5. The Optimal Allocation of Trials

In the last chapter a schema ξ was defined as potentially useful when β_ξ, the observed average performance of instances of that schema, was significantly greater than the overall average performance. However, β_ξ is basically a sample average for a random variable (or sequence of random variables) and, as such, is subject to sampling error. For any two schemata ξ and ξ', there is always a non-zero probability that $\mu_{\xi'} > \mu_\xi$ even though $\beta_\xi > \beta_{\xi'}$. This reintroduces in a sharp form the conflict of exploiting what is known vs. obtaining new information. Confidence that the ranking $\beta_\xi > \beta_{\xi'}$ reflects a true ranking $\mu_\xi > \mu_{\xi'}$ can be increased significantly only by allocating additional trials to *both* ξ and ξ'. Thus, we can allocate a trial to exploit the observed best or we can allocate a trial to reduce the probability of error as much as possible but we cannot generally do both at once. Given a string-represented domain $\mathfrak{A}$, it is important to have some idea of what proportion of trials should be allocated to each purpose as the number of trials increases.

Corresponding to each of these objectives—exploitation vs. new information—there is a source of loss. A trial allocated to the *observed* best may actually incur a loss because of sampling error, the observed best being in fact less than the best among the alternatives examined. On the other hand trials intended to maximally reduce the probability of error will generally be allocated to a schema other than the observed best. This means a performance less on the average than the best among known alternatives, when the observations reflect the true ranking. Stated succinctly, more information means a performance loss, while exploitation of the observed best runs the risk of error perpetuated.

Competing sources of loss suggest the possibility of optimization—minimizing expected losses by a proper allocation of trials. If we can determine the optimal allocation for arbitrary numbers of trials, then we can determine the minimum losses to be expected as a function of the number of trials. This in turn can be used as a criterion against which to measure the performance of suggested adaptive plans. Such a criterion will be particularly useful in determining the worth of plans

which use schemata to compare structures in $\mathcal{A}$. The objective of this chapter is to determine this criterion. In the process we will learn a good deal more about schemata and intrinsic parallelism.

1. THE 2-ARMED BANDIT

The simplest precise version of the optimal allocation problem arises when we restrict attention to two random variables, ξ and ξ', with only two possible payoffs, 0 or 1. A trial of ξ produces the payoff 1 with probability p_1 and the payoff 0 with probability $1 - p_1$; similarly ξ' produces 1 with probability p_2 and 0 with probability $1 - p_2$. (For example, such trials could be produced by flipping either of two unbalanced coins, one of which produces heads with probability p_1, the other with probability p_2.) One is allowed N trials on each of which either ξ or ξ' can be selected. The object is to maximize the total payoff (the cumulative number of heads). Clearly if we know that ξ produces payoff 1 with a higher probability then all N trials should be allocated to ξ with a resulting expected accumulation $p_1 \cdot N$. On the other hand if we know nothing initially about ξ and ξ' it would be unwise not to test both. How trials should be allocated to accomplish this is certainly not immediately obvious. (This is a version of the much-studied 2-armed bandit problem, a prototype of important decision problems. Bellman [1961] and Hellman and Cover [1970] give interesting discussions of the problem.)

If we allow the two random variables to be completely general (having probability distributions over an arbitrary number of outcomes), we get a slight generalization of the original problem which makes direct contact with our discussion of schemata. The outcome of a trial of either random variable is to be interpreted as a payoff (performance). The object once more is to discover a procedure for distributing an arbitrary number of trials, N, between ξ and ξ' so as to maximize the expected payoff over the N trials. As before, if we know for each ξ_i the mean and variance (μ_i, σ_i^2) of its distribution (actually the mean μ_i would suffice), the problem has a trivial solution (allocate all trials to the random variable with maximal mean). The conflict asserts itself, however, if we inject just a bit more uncertainty. Thus we can know the mean-variance pairs but not which variable is described by which pair; i.e., we know the pairs (μ_1, σ_1^2) and (μ_2, σ_2^2) but not which pair describes ξ.

