20

Response Latencies and Probabilities

R. DUNCAN LUCE
University of Pennsylvania

Reprint from MATHEMATICAL METHODS IN THE SOCIAL SCIENCES, 1959
Edited by Kenneth J. Arrow, Samuel Karlin, and Patrick Suppes
Published by Stanford University Press, Stanford, California

1. Introduction

In that part of mathematical psychology and economics concerned with choice behavior, far more attention has been given to the probability patterns exhibited by the responses than to the other equally obvious response measure, the response latency. From time to time papers have appeared in which a general form for the latency distribution has been derived (e.g., [3, Section 14.4], [6], [7], [8], and [16, Appendix 2]); Audley [1] has attempted to deduce the two-alternative response probabilities from a latency model; and several experimentalists (e.g., [10], [12], [13], and [14]) have pointed out that mean reaction times are roughly linear with Shannon's information measure of the probability distribution of the stimulus presentation. Undoubtedly, there are several other relevant papers in the literature, but nothing like the number one might, *a priori*, expect.

Such limited and inconclusive theorizing hardly seems to do justice to a response variable that is, of necessity, omnipresent in our empirical studies and that most experimentalists are certain contains a good deal of information beyond that available from the response probabilities. For example, a part of the Skinnerian message is that time (in the form of response rates) is an important dependent variable that can be controlled to a marked degree by the schedule of reinforcement. So far, not one mathematical study has appeared that attempts to account for these data! One can hope that a better understanding of simple latency distributions may suggest both the latency parameters that are under control in the Skinnerian experiments and possible models for the nature of the control.

Such, however, is for the future: the aims of this paper are much more modest. Only those simple choice situations are considered in which a subject must select one of several responses. Distributions over the responses and latencies are assumed to exist, and a simple model is developed that

This work, which was carried out when I was a member of the Department of Social Relations, Harvard University, was supported in part by grant NSF G 5544 from the National Science Foundation. I am indebted to Mrs. Elizabeth Shipley and to Professor Robert R. Bush for their helpful comments.

leads to constraints on them. The general philosophy is similar to that of [16], in that we will worry about the inter-relations among the distributions when different, but related, sets of responses are available. Indeed, this study began as an attempt to generalize the basic assumption in [16]. Although that aim has not been achieved, an interesting restriction is developed and its relation to the principal assumption in [16] is explored. Next, it is shown that if in addition to our restriction and the assumption of [16] it is assumed that the mean latency of a response is a continuous function of the response probability, with the response set a parameter of the function, then that function must be logarithmic. Finally, the family of gamma distributions with a fixed decay parameter are shown to satisfy the basic restriction, provided that the other two parameters of the distributions exhibit a certain additive property. It is not known what other distributions, if any, meet our basic equation.

2. The Basic Equations

Let us consider a subject who, when a certain stimulus occurs, initiates a decision process that he ultimately terminates by choosing one response from a prescribed set of possible responses. We shall be concerned with both the response chosen and the elapsed time, or latency, of the response. Given that R is the set of available responses, let $P_R(x; t)$ denote the joint *response probability density* that response x is made t time-units after the activating stimulus occurs. We shall assume that these response densities are sufficiently well-behaved mathematical functions so that certain Laplace transforms, certain limits, and

$$P_R(x) = \int_0^\infty P_R(x; t)\, dt \tag{1}$$

exist. This quantity is simply interpreted as the probability that x is chosen when R is available; it will be referred to as a *response probability*. We require, of course, that

$$\sum_{x \in R} P_R(x) = 1 \ . \tag{2}$$

It is far from clear how human subjects, or other organisms, decompose and resolve such decisions into simpler ones, but one possibility is this: The subject surveys (a portion of) the set R and isolates some response, say y, as not worthy of further consideration, thereby reducing his choice problem to the set $R - \{y\}$. He repeats this elimination process until a two-element set is reached, at which point he makes the final, and only reported, decision by either rejecting the poorer or choosing the better response—it does not matter which, for they are equivalent. In this model we assume that the over-all decision latency is simply the algebraic sum of the several rejection latencies from the ever-decreasing sets of possible responses.

It is important that one misinterpretation not be made. It is not implied either that the subject views y as the least desirable response when he discards it from R or that its latency is nearly as long as would be required to determine which response is the worst. It is sufficient for the subject to discover some other response in R that he deems better than y in order to discard y. Thus, one possible procedure is for him to select two responses at random, compare them, and discard the poorer; but we shall not assume this mechanism.

However the discarding is done, we may postulate a joint *discard probability density* $Q_R(y;\tau)$ that the first response discarded from R is y and that this occurs τ time-units after stimulation. Of course, in the usual experiments this discard density Q is not an observable in the same sense that the response density P is. Rather, it is a theoretical construct that can be given meaning and that can sometimes be calculated in terms of a particular theory.[1] As for P, we define

$$Q_R(y) = \int_0^\infty Q_R(y;\tau)\,d\tau\,, \tag{3}$$

and assume that

$$\sum_{y \in R} Q_R(y) = 1\,. \tag{4}$$

$Q_R(y)$ will be referred to as a *discard probability*.

