D(44,5) = 19476.02
D(44,6) = 17121.77
D(44,7) = 15375.80
D(44,8) = 14036.77
D(44,9) = 12984.01
D(44,10)= 12140.89
D(44,11)= 11456.48
D(44,12)= 10895.67
D(44,13)= 10433.58
D(44,14)= 10052.14
D(44,15)= 9738.008
D(44,16)= 9481.225
D(44,17)= 9274.294
D(44,18)= 9111.587
D(44,19)= 8988.931
D(44,20)= 8903.322
D(44,21)= 8852.735
D(44,22)= 8836

D(45,0) = 110642.1
D(45,1) = 56550.40
D(45,2) = 38557.09
D(45,3) = 29590.33
D(45,4) = 24235.89
D(45,5) = 20689.17
D(45,6) = 18176.91
D(45,7) = 16312.62
D(45,8) = 14881.68
D(45,9) = 13755.50
D(45,10)= 12852.36
D(45,11)= 12117.94
D(45,12)= 11514.79
D(45,13)= 11016.31
D(45,14)= 10603.20
D(45,15)= 10261.16
D(45,16)= 9979.482
D(45,17)= 9750.069
D(45,18)= 9566.797
D(45,19)= 9425.067
D(45,20)= 9321.494
D(45,21)= 9253.702
D(45,22)= 9220.174

D(46,0) = 117700.1
D(46,1) = 60129.39
D(46,2) = 40977.07
D(46,3) = 31431.27
D(46,4) = 25729.79
D(46,5) = 21952
D(46,6) = 19274.93
D(46,7) = 17287.20
D(46,8) = 15760.41
D(46,9) = 14557.64
D(46,10)= 13591.90
D(46,11)= 12805.33
D(46,12)= 12158.03
D(46,13)= 11621.65
D(46,14)= 11175.56
D(46,15)= 10804.50
D(46,16)= 10496.97
D(46,17)= 10244.27
D(46,18)= 10039.75
D(46,19)= 9878.4
D(46,20)= 9756.444
D(46,21)= 9671.161
D(46,22)= 9620.702
D(46,23)= 9604

D(47,0) = 125052.1
D(47,1) = 63856.38
D(47,2) = 43496.38
D(47,3) = 33347.22
D(47,4) = 27284.09
D(47,5) = 23265.50
D(47,6) = 20416.67
D(47,7) = 18300.30
D(47,8) = 16673.61
D(47,9) = 15391.03
D(47,10)= 14360.05
D(47,11)= 13519.14
D(47,12)= 12825.85
D(47,13)= 12250
D(47,14)= 11769.61
D(47,15)= 11368.37
D(47,16)= 11034.01
D(47,17)= 10757.17
D(47,18)= 10530.70
D(47,19)= 10349.14
D(47,20)= 10208.33
D(47,21)= 10105.22
D(47,22)= 10037.63
D(47,23)= 10004.17

D(48,0) = 132704.1
D(48,1) = 67734.37
D(48,2) = 46117.02
D(48,3) = 35339.67
D(48,4) = 28900
D(48,5) = 24630.68
D(48,6) = 21602.99
D(48,7) = 19352.68
D(48,8) = 17621.95
D(48,9) = 16256.25
D(48,10)= 15157.34
D(48,11)= 14259.87
D(48,12)= 13518.71
D(48,13)= 12901.79
D(48,14)= 12385.71
D(48,15)= 11953.13
D(48,16)= 11590.91
D(48,17)= 11289.06
D(48,18)= 11039.90
D(48,19)= 10837.50
D(48,20)= 10677.34
D(48,21)= 10556.01
D(48,22)= 10471.01
D(48,23)= 10420.67
D(48,24)= 10404

D(49,0) = 140662.1
D(49,1) = 71766.37
D(49,2) = 48841
D(49,3) = 37410.13
D(49,4) = 30578.71
D(49,5) = 26048.53
D(49,6) = 22834.75
D(49,7) = 20445.07
D(49,8) = 18606.10
D(49,9) = 17153.91
D(49,10)= 15984.33
D(49,11)= 15028
D(49,12)= 14237.05
D(49,13)= 13577.42
D(49,14)= 13024.27
D(49,15)= 12559.11
D(49,16)= 12168
D(49,17)= 11840.24
D(49,18)= 11567.61
D(49,19)= 11343.72
D(49,20)= 11163.66
D(49,21)= 11023.67
D(49,22)= 10920.97
D(49,23)= 10853.56
D(49,24)= 10820.16

D(50,0) = 148932.1
D(50,1) = 75955.36
D(50,2) = 51670.31
D(50,3) = 39560.08
D(50,4) = 32321.43
D(50,5) = 27520.06
D(50,6) = 24112.81
D(50,7) = 21578.23
D(50,8) = 19626.71
D(50,9) = 18084.61
D(50,10)= 16841.54
D(50,11)= 15824.03
D(50,12)= 14981.33
D(50,13)= 14277.32
D(50,14)= 13685.65
D(50,15)= 13186.69
D(50,16)= 12765.61
D(50,17)= 12411.01
D(50,18)= 12114.09
D(50,19)= 11868.02
D(50,20)= 11667.49
D(50,21)= 11508.39
D(50,22)= 11387.61
D(50,23)= 11302.88
D(50,24)= 11252.65
D(50,25)= 11236

Appendix

TABLE B

Optimal Sampling Plans for $L(w,a) = k(w-a)^2$

Plan's Size	Values of k	Figures
0	$0 < k \leq 36$	
1	$36 < k \leq 72$	1
2	$72 < k \leq 100$	2
3	$100 < k \leq 133.333$	3
	$133.333 < k \leq 150$	4
4	$150 < k \leq 196$	5
5	$196 < k \leq 220.5$	6
	$220.5 < k \leq 225$	7
	$225 < k \leq 261.333$	8
6	$261.333 < k \leq 313.6$	9
	$313.6 < k \leq 324$	10
7	$324 < k \leq 345.6$	11
	$345.6 < k \leq 352.8$	12
	$352.8 < k \leq 405$	13
8	$405 < k \leq 432$	14
	$432 < k \leq 450$	15
	$450 < k \leq 484$	16
9	$484 < k \leq 504.167$	17
	$504.167 < k \leq 522.667$	18
	$522.667 < k \leq 576.191$	19
	$576.191 < k \leq 578.571$	20
	$578.571 < k \leq 580.8$	21
10	$580.8 < k \leq 622.286$	22
	$622.286 < k \leq 676$	23
11	$676 < k \leq 695.314$	24

	$695.314 < k \le 726$	25
	$726 < k \le 740.571$	26
	$740.571 < k \le 756.25$	27
	$756.25 < k \le 760.5$	28
	$760.5 < k \le 788.667$	29
12	$788.667 < k \le 828.1$	30
	$828.1 < k \le 900$	31
13	$900 < k \le 901.333$	32
	$901.333 < k \le 918.75$	33
	$918.75 < k \le 920.111$	34
	$920.111 < k \le 968$	35
	$968 < k \le 980$	36
	$980 < k \le 1012.5$	37
	$1012.5 < k \le 1028.571$	38
14	$1028.571 < k \le 1066.667$	39
	$1066.667 < k \le 1102.5$	40
	$1102.5 < k \le 1104.133$	41
	$1104.133 < k \le 1152$	42
	$1152 < k \le 1156$	43
15	$1156 < k \le 1174.349$	44
	$1174.349 < k \le 1216.8$	45
	$1216.8 < k \le 1233.067$	46
	$1233.067 < k \le 1300.5$	47
16	$1300.5 < k \le 1309.091$	48
	$1309.091 < k \le 1336.364$	49
	$1336.364 < k \le 1337.657$	50
	$1337.657 < k \le 1344.444$	51
	$1344.444 < k \le 1345.164$	52
	$1345.164 < k \le 1418.727$	53
	$1418.727 < k \le 1444$	54
17	$1444 < k \le 1462.05$	
	$1462.05 < k \le 1505.636$	
	$1505.636 < k \le 1519.013$	
	$1519.013 < k \le 1541.333$	
	$1541.333 < k \le 1560.6$	
	$1560.6 < k \le 1600$	
	$1600 < k \le 1604.444$	
18	$1604.444 < k \le 1624.5$	
	$1624.5 < k \le 1640.909$	
	$1640.909 < k \le 1719.048$	
	$1719.048 < k \le 1742.4$	
	$1742.4 < k \le 1764$	

