BAI, Do Sun, 1937-
EFFICIENT SAMPLING PLANS IN A TWO-STATE
MARKOV CHAIN.

The Ohio State University, Ph.D., 1971
Statistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED
EFFICIENT SAMPLING PLANS IN A TWO-STATE MARKOV CHAIN

DISSertation

Presented in Partial Fulfillment of the requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Do Sun Bai, B.S., M.S.

The Ohio State University
1971

Approved by

J.S. Portegi
Adviser
Division of Statistics
PLEASE NOTE:

Some Pages have indistinct print. Filmed as received.

UNIVERSITY MICROFILMS
ACKNOWLEDGEMENTS

I would like to thank Professor Jagdish S. Rustagi for the time, encouragement and advice he has so generously given during the preparation of this dissertation and Professor Peter O. Anderson for many helpful discussions. I am also grateful to Professor Thomas A. Willke for the interest he has taken in my professional development throughout my graduate career.
VITA

October 11, 1937. . . Born - Haeju City, Korea

1960. . . . . . . B.S. in Engineering, Seoul National University, Seoul, Korea

1960 - 1967 . . . Member of Technical Staff, The Korea Development Bank, Seoul, Korea

1963. . . . . . . B.S. in Commerce, Seoul National University, Seoul, Korea

1968 - 1969 . . . Teaching Assistant, Department of Mathematics, The Ohio State University, Columbus, Ohio

1969. . . . . . . M.S., The Ohio State University, Columbus, Ohio

1970. . . . . . . Teaching Associate, Department of Mathematics and Division of Statistics, The Ohio State University, Columbus, Ohio

1970 - 1971 . . . Research Fellow, The Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Statistics

Real Analysis. Professor Ranko Bojanic

Probability. Professor Fredo Papangelou

Statistical Inference. Professor Austin M. Barron
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. ii

VITA ............................................................................................... iii

Chapter

I INTRODUCTION ................................................................. 1

II EFFICIENT SAMPLING PLANS .............................................. 5

  Preliminaries
  Efficient Sampling Plans for Binomial Populations
  Efficient Sampling Plans for Functions of One Parameter
  Efficient Sampling Plans for Functions of Two Parameters
  Completeness of Efficient Sampling Plans

III INFERENCE UNDER SAMPLING PLANS \( S(N_{11} + N_{10}) \) .... 30

  Sampling Plan \( S_c(X) \)
  Sampling Plan \( S_c(N_{11} + N_{10}) \)

IV INFERENCE UNDER SAMPLING PLANS \( S(N_{11}) \) ................. 49

V INFERENCE UNDER SAMPLING PLANS \( S(N_{10}) \) ............... 58

  Sampling Plan \( S_c(R) \)
  Sampling Plan \( S_c(N_{10}) \)

VI INFERENCE UNDER SAMPLING PLANS \( S(N_{10} + N_{01}) \) .... 73

  Sampling Plan \( S_{2c}(N_{10} + N_{01}) \)
  Sampling Plan \( S_{2c-1}(N_{10} + N_{01}) \)

VII DISCUSSION .......................................................................... 86

BIBLIOGRAPHY ............................................................................. 89
A Markov chain is a suitable model for many natural phenomena. It is a special stochastic process whose development may be treated as series of transitions between certain states of the process. In a Markov chain, the probability of transition to a state from a given state depends only on the given state. Recently there has been a vigorous development in the theory as well as the applications of Markov chains. The extensive treatment of Markov chains is available in many standard text books.

Statistical inference for finite Markov chains has been studied by many authors. Most of the studies have been concerned with large samples. Bartlett (1951) and Hoel (1954) have studied the problem of maximum likelihood estimation of the transition probabilities, goodness of fit tests, and tests of order of dependence. Their studies are based on one long, unbroken sequence of observations \( \{X_1, \ldots, X_n\} \) on the chain. Bartlett has shown that the transition count, \( N_{ij} \), which is the number of times that the process was in state \( i \) at time \( t - 1 \), and in state \( j \) at time \( t \) for \( t = 2, \ldots, n \), and the maximum likelihood estimate \( \hat{p}_{ij} = N_{ij} / \sum_j N_{ij} \) are asymptotically normal. Bartlett has also found the asymptotic distribution of the likelihood ratio test statistic for certain tests of hypotheses. Hoel (1954) extended some of Bartlett's results.
Anderson and Goodman (1957) considered similar problems with a different approach. Under the condition that a large number, say $r$, of repeated sequence of $n$-long observations $\{X_1, \ldots, X_n\}$ are available on the chain, they obtained the maximum likelihood estimates of the transition probabilities. The joint asymptotic distribution of the estimates in any Markov chain when $r$ tends to infinity, was found to be multivariate normal. They also obtained likelihood ratio tests and chi-square tests.

Billingsley (1961a) gave an exposition with an extensive bibliography on statistical inference for finite Markov chains. In a further monograph Billingsley (1961b) gave a more rigorous mathematical analysis of these problems and generalizes the results to denumerable Markov chains and continuous-time Markov processes.

Not much work has been done in the finite sample case, especially in the study of unbiased estimation. If we assume that the transition probabilities $p_{ij}$ and $p_{i'j'} (i \neq j)$ are not related, it can easily be seen that there do not exist unbiased estimators of functions of transition probabilities from a given sequence of observations $\{X_1, \ldots, X_N\}$ on the Markov chain when $N$ is fixed. This leads us to the study of sequential sampling plans or stopping rules. One would like to have an additional criterion of optimality such as efficiency in estimators and sampling plans. By an efficient estimator we mean a function of observations whose variance attains the lower bound of the Cramer–Rao inequality or its sequential version for all values of the parameters. The expected value of the estimator is then said to be estimable efficiently. A sampling plan is efficient if it admits an efficient estimator.
The problem of finding efficient estimators and sampling plans has been studied first by Girshick et al. (1966) and DeGroot (1959) for binomial populations. DeGroot has shown that the only efficient sampling plans for binomial populations with parameter \( p \) are single or fixed sample size sampling plans and inverse binomial sampling plans. He has also shown that the only functions of the parameter \( p \) which are estimable efficiently are \( p, \frac{1}{p}, \) and \( \frac{1}{1-p} \). Bhat and Kulkarni (1966) have extended these results to multinomial populations.

Recently Trybula (1968) has studied similar problems for certain one-parameter stochastic processes with independent increments, namely the Poisson process, negative binomial process and Wiener process.

This dissertation is concerned with the problem of finding efficient estimators and sampling plans for Markov chains. The underlying model is assumed to be a two state Markov chain, states being called success and failure. Let the probability of success given success at the previous stage be \( \alpha \) and the probability of success given failure at the previous stage be \( \beta, 0 < \alpha, \beta < 1 \). Further let \( N_{11} \) be the random variable denoting the number of transitions from success to success, \( N_{10} \) be the random variable denoting the number of transitions from success to failure, and so on for \( N_{01} \) and \( N_{00} \).

In chapter II, we show that the only efficient sampling plans for functions of one parameter are those plans for which one of the quantities \( N_{11} + N_{10}, N_{11}, N_{10}, N_{01} + N_{00}, N_{00} \), or \( N_{01} \) is held constant. We also show that the corresponding functions of one parameter which are estimable efficiently are of the forms \( \alpha, \frac{1}{\alpha}, \frac{1}{1-\alpha}, \beta, \frac{1}{1-\beta} \), and \( \frac{1}{\beta} \) respectively for each of above sampling plans. Furthermore, we show
that the only sampling plan which is efficient for functions of two parameters is the plan for which \( N_{10} + N_{01} = 2c \) for some positive integer \( c \). The corresponding functions which are estimable efficiently are of the form \( c \left( \frac{a}{1-\alpha} + \frac{b}{\beta} \right) + d \), \( a \), \( b \) and \( d \) being arbitrary constants.

In later chapters, sampling plans for which \( N_{11} + N_{10} \), \( N_{11} \) or \( N_{10} \) is constant are studied. Explicit forms of sampling distributions, moments and probability generating functions of the number of successes, number of failures, and total number of observations required are obtained in each sampling plan. Certain optimal estimators for \( \alpha \) and \( \beta \) are given for each sampling plan considered.

Some new directions for research are proposed in the last chapter. Somewhat straightforward extensions to Markov chains with more than two states are mentioned.
In this chapter various definitions are given. The notion of efficient sampling plan for a finite Markov chain is developed. The case of Bernoulli independent trials is reviewed. Certain conditions for a sampling plan to be efficient for estimating functions of transition probabilities in a Markov chain are given and a characterization of the plans is obtained.

2.1 Preliminaries

Let \( \{X_1, X_2, \ldots\} \) be a two-state Markov chain. We call these states success denoted by S or 1 and failure denoted by F or 0. Let the stationary transition probabilities be given by

\[
p(X_t = 1|X_{t-1} = 1) = \alpha, \quad 0 < \alpha < 1, \tag{2.1.1}
\]

\[
p(X_t = 1|X_{t-1} = 0) = \beta, \quad 0 < \beta < 1, \tag{2.1.2}
\]

and the initial probability distribution given by

\[
p(X_1 = 1) = p, \quad 0 < p < 1,
\]

\[
p(X_1 = 0) = q, \quad q = 1 - p.
\]

Let \( \bar{\alpha} = 1 - \alpha \) and \( \bar{\beta} = 1 - \beta \).

Given a sequence of observations \( (X_1, X_2, \ldots, X_N) \) on the Markov chain, let
(2.1.3) \( N_{ij}(t) = \begin{cases} 1 & \text{if } X_{t-1} = i \text{ and } X_t = j, \\ 0 & \text{otherwise} \end{cases} \)

and

(2.1.4) \( N_{ij} = N_{ij}(X_1, \ldots, X_N) = \sum_{t=2}^{N} N_{ij}(t). \)

that is, the transition count random variable \( N_{ij} \) is the total number of transitions from state \( i \) to state \( j \) in the sequence of observations \( (X_1, \ldots, X_N) \).

Let \( X = X(X_1, \ldots, X_N) = \sum_{t=1}^{N} X_t \)
\( Y = X(X_1, \ldots, X_N) = N - X. \)

Then, \( X \) and \( Y \) are random variables denoting the total numbers of successes and failures respectively.

The above quantities satisfy the relations which are given in the following lemma.

**Lemma 2.1.1.**

(2.1.5) \( N_{11} + N_{10} + N_{01} + N_{00} = N - 1 \)
\( X = N_{11} + N_{10} - X_1 = N_1 - X_1 \)

(2.1.6) \( Y = N_{00} + N_{01} - 1 + X_1 = N_0 - 1 + X_1 \)

(2.1.7) \( N_{10} - N_{01} = X_1 - X_N. \)

We shall denote the values taken by the random variables \( N_{ij} \) by \( n_{ij} \), \( i, j = 0, 1, \) and we let \( N \) denote the quadruple \( (N_{11}, N_{10}, N_{01}, N_{00}) \).

By a **sampling plan**, we mean a rule that specifies at each stage of the sampling process whether sampling is to cease or another observation is to be taken. The **sample size** \( N \) is a random variable whose distribution is completely specified by the sampling plan.

Let the set \( B \), called the **stopping set**, be the set of all quad-
ruples \( \underline{n} = (n_{11}, n_{10}, n_{01}, n_{00}) \) of possible transition counts under all possible sample sizes \( N = n \) of the given sampling plan.

The probability of obtaining a particular transition count \( \underline{N} = \underline{n} \), given \( X_1 = x_1 \) is

\[
(2.1.8) \quad p(\underline{n} | x_1) = k(\underline{n} | x_1) \cdot n_{11} n_{10} n_{01} n_{00} ,
\]

where \( k(\underline{n} | x_1) \) is the number of all possible sequences of \( N \) observations which yield the same transition count \( \underline{N} = \underline{n} \), given \( X_1 = x_1 \).

Therefore, the joint probability distribution of \( (X_1, \underline{N}) \) is

\[
(2.1.9) \quad p(x_1, \underline{n}) = p[X_1 = x_1, \underline{N} = \underline{n}] = k(\underline{n} | x_1) \cdot p^{x_1} q^{1-x_1} n_{11} n_{10} n_{01} n_{00} .
\]

We give below some definitions, assumptions and lemmas which are needed later. They are similar to those given by DeGroot (1959) for binomial case.

**Definition 2.1.1.** A sampling plan is said to be closed if

\[
\sum_{B^*} p(x_1, \underline{n}) = 1
\]

where

\[
B^* = \{(x_1, \underline{n}) : x_1 \in \{0, 1\}, \underline{n} \in B \text{ and } B \text{ is the stopping set of the sampling plan}\}
\]

We assume two conditions.

(i) For every sampling plan to be considered,

\[
E(n^2) = \sum_{B^*} n^2 \cdot p(x_1, \underline{n})
\]

is uniformly convergent on every closed intervals of values of \( p, \alpha, \beta \).

(ii) For every estimator \( f \), which is a real-valued function defined on \( B^* \),

\[
g(p, \alpha, \beta) = E(f) = \sum_{B^*} f(x_1, \underline{n}) \cdot p(x_1, \underline{n})
\]
is differentiable termwise with respect to $p$, $a$, and $\beta$ in the open intervals $0 < p, a, \beta < 1$ and the derived series is uniformly convergent.

Assumption (i) assures the existence of $E(N_{ij} \cdot N_{i'j'})$, $i,j,i',j'=0,1$, $E(X^2)$, $E(Y^2)$, and $E(XY)$, and it is also clear that these expectations are all less than or equal to $E(N^2)$.

**Lemma 2.1.2.** $N, X, Y$, and $N_{ij}$, considered as estimators, satisfy condition (ii).

**Proof.**

\[
E(N) = \sum_{B^*} n \cdot p(x_1, n) = \sum_{B^*} n \cdot \text{k}(n|x_1) \cdot p^1 q^{n-1} \cdot \frac{n!}{a 10} \cdot \frac{\beta 01}{\beta 00}
\]

A sufficient condition for the termwise differentiability of a series is that the termwise derivative of it be absolutely and uniformly convergent on every closed subinterval. The derivative with respect to $a$ is given by

\[
\frac{\partial E(N)}{\partial a} = \frac{1}{a \cdot \bar{a}} \sum_{B^*} n \cdot (\bar{a} n_{11} - a n_{10}) \cdot p(x_1, n).
\]

This is, in absolute value, less than $(1/a \cdot \bar{a}) \cdot E(N^2)$.

Hence by (i) this series converges uniformly on every closed intervals of $p$, $a$, and $\beta$, $0 < p, a, \beta < 1$, and therefore $E(N)$ is termwise differentiable with respect to $a$. Similarly it is termwise differentiable with respect to $\beta$ and $p$.

Similar arguments apply to $X, Y$, and $N_{ij}$.

**Lemma 2.1.3.** If $f = f(x_1, n)$ satisfies assumption (ii) and $E(f) = g(p, a, \beta) = g$, then,
\[
\begin{aligned}
E[(x_1 - p) f] &= p q \, g' \\
E[(\alpha N_{11} - \alpha N_{10}) f] &= \alpha \bar{\alpha} \, g' \\
E[(\beta N_{01} - \beta N_{00}) f] &= \beta \bar{\beta} \, g' 
\end{aligned}
\]

where \( g'_{\alpha} = \frac{\partial g}{\partial \alpha} \), \( g'_{\beta} = \frac{\partial g}{\partial \beta} \) and \( g'_{p} = \frac{\partial g}{\partial p} \).

In particular, when \( f = 1 \), we have
\[
E(x_1 - p) = E[\alpha N_{11} - \alpha N_{10}] = E[\beta N_{10} - \beta N_{00}] = 0.
\]

**Proof.** The proof is immediate by termwise partial differentiation of

\[
g(p, \alpha, \beta) = Ef = \sum_{B^*} f(x_1, n) \cdot p(x_1, n)
\]

with respect to \( p, \alpha \) or \( \beta \).

**Lemma 2.1.4.** Let \( N_{11} + N_{10} = N_1 \), \( N_{01} + N_{00} = N_0 \).

Then

\[
\begin{aligned}
(2.1.11) \quad & (a) \quad \bar{\alpha} \, E_N_{11} = \alpha \, E_N_{10} = \alpha \, \bar{\alpha} \, E_N_{11} = E[\alpha N_{11} - \alpha N_{10}]^2 \\
(2.1.12) \quad & (b) \quad \bar{\beta} \, E_N_{01} = \beta \, E_N_{00} = \beta \bar{\beta} \, E_N_{01} = E[\beta N_{01} - \beta N_{00}]^2 \\
& (c) \quad E[\alpha N_{11} - \alpha N_{10}] [\beta N_{01} - \beta N_{00}] = E(x_1 - p)[\alpha N_{11} - \alpha N_{10}] \bar{\beta} \bar{N}_{01} - \beta N_{00} = 0.
\end{aligned}
\]

**Proof.** (a) By Lemma 2.1.3., we have

\[
\bar{\alpha} \, E_N_{11} = \alpha \, E_N_{10}
\]

\[
= \alpha \, E_N_{11} - \alpha \, E_N_{11}
\]

Hence

\[
E_N_{11} = \alpha \, E_N_{11}.
\]

so that

\[
\bar{\alpha} \, E_N_{11} = \alpha \, E_N_{11}.
\]

Now, by differentiating

\[
\sum_{B^*} k(n|x_1) \, p \, q \, \frac{\partial}{\partial \alpha} \frac{\partial}{\partial \beta} \frac{\partial}{\partial p} = 1
\]
twice with respect to \( \alpha \), we get

\[
\sum \left( \frac{(\alpha n_{11} - \alpha n_{10})^2}{\alpha \bar{\alpha}} + (-n_{11} - n_{10}) \right) p(x_1, \bar{y}) = 0.
\]

Thus,

\[
\alpha \bar{\alpha} E \alpha_1 = E [\alpha n_{11} - \alpha n_{10}]^2.
\]

Similar arguments hold for (b) and (c).

Now, from (2.1.9)

\[
(2.1.13) \quad \log p(x_1, \bar{n}) = \log k(n|x_1) + x_1 \log p + (1-x_1) \log q + n_{11} \log \alpha + n_{10} \log \bar{\alpha} + n_{01} \log \beta + n_{00} \log \bar{\beta}.
\]

Hence define

\[
\begin{align*}
Z_1 &= \frac{\partial \log p}{\partial p} = \frac{x_1 - p}{pq} \\
Z_2 &= \frac{\partial \log p}{\partial \alpha} = \frac{\alpha n_{11} - \alpha n_{10}}{\alpha \bar{\alpha}} \\
Z_3 &= \frac{\partial \log p}{\partial \beta} = \frac{\beta n_{01} - \beta n_{00}}{\beta \bar{\beta}}.
\end{align*}
\]

It follows from Lemma 2.1.3 and Lemma 2.1.4 that

\[
\begin{align*}
EZ_{2} &= 0, \quad i = 1, 2, 3 \\
EZ_1^2 &= \frac{E(x_1 - p)^2}{pq} = \frac{1}{pq} \\
EZ_2^2 &= \frac{E(\alpha n_{11} - \alpha n_{10})^2}{(\alpha \bar{\alpha})^2} = \frac{E \alpha_1}{\alpha \bar{\alpha}} \\
EZ_3^2 &= \frac{E(\beta n_{01} - \beta n_{00})^2}{(\beta \bar{\beta})^2} = \frac{E \beta_0}{\beta \bar{\beta}}
\end{align*}
\]

and

\[
E(Z_i Z_j) = 0 \quad i \neq j.
\]

Hence the covariance matrix \( \Sigma = [E(Z_i Z_j)] \) is given by
\[(2.1.15) \quad \Sigma = \begin{bmatrix} 1/pq & 0 & 0 \\ 0 & EN_1 \cdot \bar{\alpha} & 0 \\ 0 & 0 & EN_0 / \beta \bar{\beta} \end{bmatrix} \]

Also
\[(2.1.16) \quad \Sigma^{-1} = \begin{bmatrix} pq & 0 & 0 \\ 0 & \alpha \bar{\alpha} / EN_1 & 0 \\ 0 & 0 & \beta \bar{\beta} / EN_0 \end{bmatrix} \]

with
\[(2.1.17) \quad |\Sigma| = \left(\frac{EN_1}{pq} \right) \left(\frac{EN_0}{\alpha \bar{\alpha} \beta \bar{\beta}}\right) > 0 .\]

We have the following theorem.

**Theorem 2.1.1. (Information Inequality)**

For any estimator \( f = f(X_1, N) \) with \( E(f) = g = g(p, \alpha, \beta) \) and sampling plan \( S \),
\[(2.1.18) \quad \sigma_f^2 = \text{Var}(f) \geq pq \left( g'_p \right)^2 + \frac{\alpha \bar{\alpha} (g'_\alpha)^2}{EN_1} + \frac{\beta \bar{\beta} (g'_\beta)^2}{EN_0} .\]

Equality holds if and only if \( f = f(x_1, n) \) is a linear function of \( z_1 = z_1(x_1), z_2 = z_2(n_{11}, n_{10}), \) and \( z_3 = z_3(n_{01}, n_{00}) \) for all \( (x_1, n) \in B^* \).