Now, if y is discarded from R at time τ, then the decision process begins anew at time τ with the reduced set $R - \{y\}$. If response x is chosen from $R - \{y\}$ in exactly $t - \tau$ units of time, then the over-all process has been a choice of x from R in t units of time. Integrating over all physically possible values of τ and summing over all possible first discards from R yields the following set of convolution equations:

$$P_R(x;t) = \sum_{y \in R-\{x\}} \int_0^t Q_R(y;\tau) P_{R-\{y\}}(x;t-\tau)\,d\tau\,. \tag{5}$$

These equations are basic to the rest of this paper.

For $\rho > 0$, the Laplace transform[2] of $P_R(x;t)$ is defined as

$$L(P_R, x;\rho) = \int_0^\infty e^{-\rho t} P_R(x;t)\,dt\,. \tag{6}$$

An analogous definition holds for $Q_R(y;\tau)$. By well-known properties of the transform, (5) can be converted into the algebraic equation

$$L(P_R, x;\rho) = \sum_{y \in R-\{x\}} L(Q_R, y;\rho) L(P_{R-\{y\}}, x;\rho)\,. \tag{7}$$

An important relation arises when $\rho \to 0$, provided that we assume the density functions are sufficiently well behaved so that the following limits exist:

---

[1] Possibly this statement is too strong for, as L. S. Shapley suggested to me in conversation, one could present the subject with a written list of alternatives and require that he cross out all but his choice. The order and times of crossing out the discarded alternatives could be observed and used to estimate Q.

[2] For an elementary discussion of Laplace transforms, see [9].

(8)
$$\begin{aligned}\lim_{\rho\to 0} L(P_R, x; \rho) &= \lim_{\rho\to 0}\int_0^\infty e^{-\rho t} P_R(x; t)\, dt \\ &= \int_0^\infty P_R(x; t)\, dt \\ &= P_R(x)\,,\end{aligned}$$

and

(9)
$$\lim_{\rho\to 0} L(Q_R, y; \rho) = Q_R(y)\,.$$

Assuming this, we find that (7) immediately implies

(10)
$$P_R(x) = \sum_{y\in R-\{x\}} Q_R(y) P_{R-\{y\}}(x)\,.$$

This equation is also important in the rest of the paper.

3. An Example

If, as will surely be the case in practice, the distributions P_R and $P_{R-\{y\}}$ are known or estimated and Q_R is unknown, then (10) and (for each ρ) (7) each consist of $r = |R|$ linear equations in their r unknowns $Q_R(x)$ and $L(Q_R, x; \rho)$, respectively. Except in those rare cases when the determinant of coefficients is 0, we may always solve for the unknowns; however, it is not certain in general that the solutions of (10) will form a probability distribution or that those of (7) will form Laplace transforms of density functions.

The following example illustrates the problem. Let $R = \{1, 2, 3\}$ and, for simplicity of notation, let

$$\boldsymbol{P} = [P_R(1), P_R(2), P_R(3)], \qquad \boldsymbol{Q} = [Q_R(1), Q_R(2), Q_R(3)]\,,$$
$$a = P(1, 2), \qquad b = P(1, 3), \qquad c = P(2, 3)\,,$$

where $P(i, j)$ stands for $P_{\{i,j\}}(i)$. Equation (10) is, therefore, the matrix equation $\boldsymbol{P} = M\boldsymbol{Q}$, where

$$M = \begin{bmatrix} 0 & b & a \\ c & 0 & 1-a \\ 1-c & 1-b & 0 \end{bmatrix}.$$

It is easily verified that M^{-1} exists if and only if

$$K = a(1-b)c + (1-a)b(1-c) \neq 0\,,$$

in which case it is given by

$$M^{-1} = \frac{1}{K}\begin{bmatrix} -(1-a)(1-b) & a(1-b) & (1-a)b \\ (1-a)(1-c) & -a(1-c) & ac \\ (1-b)c & b(1-c) & -bc \end{bmatrix}.$$

The numerical values $a = 3/4$, $b = 7/8$, $c = 2/3$, and $\boldsymbol{P} = (5/12, 4/12, 3/12)$ do not seem obviously implausible, but computing $\boldsymbol{Q} = M^{-1}\boldsymbol{P}$ we obtain $\boldsymbol{Q} = (21/39,\ 22/39,\ -4/39)$, which is not a probability distribution. Thus, either the chosen probabilities do not arise in practice or (10) does not make sense.

