Maximum Entropy and the Lottery

HAL STERN and THOMAS M. COVER*

The distribution on m-tuples of the first M integers is estimated from marginals of the distribution. This problem is of interest in determining unpopular numbers in lotto games. In Canada's Lotto 6/49 the proportion of tickets purchased in previous games containing each number is available. Under certain conditions the limiting distribution subject to the observed marginals is the constrained maximum entropy distribution. This distribution is estimated, and Monte Carlo methods are used to estimate the expected return of various lottery strategies. Tickets consisting of unpopular numbers may have expected return greater than their cost when the weekly sales are large or there are large carryover prizes (prizes not won in earlier games). The maximum entropy distribution is a rough approximation of the true distribution of tickets purchased. Certain aspects of the empirical distribution are not consistent with the maximum entropy distribution. Alternative methods, which attempt to model the behavior of ticket buyers, are considered.

KEY WORDS: Distributions on m-tuples; Generalized iterative scaling; Lotto games.

1. INTRODUCTION

Each year hundreds of millions of dollars are spent on tickets in government lotteries characterized by an extremely low probability of winning a multimillion dollar jackpot. The most common lottery game is lotto, in which participants choose m distinct numbers from among the first M integers. The jackpot prize is awarded if the m numbers chosen match the m winning numbers; smaller prizes are given for fewer matches. Much evidence exists (Ziemba, Brumelle, Gautier, and Schwartz 1986) that the only way to improve the expected return from playing lotto is to select numbers that are unpopular with the other bettors. This strategy takes advantage of the pari-mutuel prize payouts.

Here we consider several estimates of the distribution of tickets purchased by the betting public. Careful study of this distribution enables us to estimate the distribution of prizes won when a particular m-tuple is chosen. The analysis here is concentrated on the Canadian lotto game, with $m = 6$ numbers chosen from $M = 49$ choices. For this game the marginal probability of a number being selected as part of a sixtuple is available. That is, after each game the proportion of tickets purchased that contain each integer is known. Without such information little can be done to estimate the unknown distribution. Under certain conditions the distribution of tickets purchased is the maximum entropy distribution subject to the given marginal constraints. The maximum entropy distribution is used to provide an idea of the expected return from various lottery strategies. The data, however, do not seem to support the necessary assumptions, so alternatives to the maximum entropy distribution are also considered.

2. LOTTO GAMES AND TESTS OF RANDOMNESS

Players select 6 numbers from the first 49 integers on a \$1 ticket to participate in Canada's Lotto 6/49. No number may be selected more than once. Six winning numbers and a bonus number are selected at random by the Lottery Commission. If a player matches fewer than three of the winning numbers then no prize is won. Three winning numbers entitles the player to a \$10 prize. If four or more numbers match, then the player wins an amount determined as follows. First, 55% of the lottery sales is used to pay lottery costs and to provide revenue for the government. Then all of the \$10 prizes are awarded. The remaining money in the prize pool is split among four groups of prizewinners. Twenty-five percent of the pool is split among tickets with four winning numbers, 13% is split among tickets with five winning numbers that do not have the bonus number as their sixth number, 17% among tickets that include five winning numbers and the bonus number, and 45% among tickets with six winning numbers (the jackpot). Any prizes that are not won (this is usually limited to the jackpot) are added to the jackpot for the following game. In Canada the jackpot is guaranteed to be at least one million dollars.

Based on the rules described previously the obvious strategy is to try to figure out which numbers will be winning numbers. Sadly, much evidence indicates that the winning numbers are selected at random. Figure 1 shows the number of times that each number was a winning number in 161 games through July 6, 1985. They are well scattered around the mean $(6/49) \times 161 = 19.7$. The chi-squared statistic comparing observed and expected counts is 54.12 on 48 df. In this case the usual chi-squared statistic is multiplied by 48/43; this is an adjustment for sampling without replacement. There is thus no evidence that some numbers tend to be good numbers. A variety of randomness tests have been applied (Ziemba et al. 1986) and, typically, no pattern is found. These tests include examination of whether there are such things as "hot" numbers (those that have won often recently). Lottery organizers in Canada and the United States expend considerable effort to ensure that lotto games are unpredictable.