**Proof.** Let \( \lambda_1 = E f Z_1 = \text{Cov}(f, Z_1), \ i = 1, 2, 3. \)

And let \( \Lambda = (\lambda_1, \lambda_2, \lambda_3), \ Z = (Z_1, Z_2, Z_3). \)

Define
\[(2.1.19) \quad f^* = \Lambda \Sigma^{-1} Z'. \]

Then \( \text{Cov}(f, f^*) = \text{Cov}(f, \Lambda \Sigma^{-1} Z') = \Lambda \Sigma^{-1} \Lambda^t .\)
(2.1.20) \[ \text{Corr}(f, f^*) = \frac{\Lambda^{-1} \Lambda'}{\sigma_f \sqrt{\Lambda \Sigma^{-1} \Lambda'}} = (\Lambda \Sigma^{-1} \Lambda')^{1/2} \cdot 1 / \sigma_f, \]

\[ \text{Corr}^2(f, f^*) = \Lambda \Sigma^{-1} \Lambda' / \sigma_f^2. \]

Hence we have

(2.1.21) \[ \sigma_f^2 \geq \Lambda \Sigma^{-1} \Lambda', \]

and equality holds if and only if \( f = f(x_1, \eta) \) is a linear function of \( z_1, z_2, z_3 \) for all \((x_1, \eta) \in B^* \).

But from Lemma 2.1.3, \( \lambda_1 = g_\beta', \lambda_2 = g'_\alpha \) and \( \lambda_3 = g'_\beta \).

And \( \Sigma^{-1} \) is given by (2.1.16). Hence the inequality (2.1.21) is the same as (2.1.18). \( \square \)

Definition 2.1.2.

(i) For a given sampling plan, a non-constant estimator \( f = f(x_1, \eta) \) is said to be efficient at \((p_0, \alpha_0, \beta_0)\) if equality holds in (2.1.18) when \( p = p_0, \alpha = \alpha_0, \) and \( \beta = \beta_0 \).

The expected value \( E(f) = g(p, \alpha, \beta) \) is then said to be estimable efficiently at \((p_0, \alpha_0, \beta_0)\).

(ii) If \( f \) is efficient at \((p, \alpha, \beta)\) for all \( 0 < p, \alpha, \beta < 1, \)

then \( f \) is said to be efficient.

Corollary 2.1.1.

A non-constant estimator \( f = f(x_1, \eta) \) is efficient if and only if there exist functions \( a_1 = a_1(p, \alpha, \beta), \) \( i = 1, 2, 3, \) and \( b = b(p, \alpha, \beta), \) not all \( a_i \)'s zero, such that

(2.1.22) \[ f(x_1, \eta) = a_1(x_1 - p) + a_2(\bar{z}n_{11} - \alpha n_{10}) \]

\[ + a_3(\bar{z}n_{01} - \beta n_{00}) + b, \]
for all values of \((p, \alpha, \beta)\), \(0 < p, \alpha, \beta < 1\), and for all \((x_1, \eta) \in B^*\).

**Definition 2.1.3.** A sampling plan is said to be efficient if it admits a non-constant efficient estimator.

2.2. **Efficient sampling plans for binomial populations.**

In this section, we briefly describe the results obtained by DeGroot (1959) for binomial populations. The main result of interest here is that the only efficient sampling plans are the single (binomial) sampling plans and the inverse binomial sampling plans.

Suppose that independent observations are to be taken sequentially on the random variable \(U\) where

\[
\begin{align*}
P(U = 1) &= \theta, \quad 0 < \theta < 1, \\
P(U = 0) &= 1 - \theta.
\end{align*}
\]

Suppose we represent a sequential sample as a path in the two dimensional Euclidean plane, the path starting at the origin and being extended at a given stage one unit in either the horizontal or vertical direction according as the observed result at that stage is 1 or 0.

We will be using the following terms.

1. By a **point** we mean a point \(Y\) in the plane whose coordinates \(X(Y)\) and \(Y(Y)\) are non-negative integers.

2. A **sampling plan** is a function \(S\) defined on the points \(Y\) and taking only values 1 and 0, and for the original point \(Y_0\) with \(X(Y_0) = Y(Y_0) = 0, S(Y_0) = 1\). That is, if \(S(Y) = 1\), sampling is continued, and if \(S(Y) = 0\), sampling is stopped.
(3) A path to $\gamma$ is a sequence of points $\gamma_0, \gamma_1, \ldots, \gamma_n = \gamma$ such that $S(\gamma_k) = 1$ for $k = 0, \ldots, n - 1$, and either

$$X(\gamma_{k+1}) = X(\gamma_k) + 1$$

(2.2.2) and

$$Y(\gamma_{k+1}) = Y(\gamma_k)$$

or

$$X(\gamma_{k+1}) = X(\gamma_k)$$

(2.2.3) and

$$Y(\gamma_{k+1}) = Y(\gamma_k) + 1.$$

(4) A point $\gamma$ is a boundary point (or a continuation point) if there exists a path to $\gamma$ and $S(\gamma) = 0$ (or $S(\gamma) = 1$).

(5) A point $\gamma$ is an inaccessible point if there does not exist a path to $\gamma$.

(6) The sample size $N(\gamma)$ of any point $\gamma$ is the sum of its coordinates $X(\gamma) + Y(\gamma)$.

If we let $S(\gamma) = 1$, for all inaccessible points $\gamma$, then a sampling plan $S$ is completely determined by the set $B$ of all boundary points.

The probability of reaching a particular boundary point $\gamma$ is

$$k(\gamma) \Theta^X(\gamma) \Theta Y(\gamma)$$

(2.2.4)

where $k(\gamma)$ is the number of distinct paths to $\gamma$.

As in Definition 2.1.1, a sampling plan will be said to be closed if

$$\sum_{\gamma \in B} k(\gamma) \Theta^X(\gamma) \Theta Y(\gamma) = 1$$

(2.2.5)

and the terms "efficient estimator", and "efficient sampling plan"
are as defined in Definition 2.1.2 and Definition 2.1.3.

Corresponding to Theorem 2.1.1 we have following well-known information inequality for the binomial case.

Lemma 2.2.1. For any unbiased estimator $f$ of $g(\theta)$,

$$\sigma_f^2 = \text{Var}(f) \geq \frac{\theta \overline{e}[g'(\theta)]^2}{\text{EN}}$$

and equality holds at a value $\theta_0$ of $\theta$ if and only if there exist constants $a$ and $b$ such that

$$f(\gamma) = a[\theta_0 X(\gamma) - \theta_0 Y(\gamma)] + b, \quad \forall \gamma \in B.$$  

Recall that a single (binomial) sampling plan is a plan for which

$B = \{\gamma: X(\gamma) + Y(\gamma) = c\}$ for some positive integer $c$ and an inverse binomial sampling plan is a plan for which $B=\{\gamma: X(\gamma) = c\}$ or $B = \{\gamma: Y(\gamma) = c\}.$

Lemma 2.2.2. (DeGroot).

Let $S$ be a sampling plan for which there exists a non-constant estimator $f$ that is efficient at two values of $\theta$, then there exist constants $\mu$, $\nu$ and $\zeta$, $\mu$ and $\nu$ not both zero, such that

$$\mu X(\gamma) + \nu Y(\gamma) = \zeta, \quad \forall \gamma \in B.$$  

Clearly the boundary points of single sampling plans or inverse binomial sampling plans lie on straight lines as the lemma demands. Following theorem shows that they are the only such plans.

Note: We describe a more detailed proof of the theorem than that given by DeGroot (1959), since it is used frequently in our results later.

Theorem 2.2.1. (DeGroot).

Let $S$ be a given closed sampling plan for which there exist constants $\mu$ and $\nu$, not both zero, and $\zeta$ such that

$$\mu X(\gamma) + \nu Y(\gamma) = \zeta, \quad \forall \gamma \in B.$$  

Then $S$ is either a single sampling
plan or an inverse binomial sampling plan.

Proof: Without loss of generality we assume $\nu \geq 0$. We divide the proof into several cases.

(i) $\mu = \nu$

If $\zeta/\mu$ is not a positive integer, then $B = \emptyset$ and $S$ is not closed. If $\zeta/\mu = c$, a positive integer, then $B$ is a subset of $B^* = \{ \gamma : X(\gamma) + Y(\gamma) = c \}$ which is the boundary set of the single sampling plan. And if $S$ is closed, then $B = B^*$.

(ii) $\mu = 0$

If $\zeta/\nu$ is not a positive integer, then $B = \emptyset$ and $S$ is not closed.

If $\zeta/\nu = c$, a positive integer, then $B$ is a subset of $B^* = \{ \gamma : Y(\gamma) = c \}$ which is the boundary set of the inverse binomial sampling plan. If $S$ is closed, then $B = B^*$.

(iii) $\nu = 0$

As in (ii), $B$ is of the form $B = \{ \gamma : X(\gamma) = c \}$

(iv) $\mu > 0$, $\nu > 0$, $\mu \parallel \nu$

If $\zeta \leq 0$, obviously $B = \emptyset$

if $\zeta > 0$, there exists a point $\gamma$ such that

$$\mu X(\gamma) + \nu Y(\gamma) < \zeta$$

and either

$$\mu[X(\gamma) + 1] + \nu Y(\gamma) > \zeta$$

or

$$\mu X(\gamma) + \nu[Y(\gamma) + 1] > \zeta$$

thus, passing the boundary with positive probability, and $S$ is not closed.
(v) \( \mu < 0, \nu > 0 \)

Suppose \( \zeta > 0 \), then by the strong law of large numbers, given \( \epsilon > 0 \), there exists a constant \( n_0 \) such that with positive probability

\[
\left| \frac{X(y_n)}{n} - \theta \right| < \epsilon \quad \forall n \geq n_0
\]

for any path \( y_0, y_1, \ldots, y_n, \ldots \).

Thus,

\[
\theta X(y_n) - \theta Y(y_n) > - \epsilon n \quad \forall n \geq n_0.
\]

To show this inequality for \( y_1, \ldots, y_{n_0-1} \), consider a path with all 1's. Then

\[
X(y_k) = k, \quad Y(y_k) = 0 \quad \forall k = 1, \ldots, n_0-1,
\]

and this event has positive probability \( (\theta^k > 0) \). Then, with positive probability

\[
\theta X(y_k) - \theta Y(y_k) = -\epsilon k \quad \forall k = 1, \ldots, n_0-1
\]

so that given \( \epsilon > 0 \), there is a positive probability that

\[
(2.2.10) \quad \theta X(y) - \theta Y(y) > -\epsilon N(y)
\]

for every point \( y \) reached by the sample path. But for \( \theta \) sufficiently large and \( \epsilon \) sufficiently small, the line

\[
\theta X(y) - \theta Y(y) = -\epsilon N(y)
\]

lies entirely on the right of the line

\[
\mu X(y) + \nu Y(y) = \zeta. \quad \text{Hence no point satisfying the inequality (2.2.10) can be a boundary point, and there is a positive probability that the boundary will not be reached. Thus, \( S \) is not closed. Analogous arguments hold for \( \zeta < 0 \) or \( \zeta = 0 \).}
\]

\textbf{Theorem 2.2.2.} (DeGroot)

The only efficient sampling plans are single (binomial) sampling
plans and inverse binomial sampling plans.

(i) When $B = \{ \gamma: X(\gamma) + Y(\gamma) = c \}$, any non-constant function of the form $aX + b$ is an efficient estimator of $ac\theta + b$ and these are the only efficient estimators.

(ii) When $B = \{ \gamma: Y(\gamma) = c \}$, any non-constant function of the form $aN + b$ is an efficient estimator of $ac/\theta + b$, and these are the only efficient estimators.

(iii) When $B = \{ \gamma: X(\gamma) = c \}$, any non-constant function of the form $aN + b$ is an efficient estimator of $ac/\theta + b$, and these are the only efficient estimators.

**Corollary 2.2.1.** A non-constant estimator is efficient if and only if it is efficient at two distinct values of $\theta$.

DeGroot (1959) also gives following result on the relation between sampling plans and estimable functions which we state without proof.

**Theorem 2.2.3. (DeGroot)**

For a given sampling plan, a non-constant function $g(\theta)$ is estimable efficiently at $\theta_0$ if and only if there exist constants $a$ and $b$, $a \neq 0$ such that

$$g(\theta) = a(\theta - \theta_0)E_{\theta_0}N + b$$

or alternatively, if and only if there exists a constant $k \neq 0$ such that

$$E_{\theta_0}N = k \frac{g(\theta) - g(\theta_0)}{\theta - \theta_0}, \quad \theta \neq \theta_0, \quad E_{\theta_0}N = k g'(\theta_0).$$

### 2.3 Efficient sampling plans for functions of one parameter

In this section, we find efficient sampling plans for functions of $\alpha$ (or $\beta$) alone, say $g(\alpha)$ (or $g(\beta)$). The results are similar to those
of section 2.2. Specifically we show that the only efficient sampling plans for some \( g(\alpha) \) are plans \( S(N_{1}) \), \( S(N_{11}) \), and \( S(N_{10}) \), and the only functions estimable efficiently are \( g(\alpha) = \alpha, \frac{1}{\alpha}, \) and \( \frac{1}{\alpha} \)

where

\[(2.3.1) \quad S(Z) = \text{set of all sampling plans for which the random variable } Z \text{ is some constant.} \]

If we let

\[(2.3.2) \quad S_{c}(Z) = \text{sampling plan in which observations are continued until we get exactly } c \text{ number of } Z'\text{s and stop, then} \]

\[S_{c}(Z) \in S(Z) \]

For example, it will be seen in chapter III, that in \( S_{c}(N_{1}) \) we have \( N_{1} = c \), but in \( S_{c}(X) \), we have \( N_{1} = c - 1 \).

Therefore, \( S_{c}(N_{1}) \), \( S_{c}(X) \in S(N_{1}). \) But \( S_{c}(N_{1}) \notin S(X) \).

Now, we want to find efficient sampling plans for functions of \( \alpha \) alone, say \( g(\alpha) \). From Theorem 2.1.1, for any unbiased estimator \( f \) of \( g(\alpha) \) we have

\[(2.3.3) \quad \sigma_{f}^{2} = \text{Var}(f) \geq \frac{\alpha \alpha [g'(\alpha)]^{2}}{E_p, \alpha, \beta(N_{1})}. \]

By corollary 2.1.1, a non-constant estimator \( f \) is efficient for \( g(\alpha) \) if and only if there exist functions \( a = a(p, \alpha, \beta) \) \( b=b(\alpha) \) such that

\[(2.3.4) \quad f = f(n_{11}, n_{10}) = a(\alpha n_{11} - \alpha n_{10}) + b \]

for all \( (n_{11}, n_{10}) \in B_{1} \)

where \( B_{1} = \{(n_{11}, n_{10}) : n \in B\} \).

Remark. In Theorem 2.3.3, we will show that for \( g(\alpha) \) to be
efficiently estimable, \(a = a(p, \alpha, \beta)\) should be such that

\[ a(p, \alpha, \beta) \cdot E_{p, \alpha, \beta}(N_1) \]

is a function of \(\alpha\) alone.

Analogous to Lemma 2.2.2, we have

**Lemma 2.3.1.** Let \(S\) be a given sampling plan for which there exists a non-constant estimator \(f\) of \(g(\alpha)\) which is efficient at two values of \(\alpha\). Then there exist constants \(\mu\) and \(\nu\), not both zero, and \(\zeta\) such that

\[ \mu n_{11} + \nu n_{10} = \zeta \]

for all \((n_{11}, n_{10}) \in B_1\).

**Proof:** Suppose \(f\) is efficient at \(\alpha_0\) and \(\alpha_1\).

Then there exist constants \(a_0 = a(p, \alpha_0, \beta)\), \(a_1 = a(p, \alpha_1, \beta)\), \(b_0 = b(\alpha_0)\) and \(b_1 = b(\alpha_1)\) such that

\[ f = f(n_{11}, n_{10}) = a_0[ a_{00} n_{11} - a_{01} n_{10} ] + b_0 = a_1[ a_{10} n_{11} - a_{11} n_{10} ] + b_1 \]

for all \((n_{11}, n_{10}) \in B_1\).

Therefore

\[ (a_0 \alpha_0 - a_1 \alpha_1) n_{11} + (a_1 \alpha_1 - a_0 \alpha_0) n_{10} = b_1 - b_0 \]

and since \(f\) is not a constant neither \(a_0\) nor \(a_1\) is 0.

Thus both coefficients of \(n_{11}\) and \(n_{10}\) cannot be zero and (2.3.7) has the required form.

Analogous to Theorem 2.2.2 we have following.

**Theorem 2.3.1.** Let \(S\) be a given closed sampling plan with

\[ P(N_1 = 0) = 0 \]

for which there exist constants \(\mu\) and \(\nu\), not both zero, and \(\zeta\) such that

\[ \mu n_{11} + \nu n_{10} = \zeta \]

for all \((n_{11}, n_{10}) \in B_1\).
Then \( S \) is one of the sampling plans \( S(N_1), S(N_{11}) \) or \( S(N_{10}) \).

**Proof.** Define a sequence of random variables \( \{M_k, k \geq 1\} \) as follows:

\[
M_1 = \min \{t : X_t = 1\}
\]

(2.3.8)

\[
M_k = \min \{t : X_t = 1 \text{ and } t > M_{k-1}\}, k = 2, 3, \ldots
\]

Let

(2.3.9)

\[
Z_k = \begin{cases} 
1 & \text{if } M_{k+1} = M_k + 1 \\
0 & \text{if } M_{k+1} > M_k + 1
\end{cases}
\]

Then we have

(2.3.10)

\[
P(Z_k = 1) = P(M_{k+1} = M_k + 1)
\]

\[
= \sum \left[ P(M_{k+1} = M_k + 1 \mid M_k = m_k) \cdot P(M_k = m_k) \right]
\]

\[
= \sum \alpha \cdot p(M_k = m_k)
\]

\[
= \alpha
\]

Thus \( Z_k \)'s are identically distributed with

(2.3.11)

\[
P(Z_1 = 1) = \alpha
\]

\[
P(Z_1 = 0) = \overline{\alpha}
\]

To show the \( Z_k \)'s are independent, it is sufficient to show that \( Z_1 \) and \( Z_2 \) are independent. First we note that by the Markovian property,

\[
P[M_k = m_k \mid M_{k-1} = m_{k-1}, M_{k-2} = m_{k-2}, \ldots] = P[M_k = m_k \mid M_{k-1} = m_{k-1}]
\]

and
\[ P[M_3 = m_1 + 2, M_2 = m_1 + 1 | M_1 = m_1] \]

\[ = \frac{P[N_3 = m_1 + 2, M_2 = m_1 + 1, M_1 = m_1]}{P[M_1 = m_1]} \]

\[ = \frac{P[M_3 = m_1 + 2 | M_2 = m_1 + 1, M_1 = m_1] \cdot P[M_2 = m_1 + 1, M_1 = m_1]}{P[M_1 = m_1]} \]

\[ = P[M_3 = m_1 + 2 | M_2 = m_1 + 1] \cdot P[M_2 = m_1 + 1 | M_1 = m_1] \]

\[ = \alpha^2 , \]

so that

\[ P[Z_2 = 1, Z = 1] = P[M_3 = M_2 + 1, M_2 = M_1 + 1] \]

\[ = P[M_3 = M_1 + 2, M_2 = M_1 + 1] \]

\[ = \sum_{m_1} P[M_3 = m_1 + 2, M_2 = m_1 + 1 | M_1 = m_1] \cdot P[M_1 = m_1] \]

\[ = \alpha^2 \sum_{m_1} \{M_1 = m_1\} \]

\[ = \alpha^2 . \]

Also, we have

\[ P(Z_2 = 1, Z_1 = 0) = P[M_3 = M_2 + 1, M_2 > M_1 + 1] \]

\[ = \sum_{j=1} P[M_3 = M_2 + 1, M_2 = M_1 + 1 + j] \]

\[ = \sum_{m_1} \sum_{j=1} P[M_3 = m_1 + 2 + j, M_2 = m_1 + 1 + j | M_1 = m_1] \cdot P[M_1 = m_1] \]

\[ = \alpha \sum_{m_1} P[M_1 = m_1] \]

\[ = \alpha \bar{\alpha} , \]
since
\[
\sum_{j=1}^{\infty} \frac{P\{M_3 = m_1 + 2 + j, M_1 = m_1\}}{P\{M_1 = m_1\}}
\]
\[
= \sum_{j=1}^{\infty} P\{M_3 = m_1 + 2 + j, M_2 = m_1 + 1 + j, M_1 = m_1\}
\]
\[
= \alpha \sum_{j=1}^{\infty} P\{M_2 = m_1 + 1 + j | M_1 = m_1\}
\]
\[
= \alpha \sum_{j=1}^{\infty} \{M_2 > m_1 + 1 | M_1 = m_1\}
\]
\[
= \alpha \bar{\alpha}.
\]

But, by (2.3.11) we have
\[
P\{Z_2 = 1 | Z_1 = 1\} = \alpha^2
\]
\[
P\{Z_2 = 1 | Z_1 = 0\} = \alpha \bar{\alpha}.
\]

Therefore \(Z_k\) are independently and identically distributed Bernoulli random variables with \(P\{Z_1 = 1\} = \alpha, P\{Z_1 = 0\} = \bar{\alpha}\).

