Dynamic Programming and Decision Theory

D. V. Lindley

Applied Statistics, Volume 10, Issue 1 (Mar., 1961), 39-51.

DYNAMIC PROGRAMMING AND DECISION THEORY*

D. V. LINDLEY†

Statistical Laboratory, University of Cambridge

The previous article in this issue dealt with the general principles of Dynamic Programming. In this article Professor Lindley shows how Dynamic Programming links up with certain decision problems in the statistical field.

Introduction

The subject of this paper is the contribution of dynamic programming to some sequential statistical (or decision) problems. The first, and the most familiar, truly sequential procedure is Wald's[7] likelihood ratio test; this was originally designed to test one simple hypothesis against another, but is often used to deal with composite hypotheses. It is remarkable that although Wald's test is about twenty years old it is almost the only sequential procedure which is known to be *optimum* in some satisfactory sense (Wald and Wolfowitz[9]): for example, it is not known what is the optimum procedure for deciding between three (as distinct from two) simple hypotheses. This is not to say that these twenty years have not seen the introduction of useful sequential procedures, they have; my point is that these are not known to be best in any useful sense. In this paper I hope to show that dynamic programming provides a *computational* technique for finding the *optimum* sequential procedure in a wide variety of statistical situations. Examples of useful, if not necessarily optimum, procedures are the sequential *t*-test (see, for example, National Bureau of Standards[6]); applications for medical purposes devised by Armitage[1]; and the procedures for finding the best operating conditions, due to Box and Wilson.[3] In this paper I shall discuss, in particular, the problem of deciding whether the unknown mean of a normal distribution of known variance is positive or negative—a seemingly simple problem which is not completely solved. But before I do this it is necessary to make some remarks about decision problems, particularly in reference to Bayesian ideas.

Non-sequential Decision Problems

Suppose that there is some 'unknown' state of nature denoted by θ. Usually θ will consist of a single real parameter (such as a mean or variance) or a set of real parameters. Suppose further that it is possible to make an observation, resulting in the value x, and that the probability distribution of the observation depends on θ in a known way.

* Based on a paper read at a one-day conference on Dynamic Programming organised by the Birmingham Group of the Royal Statistical Society, May 1960.

† Present address: University College of Wales, Aberystwyth.

Denote the probability density of x, given θ, by $\mathrm{p}(x|\theta)$; a known function. So far this is the usual statistical situation. But now let $\pi(\theta)$ be a prior probability distribution for θ supposed to represent the statistician's beliefs about θ before making the observation. For example, if he were asked whether θ were positive or negative he would say that the odds were

$$\int_{\theta \geqslant 0} \pi(\theta)\mathrm{d}\theta \quad \text{to} \quad \int_{\theta < 0} \pi(\theta)\mathrm{d}\theta$$

that it was positive. It is not the place here to discuss the need for such a probability distribution over the states of nature; it is enough to say that its introduction is basic to dynamic programming ideas in sequential experimentation, and that it describes our knowledge of θ. No state of nature is truly 'unknown' in the sense used by most contemporary statisticians.

Now consider what happens to our knowledge of θ when the observation is made: $\pi(\theta)$ changes according to Bayes's theorem; the posterior distribution of θ, given x, is

$$\pi(\theta|x) = \mathrm{p}(x|\theta)\pi(\theta)/\mathrm{p}(x) \qquad \ldots . (1)$$

where

$$\mathrm{p}(x) = \int \mathrm{p}(x|\theta)\pi(\theta)\mathrm{d}\theta \qquad \ldots . (2)$$

That is, the posterior distribution of θ is proportional to the product of the known density of x and the prior distribution of θ. The constant of proportionality is the inverse of $\mathrm{p}(x)$, which is itself of interest, because it is the prior distribution of the observation before it is made. If the statistician were asked his views about what value the observation would assume, then he would express them through $\mathrm{p}(x)$. We can now see the role that observations play: they change the distributions over the states of nature (from prior to posterior) and, for example, it can be shown that, on the average, the distributions become, in an obvious intuitive sense, more concentrated; thus the distributions reflect an average gain in knowledge, for a concentrated distribution means that the value of θ is more accurately known than when the distribution has a large spread. The appreciation of this change in the distribution of θ with observations is important in understanding the sequential ideas below.

