*Revista Colombiana de Matemáticas*
*Vol. XXI (1987), págs. 201-212*

# A THEOREM ON THE OPTIMAL ALLOCATION OF EFFORT*

by

Bernard GROFMAN and Guillermo OWEN

**Abstract.** A limited time budget is to be allocated to several tasks, so as to maximize the probability that a majority of these tasks will be performed correctly. It is shown that in the symmetric linear case, it is optimal to allocate time equally among k of the tasks, where k is at least a majority, but may be more, depending on the actual time available. In particular, time is allocated to all tasks if there is little time available, but to only a majority of the tasks if the available amount of time is reasonably large.

§1. **Introduction.** We consider here the following problem: a student, with a limited time budget, must study for an examination will consist of several questions, one from each of several fields. The student will be successful (pass the exam) if he answers a majority of the questions correctly. The problem is to decide how much time to spend on each of the several fields.(1)

---

* This research was supported by the National Science Foundation, Grant 85-03676.

(1) The problem we consider here is mathematically analogous to problem considered by the early French mathematician, Condorcet. For historical background and parallels, see Grofman, Owen and Feld, 1982, 1983.

Mathematically, we assume the subject is divided into n fields. For $i = 1,\dots,n$ we assume a function

$$p_i = f_i(x_i)$$

gives the probability that the question on the $i^{\underline{th}}$ field will be correctly answered if the student spends $x_i$ units of time on that field. For obvious reasons, we shall assume each $f_i$ is monotone non-decreasing and continuous, and bounded below by 0 and above by 1.

Let $N = \{1,2,\dots,n\}$ be the set of all questions. If the student has probability $p_i$ of answering question i correctly, and if all these probabilities are independet, then, for gieven $S \subset N$,

$$P_S(p_1,\dots,p_n) = \prod_{i\in S} p_i \prod_{i\notin S} (1-p_i) \tag{1}$$

is the probability that the student answer all the questions in fields $i \in S$, and none of the others, correctly.

Let m, now, be the required number of correct answers. If so, then the student's probability of passing the test is

$$F(p_1,\dots,p_n) = \sum_{\substack{S \\ s\geqslant m}} \prod_{i\in S} p_i \prod_{i\notin S} (1-p_i) \tag{2}$$

where the summation is taken over all sets S with at least m elements.

The student's problem, is, then, to maximize expression (2) subject to (1) and the budget constraint

$$\sum x_i \leqslant \alpha \tag{3}$$

$$x_i \geqslant 0, \qquad i = 1,\dots,n \tag{4}$$

where $\alpha$ is the student's available time. The first order conditions for this problem are

$$\frac{\partial F}{\partial p_i}\frac{dp_i}{dx_i} = \lambda \text{ if } x_i > 0 \tag{5}$$

$$\frac{\partial F}{\partial p_i}\frac{dp_i}{dx_i} \leqslant \lambda \text{ if } x_i = 0. \tag{6}$$

In the general case, of course, this presents a complicated computation. We will consider the special case where

$$p_i(x) = \begin{cases} x & 0 \leqslant x \leqslant 1 \\ 1 & x \geqslant 1 \end{cases}. \tag{7}$$

This is not an unreasonable probability function: it represents the case where the student requires unit time (the time can of course be suitable normalized) to read each section of the textbook. In less than unit time, he can only read a proportional fraction of the section, and the probability of a correct answer is in turn proportional to that. In this case, the first-order conditions take the form

$$\frac{\partial F}{\partial p_i} \begin{cases} =\lambda & \text{if } 0 < p_i < 1 \\ \geqslant\lambda & \text{if } p_i = 1 \\ \leqslant\lambda & \text{if } p_i = 0 \end{cases}.$$

Now, it can be seen that

$$\frac{\partial F}{\partial p_i} = \sum_{\substack{S \\ i\in S \\ s=m}} \prod_{\substack{j\in S \\ j\neq i}} p_j \prod_{j\notin S} (1-p_j) \tag{8}$$

where the sum is taken over all sets S, containing i and exactly m-1 other elements. We shall use $F_i$ to denote this partial derivative.

We prove, now, that we need only consider points $(p_1,\dots,p_n)$ in which each $p_i$ has one of the three values 0, 1, and some other p.