If it could be determined through some small number of trials which of ξ and ξ' has the higher mean, then from that point on all trials could be allocated to that random variable. Unfortunately, unless the distributions are non-overlapping, no finite number of observations will establish *with certainty* which random variable has the higher mean. (E.g., given $\mu_\xi > \mu_{\xi'}$ along with a probability $p > 0$ that a trial of ξ' will yield an outcome $x > \mu_\xi$, there is still a probability p^N after N

trials of ξ' that *all* of the trials have had outcomes exceeding μ_ξ. A fortiori their average $\mu_{\xi'}$ will exceed μ_ξ with probability at least p^N, even though $\mu_{\xi'} < \mu_\xi$.) Here the tradeoff between gathering information and exploiting it appears in its simplest terms. To see it in exact form let $\xi_{(1)}(N)$ name the random variable with the highest *observed* payoff rate (average per trial) after N trials and let $\xi_{(2)}(N)$ name the other random variable. For any number of trials n, $0 \leqq n \leqq N$, allocated to $\xi_{(2)}(N)$ (and assuming overlapping distributions) there is a positive probability, $q(N-n, n)$, that $\xi_{(2)}(N)$ is actually the random variable with the highest mean, $\max\{\mu_\xi, \mu_{\xi'}\}$. The two possible sources of loss are: (1) The *observed* best $\xi_{(1)}(N)$ is really second best, whence the $N-n$ trials given $\xi_{(1)}(N)$ incur an (expected) cumulative loss $(N-n)\cdot|\mu_\xi - \mu_{\xi'}|$; this occurs with probability $q(N-n, n)$. (2) The *observed* best is in fact the best, whence the n trials given $\xi_{(2)}(N)$ incur a loss $n\cdot|\mu_\xi - \mu_{\xi'}|$; this occurs with probability $(1 - q(N-n, n))$. The total expected loss for any allocation of n trials to $\xi_{(2)}$ and $N-n$ trials to $\xi_{(1)}$ is thus

$$L(N-n, n) = [q(N-n, n)\cdot(N-n) + (1 - q(N-n, n))\cdot n]\cdot|\mu_\xi - \mu_{\xi'}|.$$

We shall soon see that, for n not too large, the first source of loss decreases as n increases because both $N-n$ and $q(N-n, n)$ decrease. At the same time the second source of loss increases. By making a tradeoff between the first and second sources of loss, then, it is possible to find for each N a value $n^*(N)$ for which the losses are minimized; i.e.,

$$L(N-n^*, n^*) \leqq L(N-n, n) \quad \text{for all } n \leqq N.$$

For the determination of n^* let us assume that *initially* one random variable is as likely as the other to be best. (This would be the case for example if the two unbalanced coins referred to earlier have no identifying external characteristics and are positioned initially at random. More generally, the result is the same if the labels of the random variables are assigned at random. The proof of the theorem will indicate the modifications necessary for cases where one random variable is initially more likely than the other to be the best.) For convenience let us adopt the convention that ξ_1 is the random variable with the highest mean and let μ_1 be that mean; accordingly ξ_2 is the other random variable with mean $\mu_2 \leqq \mu_1$. (The observer, of course, does not know this.) Using these conventions we can now establish

THEOREM 5.1: *Given N trials to be allocated to two random variables, with means $\mu_1 > \mu_2$ and variances σ_1^2, σ_2^2 respectively, the minimum expected loss results when the number of trials allocated $\xi_{(2)}(N)$ is*

$$n \leqq n^* \sim b^2 \ln [N^2/(8\pi b^4 \ln N^2)]$$

where $b = \sigma_1/(\mu_1 - \mu_2)$. *If, initially, one random variable is as likely as the other to be best,* $n = n^*$ *and the expected loss per trial is*

$$L^*(N) \sim (b^2(\mu_1 - \mu_2)/N)[2 + \ln [N^2/(8\pi b^4 \ln N^2)]].$$

(Given two arbitrary functions, $Y(t)$ and $Z(t)$, of the same variable t, "$Y(t) \sim Z(t)$" will be used to mean $\lim_{t\to\infty} (Y(t)/Z(t)) = 1$ while "$Y(t) \cong Z(t)$" means that under stated conditions the difference $(Y(t) - Z(t))$ is negligible.)

Proof: In order to select an n which minimizes the expected loss, it is necessary first to write $q(N - n, n)$ as an explicit function of n. As defined above $q(N - n, n)$ is the probability that $\xi_{(2)}(N) = \xi_1$. More carefully, given the observation, say, that $\xi' = \xi_{(2)}(N)$, we wish to determine the probability that $\xi' = \xi_1$. That is, we wish to determine

$$q(N - n, n) = Pr\{\xi' = \xi_1 \mid \xi_{(2)} = \xi'\}$$

as an explicit function of $N - n$ and n. Bayes's theorem then gives us the equation

$$Pr\{\xi' = \xi_1 \mid \xi_{(2)} = \xi'\} = \frac{Pr\{\xi' = \xi_{(2)} \mid \xi' = \xi_1\}Pr\{\xi' = \xi_1\}}{Pr\{\xi' = \xi_{(2)} \mid \xi' = \xi_1\}Pr\{\xi' = \xi_1\} + Pr\{\xi' = \xi_{(2)} \mid \xi' = \xi_2\}Pr\{\xi' = \xi_2\}}$$

Letting q', q'', and p designate $Pr\{\xi' = \xi_{(2)} \mid \xi' = \xi_1\}$, $Pr\{\xi' = \xi_{(2)} \mid \xi' = \xi_2\}$, and $Pr\{\xi' = \xi_1\}$, respectively, and using the fact that ξ' must be ξ_2 if it is not ξ_1, this can be rewritten as

$$q(N - n, n) = q'p/(q'p + q''(1 - p)).$$

(If one random variable is as likely as the other to be best, then $p = (1 - p) = \frac{1}{2}$.)

