# On a Generalized Inverse Binomial Sampling Plan

Do Sun Bai*
Seong-In Kim*
Jung-Kyun Lee**

## 1. Introduction

In many applications one is concerned with repeated Bernoulli trials whose parameter (success probability) is usually unknown and has to be estimated from a sample. The probability distribution and statistical inference on the repeated independent Bernoulli trials have been studied extensively for the cases of fixed sample size sampling plan, and inverse binomial sampling plan in which observations are continued until a preassigned number of successes are obtained. See, for example, Haldane [4], Girschick et al. [3], DeGroot [2] and Johnson and Kotz [5].

Another interesting sampling plan is the one in which observations are to be continued until at least $m_1$ successes and at least $m_2$ failures are obtained where $m_1$ and $m_2$ are preassigned numbers. The probability distribution of the sample size of this sampling plan has been studied by McCarthy [7], Bai [1] and Lee [6]. Since it is a generalization of the inverse binomial sampling plan (IBSP) in which case $m_2=0$, we will call it a generalized inverse binomial sampling plan (GIBSP).

* Department of Industrial Science, Korea Advanced Institute of Science.
**Korea Institute of Science and Technology.

In this paper, the probability distribution of a sufficient statistic for the parameter is studied for the GIBSP based on independent Bernoulli trials, and a simple unbiased estimator based on this sufficient statistic is presented. The relative efficiency of GIBSP to IBSP in unbiased estimation is then discussed for various values of the parameter.

## 2. Probability Distribution and Parameter Estimation

Let the point $(x,y)$ in $xy$-plane with nonnegative integral coordinates represent an outcome of $x+y$ independent Bernoulli trials where $x$ and $y$ denote the numbers of successes and failures respectively. A sampling plan then defines a *region* $R$ representing the set of all possible intermediate and terminal outcomes (points) in the course of $0, 1, \ldots, N$ trials, and a *boundary set* $B$ of $R$ representing the set of all possible terminal outcomes. For a sampling plan, $P\{(x,y)\}=K(x,y)p^xq^y$ where $(x,y)\in R$ and $K(x,y)$ is the number of possible ways to reach $(x,y)$ from $(0,0)$. The region $R$ defined by a sampling plan will be said to be *closed* if $\sum_{(x,y)\in B} P\{(x,y)\}=1$, and *simple* if all the points with nonnegative integral coordinates between any two points of $R$ on the line $x+y=n$ are also in $R$.

In IBSP$(m,p)$, the trials are to be continued until a preassigned number $m$ of successes are obtained in a sequence of independent Bernoulli trials with success probability $p=1-q$. The region $R$ defined by IBSP is easily seen to be closed and simple. The sample size $N$ is here a random variable with the well-known negative binomial distribution,

$$p_m(n)=P[N=n]=\binom{n-1}{m-1}p^mq^{n-m}, \qquad n=m,m+1,\ldots \tag{1}$$

The expected sample size is

$$E(N)=m/p, \tag{2}$$

and the unique minimum variance unbiased estimator of $p$ is easily obtained as

$$\hat{p}=\frac{m-1}{N-1}, \quad m\geqslant 2.$$

Haldane [4] gives the variance of $\hat{p}$ in the form of an integral which, after repeated integration by parts, becomes

$$\begin{aligned}\mathrm{Var}_m(\hat{p}) &= p^2\sum_{j=1}^{\infty}\binom{m-1+j}{j}^{-1}q^j \\ &= \frac{p^2q}{m}\left\{1+\frac{2!}{m+1}q+\frac{3!}{(m+1)(m+2)}q^2+\cdots\right\} \\ &\approx \frac{m(n-m)}{n^2(n-1)} \qquad (3)\end{aligned}$$

whereas Degroot [2] gives it in the form of

$$\mathrm{Var}_m(\hat{p}) = \frac{(m-1)p^m}{q^{m-1}}\left\{(-1)^{m-1}\log p+\sum_{i=1}^{m-2}\frac{(-1)^{m-i}}{i}\left(-\frac{q}{p}\right)^i\right\}-p^2 \quad (4)$$

In GIBSP($m_1, m_2, p$), the trials are to be continued until at least $m_1$ successes and at least $m_2$ failures are obtained where $m_1$ and $m_2$ are pre-assigned positive integers. That is, independent Bernoulli trials $U_1, U_2, \ldots$ are to be observed until

$$X=\sum_{i=1}^{N}U_i=m_1 \quad \text{and} \quad Y=N-\sum_{i=1}^{N}U_i\geqslant m_2$$

or

$$X=\sum_{i=1}^{N}U_i\geqslant m_1 \quad \text{and} \quad Y=N-\sum_{i=1}^{N}U_i=m_2.$$

The boundary set $B$ of the region defined by GIBSP($m_1, m_2, p$) can be represented as $B=B_1\cup B_0$ where

$$B_1=\{(x,y);x=m_1,\ y=m_2, m_2+1,\ldots\}$$
$$B_0=\{(x,y);y=m_2,\ x=m_1,m_1+1,\ldots\}$$

We note here that to reach a point in $B_1-(m_0, m_1)$ the last trial must result in a success ($U_N=1$), and that to reach a point in $B_0-(m_0, m_1)$ the last trial must result in a failure ($U_N=0$). The point $(m_0, m_1)$ can be reached in both ways.

The joint distribution of $(N, U_N)$ is then given by

$$p_{m_1,m_2}(n,u_n)=P[N=n,\ U_N=u_n]$$

$$= \left\{\binom{n-1}{m_1-1} p^{m_1} q^{n-m_1}\right\}^{u_n} \left\{\binom{n-1}{m_2-1} p^{n-m_2} q^{m_2}\right\}^{1-u_n}$$

$$= \binom{n-1}{m_1-1}^{u_n} \binom{n-1}{m_2-1}^{1-u_n} p^{u_n m_1 + (1-u_n)(n-m_2)} q^{u_n(n-m_1)+(1-u_n)m_2} \qquad (5)$$

where $n=m_1+m_2, m_1+m_2+1, \ldots$ and $u_n=0, 1$. The distribution of $N$ is given by

$$p_{m_1,m_2}(n) = P[N=n]$$

$$= \binom{n-1}{m_1-1} p^{m_1} q^{n-m_1} + \binom{n-1}{m_2-1} p^{n-m_2} q^{m_2} \qquad (6)$$

where $n=m_1+m_2, m_1+m_2+1, \ldots$ .

The fact that $p_{m_1, m_2}(n)$ is really a probability distribution, that is, the region $R$ defined by GIBSP$(m_1, m_2, p)$ is closed, can easily shown using following identity (See Bai [1]).

$$p^{m_1+1} \sum_{j=0}^{m_2} \binom{m_1+j}{j} q^j + q^{m_2+1} \sum_{j=0}^{m_1} \binom{m_2+j}{j} p^j \equiv 1. \qquad (7)$$

The $k$-th ascending factorial moment of $N$ is given by

$$E_{m_1,m_2}\{N^{[k]}\} = E_{m_1,m_2}\{N(N+1)\ldots(N+k-1)\}$$

$$= \sum_{n_1=m_1+m_2}^{\infty} n^{[k]} \binom{n-1}{m_1-1} p^{m_1} q^{n-m_1} + \sum_{n=m_1+m_2}^{\infty} n^{[k]} \binom{n-1}{m_2-1} p^{n-m_2} q^{m_2}$$

$$= \frac{m_1^{[k]}}{p^k} \left[1 - \sum_{j=0}^{m_2-1} \binom{m_1+k-1+j}{i} p^{m_1+k} q^j\right]$$

$$+ \frac{m_2^{[k]}}{q^k} \left[1 - \sum_{j=0}^{m_1-1} \binom{m_2+k-1+j}{j} p^j q^{m_2+k}\right]$$

$$= \frac{m_1^{[k]}}{p^k} [1 - I_p(m_1+k, m_2)] + \frac{m_2^{[k]}}{p^k} [1 - I_q(m_2+k, m_1)] \qquad (8)$$

where $I_x(a,b)$ is the incomplete bata function as tabulated by Pearson [8].