It has long been observed that players in a random lottery game with pari-mutuel prizes win large prizes if they

* Hal Stern is Assistant Professor, Department of Statistics, Harvard University, Cambridge, MA 02138. Thomas M. Cover is Professor, Departments of Electrical Engineering and Statistics, Stanford University, Stanford, CA 94305. The authors thank Persi Diaconis, William T. Ziemba, the referees, and the editors for helpful comments. They also thank the California Lottery Commission for some of the data used here. This work was partially supported by National Science Foundation Grant NCR-8520136.

Journal of the American Statistical Association
December 1989, Vol. 84, No. 408, Applications & Case Studies

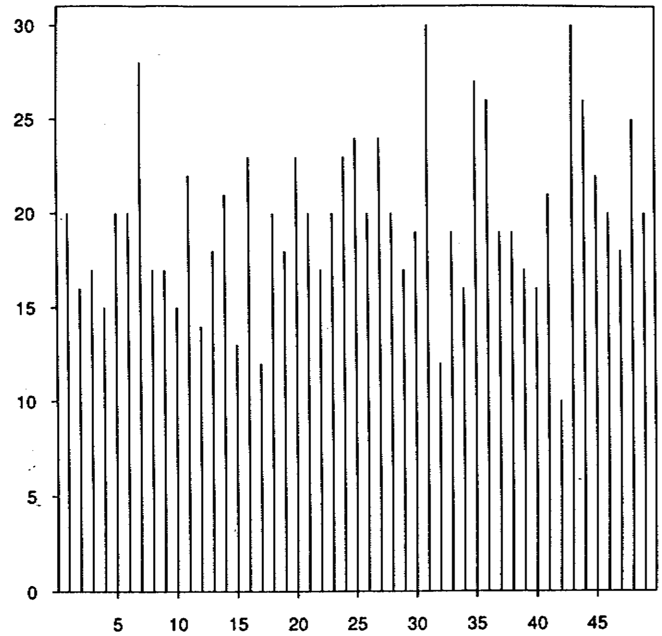

Figure 1. The Number of Times That Each Number Appeared as a Winning Number in 161 Games. These data are consistent with the hypothesis that each number is equally likely to win.

choose numbers that are not popular with other players. In the context of modern lottery games this has been studied by Chernoff (1981) and Ziemba et al. (1986). Do unpopular numbers exist in the Canadian Lottery? Figure 2 shows the number of times each number was selected in the game of July 6, 1985. This data is available to the public through the British Columbian lottery newsletter. If numbers are chosen at random by the public, then each number should appear approximately 539,000 times. The

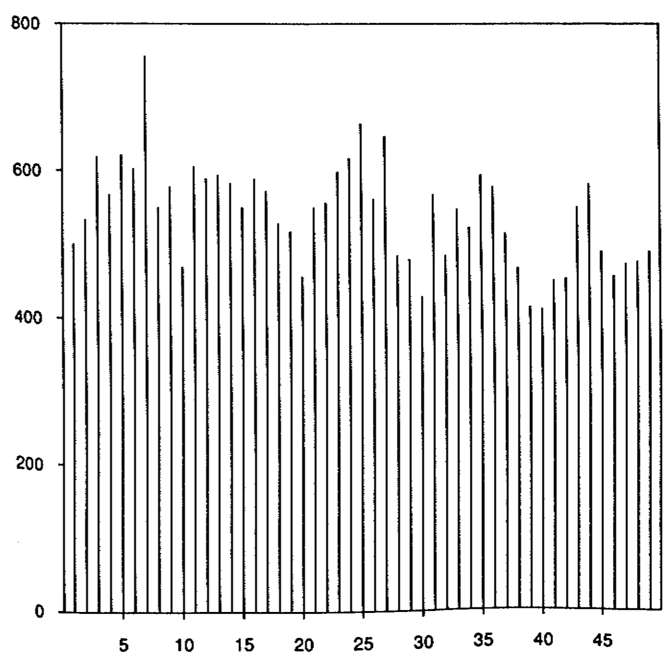

Figure 2. The Number of Times (in thousands) That Each Number Was Selected by the Public (July 6, 1985). These data reject the hypothesis that each number is equally likely to be selected. The number 7 is selected 50% more often than the average number.