Finally suppose it is necessary to take one amongst a number of decisions on the basis of one's knowledge of the state of nature. If θ were known, presumably the best decision d would be known. In default of this knowledge it is necessary to introduce a utility function $\mathrm{U}(d, \theta)$ describing the utility of taking decision d when the true state of nature is θ. For a given θ the best decision is that which maximises the utility $\mathrm{U}(d, \theta)$. Now after the observation is made, when the knowledge of θ is given by $\pi(\theta|x)$, the best decision to take is that which maximises the *expected* utility, the expectation being taken over θ, for the true utility is unknown since it depends on θ. Hence take that d which maximises

$$E_\theta\{U(d, \theta)\} = \int U(d, \theta)\pi(\theta|x)d\theta = U^*(d, x) \quad \ldots . (3)$$

say. Had the observation not been available then $\pi(\theta)$ would have had to replace $\pi(\theta|x)$ and the expectation would have been

$$\int U(d, \theta)\pi(\theta)d\theta = U^*(d)$$

say. In either case we shall take the decision which maximises the utility: with an observation this will yield $\max_d\{U^*(d, x)\} = \bar{U}(x)$, say, whereas without it the yield will be $\max_d\{U^*(d)\} = U_0$, say. Now $\bar{U}(x)$ may exceed or be less than U_0, but on the average it exceeds U_0. By 'on the average' here we mean an average over the possible observations that might arise; some observations may reduce the utility but the average effect is an increase. Now before the observation is made our prior distribution for the observation is $p(x)$, equation (2), so that the average just referred to is

$$E_x\{\bar{U}(x)\} = \int \bar{U}(x)p(x)dx = \bar{U} \quad \ldots . (4)$$

say. It is easy to prove that $\bar{U} \geq U_0$ and the difference $\bar{U} - U_0$ gives the expected gain in utility provided by the observation. The observation is only worth making if this expected gain (which is necessarily non-negative) exceeds the cost (measured in the same units as utility) of the observation. An example may clarify the situation.

*Example of a Non-sequential Decision Problem**

Suppose that θ is the mean of a normal distribution of variance unity and suppose that x is the mean of a random sample of n observations from this distribution. (Any other function of the observations besides the mean is irrelevant; in technical language the mean is sufficient.) Then

$$p(x|\theta) = \left(\frac{n}{2\pi}\right)^{\frac{1}{2}} \exp\left[-\tfrac{1}{2}n(x-\theta)^2\right] \quad \ldots . (5)$$

Suppose that $\pi(\theta)$ is normal with mean θ_0 and variance τ^2, we write $N(\theta_0, \tau^2)$; and that there are two decisions d_1 and d_2 with

$$\left.\begin{aligned} U(d_1, \theta) &= c_1 \\ \text{and} \quad U(d_2, \theta) &= c_1 + b(\theta - \bar{\theta}) \end{aligned}\right\} \quad \ldots . (6)$$

where $b > 0$. Then if θ is known, d_2 is the preferred decision if $\theta > \bar{\theta}$, otherwise d_1 is preferred. In other words we have to decide whether θ is greater or less than $\bar{\theta}$, the utility of deciding that it is greater than $\bar{\theta}$ increasing linearly with θ.

* This example was discussed in an unpublished paper which I gave to the Loughborough conference of the Royal Statistical Society in September 1959.

The calculations in this problem are much simplified by noting the following fact: the posterior distribution of θ is still normal. This is a consequence of simple calculations carried out with Bayes's formula, equation (1). The mean and variance of the posterior distribution, for those who care to check the calculations, are respectively

$$\frac{xn + \theta_0/\tau^2}{n + 1/\tau^2} \quad \text{and} \quad [n + 1/\tau^2]^{-1} \qquad \ldots . (7)$$

Notice that the mean is a weighted average of x, the mean of the observations, and θ_0, the prior mean; and that the weights are, in the usual way, inversely proportional to the variances. In other words we can think of normal observations as changing the mean of the distribution of θ in a *random* way but changing the variance *deterministically*, retaining a normal distribution for θ.