Now, we can write
\[
N_{11} = \sum_{k=1}^{\infty} Z_k
\]
\[
N_{10} = N_1 - N_{11}
\]

and by this representation we have reduced the problem to the binomial

Theorem 2.2.1. The proof then follows with \(X = N_{11}\) and \(Y = N_{10}\) respectively. □

Remark 2.3.1: We exclude any sampling plan with \(P(N_1 = 0) > 0\), that is, any sampling plan whose stopping set \(B\) contains points of the form \((0, 0, \cdots, \cdots)\), since it is clear that there will be no estimable function \(g(\alpha)\) for such a sampling plan. For example in the fixed
sample size plan, with \( N = n, \, P(N_1 = 0) = P(X = 0) = \sqrt{p^{n-1}} > 0 \).

**Theorem 2.3.2.** The only efficient sampling plans for some \( g(\alpha) \) are the sampling plans \( S(N_{10}), S(N_{11}) \) and \( S(N_{10}) \). Further

(i) If \( N_{11} + N_{10} = c \), any non-constant function of the form \( a N_{11} + b \) is an efficient estimator of \( a\alpha + b \), and these are the only efficient estimators.

(ii) If \( N_{11} = c \), any non-constant function of the form \( a N_{11} + b \) is an efficient estimator of \( \frac{a}{\alpha} + b \) and these are the only efficient estimators.

(iii) If \( N_{10} = c \), any non-constant function of the form \( a N_{11} + b \) is an efficient estimator of \( \frac{a}{\alpha} + b \) and these are the only efficient estimators.

**Proof.** From Lemma 2.3.1 and Theorem 2.3.1, we see that the only sampling plans which could be efficient are of the above forms. Now we show that these plans are efficient.

1) If \( N_{11} + N_{10} = c \) for all \( N \in B \), we have,

\[
\alpha N_{11} - \alpha N_{10} = \alpha N_{11} - \alpha (c - N_{11})
\]

\[
= N_{11} - c \alpha .
\]

Hence

\[(2.3.13) \quad N_{11} = [\alpha N_{11} - \alpha N_{10}] + c \alpha .\]

\( N_{11} \) has the required form of (2.1.22) and hence efficient for \( EN_{11} = \alpha \).

(ii) If \( N_{11} = c \), \( \forall N \in B \), we have

\[
\alpha N_{11} - \alpha N_{10} = \alpha c - \alpha (N_1. - c) = c - \alpha N_{11} .
\]

So,

\[(2.3.14) \quad N_{11} = - \frac{1}{\alpha} [\alpha N_{11} - \alpha N_{10}] + \frac{c}{\alpha} .\]

Thus, \( n_{11} \) is efficient for \( EN_{11} = \frac{c}{\alpha} .\).
(iii) If $N_{10} = c$, $\forall N = n \in B$, we have,
\[
\overline{N}_{11} - \overline{N}_{10} = \alpha(N_1, c) - \alpha = \overline{N}_{10} - c.
\]
So
\[(2.3.15) \quad N_1 = \frac{1}{\alpha} [\overline{N}_{11} - \overline{N}_{10}] + \frac{c}{\alpha}\]
Thus, $N_1$ is efficient for $E N_1 = \frac{c}{\alpha}$.

Corollary 2.3.1. A non-constant estimator is efficient for $g(\alpha)$ if and only if it is efficient at two distinct values of $\alpha$.

Theorem 2.3.3. For a given sampling plan, a non-constant function $g(\alpha)$ of $\alpha$ alone is estimable efficiently at $\alpha_0$ if and only if there exist constants $a$ and $b$, $a \not\equiv 0$, such that
\[(2.3.16) \quad g(\alpha) = a(\alpha - \alpha_0) E(\overline{N}_1) + b\]
or equivalently there exist $k \not\equiv 0$, such that
\[(2.3.17) \quad E(\alpha \overline{N}_{11} - \alpha_0 \overline{N}_{10}) = k(\overline{N}_{11} + \overline{N}_{10}) - \alpha_0(N_{11} + N_{10}) + (\alpha - \alpha_0) E(\overline{N}_1)\]
Proof. $E(\alpha \overline{N}_{11} - \alpha_0 \overline{N}_{10}) = a(\alpha - \alpha_0) E(\overline{N}_1) + c. \quad a \not\equiv 0$.

If $g(\alpha)$ is estimable efficiently at $\alpha_0$, then $g(\alpha)$ must be a constant or expectation of a function of the form
\[(2.3.18) \quad f = a[\overline{\alpha} \overline{N}_{11} - \overline{\alpha_0} \overline{N}_{10}] + b \quad a \not\equiv 0.\]
Therefore,
\[(2.3.19) \quad g(\alpha) = Ef = a(\alpha - \alpha_0) E(\overline{N}_1) + b.\]
Letting $\alpha = \alpha_0$, we get $b = g(\alpha_0)$. Differentiating both sides of
(iii) If \( N_{10} = c, \forall n = n \in B \), we have,
\[
\bar{a}N_{11} - aN_{10} - \bar{a}(N_1 - c) - ac = \bar{a}N_1 - c.
\]

So
\[
(2.3.15) \quad N_1 = \frac{1}{\bar{a}} \left[ \bar{a}N_{11} - \bar{a}N_{10} \right] + \frac{c}{\bar{a}}
\]

Thus, \( N_1 \) is efficient for \( E N_1 = \frac{c}{\bar{a}} \). \( \square \)

Corollary 2.3.1. A non-constant estimator is efficient for \( g(\alpha) \) if and only if it is efficient at two distinct values of \( \alpha \).

Theorem 2.3.3. For a given sampling plan, a non-constant function \( g(\alpha) \) of \( \alpha \) alone is estimable efficiently at \( \alpha_0 \) if and only if there exist constants \( a \) and \( b \), \( a \neq 0 \), such that
\[
(2.3.16) \quad g(\alpha) = a(\alpha - \alpha_0) E_\alpha(N_1) + b
\]
or equivalently there exists a constant \( k \), \( k \neq 0 \), such that
\[
(2.3.17) \quad E_\alpha(N_1) = \frac{g(\alpha) - g(\alpha_0)}{\alpha - \alpha_0}, \quad \alpha \neq \alpha_0
\]
\[
E_{\alpha_0}(N_1) = k \cdot g'(\alpha_0).
\]

Proof. \( E_\alpha(\bar{a}N_{11} - \alpha_0N_{10}) = E_\alpha(\bar{a}N_{11} - \alpha_0N_{10} + \alpha(N_{11} + N_{10}) - \alpha_0(N_{11} + N_{10})) = E_\alpha(\bar{a}N_{11} - \alpha_0N_{10}) + (\alpha - \alpha_0) E_\alpha(N_1) = (\alpha - \alpha_0) E_\alpha(N_1) \)

If \( g(\alpha) \) is estimable efficiently at \( \alpha_0 \), then \( g(\alpha) \) must be a constant or expectation of a function of the form
\[
(2.3.18) \quad f = a[\bar{a}N_{11} - \alpha_0N_{10}] + b \quad a \neq 0.
\]

Therefore,
\[
(2.3.19) \quad g(\alpha) = Ef = a(\alpha - \alpha_0) E_\alpha(N_1) + b.
\]

Letting \( \alpha = \alpha_0 \), we get \( b = g(\alpha_0) \). Differentiating both sides of
(2.3.19) and substituting $\alpha = \alpha_0$, \[ g'(\alpha_0) = a E_{\alpha_0}(N_1). \]

Hence, \[
\frac{g(\alpha) - g(\alpha_0)}{\alpha - \alpha_0} = a E_{\alpha}(N_1), \quad \quad \alpha \neq \alpha_0
\]

\[ g'(\alpha_0) = a E_{\alpha_0}(N_1). \]

**Remark 2.3.2.** By reasons of symmetry results analogous to Theorem 2.3.2, Corollary 2.3.1, and Theorem 2.3.3 hold for a function $g(\beta)$ of $\beta$ alone, with $\alpha, N_{11}, N_{10}, N_1$ replaced by $\bar{\beta}, N_{00}, N_{01}$ and $N_0$, respectively.

2.4. **Efficient sampling plan for functions of two parameters**.

In Theorem 2.3.2, and Remark 2.3.2, it was seen first that the only efficient sampling plans for functions of type $g(\alpha)$ are sampling plans $S(N_{11}), S(N_{10}),$ and $S(N_{10})$. Secondly we showed that the only efficient sampling plans for estimating functions of the type $g(\beta)$ are sampling plans $S(N_{00}), S(N_{00})$ and $S(N_{01})$.

We now consider the problem of finding efficient sampling plans for functions of two parameters, say $g(\alpha, \beta)$. From the structure of any sampling distribution which, from (2.1.9), has the form

\[
(2.4.1) \quad k(n|x_1)p^{l-x_1}q^{x_1-n_1-n_0-n_{10}} = n_1^{n_1}n_{10}^{n_{10}}n_{01}^{n_{01}}n_{00}^{n_{00}},
\]

it is clear that any information about $\alpha$ comes solely from $(N_{11}, N_{10})$ and not from $(N_{01}, N_{00})$, and any information about $\beta$ comes from $(N_{01}, N_{00})$, and not from $(N_{11}, N_{10})$. Therefore, the necessary condition that a function $g(\alpha, \beta)$ be estimable efficiently is that the sampling plan satisfy both conditions above simultaneously. The nine
possible combinations are

\[
\begin{align*}
(N_1 = c, N_0 = c'), (N_1 = c, N_{00} = c'), (N_1 = c, N_{01} = c') \\
(N_{11} = c, N_0 = c'), (N_{11} = c, N_{00} = c'), (N_{11} = c, N_{01} = c') \\
(N_{10} = c, N_0 = c'), (N_{10} = c, N_{00} = c'), (N_{10} = c, N_{01} = c')
\end{align*}
\]

(2.4.2)

Out of these, the only possible one that a sampling plan can satisfy is \((N_{10} = c, N_{01} = c')\).

Now from Lemma 2.1.1,

\[
(2.4.3) \quad N_{10} - N_{01} = X_1 - X_N
\]

so that

\[
(2.4.4) \quad c - c' = X_1 - X_N.
\]

The only way \(X_1 - X_N\) can be constant is that \(X_1 = X_N = 0\), or 1.

In that case \(c = c'\).

That is, \(N_{10} = N_{01} = c\).

or equivalently

\[
(2.4.5) \quad N_{10} + N_{01} = 2c.
\]

We thus have the following theorem.

**Theorem 2.4.1.** The only efficient sampling plan for functions of \(\alpha, \beta\), \(g(\alpha, \beta)\) is \(S_{2c}(N_{10} + N_{01})\), and any non-constant function \(f\) of the form \(f = a N_1 + b N_0 + d\) is an efficient estimator of \(g(\alpha, \beta) = c(\frac{a}{\alpha} + \frac{b}{\beta}) + d\). These are the only efficient estimators. In particular,

\(f = N\) is efficient for \(c(\frac{1}{\alpha} + \frac{1}{\beta}) + 1\).

**Proof.** Above arguments show that the only sampling plan which could be efficient is \(S_{2c}(N_{10} + N_{01})\) for some positive integer \(c\). We show that this plan is indeed efficient.

Since \(N_{10} = N_{01} = c\), for all \(N = n \in B\),
\[
\alpha N_{11} - \alpha N_{10} = \alpha (N_1 - c) - \alpha c \\
= \alpha N_1 - c . \tag{2.4.6}
\]

Similarly
\[
N_0^* = -\frac{1}{\beta} \left[ \beta N_{01} - \beta N_{00} \right] + \frac{c}{\beta} . \tag{2.4.7}
\]

Hence
\[
f = aN_1^* + bN_0^* + d \\
= \frac{a}{\alpha} \left[ \alpha N_{11} - \alpha N_{10} \right] + \frac{(-b)}{\beta} \left[ \beta N_{01} - \beta N_{00} \right] + c \left( \frac{a}{\alpha} + \frac{b}{\beta} \right) + d 
\]

which is of the form (2.1.22).

Therefore,
\[
f \text{ is efficient for } Ef = c \left( \frac{a}{\alpha} + \frac{b}{\beta} \right) + d = g(x, \beta) . \tag{2.4.9} \quad \Box
\]

2.5 Completeness of efficient sampling plans

In previous sections it has been shown that the sampling plans 

\( S(N_1^*), S(N_{11}), S(N_{10}), S(N_0^*), S(N_{00}), \) and \( S(N_{01}) \) are efficient plans. In this section we investigate the completeness of these plans.

Definition 2.5.1. A sampling plan is said to be complete if for an estimator \( f = f(X_1, N) \), \( Ef = 0 \) for all \( 0 < p, \alpha, \beta < 1 \) implies

\( f(x_1, n) = 0 \) for all \( (x_1, n) \in B^* \).

Theorem 2.5.1. A sufficient condition for a sampling plan belonging to \( S(N_1^*), S(N_{11}), \) or \( S(N_{10}) \) to be complete is that

\( n_{10} - n_{01} = ax_1 + b \) for all \( (x_1, n) \in B^* \)

for some constants \( a \) and \( b \).

Proof. For any sampling plan, the joint distribution of \( (X_1, N) \) is given by (2.1.9). If \( n_{10} - n_{01} = ax_1 + b \), then,
Hence, under the sampling plan considered, by reparametrization if necessary, we can express (2.5.2) as a 3-parameter exponential form whose parameter space contains a 3-dimensional rectangle. This is a sufficient condition for completeness of exponential family of distributions. (Lehman (1959), p. 132).

It is well-known that if a sampling plan is complete, then there exists at most one unbiased estimator of a given function, say \( g(\alpha, \beta) \). Thus completeness of a sampling plan guarantees the uniqueness of unbiased estimator of \( g(\alpha, \beta) \), if it exists. It will be seen in later chapters that the sampling plans \( S_c(x), S_c(N_{11}), S_c(R), S_c(N_{10}), \) and \( S_{2c}(N_{10} + N_{01}) \) satisfy condition (2.5.1).
Chapter III

INFERENCE UNDER SAMPLING PLANS $S(N_{11} + N_{10})$

In Section 2.3 it was shown that the sampling plans $S(N_{11} + N_{10}) = S(N_{11})$, that is sampling plans for which $N_{11} = c$ for some positive integer $c$, are efficient for $g(\alpha) = \alpha$. Further $N_{11}/c$ is the only efficient estimator of $\alpha$. In this chapter we study various properties of the sampling distributions of the sampling plans $S_c(N_{11})$ and $S_c(X)$ and find certain optimal estimators of the parameters under these sampling plans. Notice that $S_c(N_{11})$, $S_c(X) \in S(N_{11})$.

Sampling plan $S_c(N_{11})$ requires that we continue taking observations until we get precisely $c$ transitions from successes. This can also be accomplished simply by continuing observations until we get $c$ successes or $X = c$ and then take one more observation which could result in $S$ or $F$.

Notice that $S_c(N_{11})$ is only slightly different from $S_c(X)$ in that $S_c(N_{11})$ takes one more observation than $S_c(X)$.

We first study the sampling plan $S_c(X)$.

3.1. Sampling plan $S_c(X)$

Sampling plan $S_c(X)$ is to continue observations on $\{X_t, t \geq 1\}$ until we get exactly $c$ successes. The distributions under this plan have been studied by Rustagi and Srivastava (1968) where $c$ and $N$ can be interpreted as the number of hits it takes for a gun to destroy a target, and the total number of shots required respectively.
They give the distribution, called the Markov-dependent firing distribution and obtain the probability generating function of \( N \).

Rustagi and Laitinen (1970) studied the problem of moment estimation of parameters \( \alpha \) and \( \beta \) of this distribution under the assumption that only \( N \), not the transition counts \( N_{ij} \), is observable. Narayana and Sathe (1961) studied a special case of this distribution where \( p = \alpha \) or \( p = \beta \) in connection with coin tossing game.

**3.1.1. Distributions under sampling plan \( S_c(X) \).**

Let \( R = R(X_1, \ldots, X_N) \) be a random variable denoting the number of runs of \( S \)'s until we get \( X = c, c \geq 2 \). We recall that \( Y = Y(X_1, \ldots, X_N), N = N(X_1, \ldots, X_N) \) are total number of \( F \)'s and total number of observations required, respectively, so that \( c + Y = N \).

We first derive the joint distribution of \( (X_1, R, Y) \). There are two kinds of realizations possible, starting with \( S \) (\( X_1 = 1 \)) or \( F \) (\( X_1 = 0 \)). They are of the following form:

\[
\begin{align*}
X_1 = 1 ; & \quad S \cdots S F \cdots F S \cdots S \cdots F \cdots F S \cdots S \\
X_1 = 0 ; & \quad F \cdots F S \cdots S F \cdots F \cdots F S \cdots S
\end{align*}
\]

where \( s_1 \) and \( f_1 \) denote the length of \( i \)th runs of \( S \)'s and \( F \)'s respectively. \( X = c = \sum s_1 \), \( Y = y = \sum f_1 \).

The following relations between \( N = (N_{11}, N_{10}, N_{01}, N_{00}) \) and \( (X_1, R, Y) \) hold.

\[
\begin{align*}
N_{11} &= C - R \\
N_{10} &= R - 1 \\
N_{01} &= R - X_1
\end{align*}
\]
The number of ways we can partition an integer \( c \) into \( r \) positive integers \( s_1, s_2, \ldots, s_r \) such that \( \sum_{i=1}^{r} s_i = c \) is \( \binom{c-1}{r-1} \), and the number of ways we can partition an integer \( y \) into \( r - x_1 \) integers \( f_1, f_2, \ldots, f_{r-x_1} \) such that \( \sum_{i=1}^{r-x_1} f_i = y \) is \( \binom{y-1}{r-x_1-1} \). [Feller (1968) p. 38]

Hence total number of ways we can have \( c \) successes and \( y \) failures with \( r \) runs of \( S' \)'s is \( \binom{c-1}{r-1}\binom{y-1}{r-x_1-1} \).

Using the convention that \( \binom{n}{-1} = 1 \) and \( \binom{n}{-l} = 0 \quad \forall \, n \geq 0 \), the joint distribution of \( (X, R, Y) \) is

\[
(3.1.2) \quad p(x, r, y) = P(X = x_1, R = r, Y = y) = \binom{c-1}{r-1}\binom{y-1}{r-x_1-1} p_1^{x_1} \alpha^c \beta^{r-x_1} y^{r-x_1-1} q_0^{x_1-1-x_1} q_1^{x_1},
\]

where

\[
\begin{align*}
x_1 &= 0, 1 \\
r &= 1, \ldots, c \\
y &= r-x_1, r-x_1+1, \ldots
\end{align*}
\]

In terms of \( \mathbf{N} = (N_{11}, N_{10}, N_{01}, N_{00}) \), the distribution is

\[
(3.1.3) \quad p(x_1, n) = P(X_1 = x_1, \mathbf{N} = n) = \binom{n_0+n_{11}-1}{n_{10}} \binom{n_{10}}{n_{01}-1} p_1^{x_1} \alpha^{n_{11}} \beta^{n_{01}-n_{00}} q_0^{n_{10}-n_{11}-n_{01}+n_{00}},
\]

where

\[
\begin{align*}
n_{11} + n_{10} &= c - 1 \\
n_{01} &= n_{10} + 1 - x_1 \\
n_{00} &= 0, 1, 2, \ldots
\end{align*}
\]
The distribution of \( N = c + y \) is given by, using \((3.1.2)\),

\[
p(n) = P(N = n) = \sum_{x_1=0}^{\min(c,y)} \binom{c-1}{x_1} \binom{n-c-1}{r-1} \frac{p^n q^{n-1-1}}{r-1} \alpha^{c-r} \beta^{r-x_1} n-c-r+x_1
\]

It is interesting to note that the above is given in a different and more compact form than that given by Rustagi and Srivastava (1968).