From the form of the joint distribution (5) it can be seen that $(N, U_N)$ forms a sufficient statistic for $p$. Therefore, we present here an unbiased estimator of $p$ which is a function of $(N, U_N)$. To find it, we use the following result of Girschick et al. [3].

"*For any closed region* R, *the function* $\hat{p}(x,y)$ *defined by* $\hat{p}(x,y) = K^*(x,y)/K$

*$(x,y)$ is an unbiased estimate of $p$ where $K(x,y)$ and $K^*(x,y)$ are the numbers of all possible ways to reach $(x,y)\in B$ from $(0,0)$ and $(1,0)$ respectively.*"

In GIBSP$(m_1, m_2, p)$, we have

$$K(x,y)=\binom{n-1}{m_1-1}^{u_n}\binom{n-1}{m_2-1}^{1-u_n}$$

$$K^*(x,y)=\binom{n-2}{m_1-2}^{u_n}\binom{n-2}{m_2-1}^{1-u_n}$$

where $(x,y)\in B$ and $x+y=n$.

Hence an unbiased estimator $\hat{p}$ of $p$ is given by

$$\hat{p}=\left\{\binom{N-2}{m_1-2}\Big/\binom{N-1}{m_1-1}\right\}^{U_N}\left\{\binom{N-2}{m_2-1}\Big/\binom{N-1}{m_2-1}\right\}^{1-U_N}$$

$$=\frac{U_N(m_1-1)+(1-U_N)(N-m_2)}{N-1},\quad m_1\geqslant 2,\ m_2\geqslant 1. \tag{9}$$

The fact that $\hat{p}$ is unbiased can also be shown directly using (7) as follows.

$$E_{m_1,m_2}(\hat{p})=\sum_{n=m_1+m_2}^{\infty}\left[\frac{m_1-1}{n-1}\right]\binom{n-1}{m_1-1}p^{m_1}q^{n-m_1}+\sum_{n=m_1+m_2}^{\infty}\left[\frac{n-m_2}{n-1}\right]\binom{n-1}{m_2-1}p^{n-m_2}q^{m_2}$$

$$=p\sum_{j=m_2}^{\infty}\binom{m_1-2+j}{j}p^{m_1-1}q^j+p\sum_{j=m_1-1}^{\infty}\binom{m_2-1+j}{j}p^jq^{m_2}$$

$$=2p-p\left[1-q^{m_2}\sum_{j=0}^{m_1-2}\binom{m_2-1+j}{j}p^j\right]+pq^{m_2}\sum_{j=0}^{m_1-2}\binom{m_2-1+j}{j}p^j$$

$$=p$$

Girschick et al. [3] have shown that the necessary condition for the uniqueness of an unbiased estimator for the closed region $R$ is that $R$ be simple. Since $R$ is obviously not simple, it follows that our $\hat{p}$ may not be the only unbiased estimator. It is interesting to note that the maximum likelihood estimator of $p$ is given by $[U_Nm_1+(1-U_N)(N-m_2)]/N$ which is obviously biased.

The variance of $\hat{p}$ for $m_1, m_2\geqslant 2$ is obtained using (3) as

$$\mathrm{Var}_{m_1,m_2}(\hat{p})=\sum_{n=m_1+m_2}^{\infty}\left[\frac{m_1-1}{n-1}\right]^2\binom{n-1}{m_1-1}p^{m_1}q^{n-m_1}$$

$$+\sum_{n=m_1+m_2}^{\infty}\left[\frac{n-m_2}{n-1}\right]^2\binom{n-1}{m_2-1}p^{n-m_2}q^{m_2}-p^2$$

$$= p^2 \sum_{j=1}^{\infty} \binom{m_1-1+j}{j}^{-1} q^j + q^2 \sum_{j=1}^{\infty} \binom{m_2-1+j}{j}^{-1} p^j$$

$$- \sum_{j=0}^{m_2-1} \left[ \frac{m_1-1}{m_1-1+j} \right]^2 \binom{m_1-1+j}{j} p^{m_1} q^j$$

$$- \sum_{j=0}^{m_1-1} \left[ \frac{j}{m_2-1+j} \right]^2 \binom{m_2-1+j}{j} p^j q^{m_2} + p^2 \qquad (10)$$

Using Degroot's expression (4), it can also be written as

$$\mathrm{Var}_{m_1, m_2}(\hat{p}) = \frac{(m_1-1)p^{m_1}}{q^{m_1-1}} \left\{ (-1)^{m_1-1} \log p + \sum_{j=1}^{m_1-2} \frac{(-1)^{m_2-j}}{j} \left( \frac{q}{p} \right)^j \right\}$$

$$+ \frac{(m_2-1)q^{m_2}}{p^{m_1-1}} \left\{ (-1)^{m_2-1} \log q + \sum_{j=1}^{m_2-2} \frac{(-1)^{m_1-j}}{j} \left( \frac{p}{q} \right)^j \right\}$$

$$- \sum_{j=0}^{m_2-1} \left[ \frac{m_1-1}{m_1-1+j} \right]^2 \binom{m_1-1+j}{j} p^{m_1} q^j$$

$$- \sum_{j=0}^{m_1-1} \left[ \frac{j}{m_2-1+j} \right]^2 \binom{m_2-1+j}{j} p^j q^{m_2} - q^2 \qquad (11)$$

For the case of $m_1 \geqslant 2$, $m_2 = 1$, it can be shown that $\mathrm{Var}(\hat{p})$ is the same for both IBSP$(m_1, p)$ and GIBSP$(m_1, m_2, p)$

## 3. GIBSP vs IBSP

### 3.1 GIBSP$(m_1, m_2, p)$ vs IBSP$(m_1, p)$

In general, GIBSP's and IBSP's have similar property with respect to Var $(\hat{p})$; that is, small variances for extreme values of $p$ and large variances for medium values of $p$. On the other hand, the expected sample sizes of IBSP's decrease monotonically whereas those of GIBSP's decrease as $p$ approaches about 0.5 and then increase as $p$ increases further. In Fig.1 and Fig.2, Var $(\hat{p})$ and $E(N)$ are plotted as functions of $p$ for selected values of $m_1$ and