number 7 appears more than 750,000 times. This fact alone causes the chi-squared test to reject randomness at all the usual significance levels. In addition, both Ziemba et al. and Chernoff found that unpopular numbers tend to remain unpopular (with more publicity this may stop). In the Canadian lottery, unpopular numbers tend to be those larger than 30 and those ending in 8, 9, or 0.

Given the rules for determining prizes in Canada, unpopular numbers lead to large prizes in two ways. There are fewer $10 winners, so a large amount of money is available for other prizes. There are also fewer prize winners with whom the pari-mutuel prizes must be split. The distribution of tickets purchased must be estimated to estimate the advantage that is obtained by choosing unpopular numbers.

3. THE DISTRIBUTION ON TICKETS

Let T_j ($j = 1, \ldots, n$) be n independent tickets from the unknown distribution $P(t)$. Each ticket T_j is a random m-tuple chosen from the first M integers. The probability that T_j is the m-tuple t is given by $\Pr(T_j = t) = P(t)$. The information in Figure 2 represents the marginals of this distribution. The ith marginal r_i is the empirical probability that a ticket includes the number i. The goal here is to estimate the distribution $P(\cdot)$ consistent with the constraints implied by the marginals.

Initially, suppose that the distribution on tickets is uniform, so each ticket has the same probability of being selected. Then the marginals r_k can be treated as constraints and the conditional distribution given these constraints can be determined. This distribution is given in the limit as the number of tickets n tends to infinity by the following theorem.

Theorem 1 (Csiszár 1984; Van Campenhout and Cover 1981). If the unconditional distribution on tickets is uniform $[\Pr(T_1 = t) = 1/\binom{M}{m}$ for each $t]$, then

$$\Pr\left\{T_1 = t \;\middle|\; \frac{1}{n}\sum_{j=1}^{n} I_k(T_j) = r_k,\ k = 1, \ldots, M\right\} \to P^*(t),$$

where $P^*(t)$ maximizes $H(P) = -\sum_t P(t)\ln P(t)$ over probability mass functions $P(t)$, which satisfy

$$\sum_t P(t) I_k(t) = r_k, \qquad k = 1, \ldots, M. \tag{1}$$

The functional $H(\cdot)$ is the Shannon entropy function. It is a measure of the randomness of a discrete distribution and is a concave function over the set of probability mass functions. The form of the maximum entropy distribution can be determined by the method of Lagrange multipliers. The result is neatly given in the following theorem.

Theorem 2 (Kagan, Linnik, and Rao 1973). The distribution $P^*(t)$ that maximizes $-\sum P(t)\ln P(t)$ subject to the constraints $\sum P(t) I_k(t) = r_k$ ($k = 1, \ldots, M$) is

$$P^*(t) = \exp\left(\lambda_0 + \sum_{j=1}^{M} \lambda_j I_j(t)\right) = c \prod_{j=1}^{M} \theta_j^{I_j(t)},$$

where $\lambda_0, \lambda_1, \ldots, \lambda_M$ and $c, \theta_1, \ldots, \theta_m$ are chosen to satisfy $\sum P^*(t) = 1$ and the constraints (1).

Proof. Suppose that $Q(t)$ is another distribution that satisfies the constraints. Then

$$\begin{aligned} H(Q) &= -\sum_t Q(t)\ln Q(t) \\ &\le -\sum_t Q(t)\ln P^*(t) \\ &= -\sum_t Q(t)\left(\lambda_0 + \sum_{j=1}^{M} \lambda_j I_j(t)\right) \\ &= -\sum_t P^*(t)\left(\lambda_0 + \sum_{j=1}^{M} \lambda_j I_j(t)\right) \\ &= -\sum_t P^*(t)\ln P^*(t) = H(P), \end{aligned}$$

where the inequality follows directly from the fact that the Kullback–Leibler divergence (Kullback 1959) $\sum Q(t) \ln(Q(t)/P^*(t))$ is nonnegative. The equality on the fourth line is true, since $P^*(t)$ and $Q(t)$ each satisfy the same constraints.