**LEMMA 1.** *The maximum of the function* F, *subject to the constraints* (3)-(4), *is attained at a point* $(p_1,\dots,p_n)$

*whose components have only one value other than 0 or 1;*

***Proof.*** Let us consider the expression (8) for $F_i$. Letting $\ell \neq i$, we can write this as

$$F_i = \sum_{\substack{S \\ \ell \in S}} \prod_{j \in s} p_j \prod_{\substack{j \in S \\ j \neq i}} (1-p_j) + \sum_{\substack{S \\ \ell \notin S}} \prod_{j \in S} p_j \prod_{\substack{j \notin S \\ j \neq i}} (1-p_j)$$

where the first sum is taken over all S with $\ell \in S$, $i \notin S$, $s = m-1$, and the second over all S with $i, \ell \notin S$, $s = m-1$. We rewrite as

$$F_i = p_\ell \left[ \sum_S \prod_{\substack{j \in s \\ j \neq \ell}} p_j \prod_{\substack{j \notin S \\ j \neq i}} (1-p_j) \right] + (1-p_\ell) \left[ \sum_S \prod_{j \in S} p_j \prod_{\substack{j \notin S \\ j \neq i, \ell}} (1-p_j) \right]$$

or equivalently,

$$F_i = p_\ell \sum_S \Pi_{p_j} \Pi (1-p_j) + (1-p_\ell) \sum_S \Pi_{p_j} \Pi (1-p_j) \qquad (9)$$

where the first sum is taken over all S with $i, \ell \notin S$, $s = m-2$, and the second over all S with $i, \ell \notin S$, $s = m-1$. In each case the first product is over all $j \in S$, the second over all $j \in N - S - \{i, \ell\}$.

We have, then,

$$F_i - F_\ell = (p_\ell - p_i) \left[ \sum_S \Pi p_j \, \Pi(1-p_j) - \sum_S \Pi p_j \, \Pi(1-p_j) \right]$$

where the two sums are as in (9), or equivalently,

$$F_i - F_\ell = (p_\ell - p_i) H_{i\ell} \qquad (10)$$

where

$$H_{i\ell} = \sum_{\substack{S \\ s=m-2}} \prod_{j \in S} p_j \prod_{\substack{j \notin S \\ j \neq i, \ell}} (1-p_j) - \sum_{\substack{S \\ s=m-1}} \prod_{j \in S} p_j \prod_{\substack{j \notin S \\ j \neq i, \ell}} (1-p_j) \qquad (11)$$

where the sums in (11) are over all subsets $S \subset N - \{i, \ell\}$ with m-2 and m-1 elements respectively. We note, *inter alia,* that $H_{i\ell}$ depends on $p_j$, $j \neq i,\ell$, but *does not* depend on $p_i$ or $p_\ell$.

Let X be the set of all $p = (p_i,\ldots,p_n)$ which maximize F subject to (3)-(4). By continuity of F, X will be compact and non-empty. Then C(X), the convex hull of X, is compact and convex; moreover, the extreme points of C(X) are all points of X (though not all points of X are necessarily extreme in C(X)). We claim, now, that if $p^* = (p_1^*,\ldots,p_n^*)$ is extreme in C(X), the components $p_j^*$ will have at most one value other than 0 or 1.

In fact, suppose there is some pair of indices $i,\ell$, such that

$$0 < p_i^* < p_\ell^* < 1.$$

Since $p^* \in X$, then by (7-ii), we have

$$F_i(p^*) = F_\ell(p^*)$$

Now, by (10)

$$F_i - F_\ell = (p_\ell^* - p_i^*)H_{i\ell}(p^*).$$

However, $p_i^* < p_\ell^*$, and so we must have $H_{i\ell} = 0$.

As was pointed out above, however, $H_{i\ell}$ is independent of both $p_i$ and $p_\ell$ thus, for any t, the point $p'(t)$, given by

$$p_i'(t) = p_i^* + t$$
$$p_\ell'(t) = p_\ell^* - t$$
$$p_j'(t) = p_j^* \quad \text{for all other } j$$

will also have $H_{i\ell}(p') = 0$. For sufficiently small t (both positive and negative) p'(t) will satisfy the constraints (3)-(4). Moreover, the directional derivative in the direction of increasing t is $F_i - F_\ell$, and this will be 0 for all values of t. Thus, for sufficiently small t,

$$F(p'(t)) = F(p'(-t)) = F(p^*).$$

Since $p^*$ maximizes F, so do $p'(t)$ and $p'(-t)$. But this means both $p'(t)$ and $p'(-t)$ belong to X, and, since

$$p^* = \frac{1}{2}(p'(t) + p'(-t))$$

we conclude that $p^*$ is not extreme in C(X). This contradiction proves the lemma.