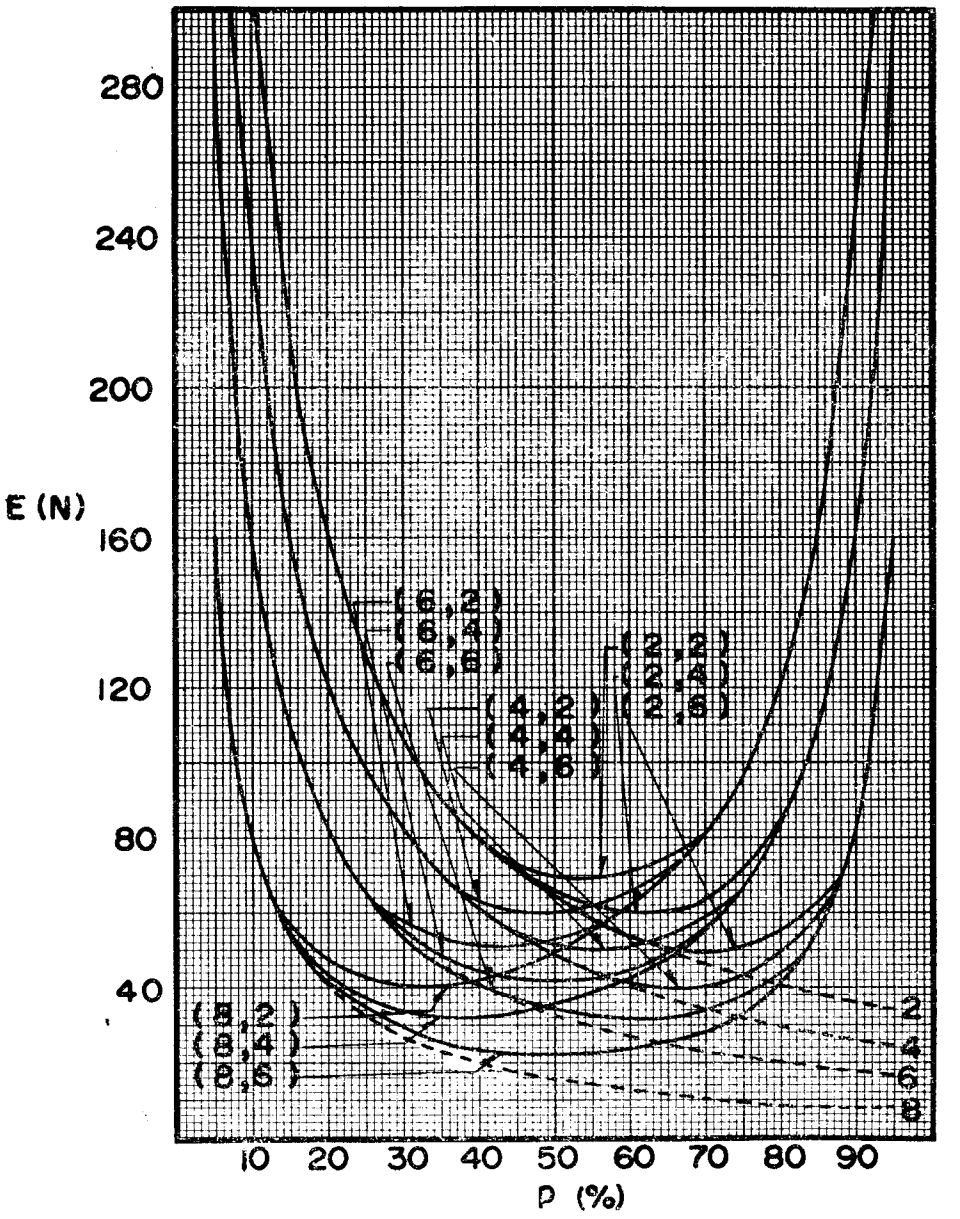

E (N)
280
240
200
160
120
80
40
10 20 30 40 50 60 70 80 90
P (%)
(6,2)
(6,4)
(6,6)
(4,2)
(4,4)
(4,6)
(2,2)
(2,4)
(2,6)
(8,2)
(8,4)
(8,6)
2
4
6
8

Fig. 1

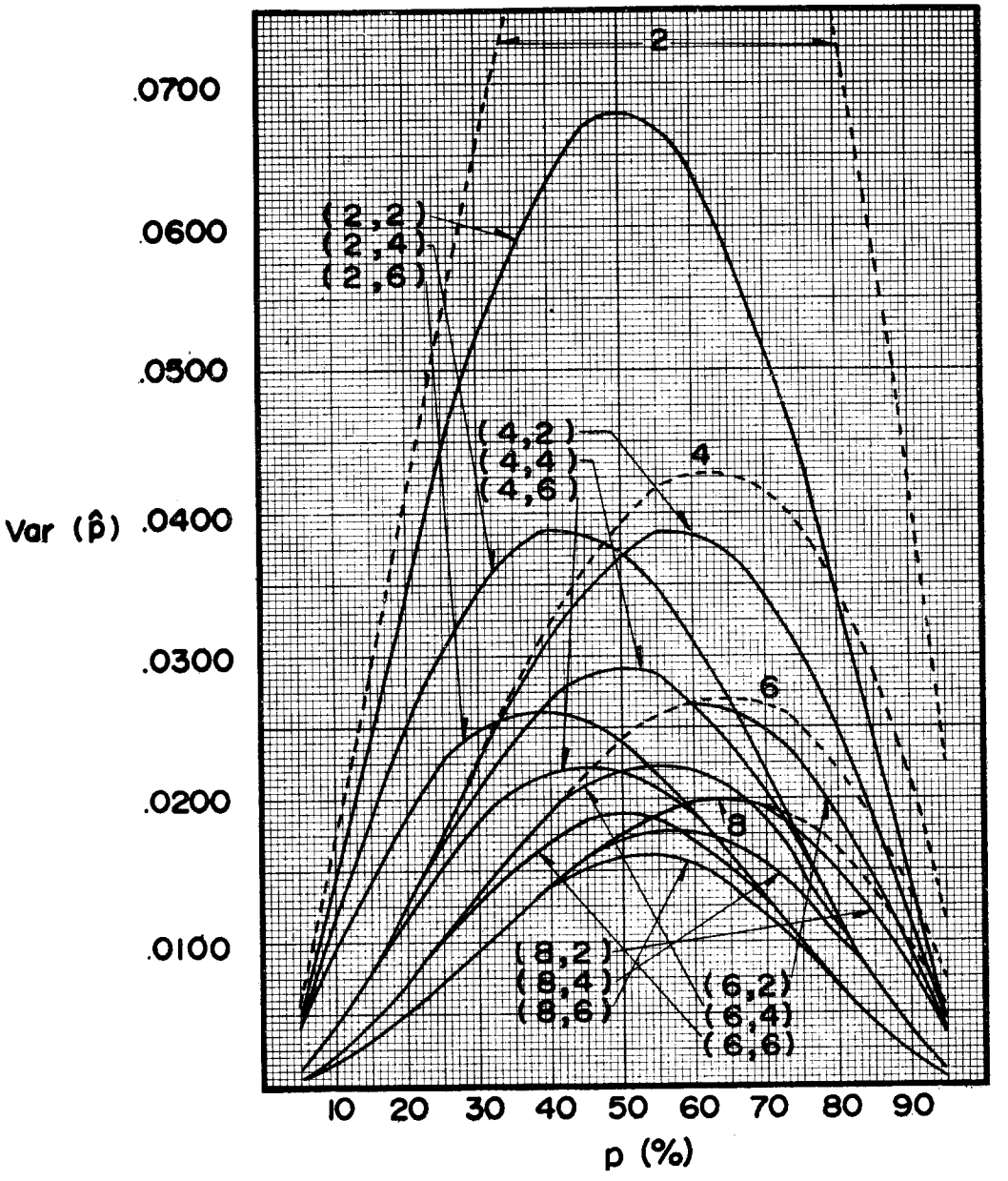

.0700
.0600
.0500
.0400
.0300
.0200
.0100
Var (p̂)
2
4
6
8
(2,2)
(2,4)
(2,6)
(4,2)
(4,4)
(4,6)
(8,2)
(8,4)
(8,6)
(6,2)
(6,4)
(6,6)
10 20 30 40 50 60 70 80 90
p (%)