Although it seems of no particular practical import, it may be worth noting that if Q_R and $P_{R-\{y\}}$ are both distributions, then the function P_R defined by (10) is also a distribution. Since all of the right-hand terms in (10) are non-negative, it is sufficient to establish that the sum over R is 1. If we define $P_{R-\{x\}}(x) = 0$, then by (10)

$$\begin{aligned}\sum_{x\in R} P_R(x) &= \sum_{x\in R}\sum_{y\in R} Q_R(y)P_{R-\{y\}}(x)\\ &= \sum_{y\in R} Q_R(y)\sum_{x\in R} P_{R-\{y\}}(x)\\ &= \sum_{y\in R} Q_R(y)\sum_{x\in R-\{y\}} P_{R-\{y\}}(x)\\ &= 1\,.\end{aligned}$$

In the following sections we shall be concerned with sufficient conditions on the P's so that (10) can be solved for Q's that form probability distributions and so that (5) can be solved for probability densities.

4. A Solution to Equation (10)

In [16], I have investigated the following assumption about the choice probabilities:

Choice Axiom. *If $P(x, y) \neq 0, 1$ for $x, y \in R$, then for every $x \in S \subset R$,*

$$P_R(x) = P_S(x)\sum_{y\in S} P_R(y)\,.$$

Theorem 1. *If $P(x, y) \neq 0, 1$ for $x, y \in R$, and if the response probabilities satisfy the choice axiom, then $Q_R(x) = [1 - P_R(x)]/(r-1)$, where $r = |R|$, is the unique solution to* (10), *and it is a probability distribution.*

Proof. By the choice axiom,

$$\begin{aligned}P_R(x) &= P_{R-\{y\}}(x)\sum_{z\in R-\{y\}} P_R(z)\\ &= P_{R-\{y\}}(x)[1 - P_R(y)]\,.\end{aligned}\tag{11}$$

It is easy to show from $P(x, y) \neq 0, 1$ and the choice axiom that all of the probabilities in (11) are different from 0 and 1. Solving (11) for $P_{R-\{y\}}(x)$ and substituting it in (10), we obtain

$$P_R(x) = \sum_{y\in R-\{x\}} \frac{Q_R(y)P_R(x)}{1 - P_R(y)}\,.$$

Dividing by $P_R(x)$ and setting $\phi(y, R) = Q_R(y)/[1 - P_R(y)]$ yields

$$1 = \sum_{y\in R-\{x\}} \phi(y, R)\,.$$

By letting x assume two different values and subtracting the resulting equations, we see that $\phi(x, R) = \phi(y, R)$; hence, $\phi(x, R) = 1/(r-1)$. By substitution, it is easy to show that $Q_R(x) = [1 - P_R(x)]/(r-1)$ is, indeed, a solution when the choice axiom holds.

It remains to establish only that this solution is a probability distribution. Since $0 \leqq P_R(x) \leqq 1$, we have $Q_R(x) \geqq 0$, hence it is sufficient to show that the Q's sum to 1:

$$\sum_{x \in R} Q_R(x) = \sum_{x \in R} \frac{1 - P_R(x)}{r-1} = \frac{r-1}{r-1} = 1 .$$

Although this form for Q suggests some sort of a random process, the following calculation shows that it is not the one, mentioned briefly in Section 2, in which two responses are picked at random and the one deemed poorer discarded. In that case, the probability that x and y are drawn is $2/r(r-1)$ and, given that they are drawn, the probability that x is discarded is $P(y, x)$. The total probability, according to this model, that x is discarded is obtained by summing the product

$$Q_R^*(x) = \frac{2}{r(r-1)} \sum_{y \in R - \{x\}} P(y, x) .$$

Using the same notation as in Section 3, suppose that $a = 3/4$, $b = 9/10$, $c = 3/4$, and $\boldsymbol{P} = (9/13, 3/13, 1/13)$. As is easily verified, these values satisfy the choice axiom, but

$$\frac{2}{13} = Q_R(1) \neq Q_R^*(1) = \frac{7}{60} .$$

5. A Property of the Means and the Variances of the Latencies

The mean latency of response x, on the assumption that x occurs, is defined to be

$$\mu(P_R, x) = \int_0^\infty t \frac{P_R(x; t)}{P_R(x)} \, dt = \frac{1}{P_R(x)} \int_0^\infty t P_R(x; t) \, dt .$$

We observe that by differentiating (6) we have

$$\frac{\partial L(P_R, x; \rho)}{\partial \rho} = -\int_0^\infty t e^{-\rho t} P_R(x; t) \, dt .$$

Thus, if the limit on the right exists as $\rho \to 0$, as we shall assume, we obtain

$$\text{(12a)} \qquad \lim_{\rho \to 0} \frac{\partial L(P_R, x; \rho)}{\partial \rho} = -\mu(P_R, x) P_R(x) .$$