From this fact it is easy to verify that d_1 is only taken when

$$x < \bar{\theta} + (\bar{\theta} - \theta_0)/n\tau^2 \qquad \ldots . (8)$$

If $\bar{\theta} = \theta_0$ this means that d_1 is preferred only when the observed mean is less than the critical value. But when $\bar{\theta} \neq \theta_0$ the additional term $(\bar{\theta} - \theta_0)/n\tau^2$ appears and allows for the prior belief, decreasing as n or τ^2 increase. It is also possible, using some rather more tedious calculations (given earlier by Grundy *et al.*[5]) to evaluate $\bar{U}$ [equation (4)] and hence the expected value of the observations, $\bar{U} - U_0$. If the cost per observation is c then n observations are worth making only if $\bar{U} - U_0 > cn$. The optimum number of observations is the value of n which maximises $\bar{U} - U_0 - cn$, the gain in utility expected from the experiment. In a sense such a procedure for deciding the number of observations is sequential, but it only becomes truly sequential when the observations are taken one at a time and at each stage the experimenter decides whether or not to take another one. To obtain the optimum procedure here involves new difficulties and we now consider these.

Sequential Decision Problems

Each stage of a truly sequential problem, with observations taken one at a time, is a decision problem of the type considered on p. 40, with the number of decisions increased by one by including the decision 'go on and take another observation'. The same remarks therefore apply to sequential problems as to non-sequential ones, with the addition that the utility of this extra decision has to be found. Now it is easy (by the methods sketched above) to find the expected utilities of stopping and taking any particular decision, but it is not easy to find the expected utility of taking one more observation. For the expected utility required is that of going on and then doing the best possible from then onwards. Consequently in order to find the best decision *now* (i.e. in particular whether to go on or not) it is necessary to know the best decision *in the future*. In other words the natural time order of working from the present to the future is not of any use because the present optimum involves the future optimum. The only method is to work backwards in time: from the optimum future behaviour to deduce the optimum present behaviour, and so on back into the past. It might be thought to be enough to consider the consequences of either stopping or taking just one further observation and then stopping; but this is not adequate because there are common situations in which, for example, two observations are worth having, but one is not. The *whole* of the future must be considered in deciding whether to go on.

For example in the problem of the normal mean on p. 41, where the observations are taken one at a time in order to decide whether θ is less than or greater than $\bar{\theta}$, it is intuitively clear that the optimum procedure will be a member of the following class: Let $\{a_n\}$, $\{b_n\}$ be two sequences of constants, with $a_n \leqslant b_n$, and let x_n be the mean of the first n observations; if $x_n < a_n$, take decision d_1; if $x_n > b_n$, take decision d_2; and if $a_n \leqslant x_n \leqslant b_n$ go on and take a further observation. It is impossible to determine, say, a_8 and b_8, until the two sequences are known for $n > 8$; because, until they are, the expected utility of going on when $n = 8$ cannot be evaluated and compared with the expected utilities of stopping and deciding finally whether θ is less than or greater than θ_0. Consequently we have to seek for an expression for a_8 and b_8 in terms of the sequences for $n > 8$. Even this is difficult, and it is here that dynamic programming ideas help. But before considering them let me repeat one remark made on p. 42: at any stage of the sequential procedure considered there the position may be summarised by two quantities, the mean and the variance of the distribution of θ at that stage. Furthermore the variance changes deterministically whilst the mean changes randomly according to a known probability law: thus one observation will change $N(\theta_0, \tau^2)$ into $N(\theta_1, [1+1/\tau^2]^{-1})$ where

$$\theta_1 = \frac{x + \theta_0/\tau^2}{1 + 1/\tau^2}$$

In practice it is easier to work with the inverse of the variance because it changes additively, namely from τ^{-2} to $1 + \tau^{-2}$. The distribution of θ_1 for given θ_0 is known since the distribution, $p(x)$, of x is known [equation (2)].

Dynamic Programming

In the previous article in this issue Dr Simpson described, in general terms, the basic important idea behind dynamic programming. In decision theory problems it is the idea of concentrating attention, not on the optimum procedure itself, but on the *utility* expected from the optimum procedure. Once the optimum utility has been found it is usually a simple problem to find that procedure which produces this utility. What we do in applying dynamic programming ideas to decision problems is to write down an equation for the optimum utility in the hope of solving it numerically or analytically.