19	$1764 < k \le 1781.818$
	$1781.818 < k \le 1799.446$
	$1799.446 < k \le 1800.692$
	$1800.692 < k \le 1837.5$
	$1837.5 < k \le 1851.282$
	$1851.282 < k \le 1897.026$
	$1897.026 < k \le 1938.462$
	$1938.462 < k \le 1940.4$
20	$1940.4 < k \le 1976.333$
	$1976.333 < k \le 2052.346$
	$2052.346 < k \le 2062.857$
	$2062.857 < k \le 2088.643$
	$2088.643 < k \le 2100$
	$2100 < k \le 2116$
21	$2116 < k \le 2133.633$
	$2133.633 < k \le 2178$
	$2178 < k \le 2188.342$
	$2188.342 < k \le 2212.364$
	$2212.364 < k \le 2215.385$
	$2215.385 < k \le 2229.429$
	$2229.429 < k \le 2286.036$
	$2286.036 < k \le 2308.364$
22	$2308.364 < k \le 2343.877$
	$2343.877 < k \le 2352$
	$2352 < k \le 2371.6$
	$2371.6 < k \le 2406.667$
	$2406.667 < k \le 2418.286$
	$2418.286 < k \le 2438.438$
	$2438.438 < k \le 2500$
23	$2500 < k \le 2517.483$
	$2517.483 < k \le 2539.2$
	$2539.2 < k \le 2571.429$
	$2571.429 < k \le 2599.2$
	$2599.2 < k \le 2642.286$
	$2642.286 < k \le 2666.667$
	$2666.667 < k \le 2667.042$
	$2667.042 < k \le 2668.05$
	$2668.05 < k \le 2708.333$
24	$2708.333 < k \le 2720.571$
	$2720.571 < k \le 2743.507$
	$2743.507 < k \le 2756.25$
	$2756.25 < k \le 2760.333$
	$2760.333 < k \le 2812.5$
	$2812.5 < k \le 2816.667$
	$2816.667 < k \le 2916$

25	2916 < k ≤ 2933.357
	2933.357 < k ≤ 2934.028
	2934.028 < k ≤ 2986.691
	2986.691 < k ≤ 2987.294
	2987.294 < k ≤ 3008.333
	3008.333 < k ≤ 3012.188
	3012.188 < k ≤ 3025.21
	3025.21 < k ≤ 3080.025
	3080.025 < k ≤ 3106.618
	3106.618 < k ≤ 3121.2
	3121.2 < k ≤ 3138.882
	3138.882 < k ≤ 3140.308
26	3140.308 < k ≤ 3175.2
	3175.2 < k ≤ 3220.941
	3220.941 < k ≤ 3247.364
	3247.364 < k ≤ 3333.333
	3333.333 < k ≤ 3353.175
	3353.175 < k ≤ 3361.976
	3361.976 < k ≤ 3364
27	3364 < k ≤ 3381.251
	3381.251 < k ≤ 3385.6
	3386.5 < k ≤ 3392.308
	3392.308 < k ≤ 3422.25
	3422.25 < k ≤ 3434.083
	3434.083 < k ≤ 3458.824
	3458.824 < k ≤ 3525.904
	3525.904 < k ≤ 3528
	3528 < k ≤ 3556.056
	3556.056 < k ≤ 3604.286
28	3604.285 < k ≤ 3638.942
	3638.942 < k ≤ 3655.125
	3655.125 < k ≤ 3663.022
	3663.022 < k ≤ 3705.293
	3705.293 < k ≤ 3706.14
	3706.14 < k ≤ 3710.294
	3710.294 < k ≤ 3760.105
	3760.105 < k ≤ 3789.474
	3789.474 < k ≤ 3822.727
	3822.727 < k ≤ 3844
29	3844 < k ≤ 3855.813
	3855.813 < k ≤ 3861.161
	3861.161 < k ≤ 3913.575
	3913.575 < k ≤ 3952.667
	3952.667 < k ≤ 3983.684
	3983.684 < k ≤ 4004.167
	4004.167 < k ≤ 4009.263
	4009.263 < k ≤ 4082.4
	4082.4 < k ≤ 4100.267

30	$4100.267 < k \leq 4106.7$
	$4106.7 < k \leq 4114.286$
	$4114.286 < k \leq 4120.9$
	$4120.9 < k \leq 4134.723$
	$4134.723 < k \leq 4138.277$
	$4138.277 < k \leq 4205$
	$4205 < k \leq 4205.402$
	$4205.402 < k \leq 4225$
	$4225 < k \leq 4247.059$
	$4247.059 < k \leq 4316.07$
	$4316.07 < k \leq 4324.5$
	$4324.5 < k \leq 4356$
31	$4356 < k \leq 4373.082$
	$4373.082 < k \leq 4425.143$
	$4425.143 < k \leq 4473.018$
	$4473.018 < k \leq 4485.333$
	$4485.333 < k \leq 4491.86$
	$4491.86 < k \leq 4500$
	$4500 < k \leq 4505.357$
	$4505.357 < k \leq 4514.721$
	$4514.721 < k \leq 4536$
	$4536 < k \leq 4576.19$
	$4576.19 < k \leq 4628.25$
32	$4628.25 < k \leq 4646.4$
	$4646.4 < k \leq 4662.533$
	$4662.533 < k \leq 4686.019$
	$4686.019 < k \leq 4693.371$
	$4693.371 < k \leq 4732.647$
	$4732.647 < k \leq 4827.429$
	$4827.429 < k \leq 4841.862$
	$4841.862 < k \leq 4900$
33	$4900 < k \leq 4914.205$
	$4914.205 < k \leq 4914.935$
	$4914.935 < k \leq 4917.014$
	$4917.014 < k \leq 4932.267$
	$4932.267 < k \leq 4968.772$
	$4968.772 < k \leq 4970.02$
	$4970.02 < k \leq 4995.03$
	$4995.03 < k \leq 4995.571$
	$4995.571 < k \leq 5029.762$
	$5029.762 < k \leq 5057.5$
	$5057.5 < k \leq 5068.8$
	$5068.8 < k \leq 5078.4$
	$5078.4 < k \leq 5187.179$
	$5187.179 < k \leq 5188.235$
34	$5188.235 < k \leq 5195.782$

	5195.782 < k ≤ 5202
	5202 < k ≤ 5222.368
	5222.368 < k ≤ 5292
	5292 < k ≤ 5348.174
	5348.174 < k ≤ 5364.015
	5364.015 < k ≤ 5372.05
	5372.05 < k ≤ 5387.13
	5387.13 < k ≤ 5400
	5400 < k ≤ 5473.409
	5473.409 < k ≤ 5476
35	5476 < k ≤ 5484.783
	5484.783 < k ≤ 5492.952
	5492.952 < k ≤ 5544.45
	5544.45 < k ≤ 5551.049
	5551.049 < k ≤ 5597.233
	5597.233 < k ≤ 5600.045
	5600.045 < k ≤ 5616.947
	5616.947 < k ≤ 5632.457
	5632.457 < k ≤ 5714.286
	5714.286 < k ≤ 5733.426
	5733.426 < k ≤ 5752.174
	5752.174 < k ≤ 5760.467
	5760.467 < k ≤ 5780.222
36	5780.222 < k ≤ 5808
	5808 < k ≤ 5814.224
	5814.224 < k ≤ 5828.167
	5828.167 < k ≤ 5852.25
	5852.25 < k ≤ 5857.524
	5857.524 < k ≤ 5883.44
	5883.44 < k ≤ 5900.417
	5900.417 < k ≤ 5933.859
	5933.859 < k ≤ 5990.412
	5990.412 < k ≤ 6006.25
	6006.25 < k ≤ 6013.636
	6013.636 < k ≤ 6084
37	6084 < k ≤ 6100.9
	6100.9 < k ≤ 6139.242
	6139.242 < k ≤ 6152.168
	6152.168 < k ≤ 6160.5
	6160.6 < k ≤ 6212.348
	6212.348 < k ≤ 6239.557
	6239.557 < k ≤ 6293.778
	6293.778 < k ≤ 6294.42
	6294.42 < k ≤ 6307.5
	6307.5 < k ≤ 6336.013
	6336.013 < k ≤ 6350.4
	6350.4 < k ≤ 6366.156

	$6366.156 < k \le 6372.206$
	$6372.206 < k \le 6400.9$
	$6400.9 < k \le 6401.515$
	$6401.515 < k \le 6404.21$
38	$6404.21 < k \le 6438.095$
	$6438.095 < k \le 6451.724$
	$6451.724 < k \le 6506.952$
	$6506.952 < k \le 6536.679$
	$6536.679 < k \le 6560.427$
	$6560.427 < k \le 6589.453$
	$6589.453 < k \le 6613.043$
	$6613.043 < k \le 6724$
39	$6724 < k \le 6740.852$
	$6740.852 < k \le 6757.92$
	$6757.92 < k \le 6760$
	$6760 < k \le 6784.615$
	$6784.615 < k \le 6791.919$
	$6791.919 < k \le 6808.173$
	$6808.173 < k \le 6823.938$
	$6823.938 < k \le 6868.167$
	$6868.167 < k \le 6878.772$
	$6878.772 < k \le 6880.235$
	$6880.235 < k \le 6912.014$
	$6912.014 < k \le 6916.945$
	$6916.945 < k \le 6919.2$
	$6919.2 < k \le 6953.143$
	$6953.143 < k \le 7004.167$
	$7004.167 < k \le 7039.5$
	$7039.5 < k \le 7060.2$
40	$7060.2 < k \le 7093.981$
	$7093.981 < k \le 7142.087$
	$7142.087 < k \le 7148.308$
	$7148.308 < k \le 7162.522$
	$7162.522 < k \le 7172.267$
	$7172.267 < k \le 7200$
	$7200 < k \le 7254.857$
	$7254.857 < k \le 7267.853$
	$7267.853 < k \le 7301.333$
	$7301.333 < k \le 7321.615$
	$7321.615 < k \le 7350$
	$7350 < k \le 7389.011$
	$7389.011 < k \le 7395.03$
	$7395.03 < k \le 7396$
41	$7396 < k \le 7412.809$
	$7412.809 < k \le 7413.21$