Similarly,

$$\text{(12b)} \qquad \lim_{\rho \to 0} \frac{\partial L(Q_R, y; \rho)}{\partial \rho} = -\mu(Q_R, y) Q_R(y) ,$$

again assuming that the limit exists. Now, if we differentiate (7) with respect to ρ and multiply through by -1, we obtain

$$-\frac{\partial L(P_R, x; \rho)}{\partial \rho} = -\sum_{y \in R-\{x\}} L(Q_R, y; \rho) \frac{\partial L(P_{R-\{y\}}, x; \rho)}{\partial \rho} - \sum_{y \in R-\{x\}} L(P_{R-\{y\}}, x; \rho) \frac{\partial L(Q_R, y; \rho)}{\partial \rho} .$$

Assuming that the limits in (8), (9), and (12) exist as $\rho \to 0$, we find that

$$\begin{aligned} \mu(P_R, x)P_R(x) &= \sum_{y \in R-\{x\}} [Q_R(y)\mu(P_{R-\{y\}}, x)P_{R-\{y\}}(x) + P_{R-\{y\}}(x)\mu(Q_R, y)Q_R(y)] \\ &= \sum_{y \in R-\{x\}} Q_R(y)P_{R-\{y\}}(x)[\mu(Q_R, y) + \mu(P_{R-\{y\}}, x)] . \end{aligned}$$

If we assume the choice axiom, Theorem 1 implies

$$Q_R(y)P_{R-\{y\}}(x) = \frac{1 - P_R(y)}{r - 1} P_{R-\{y\}}(x) = \frac{P_R(x)}{r-1} .$$

Substituting this into the preceding equation and dividing by $P_R(x)$, we see that we have proved

THEOREM 2. *If* (5) *is satisfied, if the limits in* (8), (9), *and* (12) *exist, and if the response probabilities satisfy the choice axiom, then*

$$\mu(P_R, x) = \frac{1}{r-1} \sum_{y \in R-\{x\}} [\mu(Q_R, y) + \mu(P_{R-\{y\}}, x)] .$$

If the variances $\sigma^2(P_R, x)$ and $\sigma^2(Q_R, y)$ are defined in the usual manner, then, assuming that the following limits exist, it is easy to see that

$$\lim_{\rho \to 0} \frac{\partial^2 L(P_R, x; \rho)}{\partial \rho^2} = \{\sigma^2(P_R, x) + [\mu(P_R, x)]^2\}P_R(x) \tag{13a}$$

and

$$\lim_{\rho \to 0} \frac{\partial^2 L(Q_R, y; \rho)}{\partial \rho^2} = \{\sigma^2(Q_R, y) + [\mu(Q_R, y)]^2\}Q_R(y) . \tag{13b}$$

A similar argument to that used to prove Theorem 2 leads to

THEOREM 3. *If* (5) *is satisfied, if the limits in* (8), (9), (12), *and* (13) *exist, and if the response probabilities satisfy the choice axiom, then*

$$\begin{aligned} \sigma^2(P_R, x) &= \frac{1}{r-1} \sum_{y \in R-\{x\}} [\sigma^2(Q_R, y) + \sigma^2(P_{R-\{y\}}, x)] \\ &\quad + \frac{1}{r-1} \sum_{y \in R-\{x\}} [\mu(Q_R, y) + \mu(P_{R-\{y\}}, x)]^2 - [\mu(P_R, x)]^2 . \end{aligned}$$

6. Mean Latency as a Function of Response Probability

In this section we show that if the assumptions made so far are satisfied and if, in addition, the mean latency is a function of the response set and the response probability, then that function is logarithmic.

THEOREM 4. *Suppose that families of distributions P and Q exist such that*
(i) *equation* (5) *is satisfied,*
(ii) *the limits in* (8), (9), *and* (12) *exist,*
(iii) *the response probabilities satisfy the choice axiom, and*
(iv) *the mean of $P_R(x;t)/P_R(x)$, namely $\mu(P_R, x)$, is a function of R and a continuous function of $P_R(x)$ only;*
then there are functions $A(R)$ and $B(R) > 0$ such that

$$\mu(P_R, x) = A(R) - B(R)\log P_R(x) .$$

PROOF. Denote the function mentioned in hypothesis (iv) as follows: $\mu(P_R, x) = f[P_R(x), R]$. According to the choice axiom and Theorem 1,

$$\begin{aligned} P_R(x) &= P_{R-\{y\}}(x) \sum_{z\in R-\{y\}} P_R(z) \\ &= P_{R-\{y\}}(x)[1 - P_R(y)] \\ &= P_{R-\{y\}}(x)Q_R(y)(r-1) . \end{aligned}$$

Thus, Theorem 2 can be rewritten as

$$f[P_{R-\{y\}}(x)Q_R(y)(r-1), R] = \frac{1}{r-1}\sum_{y\in R-\{x\}} \{\mu(Q_R, y) + f[P_{R-\{y\}}(x), R-\{y\}]\} .$$