The expected optimum utility at any stage is, by the discussion on pp. 40-41, a function of the distribution of θ at that stage, and, by the remark repeated at the end of the last section, this distribution depends, in many cases, on two parameters. Let $U(w, t)$ be the expected utility of the *best sequential* scheme when starting from a situation in which the parameters describing the distribution of θ are w and t; w changes randomly and t deterministically. In the normal mean situation w is the mean of the distribution of θ and t is the inverse of the variance. Let $\bar{U}(w, t)$ be the expected utility of taking the best decision now, without any further observations; $\bar{U}(w, t)$ is a known function. Then $U(w, t)$ is either $\bar{U}(w, t)$ (if it is not worth taking further observations) or can be evaluated in terms of the expected utility after taking a further observation (if an observation is taken). In the latter case w and t change to, say w' (randomly) and $t+h$ (deterministically), h is one in the example. Let $p(w'|w, t)$ be the distribution of w' given w and t: as explained above this is a known function. Lastly let $c(t, h)$ be the cost of increasing t to $t+h$. Then it is clear that if experimentation is continued

$$U(w, t) = \int U(w', t+h)p(w'|w, t)dw' - c(t, h) \qquad \ldots . (9)$$

Consequently the basic equation of dynamic programming in this sequential decision problem is

$$U(w, t) = \max\{\bar{U}(w, t) : \int U(w', t+h)p(w'|w, t)dw' - c(t, h)\} \ldots . (10)$$

in which everything is known except $U(w, t)$. If it can be solved for $U(w, t)$ then the optimum procedure is obvious: go on only when $U(w, t) > \bar{U}(w, t)$; otherwise stop and take the optimum decision obtained as in the non-sequential case. This illustrates the remark made earlier that once the optimum utility has been found it is easy to find the procedure which produces this utility.

The most important point to notice about equation (10) is that it is well suited for numerical work. If $U(w', t+h)$ is known for all w' and some single value of $t+h$ then, from (10), it is a simple matter of numerical integration and comparison with $\bar{U}(w, t)$ to calculate $U(w, t)$ for all w and the single value t. In the normal example, and typically, t corresponds to the sample size; hence one can work backwards in order of *decreasing* sample size as explained above. If the scheme is a truncated one, the point of truncation provides a convenient starting point. For example, in Armitage's medical work, it would be easy to work out for any particular case the best truncated scheme for deciding between three simple alternatives.

Satisfactory as the position is with regard to particular numerical cases, the theory and analytic solution of equation (10) are practically non-existent. I had hoped to be able to report the solution to the normal case but it has eluded me. [In the interval between delivering the lecture (May 1960) and writing it up for publication (September 1960) Chernoff[4] gave a paper in which he reached almost the same position as myself. In a personal communication he says that he hopes to perform the necessary numerical calculations.] Despite this lack of success it may be worth describing the position in that simple case. Equation (10) involves the variable w continuously whereas the variable t occurs discretely in steps of amount h. It would simplify the situation if the dependence on t could also be continuous (on the general grounds that differentials are easier to handle than differences). Now t is the inverse of the variance and so [from equation (7)] is equal to $n + \tau^{-2}$. Had each observation had variance σ^2 instead of unity, it would have been $n/\sigma^2 + \tau^{-2}$. Now allow $\sigma^2 \to \infty$ and $n \to \infty$ in such a way that n/σ^2 remains constant; then t can assume a continuous range of values. What we are doing here is to say that one observation of variance unity is equivalent to n observations each of variance n and supposing that we are taking a large (and ultimately infinite) number of observations each of large (and ultimately infinite) variance. The discontinuities caused by taking an observation are therefore removed. In this limit it is not difficult to show that the *integral* equation (9), which holds whenever it is worth taking another observation, is replaced by the familiar diffusion *differential* equation whose solution, in terms of the normal density function, is well known for given boundary conditions. Consequently in the region in the (w, t)-plane in which sampling continues the diffusion equation obtains for $U(w, t)$ and outside this region $U(w, t)$ is equal to the known function $\bar{U}(w, t)$. Furthermore the function $U(w, t)$ must be continuous on the boundary between these two regions and it is the boundary, not the function itself, that is of principal interest. Consequently we have a free-boundary problem in which the boundary is the main unknown. The natural way, therefore, to transform (10) is to derive from it an equation for the boundary. At first this does not seem difficult, but prolonged investigation has not, so far, enabled me to resolve the difficulties. I conjecture that the boundary [in the (w, t)-plane] satisfies a differential equation which is of the first order and second degree.