$7413.21 < k \le 7463.698$
$7463.698 < k \le 7492.593$
$7492.593 < k \le 7511.111$
$7511.111 < k \le 7550.083$
$7550.083 < k \le 7569$
$7569 < k \le 7569.723$
$7569.723 < k \le 7603.292$
$7603.292 < k \le 7662.678$
$7662.678 < k \le 7674.438$
$7674.438 < k \le 7748.19$

42 $7748.19 < k \le 7770.889$
$7770.889 < k \le 7781.878$
$7781.878 < k \le 7840.471$
$7840.471 < k \le 7844.014$
$7844.014 < k \le 7844.587$
$7844.587 < k \le 7844.667$
$7844.667 < k \le 7850.14$
$7850.14 < k \le 7901.299$
$7901.299 < k \le 7920.643$
$7920.643 < k \le 7938$
$7938 < k \le 7954.809$
$7954.809 < k \le 8004.762$
$8004.762 < k \le 8022.222$
$8022.222 < k \le 8053.422$
$8053.422 < k \le 8098.787$
$8098.787 < k \le 8100$

43 $8100 < k \le 8116.77$
$8116.77 < k \le 8146.385$
$8146.385 < k \le 8167.5$
$8167.5 < k \le 8181.818$
$8181.818 < k \le 8253.474$
$8253.474 < k \le 8260.267$
$8260.267 < k \le 8286.259$
$8286.259 < k \le 8316.346$
$8316.346 < k \le 8320.5$
$8320.5 < k \le 8376.923$
$8376.923 < k \le 8391.724$
$8391.724 < k \le 8415.034$
$8415.034 < k \le 8429.167$
$8429.167 < k \le 8431.348$
$8431.348 < k \le 8468.182$

44 $8468.182 < k \le 8501.786$
$8501.786 < k \le 8520.845$
$8520.845 < k \le 8520.931$
$8520.931 < k \le 8523.009$
$8523.009 < k \le 8541.176$
$8541.176 < k \le 8569.8$
$8569.8 < k \le 8651.554$

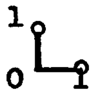
	$8651.554 < k \le 8673.887$
	$8673.887 < k \le 8740.02$
	$8740.02 < k \le 8750.893$
	$8750.893 < k \le 8791.253$
	$8791.253 < k \le 8816.667$
	$8816.667 < k \le 8816.906$
	$8816.906 < k \le 8836$
45	$8836 < k \le 8852.735$
	$8852.735 < k \le 8903.322$
	$8903.322 < k \le 8965.333$
	$8965.333 < k \le 8985.709$
	$8985.709 < k \le 8988.931$
	$8988.931 < k \le 8992.029$
	$8992.029 < k \le 9001.891$
	$9001,891 < k \le 9012.414$
	$9012.414 < k \le 9013.333$
	$9013.333 < k \le 9060.1$
	$9060.1 < k \le 9111.587$
	$9111.587 < k \le 9111.704$
	$9111.704 < k \le 9124.138$
	$9124.138 < k \le 9151.35$
	$9151.35 < k \le 9178.626$
	$9178.626 < k \le 9184.783$
	$9184.783 < k \le 9220.174$
46	$9220.174 < k \le 9234.698$
	$9234.698 < k \le 9253.702$
	$9253.702 < k \le 9274.294$
	$9274.294 < k \le 9284.21$
	$9284.21 < k \le 9321.494$
	$9321.494 < k \le 9334.286$
	$9334.286 < k \le 9413.505$
	$9413.505 < k \le 9425.067$
	$9425.069 < k \le 9481.225$
	$9481.225 < k \le 9522$
	$9522 < k \le 9564.915$
	$9564.915 < k \le 9565.432$
	$9565.432 < k \le 9566.797$
	$9566.797 < k \le 9604$
47	$9604 < k \le 9620.702$
	$9620.702 < k \le 9622.753$
	$9622.753 < k \le 9640.143$
	$9640.143 < k \le 9671.161$
	$9671.161 < k \le 9703.846$
	$9703.846 < k \le 9728.04$
	$9728.04 < k \le 9738.008$
	$9738.008 < k \le 9750.069$

	$9750.069 < k \le 9756.444$
	$9756.444 < k \le 9766.207$
	$9766.207 < k \le 9812.903$
	$9812.903 < k \le 9856.8$
	$9856.8 < k \le 9873.041$
	$9873.041 < k \le 9878.4$
	$9878.4 < k \le 9956.571$
	$9956.571 < k \le 9979.482$
	$9979.482 < k \le 10004.17$
48	$10004.17 < k \le 10037.63$
	$10037.63 < k \le 10039.75$
	$10039.75 < k \le 10052.14$
	$10052.14 < k \le 10105.22$
	$10105.22 < k \le 10121.31$
	$10121.31 < k \le 10169.50$
	$10169.50 < k \le 10192.61$
	$10192.61 < k \le 10208.33$
	$10208.33 < k \le 10209.38$
	$10209.38 < k \le 10244.27$
	$10244.27 < k \le 10261.16$
	$10261.16 < k \le 10266.675$
	$10266.675 < k \le 10296.12$
	$10296.12 < k \le 10300.24$
	$10300.24 < k \le 10349.14$
	$10349.14 < k \le 10404$
49	$10404 < k \le 10420.67$
	$10420.67 < k \le 10433.58$
	$10433.58 < k \le 10471.01$
	$10471.01 < k \le 10496.97$
	$10496.97 < k \le 10506.25$
	$10506.25 < k \le 10530.70$
	$10530.70 < k \le 10556.01$
	$10556.01 < k \le 10584$
	$10584 < k \le 10603.20$
	$10603.20 < k \le 10628.15$
	$10628.15 < k \le 10672.20$
	$10672.20 < k \le 10677.34$
	$10677.34 < k \le 10677.78$
	$10677.78 < k \le 10757.17$
	$10757.17 < k \le 10800$
	$10800 < k \le 10804.50$
	$10804.50 < k \le 10820.16$
50	$10820.16 < k \le 10820.45$
	$10820.45 < k \le 10837.50$
	$10837.50 < k \le 10847.47$
	$10847.47 < k \le 10852.45$

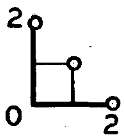
$10852.45 < k \leq 10853.56$
$10853.56 < k \leq 10864.29$
$10864.29 < k \leq 10895.67$
$10895.67 < k \leq 10920.97$
$10920.97 < k \leq 10981.94$
$10981.94 < k \leq 11016.31$
$11016.31 < k \leq 11023.67$
$11023.67 < k \leq 11034.01$
$11034.01 < k \leq 11039.90$
$11039.90 < k \leq 11163.66$
$11163.66 < k \leq 11175.56$
$11175.56 < k \leq 11232.14$
$11232.14 < k \leq 11236$

FIGURES (Optimal Sampling Plans)

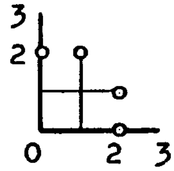
(1)



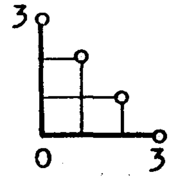
(2)



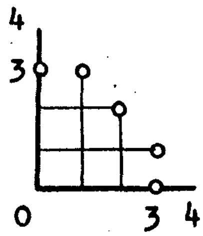
(3)



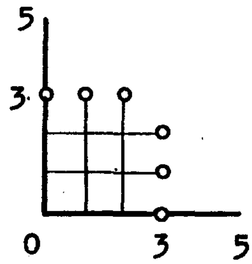
(4)



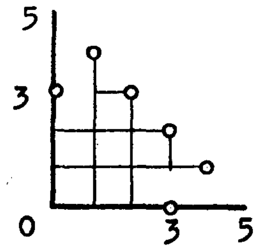
(5)



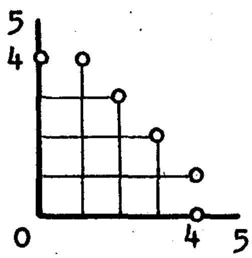
(6)



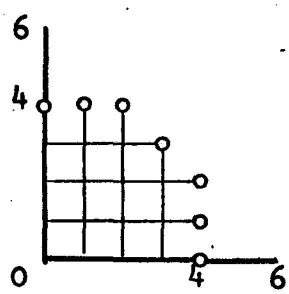
(7)



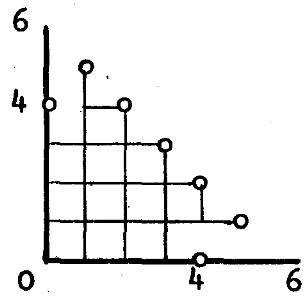
(8)



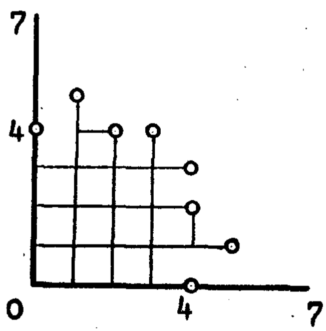
(9)