To derive q' let us assume that ξ' has received n trials out of the N total. Let S_2^{N-n} be the sum of the outcomes (payoffs) of $N - n$ trials of ξ_2 and let S_1^n be the corresponding sum for n trials of ξ_1. Since q' has $\xi' = \xi_1$ as a condition, q' is just the probability that $S_1^n/n < S_2^{N-n}/(N - n)$ or, equivalently the probability that $(S_1^n/n) - (S_2^{N-n}/(N - n)) < 0$. By the central limit theorem $S_2^{N-n}/(N - n)$ approaches a normal distribution with mean μ_2 and variance $\sigma_2^2/(N - n)$; similarly, S_1^n/n has mean μ_1 and variance σ_1^2/n. The distribution of $(S_1^n/n) - (S_2^{N-n}/(N - n))$ is given by the product (convolution) of the distributions of S_1^n/n and $-(S_2^{N-n}/(N - n))$; by an elementary theorem (on the convolution of normal distributions) this is a normal distribution with mean $\mu_1 - \mu_2$ and variance $\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{N - n}$. Thus the probability $Pr\left\{\frac{S_1^n}{n} - \frac{S_2^{N-n}}{N - n} < 0\right\}$ is the tail $1 - \Phi(x_0)$ of

a canonical normal distribution $\Phi(x)$ where

$$x = \frac{y - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{N-n}}}$$

and $-x_0$ is the value of x when $y = 0$. (I.e., $\Phi(y)$, which describes the distribution of $\frac{S_1^n}{n} - \frac{S_2^{N-n}}{N-n}$, is transformed to $\Phi(x)$ which describes the canonical normal distribution with mean 0 and variance 1.) The tail of a normal distribution is well approximated by

$$\Phi(-x) = 1 - \Phi(x) \lesssim \frac{1}{\sqrt{2\pi}} \cdot \frac{e^{-x^2/2}}{x}.$$

Thus

$$q' \lesssim \frac{1}{\sqrt{2\pi}} \cdot \frac{e^{-x_0^2/2}}{x_0} = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{N-n}}}{(\mu_1 \quad \mu_2)} \exp \frac{1}{2}\left[\frac{-(\mu_1 - \mu_2)^2}{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{N-n}}\right].$$

Using the same line of reasoning (but now with $(N - n)$ observations of ξ_1, etc.) we have

$$q'' = 1 - Pr\left\{\frac{S_1^{N-n}}{N-n} < \frac{S_2^n}{n}\right\}$$

$$\lesssim 1 - \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\frac{\sigma_1^2}{N-n} + \frac{\sigma_2^2}{n}}}{(\mu_1 - \mu_2)} \exp \frac{1}{2}\left[\frac{-(\mu_1 - \mu_2)^2}{\frac{\sigma_1^2}{N-n} + \frac{\sigma_2^2}{n}}\right].$$

From this we see that both q' and q'' are functions of the variances and means as well as the total number of trials, N, and the number of trials, n, given ξ'. More importantly, both q' and $1 - q''$ decrease exponentially with n, yielding

$$q(N-n, n) = q'p/(q'p + q''(1-p)) \sim q' \cdot (p/(1-p))$$

with the approximation being quite good even for relatively small n. For $p = \frac{1}{2}$ this reduces to

$$q(N-n, n) \sim q'$$

where the error is less than min $\{(q')^2, (1 - q'')^2\}$. (If one random variable is a priori more likely than the other to be best, i.e., if $p \neq \frac{1}{2}$, then we can see from