A general point I wish to emphasise here is that dynamic programming is, at its present stage of development, a technique which is suitable for numerical work, but that little is known about the solution of the equations, like (10), or about the properties of the solutions. It may sound old-fashioned in these days of fast machines to decry the solution of particular problems by their use; but it does seem to me that we should gain much more understanding of the problem by solving the equation analytically rather than numerically. It is true, in the normal case, that by transformations all the parameters can be removed so that a single boundary would suffice for all parameters; nevertheless a numerical solution of the differential equation—if we knew it—would be simpler than the iteration of (10) directly.

A word may be inserted here about the origin of equation (10). The phrase dynamic programming and the central ideas associated with it are obviously due to Bellman[2] but the approach had been used earlier in sequential decision problems by Wald.[8] Theorem 4.2 on page 105 of Wald's book is essentially our equation (10). I am indebted to Professor Barnard for this reference.

*Sufficiency**

In this section I show that the expected utility depends on two parameters, w and t, w changing randomly and t deterministically, in a wide class of situations. Suppose that an observation depends on a parameter θ consisting of s real numbers $(\theta_1, \theta_2, \ldots \theta_s)$, with a probability density of the following form

$$p(x|\theta) = F(x)G(\theta) \exp\left[\sum_{i=1}^{s} f_i(x)\theta_i\right] \qquad \ldots . (11)$$

i.e. a product of a function of the observation only, a function of the parameter only, and an exponential term linear in the components θ_i of the parameter. Notice that G is determined by F since the integral of (11) over x must be unity: conversely G determines F. If the prior distribution of θ is $\pi(\theta)$ the posterior distribution is proportional to

$$G(\theta)\pi(\theta) \exp\left[\sum_{i=1}^{s} f_i(x)\theta_i\right] \qquad \ldots . (12)$$

the constant of proportionality depending only on x, and being determined from the condition that the integral be unity. Now suppose n independent observations $(x_1, x_2, \ldots x_n)$ from (11) are taken. The joint density is the product of n terms like (11) and the posterior distribution $\pi(\theta|x_1, x_2, \ldots x_n)$ is proportional to

$$G(\theta)^n\pi(\theta) \exp\left[\sum_{j=1}^{n} \sum_{i=1}^{s} f_i(x_j)\theta_i\right] \qquad \ldots . (13)$$

Now this distribution of θ is characterised by $(s+1)$ parameters: the s coefficients of the θ_i, namely

$$w_i = \sum_{j=1}^{n} f_i(x_j)$$

and n, the sample size. Consequently at any stage in the sequential scheme the distribution of θ can be described in terms of $w = (w_1, w_2, \ldots w_s)$ and $t=n$. Hence the expected utility is also a function of w and t. Also w changes randomly with the addition of a further observation—in fact $w_i' = w_i + f_i(x_{n+1})$ so that it is an additive random walk—and t changes deterministically. The form (11) is that of the exponential family, the family of distributions which possess s sufficient statistics

$$\left[\sum_{j=1}^{n} f_i(x_j)\right]$$

for any size of sample. Most common distributions belong to this family: for example the normal, Poisson, binomial, and gamma distri-

* This section was not included in the conference paper, for lack of time, but is given here for completeness.

butions. (J. A. Lechner, in an unpublished Ph.D. thesis at Princeton University, has discussed the same example as ours with the substitution of the Poisson process for the normal one.)

The Marriage Problem

I conclude this paper with a very simple sequential decision problem for which the basic equation can be solved very easily. It should not be taken too seriously, and lady readers should interchange the sexes in the text. A known number, n, of ladies are presented to you one at a time in a random order. After inspecting any number r $(1 \leqslant r \leqslant n)$ of them you are able to rank them from best to worst and this order will not be changed if the $(r+1)$th lady is inspected; she will merely be inserted into the order. At any stage of the 'game' you may either propose to the lady *then being inspected* (there is no going back!), when the game stops, or inspect the next lady; however, if you reach the last lady you have to propose. All proposals are accepted. What is the optimum strategy? Before this has a definite answer we must assign a utility function: let U_i be the utility of being married to the lady with the ith true rank, so that $U_i \geq U_{i+1}$. True rank is the rank she has amongst all n ladies, as distinct from the apparent rank that she has when only some have been inspected. We can again argue in terms of two co-ordinates: r, the number of ladies inspected at any stage and s, the apparent rank of the rth lady being inspected. Now r changes deterministically in steps of one and s changes randomly; so r plays the role of t and s that of w in the above discussion.