Fig. 2

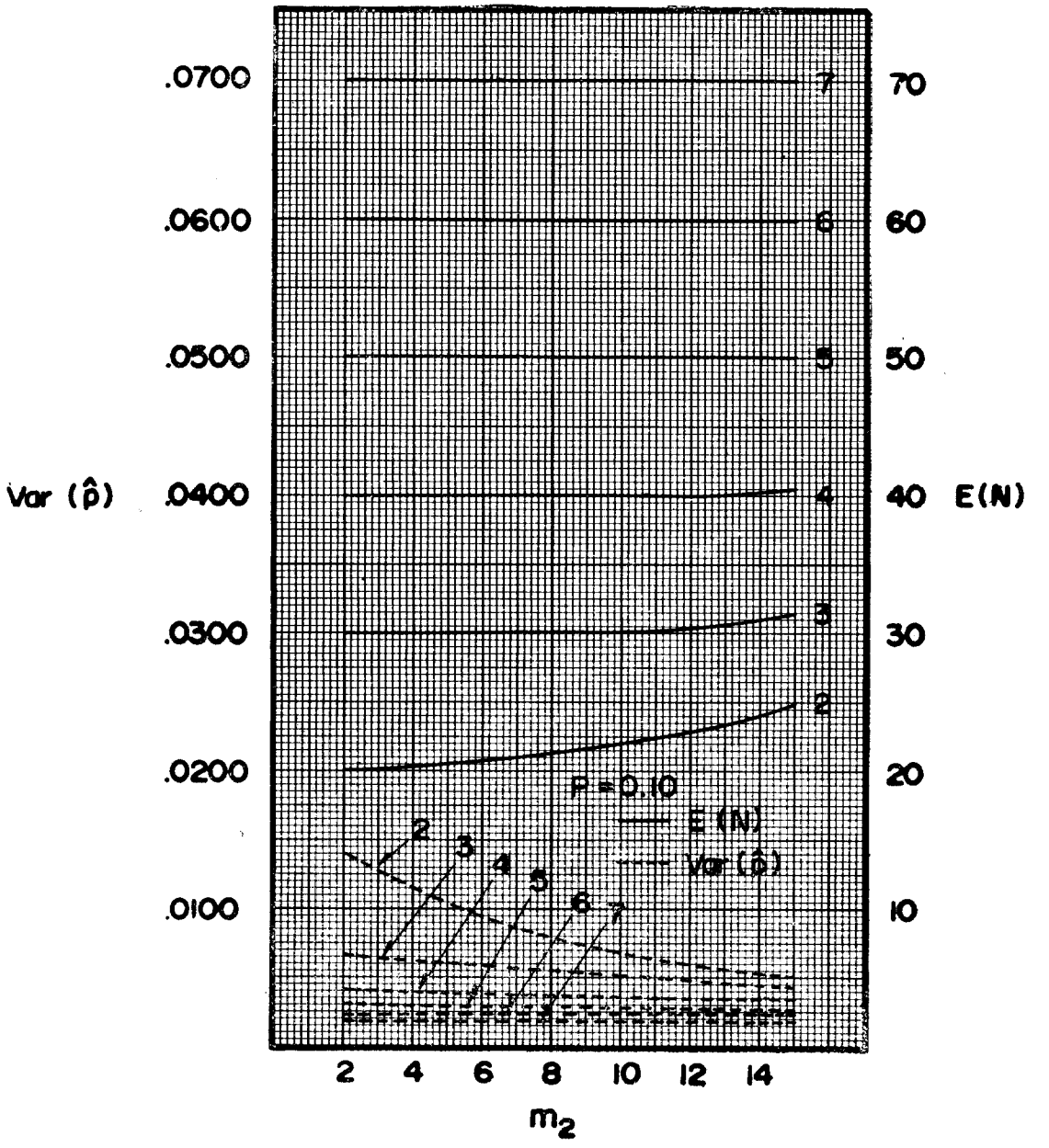

.0700
.0600
.0500
.0400
.0300
.0200
.0100
Var (p̂)
70
60
50
40
30
20
10
E(N)
7
6
5
4
3
2
P=0.10
E(N)
Var (p̂)
2
3
4
5
6
7
2
4
6
8
10
12
14
m2

