The Annals of Mathematical Statistics
1972, Vol. 43, No. 3, 871-897

A BERNOULLI TWO-ARMED BANDIT[1]

BY DONALD A. BERRY

University of Minnesota

One of two independent Bernoulli processes (arms) with unknown expectations ρ and λ is selected and observed at each of n stages. The selection problem is sequential in that the process which is selected at a particular stage is a function of the results of previous selections as well as of prior information about ρ and λ. The variables ρ and λ are assumed to be independent under the (prior) probability distribution. The objective is to maximize the expected number of successes from the n selections. Sufficient conditions for the optimality of selecting one or the other of the arms are given and illustrated for example distributions. The stay-on-a-winner rule is proved.

1. Introduction and statement of the problem. Let $\mathscr{R}$ and $\mathscr{L}$ denote independent Bernoulli processes with parameters—probabilities of success—ρ and λ respectively. Call $\mathscr{R}$ the *right arm* and $\mathscr{L}$ the *left arm*. An observation on either arm is called a *pull*. A right pull or a left pull is made at each of n stages and the result of the pull at each stage is known before a right or left pull is made at the next stage. The parameters ρ and λ associated with $\mathscr{R}$ and $\mathscr{L}$ are not known precisely but are themselves random variables. The sequences of successes and failures associated with the right and left arms are therefore not sequences of independent Bernoulli trials, but are independent conditional on the unknown quantities ρ and λ, so that pulls on the right and left arms are exchangeable—see, for example, (Feller (1966) Section VII 4)—rather than independent.

Let I_k denote the pattern of information present about $\mathscr{R}$ and $\mathscr{L}$ at stage $k + 1$; that is, after k pulls. The pattern of information or accumulated data, I_k, can always be regarded as a probability distribution on the unknown parameters ρ and λ. I_0 or I is the initial pattern of information and consists of an initial probability distribution for each of the parameters ρ and λ. Throughout this paper the parameters are assumed to be initially, and therefore also henceforth, statistically independent. The problem is to decide which arm to pull at stage $k + 1$ conditional on the accumulated data I_k; thus, the results of the first k pulls as well as the initial distributions of ρ and λ can affect the decision at stage $k + 1$.

Let $I = (R, L)$ denote a pair of arbitrary initial distributions; $R = R(\rho)$ and $L = L(\lambda)$. A success on the right arm changes R to a new distribution, say σR, and a success on the left arm changes L to σL; a failure on the right arm changes R to φR and on the left arm changes L to φL. Let $E(\rho \mid R)$ and $E(\lambda \mid L)$ represent

Received April 19, 1971; revised September 1971.

[1] This paper is adapted in part from the author's Ph. D. dissertation presented to Yale University. Partially supported by the National Science Foundation under Grant GP-11021.

the expected values of ρ and λ with respect to R and L. When the notation $E(\rho)$ or $E(\lambda)$ is used, the distribution R or L will be understood. $E(\rho)$ and $E(\lambda)$ are then the probabilities of a success on the first pull on $\mathscr{R}$ and $\mathscr{L}$, respectively, conditional on I.

The expected number of successes over the n stages is to be maximized. This objective is a natural interpretation of a gambler's desire to make as much money as possible. It amounts to assuming that the utility of money (or, success) is linear for the gambler. An arm should not be selected to be pulled only because the expected probability of success on that arm is greater than it is on the other, since a future success is worth as much as an immediate success and the other arm may offer a reasonable chance of being better in the long run. An arm can be pulled at stage $k+1$ without waste and is called *optimal* if by pulling that arm the maximum attainable expected number of successes in the remaining $n-k$ pulls is attained conditional on I_k. Let $W^{\mathscr{R}}_{n-k}(I_k)$ and $W^{\mathscr{L}}_{n-k}(I_k)$ denote, respectively, the worth of the pattern I_k with $n-k$ pulls remaining when the right and left arm is pulled at stage $k+1$ and an optimal procedure followed thereafter. Then by definition $W_0{}^{\mathscr{R}}(I_n) = W_0{}^{\mathscr{L}}(I_n) = 0$ for all patterns I_n. Let $W_n(I) = \max\{W_n{}^{\mathscr{R}}(I), W_n{}^{\mathscr{L}}(I)\}$ for all n and I; $W_n(I)$ is then the worth of I when an optimal procedure is followed. The contribution to $W_n{}^{\mathscr{R}}(I)$ of the initial pull is $E(\rho)$. If a success obtains on pulling $\mathscr{R}$ then $I_1 = (\sigma R, L)$ and if a failure obtains then $I_1 = (\varphi R, L)$. Since the former has probability $E(\rho)$ and the latter probability $\bar{E}(\rho) = 1 - E(\rho)$,

$$W_n{}^{\mathscr{R}}(I) = E(\rho) + E(\rho)W_{n-1}(\sigma R, L) + \bar{E}(\rho)W_{n-1}(\varphi R, L)\,, \tag{1.1}$$

for all $n \geqq 1$ and all $I = (R, L)$. Symmetrically,

$$W_n{}^{\mathscr{L}}(I) = E(\lambda) + E(\lambda)W_{n-1}(R, \sigma L) + \bar{E}(\lambda)W_{n-1}(R, \varphi L)\,. \tag{1.2}$$

The function

$$\Delta_n(I) = W_n{}^{\mathscr{R}}(I) - W_n{}^{\mathscr{L}}(I)\,, \tag{1.3}$$

is, therefore, the expected advantage of choosing $\mathscr{R}$ over $\mathscr{L}$ at the first stage. $\mathscr{R}$ is optimal if $\Delta_n(I) \geqq 0$ and $\mathscr{L}$ is optimal if $\Delta_n(I) \leqq 0$.

The problem treated in this paper is a *two-armed bandit*, a typical problem in dynamic programming. For a particular value of n and a particular initial pattern of information I, a program can be devised to find $W_n{}^{\mathscr{R}}(I)$ and $W_n{}^{\mathscr{L}}(I)$ using (1.1) and (1.2) recursively. Such a program is described and carried out for the example $n = 12$ and R and L both uniform distributions in Berry (1971). There are two main drawbacks to such an approach. Firstly, a large memory is required; for fixed n the problem is 4-dimensional (though my program requires on the order of $n^3/6$ storage locations). Secondly, the values of $W_n{}^{\mathscr{R}}(I)$ and $W_n{}^{\mathscr{L}}(I)$ are found only for one particular pattern I. (Actually, as Bradt, Johnson, and Karlin (1956) point out, these values depend on I only through the first n moments of R and L.)

Two-armed bandit problems more general than the one treated here differ in various ways: $\mathscr{R}$ and $\mathscr{L}$ may not be independent; the objective function can be different, as when the utility of a success depends on the stage at which it occurs (discounting the future, for example); or, the number of pulls can be infinite, in which case the problem would be uninteresting unless the objective specified is nonetheless finite.

The problem described here is set in discrete time. Chernoff (1968) considers a continuous version where $\mathscr{R}$ and $\mathscr{L}$ are time-continuous processes in (particular, independent Wiener processes with unknown means and known variances). $\mathscr{R}$ or $\mathscr{L}$ is observed, payoff accumulates equal to the values of the process, and information about that process accumulates continuously until a switch is made and the other process is observed. Observation continues until some fixed time has elapsed. A less than helpful characteristic of every optimal selection procedure in this version is that almost every switch is accompanied by an uncountable number of switches within every time interval of positive duration which includes the switch.

Quisel (1965) touches on still another variant in which there is a time delay between a pull and getting information from the pull.

A problem related to (but different from) the two-armed bandit treated here is the two-armed bandit with finite memory. See Yakowitz (1969) for a description of this problem and for additional references.

For applications of two-armed bandit problems see Bradt *et al.* (1956), Quisel (1965), and Dubins and Savage (1965), Chapter 12.

Fabius and van Zwet (1970) deserves special reference here. They characterize Bayes strategies and admissible strategies in the general case of dependent arms. Their approach delivers many instances of some results of this paper (notably, Theorems 6.4 and 6.5), but most of the present results are true only for independent arms (notably, the stay-on-a-winner rule—Theorem 6.2). Their approach, unlike the current one, uses no calculus and therefore their results are only for integer numbers of successes and failures. The current approach is not applicable for dependent arms, except where a devious argument can be used to relate a problem with dependent arms to one with independent arms. Such an argument is used in Section 8 to show that Theorem 8.3 is equivalent to the result obtained by Feldman (1962), and later generalized by Fabius and van Zwet (1970), for a problem with a special kind of dependency between the arms.

2. The initial distributions. In this section the distributions R and L will be written in a more convenient form and particular patterns of information will be considered.

Consider an arm, say the right arm for definiteness. It will prove useful to regard R as having arisen from a measure $\mu_{\mathscr{R}}$ and a number $N_{\mathscr{R}}$ of pulls on the right arm that yielded, say, r successes and $r' = N_{\mathscr{R}} - r$ failures. The numbers

r and r' are allowed to be real and not just positive integers. Since the pulls are exchangeable, only the numbers of successes and failures affect $\mu_{\mathscr{R}}$ and R can be written $\sigma^r\varphi^{r'}\mu_{\mathscr{R}}$, regardless of the order of the r successes and r' failures. If information about the right arm is regarded as having arisen in this manner, then according to Bayes' theorem,

$$dR(\rho) = \nu^{-1}(r, r'; \mu_{\mathscr{R}})\rho^r(1 - \rho)^{r'}\, d\mu_{\mathscr{R}}(\rho)\,, \tag{2.1}$$

where

$$\nu(r, r'; \mu_{\mathscr{R}}) = \textstyle\int_0^1 \rho^r(1 - \rho)^{r'}\, d\mu_{\mathscr{R}}(\rho)\,. \tag{2.2}$$

The distribution R can always be written in the form (2.1), one $(r, r'; \mu_{\mathscr{R}})$ that qualifies is $(0, 0; R)$; obviously $\nu(0, 0; R) = 1$. $\mu_{\mathscr{R}}$ can be any positive measure and r and r' any real numbers for which $\nu(r, r'; \mu_{\mathscr{R}})$ is finite. The set of (r, r') which satisfy this condition will be called the *possibility region* for $\mu_{\mathscr{R}}$. If (r, r') is in the possibility region for $\mu_{\mathscr{R}}$, then any point $(r + a, r' + b)$ is also in the possibility region for $\mu_{\mathscr{R}}$ for nonnegative a and b, provided $\mu_{\mathscr{R}}$ assigns positive measure to the interior of the unit interval. Therefore, the possibility region for any measure which assigns positive measure to the interior of the unit interval is a quadrant of the (r, r') plane (which may be a half plane or the whole plane) defined by $(r_* + a, r_*' + b)$ for some (r_*, r_*') and all positive a and b. This quadrant may be open or closed depending on $\mu_{\mathscr{R}}$; either half-line, $r = r_*$ for $r' > r_*'$ or $r' = r_*'$ for $r > r_*$, may be included, if the point (r_*, r_*') is included then both of these half-lines are included. Similarly for the left arm

$$dL(\lambda) = \nu^{-1}(l, l'; \mu_{\mathscr{L}})\lambda^l(1 - \lambda)^{l'}\, d\mu_{\mathscr{L}}(\lambda)\,. \tag{2.3}$$

Points in the interior of the possibility region for $\mu_{\mathscr{R}}$ will play a special role in Section 4. Such points (r, r') are characterized by

$$\nu(r + \delta r, r' + \delta r'; \mu_{\mathscr{R}}) < \infty \qquad \text{for some} \quad \delta r, \delta r' < 0\,. \tag{2.4}$$

Since the distribution R is determined by r, r' and $\mu_{\mathscr{R}}$ and the distribution L is determined by l, l', and $\mu_{\mathscr{L}}$, the initial pattern of information I will sometimes be written $(r, r', \mu_{\mathscr{R}}; L)$, or sometimes, $(r, r', \mu_{\mathscr{R}}; l, l'\, \mu_{\mathscr{L}})$. A success on $\mathscr{R}$ then yields the pattern $I_1 = (\sigma R, L) = (r + 1, r', \mu_{\mathscr{R}}; L)$ according to Bayes' theorem. The expected values of ρ or λ with respect to I will sometimes be written $E(\rho|r, r'; \mu_{\mathscr{R}})$ and $E(\lambda|l, l'; \mu_{\mathscr{L}})$. In this notation, for example, $E(\rho|\sigma R) = E(\rho \mid r + 1, r'; \mu_{\mathscr{R}})$.

Arbitrary patterns of information will be investigated. However, patterns for which there exist r, r' l, and l' for some $\mu_{\mathscr{R}} = \mu_{\mathscr{L}}$ are of particular interest. Two special cases of this type of pattern will be introduced here and considered again in Section 4 and again in Section 8.