In the notation used in our basic equation (10), t is the variable corresponding to the number of ladies inspected at any stage, here r; w is the variable corresponding to the rank of the lady now being inspected, here s. So $\mathrm{U}(w, t)$ is replaced by $\mathrm{U}(s, r)$ and $\bar{\mathrm{U}}(w, t)$ by $\bar{\mathrm{U}}(s, r)$. The probability that the rth lady of apparent rank s will have true rank S is easily calculated as

$$\binom{S-1}{s-1}\binom{n-S}{r-s}\Big/\binom{n}{r} = p_{S:s,r} \qquad \ldots . (14)$$

say, so that

$$\bar{\mathrm{U}}(w, t) = \bar{\mathrm{U}}(s, r) = \sum_{i=s}^{s+n-r} U_i p_{S:s,r} \qquad \ldots . (15)$$

Given the situation described by r and s, the probability that the next lady will have apparent rank s' is clearly $(r+1)^{-1}$ for all s', so that $\mathrm{p}(w'|w, t)$ in (10), now $\mathrm{p}(s'|s, r)$, is $(r+1)^{-1}$. Hence (10) is

$$\mathrm{U}(s, r) = \max\left\{\bar{\mathrm{U}}(s, r) : \sum_{i=1}^{r+1} \mathrm{U}(i, r+1)/(r+1)\right\} \qquad \ldots . (16)$$

with $\bar{\mathrm{U}}(s, r)$ a known function given by (15). There is no cost of inspection.

We consider two special cases. In the first $U_1=1$ and $U_i=0$ for $i>1$, corresponding to the attitude: nothing but the best. From (15) $\bar{U}(1, r)=r/n$ and $\bar{U}(s, r)=0$ for $s>1$. This last result means that it is no good proposing to anyone of apparent rank other than one, for to do so will result in zero utility. Consequently all we have to do is to find out when it pays to propose to a lady who has apparent rank 1. Equation (16) gives

$$U(1, r) = \max\left\{r/n: \sum_{i=1}^{r+1} U(i, r+1)/(r+1)\right\} \quad \ldots . (17)$$

and

$$U(s, r) = \sum_{i=1}^{r+1} U(i, r+1)/(r+1) \text{ for } s>1 \quad \ldots . (18)$$

Since s does not appear on the right-hand side of (18), $U(s, r)$ must, for $s>1$, be a function only of r, u_r, say. Equations (17) and (18) may then be written

$$U(1, r) = \max\{r/n, u_r\} \quad \ldots . (19)$$

and

$$u_r = \frac{1}{r+1}\{U(1, r+1)+ru_{r+1}\} \quad \ldots . (20)$$

respectively. Now suppose $U(1, r)>\bar{U}(1, r)=r/n$, that is the utility of continuing exceeds that of proposing. It follows from (19) that $U(1, r)=u_r$ and, from (20) with the value of r reduced by one, that $u_{r-1}=u_r$. Therefore since $u_r>r/n$, $u_{r-1}>(r-1)/n$ and from (19), again with the value of r reduced by one, $U(1, r-1)>(r-1)/n$. Hence if it is not worth proposing to a lady who is best out of r, it is not worth proposing to a lady who is best out of $(r-1)$, which result is intuitively obvious. Hence the best strategy must be to propose to a lady who is best out of r, provided r is large enough. We have only to find out how large r must be.