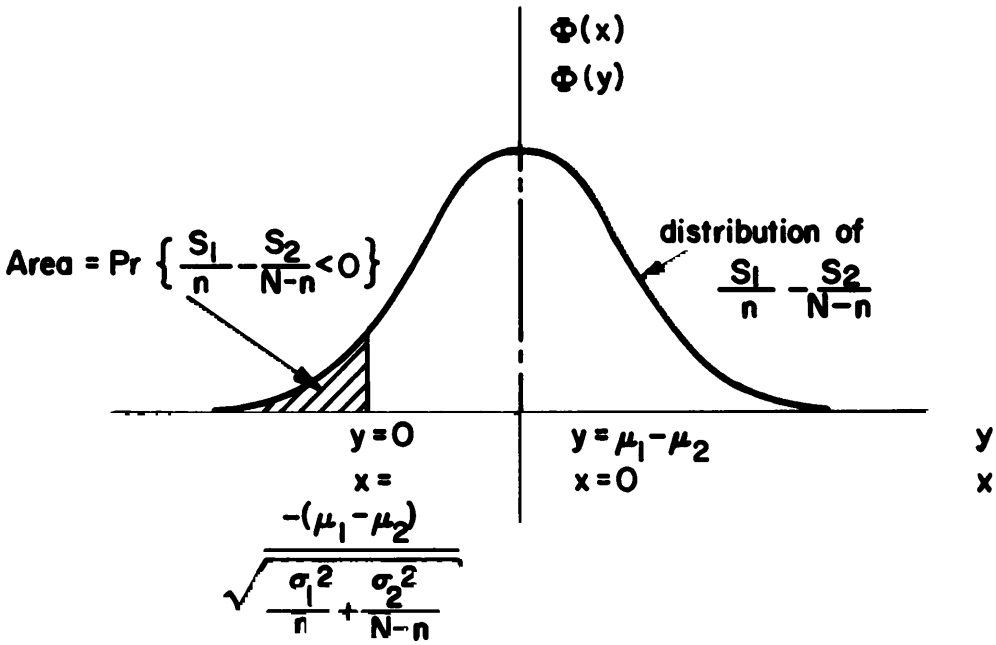

Fig. 11. The convolution of $\frac{S_1}{n}$ *with* $\frac{S_2}{N-n}$

above and from what follows that fewer trials can be allocated to attain the same reduction of $q(N-n, n)$. The expected loss is reduced accordingly.)

The observation that q' and hence $q(N-n, n)$ decreases exponentially with n makes it clear that, to minimize loss as N increases, the number of trials allocated the observed best, $N-n$, should be increased dramatically relative to n. This observation (which will be verified in detail shortly) enables us to simplify the expression for x_0. Whatever the value of σ_2, there will be an N_0 such that, for any $N > N_0$, $\sigma_2^2/(N-n) \ll \sigma_1^2/n$, for n close to its optimal value. (In most cases of interest this occurs even for small numbers of trials since, usually, σ_1 is at worst an order of magnitude or two larger than σ_2.) Using this we see that, for n close to its optimal value,

$$x_0 \lesssim \frac{(\mu_1-\mu_2)\sqrt{n}}{\sigma_1}, N > N_0.$$

We can now proceed to determine what value of n will minimize the loss $L(n)$ by taking the derivative of L with respect to n.

$$\frac{dL}{dn} = |\mu_1 - \mu_2| \cdot \left[-q + (N-n)\frac{dq}{dn} + 1 - q - n\frac{dq}{dn}\right]$$
$$= |\mu_1 - \mu_2| \cdot \left[(1-2q) + (N-2n)\frac{dq}{dn}\right]$$

where

$$\frac{dq}{dn} \lesssim \frac{1}{\sqrt{2\pi}}\left[-\frac{e^{-x_0^2/2}}{x_0^2} - e^{-x_0^2/2}\right]\frac{dx_0}{dn} = -\left[\frac{q}{x_0} + x_0 q\right]\frac{dx_0}{dn}$$

and

$\frac{dx_0}{dn} \lesssim \frac{\mu_1 - \mu_2}{2\sigma_1\sqrt{n}} = \frac{x_0}{2n}$. Thus $\frac{dL}{dn} \lesssim |\mu_1 - \mu_2| \cdot \left[(1-2q) - (N-2n)\cdot q \cdot \frac{(x_0^2+1)}{2n}\right]$.

n^*, the optimal value of n, satisfies $\frac{dL}{dn} = 0$, whence we obtain a bound on n^* as follows:

$$0 \lesssim (1-2q) - \left(\frac{N-2n^*}{2n^*}\right) \cdot q \cdot (x_0^2 + 1)$$

or

$$\left(\frac{N-2n^*}{2n^*}\right) \lesssim \frac{1-2q}{q\cdot(x_0^2+1)}$$

Noting that $1/(x_0^2+1) \lesssim 1/x_0^2$ and that $(1-2q)$ rapidly approaches 1 because q decreases exponentially with n, we see that $\frac{N-2n^*}{n^*} \lesssim \frac{2}{x_0^2 q}$, where the error rapidly approaches zero as N increases. Thus the observation of the preceding paragraph is verified, the ratio of trials of the observed best to trials of second-best growing exponentially.