In the first special case there exist positive r, r', l, $l' < \infty$ for which $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \beta$, where

$$d\beta(x) = x^{-1}(1 - x)^{-1}\, dx\,. \tag{2.5}$$

If $\mu_{\mathscr{R}} = \beta$, then R and L are beta distributions, $r_* = r_*' = 0$, and the possibility region for β does not include either of the axes $r = 0$ or $r' = 0$. The conjugate nature of the beta family of distributions is well known (Raiffa and Schlaifer (1961))—if R is beta distribution then so are σR, φR, $\sigma\varphi R$, etc. The expected value of ρ is particularly simple for this case:

$$E(\rho \mid r, r'; \beta) = \frac{r}{r + r'} = \frac{r}{N_{\mathscr{R}}} . \tag{2.6}$$

In the second special case $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \tau$ is a two-point measure, concentrating probability $\frac{1}{2}$ on each of τ_1 and τ_2, $\tau_1 < \tau_2$, with not both $\tau_1 = 0$ and $\tau_2 = 1$. (For convenience, it is assumed that τ cannot be a one-point measure—in which case $\tau_1 = \tau_2$. Many results concerning τ apply as well to one-point measures, but none are of interest if $\mu_{\mathscr{R}} = \mu_{\mathscr{L}}$.) If $\tau_1 > 0$ and $\tau_2 < 1$ then $(r_*, r_*') = (-\infty, -\infty)$ and all points in the (r, r') plane are in the possibility region for τ. If $\tau_1 = 0$ and $\tau_2 < 1$ then $(r_*, r_*') = (0, -\infty)$ and all points for which $r \geqq r_*$ are possible. If $\tau_1 > 0$ and $\tau_2 = 1$ then $(r_*, r_*') = (-\infty, 0)$ and all points for which $r' \geqq r_*'$ are possible. (If the pair $\tau_1 = 0$ and $\tau_2 = 1$ were allowed, then r_* and r_*' would both be zero and the possibility region for τ would consist of only the nonnegative axes.) The expected value of ρ is

$$E(\rho \mid r, r'; \tau) = \frac{\tau_1^{r+1}(1 - \tau_1)^{r'} + \tau_2^{r+1}(1 - \tau_2)^{r'}}{\tau_1^{r}(1 - \tau_1)^{r'} + \tau_2^{r}(1 - \tau_2)^{r'}} . \tag{2.7}$$

If $\mu_{\mathscr{R}} = \beta$ then (2.4) holds for all points in the possibility region for $\mu_{\mathscr{R}}$. If $\mu_{\mathscr{R}} = \tau$ then (2.4) holds for all points in the possibility region for $\mu_{\mathscr{R}}$ provided $\tau_1 > 0$ and $\tau_2 < 1$.

3. The function $\Delta_n(I)$. In this section the function $\Delta_n(I)$ will be defined recursively and a few simple results given. From the definition of $\Delta_n(I)$ for all nonnegative n and for any $I = (R, L)$,

$$W_n^{\mathscr{R}}(I) = W_n(I) + \Delta_n^-(I) , \tag{3.1}$$

$$W_n^{\mathscr{L}}(I) = W_n(I) - \Delta_n^+(I) , \tag{3.2}$$

where, in a slight departure from normal usage, $\Delta_n^-(I) = \min\{0, \Delta_n(I)\}$ and $\Delta_n^+(I) = \max\{0, \Delta_n(I)\}$. In view of (3.2), for $n \geqq 1$ (1.1) becomes

$$\begin{aligned} W_n^{\mathscr{R}}(I) = E(\rho) &+ E(\rho)[W_{n-1}^{\mathscr{L}}(\sigma R, L) + \Delta_{n-1}^+(\sigma R, L)] \\ &+ \bar{E}(\rho)[W_{n-1}^{\mathscr{L}}(\varphi R, L) + \Delta_{n-1}^+(\varphi R, L)] , \end{aligned} \tag{3.3}$$

and in view of (3.1), (1.2) becomes

$$\begin{aligned} W_n^{\mathscr{L}}(I) = E(\lambda) &+ E(\lambda)[W_{n-1}^{\mathscr{R}}(R, \sigma L) - \Delta_{n-1}^-(R, \sigma L)] \\ &+ \bar{E}(\lambda)[W_{n-1}^{\mathscr{R}}(R, \varphi L) - \Delta_{n-1}^-(R, \varphi L)] . \end{aligned} \tag{3.4}$$

For $n \geqq 2$, the terms

$$E(\rho) + E(\rho) W_{n-1}^{\mathscr{L}}(\sigma R, L) + \bar{E}(\rho) W_{n-1}^{\mathscr{L}}(\varphi R, L) , \tag{3.5}$$

in (3.3) amount to the worth of the procedure: Pull $\mathscr{R}$ first and $\mathscr{L}$ second, and use an optimal procedure thereafter. (Of course there can be no second pull if $n < 2$.) Likewise,

$$E(\lambda) + E(\lambda)W^{\mathscr{R}}_{n-1}(R, \sigma L) + \bar{E}(\lambda)W^{\mathscr{R}}_{n-1}(R, \varphi L) \tag{3.6}$$

is the worth of the procedure: Pull $\mathscr{L}$ first and $\mathscr{R}$ second, and use an optimal procedure thereafter. Since the pulls are exchangeable, according to this interpretation the expressions (3.5) and (3.6) are equal for $n \geqq 2$. Therefore, subtracting (3.4) from (3.3) yields for $n \geqq 2$,

$$\begin{aligned}\Delta_n(I) = {} & E(\rho)\Delta^+_{n-1}(\sigma R, L) + \bar{E}(\rho)\Delta^+_{n-1}(\varphi R, L) \\ & + E(\lambda)\Delta^-_{n-1}(R, \sigma L) + \bar{E}(\lambda)\Delta^-_{n-1}(R, \varphi L)\,.\end{aligned} \tag{3.7}$$

(Compare (2.11) in Fabius and van Zwet (1970).) This is a promising expression for $\Delta_n(I)$ since, together with the evident initial condition

$$\Delta_1(I) = E(\rho) - E(\lambda)\,, \tag{3.8}$$

(3.7) defines $\Delta_n(I)$ recursively.

It seems reasonable to expect that the vanishing of particular terms in (3.7) implies the vanishing of other terms. In fact, it will be seen in Section 6 that $\Delta_{n-1}(\sigma R, L) \geqq \Delta_{n-1}(\varphi R, L)$, and symmetrically, $\Delta_{n-1}(R, \sigma L) \leqq \Delta_{n-1}(R, \varphi L)$. Several facts about the function Δ_n are easy to verify. Three intuitive theorems will now be given without proof.

Since Δ_n is the expected advantage of choosing $\mathscr{R}$ over $\mathscr{L}$, clearly, $-1 \leqq \Delta_n \leqq 1$; in fact, more can be said.

THEOREM 3.1. *For any pattern of information $I = (R, L)$ and all n, $-\bar{E}(\rho) \leqq \Delta_n(I) \leqq \bar{E}(\lambda)$.*

If either of the distributions R or L is of a particular type, $\Delta_n(I)$ may be easy to calculate for all n. For example, if an arm yields success with probability one, then it should be pulled, and the expected loss due to pulling the other arm is the difference between 1 and the worth of that arm.

THEOREM 3.2. *If R concentrates probability one at $\rho = 1$, then for all n and L, $\Delta_n(I) = \bar{E}(\lambda)$, which is nonnegative.*

In general, the worth of pulling a particular arm consists in the net worth with respect to expected immediate payoff and with respect to worth of information, so that Δ_n is seldom given by the difference in expected immediate payoff,

$$E(\rho) - E(\lambda)\,. \tag{3.9}$$

Of course, as seen in (3.8), (3.9) gives this difference when only one pull remains, for then any information gained on the pull will not be used and therefore has no value. For $n \geqq 2$, Δ_n is given by (3.9) when and only when the worth of information is the same for both arms. This can happen when, for example, (a) both arms are the same a priori, (b) pulling neither arm has information

value, or (c) pulling either arm once will give complete information. The next theorem treats these three special patterns of information.

THEOREM 3.3. *If I is such that either*

(a) *$R = L$; that is, the arms are identical initially,*

(b) *R and L concentrate probability one on $E(\rho)$ and $E(\lambda)$; that is, the probability of success is known for both arms,*

(c) *R and L are two-point distributions, concentrating all the probability at 0 and 1, so that $\rho = 1$ and $\lambda = 1$ with probabilities $E(\rho)$ and $E(\lambda)$; that is, each arm will yield either all successes or all failures and one pull on either arm determines the quality of that arm,* *then for $n \geqq 1$, $\Delta_n(I) = E(\rho) - E(\lambda)$.*

Theorem 3.3 can be proved algebraically using (3.7), but each result holds for an intuitive reason that can be made rigorous. The conclusion is obvious in case (a) since $E(\rho) = E(\lambda)$ and $W_n^{\mathscr{R}} = W_n^{\mathscr{L}}$ when $R = L$. In case (b) the quality of both arms is known, and any pull on the inferior arm costs the difference in the quality of the arms. In case (c), the better arm to pull (if indeed one is better than the other) becomes known immediately after the first pull, whichever arm is pulled first (and will yield either all successes, or all failures if both ρ and λ are 0), therefore the difference between pulling the right and left arm is simply the difference in the expected immediate payoffs.

4. Fundamental inequalities. Thus far the possibility that r, r', l, and l' are real numbers and not necessarily integers has not been exploited. This section exploits, and it largely based on, the continuity of $\Delta_n(I)$ in (r, r') in the interior of the possibility region for $\mu_{\mathscr{R}}$.

Inequalities in $I = (r, r', \mu_{\mathscr{R}}; L)$ for the function $\Delta_n(I)$ are derived in this section when (r, r') is an interior point of the possibility region for $\mu_{\mathscr{R}}$. These inequalities will be strengthened in Section 5 and extended to all points in the possibility region. This separate treatment eliminates the need for considering in this section distributions which would unnecessarily complicate the presentation of the basic theory. Results are stated and derived in terms of the right arm; symmetric results hold for the left arm as applications of those for the right arm, with names reversed. For the purposes of this section, write $I = (R, L)$ as $(r, r', \mu_{\mathscr{R}}; L)$, where R is given by (2.1) and is a probability distribution for (r, r') in the possibility region for $\mu_{\mathscr{R}}$.

As will be seen, some important properties of $E(\rho \mid r, r'; \mu_{\mathscr{R}})$ are passed on to $\Delta_n(r, r', \mu_{\mathscr{R}}; L)$ for all n. This motivates studying the behavior of the function $\nu(r, r'; \mu_{\mathscr{R}})$, which is defined in (2.2) and which will sometimes be abbreviated to $\nu(r, r')$.

LEMMA 4.1. *For (r, r') in the interior of the possibility region for $\mu_{\mathscr{R}}$, there exist negative δr and $\delta r'$ such that*

$$(4.1)\quad \nu(r + \delta r, r' + \delta r') = \sum_{0 \leq s,t} \frac{1}{s!\, t!} \left[\int_0^1 (\log \rho)^s (\log (1 - \rho))^t \rho^r (1 - \rho)^{r'} \, d\mu_{\mathscr{R}}(\rho)\right] (\delta r)^s (\delta r')^t ,$$

where the series is absolutely convergent. Also for all $s, t \geqq 0$,

$$\frac{\partial^{s+t}}{\partial r^s \partial r'^t} \nu(r, r') = \int_0^1 (\log \rho)^s (\log (1 - \rho))^t \rho^r (1 - \rho)^{r'} \, d\mu_{\mathscr{R}}(\rho) . \tag{4.2}$$

REMARK. For many readers the asserted analyticity of ν in the pair (r, r') will be familiar, but it seems easier to give a demonstration than to provide an exactly appropriate reference.

PROOF OF LEMMA 4.1. For all $\delta r, \delta r'$,

$$\begin{aligned} \rho^{r+\delta r}(1 - \rho)^{r'+\delta r'} &= \rho^r (1 - \rho)^{r'} \exp(\delta r \log \rho) \exp(\delta r' \log (1 - \rho)) \\ &= \rho^r (1 - \rho)^{r'} \sum_{0 \leqq s,t} \frac{1}{s!\, t!} (\log \rho)^s (\log (1 - \rho))^t (\delta r)^s (\delta r')^t , \end{aligned} \tag{4.3}$$

is absolutely convergent and the partial sums of the series in (4.3) are majorized in absolute value by $\rho^{r-|\delta r|}(1 - \rho)^{r'-|\delta r'|}$. Since $\nu(r - \varepsilon, r' - \varepsilon) < \infty$ for sufficiently small ε, the Lebesgue dominated convergence theorem applies to (4.3) to prove (4.1).

Repeated differentiation of the convergent power series (4.1) yields (4.2). □

LEMMA 4.2. *For all* n *and* $I = (r, r', \mu_{\mathscr{R}}; L)$, $\Delta_n(I)$ *is continuous in* (r, r') *in the interior of the possibility region for* $\mu_{\mathscr{R}}$.