Note that the maximum entropy distribution has the form of a multinomial model. The probability of an m-tuple is obtained by multiplying the factors θ_j corresponding to the numbers in the ticket.

The maximum entropy distribution can be found numerically using the generalized iterative scaling algorithm (Darroch and Ratcliff 1972). Let $P^{(0)}(t) = \prod \phi_j^{(0)}$ for any choice of the starting values $\phi_j^{(0)}$, where the product includes each number j on the ticket t. For Canada's Lotto 6/49 the choice $\phi_j^{(0)} = 1/\binom{49}{6}^{1/6}$ for each j corresponds to the uniform distribution. The $(n + 1)$st estimate of the probability of the ticket t is computed from the nth estimate

$$P^{(n+1)}(t) = P^{(n)}(t) \prod_{k=1}^{M} \left(\frac{r_k}{r_k^{(n)}}\right)^{I_k(t)/m},$$

where $r_k^{(n)} = \Pr\{I_k(t) = 1\}$ under $P^{(n)}(t)$. At each step of the iteration the probability for every ticket is updated. We would like to avoid this, since the distribution puts mass on $\binom{M}{m}$ tickets and M may be large. A modified algorithm simply updates the ϕ's,

$$\phi_k^{(n+1)} = \phi_k^{(n)}\left(\frac{r_k}{r_k^{(n)}}\right)^{1/m}, \qquad k = 1, \ldots, M, \qquad (2)$$

and reduces the memory requirement significantly. Even with this modification, computing the maximum entropy distribution required more than 20 CPU hours on a VAX 11/780 when $M = 49$, $m = 6$, and the uniform initial estimate was used. Better initial estimates reduce this by a factor of 10. A theorem of Darroch and Ratcliff (1972) proves convergence of the algorithm whenever the constraints are consistent. At convergence, the transformation $\phi_j = c^{1/m}\theta_j$, with c chosen so that $\sum\theta_j = 1$, puts the distribution in the form given by the theorem.

4. EVALUATING LOTTO STRATEGIES

The maximum entropy distribution produces an estimate of the probability that a randomly selected lotto player will choose a particular ticket. All calculations are for the Canadian lotto game of July 6, 1985, with $\binom{49}{6}$ = 13,983,816 sixtuples. Under the uniform distribution each sixtuple has probability 7.15×10^{-8} of being selected. It follows from the form of the maximum entropy distribution that the least popular ticket in the maximum entropy distribution consists of the six least popular numbers in Figure 2. When the assumptions leading to the maximum entropy distribution are correct, the least popular ticket, containing the numbers 20, 30, 39, 40, 31, and 48, is selected with probability 1.73×10^{-8}. A couple of other interesting choices: 6-11-15-20-29-44 has probability 7.15 $\times 10^{-8}$, approximately the same as under the uniform distribution, and the most popular ticket, 3-7-9-11-25-27, has probability 2.52×10^{-7}. It seems that the least popular ticket occurs one-fourth as often as it would under the uniform distribution and the most popular approximately four times as often. A histogram of the probabilities under the assumptions leading to the maximum entropy distribution appears in Figure 3. For the uniform distribution all 13,983,816 tickets fall in one cell of the histogram.