Finally, to obtain n^* as an explicit function of N, q must be written in terms of n^*:

$$\frac{N-2n^*}{n^*} \lesssim \frac{2\sqrt{2\pi}\sigma_1}{(\mu_1 - \mu_2)} \cdot \frac{1}{\sqrt{n^*}} \cdot \exp\left[((\mu_1 - \mu_2)^2 n^*)/(2\sigma_1^2)\right].$$

Introducing $b = \sigma_1/(\mu_1 - \mu_2)$ and $N_1 = N - n^*$ for simplification, we obtain

$$N_1 \lesssim \sqrt{8\pi}\cdot b \cdot \exp\left[(b^{-2}n^* + \ln n^*)/2\right]$$

or

$$n^* + b^2 \ln n^* \gtrsim 2b^2 \cdot \ln \frac{N_1}{b\sqrt{8\pi}}$$

where the fact that $(N - 2n^*) \sim (N - n^*)$ has been used, with the inequality generally holding as soon as N_1 exceeds n^* by a small integer. We obtain a recursion for an ever better approximation to n^* as a function of N_1 by rewriting this as

$$n^* \gtrsim b^2 \ln \left[\frac{(b^{-1}N_1)^2}{8\pi n^*} \right].$$

Whence

$$\begin{aligned} n^* &\gtrsim b^2 \ln \left[\frac{(b^{-1}N_1)^2}{8\pi(b^2 \ln ((b^{-1}N_1)^2/8\pi n^*))} \right] \\ &\gtrsim b^2 \ln \left[\frac{b^{-4}N_1^2}{8\pi} \cdot \frac{1}{\ln ((b^{-1}N_1)^2/8\pi) - \ln n^*} \right] \\ &\gtrsim b^2 \ln \left[\frac{b^{-4}N_1^2}{8\pi(\ln N_1^2 - \ln (b^{-2}/8\pi))} \right] \\ &\gtrsim b^2 \ln \left[\frac{b^{-4}N_1^2}{8\pi \ln N_1^2} \right], \end{aligned}$$

where, again, the error rapidly approaches zero as N increases. Finally, where it is desirable to have n^* approximated by an explicit function of N, the steps here can be redone in terms of N/n^*, noting that N_1/n^* rapidly approaches N/n^* as N increases. Then

$$n^* \sim b^2 \ln \left[\frac{N^2}{8\pi b^4 \ln N^2} \right]$$

where, still, the error rapidly approaches zero as N increases.

The expected loss per trial $L^*(N)$ when n^* trials have been allocated to $\xi_{(2)}(\tau, N)$ is

$$\begin{aligned} L^*(N) &= \frac{1}{N} |\mu_1 - \mu_2| \cdot [(N - n^*)q(N - n, n^*) + n^*(1 - q(N - n^*, n^*))] \\ &= |\mu_1 - \mu_2| \cdot \left[\frac{N - 2n^*}{N} q(N - n^*, n^*) + \frac{n^*}{N} \right] \\ &\gtrsim |\mu_1 - \mu_2| \cdot \left[\frac{2n^*}{Nx_0^2} + \frac{n^*}{N} \right] \\ &\gtrsim \frac{b^2|\mu_1 - \mu_2|}{N} \cdot \left[2 + \ln \left(\frac{N^2}{8\pi b^4 \ln N^2} \right) \right]. \end{aligned}$$

Q.E.D.

The main point of this theorem quickly becomes apparent if we rearrange the results to give the number of trials $N^*_{(1)}$ allocated to $\xi_{(1)}$ as a function of the number of trials n^* allocated to $\xi_{(2)}$:

$$N^*_{(1)} = N - n^* \sim N \sim \sqrt{8\pi b^4 \ln N^2}\, e^{n^*/2b^2}.$$

Thus the loss rate will be optimally reduced if the number of trials allocated $\xi_{(1)}$ grows slightly faster than an exponential function of the number of trials allocated $\xi_{(2)}$. *This is true regardless of the form of the distributions defining* ξ_1 *and* ξ_2. Later we will see that the random variables defined by schemata are similarly treated by reproductive plans.

It should be emphasized that the above approximation for n^* will be misleading for small N when

(i) $\mu_1 - \mu_2$ is small enough that, for small N, the standard deviation of $\frac{S_1}{n} - \frac{S_2}{N-n}$ is large relative to $\mu_1 - \mu_2$ and, as a consequence, the approximation for the tail $1 - \Phi(x_0)$ fails,

or (ii) σ_2 is large relative to σ_1 so that, for small N, the approximation for x_0 is inadequate.

