The exact null distribution for radial maze statistics: A FORTRAN 77 program

RAPHAEL GILLETT

University of Leicester, Leicester, England

Performance in the radial-arm maze may be assessed by comparing the number of correct responses or, alternatively, the number of repetition errors with that expected if a subject were simply selecting a location at random on each choice. Previous attempts to estimate the null distribution of correct responses or repetition errors have relied on computer simulation. In the present approach, a direct analytic argument is used to derive exact expressions for the probability density functions of these random variables. Knowledge of the exact null distributions provides the researcher with enhanced analytic capabilities. For example, it supplies the means of constructing an exact test of significance to evaluate performance at the level of a single subject. A FORTRAN 77 program is provided to compute the null distribution of correct responses and repetition errors as well as their respective expectations and variances. Application of the technique to other situations is described briefly.

The radial-arm maze is widely used to study choice behavior and working memory in humans (Foreman, Foreman, Cummings, & Owens, 1990) and in animals (Brown, 1992). The apparatus consists of a number of locations arranged in a circle around a central point. The radial-arm maze derives its name from animal experiments in which corridors or arms radiate from the center to the locations on the circumference. A reward, usually in the form of food, is hidden at the end of each arm. The subject is placed in the central area and allowed to choose which locations to visit to obtain a reward. Because locations are identical in appearance, the subjects must use ambient spatial cues to discriminate among them. Once a baited location has been visited (and the reward consumed), it becomes an unbaited location.

Although the radial maze task allows the subject to sample with replacement among the spatial locations, it is clearly more efficient to sample without replacement. Thus, optimal performance occurs if a subject succeeds in visiting each of the baited locations without repetition. Commonly, however, a number of repeat visits (repetition errors) will be made to previously selected, and now empty, arms of the maze. In this event, performance may be assessed by comparing the number of repetition errors with that expected if a subject were simply selecting a location at random on each choice. Alternatively, performance may be evaluated by the number of correct choices (rewards obtained). The two measures are inter-

changeable because the number of correct choices is linearly related to the number of repetition errors.

Accordingly, it is necessary to establish the distribution of correct choices, or the distribution of repetition errors, under the hypothesis of random selection. Previous approaches (Eckerman, 1980; Olton, 1978; Spetch & Wilkie, 1980) have employed computer simulation to estimate these distributions. In the present article, a direct analytic argument is used to derive exact closed-form expressions for the probability of making R repetition errors, and the probability of making C correct choices.

Knowledge of the exact null distributions of $p(C)$ and $p(R)$ provides the researcher with enhanced analytic capabilities. It supplies the means of constructing an exact test of significance for evaluating performance at the level of a single subject. It should be noted, however, that although the random-choice model provides a useful baseline against which to measure performance, it does not, of course, offer an explanation for any nonrandom behavior that the test might uncover, such as the characteristic response biases exhibited by rats in the radial maze.

The results of the present analysis may be used to measure performance in other settings. Human activities frequently require a record of previously sampled items to be kept in working memory. Many of these situations possess a structure similar to that of the radial maze. Note that the items being sampled by the subject may be distinguished from one another by features other than spatial location. For example, a farmer may enter a pen to administer antibiotics to a small group of sick animals that are individually identifiable. As the cattle move around in the pen, the farmer may have difficulty in remembering those animals already treated, and hence may repeat the treatment a second time for some animals and omit treatment for others.

Response Probabilities Under Random Selection

Let n represent the number of arms of the radial maze, and k the number of choices that the subject is allowed to make. We wish to determine the probability $p(C = c)$ that a subject will make exactly c correct responses, that is, will obtain exactly c rewards. Under the hypothesis of random selection, the radial maze task has the same structure as the classical occupancy problem, in which k balls are distributed in n cells. Each ball is assigned to a cell chosen at random. The occupancy problem is so named because interest lies in the distribution of occupied cells (i.e., cells containing at least one ball). Thus, the probability $p(C = c)$ that exactly c locations in a radial maze are visited at least once is equivalent to the probability that exactly c cells in the occupancy model each receive at least one ball. From Johnson and Kotz (1969),

$$p(C=c) = \binom{n}{c} \sum_{i=0}^c (-1)^i \binom{c}{i} \left(\frac{c-i}{n}\right)^k. \quad (1)$$

Address correspondence to R. Gillett, Department of Psychology, University of Leicester, Leicester, LE1 7RH, Great Britain.