The probability generating function of \( N \) can be directly obtained from \((3.1.2)\) with \( N = c + Y \) and is given by

\[
G_N(t) = \left( \frac{pt + \alpha t^2}{1 - \alpha t} \right) \left( \frac{\alpha t + \beta t^2}{1 - \beta t} \right)^{c-1}.
\]

We have then,

\[
E[N] = \frac{1}{\beta} [c(\alpha + \beta) + \alpha - p] = c + \alpha Y
\]

\[
\text{Var}(N) = \frac{1}{\beta^2} [c(\beta - \alpha^2 + \alpha \beta) + \alpha^2 - \alpha \beta - \beta^2 + p \beta] = \text{Var}(Y).
\]

The marginal distribution of \((X_1, R)\) is

\[
p(X_1, R) = P(X_1 = x_1, R = r)
\]

\[
= \sum_{y=r-x_1}^\infty p(x, r, y)
\]

\[
= \left[ \binom{c-1}{r-1} \alpha^{r-1} \beta^{c-r} \right] \frac{x_1^{l-x_1}}{p^l q^{x_1}}
\]

\[
= P(R = r) P(X_1 = x_1).
\]

It shows that \( R-1 \) is distributed binomially with parameters \( c-1 \) and \( \overline{\alpha} \), that is

\[
R - 1 \sim \text{b}(c-1, \overline{\alpha}).
\]

Notice that the distribution is independent of \( X_1 \).

We have,
(3.1.8) \[ ER = c\alpha + \alpha \]
\[ \text{Var}(R) = (c-1)c\alpha. \]

Distribution of \( N_{1j} \)

1) Using the distribution (3.1.6) of \( R \) and relation (3.1.1), i.e. \( N_{1l} = c - R, N_{10} = R - 1 \), the distributions of \( N_{1l}, N_{10} \) are seen to be

\[
\begin{align*}
\{ p(n_{1l}) = P(N_{1l} = n_{1l}) = \binom{c-1}{n_{1l}} \alpha^{n_{1l}} c^{-1-n_{1l}} & \quad n_{1l} = 0, 1, \ldots, c-1 \\
p(n_{10}) = P(N_{10} = n_{10}) = \binom{c-1}{n_{10}} \alpha^{n_{10}} c^{-1-n_{10}} & \quad n_{10} = 0, 1, \ldots, c-1.
\end{align*}
\]

That is,

\[\{ \begin{align*} N_{1l} & \sim b(c-1, \alpha) \\
N_{10} & \sim b(c-1, \alpha). \end{align*}\]

This can also be deduced from the proof of Theorem 2.3.1.

Since \( R \) and \( X_{1} \) are independent and

ii) \[ N_{0l} = R - X_{1} = \begin{cases} R - 1 & \text{if } X_{1} = 1 \\
R & \text{if } X_{1} = 0 \end{cases} \]

\[ P(N_{0l} = n_{0l} | X_{1} = 1) = \binom{c-1}{n_{0l}} \alpha^{n_{0l}} c^{-1-n_{0l}} n_{0l} = 0, 1, \ldots, c-1, \]
\[ P(N_{0l} = n_{0l} | X_{1} = 0) = \binom{c-1}{n_{0l-1}} \alpha^{n_{0l-1}} c^{-n_{0l}} n_{0l} = 1, 2, \ldots, c. \]

Hence we have,

(3.1.10) \[ p(n_{0l}) = P(N_{0l} = n_{0l}) \]
\[ = \begin{cases} p^{c-1} \alpha^{n_{0l}} c^{-1-n_{0l}} & \quad n_{0l} = 0 \\
p(n_{0l}) \alpha^{n_{0l}} c^{-1-n_{0l}} + q(n_{0l-1}) \alpha^{n_{0l-1}} c^{-n_{0l}} & \quad n_{0l} = 1, \ldots, c-1 \\
q \alpha^{c-1} & \quad n_{0l} = c. \end{cases} \]
It is interesting to note that the distribution of \( N_{01} \), transition count from \( F \) to \( S \), does not depend on \( \beta \).

iii) Marginal distribution of \( (X_1, R) \) is from (3.1.6),

\[
P(x_1, r) = \binom{c-1}{r-1} p_1^{r-1} q^{c-r} q^{-r-1}
\]

Thus, the conditional dist. of \( Y \) given \( X_1 = x_1, R = r \) is

\[
P(y|x_1, r) = \frac{p(x_1, r, y)}{p(x_1, r)} = \left( \frac{r-1}{r-x_1-1} \right) \beta^{r-x_1-1} (y-r-1)^{y-r+1}
\]

Thus, the conditional distribution of \( Y \) given \( X_1 = x_1, R = r \) depends on \( x_1, r \) only through \( n_{01} = r - x_1 \).

Hence

\[
P(Y = y | n_{01}) = \binom{y-1}{n_{01}-1} \beta^{n_{01}-1} (y-n_{01})^{y-n_{01}}
\]

Remark 3.1.1. We use the following notation.

\[
W \sim NB(m, \theta)
\]

implies that

\[
P(W = w) = \binom{w-1}{m-1} \theta^m (1 - \theta)^{w-m},
\]

\( w = m, m+1, \ldots \)

and

\[
V \sim NB(m, \theta)
\]

if

\[
P(V = v) = \binom{v+m-1}{m-1} \theta^m (1 - \theta)^v, \quad v = 0, 1, 2, \ldots
\]

The conditional probability distribution of \( Y \) given that
\( N_{01} = n_{01} \) is given by,
\[
Y|N_{01} = n_{01} \geq 1 \sim \text{NB}(n_{01}, \beta)
\]
(3.1.12) and
\[
Y = 0 \quad \text{if} \quad N_{01} = 0.
\]
Since \( Y = N_{01} + N_{00}, \)
\[
P(N_{00} = n_{00}|n_{01}) = \left(\frac{n_{00} + n_{01} - 1}{n_{01} - 1}\right)^{n_{01}} \beta^{n_{01}} \beta^{-n_{00}}
\]
so that,
\[
N_{00}|N_{01} = n_{01} \geq 1 \sim \text{NB}(n_{01}, \beta)
\]
and
\[
N_{00} = 0 \quad \text{if} \quad N_{01} = 0.
\]

3.1.2 Estimation of parameters \( \alpha, \beta \)

Since our sampling plan has the property that \( N_{1} = c - 1 \)
according to Theorem 2.3.2, we expect the plan to be efficient for \( g(\alpha) = \alpha \).

Indeed, the Maximum Likelihood estimator
\[
\hat{\alpha} = \frac{N_{11}}{N_{11} + N_{10}} = \frac{N_{11}}{c - 1} = \frac{c - R}{c - 1}
\]
is such that
\[
E\hat{\alpha} = \alpha, \quad \text{Var}(\hat{\alpha}) = \frac{\alpha}{c - 1}.
\]
(lower bound of (2.1.18) with \( g(\alpha) = \alpha \))

Remark 3.1.2. Consider the well-known problem of estimating parameter
in a negative binomial distribution.
Suppose \( W \sim \text{NB}(m, p) \).

Then it is well-known that, for \( m \geq 2 \),
\begin{equation}
\hat{p} = \frac{m-1}{W-1}
\end{equation}
is the unique unbiased estimator of \( p \) and
\begin{equation}
E_{\hat{p}}^2 = p \sum_{k=0}^{\infty} \binom{k+m-1}{m-1} q^k \quad \text{(Haldane (1945))}
\end{equation}
\begin{equation}
= \frac{(m-1)p^m}{q^{m-1}} \left[ (-1)^{m-1} \log p + \sum_{k=1}^{m-2} \frac{(-1)^{m-k}}{k} \left( \frac{q}{p} \right)^k \right] \quad \text{(DeGroot (1959))}.
\end{equation}

Haldane (1945) has also shown that the ML estimator
\begin{equation}
\hat{p}^* = \frac{m}{W} \quad (m \geq 1)
\end{equation}
has expected value
\begin{equation}
E_{\hat{p}^*} = p \sum_{k=0}^{\infty} \binom{k+m-1}{m} q^k,
\end{equation}
so that it has positive bias,
\begin{equation}
E(\hat{p}^* - p) = p \sum_{k=1}^{\infty} \binom{k+m-1}{m} q^k > 0.
\end{equation}

We consider the problem of estimating \( \beta \) under the plan \( S_e(X) \).
From the forms (3.1.2), (3.1.3), it is clear that any reasonable estimator of a function \( g(\beta) \) should depend on \( (N_{01}, N_{00}) \) or \( (X_1, R, Y) \).
But if the observations yield no \( F \)'s, and therefore no transitions from \( F \)'s, that is, \( Y = N_{01} + N_{00} = 0 \), and \( X_1 = 0, R = 1 \), we do not get any information about \( \beta \). Moreover since this event has positive probability, \( P(Y = 0) = p q^{c-1} > 0 \), it follows that there does not exist any estimable function \( g(\beta) \). In particular, there does not exist any unbiased estimator of \( \beta \). The maximum likelihood estimator...
of \( \beta \) is \( \frac{N_{01}}{N_{01} + N_{00}} \) for \( N_{01} + N_{00} > 0 \), denoted by \( \beta^* \).

Let

\[
\beta^* = \begin{cases} 
\frac{N_{01}}{N_{00}} = \frac{N_{01}}{Y} & \text{if } Y = y \geq 1 \\
\alpha & \text{if } Y = 0
\end{cases}
\]  

for some arbitrary \( 0 \leq \alpha \leq 1 \).

Let \( \hat{\beta} \) be another estimate of \( \beta \) defined as

\[
\hat{\beta} = \begin{cases} 
\frac{N_{01}-1}{N_{00}} = \frac{N_{01}}{Y-1} & \text{if } Y = y \geq 2 , \\
1 & \text{if } Y = 1 , \\
\alpha & \text{if } Y = 0 .
\end{cases}
\]

Then we have

**Theorem 3.1.1.** The bias in the estimates \( \hat{\beta} \) and \( \beta^* \) is given by

\[
E(\hat{\beta} - \beta) = (a - \beta)p_0^{c-1}
\]

\[
E(\beta^* - \beta) = (a-\beta)p_0^{c-1} + \sum_{n_{01}=1}^{\infty} \left( \frac{\beta^{n_{01}-1}}{\beta} \right)^k \cdot \sum_{k=1}^{n_{01}} \left( \frac{k+n_{01}}{n_{01}} \right) \left( \frac{k+n_{01}}{n_{01}} \right)^{\beta}
\]

where \( P(N_{01} = n_{01}) \) is given by (3.1.10).

**Remark 3.1.3.** The bias term \( (a - \beta)p_0^{c-1} \) comes from guessing \( \beta \) incorrectly when the observations yield no \( F \)'s, and therefore \( \beta \) has to be guessed. The second bias term of \( \beta^* \) is strictly positive, which implies that \( \beta^* \) has a tendency to overestimate \( \beta \).

**Proof of Theorem.**

(a) From (3.1.12), we have
\[
Y|_{N_{01}} = n_{01} \geq 2 \sim NB(n_{01}, \beta)
\]
\[
Y|_{N_{01}} = 1 \sim \beta \beta^{y-1} \quad y = 1, 2, \ldots
\]
\[
Y = 0 \quad \text{if} \quad N_{01} = 0.
\]
and from (3.1.22)
\[
\begin{cases}
\hat{\beta}|_{N_{01}} = n_{01} \geq 2 = \frac{n_{01}^{-1}}{\gamma - 1} & \text{if} \quad Y = \gamma \geq n_{01} \\
\hat{\beta}|_{N_{01}} = 1 & \text{if} \quad Y = 1 \\
\hat{\beta}|_{N_{01}} = 0 = a & \text{if} \quad Y = 0
\end{cases}
\]
so that
\[
E(\hat{\beta}|_{N_{01}} = n_{01} \geq 1) = \beta
\]
\[
E(\hat{\beta}|_{N_{01}} = 0) = a.
\]

Therefore,
\[
\hat{E} \beta = E(E(\hat{\beta}|_{N_{01}}))
\]
\[
= E(\hat{\beta}|_{N_{01}} = 0) P(N_{01} = 0) + \sum_{n_{01}=1}^{c} E(\hat{\beta}|_{N_{01}} = n_{01}) P(N_{01} = n_{01})
\]
\[
= a P(N_{01} = 0) + \beta \sum_{n_{01}=1}^{c} P(N_{01} = n_{01})
\]
\[
= a P(N_{01} = 0) + \beta(1 - P(N_{01} = 0))
\]
\[
= \beta + (a - \beta) P(N_{01} = 0)
\]
\[
= \beta + (a - \beta) p \alpha^{c-1}.
\]
(b) On the other hand, from (3.1.21),
\[
\begin{align*}
\beta^*_{N_{01}} = n_{01} \geq \frac{N_{01}}{Y}, & \quad \text{if } Y = Y \geq n_{01} \\
\beta^*_{N_{01}} = 0 = a & \quad \text{if } Y = 0 .
\end{align*}
\]

Thus, from (3.1.19)

\[
E(\beta^*|_{N_{01}} = n_{01} \geq 1) = \beta \sum_{k=0}^{\infty} \left( \frac{k+n_{01} - 1}{n_{01}} \right) \beta^* \kappa = \beta \sum_{k=1}^{\infty} \left( \frac{k+n_{01} - 1}{n_{01}} \right) \beta^* \kappa ,
\]

and

\[
E(\beta^*|_{N_{01}} = 0) = a .
\]

Therefore,

\[
E \beta^* = E(\beta^*|_{N_{01}})
\]

\[
= E(\beta^*|_{N_{01}} = 0)P(N_{01} = 0) + \sum_{n_{01}=1}^{c} E(\beta^*|_{N_{01}} = n_{01})P(N_{01} = n_{01})
\]

\[
= aP(N_{01} = 0) + \sum_{n_{01}=1}^{c} \left[ \beta + \beta \sum_{k=1}^{\infty} \left( \frac{k+n_{01} - 1}{n_{01}} \right) \beta^* \kappa \right]P(N_{01} = n_{01})
\]

\[
= aP(N_{01} = 0) + \beta(1-P(N_{01} = 0)) + \beta \sum_{n_{01}=1}^{c} \sum_{k=1}^{\infty} \left( \frac{k+n_{01} - 1}{n_{01}} \right) \beta^* \kappa P(N_{01} = n_{01})
\]

\[
= \beta + (a - \beta)P \alpha^{c-1} + \beta \sum_{n_{01}=1}^{c} P(N_{01} = n_{01}) \sum_{k=1}^{\infty} \left( \frac{k+n_{01} - 1}{n_{01}} \right) \beta^* \kappa
\]

The mean square error of \( \hat{\beta} \) is obtained by noting that

\[
E((\hat{\beta} - \beta)^2|_{N_{01}} = n_{01} \geq 2) = \text{Var}(\hat{\beta}|_{N_{01}} = n_{01} \geq 2)
\]

\[
= \beta^2 \sum_{k=0}^{\infty} \left( \frac{k+n_{01} - 1}{n_{01} - 1} \right)^{-1} - \beta^2
\]

\[
= \beta^2 \sum_{k=1}^{\infty} \left( \frac{k+n_{01} - 1}{n_{01} - 1} \right)^{-1}
\]

\[ 
E((\hat{\beta} - \beta)^2 | N_{01} = 1) = \text{var}(\hat{\beta} | N_{01} = 1) = \beta \bar{\beta} \\
E((\hat{\beta} - \beta)^2 | N_{01} = 0) = (a - \beta)^2 .
\]

Hence

\[ (3.1.25) \quad \text{MSE}(\hat{\beta}) = E(E[(\hat{\beta} - \beta)^2 | N_{01}]) \]

\[ = (a-\beta)^2 p(N_{01}=0) + \beta \bar{\beta} p(N_{01}=1) \]

\[ + \sum_{n_{01}=2}^{c} \beta^2 \sum_{k=1}^{\infty} \bar{\beta}^k \binom{k+n_{01}-1}{n_{01}-1} p(N_{01} = n_{01}) \]

\[ = (a-\beta)^2 p(N_{01}=0) + \beta \bar{\beta} p(N_{01}=1) \]

\[ + \beta^2 \sum_{n_{01}=2}^{c} p(N_{01} = n_{01}) \sum_{k=1}^{\infty} \bar{\beta}^k \binom{k+n_{01}-1}{n_{01}-1} . \]

3.1.3. Testing for Independence.

The joint distribution (3.1.2) of \((X, R, Y)\) can be rewritten as

\[ p(x, r, y) = \left(\frac{c-1}{r-1}(r-x-1)\frac{\theta_2}{\theta_3}\right)^{x} \left(\frac{\theta_1}{\theta_3}\right)^{r} \beta^y \left(\frac{\theta_1}{\alpha}\right) \]

\[ = h_c(x_1, r, y, e_1 e_2 e_3) e_1^{x} e_2^{y} e_3^{x_1 r + y} \]

\[ = h_c(x_1, r, y, \eta_1 \eta_2 \eta_3) e_1^{x_1 \eta_1 + y \eta_2 + y} \eta_3 \]

where

\[ h_c(x_1, r, y) = \left(\frac{c-1}{r-1}(r-x-1)\right), \]

\[ e_1 = \frac{p_\theta}{q_\theta} , \]

\[ e_2 = \frac{\bar{\alpha}_\beta}{\alpha_\beta} , \]
\[ \theta_3 = \bar{\theta}, \]
\[ \varphi(\theta_1, \theta_2, \theta_3) = \frac{q \alpha \alpha}{\bar{\alpha}} = \frac{(1 - \theta_3)^c}{\theta_2(\theta_1 + \theta_3 - \theta_1 \theta_3)(1 - \theta_3 + \theta_2 \theta_3)^{c-1}}, \]

and
\[ \eta_i = \log \theta_i, \quad i = 1, 2, 3. \]

It is easily seen that \((X_1, R, Y)\) form a sufficient statistic. Above form indicates that \((X_1, R, Y)\) is also complete.

Now,
\[ \alpha = \beta \Rightarrow \theta_2 = 1 \Rightarrow \eta_2 = 0. \]

We wish to test
\[ H_0 : \eta_2 = 0 \ (\alpha = \beta) \ vs \ \eta_2 \neq 0 \ (\alpha \neq \beta). \]

Under \(H_0\),
\[ P_{H_0}(x_1, r, y) = \frac{(c - 1)(y - 1)}{(r - 1)(r - x_1 - 1)} \frac{x_1^{1 - x_1} \alpha^{1 - \alpha}}{\alpha} \]
\[ = p q \frac{x_1^{1 - x_1} \alpha^{1 - \alpha}}{\alpha} \sum_{r=1}^{x_1} \frac{(c - 1)(y - 1)}{(r - 1)(r - x_1 - 1)} \]
\[ = \frac{(y + c - 2)}{(c - x_1 - 1)} \frac{x_1^{1 - x_1} \alpha^{1 - \alpha}}{\alpha} \]

Thus we have,
\[ (3.1.26) \quad P_{H_0}(r | x_1, y) = \frac{(c - 1)(y - 1)}{(r - 1)(r - x_1 - 1)} \frac{y + c - 2}{(c - x_1 - 1)} \]
\[ 1 \leq r \leq \text{max}(c, y + x_1) \]
\[ y = 1 - x_1, 2 - x_1, \ldots \]
\[ x_1 = 0, 1. \]

Then by a theorem in Lehman (1959) (p-136 Theorem 3), we have
Theorem 3.1.2. For testing $H_o: \alpha = \beta$ vs $\alpha \neq \beta$, there exists a

UMP unbiased level $\delta$ test given by

$$\phi(x_1, r, y) = \begin{cases} 
1 & r < c_1(x_1, y) \text{ or } r > c_2(x_1, y) \\
\gamma_i(x_1, y) & r = c_i(x_1, y), i = 1, 2 \\
0 & c_1(x_1, y) < r < c_2(x_1, y) 
\end{cases}$$

where $c_i$'s and $\gamma_i$'s are determined by

$$E_{H_0}(\phi(x_1, R, Y) | x_1, y) = \delta \quad \text{for all} \ (x_1, y)$$

$$E_{H_0}(R \phi(x_1, R, Y) | x_1, y) = \delta E_{H_0}(R | x_1, y) \quad \text{for all} \ (x_1, y).$$

Condition (3.1.28) is

$$\begin{align*}
\sum_{r=c_1+1}^{c_2-1} \sum_{i=1}^{2} P_{H_0}(r | x_1, y) + \sum_{i=1}^{2} (1 - \gamma_i) P_{H_0}(c_i | x_1, y) &= 1 - \delta. \\
\end{align*}$$

That is

$$\begin{align*}
\sum_{r=c_1+1}^{c_2-1} \sum_{i=1}^{2} (r - 1)(r - x_1 - 1) + \sum_{i=1}^{2} \gamma_i \sum_{i=1}^{2} c_i(1 - \gamma_i) (c_i | x_1 - 1) &= (1 - \delta)(c_2 - 1).
\end{align*}$$

Condition (3.1.29) is

$$\begin{align*}
\sum_{r=c_1+1}^{c_2-1} \sum_{i=1}^{2} r P_{H_0}(r | x_1, y) + \sum_{i=1}^{2} c_i(1 - \gamma_i) P_{H_0}(c_i | x_1, y) &= \delta E_{H_0}(R | x_1, y) \\
\end{align*}$$

That is

$$\begin{align*}
\sum_{r=c_1+1}^{c_2-1} \sum_{i=1}^{2} r(r - x_1 - 1)(r - x_1 - 1) + \sum_{i=1}^{2} c_i(1 - \gamma_i) (c_i | x_1 - 1) &= (1 - \delta)(c_2 - 1) E_{H_0}(R | x_1, y).
\end{align*}$$
where

\[
E(R-1|x_1, y) = \sum_{r=1}^{y+c-2} \frac{(r - 1)(y-c-1)(y-x_1-1)}{(c-x_1-1)}
\]

\[
= \frac{(c - 1)}{y+c-2} \sum_{r=2}^{y+c-2} \frac{(r - 2)(y-c-1)}{(c-x_1-1)}
\]

\[
= \frac{(c - 1)}{y+c-2} \sum_{j=1}^{y+c-2} \frac{(j - 1)(y-c-1)}{(c-x_1-1)}
\]

\[
= \frac{(c - 1)}{y+c-2} \frac{y+c-3}{(c-x_1-1)}
\]

Thus

(3.1.32) \[
E(R|x_1, y) = 1 + \frac{(c-1)(y-x_1-1)}{(c-x_1-1)}
\]

\[
= 1 + \frac{(c-1)(y-x_1-1)}{(y+c-2)}
\]

3.1.4 Special cases

Narayana and Sathe (1961) studied parameter estimation problem in the following coin tossing game.