Suppose now that $U(1, r)=\bar{U}(1, r)=r/n$; that is, it is worth proposing to the best lady out of r, and therefore $U(1, r')=r'/n$ for all $r'\geq r$, by what we have just proved. Equation (20) gives

$$u_r = \frac{1}{r+1}\left\{\frac{r+1}{n}+ru_{r+1}\right\}$$

or, if $v_r=u_r/r$,

$$v_r = \frac{1}{nr}+v_{r+1}$$

and the same equation will obtain for all larger r. Adding both sides of all these equations we have

$$v_r = \frac{1}{n}\left\{\frac{1}{r}+\frac{1}{r+1}+\ldots+\frac{1}{n-1}\right\} \quad \ldots . (21)$$

But we know that u_r must be less than r/n since $U(1, r)=r/n$ [equation (19)], so that $v_r<n^{-1}$. Hence

$$\frac{1}{r}+\frac{1}{r+1}+\ldots+\frac{1}{n-1}<1 \qquad \ldots . (22)$$

The argument may be reversed and consequently it is only worth proposing to the apparently best lady out of r if (22) obtains. Let $r=R$ be the least such value; that is,

$$\frac{1}{R}+\frac{1}{R+1}+\ldots+\frac{1}{n-1}<1<\frac{1}{R-1}+\frac{1}{R}+\frac{1}{R+1}+\ldots+\frac{1}{n-1}$$

then from (20) with $r=R-1$

$$u_{R-1}=\frac{1}{R}\left\{\frac{R}{n}+(R-1)u_R\right\}$$

$$=\frac{R-1}{n}\left\{\frac{1}{R-1}+\frac{1}{R}+\ldots+\frac{1}{n-1}\right\} \ldots . (23)$$

from (21). Again from (20) with $r<(R-1)$ we have $u_r=u_{r+1}$, so that u_1, the utility of the best strategy starting at the beginning, is given by (23). If n is large the value of R is given by $\int_R^n dx/x=1$, if the series in (22) is replaced by an integral, and hence $n/R=\mathrm{e}$, the base of natural logarithms. Hence for large n the optimum rule is to inspect until a proportion e^{-1} $(=0{\cdot}368)$ of the ladies have been inspected and then propose to any subsequent lady of apparent rank one; the expected utility is, from (23), only e^{-1}. If, in real life, this process works between 18 and 40 (i.e. for 22 years) one should never propose until age $18+0{\cdot}368\times 22=26$, approximately. Either many people do not pursue an optimum strategy or else they have a different utility function.

To learn something of the effect of a change in utility we consider a second case in which $U_i=n-i$; that is, linear in the rank. It is a simple matter to calculate

$$\bar{\mathrm{U}}(s,r)=n-\left(\frac{n+1}{r+1}\right)s \qquad \ldots . (24)$$

and this value has to be inserted in (16). Now since $\bar{\mathrm{U}}(s,r)$ decreases as s increases and

$$\sum_{i=1}^{r+1}\mathrm{U}(i,r+1)/(r+1)$$

in (16) does not depend on s it is clear from (16) that if it is worth proposing to a lady of apparent rank s out of r, it is certainly worth proposing to one of apparent rank s' out of r, where $s'<s$; this is intuitively obvious. So the optimum strategy can be defined by a function $\mathrm{s}(r)$ with the rule: propose if and only if $s\leqslant \mathrm{s}(r)$; $\mathrm{s}(r)$ is the boundary in the sense used on p. 45. It follows from (16), again because

$$\sum_{i=1}^{r+1}\mathrm{U}(i,r+1)/(r+1)$$

does not depend on s, that $U(s, r)$, for $s > s(r)$, does not depend on s; so let us write it $n - c(r)$, say. From (16) we obtain the two equations

$$U(s, r) = \max\left\{n - \left(\frac{n+1}{r+1}\right)s,\ n - c(r)\right\} \qquad \ldots . (25)$$

and

$$n - c(r-1) = \sum_{i=1}^{r} U(i, r)/r \qquad \ldots . (26)$$

where, in (26), the value of r has been decreased by 1. Since one will propose to a lady of rank $s(r)$ but not of rank $s(r)+1$, (25) gives

$$\left(\frac{n+1}{r+1}\right)[s(r)+1] \geq c(r) \geq \left(\frac{n+1}{r+1}\right)s(r) \qquad \ldots . (27)$$

and (26) gives

$$c(r-1) = \left\{\left(\frac{n+1}{r+1}\right)[1+2+3+\ldots+s(r)] + [r - s(r)]c(r)\right\}\frac{1}{r} \qquad \ldots . (28)$$