Summing over $y \in R - \{x\}$ and rewriting yields

$$\sum_{y\in R-\{x\}} \{f[P_{R-\{y\}}(x)Q_R(y)(r-1), R] - \mu(Q_R, y) - f[P_{R-\{y\}}(x), R-\{y\}]\} = 0 . \tag{14}$$

Consider (14) in the special case in which $P_{R-\{y\}}(x) \to 1$ for all y; then

$$\sum_{y\in R-\{x\}} \mu(Q_R, y) = \sum_{y\in R-\{x\}} \{f[Q_R(y)(r-1), R] - f[1, R-\{y\}]\} . \tag{15}$$

If we set $P_y = P_{R-\{y\}}(x)$ and $Q_y = Q_R(y)(r-1)$ and substitute (15) in (14), then

$$\sum_{y\in R-\{x\}} [f(P_yQ_y, R) - f(Q_y, R) + f(1, R-\{y\}) - f(P_y, R-\{y\})] = 0 . \tag{16}$$

Define $F(P_y, R) = f(P_y, R) - f(1, R)$; then $F(1, R) = 0$ and (16) can be written

$$\sum_{y\in R-\{x\}} [F(P_yQ_y, R) - F(Q_y, R) - F(P_y, R-\{y\})] = 0 . \tag{17}$$

Again, we consider a special case, namely $Q_R(y) = 1/(r-1)$, for which (17) implies

$$\sum_{y\in R-\{x\}} F(P_y, R) = \sum_{y\in R-\{x\}} F(P_y, R-\{y\}) .$$

Substituting this in (17) yields

$$\sum_{y\in R-\{x\}} [F(P_yQ_y, R) - F(P_y, R) - F(Q_y, R)] = 0 . \tag{18}$$

Finally, we consider the special case in which $P_y = P$ and $Q_y = Q$ for all

$y \in R - \{x\}$, for which (18) reduces to

$$F(PQ, R) = F(P, R) + F(Q, R) . \tag{19}$$

Because F is continuous in the first variable for each fixed R and satisfies (19), there exists a function $B(R)$ such that $F(P, R) = -B(R) \log P$. Hence, by the definition of F and f,

$$\mu(P_R, x) = f[P_R(x), R] = A(R) - B(R) \log P_R(x) ,$$

where $A(R) = f(1, R)$.

Because $\mu(P_R, x) > 0$, it follows that $B(R) > 0$, which concludes the proof.

The conclusion of Theorem 4 is, of course, empirically testable. So also is the following weaker assertion about expected latencies over all the responses:

$$\begin{aligned} E(L) &= \sum_{x \in R} P_R(x) \mu(P_R, x) \\ &= A(R) - B(R) \sum_{x \in R} P_R(x) \log P_R(x) \\ &= A(R) + B(R) H(P_R) , \end{aligned}$$

where

$$H(P_R) = -\sum_{x \in R} P_R(x) \log P_R(x)$$

denotes Shannon's information measure (or uncertainty) of the *response* probability distribution.

Thus, the present theory leads to the prediction that the average latency over all responses will be a linear function of the uncertainty of the responses, with the two parameters dependent upon the response set. Considerable experimental evidence exists (see [10], [12], [13], and [14]) that shows the average latency to be an approximately linear function of the uncertainty of the stimulus presentation, but so far as I know, no one has made the corresponding comparison with the response uncertainty. But because there is a fairly general tendency for subjects to respond in such a way that the response probabilities roughly match the stimulus probabilities, it is unlikely that the two information measures will be very different. So one may conclude that the existing information-theory evidence tends to support our conclusions.

Against the conclusion of Theorem 4 are two sets of data with which I am familiar. The more recent, Berlyne's [2], does not, it seems to me, really apply, because two somewhat different experimental procedures were compared. One anticipates that both $A(R)$ and $B(R)$ will depend upon the experimental details, and so an adequate test of our conclusions can be had only by holding these details fixed. Kellogg's [15] earlier work on response latencies when the subjects were making a brightness discrimination between two patches of light appears to be suited to test our results. A constant stimulus and six variable stimuli were employed as in a standard discrimination experiment, and so far as the subjects knew their judgments were the only data collected; actually, reaction times were also observed.