Neither of these cases is important for our objectives here. The first is unimportant because the cumulative losses will be small until N is large since the cost of trying ξ_2 is just $\mu_1 - \mu_2$. The second is unimportant because the uncertainty and therefore the expected loss depends primarily on ξ_1 until $N - n^*$ is large; hence the expected loss rate will be reduced near optimally as long as $N - n \cong N$ (i.e., most trials go to $\xi_{(1)}$), as will be the case if n is at least as small as the value given by the approximation for n^*.

Finally, to get some idea of n^* when σ_1 is not known, note that for bounded payoff, $\xi_i : \mathcal{A} \to [r_0, r_1]$, the maximum variance occurs when all payoff is concentrated at the extremes, i.e., $p(r_0) = p(r_1) = \frac{1}{2}$. Then

$$\sigma_1^2 \leqq \sigma^2_{\max} = \left(\frac{1}{2}r_1^2 + \frac{1}{2}r_0^2\right) - \left(\frac{1}{2}r_1 + \frac{1}{2}r_0\right)^2 = \left(\frac{r_1 - r_0}{2}\right)^2.$$

2. REALIZATION OF MINIMAL LOSSES

This section points out, and resolves, a difficulty in using $L^*(N)$ as a performance criterion. The difficulty occurs because, in a strict sense, the minimal expected loss rate just calculated cannot be obtained by any feasible plan for allocating trials in terms of observations. As such $L^*(N)$ constitutes an unattainable lower bound and,

if it is too far below what *can* be attained, it will not be a useful criterion. However, we will see here that such loss rates can be approached quite closely (arbitrarily closely as N increases) by feasible plans, thus verifying $L^*(N)$'s usefulness.

The source of the difficulty lies in the expression for n^*, which was obtained on the assumption that the n^* trials were allocated to $\xi_{(2)}(N)$. However there is *no* realizable plan (sequential algorithm) which can "foresee" in all cases which of the two random variables will be $\xi_{(2)}(N)$ at the end of N trials. No matter what the plan τ, there will be some observational sequences for which it allocates $n > n^*$ trials to a random variable ξ (on the assumption that ξ will be $\xi_{(1)}(N)$) only to have ξ turn out to be $\xi_{(2)}(N)$ after N trials. (For example, the observational sequence may be such that at the end of $2n^*$ trials τ has allocated n^* trials to each random variable. τ must then decide where to allocate the next trial even though each random variable has a positive probability of being $\xi_{(2)}(N)$.) For these sequences the loss rate will perforce exceed the optimum. Hence $L^*(N)$ is not attainable by any realizable τ—there will always be payoff sequences which lead τ to allocate too many trials to $\xi_{(2)}(N)$.

There is, however, a realizable plan $\tau_{(\sim)}$ for which the expected loss per trial $L(\tau_{(\sim)}, N)$ quickly approaches $L^*(N)$, i.e.,

$$\lim_{N\to\infty} (L(\tau_{(\sim)}, N)/L^*(N)) = 1.$$

$\tau_{(\sim)}$ proceeds by initially allocating n^* trials to each random variable (in any order) and then allocates the remaining $N - 2n^*$ trials to the random variable with the highest observed payoff rate at the end of the $2n^*$ trials. It is important to note that n^* is *not* recalculated for $\tau_{(\sim)}$; it is the value determined above.

COROLLARY 5.2: *Given N trials, $\tau_{(\sim)}$'s expected loss $L(\tau_{(\sim)}, N)$ approaches the optimum $L^*(N)$; i.e., $L(\tau_{(\sim)}, N) \sim L^*(N)$.*

Proof: The expected loss per trial $L(\tau_{(\sim)}, N)$ for $\tau_{(\sim)}$ is determined by applying the earlier discussion of sources of loss to the present case.

$$L(\tau_{(\sim)}, N) = \frac{1}{N} \cdot (\mu_1 - \mu_2) \cdot [(N - n^*)q(n^*, n^*) + n^*(1 - q(n^*, n^*))]$$

where q is the same function as before, but here the probability of error is irrevocably determined after $2n^*$ trials. That is

$$q(n^*, n^*) \sim \left(\sqrt{\frac{\sigma_1^2}{n^*} + \frac{\sigma_2^2}{n^*}} \Big/ \sqrt{2\pi}(\mu_1 - \mu_2)\right) \cdot \exp \frac{1}{2}\left[-(\mu_1 - \mu_2)^2 \Big/ \left(\frac{\sigma_1^2}{n^*} + \frac{\sigma_2^2}{n^*}\right)\right] \cdot$$