We see, then, that the maximum of F will always be found at a point of the form

$$p_j = \begin{cases} 1 & j \in M_1 \\ p & j \in M_2 \\ 0 & j \in M_3 \end{cases} \tag{12}$$

where $M_1$, $M_2$, $M_3$ are disjoint sets whose union is N, with cardinalities $m_1$, $m_2$, and $m_3$, while $0 < p < 1$. We have then

$$m_1 + m_2 + m_3 = n. \tag{13}$$

$$m_1 + m_2 p = \alpha. \tag{14}$$

It is easy to see that, in this case, we will have

$$F = \sum_{s=m-m_1}^{m_2} \binom{m_2}{s} p^s (1-p)^{m_2-s} \tag{15}$$

In fact, all members of $M_1$ are always correct, and all members of $M_3$ are always wrong. Thus the student will pass the exam if and only if at least $m-m_1$ of the members of $M_2$ are answered correctly.

**LEMMA 2.** *If* $\alpha \geq m$, *then* F *is maximized by setting* $m_1 \geq m$. *If* $\alpha < m$, *then* F *is maximized by setting* $m_1 = 0$, *i.e.*, $M_1 = \emptyset$.

*Proof.* If $\alpha \geq m$, it is easy to see that F can be made equal to 1 simply by letting $m_1 \geq m$. This is clearly a maximum.

Suppose, in fact, that $\alpha < m$, but $M_1 \neq \emptyset$. Then $m_1 \leq \alpha < m$, so $m_2 > 0$ as otherwise we would have $F = 0$. Let $i \in M_1$, $\ell \in M_2$: then $p_i = 1$ and $0 < p_\ell < 1$, so assuming p is optimal, we must have

$$F_i \geqslant F_\ell .$$

Now, however,

$$F_i = \binom{m_2}{m-m_1} p^{m-m_1}(1-p)^{m_1+m_2-m}$$

$$F_\ell = \binom{m_2-1}{m-m_1-1} p^{m-m_1-1}(1-p)^{m_1+m_2-m}$$

(since as we saw before, $F_j$ is simply the probability that exactly m-1 answers other than j be correct). Thus we have

$$\binom{m_2}{m-m_1} p^{m-m_1}(1-p)^{m_1+m_2-m} \geqslant \binom{m_2-1}{m-m_1-1} p^{m-m_1-1}(1-p)^{m_1+m_2-m}$$

which reduces to

$$\frac{m_2 p}{m-m_1} \geqslant 1$$

or

$$m_2 p \geqslant m-m_1 .$$

By (14), however, this gives us $\alpha \geqslant m$ which is a contradiction. Thus, if $\alpha < m$, then at the optimum, $M_1 = \emptyset$ as claimed. Q.E.D.

From Lemma 2 we see, then, that in the "difficult" case, $\alpha < m$, we have $m_1 = 0$. Denote $M_2$ by K, then $M_3 = N-K$, and so the optium will be obtained at a point

$$p_j = \begin{cases} \frac{\alpha}{k} & \text{if} \quad j \in K \\ 0 & \text{if} \quad j \notin K \end{cases}$$

where K has k elements. In this case

$$F = \sum_{s=m}^{k} \binom{k}{s}\left(\frac{\alpha}{k}\right)^s \left(\frac{k-\alpha}{k}\right)^{k-s}$$

and we look for the value of k, $m \leqslant k \leqslant n$, which maximizes this expression:

$$F_{max} = \max_{m \leqslant k \leqslant n} \sum_{s=m}^{k} \binom{k}{s}\left(\frac{\alpha}{k}\right)^s \left(\frac{k-\alpha}{k}\right)^{k-s} \tag{16}$$

In general, we can obtain this number from tables of the cumulative normal distribution. To get an idea of its behavior, however, we let

$$q_k(s) = \binom{k}{s}\left(\frac{\alpha}{k}\right)^s \left(\frac{k-\alpha}{k}\right)^{k-s} \tag{17}$$

be the probability of exactly s correct answer, assuming that the student divided his time among k sections. Then

$$\frac{q_k(s)}{q_{k-1}(s)} = \frac{k}{k-s} \cdot \frac{(k-1)^{k-1}}{k^k} \cdot \frac{(k-\alpha)^{k-s}}{(k-\alpha-1)^{k-s-1}} \tag{18}$$

As $\alpha \to 0$, this expression approaches the limit

$$L_k(s) = \frac{k}{k-s}\left(\frac{k-1}{k}\right)^s . \tag{19}$$

Now, it is easy to see that, for $k > 0$, and $s > 1$,

$$1 - \frac{s}{k} < \left(1 - \frac{1}{k}\right)^s$$

and so $L_k(s) > 1$ for $s > 1$. We conclude that, for small values of $\alpha$, $q_k(s) > q_{k-1}(s)$ for all k and $s > 1$, and so k should be chosen as large as possible, i.e. $k = n$. On the other hand, if $\alpha$ is large, i.e., sufficiently close to m, we know it is best to choose $k = m$.