These two results are the most convenient form for calculation purposes. But the calculations are rather tedious and, as emphasised on p. 45, it would be highly desirable to have a result for $s(r)$ alone. Although it is difficult to obtain a differential equation for the boundary in the normal case on p. 45, it is easy to do so here. Suppose n is large and let $x = r/n$; x is then the proportion of ladies so far inspected. Allow n to approach infinity but x remain fixed (so that also $r \to \infty$). Let $C(x) = c(r)$ and $S(x) = s(r)$. Then (27) is approximately

$$C(x) = S(x)/x \qquad \ldots . (29)$$

and (28) is approximately

$$c(r-1) - c(r) = \left\{\frac{1}{x}\cdot\frac{[S(x)]^2}{2} - S(x)C(x)\right\}\frac{1}{xn} \qquad \ldots . (30)$$

(In the last result we have used the fact that the sum of the first n natural numbers is $n(n+1)/2$). But $c(r-1) = C(x - n^{-1})$ so that, from (30),

$$n\{C(x - n^{-1}) - C(x)\} = -\tfrac{1}{2}[C(x)]^2$$

or, as $n \to \infty$,

$$dC/dx = \tfrac{1}{2}[C(x)]^2 \qquad \ldots . (31)$$

The boundary condition is $C(1) = c(n) = \frac{1}{2}n$, and therefore the solution to (31) is

$$C(x) = \{2n^{-1} + \tfrac{1}{2}(1-x)\}^{-1} \qquad \ldots . (32)$$

and therefore, from (29)

$$S(x) = x\{2n^{-1} + \tfrac{1}{2}(1-x)\}^{-1} \qquad \ldots . (33)$$

This approximation ignores the discrete form of $s(r)$, but numerical computations suggest that it is quite good. For example, with $n = 100$

and $r = 90$, so that $x = 0{\cdot}9$, $S(x) = 90/7$, or about 13. The exact value is 14, though the change in utility by using 13 instead of 14 is negligible. Notice that the approximation is only valid for fixed x as n gets large. No proposals should be made until $S(x) > 1$, that is until $x > 1/3 + 4/3n$. For large n this gives $x > 1/3$, agreeing very well with the strategy in the first case where the corresponding value was $0{\cdot}368$. $C(0) = 2n/(4+n)$ so that the expected utility at the start is $n - 2n/(4+n) = n(n+2)/(n+4)$. This is only accurate to order 1 and to this order is $(n-2)$, so that the expected utility is nearly equal to that of the best lady, namely $(n-1)$. This approximation is, however, suspect. One could also argue that since no proposals are to be made until $x > 1/3 + 4/3n$ the expected utility is constant until this inequality is first satisfied and therefore that $n - C(1/3 + 4/3n)$ is an equally good approximation to the expected utility at the start. To order 1 this gives $(n-3)$.* Mr B. N. Barnett has suggested alternative approximations to the behaviour of $C(x)$ for small x which appear to be better, but more work remains to be done before a completely satisfactory result can be obtained.

This problem also arises in the game of 'googol' mentioned in the *Scientific American* for February 1960. One player writes n different, but otherwise arbitrary, numbers on n pieces of paper and places them face downwards on a table. His opponent turns them up one at a time and can stop at any stage and say that the one just turned up is the biggest. If correct he wins, otherwise he loses. The first utility function is relevant.

* I am indebted to the Editor for this idea.

REFERENCES

[1] Armitage, P. (1960). *Sequential Medical Trials*. Blackwell, Oxford.

[2] Bellman, R. (1957). *Dynamic Programming*. Oxford University Press.

[3] Box, G. E. P. and Wilson, K. B. (1951). 'On the experimental attainment of optimum conditions', *J. R. Statist. Soc.*, B, **13**, 1.

[4] Chernoff, H. (1960). *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*. (To appear.)

[5] Grundy, P. M., Healy, M. J. R. and Rees, D. H. (1956). 'Economic choice of the amount of experimentation', *J. R. Statist. Soc.*, B, **18**, 32.

[6] National Bureau of Standards (1951). *Tables to Facilitate Sequential t-Tests*. Appl. Math. Series 7. Government Printing Office, Washington.

[7] Wald, A. (1947). *Sequential Analysis*. Wiley, New York.

[8] Wald, A. (1950). *Statistical Decision Functions*. Wiley, New York.

[9] Wald, A. and Wolfowitz, J. (1948). 'Optimum character of the sequential probability ratio test', *Ann. Math. Statist.*, **19**, 326.