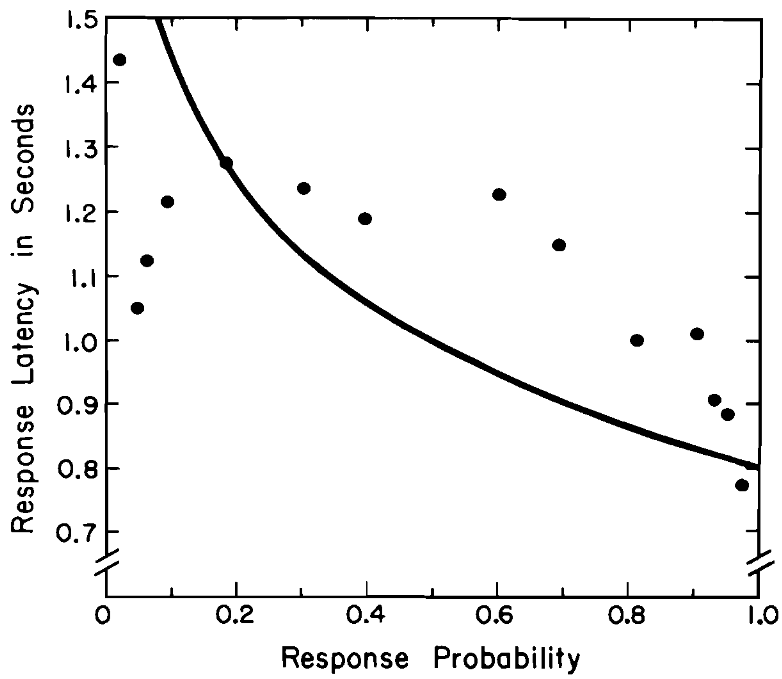

Figure 1. Mean response latency vs. mean response probability. The data points are adapted from [15] and are the average of five subjects. Each point represents $240p$ observations, where p is the response probability. The smooth curve is the function $0.80 - 0.28 \log_e p$.

Not enough raw data is presented in Kellogg's paper so that the mean response latency can be plotted as a function of the response frequency for each of the five subjects separately; however, the average function for the five subjects can be obtained. These data points are shown in Figure 1. Each point is based upon $240p$ observations, where p is the choice frequency. Thus, for $p < 0.1$ the mean latencies are probably not very reliable, and so I would not take very seriously the four points at the left of the plot. Nonetheless, it is clear that for $p > 0.1$ the data points trace out a convex function, whereas Theorem 4 predicts a concave function. The theoretical function drawn was chosen to be approximately correct for $p = 0.2$ and $p = 1.0$; it is correct nowhere else.

Although Kellogg's data cast considerable doubt upon the validity of Theorem 4 for psychophysical discrimination,[3] unfortunately they do not

[3] In conversation, Professor F. W. Irwin has questioned this conclusion. He feels that Kellogg's experiment is no more suitable for testing this theory than Berlyne's because, from our point of view, it is seven distinct experiments—one for each of the seven variable stimuli. The only fully satisfactory test would be an experiment having a fixed set of, say, five or six responses always available, but with the response probabilities as far from equal as possible.

allow us to decide just where the trouble lies. Additional studies are needed to ascertain whether (5), the choice axiom, or condition (iv) of the theorem is at fault.

7. A Set of Choice-Latency Distributions Satisfying Equation (5)

The so-called gamma, or Pearson Type III, distributions (including the simple exponential) have probably been more often suggested than any others as the appropriate form for latency distributions ([3], [7], [8], [16]). In this section I shall show that they satisfy (5), provided that the parameters are properly restricted. We define the distributions as follows:

$$(20)\quad P_R(x;t)=\begin{cases}0 & t<t_0(x,R)\,,\\ \dfrac{P_R(x)\lambda}{\Gamma[a(x,R)]}\{\lambda[t-t_0(x,R)]\}^{a(x,R)-1}e^{-\lambda[t-t_0(x\;R)]} & t\geqq t_0(x,R)\,,\end{cases}$$

and

$$(21)\quad Q_R(y;\tau)=\begin{cases}0 & \tau<\tau_0(y,R)\,,\\ \dfrac{Q_R(y)\lambda}{\Gamma[\alpha(y,R)]}\{\lambda[\tau-\tau_0(y,R)]\}^{\alpha(y,R)-1}e^{-\lambda[\tau-\tau_0(y,R)]} & \tau\geqq\tau_0(y,R)\,,\end{cases}$$

where Γ denotes the gamma function. The parameters $a(x, R)$ and $t_0(x, R)$ may depend upon $P_R(x)$ as well as on x and R, and $\alpha(y, R)$ and $\tau_0(y, R)$ may depend upon $Q_R(y)$. Observe that the notation is chosen to be consistent with (1) and (3).

Issues of parameter estimation are discussed in [3], [11], and [17].

THEOREM 5. *If the distributions defined by* (20) *and* (21) *are such that*
(i) P_R *and* Q_R *satisfy* (10),
(ii) $a(x, R) = \alpha(y, R) + a(x, R - \{y\})$ *for all* $x, y \in R$ $(x \neq y)$,
(iii) $t_0(x, R) = \tau_0(y, R) + t_0(x, R - \{y\})$ *for all* $x, y \in R$ $(x \neq y)$,
then they satisfy (5).