The probability of winning each prize and the expected value of each prize are required to compute the expected return for a particular sixtuple. The probability of winning various prizes can be computed from the hypergeometric distribution. The probability of matching j of the six winning numbers is $\binom{6}{j}\binom{43}{6-j}/\binom{49}{6}$. The probability of winning any prize at all is less than .02. These probabilities are the

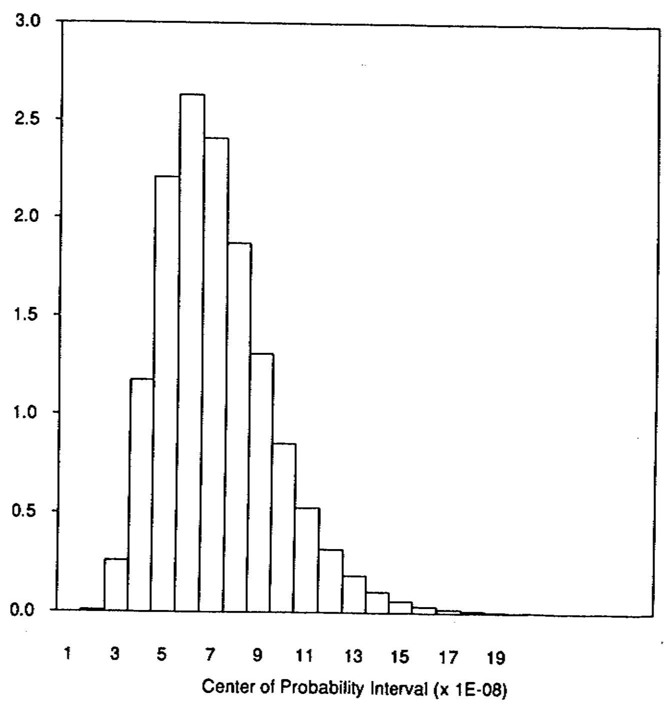

Figure 3. Histogram of Ticket Probabilities Under the Maximum Entropy Distribution. Frequencies are measured in millions of tickets. There is a total of 13,983,816 tickets.

same for any selected sixtuple, since the winning numbers are selected at random. The expected prizes depend on the ticket being considered. The expected prize for matching a given subset of the winning numbers cannot be easily computed analytically. Approximations can be made, but these are still computationally intensive. To avoid this, we use a simulation to compute the results of sample lottery games. We first ignore the special rules for the jackpot (the imposed minimum and the carryover from previous games).

The first step in the simulation is the choice of the simulation ticket. This is the sixtuple of numbers for which the expected prizes are to be computed. The expected value of each prize is computed separately. As an example, to compute the expected prize for four matching numbers the following procedure is used. Let X be the number of tickets purchased in the game. A set of winning numbers is chosen at random from among all sixtuples that contain four of the six numbers on the simulation ticket. A bonus number is also chosen. For the six winning numbers and the bonus number chosen in the simulation, the maximum entropy distribution P^* is used to compute the probabilities p_3, p_4, p_5, p_{5+b}, p_6 that a randomly selected ticket matches 3, 4, 5, 5, and the bonus, or six of the winning numbers. These probabilities are different for each simulated set of winning numbers. It is assumed that $X - 1$ tickets are selected according to P^* and the remaining ticket is the simulation ticket. The number of winners of the various prizes has the multinomial distribution with probabilities p_3, p_4, p_5, p_{5+b}, p_6 and $1 - p_3 - p_4 - p_5 - p_{5+b} - p_6$. This last probability is the probability of not winning any prize. An observation is drawn from this multinomial distribution using a sequence of binomial random variables. The number of tickets that match three of the winning numbers has the binomial distribution with $X - 1$ trials and probability of success p_3. Call this random variable N_3. Conditioned on the value of N_3, the number of tickets with four winning numbers is a binomial with $X - 1 - N_3$ trials and probability $p_4/(1 - p_3)$. This procedure is repeated to obtain the complete multinomial observation. Poisson random variables generated using the IMSL routine GGPON have been used to approximate the binomial random variables. The prizes are then determined according to the lotto rules. One hundred sets of simulated winning numbers are used to find the mean prize for four winning numbers and an estimate of the standard error. The expected value of other prizes is determined in the same manner.