PROOF. According to (3.8) and the definition of ν, $\Delta_1(I) = \nu(r + 1, r')/\nu(r, r) - E(\lambda)$, which is continuous in (r, r') in view of Lemma 4.1 for (r, r') in the interior of the possibility region for $\mu_{\mathscr{R}}$. The lemma follows from (3.7) by induction. □

Though continuous, $\Delta_n(I)$ is not necessarily everywhere differentiable. While $\Delta_n(I)$ is regular in (r, r') except along certain curves, the regularity of $\Delta_n(I)$ will not here be analyzed beyond the extent essential for later demonstrations. The required regularity is provided by Lemma 4.4. For the proof of the next result, see any advanced calculus text, for example, (Widder (1961), Theorem 9, page 40).

LEMMA 4.3. *If both partial derivatives of a function* $g(x, x')$ *exist and are continuous at a point then at that point the directional derivative of* g *along a vector* (a, b) *is given by*

$$D_{(a,b)} g(x, x') = \left(a \frac{\partial}{\partial x} + b \frac{\partial}{\partial x'}\right) g(x, x') . \tag{4.4}$$

LEMMA 4.4. *For all* n *and* $I = (r, r', \mu_{\mathscr{R}}; L)$, $D_{(a,b)}\Delta_n(I)$ *exists along every vector* (a, b) *at every point* (r, r') *in the interior of the possibility region for* $\mu_{\mathscr{R}}$.

PROOF. From (4.2),

$$\nu(r + \delta r, r' + \delta r') = \nu(r, r') + \frac{\partial}{\partial r} \nu(r, r') \delta r + \frac{\partial}{\partial r'} \nu(r, r') \delta r' + o(|\delta r| + |\delta r'|) .$$

In view of (3.8),

$$D_{(a,b)}\Delta_1(I) = D_{(a,b)}E(\rho \mid r, r'; \mu_{\mathscr{R}}) = D_{(a,b)}\frac{\nu(r+1, r')}{\nu(r, r')}$$
$$= \left(a\frac{\partial}{\partial r} + b\frac{\partial}{\partial r'}\right)\frac{\nu(r+1, r')}{\nu(r, r')},$$

according to Lemmas 4.1 and 4.3, which proves the lemma for $n = 1$. The lemma follows from (3.7) by induction. □

(It is clear from the proof of Lemma 4.4 that the directional derivative of $\Delta_n(I)$ can be expressed linearly in terms of its partial derivatives except possibly at points where $\Delta_{n-k}(I_k) = 0$ for some pattern of information I_k which can occur after k pulls when starting from I.)

Lemmas 4.2 and 4.4 will be used to prove Theorem 4.1, which will then be extended by Theorems 5.1 and 5.2. It seems reasonable to expect that when r is increased, the advantage of pulling $\mathscr{R}$ over $\mathscr{L}$ does not decrease, for then $\mathscr{R}$ promises to be at least as successful as before. Correspondingly, if r' is increased, the advantage of pulling $\mathscr{R}$ over $\mathscr{L}$ ought not increase, for then $\mathscr{R}$ promises to be no more successful than before. The next theorem says this and more when (r, r') is an interior point of the possibility region for $\mu_{\mathscr{R}}$: if r and r' increase simultaneously, the advantage of pulling $\mathscr{R}$ over $\mathscr{L}$ does not decrease if the rate of change of r compared with the rate of change of r' is larger than a particular bound and does not increase if this ratio is smaller than another (obviously, smaller) bound. These bounds are implied by the following statements, which put propositions $J(n)$ and $K(n)$ of the theorem into words. If the probability of a success on $\mathscr{R}$ conditional on having already observed $n - 1$ successes in $n - 1$ pulls on $\mathscr{R}$ (this probability is given by the expected value of ρ with respect to $\sigma^{n-1}R$) does not decrease for a particular direction from a particular point in the (r, r') plane, then Δ_n at that point does not decrease for the same direction. If the probability of a failure on $\mathscr{R}$ conditionally on having already observed $n - 1$ failures in $n - 1$ pulls on $\mathscr{R}$ (given by the expected value of $1 - \rho$ with respect to $\varphi^{n-1}R$) does not increase for a particular direction from a particular point in the (r, r') plane, then Δ_n at that point does not increase for the same direction.

THEOREM 4.1. *Provided* (2.4) *holds, the following statements are true for* $n \geqq 1$, *for* $I = (r, r', \mu_{\mathscr{R}}; L)$, *and for* a *and* b *nonnegative and not both* 0*:*

$$J(n)\colon D_{(a,b)}\Delta_n(I) \geqq 0 \quad \text{if} \quad D_{(a,b)}E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}}) \geqq 0,$$
$$K(n)\colon D_{(a,b)}\Delta_n(I) \leqq 0 \quad \text{if} \quad D_{(a,b)}E(\rho \mid r, r' + n - 1; \mu_{\mathscr{R}}) \leqq 0.$$

For the proof of Theorem 4.1, which will be presented gradually, the behavior of $E(\rho \mid r, r'; \mu_{\mathscr{R}})$ in (r, r') will be needed. Though for $n \geqq 2$ the partial derivatives of $\Delta_n(r, r', \mu_{\mathscr{R}}; L)$ with respect to r and r' do not always exist, the partial derivatives of $E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}})$ and $E(\rho \mid r, r' + n - 1; \mu_{\mathscr{R}})$ do exist

and are continuous. In view of this fact and Lemma 4.3, for $n \geqq 1$ the latter directional derivative in $J(n)$ and $K(n)$ of Theorem 4.1 can be written as in (4.4), making these hypotheses easier to manipulate: for $\mu_{\mathscr{R}}$ not a one-point measure,

$$J(n)\colon D_{(a,b)}\Delta_n(I) \geqq 0 \qquad \text{if} \quad \frac{a}{b} \geqq A(r+n-1, r'; \mu_{\mathscr{R}}),$$

$$K(n)\colon D_{(a,b)}\Delta_n(I) \leqq 0 \qquad \text{if} \quad \frac{a}{b} \leqq A(r, r'+n-1; \mu_{\mathscr{R}}),$$

where

$$A(r, r'; \mu_{\mathscr{R}}) = \frac{-\frac{\partial}{\partial r'} E(\rho \mid r, r'; \mu_{\mathscr{R}})}{\frac{\partial}{\partial r} E(\rho \mid r, r'; \mu_{\mathscr{R}})}. \tag{4.5}$$

The function A would not be defined for one-point measures $\mu_{\mathscr{R}}$. Where there can be no confusion, $A(r, r'; \mu_{\mathscr{R}})$ will be abbreviated to $A(r, r')$.

Because they are simple, and therefore potentially helpful for following later arguments, the versions of $J(n)$ and $K(n)$ for two special cases considered earlier will now be given. First, where $\beta(\rho)$ is defined by (2.5),

$$A(r, r'; \beta) = \frac{r}{r'} \qquad \text{for} \quad r, r' > 0. \tag{4.6}$$

If $I = (r, r', \beta; L)$, $J(n)$ and $K(n)$ of Theorem 4.1 become

$$J_\beta(n)\colon D_{(a,b)}\Delta_n(I) \geqq 0 \qquad \text{if} \quad \frac{a}{b} \geqq \frac{r+n-1}{r'},$$

$$K_\beta(n)\colon D_{(a,b)}\Delta_n(I) \geqq 0 \qquad \text{if} \quad \frac{a}{b} \leqq \frac{r}{r'+n-1}.$$

Second, where R is an interior two-point distribution (so that $\mu_{\mathscr{R}} = \tau$) $J(n)$ and $K(n)$ together completely determine the gradient of $\Delta_n(r, r', \tau; L)$ in (r, r'). If τ concentrates mass $\frac{1}{2}$ on τ_1 and $\frac{1}{2}$ on τ_2, $0 < \tau_1 < \tau_2 < 1$, then,

$$A(r, r'; \tau) = A(\tau) = -\log[(1-\tau_1)/(1-\tau_2)]/\log[\tau_1/\tau_2], \tag{4.7}$$

a constant in (r, r'). If $I = (r, r', \tau; L)$, $J(n)$ and $K(n)$ of Theorem 4.1 become

$$J_\tau(n)\colon D_{(a,b)}\Delta_n(I) \geqq 0 \qquad \text{if} \quad \frac{a}{b} \geqq A(\tau),$$

$$K_\tau(n)\colon D_{(a,b)}\Delta_n(I) \leqq 0 \qquad \text{if} \quad \frac{a}{b} \leqq A(\tau).$$

The proof of Theorem 4.1 will depend on the behavior of $A(r, r'; \mu_{\mathscr{R}})$. Using the notation: $\operatorname{Cov}(U, V) = E(UV) - E(U)E(V)$, for real U and V on $[0, 1]$, where unconditional expectation is as usual with respect to $I = (R, L)$,

$$\frac{\partial}{\partial r} E(\rho \mid r, r'; \mu_{\mathscr{R}}) = \operatorname{Cov}(\rho, \log \rho), \quad \frac{\partial}{\partial r'} E(\rho \mid r, r'; \mu_{\mathscr{R}}) = \operatorname{Cov}(\rho, \log(1-\rho)),$$

in view of (4.2). Therefore,

$$A(r, r'; \mu_{\mathscr{R}}) = \frac{-\operatorname{Cov}(\rho, \log(1-\rho))}{\operatorname{Cov}(\rho, \log\rho)} = \frac{\operatorname{Cov}(1-\rho, \log(1-\rho))}{\operatorname{Cov}(\rho, \log\rho)} . \tag{4.8}$$

The next lemma follows from a well-known principle: If R is not a one-point distribution, the covariance with respect to R of a strictly increasing function is positive, (Lehmann (1966)).

LEMMA 4.5. *Provided $\mu_{\mathscr{R}}$ is not a one-point measure, $A(r, r'; \mu_{\mathscr{R}})$ is positive and finite. In fact, both numerator and denominator of* (4.8) *are positive and finite.*

LEMMA 4.6. *Provided $\mu_{\mathscr{R}}$ is not a one-point measure,*

$$\frac{\partial}{\partial r} A(r, r'; \mu_{\mathscr{R}}) \geqq 0 , \qquad \frac{\partial}{\partial r'} A(r, r'; \mu_{\mathscr{R}}) \leqq 0 , \tag{4.9}$$

with equality if and only if $\mu_{\mathscr{R}}$ is a two-point measure.

PROOF. The first inequality in (4.9) holds whenever

$$\begin{aligned} &\operatorname{Cov}(\rho, \log(1-\rho)) \frac{\partial}{\partial r} \operatorname{Cov}(\rho, \log\rho) \\ &\qquad - \operatorname{Cov}(\rho, \log\rho) \frac{\partial}{\partial r} \operatorname{Cov}(\rho, \log(1-\rho)) \geqq 0 , \end{aligned} \tag{4.10}$$

unless $\operatorname{Cov}(\rho, \log\rho) = 0$, which is excluded since R is not a one-point distribution. After the indicated differentiation, inequality (4.10) becomes:

$$\begin{aligned} &\operatorname{Cov}(\rho, \log(1-\rho)) E([\rho - E(\rho)][\log\rho - E(\log\rho)]\log\rho) \\ &\qquad - \operatorname{Cov}(\rho, \log\rho) E([\rho - E(\rho)][\log(1-\rho) - E(\log(1-\rho))]\log\rho) \\ &= E(H(\rho)[\rho - E(\rho)]\log\rho) \geqq 0 , \end{aligned} \tag{4.11}$$

where

$$\begin{aligned} H(\rho) = {} & [\log\rho - E(\log\rho)] \operatorname{Cov}(\rho, \log(1-\rho)) \\ & - [\log(1-\rho) - E(\log(1-\rho))] \operatorname{Cov}(\rho, \log\rho) . \end{aligned}$$

Since $E(H(\rho)[\rho - E(\rho)]) = 0$ according to the definition of H, for any constant C, (4.11) can be written

$$E(H(\rho)[\rho - E(\rho)][\log\rho - C]) \geqq 0 . \tag{4.12}$$

H is strictly convex since it is the sum of two strictly convex functions, and the expected value of H is zero; therefore, since R is not a one-point distribution, H has exactly two zeros in (0, 1), call them ρ_1 and ρ_2 with $\rho_1 < \rho_2$: $H(\rho_1) = H(\rho_2) = 0$. According to Jensen's inequality (Hardy, *et al.* (1934) Chapter III),

$$\rho_2 > E(\rho) > \rho_1 . \tag{4.13}$$

The number

$$\begin{aligned}\operatorname{Cov}(H(\rho), \log \rho) &= E(H(\rho)[\log \rho - C]) \\ &= \operatorname{Cov}(\log \rho, \log \rho) \operatorname{Cov}(\rho, \log(1-\rho)) \\ &\quad - \operatorname{Cov}(\log \rho, \log(1-\rho)) \operatorname{Cov}(\rho, \log \rho),\end{aligned}$$

may be of either sign; correspondingly, two cases will be considered.