Two coins 1 and 2 with probabilities \( \beta \) and \( \alpha \) for heads respectively are tossed as follows;

i) the first trial is made with coin 1

ii) the \( n \)th trial is made with coin 1 or 2 according as (n-1)th trial results in tail or head \( (n > 1) \)

iii) the trials are continued until we get exactly \( c \) heads.

This is a special case of our model where the initial probability \( p \) of success (head) is the same as \( \beta \), with the same sampling plan \( S_c(X) \),
In this case \((3.1.2)\) reduces to
\[
P(x_1, r, y) = \binom{c-1}{r-1}(r-x_1-1)^{y-1} \alpha^{c-r} \beta^{r} \bar{\beta}^{y-r+1},
\]
so that
\[
(3.1.33) \quad P(r, y) = \binom{c-1}{r-1}(r-1)^{y-1} \alpha^{c-r} \beta^{r} \bar{\beta}^{y-r+1}
\]
\[y = r-1, r, \ldots \]
\[r = 1, \ldots, c.\]

From this it is seen that \((R, Y)\) is a complete sufficient statistic, and they find MVUE's of \(\alpha, \beta\) as
\[
\hat{\alpha} = \binom{c-2}{r-1}/\binom{c-1}{r-1} = \frac{c-r}{c-1}
\]
\[
\hat{\beta} = \binom{y-1}{r-1}/\binom{y}{r-1} = \begin{cases} \frac{r-1}{y} & y \geq 1 \\ 1 & y = 0 \end{cases}
\]

Remark 3.1.4

(i) Here we do not have the kind of non-estimability of \(g(\beta)\) as in our original model since, in this model, initial distribution is \(P(X_1 = 1) = \beta, \ p(X_1 = 0) = \bar{\beta}\), and even if we get \(Y = 0\), (no \(F's\)) we have at least some information from first trial. (in that case we can estimate \(\beta\) to be \(1\))

(ii) Suppose the game is modified so that the first trial is made with coin 2. Then we have \(p = \alpha\).

Then the same difficulty arises as in our original model and there does not exist any estimable function \(g(\beta)\).

3.2. Sampling plan \(S_c(N_{11} + N_{10})\)

3.2.1 Distributions under sampling plan \(S_c(N_{11} + N_{10})\)
Thus

\[(3.1.32) \quad E(R|\mathbf{x}_1, y) = 1 + \frac{(c-1)(y+c-3)}{(y+c-2)} \left(\frac{1}{c-1} - 1\right)\]

\[= 1 + \frac{(c-1)(y+c-1)}{(y+c-2)}.\]

3.1.4 Special cases

Narayana and Sathe (1961) studied parameter estimation problem in the following coin tossing game.

Two coins 1 and 2 with probabilities \( \beta \) and \( \alpha \) for heads respectively are tossed as follows:

i) the first trial is made with coin 1

ii) the \( n \)th trial is made with coin 1 or 2 according as \( (n-1) \)th trial results in tail or head \((n > 1)\)

iii) the trials are continued until we get exactly \( c \) heads.

This is a special case of our model where the initial probability \( p \) of success (head) is the same as \( \beta \), with the same sampling plan \( S_c(X) \),
In this case (3.1.2) reduces to
\[ P(x_1,r,y) = \binom{c-1}{r-1}(r-x_1-1)\alpha^{c-r} r^{-1} \beta^{r-1} y^{-r+1}, \]
so that
\[ (3.1.33) \quad P(r,y) = \binom{c-1}{r-1}(r-y-1)\alpha^{c-r} r^{-1} \beta^{r-1} y^{-r+1} \]
for \( y = r-1, r, \ldots \)
for \( r = 1, \ldots, c. \)

From this it is seen that \((R, Y)\) is a complete sufficient statistic, and they find MNUE's of
\[ \alpha = \frac{\binom{c-2}{r-1}}{\binom{c-1}{r-1}} \]
\[ (3.1.34) \quad \beta = \frac{\binom{y-1}{r-1}}{\binom{y-1}{r-1}} \]

**Remark 3.1.4**

(i) Here we do not have the kind of indispensibility of \( g(\beta) \) as in our original model since, in this model, initial distribution is \( P(X_1 = 1) = \beta, \ p(X_1 = 0) = \bar{\beta} \), and even if we get \( Y = 0 \), (no F's) we have at least some information from first trial. (in that case we can estimate \( \beta \) to be 1)

(ii) Suppose the game is modified so that the first trial is made with coin 2. Then we have \( p = \alpha \).

Then the same difficulty arises as in our original model and there does not exist any estimable function \( g(\beta) \).

3.2. **Sampling plan** \( S_c(N_{11} + N_{10}) \)

3.2.1 **Distributions under sampling plan** \( S_c(N_{11} + N_{10}) \)
The only difference between $S_c(N_{11} + N_{10}) = S_c(N_{11})$ and $S_c(x)$ is that in $S_c(N_{11})$ we take one more observation after $c$th success whereas in $S_c(x)$ we stop sampling with $c$th success.

Let the random variables $R, Y, N$ be the same as in Section 3.1 and let $X', Y', N'$ be the r.v.'s denoting total numbers of success, failures and sample sizes respectively under $S_c(N_{11})$.

Then we have following relations.

\begin{align*}
X' &= c + Z_c \\
Y' &= Y + 1 - Z_c \\
N' &= X' + Y' = N + 1
\end{align*}

(3.2.1)

and

\begin{align*}
N_{11} &= c - R + Z_c \\
N_{10} &= (R-1) + (1-Z_c) = R - Z_c \\
N_{01} &= R - X_1 \\
N_{00} &= Y - R + X_1
\end{align*}

(3.2.2)

and

\begin{align*}
N_{11} &= c , \\
N_{01} &= y
\end{align*}

where the random variable $Z_c$ is defined as in Theorem 2.3.1.

The joint distribution of $(X_1, R, Y, Z_c)$ is

\begin{align*}
P(x_1, r, y, z_c) &= P(x_1, r, y) \alpha^c \beta^{1-z_c} \\
&= \binom{c}{r-1} \binom{y-1}{x_1-1} p^{x_1} q^{c-r+z_c} \alpha^{r-z_c} \\
&= \binom{r-x_1-1}{y-r+x_1} \beta^{y-r+x_1} \\
&= \binom{y-1}{x_1-1} \binom{2-x_1}{1} \beta^{r-x_1} \\
&= \binom{1}{x_1} \binom{1}{z_c} \beta^{r-x_1} \\
y &= 1-x_1, 2-x_1, \ldots \\
r &= 1, \ldots, c \\
x_1, z_c &= 0, 1.
The probability generating function and moments of \( N' \) are

\[
G_{N'}(t) = tG_N(t) = t\left\{ pt + \frac{qpt^2}{1 - \beta t}\right\}\alpha t + \frac{qpt^2}{1 - \beta t}
\]

\[
\begin{align*}
EN' &= EN + 1 \\
\text{Var}(N') &= \text{Var}(N).
\end{align*}
\]

**Distributions of \( N_{i,j} \)**

1. Let \( c - R = W \).

Then from (3.1.9) \( W \sim b(c-1, \alpha) \).

Now, \( W \) and \( Z_c \) are independent and

\[
N_{11} = \begin{cases} 
W + 1 & \text{if } Z_c = 1 \\
W & \text{if } Z_c = 0
\end{cases}
\]

Thus

\[
P(N_{11} = n_{11}) = \binom{c-1}{n_{11}-1} \alpha^{n_{11}-1} \frac{c-1}{\alpha}^{n_{11}-1-(n_{11}-1)} + c^{-1} \binom{c-1}{n_{11}} \alpha^{n_{11}} c^{-n_{11}}
\]

\[
= \binom{c}{n_{11}} \alpha^{n_{11}} c^{-n_{11}} \quad n_{11} = 0, \ldots, c.
\]

Similarly

\[
P(N_{10} = n_{10}) = \binom{c}{n_{10}} \alpha^{n_{10}} \frac{c}{\alpha}^{n_{10}} \alpha^{c-n_{10}} \quad n_{10} = 0, \ldots, c.
\]

Thus

\[
(3.2.6) \quad \begin{align*}
N_{11} &\sim b(c, \alpha) \\
N_{10} &\sim b(c, \alpha).
\end{align*}
\]

2. The distributions of \( N_{01} \) and \( N_{00} | n_{01} \) are the same as given in Section 3.1.

**3.2.2. Parameter estimation**

1. Since the plan \( S_c(N_{1.}) \) has the property that \( N_{1.} = c \) according to Theorem 2.3.2, we expect it to be efficient for \( g(\alpha) = \alpha \).
Indeed since $N_{11} \sim b(c, \alpha)$

$$\hat{\alpha} = \frac{N_{11}}{N_{11} \mp N_{10}} \cdot \frac{N_{11}}{c}.$$

is an efficient estimator of $\alpha$.

Thus any non-constant estimator of the form $f = aN_{11} + b$ is an efficient estimator of $g(\alpha) = a\alpha + b$ and these are the only efficient estimators.

2. The statements in Section 3.1.2 concerning estimation of $g(\beta)$ apply to this plan without modification.
CHAPTER IV

INFERENCE UNDER SAMPLING PLANS \( S(N_{11}) \)

In Section 2.3, Chapter II, it was shown that sampling plans of type \( S(N_{11}) \) are efficient for \( g(\alpha) = 1/\alpha \) and \( N_{1l} = N_{11} + N_{10} \) is the only efficient estimator of \( c/\alpha \) where \( N_{11} = c \).

In this chapter sampling plans \( S_c(N_{11}) \) are studied in which we continue observation until we get exactly \( c \) transitions from \( S \) to \( S \). The minimum variance unbiased estimator of \( \alpha \) is given. An estimator of \( \beta \) is given and its mean squared error is obtained.

4.1.1. Distributions Under Sampling Plan \( S_c(N_{11}) \)

We first derive the joint distribution of \((X_1, X, Y)\). As seen in Chapter II, there are two types of realizations possible.

\[
\begin{align*}
X_1 = 1 : S \ldots S F \ldots F S \ldots S \ldots F \ldots F S \ldots S | S \\
&= s_1 f_1 s_2 f_{r-1} s_r \\
X_1 = 0 : F \ldots F S \ldots S F \ldots F \ldots F S \ldots S | S \\
&= f_1 s_1 f_2 f_r s_r
\end{align*}
\]

Note that under this sampling scheme the last two observations result in successes, so as to obtain \( c \) th transition from \( S \) to \( S \).

Then as in (3.1.1), we have the following relations,

\[
\begin{align*}
N_{1l} &= (X - 1) - R + 1 = X - R \\
N_{10} &= R - 1 \\
N_{01} &= R - X_1 \\
N_{00} &= Y - R + X_1.
\end{align*}
\]
But since \( N_{11} = c \), we have \( R = X - c \), and we obtain,

\[
N_{11} = c \\
N_{10} = X - c - 1 \\
N_{01} = X - c - X_1 \\
N_{00} = Y - X + c + X_1 \\
N_{10} = X - 1 \quad \text{and} \quad N_0 = Y.
\]

By similar arguments as given in section 5.1.1, the total number of ways we can have \( X_1 = x_1 \), \( X = x \), \( Y = y \) and \( N_{11} = c \) is

\[
\binom{x - 2}{c - 1} \binom{y - 1}{x - c - x_1 - 1} = \binom{x - 2}{c - 1} \binom{y - 1}{x - c - x_1 - 1}.
\]

The joint distribution of \((X_1, X, Y)\) is then given by

\[
p(x_1, x, y) = \binom{x - 2}{c - 1} \binom{y - 1}{x - c - x_1 - 1} p^{x - c - x_1} q^{x - 1} \alpha^c \beta^{x - c - x_1} \\
y \geq x - c - x_1
\]

where \( x \geq c + 1 \)

\( x_1 = 0, 1 \).

The marginal distribution of \((x_1, x)\) is

\[
p(x_1, x) = \sum_{y = x - c - x_1}^{x - c - 1} p(x_1, x, y) \\
= \binom{x - 2}{c - 1} x_1^{1 - x_1} \alpha^c \beta^{(x - 1) - c} \\
= p(x_1) p(x) \quad x = c + 1, \ldots \\
x_1 = 0, 1.
\]
And the conditional distribution of $y$ given $X_1 = x_1, X = x$ is

$$p(y|x_1, x) = \binom{y-1}{x-c-x_1-1} \left(\frac{x-c-x_1}{y-x(x-c-x_1)}\right)^{x-c-x_1} \left(1 - \frac{x-c-x_1}{y-x(x-c-x_1)}\right)^{y-x(x-c-x_1)} \quad y \geq x-c-x_1$$

The above results are summarized in the following proposition.

**Proposition 4.1.1**
(a) The probability distribution of $W = X - 1$ is $\overline{NB}(c, \alpha)$ and $W$ is independent of $X_1$.
(b) The conditional distribution of $Y$ given $X_1 = x_1, X = x$ is

$$\overline{NB}(x-c-x_1, \beta) \text{ for } x-c-x_1 \geq 1.$$ 

The probability generating function of $X, Y$ and $N$ are given below. Then means and variances are derived for later use.

$$G_X(t) = t^{\frac{\alpha t}{1 - \alpha t}}^c,$$

$$E[X] = 1 + \frac{c}{\alpha}, \text{ var}(X) = \frac{\alpha \beta}{\alpha^2}$$

(ii) $G_Y|X(t) = \sum_{x_1=0}^{x_1} p_{x_1} \left(1-x_1\right)\left(\frac{\beta t}{1-\beta t}\right)^{x-c-x_1}$

$$= (p + \frac{\alpha t}{1 - \beta t})(\frac{\beta t}{1 - \beta t})^{x-c-1}$$

$$G_Y(t) = \sum_{x=c+1}^{\infty} p(x) \left(\frac{\beta t}{1 - \beta t}\right)^{x-c-1}$$

After some algebraic simplification we get

$$G_Y(t) = \left(p + \frac{\alpha t}{1 - \beta t}\right)^c \left(\frac{1 - \beta t}{1 - (1-\alpha) t}\right)^c.$$

Using $G_Y(t)$, or directly from the proposition 4.1.1 using the fact that
\[ E(Y | x_1, x) = \frac{x - c - x_1}{\beta}, \]
\[ V(Y | x_1, x) = \frac{(x - c - x_1)^\beta}{\beta^2}, \]

we get,

\[ E_Y = \frac{\alpha + \alpha x}{\alpha \beta}, \]

(4.1.10)

\[ \text{Var}(Y) = \frac{1}{(\alpha \beta)^2} \{\alpha \beta (\alpha x + \alpha) + \alpha + p \alpha x^2\}. \]

(iii) The probability generating function of \( N = x + y \) can be obtained directly from (4.1.2).

\[
G_N(t) = \sum_{x=c+1}^{\infty} \frac{1}{\sum_{x_1=t}^{x} \sum_{y=x-c-x_1}^{y} p(x_1, x, y) t^x t^y}
\]

\[
= \sum_{x=c+1}^{\infty} \left(\frac{x-2}{c-1}\right) \alpha^c \frac{x-c-1}{\alpha} \frac{1}{\sum_{x_1=0}^{x_1} p \cdot q^{x_1} t^{x_1-y-x_1}}
\]

\[
= \beta x - c - x_1 \cdot \sum_{y=x-c-x_1}^{\infty} \left(\frac{y-1}{x-c-x_1}\right) \left(\frac{\alpha t}{\alpha t - \beta t}\right)^y - x + x_1 + c
\]

After algebraic simplification, we get

(4.1.11) \[ G_N(t) = \left( pt + \frac{q \beta t^2}{1-\beta t}\right) \left(\frac{\alpha t (1-\beta t)}{1-\beta t - \alpha \beta t^2}\right). \]

Let

\[ A(t) = pt + \frac{q \beta t^2}{1-\beta t}, \]

\[ B(t) = \frac{\alpha t (1-\beta t)}{1-\beta t - \alpha \beta t^2}. \]

Then,

\[ A'(1) = 1 + \frac{q}{\beta}, \quad A''(1) = \frac{2q}{\beta^2}. \]
Using the relations,

\[ EN = A' (1) + cB' (1), \]
\[ g'' (1) = A'' (1) + 2c A' (1) B' (1) + C(c-1)(B' (1))^2 + B'' (1), \]
\[ \text{Var}(N) = g'' (1) - [g' (1)]^2 + g' (1), \]

we get,

\[ \text{(4.1.12)} \quad EN = 1 + \frac{g}{\alpha} + \frac{c(\alpha + \beta)}{\alpha \beta}, \]
\[ \text{Var}(N) = \frac{g(p + \beta)}{\beta^2} + \frac{c(\alpha + \beta) - 2(\alpha - \beta)}{\alpha \beta} + \frac{[2(\beta + \alpha \beta) - c(\alpha + \beta)](\alpha + \beta)}{(\alpha \beta)^2}. \]

**Distributions of \( N_{11} \)**

From the joint distribution (4.1.2) of \((X_1, X, Y)\) and relation (4.1.1) we get

\[ \text{(4.1.13)} \quad p(x_1, n_{10}, n_{01}, n_{00}) = \left( \begin{array}{c} n_{10} + c - 1 \\ c - 1 \end{array} \right) \left( \begin{array}{c} n_{00} + n_{01} - 1 \\ n_{01} - 1 \end{array} \right) p_1 q_1^{1-x_1} \alpha^c \frac{n_{10}}{\alpha} \frac{n_{01}}{\beta} \frac{n_{00}}{\beta}. \]

Since \( n_{01} = n_{10} + 1 - x_1 \), we have

\[ p(x_1, n_{10}, n_{00}) = \left( \begin{array}{c} n_{10} + c - 1 \\ c - 1 \end{array} \right) \left( \begin{array}{c} n_{00} + n_{10} - x_1 \\ n_{10} - x_1 \end{array} \right) p_1 q_1^{1-x_1} \alpha^c \frac{n_{10}}{\alpha} \frac{n_{00}}{\beta} \frac{n_{10} - x_1}{\beta} \frac{n_{00}}{\beta}. \]

The marginal distribution of \((x_1, N_{10})\) is then given by

\[ p(x_1, n_{10}) = \left( \begin{array}{c} n_{10} + c - 1 \\ c - 1 \end{array} \right) \alpha^c \frac{n_{10}}{\alpha} p_1 q_1^{1-x_1} \frac{n_{10} + 1 - x_1}{\beta} \sum_{n_{00}=0}^{n_{00}=n_{10} - x_1} \left( \begin{array}{c} n_{00} + n_{10} - x_1 \\ n_{10} - x_1 \end{array} \right) \frac{n_{00}}{\beta}. \]

\[ = \left( \begin{array}{c} n_{10} + c - 1 \\ c - 1 \end{array} \right) \alpha^c \frac{n_{10}}{\alpha} \left[ p_1 q_1^{1-x_1} \right]. \]

The above results are stated in the following.
(a) The distribution of $N_{10}$ is given by

\[(4.1.15)\quad N_{10} \sim NB(c, \alpha) \text{ and independent of } X_1.\]

(b) The conditional distribution of $N_{00}$ given $N_{01}$ is

\[(4.1.16)\quad N_{00} \mid N_{01} = n_{01} \geq 1 \sim NB(n_{01}, \beta) \]

or

\[
N_{00} \mid N_{10} = n_{10} \sim NB(n_{10} + l - x_1, \beta) \quad \text{for } n_{10} + l - x_1 \geq 1.
\]

The unconditional distribution of $N_{01}$ is given by

\[(4.1.17)\quad p(n_{01}) = \begin{cases} p \alpha^c & \text{for } n_{01} = 0 \\ \frac{p}{(n_{01} + c - 1)^{n_{01}} (c - 1)^{n_{01} + c - 2}} \alpha^{n_{01} + c - 2} \left(\alpha + q\right)^{-c - 1} & \text{for } n_{01} \geq 1. \end{cases}
\]

### 4.1.2. Estimation of parameters $\alpha$, $\beta$

Since our sampling plan $S_{c(N_{11})}$ is efficient for $g(\alpha) = \frac{ac}{\alpha} + b$ (where $a$, $b$ are constants) and the only efficient estimators are of the form $a N_1 + b = a(X - 1) + b$, there does not exist efficient estimator for $\alpha$ under $S_{c(N_{11})}$. However, minimum variance unbiased estimators do exist.