The simulation has been repeated for several sixtuples of interest. The results for the least popular, most popular, and an "average" ticket are given in Table 1. The "average" ticket is a sixtuple that has approximately the same probability of being selected under the uniform and the maximum entropy distribution. Estimated standard errors for the mean prize are given in parentheses. Prizes for unpopular tickets are much larger than those for popular tickets. The last number in each column represents the expected return for a one dollar ticket. Standard errors for the estimated mean ticket return are approximately 1.5 cents. This standard error should be distinguished from the standard deviation of the return on a single ticket, which is quite large. X is taken to be 10 million tickets per game (this is typical of the Canadian lottery). The least popular ticket returns \$.30 more than the "average" ticket. Note that the return for the "average" ticket is less than we might expect. We suspect that this is a result of ignoring the jackpot minimum and the jackpot carryover. The effect of ignoring these jackpot rules is eliminated by considering $X = 1$ billion tickets. In that case the least popular ticket returns \$1.28 on average, the most popular ticket returns \$.25, and the "average" ticket returns \$.45. This last figure matches the 45% return we expect, since 55% of ticket sales are not returned as prizes. A favorable strategy has been determined for those readers who insist on playing the lottery—choose an unpopular ticket and convince a billion of your friends to enter the lottery.

The jackpot carryover can be analyzed using simulation as above. The least popular ticket benefits from large carryovers, since there are unlikely to be many winners when the least popular ticket wins the jackpot. For the least popular ticket and sales of 10 million per week a carryover of about 5 million dollars leads to an expected return of more than one dollar. This calculation ignores the increase in the number of bettors when the jackpot carryover is large. Taxes and deferred payments must also be considered in determining the break-even point.

Of course the chance of winning any prize is less than 1 in 50. If only one ticket is purchased each week, then many thousands of games are required for a player to have a chance of ending up ahead. This is due to the fact that the improved return is mainly a result of infrequent, extremely large prizes. An alternative is to have several players pool their money and bet on many unpopular combinations. The least popular 7,710 combinations pictured in Figure 3 each have expected returns 40% higher than the average ticket.

5. EVALUATING THE FIT OF THE MAXIMUM ENTROPY DISTRIBUTION

The maximum entropy distribution is a first approximation to the distribution on lotto tickets; it is close to the uniform distribution and matches the observed marginals. The theoretical justification of the maximum entropy distribution assumes that the unconditional distribution on tickets is the uniform distribution. There are several reasons, however, to suspect that the unconditional distribution may not be uniform. Most states make it possible for players to have a computer pick a random ticket for them. Typically, less than half of the players use this option. The rest are a conglomeration of players who tend to choose birthdays, lucky numbers, and numbers with convenient locations on the ticket. In addition, there are countless over-the-counter devices sold to lottery players as randomization devices that have built-in biases. Commercially available dice in Massachusetts, which are supposed to generate a random sixtuple, actually generate

Table 1. Expected Prizes and Expected Returns for Sample Tickets Under the Maximum Entropy Distribution

Number of winning nos.	Least popular (20 30 39 40 41 48)	Average (6 11 15 20 29 44)	Most popular (3 7 9 11 25 27)
3	10.0 (.0)	10.0 (.0)	10.0 (.0)
4	151.3 (2.2)	72.4 (1.55)	33.0 (.76)
5	6,305.0 (111)	2,047.7 (50)	597.4 (14.4)
5 + bonus	375,032.0 (18,073)	108,050.0 (6,113)	25,120.0 (787)
6	1,482,952.0 (28,626)	934,802.0 (34,381)	269,851.0 (16,541)
Expected return	.703	.397	.249

NOTE: Estimated standard errors are in parentheses.

less than 3% of the possible combinations. In addition, recall from Figure 2 that the marginals are not consistent with the uniform distribution. For these reasons we expect that the maximum entropy distribution will dramatically underestimate the popularity of the most popular tickets.