Case 1. $\operatorname{Cov}(H(\rho), \log \rho) \geqq 0$. In (4.12), let $C = \log \rho_1$, then

$$\begin{aligned}&E(H(\rho)[\rho - E(\rho)][\log \rho - \log \rho_1]) \\ &= E(H(\rho)[\rho - \rho_2][\log \rho - \log \rho_1]) \\ &\quad + [\rho_2 - E(\rho)]E(H(\rho)[\log \rho - \log \rho_1]).\end{aligned} \tag{4.14}$$

The second term of the right side of (4.14) is nonnegative for this case in view of (4.13). The first term is nonnegative since according to the following argument, $H(\rho)[\rho - \rho_2][\log \rho - \rho_1] \geqq 0$ for all ρ: for $\rho \leqq \rho_1$, $H(\rho) \geqq 0$, $\rho - \rho_2 < 0$, and $\log \rho - \log \rho_1 \leqq 0$; for $\rho_1 < \rho < \rho_2$, $H(\rho) \leqq 0$, $\rho - \rho_2 \leqq 0$, and $\log \rho - \log \rho_1 \geqq 0$; and for $\rho_2 < \rho$, $H(\rho) \geqq 0$, $\rho - \rho_2 \geqq 0$, and $\log \rho - \log \rho_1 \geqq 0$.

Case 2. $\operatorname{Cov}(H(\rho), \log \rho) < 0$. In (4.12), let $C = \log \rho_2$, then

$$\begin{aligned}&E(H(\rho)[\rho - E(\rho)][\log \rho - \log \rho_2]) \\ &= E(H(\rho)[\rho - \rho_1][\log \rho - \log \rho_2]) \\ &\quad + [\rho_1 - E(\rho)]E(H(\rho)[\log \rho - \log \rho_2]),\end{aligned} \tag{4.15}$$

which is nonnegative according to an argument similar to the one in Case 1.

The symmetry of the form of A makes it clear that the second inequality in (4.9) is an instance of the first.

In view of (4.7), the inequalities in (4.9) are equalities if $\mu_{\mathscr{R}}$ is a two-point measure. If $\mu_{\mathscr{R}}$ is concentrated on more than two points then both terms on the right side of (4.14) and (4.15) are positive, so that the inequalities in (4.9) are strict. □

PROOF OF THEOREM 4.1. Since, by definition, $\Delta_1(I) = E(\rho) - E(\lambda)$,

$$D_{(a,b)}\Delta_1(r, r', \mu_{\mathscr{R}}; L) = D_{(a,b)}E(\rho \mid r, r'; \mu_{\mathscr{R}});$$

so that $J(1)$ and $K(1)$ both hold, and Theorem 4.1 determines the sign of the derivatives of $\Delta_1(I)$ for every direction in the (r, r') plane from points in the interior of the possibility region for $\mu_{\mathscr{R}}$.

If $\mu_{\mathscr{R}}$ is a one-point measure then the theorem is obvious since a one-point measure is not affected by changes in r or r'; $\mu_{\mathscr{R}}$ is assumed not to be a one-point measure for the remainder of the proof. This assumption will frequently be used implicitly since Lemmas 4.5 and 4.6 which depend on it will frequently be used.

The proof will be accomplished inductively by differentiating in (3.7). The signs of the last two terms of (3.7) do not materially affect the proof. $\Delta_{n-1}(\sigma R, L)$ can be $\geqq 0$ or < 0, as can $\Delta_{n-1}(\varphi R, L)$, so that there are four cases to be considered. However, one of these cases is vacuous, as will now be shown.

In view of Lemma 4.5, $a/b > A(r + n - 1, r')$ when $b = 0$ and $a/b < A(r, r' + n - 1)$ when $a = 0$, for all n. Therefore, for $n \geqq 2$, $J(n-1)$ implies that $\Delta_{n-1}(I)$ does not decrease as r increases and $K(n-1)$ implies that $\Delta_{n-1}(I)$ does not increase as r' increases. In view of these two facts (for $n \geqq 2$),

$$\Delta_{n-1}(r+1, r', \mu_{\mathscr{R}}; L) \geqq \Delta_{n-1}(r, r', \mu_{\mathscr{R}}; L) \geqq \Delta_{n-1}(r, r'+1, \mu_{\mathscr{R}}; L) . \tag{4.16}$$

This relationship implies:

$$\Delta_{n-1}(\sigma R, L) \geqq {}_{n-1}(\varphi R, L) . \tag{4.17}$$

Inequality (4.17) will be required in a critical point of the proof, for the present it serves to show that the following three cases are exhaustive:

$$\begin{aligned}
&\Delta_{n-1}(\sigma R, L) < 0 \quad (\text{and } \Delta_{n-1}(\varphi R, L) < 0) ; \\
&\Delta_{n-1}(\sigma R, L) \geqq 0 \quad \text{and } \Delta_{n-1}(\varphi R, L) < 0 ; \\
&\Delta_{n-1}(\varphi R, L) \geqq 0 \quad (\text{and } \Delta_{n-1}(\sigma R, L) \geqq 0) .
\end{aligned}$$

Differentiating both sides of (3.7),

$$\begin{aligned}
&D_{(a,b)}\Delta_n(I) \\
&\quad = E(\rho) D_{(a,b)}\Delta^+_{n-1}(r+1, r', \mu_{\mathscr{R}}; L) \\
&\quad\quad + \bar{E}(\rho) D_{(a,b)}\Delta^+_{n-1}(r, r'+1, \mu_{\mathscr{R}}; L) + [\Delta^+_{n-1}(r+1, r', \mu_{\mathscr{R}}; L) \\
&\quad\quad - \Delta^+_{n-1}(r, r'+1, \mu_{\mathscr{R}}; L)] D_{(a,b)}E(\rho \mid r, r'; \mu_{\mathscr{R}}) \\
&\quad\quad + E(\lambda) D_{(a,b)}\Delta^-_{n-1}(r, r', \mu_{\mathscr{R}}; \sigma L) + \bar{E}(\lambda) D_{(a,b)}\Delta^-_{n-1}(r, r', \mu_{\mathscr{R}}; \varphi L) .
\end{aligned} \tag{4.18}$$

To show $J(n)$ assume $a/b \geqq A(r + n - 1, r')$. Since $A(r + n - 1, r') \geqq A(r + (n-1) - 1, r')$ according to Lemma 4.6, $J(n-1)$ applies to show $D_{(a,b)}\Delta_{n-1}(r, r', \mu_{\mathscr{R}}; \sigma L) \geqq 0$ and $D_{(a,b)}\Delta_{n-1}(r, r', \mu_{\mathscr{R}}; \varphi L) \geqq 0$. Therefore, the last two terms in (4.18) are nonnegative since $D_{(a,b)}\Delta^-_{n-1}(I) \geqq 0$ whenever $D_{(a,b)}\Delta_{n-1}(I) \geqq 0$.

Since $A(r + n - 1, r') = A((r+1) + (n-1) - 1, r')$, $J(n-1)$ applies to show that $D_{(a,b)}\Delta_{n-1}(r+1, r', \mu_{\mathscr{R}}; L) \geqq 0$. If $\Delta_{n-1}(\sigma R, L) > 0$ then $\Delta^+_{n-1}(r+1, r', \mu_{\mathscr{R}}; L)$ in (4.18) can be replaced by $\Delta_{n-1}(r+1, r', \mu_{\mathscr{R}}; L)$; if $\Delta_{n-1}(\sigma R, L) < 0$ then it can be replaced by 0; and if $\Delta_{n-1}(\sigma R, L) = 0$ then it may be equal to either $\Delta_{n-1}(\sigma R, L)$ or 0, depending on (a, b). In each case the first term in (4.18) is nonnegative.

Since $A(r + n - 1, r') \geqq A(r + (n-1) - 1, r') \geqq A(r + (n-1) - 1, r' + 1)$ according to Lemma 4.6, $J(n-1)$ applies to show $D_{(a,b)}\Delta_{n-1}(r, r'+1, \mu_{\mathscr{R}}; L) \geqq 0$. If $\Delta_{n-1}(\varphi R, L) > 0$ then $\Delta^+_{n-1}(r, r'+1, \mu_{\mathscr{R}}; L)$ in (4.18) can be replaced by $\Delta_{n-1}(r, r'+1, \mu_{\mathscr{R}}; L)$; if $\Delta_{n-1}(\varphi R, L) < 0$ then it can be replaced by 0; and if $\Delta_{n-1}(\varphi R, L) = 0$ then it may be equal to either $\Delta_{n-1}(\varphi R, L)$ or 0. In each case the second term in (4.18) is nonnegative.

$J(n-1)$ and $K(n-1)$ apply to show that the third term is nonnegative. The factor in square brackets is nonnegative in view of (4.17) and $D_{(a,b)}E(\rho \mid r, r'; \mu_{\mathscr{R}})$ is nonnegative according to $J(1)$.

Therefore $J(n)$ of the theorem is proved. The proof of $K(n)$ is similar by assuming $a/b \leqq A(r, r' + n - 1)$, $K(1)$, $K(n-1)$ and $J(n-1)$, and arguing that each term in (4.18) is nonpositive. □

5. Fundamental inequalities; extensions. Theorem 4.1, which deals with the sign of the gradient of $\Delta_n(r, r', \mu_{\mathscr{R}}; L)$ along curves in the interior of the possibility region for $\mu_{\mathscr{R}}$, is extended by the three main results on the present section. Lemma 5.1 is a macroscopic version of Theorem 4.1. The contours of $E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}})$ and $E(\rho \mid r, r' + n - 1; \mu_{\mathscr{R}})$ are shown to be lines of nondecrease and nonincrease of $\Delta_n(r, r', \mu_{\mathscr{R}}; L)$ in any direction of nondecreasing r and r' in the interior of the possibility region for $\mu_{\mathscr{R}}$. Theorem 5.1 extends Lemma 5.1 to include the edges of the possibility region for $\mu_{\mathscr{R}}$ (the region may have no edges, one edge, or two edges). Finally, Theorem 5.2 shows that for fixed $\mu_{\mathscr{R}}$, L, and n, $\Delta_n(r, r', \mu_{\mathscr{R}}; L)$ is strictly increased or decreased if $E(\rho \mid r + n + 1, r'; \mu_{\mathscr{R}})$ or $E(\rho \mid r, r' + n - 1; \mu_{\mathscr{R}})$ is increased or decreased.

LEMMA 5.1. *Provided* (2.4) *holds, the following statements are true for* $n \geqq 1$, *for* $I = (r, r', \mu_{\mathscr{R}}; L)$, *and for all* $\delta r, \delta r' \geqq 0$:

$$\hat{J}(n): \Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) \geqq \Delta_n(r, r', \mu_{\mathscr{R}}; L)$$
$$\text{if} \quad E(\rho \mid r + \delta r + n - 1, r' + \delta r'; \mu_{\mathscr{R}}) \geqq E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}});$$
$$\hat{K}(n): \Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) \leqq \Delta_n(r, r', \mu_{\mathscr{R}}; L)$$
$$\text{if} \quad E(\rho \mid r + \delta r, r' + \delta r' + n - 1; \mu_{\mathscr{R}}) \leqq E(\rho \mid r, r' + n - 1; \mu_{\mathscr{R}}).$$

PROOF. The implicit function theorem (Widder (1961) Theorem 14, page 56) applies to show that on expressing the contours of $E(\rho \mid r, r'; \mu_{\mathscr{R}})$ as $(r(N_{\mathscr{R}}), r'(N_{\mathscr{R}}))$ in the parameter $N_{\mathscr{R}} = r + r'$, each contour extends uninterrupted for all $N_{\mathscr{R}}$; the slope of the contours of $E(\rho \mid r, r'; \mu_{\mathscr{R}})$ is $A(r, r'; \mu_{\mathscr{R}})$, which is defined by (4.7).