Fig 3

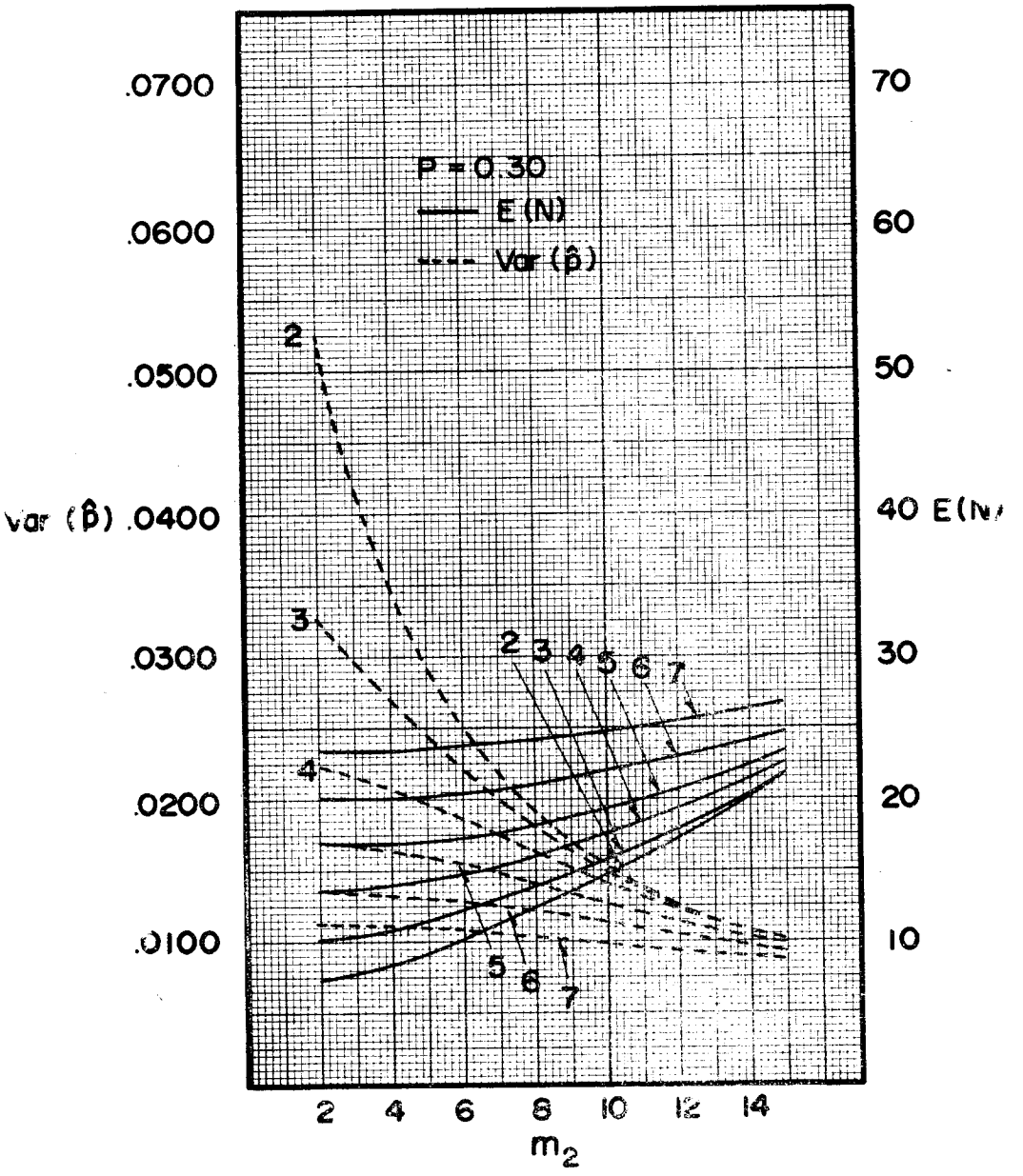

P = 0.30
E (N)
Var (p̂)
.0700
.0600
.0500
.0400
.0300
.0200
.0100
Var (p̂)
70
60
50
40
E(N)
30
20
10
2
3
4
5
6
7
2 4 6 8 10 12 14
m2

Fig. 4

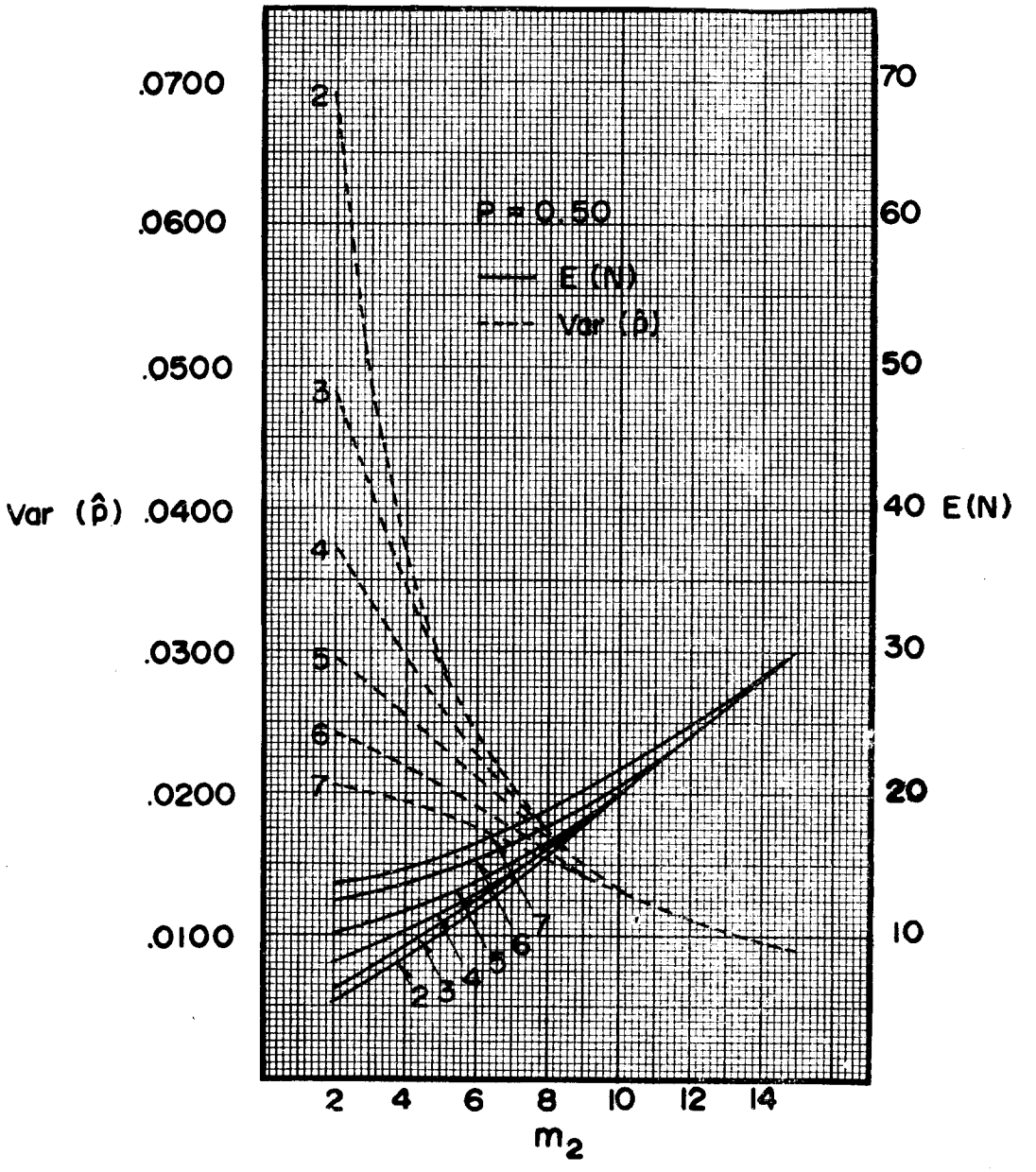

.0700
.0600
.0500
Var (p̂) .0400
.0300
.0200
.0100
70
60
50
40 E(N)
30
20
10
P = 0.50
E (N)
Var (p̂)
2
3
4
5
6
7
2 4 6 8 10 12 14
m2