We conclude, then, that for small $\alpha$ the student should study some of each section; for large $\alpha$ (i.e., near m) he should concentrate his studying on m of the sections. What is not clear is (a) whether any intermediate values of k (i.e., $m < k < n$) are ever optimal.

To look at this problem in some detail, we consider the case $n = 13$, $m = 7$. Figure 1 shows the result of our calculations: $k = 13$ is optimal for all $\alpha < 6.16$, while $k = 7$ is optimal for $\alpha > 6.30$. In between there seem to be five small subintervals where $k = 12, 11, 10, 9, 8$ are successively optimal.

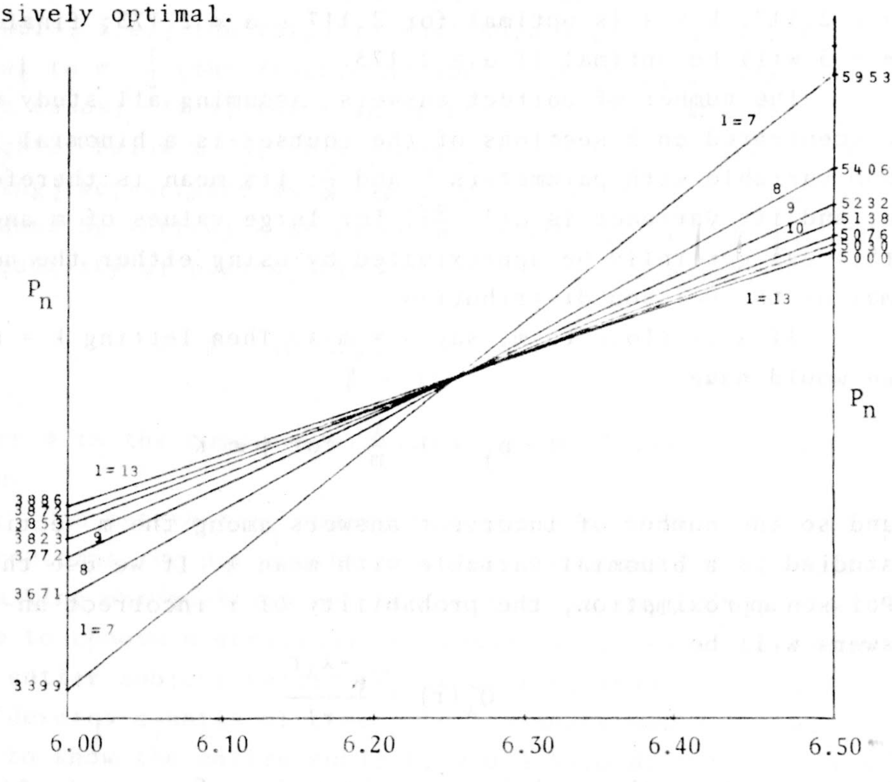


*Figure 1.* The Impact of Concentrating Competence on $P_n$, for n = 13.

It is not clear whether this type of behavior always holds, though in the several cases studied by the authors this is indeed the case. If we look at expression (18), we note that, as a function of $\alpha$, these ratios are convex, i.e.,

$$\frac{\partial^2}{\partial\alpha^2}\left(\frac{q_k(s)}{q_{k-1}(s)}\right) < 0.$$

This suggests (though it does not prove) that this type of behavior will usually hold.

For small values of m, it is not difficult to show that this is indeed the case. For example, in the case $n = 3$, $m = 2$, we find $k = 3$ is optimal for $\alpha \leqslant 1.125$, with $k = 2$ optimal if $\alpha \geqslant 1.125$.

For $n = 5$, $m = 3$, we find that $k = 5$ is optimal if $\alpha \leqslant 2.117$; $k = 4$ is optimal for $2.117 \leqslant \alpha \leqslant 2.173$; finally, $k = 3$ will be optimal if $\alpha \geqslant 2.173$.