PROOF. Define

$$P_R^*(x;t)=P_R(x;t)e^{\lambda[t-t_0(x,R)]},\qquad Q_R^*(y;\tau)=Q_R(y;\tau)\,e^{\lambda[\tau-\tau_0(y,R)]}\,.$$

Then we show, using hypothesis (iii), that P_R and Q_R satisfy (5), provided that P_R^* and Q_R^* do:

$$\begin{aligned}\sum_{y\in R-\{x\}}&\int_0^t Q_R(y;\tau)P_{R-\{y\}}(x;t-\tau)\,d\tau\\ &=\sum_{y\in R-\{x\}}\int_0^t Q_R^*(y;\tau)e^{-\lambda[\tau-\tau_0(y,R)]}P_{R-\{y\}}^*(x;t-\tau)e^{-\lambda[t-\tau-t_0(x,R-\{y\})]}\,d\tau\\ &=\sum_{y\in R-\{x\}}\int_0^t Q_R^*(y;\tau)P_{R-\{y\}}^*(x;t-\tau)\,d\tau\;e^{-\lambda[t-\tau_0(y;R)-t_0(x,R-\{y\})]}\\ &=P_R^*(x;t)\,e^{-\lambda[t-t_0(x,R)]}\\ &=P_R(x;t)\,.\end{aligned}$$

It is well known that

$$L(P_R^*,x;\rho)=P_R(x)\,e^{t_0(x,R)(\lambda-\rho)}\left(\frac{\lambda}{\rho}\right)^{a(x,R)}$$

and

$$L(Q_R^*, y; \rho) = Q_R(y)\, e^{\tau_0(y,R)(\lambda-\rho)}\left(\frac{\lambda}{\rho}\right)^{\alpha(y,R)}.$$

From this and our hypotheses we obtain

$$\begin{aligned}
&\sum_{y\in R-\{x\}} L(Q_R^*, y; \rho) L(P_{R-\{y\}}^*, x; \rho) \\
&\quad= \sum_{y\in R-\{x\}}{}' Q_R(y) P_{R-\{y\}}(x)\, e^{[\tau_0(y,R)+t_0(x,R-\{y\})](\lambda-\rho)}\left(\frac{\lambda}{\rho}\right)^{[\alpha(y,R)+a(x,R-\{y\})]} \\
&\quad= e^{t_0(x,R)(\lambda-\rho)}\left(\frac{\lambda}{\rho}\right)^{a(x,R)} \sum_{y\in R-\{x\}} Q_R(y) P_{R-\{y\}}(x) \\
&\quad= P_R(x)\, e^{t_0(x,R)(\lambda-\rho)}\left(\frac{\lambda}{\rho}\right)^{a(x,R)} \\
&\quad= L(P_R^*, x; \rho).
\end{aligned}$$

So the transforms satisfy (7), which is the transform of (5).

This set of solutions has the intuitively desirable feature that all of the distributions are of the same mathematical form. One might suppose that if decision latency is a basic psychological concept, then all choices, including acts of discarding, should result in latency distributions of the same form. Now it is clear that if a family of distributions is chosen for the P's, then (5) will not in general lead to the same form for the Q's, assuming that it leads to a distribution at all. An interesting problem is to formulate rigorously the requirement that the distributions all have the same form and then to determine all sets of possible solutions to (5) that meet the requirement; presumably, there are very few.

8. A Comment on Audley's Work

In [16] it was shown that the choice axiom is equivalent to the assumption that there exists a positive ratio scale v over the responses such that for all $S \subset R$,

$$P_S(x) = \frac{v(x)}{\sum_{y\in S} v(y)}.$$

Several of us who have been working on that theory have wondered if there might not be some comparatively simple relation between the v-scale and some of the parameters of the latency distribution. An idea for making such a connection is contained in [1], but it appears to be inconsistent with the present notions.

Audley [1] develops a stochastic two-choice learning model, in part, by assuming latency distributions of the form given in (20) with $a(1, R) = a(2, R) = 1$ and $t_0(1, R) = t_0(2, R) = 0$, where $R = \{1, 2\}$. Then, in essence, he defines $v(1) = \lambda P_R(1)$ and $v(2) = \lambda P_R(2)$, from which it follows that $\lambda = v(1) + v(2)$ and $P_R(i) = v(i)/[v(1) + v(2)]$. This can be generalized to arbitrary R, provided that the response probabilities in (20) satisfy the choice axiom (no assumptions need be made about a and t_0), simply by setting $v(x) = \lambda P_R(x)$ and noting that

$$\sum_{x \in R} \lambda P_R(x) = \lambda = \sum_{x \in R} v(x) .$$

This, however, does not seem satisfactory for two reasons. First, it makes λ a function of, at least, R, and so the distributions will no longer satisfy (5). Second, it means that the exponential decay part of the distribution becomes stronger the larger the set of responses, and this seems dubious. Neither of these facts is, however, at all conclusive, and it may be worthwhile to pursue this tack further; I will not do so here.