Fig. 5

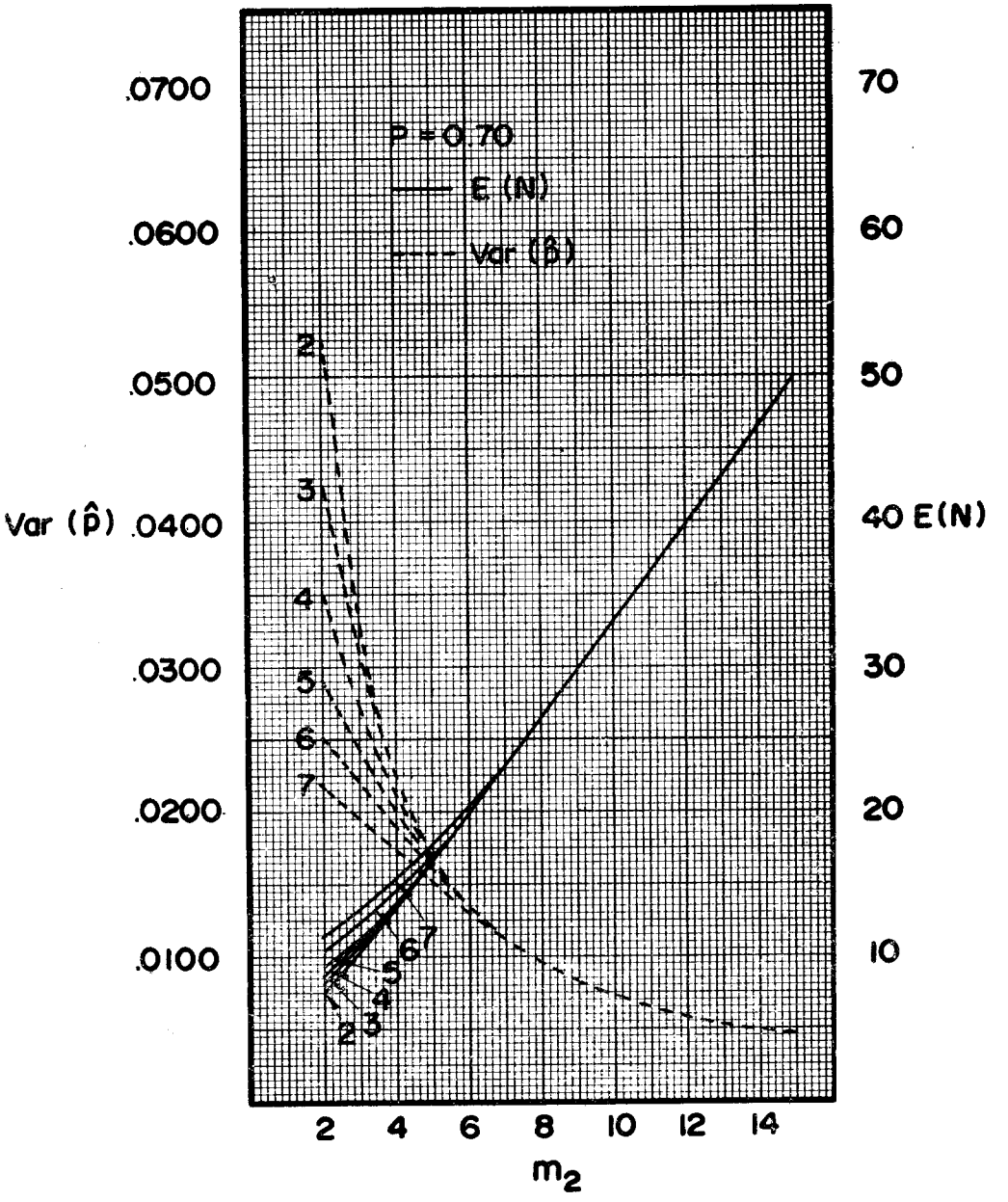

P = 0.70
E (N)
Var (p̂)
Var (p̂)
.0700
.0600
.0500
.0400
.0300
.0200
.0100
70
60
50
40 E(N)
30
20
10
2
3
4
5
6
7
2 4 6 8 10 12 14
m2

Fig. 6

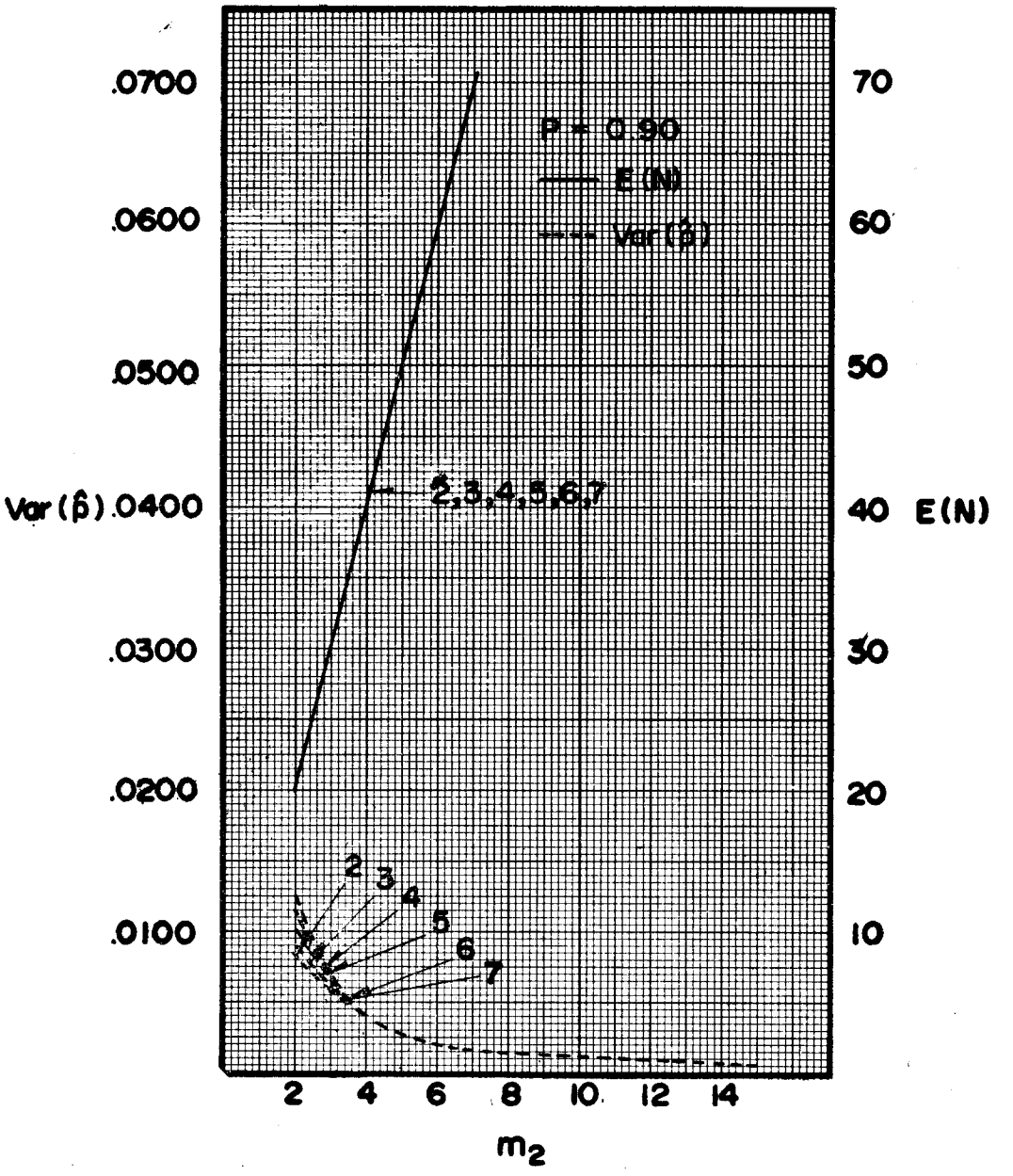

.0700
.0600
.0500
Var(p̂) .0400
.0300
.0200
.0100
70
60
50
40 E(N)
30
20
10
P = 0.90
E(N)
Var(p̂)
2,3,4,5,6,7
2
3
4
5
6
7
2 4 6 8 10 12 14
m2

Fig. 7

$(m_1, m_2)$. Fig.3 to Fig.7 show the values of $\mathrm{Var}_{m_1}(\hat{p})$ and $E_{m_1, m_2}(N)$ as $m_2$ varies for selected values of $m_1$ and $p$.

From Fig.1 to Fig.7, it is seen that, in general, $\mathrm{Var}_{m_1, m_2}(\hat{p}) \leqslant \mathrm{Var}_{m_1}(\hat{p})$ and $E_{m_1,m_2}(N) \geqslant E_{m_1}(N)$. Above figures also show that, for ver ysmall values of $p$, $\mathrm{Var}_{m_1, m_2}(\hat{p}) \approx \mathrm{Var}_{m_2}(\hat{p})$ and $E_{m_1, m_2}(N) \approx E_{m_1}(N)$, and both plans are almost equally effective. For intermediate values of $p$, $\mathrm{Var}_{m_1, m_2}(\hat{p})$ decreases considerably as $m_2$ increases whereas $E_{m_1, m_2}(N)$ increases moderately. For large values of $p$, $E_{m^1, m^2}(N)$ increases sharply as $m_2$ increases while the reduction in $\mathrm{Var}_{m^1, m_2}$ $(\hat{p})$ is small. However, since IBSP$(m_1, p)$ is primarily designed to be used when $p$ is not large, it will be more appropriate to campare GIBSP$(m_1, m_2,$ $q)$ with GIBSP$(m_2, q)$ when $p$ is large, in which case the two plans are similar.