Since $W = X - 1 \sim NB(c, \alpha)$ the minimum variance unbiased estimator for $\alpha$ for $c \geq 2$ is given by

\[(4.1.18)\quad \hat{\alpha} = \frac{c - 1}{X - 2}, \quad X \geq c + 1.
\]

Notice that the variance of $\hat{\alpha}$ is given by

\[(4.1.19)\quad \text{Var}(\hat{\alpha}) = \text{Var}\left(\frac{c - 1}{W - 1}\right)
\]

\[
= \alpha^2 \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} (k+c-1)^{-1} (\alpha + q)^{k+c-2} \sum_{k=1}^{\infty} \frac{\alpha^k}{k!} (k+c-1)^{-1}.
\]
For functions of $\beta$, $g(\beta)$, we have the same problem of non-estimability discussed in Section 4.2, Chapter III. In particular there does not exist any unbiased estimator of $\beta$.

We now consider two estimators, the maximum likelihood estimator $\beta^*$, and $\hat{\beta}$ defined in the following.

\begin{align*}
(4.1.20) \quad \beta^* &= \begin{cases} 
\frac{N_{01}}{N_{01} + N_{00}} = \frac{N_{01}}{Y} & \text{if } Y = y \geq 1 \\
\alpha & \text{if } Y = 0
\end{cases}
\end{align*}

and

\begin{align*}
(4.1.21) \quad \hat{\beta} &= \begin{cases} 
\frac{N_{01} - 1}{N_{01} + N_{00} - 1} = \frac{N_{01} - 1}{Y - 1} & \text{if } Y = y \geq 2 \\
1 & \text{if } Y = 1 \\
\alpha & \text{if } Y = 0 .
\end{cases}
\end{align*}

where $\alpha$ is an arbitrary constant $0 \leq \alpha \leq 1$.

The bias of the above estimates is given by the following theorem.

**Theorem 4.1.3.**

\begin{align*}
(4.1.22) \quad \text{i) } E(\hat{\beta} - \beta) &= (\alpha - \beta) p \alpha^c \\
\text{(4.1.23) \quad ii) } E(\beta^* - \beta) &= (\alpha - \beta) p \alpha^c + \beta \sum_{n_{01}=1}^{\infty} P(N_{01} = n_{01}) \sum_{k=1}^{\infty} \beta^k (n_{01})^{-1}
\end{align*}

where $p(n_{01}) = P(N_{01} = n_{01})$ is given by (4.1.11).

**Proof.** i) From (4.1.16) we have

\begin{align*}
Y = N_{00} \mid N_{01} = n_{01} \geq 1 & \sim NB(n_{01}, \beta) \\
Y = 0 & \text{ if } N_{01} = 0 ,
\end{align*}

and from (4.1.21)
\[
\hat{\beta} |_{N_{01} = 0 \geq 2} = \frac{n_{01} - 1}{y - 1} 
\]
\[
\hat{\beta} |_{N_{01} = 1} = \begin{cases} 
1 \\
0
\end{cases}
\]
\[
\hat{\beta} |_{N_{01} = 0} = \alpha
\]

So that, we have
\[
\begin{cases}
E(\hat{\beta} |_{N_{01} = 0 \geq 1}) = \beta \\
E(\hat{\beta} |_{N_{01} = 0}) = \alpha
\end{cases}
\]

Therefore,
\[
E(\hat{\beta}) = E\{E(\hat{\beta} |_{N_{01}})\} = a \, P(N_{01} = 0) + \sum_{n_{01} = 1}^{\infty} \hat{\beta} \, P(N_{01} = n_{01})
\]
\[
= a \, P(N_{01} = 0) + \beta \, (1 - P(N_{01} = 0))
\]
\[
= \beta + (a - \beta) \, p \, \alpha^c.
\]

ii) From (4.1.20)
\[
\begin{cases}
\hat{\beta}^* |_{N_{01} = n_{01} \geq 1} = \frac{n_{01}}{Y} \\
\hat{\beta}^* |_{N_{01} = 0} = \alpha
\end{cases}
\]

If \(Y = y \geq 1\)

Hence, we get
\[
E(\hat{\beta}^*) = E\{E(\hat{\beta}^* |_{N_{01}})\}
\]
\[
= a \, P(N_{01} = 0) + \sum_{n_{01} = 1}^{\infty} \hat{\beta} \, E(\hat{\beta}^* |_{N_{01} = n_{01}}) \, P(N_{01} = n_{01})
\]
\[
= a \, P(N_{01} = 0) + \beta \, (1 - P(N_{01} = 0)) + \hat{\beta} \, \sum_{n_{01} = 1}^{\infty} \frac{c}{k} \, P(N_{01} = n_{01}) \, \sum_{k=1}^{\infty} \beta^k \left( \frac{n_{01}}{n_{01}} \right)
\]
\[
= \beta + (a - \beta) \, p \, \alpha^c + \beta \, \sum_{n_{01} = 1}^{\infty} \frac{c}{k} \, P(N_{01} = n_{01}) \, \sum_{k=1}^{\infty} \beta^k \left( \frac{k + n_{01}}{n_{01}} \right).
\]

□
The mean squared error of the estimator $\hat{\beta}$ is then given by

$$E\{(\hat{\beta} - \beta)^2 | N_{01} = n_{01} \geq 2\} = Var(\hat{\beta} | N_{01} = n_{01} \geq 2) = \beta^2 \sum_{k=1}^{\infty} \frac{k}{n_{01} - 1} \left( k + n_{01} - 1 \right)^{-1}$$

$$E\{(\hat{\beta} - \beta)^2 | N_{01} = 1\} = Var(\hat{\beta} | N_{01} = 1) = \beta \bar{\beta}$$

$$E\{(\hat{\beta} - \beta)^2 | N_{01} = 0\} = (a - \beta)^2,$$

so that,

$$(4.1.24) \quad \text{MSE}(\hat{\beta}) = (a - \beta)^2 \sigma_c + \beta \bar{\beta} P(N_{01} = 1) + \beta^2 \sum_{n_{01} = 1}^{\infty} P(N_{01} = n_{01}) \sum_{k=1}^{\infty} \frac{k}{n_{01} - 1} \left( k + n_{01} - 1 \right)^{-1}. $$
Chapter V

INFERENCES UNDER SAMPLING PLANS $S(N_{10})$

In Chapter II, Section 2.3, it was shown that the sampling plans $S(N_{10})$ are efficient for $g(\alpha) = \frac{1}{\alpha}$ and $N_1 = N_{11} + N_{10}$ is the only efficient estimator of $\frac{c}{\alpha}$ where $N_{10} = c$.

In this chapter we study two of sampling plans, $S_c(N_{10})$ and $S_c(R)$, show that given first observation $X_1$, these plans are also efficient for $g(\beta) = \frac{1}{\beta}$, and finally find minimum variance unbiased estimates for $\alpha$ and $\beta$ under these sampling plans.

Sampling plan $S_c(N_{10})$ is to continue observations on $\{X_t, t \geq 1\}$ until we get exactly $c$ transitions from $S$ to $F$. This can also be accomplished by continuing observations until we get $c$ runs of successes which is $S_c(R)$, where $R$ is the number of the runs, and then take additional observations until we get one more $F$.

We first investigate the sampling plan $S_c(R)$.

5.1 Sampling Plan $S_c(R)$

5.1.1. Distributions under $S_c(R)$

We first derive the joint distribution of $(X_1, X, Y)$. Now, we have two kinds of realizations possible, starting with $S$ or $F$.

$X_1 = 1: \quad S...S \quad F...F \quad S...S \quad \ldots \quad S...S \quad F...F|3$

\[ s_1 \quad f_1 \quad s_2 \quad s_{c-1} \quad f_{c-1} \quad s_c \]
We note that the last observation should result in \( S \) preceded by \( F \) so that it constitutes \( c \)th run of \( S \)'s. Then we have the following relations.

\[
\begin{align*}
N_{11} &= (X - 1) - (c - 1) = X - c \\
N_{10} &= c - 1 \\
N_{01} &= c - X_1 \\
N_{00} &= Y - c + X_1 \\
N_{11} &= c - 1 \text{ and } N_{00} = Y
\end{align*}
\]

(5.1.1)

By usual arguments, the total number of possible ways we can have

- \((c-1)\) runs of successes \( s_1, s_2, \ldots, s_{c-1} \) such that \( X - 1 = x - 1 = c - 1 \)
- \((c-x_1)\) runs of failures, \( f_1, f_2, \ldots, f_{c-x_1} \) such that \( Y = y = \sum_{i=1}^{c-x_1} f_i \), with \( X_1 = x_1 \), is

\[
\frac{x^2 y^{c-2}}{(c-2)(c-x_1-1)}
\]

The joint distribution of \((X_1, X, Y)\) is

\[
n(x_1, x, y) = \binom{x-2}{c-2} \binom{y-1}{c-x_1-1} x_1^{1-x_1} a^{x-c} c^{c-x_1} y^{c+x_1} p^{x_1} q^{y}
\]

where

\[
x = c, c+1, \ldots,
\]

\[
y = c-x_1, c-x_1+1, \ldots,
\]

\[
x_1 = 0, 1.
\]

The marginal distributions of \( X \) and \((X_1, Y)\) are seen to be

\[
p(x) = \sum_{x_1} \sum_{y} p(x_1, x, y) = \binom{x-2}{c-2} \alpha^{c-1} \alpha^{x-c}
\]

\[
x = c, c+1, \ldots
\]
(5.1.4) \[ p(x_1, y) = P(X_1=x_1, Y=y) = \binom{y-1}{c-x_1-1} p^{x_1-1} q^{c-x_1} \beta^c y-c+x_1 \]

\[ y = c-x_1, c-x_1+1, \ldots, \]

\[ x_1 = 0, 1. \]

Thus we have

\[ p(x_1, x, y) = p(x) \cdot p(x_1, y) = p(x) \cdot p(y|x_1)p(x_1) \]

The above results are summarized in the following proposition.

**Proposition 5.1.1.** Under sampling plan \( S_c(R) \)

1. \( X \) and \( (X_1, Y) \) are independent,

2. \( W = X - 1 \sim NB(c-1, \alpha) \),

3. \( Y|X_1 = x_1 \sim NB(c-x_1, \beta) \).

Recall that if \( Z \sim NB(m; \theta) \), then

\[ EZ = m/\theta, \quad V(Z) = m\theta^2/\theta^2 \]

and the probability generating function of \( Z \) is given by

\[ G_z(t) = \left[ \frac{\theta t}{1-\theta t} \right]^m \]

Then using proposition 5.1.1., we have the mean, variance and probability generating function of \( X \) as follows.

(5.1.7) \[ EX = 1 + \frac{c-1}{\alpha} \quad \text{and} \quad \text{Var}(X) = \frac{(c-1)\alpha}{\alpha^2} \]

(5.1.8) \[ G_x(t) = t^{c-1} \left[ \frac{\alpha t}{1-\alpha t} \right] \]

The conditional mean, conditional variance and probability generating function of \( Y \) given \( X_1 \) are

\[ E(Y|x_1) = \frac{c-x_1}{\beta}, \quad \text{Var}(Y|x_1) = \frac{(c-x_1)\beta}{\beta^2} \]
We find the unconditional mean, variance and probability generating function from the above

\[
E(Y) = \frac{c-p}{\beta},
\]

(5.1.9)

\[
\text{Var}(Y) = \text{E}[\text{Var}(Y|X_1)] + \text{Var}(\text{E}(Y|X_1))
= \frac{(c-p)\beta}{\beta^2} + \frac{pq}{\beta^2} = \frac{1}{\beta^2} \left( (c-p)\beta + pq \right),
\]

(5.1.10)

\[
G_Y(t) = \sum_{x_1=0}^{1} \sum_{q=p}^{1-x_1} g_y|_{x_1}(t)
= (p + \frac{\alpha \beta t}{1-\beta t})(-\frac{\alpha \beta t}{1-\beta t})^{c-1}.
\]

Also, since \( E(XY) = \sum_{x_1} \sum_{y} x_1 y p(x, y) = \frac{p(c-1)}{\beta} \),

(5.1.11)

\[
\text{Cov}(X_1, Y) = \frac{p(c-1)}{\beta} - p \frac{(c-p)}{\beta} = -\frac{pq}{\beta}.
\]

The probability generating function of \( N \) is obtained by noting from Proposition 5.1.1 that \( X \) and \( Y \) are independent, so that we have,

(5.1.12)

\[
G_N(t) = G_X(t) \cdot G_Y(t)
= \left[ pt + \frac{\alpha \beta t^2}{1-\beta t} \right] \left( \frac{\alpha t}{1-\alpha t} \right)^{c-1} \left( \frac{\beta t}{1-\beta t} \right)^{c-1}.
\]

Then we have,

(5.1.13)

\[
\text{E}(N) = 1 + \frac{c-1}{\alpha} + \frac{c-p}{\beta} = \frac{\alpha \beta + \beta(c-1) + \alpha(c-p)}{\alpha \beta},
\]

\[
\text{Var}(N) = \text{Var}(X) + \text{Var}(Y) = \frac{(c-1)\alpha}{\alpha^2} + \frac{(c-p)\beta + pq}{\beta^2}.
\]
From (5.1.1) and (5.1.3), we get the joint distribution of 
\((x_1, N)\), given below.

\[(x^1, N), \text{ given below.}\]

\[\begin{align*}
\text{(5.1.14)} & \quad p(x_1, n_{11}, n_{01}, n_{00}) = \binom{n_{11} + c - 2}{c - 2} \binom{n_{00} + n_{11} - 1}{n_{01} - 1} \beta \frac{x_1^{1 - x_1} n_{11}^{-c - 1}}{\gamma}.
\end{align*}\]

But, since \(N_{01} = c - x_1\)

\[\begin{align*}
\text{(5.1.15)} & \quad p(x_1, n_{11}, n_{00}) = \\
& = \binom{n_{11} + c - 2}{c - 2} \binom{n_{00} + c - x_1 - 1}{c - x_1 - 1} \beta \frac{x_1^{1 - x_1} n_{11}^{-c - 1}}{\gamma}.
\end{align*}\]

We summarize the above results in the following

(a) \(N_{11}\) and \((N_{01}, N_{00})\), or \(N_{11}\) and \((x_1, N_{00})\) are indep.,

(b) \(N_{11} \sim NB(c - 1, \beta)\),

(c) \(N_{00} \mid x_1 = x_1 \sim NB(c - x_1, \beta)\), or

\[N_{00} \sim p \text{NB}(c - 1, \beta) + q \text{NB}(c, \beta).\]

5.1.2. Efficient Estimation

Since our plan \(S_c(R)\) has the property that \(N_{10} = c - 1\), from Section 2.3, Chapter II, we expect it to be efficient for functions of the form \(g(\alpha) = 1/\alpha\).

Indeed, by Proposition 5.1.1, we have

\[N_{11} = X - 1 \sim NB(c - 1, \alpha),\]

\[E \left( \frac{N_{11}}{c - 1} \right) = \frac{1}{\alpha}\]

\[V \left( \frac{N_{11}}{c - 1} \right) = \frac{1}{(c - 1)^2} \text{Var}(N_{11}) = \frac{\alpha}{(c - 1)\alpha^2}.\]

(lower bound given by (2.1.18))
Hence our plan is efficient for \( g(\alpha) = 1/\alpha \), and any non-constant estimator of the form \( a(N_1, t) + b = a(x-1) + b \) is efficient for \( a(c-1)/\alpha + b \).

Now, consider a conditional estimator

\[ f|x_1 = \frac{y}{c-x_1} . \]

Thus, by Proposition 5.1.1

\[ Y|x_1 \sim NB(c-x_1, \beta) \]

\[ E(f|x_1) = \frac{E(Y|x_1)}{c-x_1} = \frac{1}{\beta} . \]

\[ \text{Var}(f|x_1) = \frac{1}{(c-x_1)^2} \quad V(Y|x_1) = \frac{\beta}{(c-x_1)\beta^2} . \]

By (2.1.18)

\[ \text{Var}(f|x_1) = \frac{\beta}{E(N_0|x_1)} \left[ \frac{1}{\beta} \right]^2 = \frac{\beta}{(c-x_1)\beta^2} . \]

Hence, given \( x_1 = x_1 \), the plan \( S_c(R) \) is also efficient for \( g(\beta) = 1/\beta \), and any non-constant estimator of the form \( aN_0 + b = aY + b \) is efficient for \( \frac{a(c-x_1)}{\beta} + b \).

Thus, we have the following theorem.

**Theorem 5.1.1.** Given initial observation \( X_1 = x_1 \), sampling plan \( S_c(R) \) is efficient, and any non-constant functions of the form

(5.1.18) \[ f = a(N_1) + b(N_0) + d \]

are efficient estimators of

(5.1.19) \[ \frac{a(c-1)}{\alpha} + \frac{b(c-x_1)}{\beta} + d \]

and these are the only efficient estimators.

Note that another way of proving the above theorem is by showing
that

\( f = a(N_{11}) + b(N_{00}) + d \)

\( = \frac{a}{\alpha} [\alpha N_{11} - \alpha N_{10}] + \frac{b}{\beta} [\beta N_{01} + \beta N_{00}] + \frac{sc}{\alpha} + \frac{b(c-x_1)}{\beta} + d. \)

and using Corollary 2.3.1.

5.1.3. Estimation of parameters

The forms of the joint distributions (5.1.2) of \((X_1, X, Y)\) or (5.1.15) of \((X_1, N_{11}, N_{00})\) show that they are equivalent complete sufficient statistics. We now find minimum variance unbiased estimator of \(\alpha\) and \(\beta\). In Section 5.1.1 it was shown that \(X\) and \((X_1, Y)\) or equivalently \(N_{11}\) and \((X_1, N_{00})\) are independent and their distribution depends only on \(\alpha\) and \((p, \beta)\) respectively.

That is,

\( N_{11} = N_{11} + c = X - 1 \sim NB(c-1, \alpha) \)

\( N_{00} = Y|_{x_1} \sim NB(c-x_1, \beta) \).

For \(c \geq 3\), let

\[ \hat{\alpha} = 1 - \frac{(c-1)-1}{N_{11} - 1} = \frac{N_{11} - c + 1}{N_{11} - 1} = \frac{X - c}{X - 2} \quad X = x \geq c \]

and for \(c = 2\) let

\[ \hat{\alpha} = \begin{cases} 1 & X = x \geq 3 \\ 0 & X = 2 \end{cases} \]

Then \(\hat{\alpha}\), for \(c \geq 2\), is an unbiased estimator of \(\alpha\) which is a function of sufficient statistic.

Similarly for \(c \geq 3\), let

\[ \hat{\beta} = \frac{N_{01} - 1}{N_{00} - 1} = \frac{c - X - 1}{X - c} \quad Y + X_1 = y + x_1 \geq c. \]
Then \( E(\beta|x_1) = \beta \) so \( \hat{E}\beta = \beta \).

Also for \( c = 2 \), let

\[
\beta = \begin{cases} 
\frac{1-x_1}{Y-1} & Y = y \geq 2 \\
1 & Y = 1, x_1 = 1 
\end{cases}
\]

The above definition of \( \hat{\beta} \) shows that,

\[
\hat{\beta}|_{x_1=0} = \begin{cases} 
\frac{1}{Y-1} & Y = y \geq 2 \\
1 & Y = 1 
\end{cases}
\]

The conditional distributions of \( Y \) are given by,

\[
Y|_{x_1=0} \sim \text{NB}(2, \beta) \\
Y|_{x_1=0} \sim \text{NB}(1, \beta) = \beta^{-1} y^{-1}
\]

Hence, we find

\[
E(\hat{\beta}|x_1) = \beta
\]

giving

\[
\hat{E}\beta = \beta.
\]

Hence for \( c \geq 2 \), \( \hat{\beta} \) is also an unbiased estimator of \( \beta \) and \( \hat{\beta} \) is a function of complete sufficient statistic. These results are summarized in the following theorem.

**Theorem 5.1.2.** The MVUE's of \( \alpha \), and \( \beta \) are given by

\[
\hat{\alpha} = \begin{cases} 
\frac{1-x}{N+c+1} = \frac{x-c}{x-2} & x = x \geq c, \ c \geq 3 \\
1 & x = x \geq 3, \ c = 2 \\
0 & x = 2, \ c = 2
\end{cases}
\]

(5.1.21)
The minimum variances of $\hat{\alpha}$ and $\hat{\beta}$ are obtained now.