For $\delta r' \geqq 0$, consider two points $(r, r'; \mu_{\mathscr{R}})$ and $(r + h, r' + \delta r'; \mu_{\mathscr{R}})$ on a contour of $E(\rho \mid x + n - 1, x'; \mu_{\mathscr{R}})$; that is,

$$E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}}) = E(\rho \mid r + h + n - 1, r' + \delta r'; \mu_{\mathscr{R}}). \tag{5.1}$$

According to $J(n)$ of Theorem 4.1, $\Delta_n(x, x', \mu_{\mathscr{R}}; L)$ is nondecreasing along such a contour in the interior of the possibility region for $\mu_{\mathscr{R}}$ for any L, so that

$$\Delta_n(r + h, r' + \delta r', \mu_{\mathscr{R}}; L) \geqq \Delta_n(r, r', \mu_{\mathscr{R}}; L). \tag{5.2}$$

Consider a third point $(r + \delta r, r' + \delta r'; \mu_{\mathscr{R}})$ satisfying the condition in $\hat{J}(n)$, so that

$$E(\rho \mid r + \delta r + n - 1, r' + \delta r'; \mu_{\mathscr{R}}) \geqq E(\rho \mid r + h + n - 1, r' + \delta r'; \mu_{\mathscr{R}}) \tag{5.3}$$

according to (5.1). According to $J(n)$ of Theorem 4.1 for $b = 0$,

$$\Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) \geqq \Delta_n(r + h, r' + \delta r', \mu_{\mathscr{R}}; L). \tag{5.4}$$

$\hat{J}(n)$ of the lemma follows from (5.2) and (5.4).

$\hat{K}(n)$ of the lemma is proved in a similar fashion by considering points on a contour of $E(\rho \mid x, x' + n - 1; \mu_{\mathscr{R}})$ and applying $K(n)$ of Theorem 4.1. □

Lemma 5.1 does not apply for any distribution $R = (r, r'; \mu_{\mathscr{R}})$ which corresponds to a point on an edge of the possibility region for $\mu_{\mathscr{R}}$. For such a distribution, $\nu(r + \delta r, r' + \delta r'; \mu_{\mathscr{R}}) = \infty$ for $\delta r < 0$ if (r, r') is on the vertical edge $(r = r_*)$, and for $\delta r' < 0$ if (r, r') is on the horizontal edge $(r' = r_*')$. Lemma 5.1 will be extended to arbitrary distributions $R = (r, r'; \mu_{\mathscr{R}})$ by showing first that $\Delta_n(R, L)$ can be approximated arbitrarily closely by replacing $\mu_{\mathscr{R}}$ with a measure which satisfies (2.4).

LEMMA 5.2. *For each $I = (r_0, r_0', \mu_{\mathscr{R}}; L)$ for which (r_0, r_0') is in the possibility region of $\mu_{\mathscr{R}}$ and for which $\mu_{\mathscr{R}}$ is not confined to the two points $\{0, 1\}$, there exists a family of measures m_ε such that for all real r and r',*

$$\nu(r, r'; m_\varepsilon) < \infty \tag{5.5}$$

for each measure m_ε with $\varepsilon > 0$, and

$$\lim_{\varepsilon \downarrow 0} \Delta_n(r, r', m_\varepsilon; L) = \Delta_n(r, r', \mu_{\mathscr{R}}; L) \tag{5.6}$$

for $n \geqq 1$ and every r and r' for which $r \geqq r_0$ and $r' \geqq r_0'$.

REMARKS. The convergence in (5.6) is not necessarily uniform in (r, r') or in n.

Any measure which satisfies (5.5) also satisfies (2.4), so that for $\varepsilon > 0$ Lemma 5.1 applies to $I = (r, r', m_\varepsilon; L)$.

PROOF OF LEMMA 5.2. To prove the lemma a family of measures m_ε which depends on $\mu_{\mathscr{R}}$ and for which each of the m_ε with $\varepsilon > 0$ satisfies (5.5) has to be exhibited. For $\varepsilon < \frac{1}{2}$, m_ε will be constructed from $\mu_{\mathscr{R}}$ in a completely explicit and very simple way; the definition of m_ε for $\varepsilon \geqq \frac{1}{2}$ is of course almost immaterial, so for $\varepsilon \geqq \frac{1}{2}$ let m_ε concentrate measure 1 on the point $\frac{1}{2}$. It is enough to prove the lemma for $r_0 = r_0' = 0$, because if $d\mu_{\mathscr{R}}^*(\rho) = \rho^{r_0}(1 - \rho)^{r_0'} d\mu_{\mathscr{R}}(\rho)$, then $(r, r', \mu_{\mathscr{R}}; L) = (r - r_0, r' - r_0', \mu_{\mathscr{R}}^*; L)$.

For $r_0 = r_0' = 0$, let m_ε (for $\varepsilon \leqq \frac{1}{2}$) be the result of shrinking $\mu_{\mathscr{R}}$ toward $\rho = \frac{1}{2}$ by the factor $1 - 2\varepsilon$. For any set $S \subset [0, 1]$ and $0 < \varepsilon < \frac{1}{2}$ define the set S_ε to be $\varepsilon + (1 - 2\varepsilon)S$ in the usual algebraic sense, so that for all $\rho \in [0, 1]$, $\rho \in S_\varepsilon$ iff $(\rho - \varepsilon)/(1 - 2\varepsilon) \in S$. Define $m_\varepsilon(S) = \mu_{\mathscr{R}}(S_\varepsilon)$ for any set $S \subset [0, 1]$ such that S is Borel measurable. Then $m_\varepsilon([0, \varepsilon)) = m_\varepsilon((1 - \varepsilon, 1]) = 0$, and (5.5) holds as long as $\varepsilon > 0$.

To see that (5.6) holds at $n = 1$ for the family of measures m_ε, write

$$\begin{aligned} \nu(r, r'; m_\varepsilon) &= \int_0^1 \rho^r (1 - \rho)^{r'} \, dm_\varepsilon(\rho) = \int_\varepsilon^{1-\varepsilon} \rho^r (1 - \rho)^{r'} \, d\mu_{\mathscr{R}}\left(\frac{\rho - \varepsilon}{1 - 2\varepsilon}\right) \\ &= \int_0^1 [x + \varepsilon(1 - 2x)]^r [1 - x - \varepsilon(1 - 2x)]^{r'} \, d\mu_{\mathscr{R}}(x) \\ &= \int_0^1 x^r (1 - x)^{r'} \, d\mu_{\mathscr{R}}(x) + o(1) , \end{aligned} \tag{5.7}$$

for $r, r' \geqq 0$ according to the Lebesgue dominated convergence theorem. Therefore,

$$E(\rho \mid r, r'; m_\varepsilon) = \frac{\nu(r+1, r'; m_\varepsilon)}{\nu(r, r'; m_\varepsilon)} = \frac{\nu(r+1, r'; \mu_{\mathscr{R}}) + o(1)}{\nu(r, r'; \mu_{\mathscr{R}}) + o(1)} \tag{5.8}$$
$$= E(\rho \mid r, r'; \mu_{\mathscr{R}}) + o(1)\,,$$

which proves (5.6) for $n = 1$.

For $n \geqq 2$, in view of (3.7) and (5.7)

$$\begin{aligned}\Delta_n(r, r', m_\varepsilon; L) &= E(\rho \mid r, r'; m_\varepsilon)\Delta^+_{n-1}(r+1, r', m_\varepsilon; L) \\ &\quad + \bar{E}(\rho \mid r, r'; m_\varepsilon)\Delta^+_{n-1}(r, r'+1, m_\varepsilon; L) \\ &\quad + E(\lambda)\Delta^-_{n-1}(r, r', m_\varepsilon; \sigma L) + \bar{E}(\lambda)\Delta^-_{n-1}(r, r', m_\varepsilon; \varphi L) \\ &= E(\rho \mid r, r'; \mu_{\mathscr{R}})\Delta^+_{n-1}(r+1, r', m_\varepsilon; L) \qquad (5.9) \\ &\quad + \bar{E}(\rho \mid r, r'; \mu_{\mathscr{R}})\Delta^+_{n-1}(r, r'+1, m_\varepsilon; L) \\ &\quad + E(\lambda)\Delta^-_{n-1}(r, r', m_\varepsilon; \sigma L) \\ &\quad + \bar{E}(\lambda)\Delta^-_{n-1}(r, r', m_\varepsilon; \varphi L) + o(1)\,.\end{aligned}$$

Assuming that (5.6) holds at $n - 1$ for $r, r' \geqq 0$ and for all L and, in view of (5.9),

$$\begin{aligned}\lim_{\varepsilon \downarrow 0} \Delta_n(r, r', m_\varepsilon; L) &= E(\rho \mid r, r'; \mu_{\mathscr{R}})\Delta^+_{n-1}(r+1, r', \mu_{\mathscr{R}}; L) \\ &\quad + \bar{E}(\rho \mid r, r'; \mu_{\mathscr{R}})\Delta^+_{n-1}(r, r'+1, \mu_{\mathscr{R}}; L) \qquad (5.10) \\ &\quad + E(\lambda)\Delta^-_{n-1}(r, r', \mu_{\mathscr{R}}; \sigma L) \\ &\quad + \bar{E}(\lambda)\Delta^-_{n-1}(r, r', \mu_{\mathscr{R}}; \varphi L) = \Delta_n(r, r', \mu_{\mathscr{R}}; L). \ \square\end{aligned}$$

Lemma 5.2 will be used to prove the next theorem, which extends Lemma 5.1 to include arbitrary measures. The proof of $\hat{J}(n)$ depends only on $\hat{J}(n)$ of Lemma 5.1 and the proof of $\hat{K}(n)$ depends only on $\hat{K}(n)$ of Lemma 5.1.

THEOREM 5.1. *$\hat{J}(n)$ and $\hat{K}(n)$ of Lemma 5.1 hold for $n > 1$, for all $I = (r, r', \mu_{\mathscr{R}}; L)$ and for all $\delta r, \delta r' \geqq 0$.*

REMARK. It was noted in Section 2 that for all $\delta r, \delta r' \geqq 0$, $(r + \delta r, r' + \delta r')$ is in the possibility region for $\mu_{\mathscr{R}}$ whenever (r, r') is, unless $\mu_{\mathscr{R}}((0, 1)) = 0$, that is, unless $\mu_{\mathscr{R}}(0) + \mu_{\mathscr{R}}(1) = 1$. In the latter event the possibility region for $\mu_{\mathscr{R}}$ consists at most of the nonnegative axes. For such measures $\hat{J}(n)$ and $\hat{K}(n)$ may be meaningless, depending on δr and $\delta r'$. The convention is adopted here that $\hat{J}(n)$ and $\hat{K}(n)$ have content only if $(r + \delta r, r' + \delta r')$ is in the possibility region for $\mu_{\mathscr{R}}$. This convention does not exclude the extreme one-point or two-point measures from consideration in the theorem, but it does eliminate consideration of any direction out from the possibility region for $\mu_{\mathscr{R}}$. These easy special cases are not explicitly covered in the proof below.

PROOF OF THEOREM 5.1. Assume that L is a distribution for which $\hat{J}(n)$ is false. Say for $\delta r, \delta r' \geqq 0$,

$$E(\rho \mid r + \delta r + n - 1, r' + \delta r'; \mu_{\mathscr{R}}) - E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}}) \geqq 0\,, \tag{5.11}$$

while

$$\Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) - \Delta_n(r, r', \mu_{\mathscr{R}}; L) < 0 . \tag{5.12}$$

Unless $\mu_{\mathscr{R}}$ is a one-point measure, in which case the theorem is already known to hold, or unless $\mu_{\mathscr{R}}((0, 1)) = 0$, which is a case not currently under discussion, if (5.11) and (5.12) can hold at all, they hold with strict inequality in (5.11), as will now be argued.

Either δr or $\delta r'$ is positive. Say for definiteness that $\delta r > 0$; the other possibility is similar. If δr is replaced by a slightly large value (5.12) will not be lost since $\Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L)$ is continuous in δr for $\delta r > 0$, according to Lemma 4.2. But if δr is increased, (5.11) will be rendered strict according to Lemma 4.6.

In view of (5.12) and Lemma 5.2 there is a measure m_ε which satisfies (2.4) and which approximates $\mu_{\mathscr{R}}$ sufficiently well to guarantee that

$$\Delta_n(r + \delta r, r' + \delta r', m_\varepsilon; L) - \Delta_n(r, r' m_\varepsilon; L) < 0 \tag{5.13}$$

for sufficiently small ε and also, since (5.11) is now supposed to hold with strict inequality,

$$E(\rho \mid r + \delta r + n - 1, r' + \delta r'; m_\varepsilon) - E(\rho \mid r + n - 1, r'; m_\varepsilon) > 0 . \tag{5.14}$$

This inequality contradicts $\hat{J}(n)$ of Lemma 5.1.

A similar argument delivers $\hat{K}(n)$. □

The next theorem strengthens Theorem 5.1 to show that a strict increase in $E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}})$ or a strict decrease in $E(\rho \mid r, r' + n - 1; \mu_{\mathscr{R}})$ guarantees a strict increase or decrease in $\Delta_n(r, r', \mu_{\mathscr{R}}; L)$ for all L and n.