Therefore, GIBSP$(m_1, m_2, p)$ can be advantegeously used in place of IBSP $(m_1, p)$ when a) high precision is required in estimating $p$, b) sampling is not expensive, and c) it is known a priori that $p$ is not large.

### 3.2 GIBSP$(m_1, m_2, p)$ vs IBSP$(m, p)$

If for some value $p^*$ of $p$, there exist $m_1 = m_1^*, m_2 = m_2^*$ and $m = m^*$ satisfying

$$\begin{aligned} &\mathrm{Var}_{m_1^*, m_2^*}(\hat{p}) \leqslant \mathrm{Var}_{m^*}(\hat{p}) \\ &E_{m_1^*, m_2^*}(N) \leqslant E_{m^*}(N) \end{aligned} \tag{12}$$

then GIBSP$(m_1^*, m_2^*, p)$ will be said to be more efficient than IBSP$(m^*, p)$ at $p = p^*$. If the inequalities (12) are reversed, vice versa. Otherwise, they are not compatible.

However it is difficult or even impossible to solve (12) mathematically. Thus comparisons were made using electronic computer. Fig.8 illustrates how to compare relative efficiency of the two types of sampling plans for a given value of $p$. Suppose there are five IBSP's corresponding to five different values of $m$ represented by points 1 to 5 respectively in the plane whose $x$-axis represents $\mathrm{Var}(\hat{p})$ and $y$-axis represents $E(N)$. For each GIBSP represented by a point below the shaded area, there exists at least one IBSP which is less efficient. For example, GIBSP A is more efficient than IBSP 3 or 4. We draw adverse conclusion for the plan represented by a point above the shaded area, and no conclusion for the shaded area.

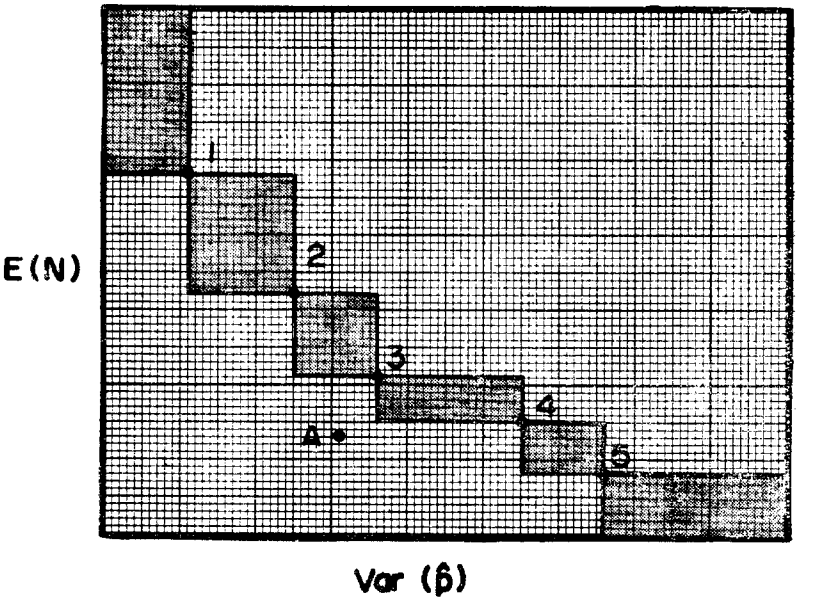

E(N)
1
2
3
4
5
A
Var (p̂)

Fig. 8

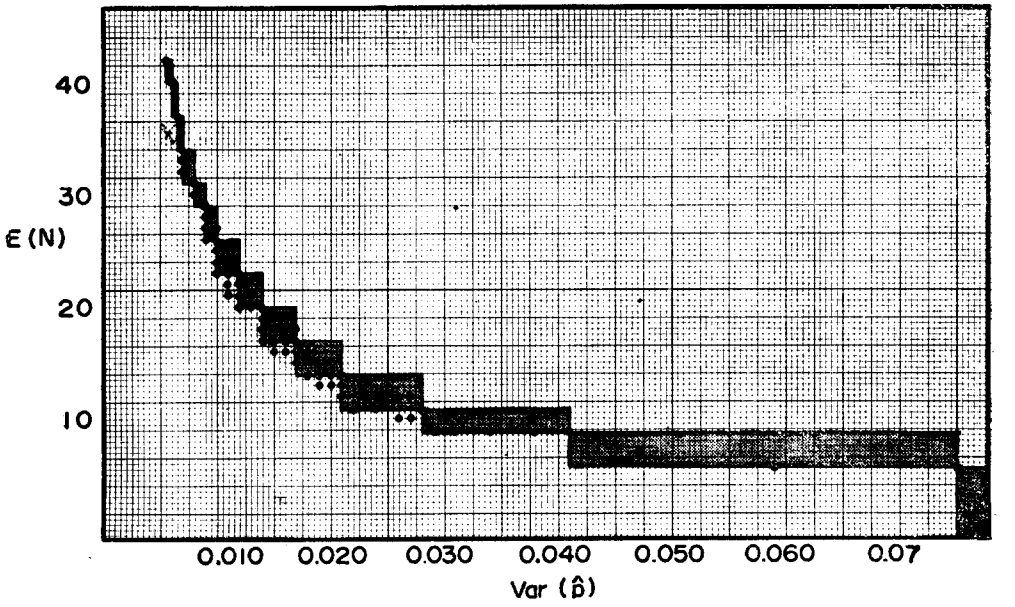
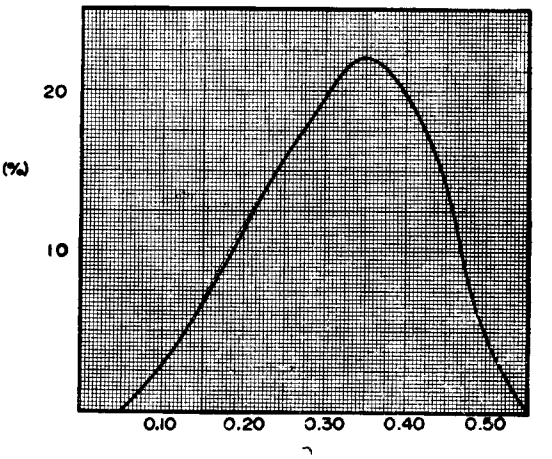

40
30
E(N)
20
10
0.010
0.020
0.030
0.040
0.050
0.060
0.07
Var (p̂)

Fig. 9

Table 1

| $p$ | Below | In | Above |
|---|---|---|---|
| 0.05 | 0 | 196 | 0 |
| 0.10 | 6 | 190 | 0 |
| 0.15 | 13 | 183 | 0 |
| 0.20 | 23 | 173 | 0 |
| 0.25 | 31 | 165 | 0 |
| 0.30 | 37 | 159 | 0 |
| 0.35 | 43 | 153 | 0 |
| 0.40 | 38 | 158 | 0 |
| 0.45 | 28 | 158 | 0 |
| 0.50 | 8 | 188 | 0 |
| 0.55 | 0 | 161 | 35 |
| 0.60 | 0 | 82 | 114 |

Fig. 10




19 values of $p$ ranging from 0.05 to 0.95 are considered. For each value of $p$, $\mathrm{Var}(\hat{p})$ and $E(N)$ are computed for 29 IBSP's with $m$ ranging from 2 to 30, and for 196 GIBSP's with $m_1$ and $m_2$ ranging from 2 to 15. Fig.9 shows the plans for $p=0.35$. Among 196 GIBSP's, 43 fall below the shaded area and none above the shaded area (See Appendix). Table 1 shows the results for $p$ from 0.05 to 0.60. It indicates that GIBSP's are effective for the values of $p$ less than 0.5. The percentage of efficient plans among 196 GIBSP's appears in Fig.10. As $p$ increases, the percentage increases and has peak at about 0.35 and then decreases. GIBSP is characterized by adding another stopping rule $m_2$, and this will increase $E(N)$ and decrease $\mathrm{Var}(\hat{p})$. For small values of $p$, the effect of $m_2$ on decreasing $\mathrm{Var}(\hat{p})$ will be larger compared to the effect on increasing $E(N)$. As $p$ increases, the net effect will increase, peak, and then decrease. This explains the reason why we get such a curve as Fig.10.