For $c \geq 3$, by (3.1.11), we note that

$$\text{Var}(\hat{\alpha}) = \text{Var}(\hat{\alpha}) = \text{Var}\left(\frac{(c-1)}{n_{1.1} - 1}\right)$$

$$= \sum_{k=1}^{c-1} \alpha^k \binom{k+c-2}{c-2} \cdot$$

Trivially for $c = 2$, $\text{Var}(\hat{\alpha}) = \alpha \alpha$.

For $c \geq 3$,

$$\text{Var}(\hat{\beta} \mid X_1) = \beta^2 \sum_{k=1}^{c-1} \beta^k \binom{k+c-1}{c-1} \cdot$$

since $E(\hat{\beta} \mid X_1) = \beta$, $V(E(\hat{\beta} \mid X_1)) = 0$.

Hence we obtain $\text{Var}(\hat{\beta}) = E(\text{Var}(\hat{\beta} \mid X_1))$ and after simplification this reduces to

$$\text{Var}(\hat{\beta}) = \beta^2 \sum_{k=1}^{c-1} \beta^k \binom{k+c-1}{c-1} \left(1 + \frac{pk}{c-1}\right) \cdot$$

When $c = 2$, we have

$$\text{Var}(\hat{\beta} \mid X_1 = 1) = \beta \beta$$

$$\text{Var}(\hat{\beta} \mid X_1 = 0) = \beta^2 \sum_{k=1}^{c-1} \beta^k \binom{k+1}{c-1} \cdot$$

which simplifies to

$$V(\hat{\beta} \mid X_1 = 0) = \frac{\beta^2}{\beta} [\beta + \log \beta]$$

or,
(5.1.25) \[ \text{Var}(\beta) = p\beta^2 - \frac{q\theta^2}{\beta} (\bar{\beta} + \log \beta). \]

5.2. Sampling plan \(S_c(N_{10})\)

5.2.1. Distributions under \(S_c(N_{10})\)

Let the random variables \(X, Y, N\) be defined as in Section 5.1 and define new random variable \(N^*\) to be the total number of observations required under \(S_c(N_{10})\). Let \(X^* = X^*(1, \ldots, X_{N^*}), Y^* = Y^*(X^*_1, \ldots, X^*_{N^*})\) be the total numbers of successes and failures respectively.

Two types of sequences of success and failures are possible in this case given as follows.

\[
\begin{align*}
X_1 = 1 : & \quad S \ldots S \ F \ldots F \ S \ldots S \quad \ldots \quad S \ldots S \ F \ldots F \quad S \ S \ldots S | F \\
& \quad s_1 \quad f_1 \quad s_2 \quad \ldots \quad s_{c-1} \quad f_{c-1} \quad s_c
\end{align*}
\]

\[
\begin{align*}
X_1 = 0 : & \quad F \ldots F \ S \ldots S \ F \ldots F \ \ldots \quad S \ldots S \ F \ldots F \quad S \ S \ldots S | F \\
& \quad f_1 \quad s_1 \quad f_2 \quad \ldots \quad s_{c-1} \quad f_c \quad s_c
\end{align*}
\]

Notice that the last observation should yield an \(F\) preceded by an \(S\) so that it constitutes \(c\)th transition from \(S\) to \(F\). The following relations hold.

\[
\begin{align*}
N_{11} &= X^* - c \\
N_{10} &= c \\
N_{01} &= c - X_1 \\
N_{00} &= (Y^*-1) - (c-X_1) = Y^* - c - 1 + X_1 \\
N_{10} &= X^*, \text{ and } N_{00} = Y^* - 1.
\end{align*}
\]

As discussed earlier, the total number of ways of obtaining \(X_1 = x_1\)
\(X^* = x^*, Y^* = y^*\), and \(N_{10} = c\) is given by \(\binom{x^*-1}{c-1}\binom{y^*-2}{c-x_1-1}\).

The joint distribution of \((X_1, X^*, Y^*)\) is, therefore
\( p(x_1, x^*, y^*) = \frac{(x^*-1) \binom{y^*-2}{c-1} x_1^{1-x_1} \beta^{c-x_1} y^*-c-1+x_1 \beta}{\binom{c-2}{c-x_1-1} \alpha^{x^*} \alpha^{x^*+c-x_1-2}} \)

where \( x^* = c, c+1, \ldots \)
\( y^* = c-x_1+1, c-x_1+2, \ldots \)

Notice that the distribution \( p(x_1, x^*, y^*) \) can also be found from (5.1.2) of plan \( S_c(R) \).

Let \( X_c \) = number of successes before we get first \( F \), given \( S \).

Then the probability density of \( X_c \) is given by
\( p(x_c) = \frac{x_c^c}{\alpha^c} x_c = 0, 1, \ldots \)

Also,
\( x^* = X + X_c \)
\( y^* = Y + 1 \)
\( N^* = N + X_c + 1 \)

Hence using (5.1.2), we have
\( P(x_1, x^*, y^*, x_c) = \)
\( = \frac{(x^*-x_c-2) \binom{y^*-2}{c-2} x_1^{1-x_1} \beta^{c-x_1} y^*-c-1+x_1 \beta}{\binom{c-2}{c-x_1-1} \alpha^{x^*} \alpha^{x^*-c-x_1-2}} \)

Using the fact that
\( \sum_{x_c=0}^{x^*-c} \binom{x^*-c-2}{c-2} = \binom{x^*-1}{c-1} \)
we can obtain the joint probability distribution of \( X_1, X^* \) and \( Y^* \).

The results are stated in the following proposition.

**Proposition 5.2.1.** The random variables \( X^* \) and \( (X_1, Y^*) \) are independent. The distribution of \( X^* \) is negative binomial given by
\( x^* \sim NB(c, \alpha) \)
and the conditional distribution of \( Y^*-1 \) given \( X_1 \) is also negative
given by

\[ N_{0*} = Y^* - 1|_{x_1} \sim \overline{NB}(c-x_1, \beta). \]

The following results can be obtained from the proposition:

(a) \[ G_{x^*}(t) = \left[ \frac{\alpha t}{1-\alpha t} \right]^c, \]

(b) \[ \frac{\alpha}{\beta}, \quad \text{Var}(X^*) = \frac{c\alpha}{\beta^2}, \]

\[ \frac{\alpha}{\beta}, \quad \text{Var}(Y^*) = \frac{(c-x_1)\beta}{\beta^2}, \]

\[ G_{y^*|x_1}(t) = t\left[ \frac{\beta t}{1-\beta t} \right]^{c-x_1}, \]

so that the unconditional mean and variance are given by

\[ \text{E}X^* = 1 + \frac{c-p}{\beta}, \]

\[ \text{Var}(X^*) = \frac{(c-p)\beta + pq}{\beta^2}. \]

Also

\[ G_{y^*}(t) = (pt + \frac{q\beta t^2}{1-\beta t})\left( \frac{\beta t}{1-\beta t} \right)^{c-1}. \]

(c) Since \( X^* \) and \( Y^* \) are independent and \( N^* = X^* + Y^* \)

\[ G_{n^*}(t) = (pt + \frac{q\beta t^2}{1-\beta t})\left( \frac{\beta t}{1-\beta t} \right)^{c-1}, \]

giving the following.

\[ \text{E}N^* = 1 + \frac{c}{\alpha} + \frac{c-p}{\beta} = \overline{\alpha} + \overline{\beta} + \frac{(c-p)\alpha}{\alpha \beta}, \]

\[ \text{Var}(N^*) = \frac{c\alpha}{\alpha^2} + \frac{(c-p)\beta + pq}{\beta^2}, \]

\[ \text{Cov}(X, Y) = -pq/\beta, \]
The joint distributions of $X_1$, $N_{11}$, $N_{01}$ and $N_{00}$ are obtained from (5.2.1) and (5.2.3) as follows:

$$
(5.2.15) \quad P(x_1, n_{11}, n_{01}, n_{00}) = \left( \begin{array}{l}
\frac{n_{11} + c - 1}{c - 1} \frac{n_{00} + n_{01} - 1}{n_{01} - 1}
\end{array} \right) \left( \begin{array}{l}
\frac{x_1}{p} \frac{1 - x_1}{q} \frac{n_{11} - c}{\beta} \frac{n_{01} - n_{00}}{\beta}
\end{array} \right).
$$

The marginal distribution of $X_1$, $N_{11}$ and $N_{00}$ is given by

$$
(5.2.16) \quad P(x_1, n_{11}, n_{00}) = \left( \begin{array}{l}
\frac{n_{11} + c - 1}{c - 1} \frac{n_{00} + c - x_1 - 1}{c - x_1 - 1}
\end{array} \right) \left( \begin{array}{l}
\frac{x_1}{p} \frac{1 - x_1}{q} \frac{n_{11} - c}{\beta} \frac{c - x_1 - n_{00}}{\beta}
\end{array} \right).
$$

Notice that $N_{01} = c - x_1$. We can summarize the above results in the following.

The random variables $N_{11}$ and $(N_{01}, N_{00})$ or $N_{11}$ and $(X_1, N_{00})$ are independent. $N_{11}$ has negative binomial distribution given by

$$
(5.2.17) \quad N_{11} \sim NB(c, \alpha)
$$
and the conditional distribution of $N_{00}$ given $X_1$ is also negative binomial given by

$$
(5.2.18) \quad N_{00 | x_1} \sim NB(c - x_1, \beta).
$$

### 5.2.2. Efficient Estimation

The sampling plan $S_c(N_{10})$ is such that $N_{10} = c$, hence it is efficient for $g(\alpha) = 1/\alpha$. Also since $N_{11}$ is negative binomial, we have the mean and variance of $N_{11}$ given by,

$$
\{ \begin{align*}
E \left( \frac{N_{11}}{c} \right) &= \frac{1}{\alpha} \\
Var \left( \frac{N_{11}}{c} \right) &= \frac{\alpha}{c^2 \alpha^2}.
\end{align*}
\$$
Any non-constant estimators of the form \( a(N_1) + b = aX + b \) are efficient for \( \frac{ac}{\alpha} + b \).

Again since the conditional distribution of \( N_{01} \) is negative binomial, we have

\[
E\left( \frac{N_{01}}{c-x_1} \mid x_1 \right) = \frac{1}{\beta},
\]

\[
\text{Var}\left( \frac{N_{01}}{c-x_1} \mid x_1 \right) = \frac{\beta}{(c-x_1)\beta^2}.
\]

In view of (2.1.18) this is the lower bound of the variance given \( x_1 = x_1 \).

Hence given \( x_1 = x_1 \), any nonconstant estimator \( a(N_{01}) + b = a(Y + b = a(Y^*-1) + b \) is efficient for \( \frac{ac}{\alpha} + b \).

We state the above results in the following theorem.

**Theorem 5.2.1.** Given \( x_1 = x_1 \), sampling plan \( S_c(N_{10}) \) is efficient and any non-constant estimator of the form

\[(5.2.19) \quad f = a(N_{01}) + b(N_{01}) + d = aX + b(Y^*-1) + d \]

is efficient for its mean

\[(5.2.20) \quad \frac{ac}{\alpha} + \frac{b(c-x_1)}{\beta} + d\),

and these are the only efficient estimators.

5.2.3. **Estimation of parameters**

Since the distribution of \( N_1 \) and conditional distribution of \( N_{01} \) given \( x_1 \) are negative binomial, we can state the following theorem analogous to theorem 5.1.2.

**Theorem 5.2.2.** The MVUE's of \( \alpha \) and \( \beta \) are given by
\[
\hat{\alpha} = \frac{N_0 - c}{N_0 - 1} = \frac{x^* - c}{x^* - 1} \quad \text{if } x^* \geq c, \quad c \geq 2
\]

\[
\hat{\beta} = \begin{cases} 
\frac{1 - x_1}{y^* - 2} & \text{if } y^* = y^* \geq 3, \quad c = 2 \\
1 & \text{if } y^* = 2, \quad x_1 = 1, \quad c = 2 
\end{cases}
\]

The minimum variances of \( \hat{\alpha} \) is for all \( c \geq 2 \),

\[
\text{Var}(\hat{\alpha}) = \alpha^2 \sum_{k=1}^{\infty} \alpha^k \binom{k+c-1-1}{c-1}.
\]

The variances of \( \hat{\beta} \) are given by (5.1.24) for \( c \geq 3 \) and (5.1.25) for \( c = 2 \).
Chapter VI

INFERENCE UNDER SAMPLING PLANS S(N₁₀⁺ N₀₁)

In this chapter we study two sampling plans \( S₂c(N₁₀⁺ N₀₁) \) and \( S₂c⁻¹(N₁₀⁺ N₀₁) \) which constitute \( S(N₁₀⁺ N₀₁) \). Specifically we study the sampling distributions, find MVUE's of \( \alpha \) and \( \beta \), and show that given initial observation \( X₁ = x \), the plan \( S₂c⁻¹(N₁₀⁺ N₀₁) \) is efficient for two-parameter functions \( g(\alpha, \beta) \). Note that \( S₂c(N₁₀⁺ N₀₁) \) is efficient for two-parameter functions \( g(\alpha, \beta) = \frac{a}{\alpha} + \frac{b}{\beta} \) as has been shown in Section 2.4 of Chapter II.

6.1 Sampling plan \( S₂c(N₁₀⁺ N₀₁) \)

We first find the joint distribution of \( (X₁, X, Y) \) where the random variables \( X₁, X, Y \) are defined as before.

There are two possible realizations, starting with \( S \) and terminating with \( S \) and starting with \( F \) and terminating with \( F \):

\[
\begin{align*}
X₁ = 1 : & S \ldots S \ F \ldots F \ S \ldots S \ldots S \ S \ldots S \ F \ldots F \\
& \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \\
& s₁ \ f₁ \ s₂ \ s₂ \ s₂ \ f₁ \\
X₁ = 0 : & F \ldots F \ S \ldots S \ F \ldots F \ldots F \ F \ldots F \ S \ldots S \\
& \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \ \downarrow \\
& f₁ \ s₁ \ f₂ \ f₂ \ f₂ \ s₁
\end{align*}
\]

Thus we have following relations

\[
\begin{align*}
N₁₁ &= X - X₁ - c \\
N₁₀ &= N₀₁ = c \\
N₀₀ &= Y - (1 - X₁) - c
\end{align*}
\]

(6.1.1)
By usual arguments, the total number of ways to partition an integer \( x - x_1 \), into \( c \) integers \( s_1, \ldots, s_i, s_i > 0 \), such that \( \sum s_i = x - x_1 \), and an integer \( y - 1 + x_1 \) into \( c \) integers \( f_1, \ldots, f_c \), \( f_i > 0 \) such that \( \sum f_i = y - 1 + x_1 \) is given by,

\[
\binom{x-x_1-1}{c-1} \binom{y-(1-x_1)-1}{c-1} = \binom{x-x_1-1}{c-1} \binom{y+x_1-2}{c-1}.
\]

Hence the joint distribution of \((X_1, X, Y)\) is

\[
p(x_1, x, y) = \binom{X_1-1}{c-1} \binom{y-x_1-2}{c-1} p^{-\alpha} q^{\alpha} \beta^{x_1-1} \gamma^{c-1+x_1}
\]

where

\[
\begin{cases}
x \geq c + x_1 \\
y \geq c + 1 - x_1 \\
x_1 = 0.1
\end{cases}
\]

Since \( N_{11} = x - x_1 - c \),

and \( N_{00} = y - 1 + x_1 - c \),

we can write (6.1.2) as follows.

\[
P(x_1, n_{11}, n_{00}) = \left[ \binom{n_{11}+c-1}{c-1} \alpha^{n_{11}} \beta^{c} \right] \left[ \binom{n_{00}+c-1}{c-1} \beta^{c} \right] = \binom{x_1-1}{c-1} p^{-\alpha} q^{\alpha} \beta^{x_1-1} \gamma^{c-1+x_1}
\]

or

\[
P(x_1, n_{11}, n_{00}) = \left[ \binom{n_{11}+c-1}{c-1} \alpha^{n_{11}} \beta^{c} \right] \left[ \binom{n_{00}+c-1}{c-1} \beta^{c} \right] \gamma^{c-1+x_1} = \binom{x_1-1}{c-1} p^{-\alpha} q^{\alpha} \beta^{x_1-1} \gamma^{c-1+x_1}
\]

so that

\[
P(x_1, n_{11}, n_{00}) = p(n_{11}) \cdot p(n_{00}) \cdot P(x_1)
\]

The above results are summarized in the following proposition.
Proposition 6.1.1. (a) The random variables \( N_{11}(N_1), N_{00}(N_0) \) and \( X_1 \) are mutually independent.

(b) The probability distribution of \( N_{11}, N_{11}, N_{01} \) and \( N_0 \) are given by

\[
\begin{align*}
N_{11} & \sim NB(c, \alpha) \\
N_{11} & \sim NB(c, \alpha) \\
N_{01} & \sim NB(c, \beta) \\
N_0 & \sim NB(c, \beta)
\end{align*}
\]

(6.1.5)

We also have the marginal distributions \((X_1, X)\) and \((X_1, Y)\) given by the following.

\[
p(x_1, x) = \binom{x-x_1-1}{c-1} p^{1-x_1} q^{x_1} \alpha^{x_1} \\
p(x_1, y) = \binom{y+x_1-2}{c-1} p^{1-x_1} q^{x_1} \beta^{y-(c+1-x_1)}
\]

(6.1.6)

The above relations provide various moment and probability generating functions of the random variables considered.

1. \( E(N_{11}) = c/\alpha \), \( \text{Var}(N_{11}) = \frac{\alpha}{\alpha^2} \), \( G_{N_{11}}(t) = \left( \frac{\alpha t}{1-\alpha t} \right)^2 \)

and since \( N_{11} = X - X_1 \) we get

\[
\begin{align*}
EX & = p + c/\alpha \\
\text{Var}(X) & = \text{Var}(N_{11} + X_1) = \frac{\alpha}{\alpha^2} + pq \\
\text{Cov}(X_1, X) & = (p + \frac{pc}{\alpha}) - p(p + c/\alpha) = pq,
\end{align*}
\]

and \( G_X(t) = G_{N_{11} + X_1}(t) = G_{X_1}(t)G_{N_{11}}(t) \), so that

\[
G_X(t) = (pt + q)\left( \frac{\alpha t}{1-\alpha t} \right)^c.
\]

(6.1.8)
(ii) \[ E(N_0) = \frac{c}{\beta}, \quad \text{var}(N_0) = \frac{c\beta}{\beta^2} \]

and since \( N_0 = Y - 1 + X_1 \), we get

\[
\begin{align*}
\text{EX} &= q + \frac{c}{\beta} \\
\text{Var}(Y) &= \text{Var}(N_0 + (1 - X_1)) = \frac{c\beta}{\beta^2} + pq \\
\text{Cov}(X_1, Y) &= \frac{p\beta}{\beta} - p(q + \frac{c}{\beta}) = -pq.
\end{align*}
\]

and

\[ G_Y(t) = G_{N_0 + 1 - X_1}(t) = G_{1 - X_1}(t) \cdot G_{N_0}(t) \]

so that,

\[ G_Y(t) = (p + qt)\left(\frac{\beta t}{1 - \beta t}\right)^c. \]

(iii) Since \( N_1, N_0, \) and \( X_1 \) are independent and \( N = N_1 + N_0 + 1 \)

the probability generating function of \( N \) is given by

\[ G_N(t) = G_{N_1 + N_0}(t) = tG_{N_1}(t) \cdot G_{N_0}(t) \]

\[ = t\left(\frac{\alpha}{1 - \alpha t}\right)^c \cdot \left(\frac{\beta t}{1 - \beta t}\right)^c. \]

The mean and variance of \( N \) can then be obtained from (6.1.11) or directly from (i) and (ii) as follows.

\[
\begin{align*}
E(N) &= 1 + \frac{c\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)}{\alpha} \\
\text{Var}(N) &= c\left(\frac{\alpha}{\alpha^2} + \frac{\beta}{\beta^2}\right)
\end{align*}
\]

The covariance of \( X, Y \) is obtained by

\[ E(XY) = E(N_1 + X_1)(N_0 + 1 - X_1) = \frac{ac}{\alpha} + \frac{pc}{\beta} + \frac{c^2}{\alpha\beta}, \]

so that,
\[ \text{Cov}(X, Y) = -pq. \]

The necessary and sufficient condition that a sampling plan be efficient for a function of two parameters \( \alpha \) and \( \beta \), say \( g(\alpha, \beta) \), is that the sampling plan be \( S_{2c}(N_{10} + N_{01}) \) for some positive integer \( c \) as shown in Section 2.4. The function \( g(\alpha, \beta) \), estimable efficiently, has the following form.

\[ g(\alpha, \beta) = c\left(\frac{a}{\alpha} + \frac{b}{\beta}\right) + d \]

The efficient estimator of \( g(\alpha, \beta) \) is given by

\[ f = a(N_{1*}) + b(N_{0*}) + d. \]

The fact can also be noted as follows.