THEOREM 5.2. *The following statements are true for* $n \geqq 1$, *for all* $I = (r, r', \mu_{\mathscr{R}}; L)$, *and for all* $\delta r, \delta r' \geqq 0$:

$$J^*(n): \Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) > \Delta_n(r, r', \mu_{\mathscr{R}}; L)$$
$$\text{if} \quad E(\rho \mid r + \delta r + n - 1, r' + \delta r'; \mu_{\mathscr{R}}) > E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}}) ;$$
$$K^*(n): \Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) < \Delta_n(r, r', \mu_{\mathscr{R}}; L)$$
$$\text{if} \quad E(\rho \mid r + \delta r, r' + \delta r' + n - 1; \mu_{\mathscr{R}}) < E(\rho \mid r, r' + n - 1; \mu_{\mathscr{R}}) .$$

REMARKS. The theorem is true for all distributions $R = (r, r'; \mu_{\mathscr{R}})$, but the conditions in $J^*(n)$ and $K^*(n)$ clearly indicate that the theorem has no content if $\mu_{\mathscr{R}}$ is a one-point measure, for in that case R is not affected by changes in r or r'.

The proof of Theorem 5.2 can be viewed as a modification of the proof of Theorem 4.1 (with differences playing the role of derivatives). The key to the modification is the demonstration that under the condition in $J^*(n)$ or in $K^*(n)$ the four terms in (3.7) cannot vanish simultaneously—they may all vanish if R is a one-point distribution (but only when L is the same one-point distribution).

Like Theorem 5.1, Theorem 5.2 can easily be interpreted as true in case $\mu_{\mathscr{R}}$ is confined to the two extreme points $\{0, 1\}$, but this possibility will not be attended to in the following proof.

Suppose δr and $\delta r'$ are positive and equality holds in the second comparison in $J^*(n)$ (or in $K^*(n)$); can the first inequality nonetheless be concluded? No, not if $\mu_{\mathscr{R}}$ is carried by at most two points, as the attentive reader may perceive (in view of Lemma 4.6), but otherwise it does, though this extension of the theorem will not be carried out in the present paper.

PROOF OF THEOREM 5.2. The theorem will be proved by induction, starting at $n = 1$, where it is trivial.

In view of (3.7), for $n \geqq 2$,

$$
\begin{aligned}
&\Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) - \Delta_n(r, r', \mu_{\mathscr{R}}; L) \\
&\quad = E(\rho \mid r + \delta r, r' + \delta r'; \mu_{\mathscr{R}})\Delta^+_{n-1}(r + \delta r + 1, r' + \delta r' \mu_{\mathscr{R}}; L) \\
&\qquad + \bar{E}(\rho \mid r + \delta r, r' + \delta r'; \mu_{\mathscr{R}})\Delta^+_{n-1}(r + \delta r, r' + \delta r' + 1, \mu_{\mathscr{R}}; L) \\
&\qquad + E(\lambda)\Delta^-_{n-1}(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; \sigma L) \\
&\qquad + \bar{E}(\lambda)\Delta^-_{n-1}(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; \varphi L) \\
&\qquad - E(\rho \mid r, r'; \mu_{\mathscr{R}})\Delta^+_{n-1}(r + 1, r', \mu_{\mathscr{R}}; L) \\
&\qquad - \bar{E}(\rho \mid r, r'; \mu_{\mathscr{R}})\Delta^+_{n-1}(r, r' + 1, \mu_{\mathscr{R}}; L) \\
&\qquad - E(\lambda)\Delta^-_{n-1}(r, r', \mu_{\mathscr{R}}; \sigma L) - \bar{E}(\lambda)\Delta^-_{n-1}(r, r', \mu_{\mathscr{R}}; \varphi L) .
\end{aligned} \tag{5.15}
$$

After some algebra, the right side of (5.15) becomes

$$
\begin{aligned}
&E(\rho \mid r, r'; \mu_{\mathscr{R}})[\Delta^+_{n-1}(r + \delta r + 1, r' + \delta r', \mu_{\mathscr{R}}; L) - \Delta^+_{n-1}(r + 1, r', \mu_{\mathscr{R}}; L)] \\
&\quad + \bar{E}(\rho \mid r, r'; \mu_{\mathscr{R}})[\Delta^+_{n-1}(r + \delta r, r' + \delta r' + 1, \mu_{\mathscr{R}}; L) - \Delta^+_{n-1}(r, r' + 1, \mu_{\mathscr{R}}; L)] \\
&\quad + (E(\rho \mid r + \delta r, r' + \delta r'; \mu_{\mathscr{R}}) - E(\rho \mid r, r'; \mu_{\mathscr{R}})) \\
&\quad \times [\Delta^+_{n-1}(r + \delta r + 1, r' + \delta r', \mu_{\mathscr{R}}; L) - \Delta^+_{n-1}(r + \delta r, r' + \delta r' + 1, \mu_{\mathscr{R}}; L)] \\
&\quad + E(\lambda)[\Delta^-_{n-1}(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; \sigma L) - \Delta^-_{n-1}(r, r', \mu_{\mathscr{R}}; \sigma L)] \\
&\quad + \bar{E}(\lambda)[\Delta^-_{n-1}(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; \varphi L) - \Delta^-_{n-1}(r, r', \mu_{\mathscr{R}}; \varphi L)] .
\end{aligned} \tag{5.16}
$$

For $n \geqq 2$, $J^*(n-1)$ and $K^*(n-1)$ apply to show that

$$
\Delta_{n-1}(\sigma R, L) > \Delta_{n-1}(R, L) > \Delta_{n-1}(\varphi R, L) , \tag{5.17}
$$

which is a strict inequality version of (4.16). (5.17) implies that at least one term of the right side of (3.7) is nonzero. For, in view of (5.17), Theorem 5.1, and the fact that $\Delta_n(R, L) = -\Delta_n(L, R)$,

$$
\Delta_{n-1}(\sigma R, L) > \Delta_{n-1}(R, L) = -\Delta_{n-1}(L, R) \geqq -\Delta_{n-1}(\sigma L, R) = \Delta_{n-1}(R, \sigma L) . \tag{5.18}
$$

Therefore, either $\Delta_{n-1}(\sigma R, L) > 0$ or $\Delta_{n-1}(R, \sigma L) < 0$.

Assume $E(\rho \mid r + \delta r + n - 1, r' + \delta r'; \mu_{\mathscr{R}}) \geqq E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}})$, then $J^*(n-1)$ implies that the bracketed portion of the first term of (5.16) is positive when $\Delta_{n-1}(r + 1, r', \mu_{\mathscr{R}}; L) > 0$ and the bracketed portion of the fourth term of (5.16) is positive when $\Delta_{n-1}(r, r', \mu_{\mathscr{R}}; \sigma L) < 0$. Therefore, the first term of (5.16) is positive when $\Delta_{n-1}(r + 1, r', \mu_{\mathscr{R}}; L) > 0$ since $E(\rho \mid r, r'; \mu_{\mathscr{R}})$ cannot then be zero and the fourth term of (5.16) is positive when $\Delta_{n-1}(r, r', \mu_{\mathscr{R}}; \sigma L) < 0$ since $E(\lambda)$ cannot then be zero. In either case the remaining terms of (5.16)

are nonnegative in view of $\hat{J}(n-1)$ of Theorem 5.1, so that $J^*(n-1)$ implies $J^*(n)$.

A similar argument uses $K^*(n-1)$ of the theorem and $\hat{K}(n-1)$ of Theorem 5.1 to deliver $K^*(n)$. □

6. Results that hold for all n. In this section, the inequalities derived in the previous section will be used to examine parts of the domain space where the sign of $\Delta_n(I)$ is the same for all n; Theorem 6.3 is the only result in this section which depends on n. The conclusions rest on the principal theorems, Theorems 5.1 and 5.2, but only for the special cases $\delta r' = 0$ in $\hat{J}(n)$ and $J^*(n)$ and $\delta r = 0$ in $\hat{K}(n)$ and $K^*(n)$. These theorems will be used in their full generality in the next section.

THEOREM 6.1. *For all $I = (R, L)$ and $n \geqq 2$,*

$$\Delta_n(R, L) \leqq \Delta^+_{n-1}(\sigma R, L), \tag{6.1}$$

with strict inequality if $0 \leqq \Delta_n(R, L)$ and R is not a one-point distribution.

PROOF: According to (3.7),

$$\Delta_n(R, L) \leqq E(\rho)\Delta^+_{n-1}(\sigma R, L) + \bar{E}(\rho)\Delta^+_{n-1}(\varphi R, L), \tag{6.2}$$

and the inequality is strict unless $0 \leqq \Delta_{n-1}(R, \sigma L)$. The right side of (6.2) is

$$\Delta^+_{n-1}(\sigma R, L) - \bar{E}(\rho)(\Delta^+_{n-1}(\sigma R, L) - \Delta^+_{n-1}(\varphi R, L)) \leqq \Delta^+_{n-1}(\sigma R, L), \tag{6.3}$$

in view of (4.17), which holds for all R according to Theorem 5.1. In view of (5.17), the strict-inequality version of (4.17), inequality (6.3) is strict when R is not a one-point distribution unless $\Delta_{n-1}(\sigma R, L) \leqq 0$. But if R is not a one-point distribution, $\Delta_{n-1}(\sigma R, L) \leqq 0 \leqq \Delta_{n-1}(R, \sigma L)$ cannot be satisfied in view of (5.18); therefore, in this case either inequality (6.2) or inequality (6.3) is strict. □

Bradt, Johnson, and Karlin (1956) prove the following result (which they call the "stay-on-a-winner-rule") for the one-armed bandit problem. (A two-armed bandit is called a *one-armed bandit* if ρ or λ is known with probability one; that is, if R or L is a one-point distribution.) Quisel (1965) offers a proof of this result for the two-armed bandit that is different from the present proof. It is easy to see that the stay-on-a-winner-rule is not optimal if ρ and λ are dependent (Bradt *et al.* (1956) or Fabius and van Zwet (1970)). Theorem 6.2 is an immediate corollary of Theorem 6.1; it means that if an arm is optimal and pulled and yields a success, then it is optimal on the next pull as well.

THEOREM 6.2. *For all patterns $I = (R, L)$ for which R is not a one-point distribution and $n \geqq 2$, $\Delta_n(R, L) \geqq 0$ implies $\Delta_{n-1}(\sigma R, L) > 0$. If R is a one-point distribution then $\Delta_n(R, L) \geqq 0$ implies $\Delta_{n-1}(\sigma R, L) \geqq 0$ and $\Delta_n(R, L) > 0$ implies $\Delta_{n-1}(\sigma R, L) > 0$.*

Nothing can be said in general about the relationship between $\Delta_n(R, L)$ and $\Delta_{n-1}(\varphi R, L)$, either can be less than the other; in fact, either can be positive

and the other negative. For example, suppose $n = 2$ and L is determined by $\mu_{\mathscr{L}}(\lambda) = \beta(\lambda) = \lambda^{-1}(1 - \lambda)^{-1}$ and $l = l' = 1$. If R is such that $\mu_{\mathscr{R}} = \beta$ and $r = r' = \frac{1}{2}$, then using the notation $N_{\mathscr{R}} = r + r'$ and $N_{\mathscr{L}} = l + l'$,

$$\Delta_2(R, L) = E(\rho^2) - E(\lambda^2) = \frac{r}{N_{\mathscr{R}}}\frac{r+1}{N_{\mathscr{R}}+1} - \frac{l}{N_{\mathscr{L}}}\frac{l+1}{N_{\mathscr{L}}+1} = \frac{3}{8} - \frac{1}{3} = \frac{1}{24};$$

while, according to (3.8),

$$\begin{aligned}\Delta_1(\varphi R, L) &= \frac{E(\rho) - E(\rho^2)}{1 - E(\rho)} - E(\lambda) = \frac{r}{N_{\mathscr{R}}+1} - \frac{l}{N_{\mathscr{L}}} \\ &= \frac{1}{4} - \frac{1}{2} \doteq -\frac{1}{4} < \Delta_2(R, L).\end{aligned}$$

If, however, $r = 11$ and $r' = 9$, then

$$\begin{aligned}\Delta_2(R, L) &= E(\rho) - E(\lambda) - E(\lambda^2) + E(\rho)E(\lambda) \\ &= \frac{r}{N_{\mathscr{R}}} - \frac{l}{N_{\mathscr{L}}} - \frac{l}{N_{\mathscr{L}}}\frac{l+1}{N_{\mathscr{L}}+1} + \frac{r}{N_{\mathscr{R}}}\frac{l}{N_{\mathscr{L}}} \\ &= \frac{11}{20} - \frac{1}{2} - \frac{1}{2}\frac{2}{3} + \frac{11}{20}\frac{1}{2} = -\frac{1}{120};\end{aligned} \tag{6.4}$$

while,

$$\Delta_1(\varphi R, L) = \frac{r}{N_{\mathscr{R}}+1} - \frac{l}{N_{\mathscr{L}}} = \frac{11}{21} - \frac{1}{2} = \frac{1}{42} > \Delta_2(R, L). \tag{6.5}$$

That $\Delta_n(R, L)$ and $\Delta_{n-1}(\varphi R, L)$ are not related in the way that $\Delta_n(R, L)$ and $\Delta_{n-1}(\sigma R, L)$ are, is a manifestation of the asymmetry of the two-armed bandit problem in successes and failures, an asymmetry not evinced by Theorems 5.1 and 5.2. Heuristically, a success on an optimal arm never decreases (and typically increases) the inclination to pull that arm again, while a failure on an optimal arm (obviously could decrease, but also) can increase the inclination to pull the arm again. The number of pulls remaining has been lessened by 1, leaving less time to take advantage of anything learned.