9. Summary

Assuming that an organism makes its response from a set of possible responses by discarding, one by one, those it considers undesirable, and assuming that the total response latency is the simple sum of the discard latencies, we developed the following relation [equation (5)] between the response densities $P_R(x; t)$ and the postulated discard densities $Q_R(y; \tau)$:

$$P_R(x; t) = \sum_{y \in R - \{x\}} \int_0^t Q_R(y; \tau) P_{R-\{y\}}(x; t - \tau) \, d\tau .$$

From this equation, a similar relation was derived [equation (10)] for the response and discard probabilities:

$$P_R(x) = \sum_{y \in R - \{x\}} Q_R(y) P_{R-\{y\}}(x) .$$

Theorem 1 established that the second equation has the simple solution $Q_R(y) = [1 - P_R(y)]/(r - 1)$, provided that the response probabilities satisfy the choice axiom discussed in [16]. In addition, it was shown in Theorem 2 that the mean latencies of specific responses and discards must exhibit a simple additive property when the response probabilities satisfy the choice axiom.

If one adds to the conditions needed in Theorem 2 the assumption that the mean latency is a continuous function of the response probability, with the response set a parameter of the function, then it was shown in Theorem 4 that $\mu(P_R, x) = A(R) - B(R) \log P_R(x)$, where $A(R)$ and $B(R) > 0$ are functions of the response set R. The average latency over all responses is, therefore, $A(R) + B(R)H(P_R)$, where $H(P_R)$ is Shannon's information measure of the *response* probability distribution. In most situations, this prediction is similar to the empirical generalization that the average latency is linear with the information measure of the *stimulus* probability distribution. However, the more precise prediction that the mean latency of a response is linear with the logarithm of the response probability was shown to be inconsistent with existing data [15]. What assumptions are at fault is not known.

Returning to the first basic equation, we showed in Theorem 5 that the familiar gamma distributions [equations (20) and (21)] satisfy (5), provided that the response and discard probabilities satisfy the second equation above [equation (10)], that the exponential decay parameter is the same in all the distributions, and that the other two sets of distribution parameters each meet a simple additivity requirement. It is not known what, if any, other families of distribution satisfy (5).

References

[1] AUDLEY, R. J. "A Stochastic Description of the Learning Behaviour of an Individual Subject," *Quarterly Journal of Experimental Psychology*, **9** (1957), 12–20.

[2] BERLYNE, D. E. "Conflict and Choice Time," *The British Journal of Psychology*, **48** (1957), 106–18.

[3] BUSH, R. R., and F. MOSTELLER. *Stochastic Models for Learning*, New York: John Wiley and Sons, 1955.

[4] CARTWRIGHT, D. "The Relation of Decision-Time to the Categories of Response," *American Journal of Psychology*, **54** (1941), 174–96.

[5] CARTWRIGHT, D. "Decision-Time in Relation to the Differentiation of the Phenomenal Field," *Psychological Review*, **48** (1941), 425–42.

[6] CARTWRIGHT, D., and L. FESTINGER. "A Quantitative Theory of Decision," *Psychological Review*, **50** (1943), 595–621.

[7] CHRISTIE, L. S. "The Measurement of Discriminative Behavior," *Psychological Review*, **59** (1952), 443–52.

[8] CHRISTIE, L. S., and R. D. LUCE. "Decision Structure and Time Relations in Simple Choice Behavior," *Bulletin of Mathematical Biophysics*, **18** (1956), 89–112.

[9] CHURCHILL, R. V. *Modern Operational Mathematics in Engineering*, New York: McGraw-Hill, 1944.

[10] CROSSMAN, E. R. F. W. "Entropy and Choice Time: The Effect of Frequency Unbalance on Choice-Response," *Quarterly Journal of Experimental Psychology*, **5** (1953), 41–52.

[11] FESTINGER, L. "A Statistical Test for Means of Samples from Skew Populations," *Psychometrika*, **8** (1943), 205–10.

[12] HICK, W. E. "On the Rate of Gain of Information," *Quarterly Journal of Experimental Psychology*, **4** (1952), 11–26.

[13] HICK, W. E. "The Impact of Information Theory on Psychology," *The Advancement of Science*, **40** (1954), 397–402.

[14] HYMAN, R. "Stimulus Information as a Determinant of Reaction-Time," *Journal of Experimental Psychology*, **45** (1953), 188–96.

[15] KELLOGG, W. M. "The Time of Judgment in Psychometric Measures," *American Journal of Psychology*, **43** (1931), 65–86.

[16] LUCE, R. D. *Individual Choice Behavior*, New York: John Wiley and Sons, 1959.

[17] PERRY, N. C. "Note Concerning Some Empirical Results on Sampling from Skewed Populations," *Psychological Reports*, **5** (1959), 161.