We see from (6.1.5) that \( N_{1*} \) and \( N_{0*} \) are independent and have negative binomial distributions, so that

\[ \text{Var}(f) = c\left(\frac{a^2}{\alpha^2} + \frac{b^2}{\beta^2}\right). \]

(6.1.6) gives the lower bound as given by (2.1.18) for \( g(\alpha, \beta) \) with \( E_{N_{1*}} = c/\alpha \), \( E_{N_{0*}} = c/\beta \).

The uniqueness of estimable function \( g(\alpha, \beta) \) except for constants \( a, b \) and \( d \) comes from the fact that if an efficient estimator exists, it will exist for only one specific function of \( \alpha \) and \( \beta \), and for no other function of \( \alpha \) and \( \beta \). We consider an example to show this uniqueness.

**Example:** Consider a function

\[ g_1(\alpha, \beta) = \frac{a}{\alpha \beta} + b \]

Then an unbiased estimator \( f_1 \) of \( g_1 \) is given by
(6.1.18) \[ f_1 = \frac{a}{c} (N_{1.}) (N_{0.}) + b \]

The variance of \( f_1 \) is given by
\[
\text{Var}(f_1) = \frac{a^2}{c^4} \text{Var}(N_{1.} \cdot N_{0.}) = \frac{a^2}{c^2} (EN_{1.}^2 N_{0.}^2 - (EN_{1.} N_{0.})^2)
\]

Using the fact that \( N_{1.} \) and \( N_{0.} \) are independent and after algebraic simplification, we get
\[
(6.1.19) \quad \text{Var}(f_1) = \frac{a^2 (\alpha + \beta)}{c \alpha \beta \bar{\beta}^2} + \frac{a^2 \alpha \beta}{c^2 \alpha \beta \bar{\beta}^2}
\]

However, the lower bound of \( g_1 \) as given by (2.1.18) is
\[
(\frac{a \alpha}{EN_{1.}}) [g_\alpha'(\alpha, \beta)]^2 + \frac{\beta \alpha}{EN_{0.}} [g_\beta'(\alpha, \beta)]^2 = \frac{a \alpha \beta}{c \alpha \beta \bar{\beta}^2} \cdot \frac{a^2}{c^2 \alpha \beta} + \frac{\beta \alpha}{c \beta \alpha \bar{\beta}^2} \cdot \frac{a^2}{c^2 \alpha \beta}
\]
\[
= \frac{a^2 (\alpha + \beta)}{c \alpha \beta \bar{\beta}^2} < V(f_1)
\]

showing uniqueness of \( g \).

The estimates of parameters \( \alpha \) and \( \beta \) are, then obtained in terms of \( N_{1.} \) and \( N_{0.} \) from the negative binomial distribution. The results are given in the following theorem.

**Theorem 6.1.1.** For all \( c \geq 2 \), MVUE's of \( \alpha \) and \( \beta \) exist and are given by
\[
(6.1.20) \quad \hat{\alpha} = 1 - \frac{c-1}{N_{1.} - 1} = \frac{N_{1.} - c}{N_{1.} - 1} = \frac{X - X_1 - c}{X - X_1 - 1},
\]
\[
(6.1.21) \quad \hat{\beta} = \frac{c-1}{N_{0.} - 1} = \frac{c-1}{Y + X_1 - 2}.
\]

The minimum variances of \( \hat{\alpha} \) and \( \hat{\beta} \) are given by,
\[
(6.1.22) \quad \text{Var}(\hat{\alpha}) = \frac{a^2}{c^2} \sum_{k=1}^{\infty} \alpha^k \binom{k+c-1}{c-1}^{-1},
\]
\[
\begin{align*}
(6.1.23) \quad \text{Var}(\beta) &= \beta^2 \sum_{k=1}^{c} \beta^k \left(\frac{k+c-1}{c-1}\right)^{-1}.
\end{align*}
\]

6.2. **Sampling Plan** \(S_{2c-1}(N_{10} + N_{01})\)

Since we have \(|N_{11} - N_{00}| = 1\), we have two kinds of possible outcomes. Either \(N_{10} = c\) and \(N_{01} = c-1\) or \(N_{10} = c-1\) and \(N_{01} = c\). We consider these cases in the following.

**Case (i)** \(N_{10} = c, N_{01} = c-1\).

The sampling scheme has following form of outcomes. Starting with an \(S\) and terminating with an \(F\) which is preceded by an \(S\), a typical sequence has the form,

\[
\underbrace{S \ldots S}_{s_1} \underbrace{F \ldots F}_{f_1} \underbrace{S \ldots S}_{s_2} \underbrace{F \ldots F}_{f_{c-1}} \underbrace{S \ldots S}_{s_c} | F.
\]

The following relations between \(N_{11}, N_{00}\) and \(X, Y\) hold:

\[
\begin{align*}
N_{11} &= X - c \\
N_{00} &= (Y-1) - (c-1) = Y - c.
\end{align*}
\]

By same arguments as given earlier the total number of ways of obtaining a sequence is \(\binom{x-1}{c-1}\binom{y-2}{c-2}\).

**Case (ii)** \(N_{10} = c-1, N_{01} = c\). The sampling scheme now has the following form of outcomes. Starting from \(F\) and terminating with an \(F\) which is preceded by an \(F\), we have

\[
\underbrace{F \ldots F}_{f_1} \underbrace{S \ldots S}_{s_1} \underbrace{F \ldots F}_{f_1} \ldots \underbrace{S \ldots S}_{s_{c-1}} \underbrace{F \ldots F}_{f_{c}} | S.
\]

The following relations among \(N_{11}, N_{00}\) and \(X, Y\) hold.
Again the total number of ways of obtaining such a sequence is given by \( \binom{x-2}{c-2} \binom{y-1}{c-1} \).

Combining the results in Cases (i) and (ii), the following relations hold

\[
\begin{align*}
N_{11} &= X - c \\
N_{10} &= c + X_1 - 1 \\
N_{01} &= c - X_1 \\
N_{00} &= Y - c \\
N_{1*} &= X + X_1 - 1, \text{ and } N_{0*} = Y - X_1;
\end{align*}
\]

The total number of ways such that \( X_1 = x_1, X = x, Y = y \) and \( N_{10} + N_{01} = 2c - 1 \) is \( \binom{x+x_1-2}{c+x_1-2} \binom{y-x_1-1}{c-x_1-1} \).

The joint distribution of \((X_1, X, Y)\), therefore is given by,

\[
(6.2.2) \quad p(x_1, x, y) = \binom{x+x_1-2}{c+x_1-2} \binom{y-x_1-1}{c-x_1-1} p(x_1 \geq 0)^{x_1 - x_1 - 1} \alpha^{x_1} \beta^{-y_1 - 1} y_1^{-x_1} a^{-c} b^{c+x_1-1} c-x_1^{-1} y_1^{-c},
\]

\[ x = c, c+1, ..., \]

\[ y = c, c+1, ... \]

The conditional distributions of \( X \) and \( Y \) given \( X_1 \) are similarly given by

\[
(6.2.3) \quad p(x | x_1) = \binom{x+x_1-2}{c+x_1-2} \alpha^{c+x_1-1} \alpha^{x_1} x_1^{-c}, \quad x = c, c+1, ...
\]

and

\[
(6.2.4) \quad p(y | x_1) = \binom{y-x_1-1}{c-x_1-1} \beta^{c-x_1} y_1^{-y_1} y_1^{-c}, \quad y = c, c+1, ...
\]

Hence we see that \( X \) and \( Y \) are conditionally independent since
\( p(x, y|x_1) = p(x|x_1)p(y|x_1) \).

From (6.2.1) and (6.2.2), we get the joint distribution of \( X_1, N_1 \), and \( N_0 \) as follows.

\[
(6.2.5) \quad p(x_1, n_1, n_0) = \left[ \frac{n_1!}{(c+x_1-2)! \alpha^{c+x_1-1}} \right] \left[ \frac{n_0!}{(c-x_1-1)! \beta^{c-x_1}} \right] p(x_1) p(n_1) p(n_0).
\]

Notice that \( P(x_1, n_1, n_0) = P(n_1|x_1)P(n_0|x_1)p(x_1) \).

The above results are summarized in the following proposition.

**Proposition 6.2.1.** Given \( X_1 \) the random variables \( X \) and \( Y \) or random variables \( N_1 \) and \( N_0 \) are independent.

The conditional distributions of \( N_1 \) and \( N_0 \) are negative binomial given by

\[
(6.2.6) \quad N_1|x_1 \sim \text{NB}(c+x_1-1, \alpha)
\]

\[
(6.2.7) \quad N_0|x_1 \sim \text{NB}(c-x_1, \beta).
\]

The following properties can be easily obtained using Proposition 6.2.1.

(a) \( E(N_1|x_1) = \frac{c+x_1-1}{\alpha} \), \( \text{Var}(N_1|x_1) = \frac{(c+x_1-1)\alpha}{\alpha^2} \).

\( G_{N_1}(t) = (\frac{\alpha t}{1-\alpha t})^{c+x_1-1} \).

But since \( N_1 = X + X_1 - 1 \)

\( E(X|x_1) = 1 - x_1 + \frac{c+x_1-1}{\alpha} \), \( V(X|x_1) = \frac{(c+x_1-1)\alpha}{\alpha^2} \).

Hence

\[
(6.2.8) \quad \begin{cases}
E(X) = q + \frac{c-q}{\alpha} \\
V(X) = E(V(X|x_1)) + V(E(X|x_1)) = \frac{(c-q)\alpha}{\alpha^2} + \frac{\alpha^2 pq}{\alpha^2},
\end{cases}
\]
\[ G_{x|x_1}(t) = t^{1-x_1} G_{N_1|x_1}(t) = t^{1-x_1} \left( \frac{\alpha t}{1-\alpha t} \right)^{c-x_1-1}, \]

(6.2.9) \[ G_x(t) = \left( q t + \frac{c-x_1 t}{1-\alpha t} \right)^{c-1} \]

and

(6.2.10) \[ \text{Cov}(X, X) = \frac{pc}{\alpha} - p(q + \frac{c+q}{\alpha}) = \frac{pq\alpha}{\alpha}. \]

(b) \[ E(N_0|x_1) = \frac{c-x_1}{\beta}, \quad \text{Var}(N_1|x_1) = \frac{(c-x_1)^2}{\beta^2} \]

\[ G_{N_0}(t) = \left( \frac{\beta t}{1-\beta t} \right)^{c-x_1}. \]

But since \( N_0 = c - x_1 \)

\[ E(X|x_1) = x_1 + \frac{c-x_1}{\beta}, \quad \text{Var}(Y|x_1) = \frac{(c-x_1)^2}{\beta^2} \]

Hence

(6.2.11) \[ \begin{cases} E(Y) = p + \frac{c-p}{\beta} \\ \text{Var}(Y) = \frac{(c-p)^2}{\beta^2} + \frac{pq}{\beta^2} \end{cases} \]

\[ G_{Y|x_1}(t) = t^{x_1} G_{N_0|x_1}(t) = t^{1-x_1} \left( \frac{\beta t}{1-\beta t} \right)^{c-x_1} \]

Thus

(6.2.12) \[ G_Y(t) = \left( pt + \frac{q\beta t}{1-\beta t} \right)^{c-1} \left( \frac{\beta t}{1-\beta t} \right)^{c-x_1} \]

and

(6.2.13) \[ \text{Cov}(X, Y) = -\frac{pq\beta}{\beta} \]

(c) Since given \( X_1, X \) and \( Y \) are independent, with \( N = X + Y \)

we have
\[ G_N(t) = G_X(t) \cdot G_Y(t) = t^{\frac{\bar{c}t}{1-\alpha t}}(\frac{\beta t}{1-\beta t})^{c-x_1} \]

Hence

\[ (6.2.14) \quad G_N(t) = t^{\frac{\bar{c}t}{1-\alpha t} + \frac{\alpha t}{\beta t}}(\frac{\bar{c}t}{1-\alpha t})^{c-1}(\frac{\beta t}{1-\beta t}) \]

From this p.g.f. of \( N \) or directly from (a) and (b) we get

\[ \begin{align*}
\text{EN} &= 1 + \frac{c-\alpha}{\bar{c}} + \frac{(c-p)}{\beta} \\
\text{Var}(N) &= \frac{(c-\alpha)\alpha}{\bar{c}^2} + \frac{(c-p)\beta}{\beta^2} + pq\left(\frac{1}{\alpha} - \frac{1}{\beta}\right)^2
\end{align*} \]

and

\[ E(XY|x_1) = E(X|x_1) \cdot E(Y|x_1) = \left(1-x_1+\frac{c+x_1-1}{\alpha}\right)(x_1+\frac{c-x_1}{\beta}) \]

hence

\[ EXY = \frac{\alpha c}{\bar{c}} + \frac{\alpha c}{\beta} + \frac{c(c-1)}{\bar{c}^2} \]

Thus

\[ \text{Cov}(X, Y) = -pq\left(\frac{1}{\alpha} - 1\right)\left(\frac{1}{\beta} - 1\right) \]

6.2.2. **Efficient Estimation**

From (6.2.1), we see that the plan \( S_{2c-1}(N_{10}+N_{01}) \) does not belong to any of the plans \( S(N_{11}+N_{10}), S(N_{11}), S(N_{10}), S(N_{01}+N_{00}), S(N_{00}) \) or \( S(N_{01}) \). Therefore, it follows that this plan is not efficient for any \( g(\alpha, \beta) \).

We shall show that when \( X_1 \) is given, \( S_{2c-1}(N_{10}+N_{01}) \) is efficient. From Section 6.2.1, we find the following.

\[ E\left(\frac{N_{1}^{*}}{c+x_1-1} \mid x_1\right) = \frac{1}{\alpha} \]

\[ \text{Var}\left(\frac{N_{1}^{*}}{c+x_1-1} \mid x_1\right) = \frac{\alpha}{(c+x_1-1)^2 \alpha^2} \]
\[ E\left( \frac{N_{0+}}{c-x_1} \mid x_1 \right) = \frac{1}{\beta} \]

\[ \text{Var}\left( \frac{N_{0+}}{c-x_1} \mid x_1 \right) = \frac{\beta}{(c-x_1)^2} \]

Also for any estimator \( f \) such that

\[ E(f \mid x_1) = \frac{a}{\alpha} + \frac{b}{\beta} + d = g(\alpha, \beta) \]

We have from (2.1.18)

\[ \text{Var}(f \mid x_1) \geq \frac{a^2}{(c+x_1-1)\alpha^2} + \frac{b^2}{(c-x)\beta^2} \]

The above results can be stated in the following theorem.

**Theorem 6.2.1.** Given \( X_1 = x_1 \), sampling plan \( S_{2c-1}(N_{10} + N_{01}) \) is efficient and any non-constant estimator \( f \) of the form

\[ f = a\left(\frac{N_{10}}{c+x_1-1}\right) + b\left(\frac{N_{0+}}{c-x_1}\right) + d \]

is an efficient estimator for

\[ g(\alpha, \beta) = \left(\frac{a}{\alpha} + \frac{b}{\beta}\right) + d. \]

### 6.2.3 Estimation of parameters

Since \( N_{1+} \) and \( N_{0+} \) are distributed as negative binomial for a given \( X_1 \), we have from (6.2.6) and (6.2.7) and \( c \geq 3 \),

\[ \alpha = 1 - \frac{N_{10}-1}{N_{1-1}} = \frac{N_{11}}{N_{1-1}} = \frac{x-c}{x+x_1-2} \quad X = x = c, c+1, \ldots \]

\[ \beta = \frac{N_{01}-1}{N_{0-1}} = \frac{c-x_1-1}{x_1-1} \quad Y = y = c, c+1, \ldots \]

Notice that \( \hat{\alpha} \) and \( \hat{\beta} \) are unbiased estimates of \( \alpha \) and \( \beta \) respectively for \( c \geq 3 \).
For \( c = 2 \), we define \( \hat{\alpha} \) and \( \hat{\beta} \) as follows.

\[
(6.2.20) \quad \hat{\alpha} = \begin{cases} 
\frac{x-2}{x-1} & x = x \geq 2, \ x_1 = 1 \\
1 & x = x \geq 3, \ x_1 = 0 \\
0 & x = 2 \end{cases}
\]

\[
(6.2.21) \quad \hat{\beta} = \begin{cases} 
\frac{1}{y-1} & y = y \geq 2, \ x_1 = 0 \\
1 & y = 2, \ x_1 = 1 \\
0 & y = y \geq 3 \end{cases}
\]

Also since, given \( X_1, X \) and \( Y \) are distributed an negative binomial \( \text{NB}(2, \alpha) \), and \( \text{NB}(2, \beta) \) respectively, we find after simplification that for \( c = 2 \), \( \hat{\alpha} \) and \( \hat{\beta} \) are also unbiased estimates of \( \alpha \) and \( \beta \).

These results are stated in the following theorem.

**Theorem 6.2.2.** Under the sampling plans \( S(N_{10} + N_{01}) \) MVUE's for \( \alpha \) and \( \beta \) exist and they are given by (6.2.18) and (6.2.19) for \( c \geq 3 \) and by (6.2.20) and (6.2.21) for \( c = 2 \).
CHAPTER VII

DISCUSSION

The study of unbiased estimation in Markov chains leads us to the problem of sequential estimation. It entails the problem of finding optimal stopping rule or sampling plan. The usual criterion of optimality in unbiased estimation is in terms of the variances of the estimators. The bound of the variance of an estimator is given by the information inequality and the equality is attained for some "efficient" sampling plan $S$, "efficient" estimator $f$, and "efficiently estimable" function $g = Ef$ of the parameters.

The main purpose of this dissertation has been to find and study all possible triplets $(S, f, g)$ for which the information inequality becomes equality in a two-state Markov chain.

In Chapter II, when $g$ is a function of $\alpha$ alone, such triplets were found to be

$[S(N_1), aN_1 + b, ac_1 + b]$  
$[S(N_11), aN_1 + b, ac'/a + b]$  
$[S(N_10), aN_1 + b, ac'/a + b]$  

where $c$ is the constant for which $Z = c$ in plan $S(Z)$ and $a$, $b$ are arbitrary constants. Furthermore, when $g$ is a function of both parameters $\alpha$ and $\beta$,

$[S_{2e}(N_1 + N_0), aN_1 + bN_0 + d, c(a/\alpha + b/\beta) + d]$  

86
was found to be the only triplet for which equality holds in the information inequality.

In later chapters, sampling distributions and other properties of efficient sampling plans were obtained. In each plan certain optimal estimators of $\alpha$ and $\beta$ were given.

Further study in this direction seems promising for more general Markov processes. In the area of continuous-time Markov process, some limited work has already been done. Trybula (1968) considered the problem for certain one-parameter processes with independent increments, namely Poisson process, negative binomial process and Wiener process, and has shown that simple, inverse and step sampling plans are the only efficient ones. For example, in the case of Poisson process with parameter $\lambda$, he has shown that the only efficiently estimable functions are

$$g_1(\lambda) = a\lambda + b, \quad g_2(\lambda) = a/\lambda + b, \quad g_3(\lambda) = a/(1-r\lambda) + b.$$  

$g_1$ is estimated efficiently only by a simple plan in which observation is continued for a fixed time interval. $g_2$ is estimated efficiently only by an inverse plan in which observation is continued until the time when the process attains the prescribed value. $g_3$ is estimated efficiently only by a step plan in which the observation is continued until its sample function crosses the line $k = (1/r)(t - s)$ where $r$ and $s$ are positive constants and $k$ and $s$ stand for the value of the process and time respectively. However, whether similar conclusions can be drawn for any process with independent increments is still an open question.
In Markov chains with more than two states, it is not clear whether there can be a sampling plan that is efficient for functions of the whole transition probability matrix, or functions of two or more rows of transition probability matrix. In view of the discussions made in section 2.4 it seems unlikely. However, it is possible to obtain efficient estimators of certain functions of a row of the transition probability matrix by extension of results in section 2.3. The results depend in part on certain results in the paper of Bhat and Kulkarni (1966).

So far our study has been based on the implicit assumption that the transition probabilities $p_{ij}$ and $p'_{ij}$, $(i|i')$ are not related. Another interesting model is a Markov chain in which the transition probabilities are functionally related as $p_{ij} = p_{ij}(\theta)$, where $\theta = (\theta_1, \ldots, \theta_k)$ is an unknown vector parameter, and the functional relation may or may not be known. The problem is to find efficient triplets $(S, f, g(\theta))$ under some known functional relationship, or to find some conditions under which there exist efficiently estimable functions under the sampling plan considered. It seems possible to solve the above problem with similar techniques as discussed in this dissertation.
BIBLIOGRAPHY