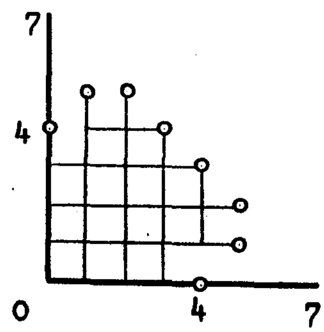
(10)



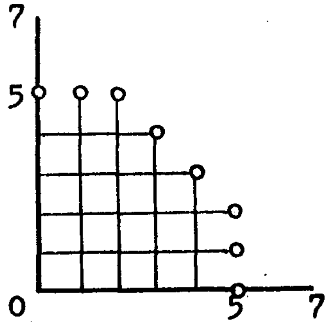
(11)



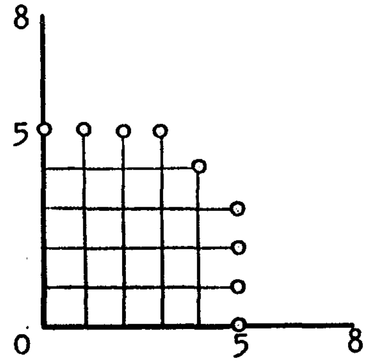
(12)



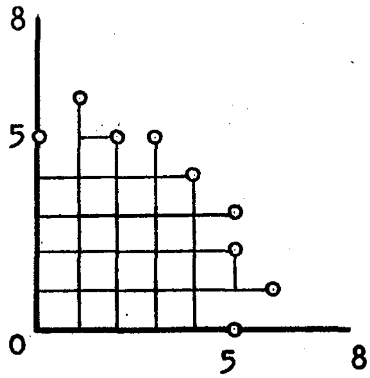
(13)



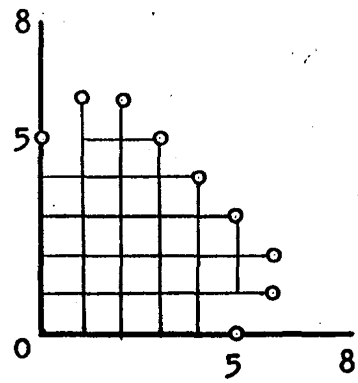
(14)



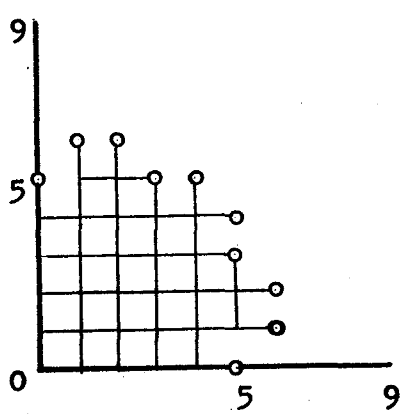
(15)



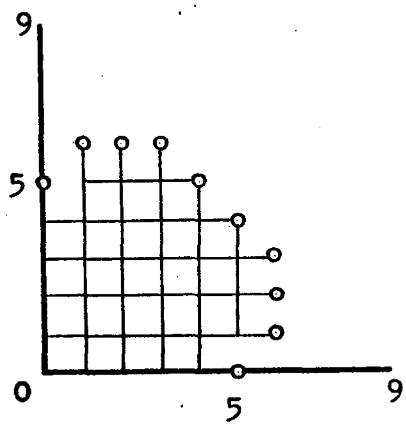
(16)



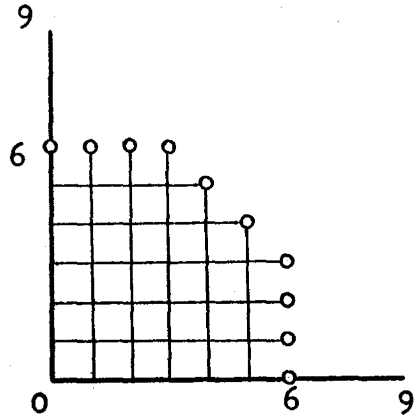
(17)



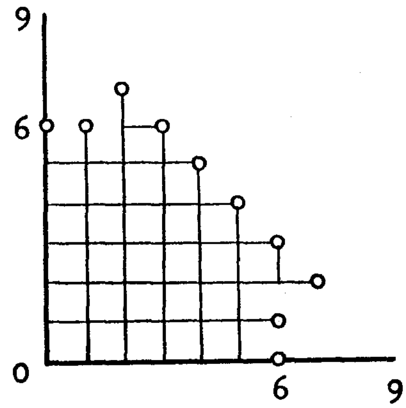
(18)



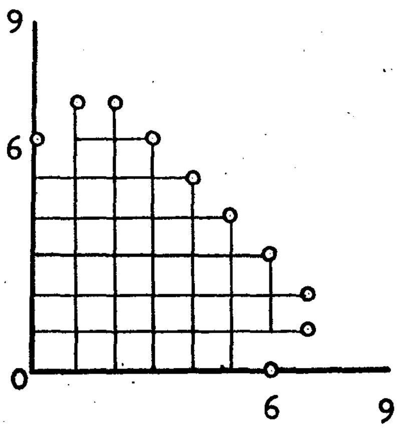
(19)



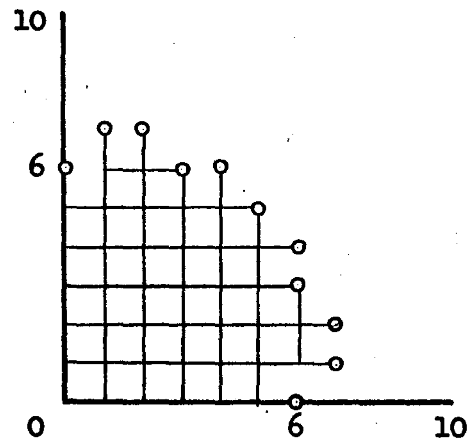
(20)



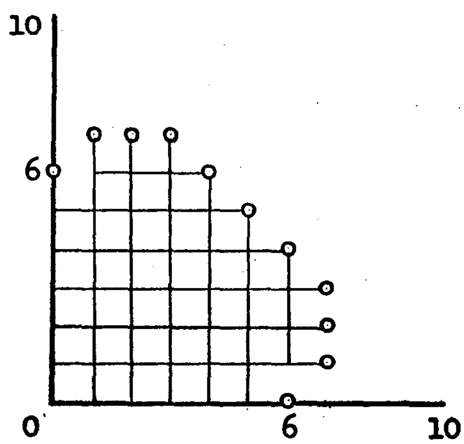
(21)



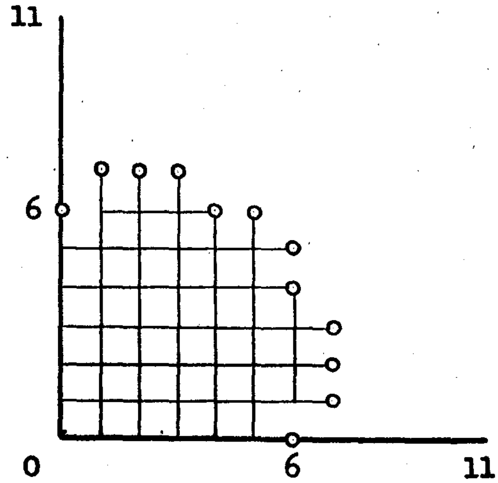
(22)



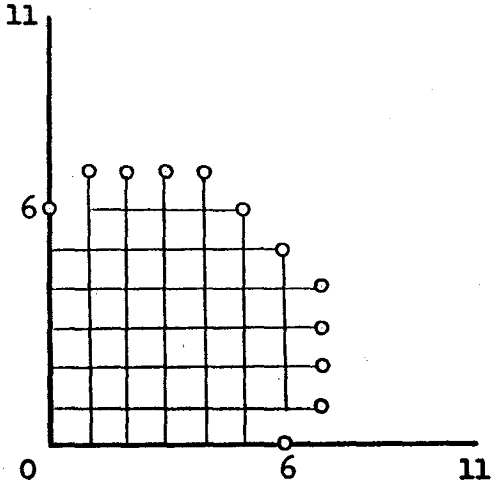
(23)



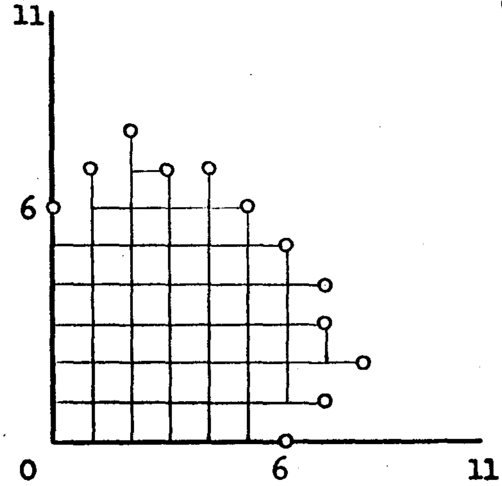
(24)