It is necessary to determine the consequences of the nonrandom ticket buyers before applying the simulation results of the previous section. Table 2 indicates the number of sixtuples that were selected with particular frequencies in a sample of 5,717,817 tickets from a recent game in the California lottery (which has rules similar to those of the Canadian lotto game). Table 2 also contains the expected number of tickets with particular frequencies under the maximum entropy distribution. The maximum entropy distribution from the Canadian game was used in Table 2, since the range of popular and unpopular numbers is similar to those in California. As expected, maximum entropy underestimates the number of sixtuples selected more than twice. The maximum entropy distribution does not do as poorly for unpopular tickets. In fact, the number of sixtuples that were not selected is underestimated. This indicates that the maximum entropy distribution may be conservative in evaluating the benefits of selecting unpopular tickets.

6. ALTERNATIVE APPROACHES

The bottom line in examining the fit of the maximum entropy distribution is that a multinomial distribution on nearly 14 million cells is fit using 49 parameters. The most logical way to improve the fit would be to gather more data. If, for example, the pairwise marginals q_{ij} are available (the proportion of tickets purchased that contain an i and a j), then the arguments given earlier can be repeated to find a maximum entropy distribution with $\binom{49}{2} = 1,176$ parameters. In this case the probability of a ticket is obtained by multiplying factors τ_{ij} corresponding to pairs of numbers, i and j ($i < j$), on the ticket. This would allow the introduction of correlation into players' selections. Note that although the formal assumption behind the maximum entropy distribution would still be suspect this increased information would lead to a better fit.

Table 2. Number of Sixtuples Selected With Different Frequencies

Frequency	Actual	Maximum entropy
0 times	10,324,457	9,372,767
1 time	2,695,138	3,670,944
2 times	639,870	794,660
3 times	181,819	126,627
4 times	65,537	16,668
5 times	29,251	1,926
10 times	2,677	0
20 times	371	0
>20 times	5,078	0

Joe (1987) considered other models for estimating the distribution on tickets. A class of minimally dependent distributions was suggested for games like lotto. This class includes the maximum entropy distribution. Ziemba et al. (1986) used regression to estimate directly the expected return on a ticket instead of estimating the unknown distribution. They found that unpopular tickets may have expected returns as high as \$1.50, with no carryover.

Another approach to estimating the distribution of tickets purchased by the public is to model the public's behavior. For example, suppose that a proportion P_1 of the ticket buyers chooses numbers at random and the remainder (the nonrandom ticket buyers) select tickets from a fraction, P_2, of the nearly 14 million possibilities. Let NR represent the subset of sixtuples considered by the nonrandom ticket buyers. This would include popular numbers and common birthdays. Let N_t be the number of times the sixtuple t is selected in a sample of N tickets. Under this model the probability that $N_t = k$ is computed by conditioning on whether the sixtuple is part of the nonrandom subset NR. Let λ_1 represent the expected number of times a ticket is selected by the random ticket buyers, and let λ_2 represent the expected number of times a ticket in the NR subset is selected by either random or nonrandom ticket buyers. Then, using the Poisson distribution for N_t leads to

$$\begin{aligned}\Pr(N_t = k) &= \Pr(t \notin \text{NR})\Pr(N_t = k \mid t \notin \text{NR}) \\ &\quad + \Pr(t \in \text{NR})\Pr(N_t = k \mid t \in \text{NR}) \\ &= (1 - P_2)\exp(-\lambda_1)\frac{\lambda_1^k}{k!} + P_2\exp(-\lambda_2)\frac{\lambda_2^k}{k!},\end{aligned}$$

where

$$\lambda_1 = NP_1/13983816$$

and

$$\lambda_2 = \frac{NP_1}{13983816} + \frac{N(1 - P_1)}{13983816\, P_2}.$$

The maximum likelihood estimates for P_1 and P_2 are computed from the Southern California data using a method described in Efron and Thisted (1976). The probabilities $\Pr(N_t = k)$ for $k = 0, \ldots, 20$ are used as probabilities in a multinomial likelihood (conditioning on 20 or fewer selections). The observed Southern California frequencies are used to determine the maximum likelihood estimates for P_1, P_2. In this case the estimates are $\hat{P}_1 = .711$ and $\hat{P}_2 = .041 \approx \binom{30}{6}/\binom{49}{6}$. This simple model suggests that 71% of the public chooses tickets at random and 29% chooses tickets from a subset of 30 numbers.