Table 1
Probability of Exactly c Correct Choices and Probability of
at Least c Correct Choices When a Subject is Allowed
 $k=12$ Choices in a Radial Maze with $n=16$ Arms

c	1	2	3	4	5	6	7	8	9	10	11	12
$p(C=c)^*$	.000	.000	.000	.003	.027	.129	.293	.329	.176	.041	.003	
$p(C \geq c)$	1.00	1.00	1.00	1.00	1.00	.997	.970	.842	.549	.220	.044	.003

The expectation (or mean) and variance of C are

$$E(C) = n \left[1 - \left(1 - \frac{1}{n} \right)^k \right] \quad (2)$$

and

$$\begin{aligned} \text{Var}(C) = n \left(1 - \frac{1}{n} \right)^k + n(n-1) \left(1 - \frac{2}{n} \right)^k \\ - n^2 \left(1 - \frac{1}{n} \right)^{2k}. \end{aligned} \quad (3)$$

The number of repetition errors is given by $R = k - C$. Hence, the probability $p(R = r)$ that a subject will make exactly r repetition errors is

$$p(R=r) = \binom{n}{k-r} \sum_{i=0}^{k-r} (-1)^i \binom{k-r}{i} \left(\frac{k-r-i}{n} \right)^k. \quad (4)$$

The expectation (or mean) and variance of R are

$$E(R) = k - n \left[1 - \left(1 - \frac{1}{n} \right)^k \right] \quad (5)$$

and

$$\text{Var}(R) = \text{Var}(C). \quad (6)$$

Some of the formulas can be time consuming to evaluate, so it would be helpful to researchers to use a program to perform the calculations on a microcomputer. Accordingly, a FORTRAN 77 program titled RADMAZE is provided in the Appendix to compute the exact null distributions of $p(C)$ and $p(R)$ for $n \leq 36$ locations and any number of choices k . Note that k may be greater or less than n . To run the program on an IBM PC or compatible microcomputer, simply type RADMAZE and press the "enter" key. The program prompts the user for values of n and k .

Test of Significance for a Single Subject

Suppose that a rat has been allowed to make $k = 12$ choices in a radial maze with $n = 16$ arms, and that 10 of these choices were correct responses (i.e., visits to baited locations). Table 1 provides both the probability of exactly c correct choices and the probability of at least c correct choices, for $c = 1, 2, \dots, 12$, under the null hypothesis that the rat chooses locations at random. The values in Table 1 were supplied by RADMAZE, which also yields values for the expectation and variance of the number of correct choices: $E(C) = 8.625$, and $\text{Var}(C) = 1.321$.

From Table 1 it is apparent that the probability of obtaining at least 10 correct choices is $p(C \geq 10) = .22$.

Because this p value is greater than $\alpha = .05$ (one-tailed), it follows that the null hypothesis cannot be rejected. By simply selecting locations at random, a rat could achieve 10 correct responses in a 16-arm radial maze relatively frequently. Therefore, it can be concluded that the rat's performance does not significantly differ from a pattern of choice in which locations are selected at random.

Accuracy

All probability values, expectations, and variances are accurate to at least three decimal places. Researchers will seldom wish to employ values of n greater than the program limit of 36. If the array limit in the program were raised beyond 36 for any purpose, the accuracy of the expectations and variances would be unaffected, but the accuracy of the probability values would diminish as n became large.

As might be expected, a comparison of the output of RADMAZE with the results of earlier attempts (Eckerman, 1980; Olton, 1978; Spetch & Wilkie, 1980) to obtain the expected number of correct choices $E(C)$ using simulation techniques reveals some minor discrepancies (mostly in the second decimal place) in the figures supplied by these earlier studies.

Availability

A copy of the program in the Appendix can be obtained by e-mail (rtg@leicester.ac.uk), or by mail from R. Gillett, Department of Psychology, University of Leicester, Leicester LE1 7RH, Great Britain. In the latter case, please send a formatted (MS-DOS 3.3 or later) disk, a self-addressed disk mailer, and return postage (e.g., international reply coupons, obtainable at any post office).

REFERENCES

- BROWN, M. F. (1992). Does a cognitive map guide choices in the radial-arm maze? *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 56-66.
- ECKERMAN, D. A. (1980). Monte Carlo estimation of chance performance for the radial arm maze. *Bulletin of the Psychonomic Society*, **15**, 93-95.
- FOREMAN, N., FOREMAN, D., CUMMINGS, A., & OWENS, S. (1990). Locomotion, active choice, and spatial memory in children. *Journal of General Psychology*, **117**, 215-232.
- JOHNSON, N. L., & KOTZ, S. (1969). *Distributions in statistics: Discrete distributions*. New York: Wiley.
- OLTON, D. S. (1978). Characteristics of spatial memory. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 341-373). Hillsdale, NJ: Erlbaum.
- SPETCH, M. L., & WILKIE, D. M. (1980). A program that simulates random choice in radial arm mazes and similar choice situations. *Behavior Research Methods & Instrumentation*, **12**, 377-378.