The number of correct answers--assuming all study was concentrated on k sections of the course--is a binomial random variable with parameters k and $\frac{\alpha}{k}$; its mean is therefore $\alpha$, and its variance is $\alpha(1 - \frac{\alpha}{k})$. For large values of m and n, this can generally be approximated by using either the normal or the Poisson distribution.

If $\alpha$ is close to m, say $\alpha = m - \lambda$. Then letting $k = m$, we would have

$$1 - p_j = 1 - \frac{\lambda}{m} \quad \text{for } j \in K$$

and so the number of incorrect answers among the m sections studied is a binomial variable with mean $\lambda$. If we use the Poisson approximation, the probability of r incorrect answers will be

$$Q_\lambda(r) = \frac{e^{-\lambda}\lambda^r}{r!}$$

In particular, the probability of passing the exam is $Q(0)$, or $p^{-\lambda}$.

As against this, if the student studies $m + 1$ sections, the number of incorrect answers among the sections studied will also be approximately Poisson with mean $\lambda + 1$. To pass, at most one can be incorrect; the probability of passing is then

$$Q_{\lambda+1}(0) + Q_{\lambda+1}(1) = \ell^{-\lambda+1}(1+\lambda+1)$$

and this will be greater than $p^{-\lambda}$ only if $\lambda \geqslant \ell - 2$, i.e., if $\alpha \leqslant M+2-\ell$, or about $\alpha \leqslant M-0.718$. Thus $k = M$ is optimal if $\alpha \geqslant M-0.718$.

Suppose, on the other hand, $\alpha$ is considerably smaller than M. In this case concentration on k sections gives us a binomial variable which can best be approximated by a normal variable with mean $\alpha$ and variance $\alpha(1-\frac{\alpha}{k})$. To pass the examination, the student requires at least m correct answers, i.e., the variable must have a value at least equal to $m-\frac{1}{2}$ (the fractional modification is standard in such cases). If $\alpha$, the mean of the variable, is more than slightly below $m-\frac{1}{2}$, this probability will be maximized by making the variance as large as possible. With $\alpha$ fixed, this is done by setting k as large as possible, i.e., $k = n$. The probability of passing the exam will then be given by

$$P = \Phi\left(\frac{\alpha - m + \frac{1}{2}}{\sqrt{\alpha(1-\alpha/n)}}\right)$$

where $\Phi$ is the cumulative standard normal distribution function.

One interesting observation remains to be made, and it concerns the person who makes up the exam. If, instead of asking one question on each section of the course, he were to choose n questions at random (independently) from the entiee subject matter of the course, then the student who devotes $\alpha$ units of time (where n units would be required to know the entire subject) would have probability $\alpha/n$ on each question. In effect, this is the same as if the student had devoted $\alpha/n$ units to each of the n sections of the course. But we have seen that this is precisely the optimal study strategy for the student who spends a relatively small time preparing for this exam. Thus, such a strategy on the part of the examiner will penalize only the students who spend a relatively long time preparing, i.e., the conscientious students. In other words, the student who knows, e.g., 80% of the course material will get a grade of 80% if there is one question from each section, but might fail if the questions are chosen randomly from the entire course matter. The student who knows only 30% of the course matter has the same probability of passing under either model of examination.

## REFERENCES


Barlow, R.E. and Proschan, F., *Mathematical Theory of Reliability*, Wiley (New York) 1965.

Boland, P.J. and Proschan, F., *The Reliability of* k *out of* n *Systems*, **Annals of Probability**, Vol. 11 (1983) 760-764.

Feller, W., *An Introduction to Probability Theory and its Applications*, New York (Wiley) 1957.

Gleser, L., *On the Distribution of the Numer of Successes in Independent Trials*, **Annals of Probability**, 3 (1975) 182-188.

Grofman, B., Owen, G. and Feld, S., *Average Competence Variation in Competence, and the Behavior of a Group Majority*, **Psychological Reports**, 50 (1982) 683-688.

Grofman, B., Owen, G. and Feld, S., *Thirteen Theorem in Search of the Truth*, **Theory and Decision**, 1983, 261-278.

Marshall, A.W. and Olkin, I., *Inequalities: Theory of Majorization and Its Applications*, New York, (Academic Press), 1979.

Owen, G., *Multilinear Extension of Games*, **Management Science**, 1972, P64-P79.



*School of Social Sciences*
*University of California*
*Irvine, CA, USA.*

*Department of Mathematics*
*Naval Posgraduate School*
*Monterey, CA 93943, USA.*



(Recibido en noviembre de 1986).