The one-armed bandit can be instructive in this regard. Suppose that R is a one-point distribution and that L is not a one-point distribution. In this case $\varphi R = \sigma R = R$ and (6.1) yields

$$\Delta_n(R, L) \leqq \Delta^+_{n-1}(\varphi R, L) = \Delta^+_{n-1}(R, L).$$

It can easily happen that the left arm is worth pulling on the first of n pulls remaining on the chance that it is really better than the right arm, and not worth pulling on the first of $n - 1$ pulls remaining. The latter example above (with calculations in (6.4) and (6.5)) is much like this one-armed bandit since $N_{\mathscr{R}}$ is large relative to $N_{\mathscr{L}}$.

The next theorem, Theorem 6.3, is the only result in this section which depends on n, but it is really a corollary of Theorem 6.1 which is true for all n.

The intuitive notion of Theorem 6.3 is that an arm should be pulled at the last stage (that is, when $n = 1$) if it was optimal at some previous stage and has since yielded all successes. The theorem gives a crude but easily computable sufficient condition on the distributions R and L for the optimality of $\mathscr{L}$; and, of course, there is a symmetric condition for the optimality of $\mathscr{R}$.

THEOREM 6.3. *For all n and $I = (r, r', \mu_{\mathscr{R}}; L)$, if*

$$E(\lambda) \geqq E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}}), \tag{6.6}$$

then $\Delta_n(I) \leqq 0$, and $\Delta_n(I) < 0$ if R is not a one-point distribution.

REMARK. In the special case $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \beta$, condition (6.6) becomes:

$$\frac{l}{N_{\mathscr{L}}} \geqq \frac{r + n - 1}{N_{\mathscr{R}} + n - 1}.$$

PROOF OF THEOREM 6.3. Assume $\Delta_n(R, L) > 0$. Applying Theorem 6.1 $n - 1$ times, $\Delta_{n-1}(\sigma R, L) > 0$, $\Delta_{n-2}(\sigma^2 R, L) > 0, \cdots, \Delta_1(\sigma^{n-1}R, L) = E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}}) - E(\lambda) > 0$. Thus $\Delta_n(R, L) \leqq$ follows by contradiction. If R is not a one-point distribution and $n \geqq 2$ then $\Delta_n(R, L) \geqq 0$ is similarly contradicted. □

Condition (6.6) is more easily satisfied for small n since $E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}})$ is nondecreasing (and typically increasing) in n. Moreover, if R associates probability to all intervals $(1 - \varepsilon, 1]$, $\varepsilon > 0$, then

$$\lim_{n\to\infty} E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}}) = 1;$$

and if R is such a distribution, (6.6) would be satisfied for very large n only if, under L, $\lambda = 1$ with probability one. For fixed n and $E(\lambda)$, (6.6) is more easily satisfied for distributions R that concentrate probability near $E(\rho)$. For example, if R is a one-point distribution, $E(\rho \mid r + n - 1, r'; \mu_{\mathscr{R}}) = E(\rho \mid r, r'; \mu_{\mathscr{R}})$ and the problem is a one-armed bandit, then Theorem 6.3 implies that $\mathscr{L}$ is optimal for all n whenever $E(\lambda) \geqq E(\rho)$. This application of Theorem 6.3 is intuitive since a left arm which will yield at least as much immediate expected payoff and at least as much information as the right arm would seem to be optimal.

In the remaining results of this section, R and L are assumed to be conjugate with respect to each other; that is, given R and L there exist $\mu_{\mathscr{R}}$ and $\mu_{\mathscr{L}}$ such that $\mu_{\mathscr{R}} = \mu_{\mathscr{L}}$. The next result means that whenever one of the two comparable arms has a greater "effective number" of successes and a smaller "effective number" of failures, it is optimal. Many instances of the next theorem follow from Theorem 5 of Fabius and van Zwet (1970) which is a much more general result, applying to possibly dependent arms.

THEOREM 6.4. *Provided $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \mu$, if $r \geqq l$ and $r' \leqq l'$, then $\Delta_n(I) \geqq 0$ for all n and I.*

PROOF: In view of the conditions, l and l' can be written $r - \delta r$ and $r' + \delta r'$ for δr, $\delta r' \geqq 0$. Applying first $\hat{J}(n)$ of Theorem 5.1 for $\delta r' = 0$, then $\hat{K}(n)$ of

Theorem 5.1 for $\delta r = 0$ then symmetry,

$$\Delta_n(r, r', \mu; L) \geqq \Delta_n(r - \delta r, r', \mu; L)$$
$$\geqq \Delta_n(r - \delta r, r' + \delta r', \mu; L) = \Delta_n(l, l', \mu; L) = 0 . \square$$

Theorem 6.4 will be discussed in the form of the following immediate corollaries; the first gives a sufficient condition for the optimality of $\mathscr{L}$ and the second gives a sufficient condition for the optimality of $\mathscr{R}$, both under the additional condition that $N_{\mathscr{R}} \leqq N_{\mathscr{L}}$.

COROLLARY 1. *If $N_{\mathscr{R}} \leqq N_{\mathscr{L}}$ and $r' \geqq l'$, then $\Delta_n(I) \leqq 0$ for all n and I, provided $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \mu$.*

COROLLARY 2. *If $N_{\mathscr{R}} \leqq N_{\mathscr{L}}$ and $r \geqq l$, then $\Delta_n(I) \geqq 0$ for all n and I, provided $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \mu$.*

If, in addition to the conditions of Theorem 6.4, $n \geqq 2$ and $\mu_{\mathscr{R}}$ is not a one-point measure, then Theorem 5.2 can be applied to strengthen 6.4. The next theorem is given for completeness, its proof, which will not be given explicitly, uses Theorem 5.2 in the same way that the proof of Theorem 6.4 uses Theorem 5.1.

THEOREM 6.5. *Provided $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \mu$ is not a one-point measure, if $r > l$ and $r' \leqq l'$ or $r \geqq l$ and $r' < l'$, then $\Delta_n(I) > 0$ for all n and I.*

It seems intuitive that Corollary 1 of Theorem 6.4 cannot be improved; that is, for given $\mu = \mu_{\mathscr{R}} = \mu_{\mathscr{L}}$, if $N_{\mathscr{R}} \leqq N_{\mathscr{L}}$ then the only patterns I for which $\Delta_n(I) \leqq 0$ for all n have $r' \geqq l'$. Because for large n the effective numbers of successes, r and l, would seem to matter less than they do for small n. If not more is known about $\mathscr{R}$ than about $\mathscr{L}$ (which in a sense is expressed by $N_{\mathscr{R}} \leqq N_{\mathscr{L}}$ whenever $\mu_{\mathscr{R}} = \mu_{\mathscr{L}}$) and n is large, then obtaining a success on the current pull matters little compared to the possibility of learning something on the current pull about $\mathscr{R}$ that will increase the number of future pulls on the better arm, except that learning something about arm $\mathscr{R}$ if arm $\mathscr{L}$ must be used eventually in any case (and it will if $r' \geqq l'$) can hardly be very worthwhile. I conjecture that Corollary 1 barely holds in the limit as $n \to \infty$; and more, that for a large number of remaining pulls, the only criterion for optimality is the difference between the effective numbers of failures on the two arms. That is, Conjecture A: For any $\mu = \mu_{\mathscr{R}} = \mu_{\mathscr{L}}$ and all sufficiently large n, $\Delta_n(I)$ has the same sign as $l' - r'$ independent of r and l.

On the other hand, Corollary 2 of Theorem 6.4 seems very weak compared to what should be true for all n. For, whenever less is known about arm $\mathscr{R}$ (and therefore, more information is gained by pulling $\mathscr{R}$) and $\mathscr{R}$ offers greater expected immediate payoff, then $\mathscr{R}$ should be optimal. This is supported by many unsuccessful searches for counterexamples. Conjecture B: For all n and I, if $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \mu$, $N_{\mathscr{R}} \leqq N_{\mathscr{L}}$, and $E(\rho) \geqq E(\lambda)$, then $\Delta_n(I) \geqq 0$.

Conjecture B is implied by the notion that as more becomes known about

arm $\mathscr{R}$, say, and the expected immediate payoff on $\mathscr{R}$ remains the same ($= E(\rho \mid R)$), the advantage of $\mathscr{R}$ over $\mathscr{L}$ does not increase. This notion can be shown to be equivalent to the following, which is stated in a manner so as to emphasize that it is stronger than $\hat{K}(n)$ of Theorem 5.1. Conjecture C: For all n, for all $I = (r, r', \mu_{\mathscr{R}}; L)$, and for $\delta r, \delta r' \geqq 0$, $\Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) \leqq \Delta_n(r, r', \mu_{\mathscr{R}}; L)$ if $E(\rho \mid r + \delta r, r' + \delta r'; \mu_{\mathscr{R}}) \leqq E(\rho \mid r, r'; \mu_{\mathscr{R}})$.

Conjecture C would also imply many instances of the following conjecture in Chernoff (1968): Let R and L be arbitrary distributions, and R^* a degenerated R, the one-point distribution that concentrates probability one on $E(\rho)$; then $\Delta_n(R, L) \geqq 0$ if $\Delta_n(R^*, L) \geqq 0$ for all n. This would mean that the solution of the two-armed bandit problem is partially determined by the solution of a corresponding one-armed bandit problem. For any point R in the possibility region of $\mu_{\mathscr{R}}$, the corresponding R^* is in the direction (a, b) defined by $D_{(a,b)}E(\rho \mid r, r'; \mu_{\mathscr{R}}) = 0$ (provided R^* is in the possibility region for $\mu_{\mathscr{R}}$; that is, provided R is such that $R(E(\rho) - \varepsilon, E(\rho) + \varepsilon) > 0$ for all $\varepsilon > 0$), so that Conjecture C would imply $\Delta_n(R, L) \geqq \Delta_n(R^*, L)$.

Conjecture C would also imply that the solution of the two-armed bandit is partially determined (in the other direction) by the solution of a particular two-armed bandit, one in which one of the arms, say $\mathscr{R}$, produces either all successes (with probability $E(\rho)$) or all failures (with probability $\bar{E}(\rho)$), and one pull on $\mathscr{R}$ will, with probability one, reveal which. Let R be an arbitrary distribution with expected value $E(\rho)$, and R_* the distribution which concentrates probabilities $E(\rho)$ and $\bar{E}(\rho)$ at $\rho = 1$ and $\rho = 0$, respectively, then the direction (a, b) in the (r, r') plane from R_* to R (provided R_* is in the possibility region for $\mu_{\mathscr{R}}$) is defined by $D_{(a,b)}E(\rho \mid r, r'; \mu_{\mathscr{R}}) = 0$, and Conjecture C would imply $\Delta_n(R_*, L) \geqq \Delta_n(R, L)$. ($R_*$ is in the possibility region for $\mu_{\mathscr{R}}$ provided R is such that $R[0, \varepsilon) > 0$ and $R(1 - \varepsilon, 1] > 0$; if $\mu_{\mathscr{R}}$ is not such a measure, then this inequality would follow from Conjecture C, but for an R_* different from the one defined here.)

7. Results that depend on n. In the previous section, $\hat{J}(n)$ and $J^*(n)$ of Theorems 5.1 and 5.2 are applied for $\delta r' = 0$ and $\hat{K}(n)$ and $K^*(n)$ for $\delta r = 0$, particularly when $\mu_{\mathscr{R}} = \mu_{\mathscr{L}}$. In the present section, Theorems 5.1 and 5.2 are applied in their full generality when $\mu_{\mathscr{R}} = \mu_{\mathscr{L}}$. Theorem 6.4 determines the sign of $\Delta_n(I)$ when $r \geqq l$ and $r' \leqq l'$ (and, of course, when $r \leqq l$ and $r' \geqq l'$); each of the theorems in this section determines the sign of $\Delta_n(I)$ when $r \leqq l$ and $r' \leqq l'$ under an additional condition, which depends on n. Theorem 7.1 uses $\hat{J}(n)$ of Theorem 5.1 to determine a sufficient condition for the optimality of $\mathscr{L}$ and the very closely parallel Theorem 7.2 uses $\hat{K}(n)$ of Theorem 5.1 to determine sufficient conditions for the optimality of $\mathscr{R}$.