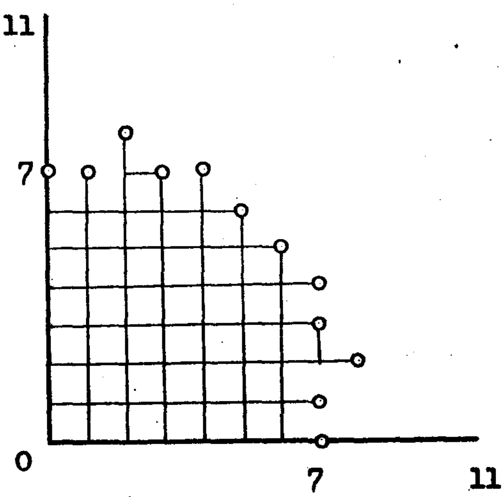
(25)



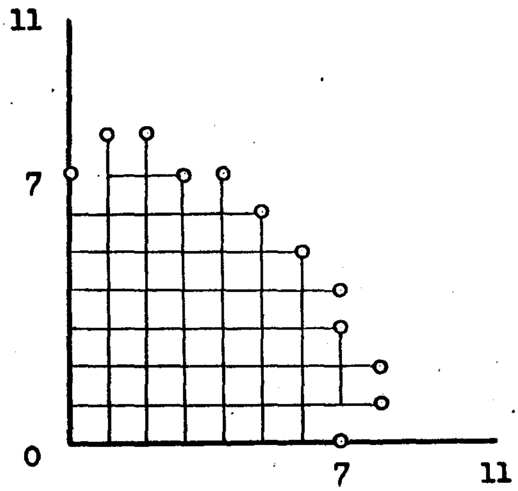
(26)



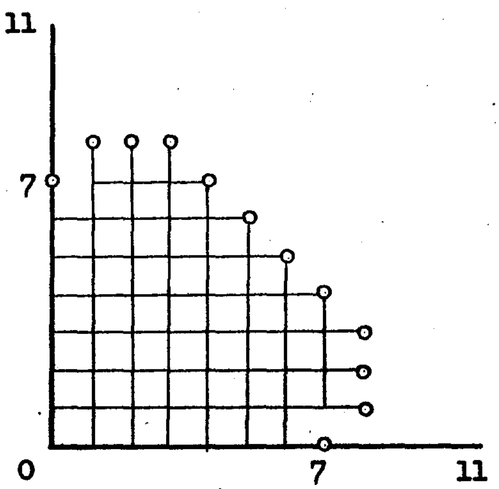
(27)



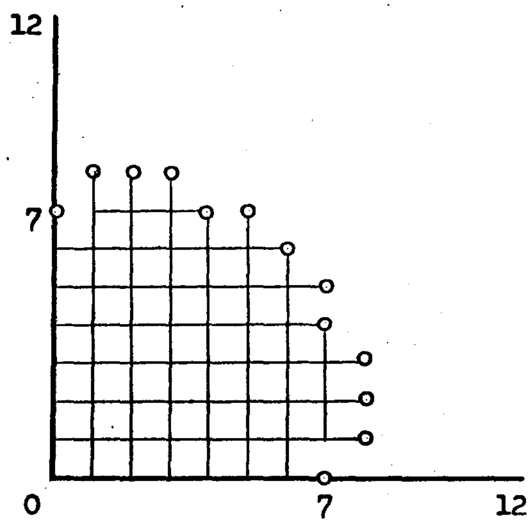
(28)



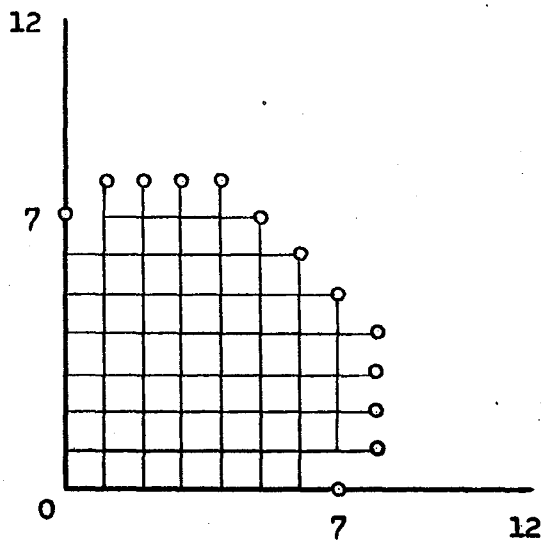
(29)



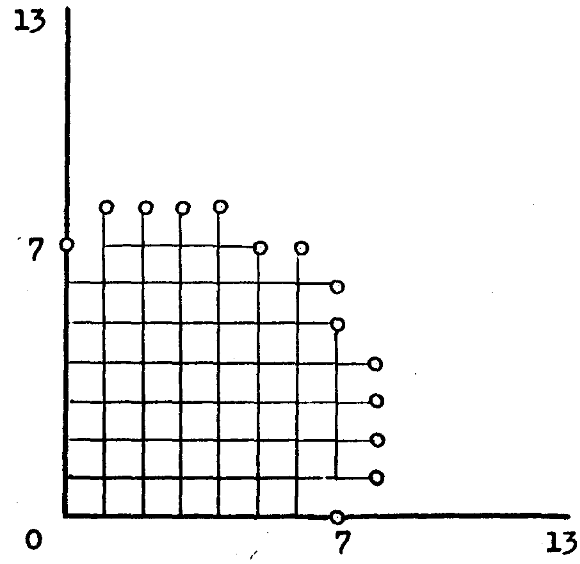
(30)



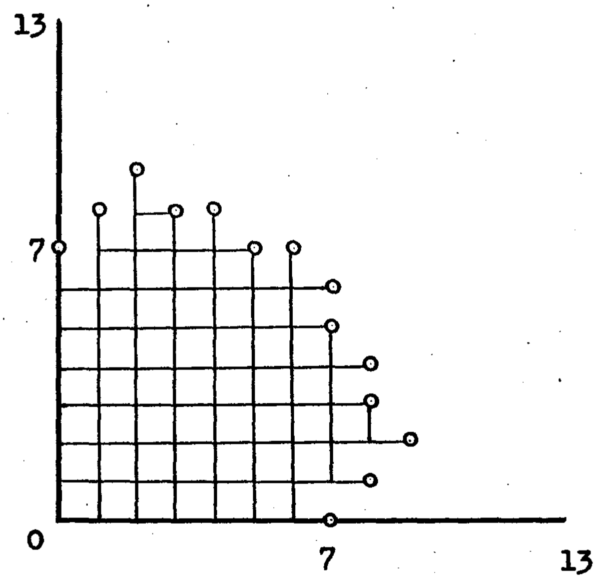
(31)



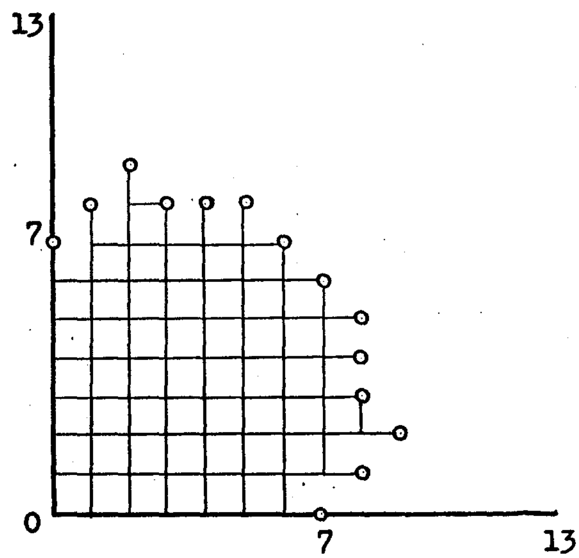
(32)



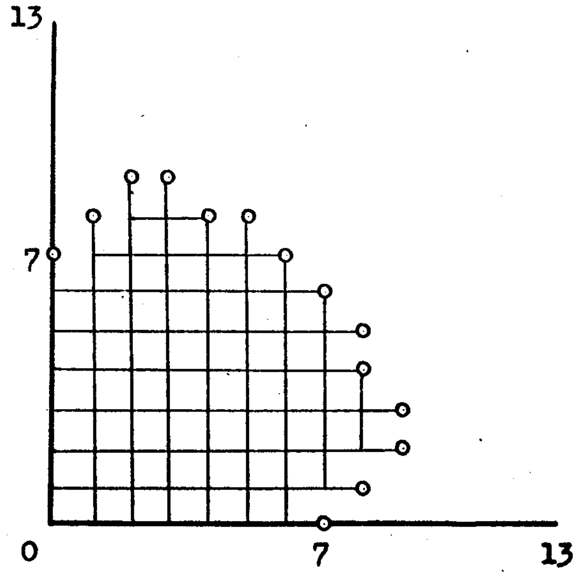
(33)