Rewriting $L(\tau_{(\sim)}, N)$ we have

$$L(\tau_{(\sim)}, N) = (\mu_1 - \mu_2)\left[\frac{N - 2n^*}{N} q(n^*, n^*) + \frac{n^*}{N}\right].$$

Since, asymptotically, q decreases as rapidly as N^{-1}, it is clear that the second term in the brackets will dominate as N grows. Inspecting the earlier expression for $L(N)$ we see the same holds there. Thus, since the second terms are identical

$$\lim_{N\to\infty} (L(\tau_{(\sim)}, N)/L^*(N)) = 1. \qquad \text{Q.E.D.}$$

From this we see that, given the requisite information (μ_1, σ_1) and (μ_2, σ_2), there exist plans which have loss rates closely approximating $L^*(N)$ as N increases.

3. MANY OPTIONS

The function $L^*(N)$ sets a very stringent criterion when there are *two* uncertain options, specifying a high goal which can only be approached where uncertainty is very limited. Adaptive plans, however, considered in terms of testing schemata, face many more than two uncertain options at any given time. Thus a general performance criterion for adaptive plans must treat loss rates for arbitrary numbers of options. Though the extension from two options to an arbitrary number of r options is conceptually straightforward, the actual derivation of $L^*(N)$ is considerably more intricate. The derivation proceeds by indexing the r random variables $\xi_1, \xi_2, \ldots, \xi_r$ so that the means are in decreasing order $\mu_1 > \mu_2 > \cdots > \mu_r$ (again, without the observer knowing that this ordering holds).

THEOREM 5.3: *Under the same conditions as for Theorem 5.1, but now with r random variables, the minimum expected loss after N trials must exceed*

$$(\mu_1 - \mu_2)\cdot(r - 1)b^2\left[2 + \ln\left(\frac{N^2}{8\pi(r-1)^2 b^4 \ln N^2}\right)\right]$$

where $b = \sigma_1/(\mu_1 - \mu_r)$.

Proof: Following Theorem 5.1 we are interested in the probability that the average of the observations of any ξ_i, $i > 1$, exceeds the average for ξ_1. This probability of error is accordingly

$$q(n_1, \ldots, n_r) \sim Pr\left\{\left(\frac{S_1}{n_2} < \frac{S_2}{n_1}\right) \text{ or } \left(\frac{S_1}{n_3} < \frac{S_3}{n_1}\right) \text{ or } \ldots \text{ or } \left(\frac{S_1}{n_r} < \frac{S_r}{n_1}\right)\right\},$$

where n_i is the number of trials given ξ_i, and the loss ranges from $(\mu_1 - \mu_2)$ to $(\mu_1 - \mu_r)$ depending on which ξ_i is mistakenly taken for best.

Let $n = \sum_{i=2}^{r} n_i$, let $m = \min \{n_2, n_3, \ldots, n_r\}$, and let j be the largest index of the random variables (if more than one) receiving m trials.

The proof of Theorem 5.1 shows that a lower bound on the expected loss is attained by minimizing with respect to any lower bound on the probability q (a point which will be verified in detail for r variables). In the present case q must exceed

$$Pr\left\{\frac{S_1}{m} < \frac{S_j}{N-n}\right\} > q' = \frac{1}{\sqrt{2\pi}} \frac{\sigma_1}{(\mu_1 - \mu_r)\sqrt{m}} \exp\left[\frac{(\mu_1 - \mu_r)^2 m}{2\sigma_1^2}\right]$$

using the fact that $(\mu_1 - \mu_r) \geqq (\mu_1 - \mu_j)$ for any $j > 1$. By the definition of q

$$L_{N,r}(n) > L'_{N,r}(n) = (\mu_1 - \mu_2)[(N-n)q + n(1-q)]$$

using the fact that $(\mu_1 - \mu_2) \leqq (\mu_1 - \mu_i)$ for $i \geqq 2$. Moreover the same value of n minimizes both $L_{N,r}(n)$ and $L'_{N,r}(n)$. To find this value of n, set

$$\frac{dL'_{N,r}}{dn} = (\mu_1 - \mu_2)\left[(N-2n)\frac{dq}{dn} + (1-2q)\right] = 0.$$