THEOREM 7.1. *For all n and $I = (r, r', \mu; l, l', \mu)$, if $r \leqq l$ and $r' \leqq l'$, and $E(\rho \mid r + n - 1, r'; \mu) \leqq E(\lambda \mid l + n - 1, l'; \mu)$ then $\Delta_n(I) \leqq 0$.*

PROOF. In view of the first two conditions of the theorem, l and l' can be

written $r + \delta r$ and $r' + \delta r'$ for $\delta r, \delta r' \geqq 0$. The third condition of the theorem then becomes the condition in $\hat{J}(n)$ of Theorem 5.1; therefore,

$$0 = \Delta_n(l, l', \mu; L) = \Delta_n(r + \delta r, r' + \delta r', \mu; L) \geqq \Delta_n(r, r', \mu; L) . \quad \square$$

THEOREM 7.2. *For all n and $I = (r, r', \mu; l, l', \mu)$, if $r \leqq l$ and $r' \leqq l'$ and $E(\rho \mid r, r' + n - 1; \mu) \geqq E(\lambda \mid l, l' + n - 1; \mu)$, then $\Delta_n(I) \geqq 0$.*

The proof of the latter theorem is strictly parallel to that of Theorem 7.1 with $\hat{K}(n)$ of Theorem 5.1 playing the role of $\hat{J}(n)$.

When μ is not a one-point measure and $n \geqq 2$, Theorems 7.1 and 7.2 can be strengthened just as Theorem 6.4 is strengthened by Theorem 6.5. The next two theorems accomplish this. Their proofs will not be given explicitly; they can be proved by applying Theorem 5.2 in the same way that the proofs of Theorems 7.1 and 7.2 apply Theorem 5.1.

THEOREM 7.3. *Provided μ is not a one-point measure and $n \geqq 2$, if $I = (r, r', \mu; l, l', \mu)$, $r < l$ and $r' \leqq l'$ or $r \leqq l$ and $r' < l'$, and $E(\rho \mid r + n - 1, r'; \mu) \leqq E(\lambda \mid l + n - 1, l'; \mu)$ then $\Delta_n(I) < 0$.*

THEOREM 7.4. *Provided μ is not a one-point measure and $n \geqq 2$, if $I = (r, r', \mu; l, l', \mu)$, $r < l$ and $r' \leqq l'$ or $r \leqq l$ and $r' < l'$, and $E(\rho \mid r, r' + n - 1; \mu) \geqq E(\lambda \mid l, l' + n - 1; \mu)$ then $\Delta_n(I) > 0$.*

Theorem 7.1 and 7.2 will be applied to two previously discussed example two-armed bandit problems in the next section: $\mu = \beta$ in the first example and $\mu = \tau$ in the second.

8. Two important applications. If $\mu = \beta$ as defined by (2.5), the application of Theorems 7.1 and 7.2 is particularly simple. If $R = (r, r'; \beta)$ and $n - 1$ successes are subsequently observed in $n - 1$ pulls on $\mathscr{R}$, the probability of success on the next pull is

$$\begin{aligned} E(\rho \mid r + n - 1, r'; \beta) &= \frac{\nu(r, r'; \beta) \int_0^1 \rho^{r+n}(1 - \rho)^{r'} \, d\beta(\rho)}{\nu(r, r'; \beta) \int_0^1 \rho^{r+n-1}(1 - \rho)^{r'} \, d\beta(\rho)} \\ &= \frac{\int_0^1 \rho^{r+n-1}(1 - \rho)^{r'-1} \, d\rho}{\int_0^1 \rho^{r+n-2}(1 - \rho)^{r'-1} \, d\rho} \\ &= \frac{r + n - 1}{r + r' + n - 1} = \frac{r + n - 1}{N_{\mathscr{R}} + n - 1} ; \end{aligned}$$

for this formula, see the topic of beta integrals in any advanced calculus text, (for example, Widder (1961) Section 11.2). Similarly, if $R = (r, r'; \beta)$ and $n - 1$ failures are observed in $n - 1$ pulls on $\mathscr{R}$, the probability of a failure on the next pull (equals one minus the probability of a success) is

$$E(1 - \rho \mid r, r' + n - 1; \beta) = \frac{r' + n - 1}{N_{\mathscr{R}} + n - 1} = 1 - \frac{r}{N_{\mathscr{R}} + n - 1} .$$

Theorems 8.1 and 8.2 apply Theorems 7.1 and 7.2 in a way that complements

Theorem 6.4. As in the corollaries of Theorem 6.4, it is assumed for definiteness that the effective number of pulls on $\mathscr{L}$ is not smaller than the effective number on $\mathscr{R}$. Strict inequality versions of these theorems follow from Theorems 6.5, 7.3, and 7.4.

THEOREM 8.1. *For all n and I, provided $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \beta$, if $N_{\mathscr{R}} \leqq N_{\mathscr{L}}$ and $(r + n - 1)/(N_{\mathscr{R}} + n - 1) \leqq (l + n - 1)/(N_{\mathscr{L}} + n - 1)$, then $\Delta_n(I) \leqq 0$.*

PROOF. First, assume $r' \geqq l'$. In this case, $r + r' \leqq l + l'$ implies $r \leqq l$ and, therefore, $\Delta_n(I) \leqq 0$ according to Theorem 6.4.

Now, assume $r' < l'$. In this case, $(r + n - 1)/(N_{\mathscr{R}} + n - 1) \leqq (l + n - 1)/(N_{\mathscr{L}} + n - 1)$ implies $r < l$ and, therefore, $\Delta_n(I) \leqq 0$ according of Theorem 7.1. □

THEOREM 8.2. *For all n and I, provided $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \beta$, if $N_{\mathscr{R}} \leqq N_{\mathscr{L}}$ and $r/(N_{\mathscr{R}} + n - 1) \geqq l/(N_{\mathscr{L}} + n - 1)$, then $\Delta_n(I) \geqq 0$.*

The proof of the latter theorem is strictly parallel to that of Theorem 8.1 with Theorem 7.2 playing the role of Theorem 7.1

For a second application assume $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \tau$, then Theorem 5.1 completely resolves the question of which is the better arm to pull. There is an intuitive reason why this problem is so readily solvable. Ordinarily, $(r, r'; \mu_{\mathscr{R}})$ is a two-parameter family of distributions. If $\mu_{\mathscr{R}}$ is not a one-point or two-point measure, then the distribution $(r_1, r_1'; \mu_{\mathscr{R}})$ is different from the distribution $(r_2, r_2'; \mu_{\mathscr{R}})$, unless $r_1 = r_2$, $r_1' = r_2'$. But in case $R = (r, r'; \tau)$,

$$\frac{R(\tau_1)}{R(\tau_2)} = \left(\frac{\tau_1}{\tau_2}\right)^r \left(\frac{1 - \tau_1}{1 - \tau_2}\right)^{r'} = \left(\frac{\tau_1}{\tau_2}\right)^{r - r'A(\tau)}, \tag{8.1}$$

where $A(\tau)$ is given by (4.7).

In view of (8.1), the whole family of distributions R depends only on the parameter $\bar{r} = r - r'A(\tau)$. Therefore, for all n, $\Delta_n(r, r', \tau; L)$ depends on (r, r') through $\bar{r}$ alone and has straight parallel contours in (r, r'). The slope of these contours is $A(\tau)$, the proportion of successes to failures on $\mathscr{R}$ which does not change Δ_n; for example, if $\tau_1 = 1 - \tau_2$ then $A(\tau) = 1$ and the contours of Δ_n in (r, r') are all parallel to the line $r = r'$.

As previously noted, Theorem 5.1 provides a complete specification of the gradient of $\Delta_n(r, r', \tau; L)$ in (r, r'). Therefore, the sign of $\Delta_n(I)$ is completely determined when $\mu_{\mathscr{R}}$ and $\mu_{\mathscr{L}}$ are the same two-point measure. If the same two numbers τ_1 and τ_2 are the only possible probabilities of success on either arm, then it seems clear that that arm should be pulled which is more likely to be the one associated with τ_2, the larger of the two probabilities; that is, the one which is more likely to be successful on the first pull. The next theorem says that this is the case.

THEOREM 8.3. *For all n and I, provided $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \tau$, $\Delta_n(I)$ has the same sign as $E(\rho) - E(\lambda)$.*

REMARK. Theorems 6.4, 7.1, and 7.2 can be cited as in the proofs of Theorems 8.1 and 8.2 to prove Theorem 8.3. However, the fact that $A(r, r'; \tau)$ does not depend on r and r' can be employed more simply to prove the theorem by appealing directly to Theorem 5.1. The algebra is straightforward though cumbersome and is omitted.

Feldman (1962) solved a closed related problem and obtained a similar solution; see also, (Degroot (1970) Section 14.7). In Feldman's problem there are two possible probabilities of success, but the larger is associated with one of the arms and the smaller with the other; which is the better arm is not known. This dependence between the arms is very strong; nevertheless, it will be seen that the solution of Feldman's problem and the solution of the independent two-armed bandit considered in this section can be used to obtain each other. Fabius and van Zwet (1970) obtain a generalization of Feldman's solution to arbitrary joint prior distributions on (ρ, λ).

If $\mu_{\mathscr{R}} = \mu_{\mathscr{L}} = \tau$ then it is possible that $\rho = \lambda = \tau_1$ (the appropriate probability is $R(\tau_1)L(\tau_1)$ since ρ and λ are independent) or that $\rho = \lambda = \tau_2$ (the probability is $R(\tau_2)L(\tau_2)$). If it is known a priori that the arms are identical (that is, either $\rho = \lambda = \tau_1$ or $\rho = \lambda = \tau_2$), then neither arm would be strictly preferred. The only possibilities that influence the size of Δ_n, which determines the preference between the right and left arms, have $\rho \neq \lambda$ (that is, either $\rho = \tau_1$, $\lambda = \tau_2$ or $\rho = \tau_2$, $\lambda = \tau_1$). Therefore, $\Delta_n > 0$ when and only when it is a priori more likely that $\rho = \tau_2$, $\lambda = \tau_1$ than that $\rho = \tau_1$, $\lambda = \tau_2$.

If it is known a priori that either $\rho = \tau_1$, $\lambda = \tau_2$ or $\rho = \tau_2$, $\lambda = \tau_1$, the problem is identical with the one considered by Feldman (1962). Therefore, Feldman's result implies and is implied by Theorem 8.3.

Acknowledgment. I wish to thank Professor L. J. Savage for his guidance and many invaluable suggestions during the preparation of this paper. I am also grateful for many discussions with Professor J. B. Kadane and the suggestions of Professors P. Feder, T. Cover, and a referee.

REFERENCES

BERRY, D. A. (1971). A Bernoulli two-armed bandit. Technical Report No. 151, School of Statistics, Univ. of Minnesota.

BRADT, R. N., JOHNSON, S. M., and KARLIN, S. (1956). On sequential designs for maximizing the sum of n observations. *Ann. Math. Statist* **27** 1060–1070.

CHERNOFF, H. (1968). Optimal stochastic control. *Sankhyā Ser. A* **30** 221–252.

DEGROOT, M. (1970). *Optimal Statistical Decisions*. McGraw-Hill, New York.

DUBINS, L. E. and SAVAGE, L. J. (1965). *How to Gamble If You Must: Inequalities for Stochastic Processes*. McGraw-Hill, New York.

FABIUS, J. and van ZWET, W. R. (1970). Some remarks on the two-armed bandit. *Ann. Math. Statist.* **41** 1906–1916.

FELDMAN, D. (1962). Contributions to the "two-armed bandit" problem. *Ann. Math. Statist.* **33** 847–856.

FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

HARDY, G. H., LITTLEWOOD, J. E., and POLYA, G. (1934). *Inequalities*. Cambridge Univ. Press, New York.

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

QUISEL, K. (1965). Extensions of the two-armed bandit and related processes with on-line experimentation. Technical Report No. 137, Inst. for Math. Studies in the Social Sciences, Stanford Univ.

RAIFFA, H. and SCHLAIFER, R. (1961). *Applied Statistical Decision Theory*. Graduate School of Business Administration, Harvard Univ. Press.

WIDDER, D. V. (1961). *Advanced Calculus*, 2nd. ed. Prentice-Hall, Englewood Cliffs.

YAKOWITZ, S. J. (1969). *Mathamatics of Adaptive Control Processes*. American Elsevier Publishing Co., New York.

Note added in proof. Mr. Charles Schoenfeldt found a counterexample to Conjecture C on page 893 for $n = 3$ after this paper went to the printer. The additional condition $\Delta_n(r + \delta r, r' + \delta r', \mu_{\mathscr{R}}; L) \leqq 0$ may be sufficient to make the conjecture correct.