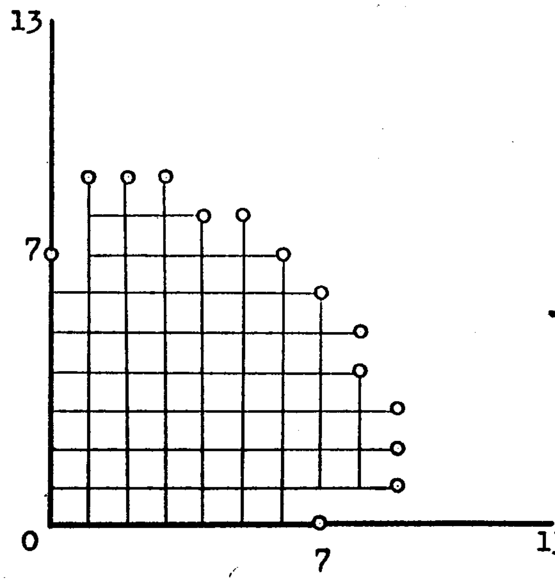
(34)



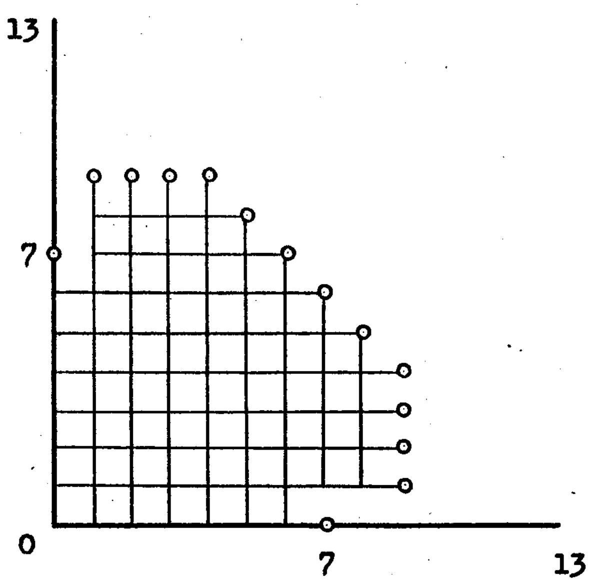
(35)



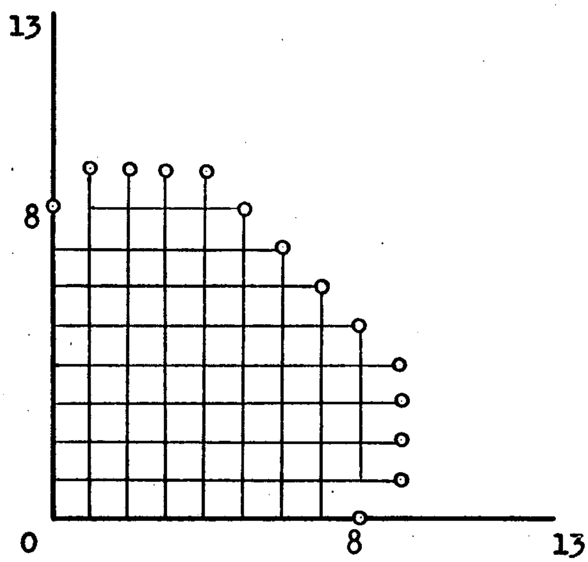
(36)



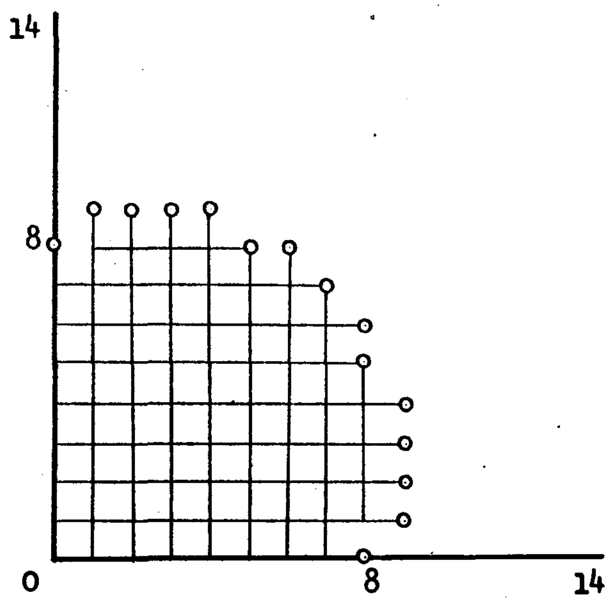
(37)



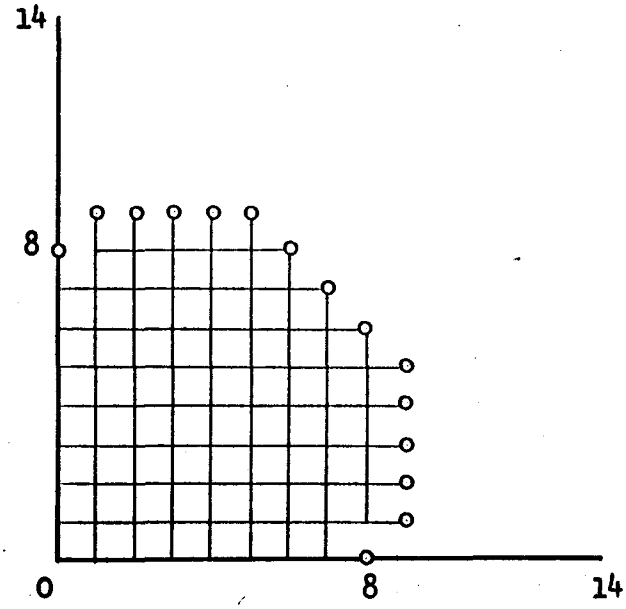
(38)



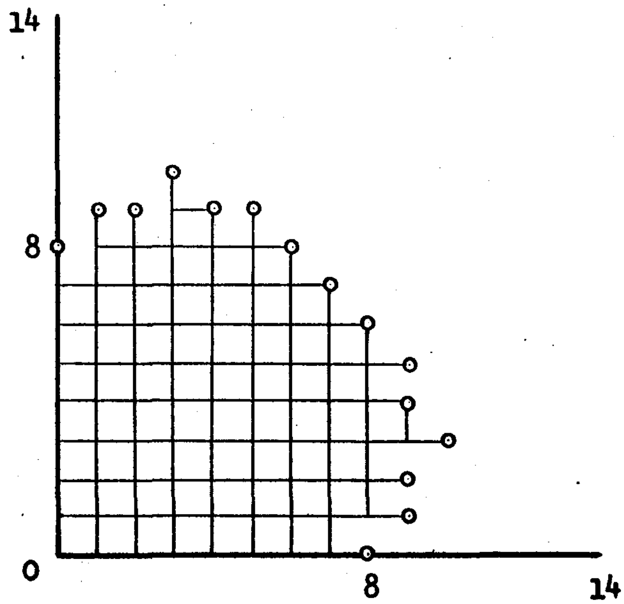
(39)



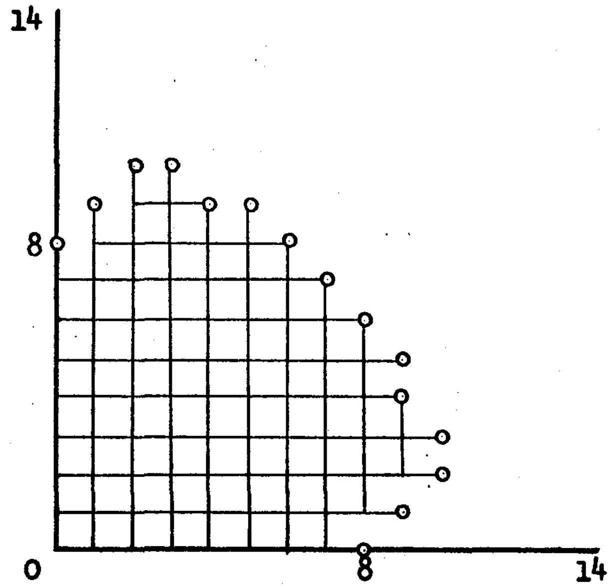
(40)



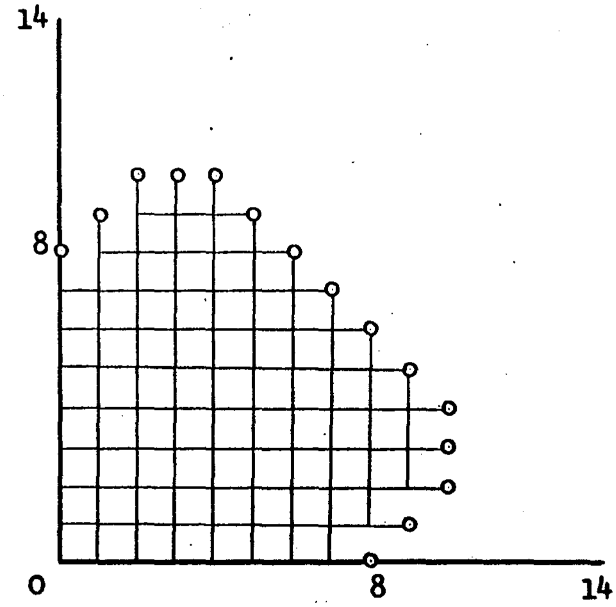
(41)



