A learning model for signal detection theory - temporal invariance of learning parameters

MICHAEL BIDERMAN

University of Tennessee, Chattanooga, Tennessee 37401

and

DONALD D. DORFMAN and JOHN C. SIMPSON

University of Iowa, Iowa City, Iowa 52242

Temporal invariance of learning parameters of an additive-operator model was evaluated on signal-detection data collected by Dorfman and Biderman (1971). Their study was a two-interval forced-choice auditory detection task in which the presentation probability of the signal in Interval 1 was manipulated. Parameters of the model were estimated separately in the nonstationary and stationary phases of the learning process. The results showed quite respectable temporal invariance of the learning parameters and of sensitivity. Pure error correction was observed in both phases, which agreed with general conclusions of Dorfman, Saslow, and Simpson (1975).

Dorfman and Biderman (1971) and Dorfman, Saslow, and Simpson (1975) have investigated the feasibility of an additive-operator learning process for the decision criterion of signal detection theory. The learning axiom they studied derives from Kac (1962, 1969), who proposed the special case of error correction. In particular, assume a sensory random variable, x, two stimuli, S_i (i = 1, 2), two responses, R_j (j = 1, 2), and knowledge of results, E_k (k = 1, 2). Let X_n be the decision cutoff on Trial n so that if

$$\text{(i) } x < X_n \text{ response is } R_2$$

$$\text{(ii) } x > X_n \text{ response is } R_1.$$

If $S_i R_j E_k$ occurs on Trial n, then the cutoff changes according to,

$$X_{n+1} = X_n + \delta_k \Delta_{jk},$$

where

$$\delta_k = \begin{cases} -1 \text{ if k = 1} \\ +1 \text{ if k = 2.} \end{cases}$$

In Dorfman and Biderman (1971), correct feedback was given, so that $S_i R_j E_k$ was denoted as $S_i R_j$, and to each $S_i R_j$ there was a corresponding learning parameter, Δ_{ij}. This notation is used here.

The purpose of this investigation was to evaluate the temporal invariance of the learning parameters on the data of Dorfman and Biderman (1971). Recall that their study was a two-interval forced-choice simple

auditory-detection task with noise continuously present during the experimental session. The probability that the signal appeared in Interval 1 was manipulated with $P(S_1)$ = 0.1, 0.5, 0.9. There was a different $P(S_1)$ on each test session, and each test session had 400 trials given without interruption. Eighteen students served as subjects.

DATA ANALYSIS

Let X_{1m} denote the initial cutoff for the mth schedule of a particular block of trials, where m = 1, 2, 3 for $P(S_1)$ = 0.1, 0.5, and 0.9, respectively. Furthermore, let $f(x \mid S_1)$ be N(a/2, 1) and $f(x \mid S_2)$ be N(−a/2, 1).

The analyses of Dorfman et al. (1975) suggest that Trials 1-200 are predominantly the nonstationary phase of the process and Trials 201-400 are essentially the asymptotic or stationary phase. Hence, the vector of parameters $(X_{11}, X_{12}, X_{13}, a, \Delta_{11}, \Delta_{12}, \Delta_{21}, \Delta_{22})$ was estimated from the data separately for these two trial blocks. The method of maximum likelihood was used, and all solutions were strict absolute maximum points (see Dorfman, 1973).

RESULTS AND CONCLUSIONS

Table 1 gives the mean maximum likelihood estimates (MLEs) and their standard errors for the two trial blocks. The results on Trials 201-400 have been reported elsewhere (Dorfman et al., 1975), whereas the results on Trials 1-200 are new. Note that for Trials 201-400, X_{1m} is $X_{201,m}$.

Table 2 presents the results of two ANOVAs, one performed on the MLEs of the learning parameters for successes $(\Delta_{11}, \Delta_{22})$ and the other on the MLEs for errors $(\Delta_{12}, \Delta_{21})$. First, we see that the learning process is essentially pure error-correction over both trial blocks

Table 1
Mean Parameter Estimates and Standard Errors for Trial Block 1-200 and Trial Block 201-400*

	Mean Parameter Estimates							
	a	Δ_{11}	Δ_{12}	Δ_{21}	Δ_{22}	X_{11}	X_{12}	X_{13}
Trials 1-200	1.673 (.089)	−.002 (.002)	.146 (.021)	.118 (.021)	.001 (.003)	−.090 (.124)	.013 (.101)	.068 (.092)
Trials 201-400	1.610 (.082)	−.003 (.003)	.192 (.023)	.184 (.026)	−.000 (.003)	.840 (.144)	.135 (.129)	−.900 (.166)

*Standard errors are in parentheses.