Solving this for n^* and noting that $1 - 2q$ rapidly approaches 1 as N increases, gives

$$n^* \sim \frac{N}{2} + \left(\frac{dq}{dn}\right)^{-1}.$$

Noting that q must decrease less rapidly than q' with increasing n, we have $(dq'/dn) < (dq/dn)$ and, taking into account the negative sign of the derivatives,

$$n^* > \frac{N}{2} + \left(\frac{dq'}{dn}\right)^{-1}.$$

(This verifies the observation at the outset, since the expected loss approaches n^* as N increases—see below.) Finally, noting that $n > (r-1)m$, we can proceed as in the two-variable case by using $(r-1)m$ in place of n and taking the derivative of q' with respect to m instead of n. The result is

$$n^* > (r-1)m^* \sim (r-1)b^2 \ln\left(\frac{N^2}{8\pi(r-1)^2 b^4 \ln N^2}\right)$$

where $b = \sigma_1/(\mu_1 - \mu_r)$. Accordingly,

$$L_{N,r}(n^*) > L'_{N,r}((r-1)m^*)$$
$$> (\mu_1 - \mu_2)\cdot(r-1)b^2\left[2 + \ln\left(\frac{N^2}{8\pi(r-1)^2b^4\ln N^2}\right)\right]. \quad \text{Q.E.D.}$$

4. APPLICATION TO SCHEMATA

We are ready now to apply the criterion just developed to the general problem of ranking schemata. The basic problem was rephrased as one of minimizing the performance losses inevitably coupled with any attempt to increase confidence in an *observed* ranking of schemata. The theorem just proved provides a guideline for solving this problem by indicating how trials should be allocated among the schemata of interest.

To see this note first that the central limit theorem, used at the heart of the proof of Theorem 5.3, applies to any sequence of independent random variables having means and variances. As such it applies to the observed average payoff $\hat{\mu}_\xi$ of a sequence of trials of the schema ξ under any probability distribution P over $\mathcal{A}$ (cf. chapter 4). It even applies when the distribution over $\mathcal{A}$ changes with time (a fact we will take advantage of with reproductive plans). In particular, then, Theorem 5.3 applies to any given set of r schemata. It indicates that under a good adaptive plan the number of trials of the (observed) best will increase exponentially relative to the total number of trials allocated to the remainder.

Near the end of chapter 4 it was proposed that the observed performance rankings of schemata be stored by selecting an appropriate (small) set of elements $\mathcal{B}$ from $\mathcal{A}$ so that the rank of each schema would be indicated by the relative number of instances of ξ in $\mathcal{B}$. Theorem 5.3 suggests an approach to developing $\mathcal{B}$, or rather a sequence $\mathcal{B}(1), \mathcal{B}(2), \ldots, \mathcal{B}(t)$, according to the sequence of observations of schemata. Let the number of instances of ξ in the set $\mathcal{B}(t)$ represent the number of observations of ξ at time t. Then the number of instances of ξ in the set $\bigcup_{t=1}^{T}\mathcal{B}(t)$ represents the total number of observations of ξ through time T. If schema ξ should persist as the observed best, Theorem 5.3 indicates that ξ's portion of $\bigcup_{t=1}^{T}\mathcal{B}(t)$ should increase exponentially with respect to the remainder. We can look at this in a more "instantaneous" sense. ξ's portion of $\mathcal{B}(t)$ corresponds to the rate at which ξ is being observed, i.e., to the "derivative" of the function giving ξ's increase. Since the derivative of an exponential is an exponential, it seems natural to have ξ's portion $M_\xi(t)$ of $\mathcal{B}(t)$ increase exponentially with t (at least until ξ

occupies most of $\mathcal{B}(t)$). This will be the case if ξ's rate of increase is proportional to the observed average payoff $\hat{\mu}_\xi(t)$ of instances of ξ at time t or, roughly,

$$dM_\xi(t)/dt = \hat{\mu}_\xi(t)M_\xi(t).$$

It will still be the case if the rate is proportional to the schema's "usefulness," the difference between $\hat{\mu}_\xi(t)$ and the overall average performance $\hat{\mu}(t)$ of instances in $\mathcal{B}(t)$, so that $dM_\xi(t)/dt = (\hat{\mu}_\xi(t) - \hat{\mu}(t))M_\xi(t)$. (In genetics $\hat{\mu}_\xi(t) - \hat{\mu}(t)$ is called the "average excess" of ξ when ξ is defined on a single locus, i.e., when ξ is a specific allele.)

The discussion of "intrinsic parallelism" in chapter 4 would imply here that *each* ξ represented in $\mathcal{B}(t)$ should increase (or decrease) at a rate proportional to *its* observed "usefulness" $\hat{\mu}_\xi(t) - \hat{\mu}(t)$. If this could be done consistently then each ξ would be automatically and properly ranked within $\mathcal{B}(t)$ as t increases. The reasoning behind this, as well as the proof that reproductive plans accomplish the task, will be developed in full in the next two chapters.