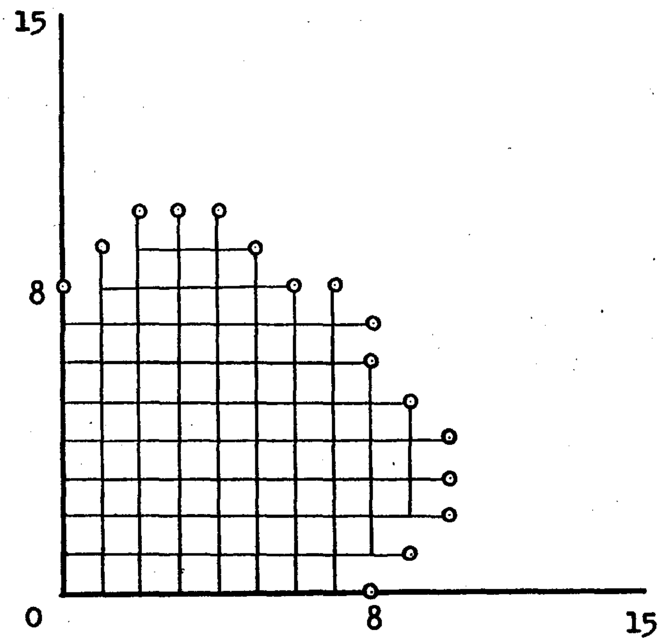
(42)



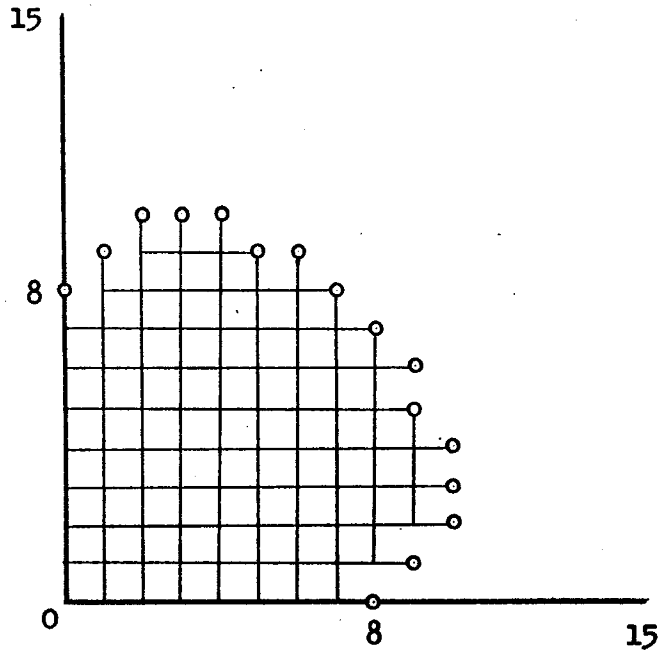
(43)



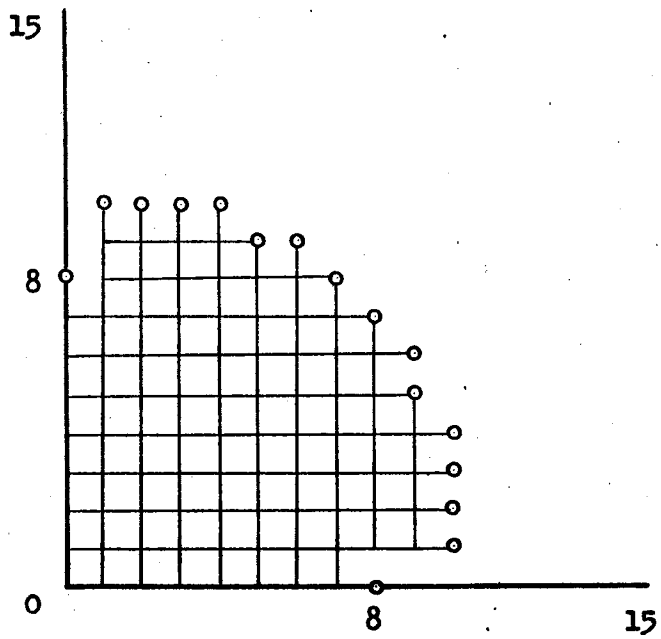
(44)



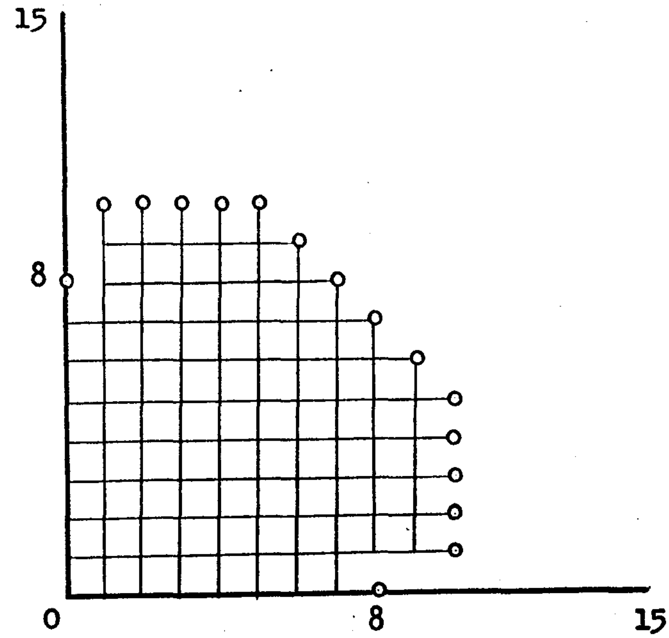
(45)



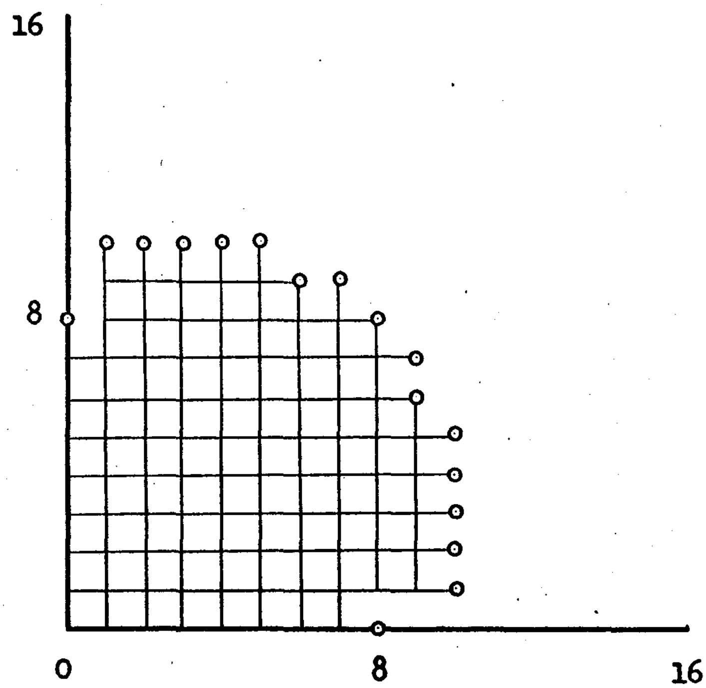
(46)



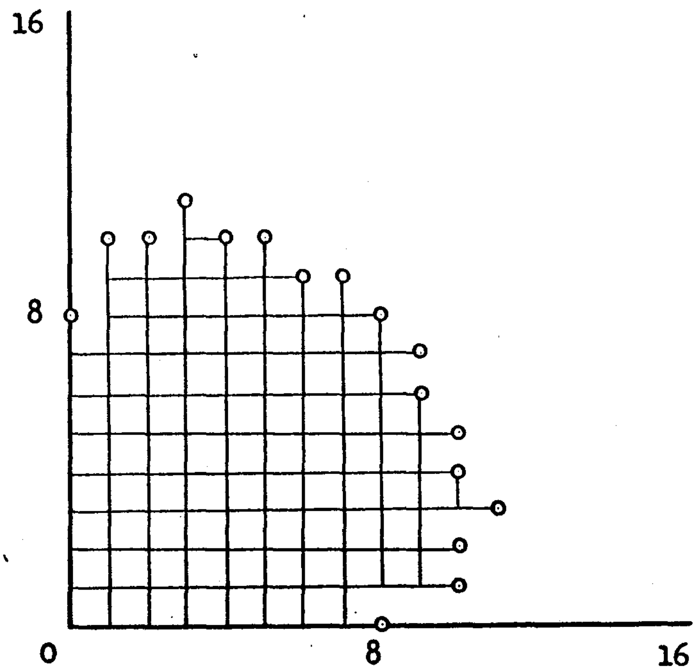
(47)