In conclusion, the use of a GIBSP may be justified at the lower half values of $p$. We also note that relative efficiency of any two IBSP's cannot be compared. However, this is not the case for GIBSP's. A GIBSP can be more efficient than another GIBSP.

## 4. Appendix

$p=0.35$

| IBSP | | | GIBSP | | |
|---|---|---|---|---|---|
| $m$ | $E(N)$ | $\mathrm{Var}_m(\hat{p})$ | $\mathrm{Var}_{m_1,m_2}(\hat{p})$ | $E(N)$ | $m_1, m_2$ |
| 3 | 8.57 | 0.04135 | 0.03754 | 7.89 | 2, 4 |
| 4 | 11.43 | 0.02785 | 0.02623 | 10.11 | 2, 6 |
| 4 | 11.43 | 0.02785 | 0.02259 | 11.39 | 2, 7 |
| 4 | 11.43 | 0.02785 | 0.02690 | 10.46 | 3, 5 |
| 4 | 11.43 | 0.02785 | 0.02378 | 11.34 | 3, 6 |
| 5 | 14.29 | 0.02084 | 0.01977 | 12.75 | 2, 8 |
| 5 | 14.29 | 0.02084 | 0.01753 | 14.15 | 2, 9 |
| 5 | 14.29 | 0.02084 | 0.01889 | 13.49 | 3, 8 |
| 5 | 14.29 | 0.02084 | 0.02063 | 13.07 | 4, 6 |
| 5 | 14.29 | 0.02084 | 0.01912 | 13.82 | 4, 7 |
| 6 | 17.14 | 0.01660 | 0.01573 | 15.60 | 2, 10 |
| 6 | 17.14 | 0.01660 | 0.01425 | 17.07 | 2, 11 |
| 6 | 17.14 | 0.01660 | 0.01541 | 16.03 | 3, 10 |
| 6 | 17.14 | 0.01660 | 0.01609 | 15.71 | 4, 9 |
| 6 | 17.14 | 0.01660 | 0.01480 | 16.82 | 4, 10 |
| 6 | 17.14 | 0.01660 | 0.01591 | 16.38 | 5, 8 |
| 7 | 20.00 | 0.01377 | 0.01301 | 18.56 | 2, 12 |
| 7 | 20.00 | 0.01377 | 0.01289 | 18.80 | 3, 12 |
| 7 | 20.00 | 0.01377 | 0.01365 | 18.02 | 4, 11 |
| 7 | 20.00 | 0.01377 | 0.01263 | 19.30 | 4, 12 |
| 7 | 20.00 | 0.01377 | 0.01303 | 19.03 | 5, 11 |
| 7 | 20.00 | 0.01377 | 0.01357 | 18.97 | 6, 9 |
| 7 | 20.00 | 0.01377 | 0.01290 | 19.64 | 6, 10 |
| 8 | 22.86 | 0.01176 | 0.01108 | 21.57 | 2, 14 |
| 8 | 22.86 | 0.01176 | 0.01104 | 21.72 | 3, 14 |
| 8 | 22.86 | 0.01176 | 0.01712 | 20.63 | 4, 13 |
| 8 | 22.86 | 0.01176 | 0.01092 | 22.01 | 4, 14 |
| 8 | 22.86 | 0.01176 | 0.01141 | 21.30 | 5, 13 |
| 8 | 22.86 | 0.01176 | 0.01070 | 22.55 | 5, 14 |
| 8 | 22.86 | 0.01176 | 0.01158 | 21.32 | 6, 12 |
| 8 | 22.86 | 0.01176 | 0.01096 | 22.31 | 6, 13 |
| 8 | 22.86 | 0.01176 | 0.01135 | 22.20 | 7, 11 |
| 9 | 25.71 | 0.01026 | 0.01021 | 23.43 | 4, 15 |

(continued)

| | IBSP | | | GIBSP | |
|---|---|---|---|---|---|
| $m$ | $E(N)$ | $\mathrm{Var}_m(\hat{p})$ | $\mathrm{Var}_{m_1,m_2}(\hat{p})$ | $E(N)$ | $m_1,m_2$ |
| 9 | 25.71 | 0.01026 | 0.01006 | 23.86 | 5, 15 |
| 9 | 25.71 | 0.01026 | 0.00981 | 24.56 | 6, 15 |
| 9 | 25.71 | 0.01026 | 0.00992 | 24.57 | 7, 14 |
| 9 | 25.71 | 0.01026 | 0.00947 | 25.58 | 7, 15 |
| 9 | 25.71 | 0.01026 | 0.01010 | 24.79 | 8, 12 |
| 9 | 25.71 | 0.01026 | 0.00976 | 25.41 | 8, 13 |
| 10 | 28.57 | 0.00910 | 0.00909 | 27.42 | 8, 15 |
| 10 | 28.57 | 0.00919 | 0.00909 | 27.42 | 9, 13 |
| 10 | 28.57 | 0.00919 | 0.00883 | 27.97 | 9, 14 |
| 11 | 31.43 | 0.00817 | 0.00806 | 30.57 | 10, 15 |

## REFERENCES


[1] Bai, Do Sun, "A Note on a Binomial Identity," *Journal of the Korean Statistical Society*, Vol. 4, No. 2, 109-112, 1975.

[2] DeGroot, M. H. "Unbiased Sequential Estimation for Binomial Population," *Annals of Mathematical Statistics*, Vol. 30, No. 1, 80-101, 1959.

[3] Girschick, M. A., Mosteller, F. and Savage, L. J., "Unbiased Etimation for Certain Binomial Sampling Problems with Applications," *Annals of Mathematical Statistics*, Vol. 13, No. 1, 13-23, 1946.

[4] Haldane, J. B. S., "On a Method of Estimating Frequencies," *Biometrika*, Vol. 33, 222-225, 1945.

[5] Johnson, N. L. and Kotz, S., *Discrete Distributions*, John Wiley and Sons Inc., New York, 1970.

[6] Lee, Jung-kyun,"A Study of Bernoulli Trials," Unpublished M. S. Thesis, Korea Advanced Institute of Science, Seoul, 1976.

[7] McCarthy, P. J., "Approximate Solutions for Means and Variances in a Certain Class of Box Problems," *Annals of Mathematical Statistics*, Vol. 18, No. 3, 349-383, 1947.

[8] Pearson, K., *Tables of the Incomplete Beta Function*, Cambridge University Press, London, 1934.