The Efron and Thisted (1976) approach to modeling Shakespeare's word usage provides another model for ticket frequencies. The number of occurrences of the ith ticket is taken to be a Poisson random variable with mean λ_i. Each ticket is assumed to have its own mean rate of occurrence. A gamma distribution with shape parameter α and scale parameter β is used as a prior distribution for the λ's. The unknown parameters α, β are estimated using the multinomial likelihood approach. The fit of the two models described here is compared with the actual data and the maximum entropy fit in Table 3. Other models produced similar results.

The models described here fit the Southern California data better than the maximum entropy distribution and use only two parameters each. Unfortunately, they provide little help in identifying sixtuples that are unpopular. They merely provide a model that is consistent with the presence of popular and unpopular numbers. Other sources of information are required to determine which tickets are likely to be unpopular.

Table 3. Number of Sixtuples Appearing k Times Under Each Model (in thousands)

Frequency	Actual	Maximum entropy	P_1, P_2 Model	Gamma
0 times	10,324	9,372	10,049	10,390
1 time	2,695	3,671	2,991	2,455
2 times	640	795	544	754
3 times	182	127	169	249
4 times	65	17	104	85
5 times	29	2	65	30
10 times	2.7	0	.7	.2
20 times	.3	0	0	0

7. SUMMARY

The lottery numbers selected by players are not chosen at random. The sizes of the prizes are larger for combinations of numbers that are unpopular than for other combinations. A theorem shows that if the unconditional distribution of tickets selected is uniform, then the conditional distribution given a set of marginal constraints tends to the maximum entropy distribution. In this case simulations indicate that choosing unpopular numbers improves the expected return on a one dollar ticket by more than 50% over a random selection. The unpopular numbers include 20, 30, 38, 39, 40, 41, 42, 46, 48, and 49. The most popular numbers, which should be avoided, are 3, 7, 9, 11, 25, 27. Most of the evidence seems to indicate that the assumption behind the maximum entropy result is suspect. This introduces doubts about the use of simulation results in designing a lotto strategy. Other modeling approaches are discussed. These alternatives do not provide as much detailed information as the maximum entropy distribution but seem to better capture the range of popular and unpopular tickets. These approaches do not provide strategic advice. In general, lottery purchases for profit are still not recommended. By choosing unpopular numbers and waiting for carryover jackpots, however, it is possible to turn a lottery bet into an even or favorable bet.

[*Received February 1987. Revised May 1989.*]

REFERENCES

Chernoff, H. (1981), "An Analysis of the Massachusetts Numbers Game," *Mathematical Intelligencer*, 3, 166–172.

Csiszár, I. (1984), "Sanov Property, Generalized *I*-Projection and a Conditional Limit Theorem," *The Annals of Probability*, 12, 768–793.

Darroch, J. N., and Ratcliff, D. (1972), "Generalized Iterative Scaling for Log-Linear Models," *The Annals of Mathematical Statistics*, 43, 1470–1480.

Efron, B., and Thisted, R. (1976), "Estimating the Number of Unknown Species: How Many Words Did Shakespeare Know?" *Biometrika*, 63, 435–447.

Joe, H. (1987), "An Ordering of Dependence for Distributions of *k*-tuples, With Applications to Lotto Games," *Canadian Journal of Statistics*, 15, 227–238.

Kagan, A. M., Linnik, Y. V., and Rao, C. R. (1973), *Characterization Problems in Mathematical Statistics*, New York: John Wiley.

Kullback, S. (1959), *Information Theory and Statistics*, New York: John Wiley.

Van Campenhout, J. M., and Cover, T. M. (1981), "Maximum Entropy and Conditional Probability," *IEEE Transactions on Information Theory*, 27, 483–489.

Ziemba, W. T., Brumelle, S. L., Gautier, A., and Schwartz, S. L. (1986), *Dr. Z's 6/49 Lotto Guidebook*, Vancouver: Dr. Z Investments.