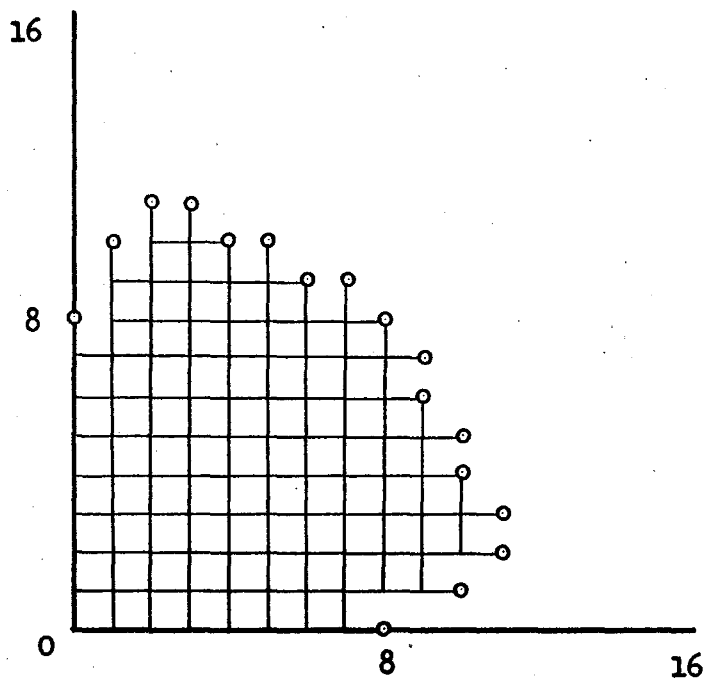
(48)



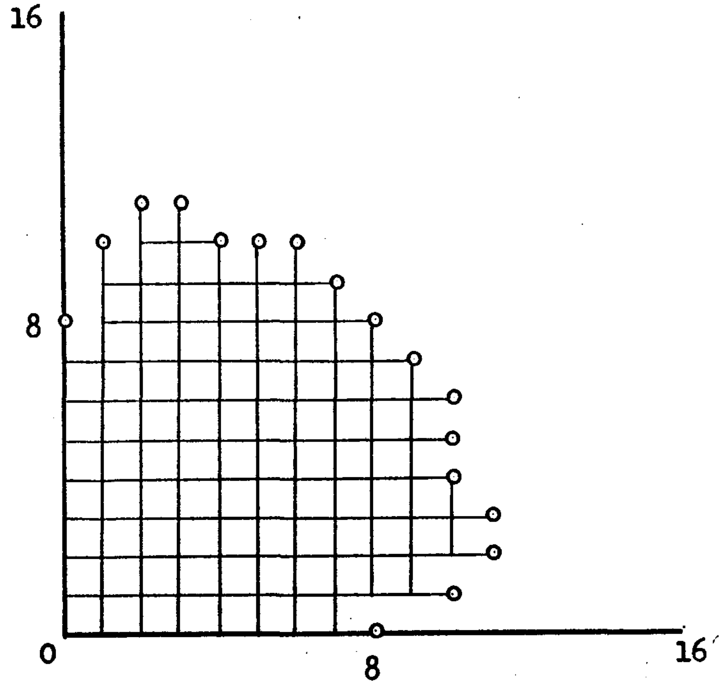
(49)



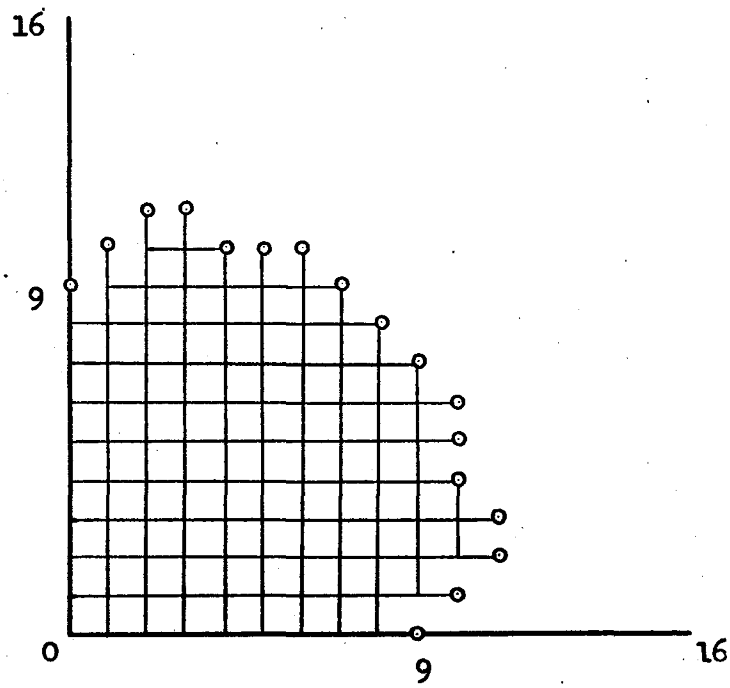
(50)



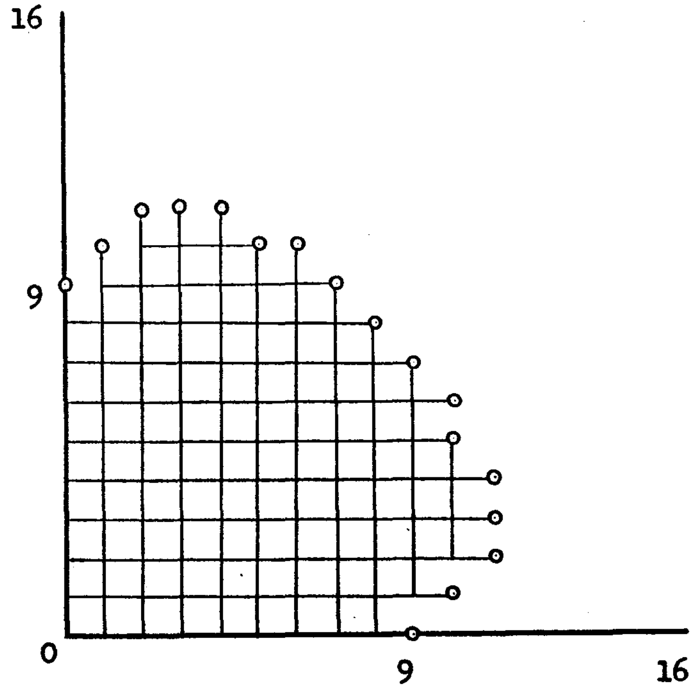
(51)



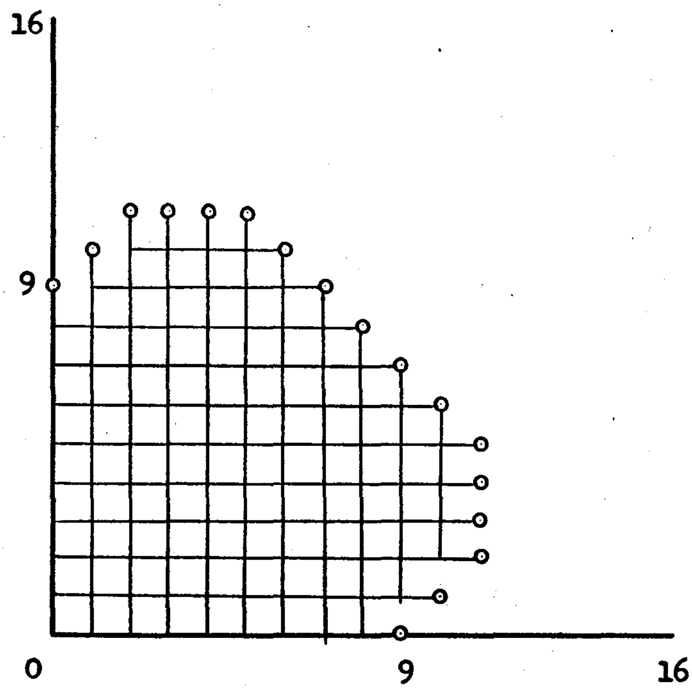
(52)



(53)



(54)



References

[1] M. A. Girshick, F. Mosteller, and L. J. Savage, "Unbiased estimates for certain binomial sampling problems with applications". Ann. Math. Stat., Vol. 17 (1946) pp. 13-23.

[2] M. H. DeGroot, "Unbiased sequential estimation for binomial populations", Ann. Math. Stat., Vol. 30 (1959) pp. 80-101.

[3] B. Brainerd and T. V. Narayana, "A note on binomial sampling plans", unpublished.

[4] D. Blackwell and M. A. Girshick, Theory of Games and Statistical Decisions, John Wiley and Sons, New York, 1954.

[5] J. Wolfowitz, "On sequential binomial estimation", Ann. Math. Stat., Vol. 17 (1946) pp. 489-493.

[6] L. J. Savage, "A uniqueness theorem for unbiased sequential estimation", Ann. Math. Stat., Vol. 18 (1947) pp. 295-297.

[7] J. Wolfowitz, "The efficiency of sequential estimates and Wald's equation for sequential processes", Ann. Math. Stat., Vol. 18 (1947), pp. 215-230.

[8] L. J. Savage, The foundations of statistics, John Wiley and Sons, New York, 1954.

[9] M. C. K. Tweedie, "Inverse statistical variates", Nature, Vol. 155 (1945) pp. 453.

[10] J. B. S. Haldane, "A labour-saving method of sampling", Nature, Vol. 155 (1945), pp. 49-50.

[11] J. B. S. Haldane, "On a method of estimating frequencies", Biometrika, Vol. 33 (1945) pp. 222-225.

[12] B. Brainerd and T. V. Narayana, "A note on simple Binomial sampling plans", Ann. Math. Stat., Vol. 32 (1961) pp. 906-908.