(Table 1), which agrees with the conclusions of Dorfman et al. (1975). Table 2 shows only one significant source of variance. The mean of the learning-parameter estimates for errors was smaller for Trials 1-200 than for Trials 201-400. Hence, there is no evidence that $\Delta_{ij} \to 0$ as $n \to \infty$, again in accord with findings of Dorfman et al. (1975). A reasonable explanation for the larger values of $(\Delta_{12}, \Delta_{21})$ for Trials 201-400 is based upon the bias of the small-sample ML estimator. Monte Carlo simulations

Table 2
Results of Analyses of Variance of MLEs of Learning Parameters

Source	Successes p	Errors p
Trial Block (T)	.496	.006*
SR Event (E)	.262	.402
T by E	.871	.642

*Statistically significant.

Table 3
Mean Parameter Estimates and Standard Errors for Trial Blocks 1-200 and 201-400 of Monte Carlo Simulation 0*

	Mean Parameter Estimates							
	a	Δ_{11}	Δ_{12}	Δ_{21}	Δ_{22}	X_{11}	X_{12}	X_{13}
Trials 1-200	1.69 (.035)	−.004 (.002)	.184 (.012)	.183 (.013)	−.004 (.001)	−.156 (.071)	−.057 (.069)	.268 (.096)
Trials 201-400	1.595 (.038)	.001 (.001)	.223 (.020)	.226 (.022)	−.001 (.002)	.821 (.095)	−.113 (.138)	−.698 (.118)

*Standard errors are in parentheses.

have shown that for $(\Delta_{12}, \Delta_{21})$, $\hat{\theta}_n = \theta_0 + \epsilon_n$, where θ_0 is the population parameter, $\epsilon_n > 0$ and $\epsilon_n \to 0$ in probability as $n \to \infty$. The value of ϵ_n decreases with the number of errors, and there were fewer errors on Trials 201-400.

To evaluate this argument, the data of Monte Carlo simulation 0 from Dorfman et al. (1975) were used. MLEs were obtained separately for Trials 1-200 and Trials 201-400. The results are shown in Table 3. The population parameter was $\tilde{X} = (0, 0, 0)$; a = 1.6; $\Delta_{11} = \Delta_{22} = 0$; $\Delta_{12} = \Delta_{21} = .09$. The simulation data support the hypothesis.

Table 1 also reveals that sensitivity (a) was larger for Trials 1-200 than for Trials 201-400. Nevertheless, the magnitude of the difference is quite small. In particular, consider the standard condition in which the cutpoint is at the solution to $f(x \mid S_1) = f(x \mid S_2)$. Percent of correct responses will be 79.8 for a = 1.673, and 79.0 for a = 1.610 at $P(S_1) = 0.5$. Note that $F(1.673/2) \cong 0.798$ and $F(1.610/2) \cong 0.790$.

The results on the initial cutoffs are also of interest. For Trials 1-200, the initial cutoffs are independent of the asymptotic expected cutoffs. This suggests that the model gives a good account of the learning process for short trial lengths, since the estimated values of the initial cutoffs are very sensitive to false assumptions.

An alternative additive learning model worthy of

study is Thomas' (1973) recent model. His model assumes a probabilistic rather than a deterministic shift in X_n, given $R_{jn}E_{kn}$. Unfortunately, the learning parameters of the model are not currently estimable because the learning states lack identifiability. In conclusion, the additive learning model of Dorfman and Biderman (1971) gives a very respectable temporal parameter invariance.

REFERENCES

DORFMAN, D. D. The likelihood function of additive learning models: Sufficient conditions for strict log-concavity and uniqueness of maximum. *Journal of Mathematical Psychology*, 1973, 10, 73-85.

DORFMAN, D. D., & BIDERMAN, M. A learning model for a continuum of sensory states. *Journal of Mathematical Psychology*, 1971, 8, 264-284.

DORFMAN, D. D., SASLOW, C. F., & SIMPSON, J. C. Learning models for a continuum of sensory states re-examined. *Journal of Mathematical Psychology*, 1975, 12, 178-211.

KAC, M. A note on learning signal detection. *I.R.E. Transactions on Information Theory*, 1962, IT-8,126-128.

KAC, M. Some mathematical models in science. *Science*, 1969, 166, 695-699.

THOMAS, E. A. C. On a class of additive learning models: Error-correcting and probability matching. *Journal of Mathematical Psychology*, 1973, 10, 241-264.