APPENDIX
Listing of Program RADMAZE

```

PROGRAM RADMAZE
C
C   THE EXACT NULL DISTRIBUTION FOR RADIAL ARM MAZE STATISTICS
C
C   RAPHAEL GILLETT, DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF
C   LEICESTER, LEICESTER LE1 7RH, U.K.
C
  IMPLICIT REAL*8 (A-H,O-Z)
  REAL*8 P(0:36), PGE(0:36)
C
C   INPUT MAZE PARAMETERS
C
10  WRITE(6,*)'ENTER N (NUMBER OF ARMS IN RADIAL MAZE):'
    READ(5,*)N
    IF(N.GT.36)THEN
      WRITE(6,*)'N SHOULD BE LESS THAN OR EQUAL TO 36'
      GOTO 10
    ENDIF
    AN=N
    WRITE(6,*)'ENTER K (NUMBER OF CHOICES):'
    READ(5,*)K
    AK=K
C
C   COMPUTE P(J) THE PROBABILITY OF EXACTLY J CORRECT CHOICES
C   AND PGE(J) THE PROBABILITY OF AT LEAST J CORRECT CHOICES
C
    D=1.0
    P(0)=0
    PGE(0)=1
    J2=MIN(N,K)
    DO 200 J=1,J2
      AJ=J
      D=D*(AN-AJ+1.0)/AJ
      P(J)=(AJ/AN)**AK
      B=1.0
      DO 100 I=1,J
        AI=I
        B=-B*(AJ-AI+1.0)/AI
        P(J)=P(J)+B*((AJ-AI)/AN)**AK
100    CONTINUE
      P(J)=D*P(J)
      PGE(J)=PGE(J-1)-P(J-1)
200    CONTINUE
    WRITE(6,*)
    WRITE(6,*)'PROBABILITY OF EXACTLY C CORRECT CHOICES, P(C), AND'
    WRITE(6,*)'PROBABILITY OF AT LEAST C CORRECT CHOICES, P(>=C).'
    WRITE(6,300)'C',(J,J=0,J2)
300    FORMAT(1X,A1,5X,12I6/, (7X,12I6))
    WRITE(6,450)'P(C)',(P(J),J=0,J2)
450    FORMAT(1X,A4,2X,12F6.3/, (7X,12F6.3))
    WRITE(6,455)'P(>=C)',(PGE(J),J=0,J2)
455    FORMAT(1X,A6,12F6.3/, (7X,12F6.3))
    WRITE(6,*)

```

```

C
C   EXPECTATION AND VARIANCE OF NUMBER OF CORRECT CHOICES
C
E=AN*(1.0-(1.0-1.0/AN)**AK)
V=AN*((1-1/AN)**AK+(AN-1)*(1-2/AN)**AK-AN*(1-1/AN)**(2*AK))
WRITE(6,500)'EXPECTED VALUE = ',E,'VARIANCE = ',V
500 FORMAT(1X,A17,F7.3,10X,A11,F7.3)
WRITE(6,*)
WRITE(6,*)

C
C   REPETITION ERROR PROBABILITIES AND MOMENTS ARE CALCULATED
C   USING THE RELATIONSHIP  $R = K - J$ 
C
WRITE(6,*)'PROBABILITY OF EXACTLY R REPETITION ERRORS, P(R), AND'
WRITE(6,*)'PROBABILITY OF AT MOST R REPETITION ERRORS, P(<=R).'
WRITE(6,300)'R',(J,J=K-J2,K)
WRITE(6,450)'P(R)',(P(J),J=J2,0,-1)
WRITE(6,455)'P(<=R)',(PGE(J),J=J2,0,-1)
WRITE(6,*)
E=AK-AN*(1.0-(1.0-1.0/AN)**AK)
WRITE(6,500)'EXPECTED VALUE = ',E,'VARIANCE = ',V
END

```

(Manuscript received March 10, 1993;
revision accepted for publication September 3, 1993.)

Erratum

Brooks, J. O., III. The Trilinear program: A program to analyze longitudinal change on the Macintosh. *Behavior Research Methods, Instruments, & Computers*, 1993, 25(4), 485-486—On page 486, the third sentence should read:

Thus, the trilinear model will always fit data *at least* as well as the linear model, because if periods of stability are not present in a given subject's data, the estimates of change provided by the linear and trilinear models are equivalent.