Best
Available
Copy

62-4-4

MEMORANDUM
RM-3184-PR
JUNE 1962

INFERENCE OF MONTE CARLO PROPERTIES FROM THE SOLUTION OF A KNOWN PROBLEM

J. I. Marcum

PREPARED FOR:
UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM
RM-3184-PR
JUNE 1962

INFERENCE OF MONTE CARLO PROPERTIES FROM THE SOLUTION OF A KNOWN PROBLEM

J. I. Marcum

This research is sponsored by the United States Air Force under Project RAND — Contract No. AF 49(638)-700 — monitored by the Directorate of Development Planning, Deputy Chief of Staff, Research and Technology, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force. Permission to quote from or reproduce portions of this Memorandum must be obtained from The RAND Corporation.

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

PREFACE

RAND, among other organizations, pioneered in the use of the Monte Carlo method for the solution of complicated problems involving the transport of nuclear radiation such as gamma rays and neutrons. This method has been useful chiefly when solution by other means was prohibitively complicated.

One of the main disadvantages of the Monte Carlo method or, for that matter, any other method of random sampling is that one never quite knows the accuracy of the answers. In this report we solve by means of Monte Carlo a problem with a known analytical solution. This problem is similar to those occurring in actual RAND Monte Carlo codes. Comparison of the Monte Carlo solutions with the known solutions allows us to get a better feeling for the accuracy of Monte Carlo results in practical problems.

SUMMARY

It is shown that the probability density function

$$dp = \frac{2dx}{x^3} \qquad x > 1$$

is a reasonable approximation to sampling distributions found in several Monte Carlo transport codes.

A machine code was set up which was used to obtain 10,000 runs of 10 samples each from the above distribution function. Histograms of the probability density function for the mean value, of the standard deviation of the mean and of Student's ratio of 10 samples were obtained. Because the standard deviation has an infinite expected value, it is of doubtful usefulness in measuring the probable error of, or placing confidence limits on, the mean value of the sample.

The results obtained in this paper, however, show that with proper precautions the sample standard deviation may be used to establish confidence limits on the sample mean in spite of the fact that it has an infinite expected value.

A further examination of the results of this problem sheds light on other properties of Monte Carlo solutions, including the fact that the results are biased. Although the results of this problem do not allow one to make quantitative improvements in the interpretation of Monte Carlo results, they do uncover several qualitative features of the Monte Carlo results that allow for more lucid interpretation.

CONTENTS

I. INTRODUCTION

Over the past several years questions have continuously arisen as to the accuracy of solutions obtained from various transport codes that use the Monte Carlo method of solution. The first attempts to delineate the accuracy made use of the calculation of the sample variance of the sample mean as well as the sample mean itself for each point in the sample space. All present RAND Monte Carlo transport codes calculate the sample variance of the sample mean, and this quantity is used to estimate the error in the sample mean. However, it has been realized that this estimate of the error may oftentimes be misleading.(1) If the distribution function being sampled is Gaussian or a reasonable facsimile thereof, then the sample variance, or more correctly the sample standard deviation, will indeed give a known measure of the error. If the distribution being sampled deviates from a Gaussian distribution, the estimate of the error which can be inferred from the sample standard deviation of the sample mean becomes more muddled. In particular, when the distribution function being sampled has long tails, the difficulty mentioned above may become quite aggravated.

In this report we will set up a distribution function which is similar to ones encountered in the RAND neutron and gamma ray transport codes. However, with simplifications, this distribution function will have a known or calculable expected value. The Monte Carlo method is then used to obtain estimates of this expected value. The sample standard deviations are also estimated, and their value in estimating the error is then examined.

II. APPROXIMATION TO THE SOLUTIONS OBTAINED BY MONTE CARLO

The RAND neutron and gamma ray transport codes have been used to obtain Monte Carlo solutions for a large number of problems which involve point sources of neutrons or gamma rays. All of the solutions obtained with these Monte Carlo codes have, in general, the properties that at close distances to the source, the solutions have small fluctuations from one range interval to the next. These fluctuations increase as the range increases, finally becoming so very large that the solutions are useful only out to a certain maximum range, which is commonly of the order of 10 mean free paths of the radiation being considered. Oftentimes results are needed or desired at the very maximum ranges at which the Monte Carlo solutions are useful, and various so-called tricks are used to make deeper penetrations possible.

Even by these methods it commonly happens that the fluctuations or statistical uncertainties in the solutions at large ranges are sizeable. Since in many cases it is necessary to live with these large errors, one would like to interpret them as precisely as possible.

For any particular element of the sampling space, the Monte Carlo estimates a sample mean in accordance with the following formula:

$$\bar{x} = \frac{\Sigma x_i}{N} , \qquad (1)$$

where the x_i are the individual sample values of the quantity being estimated and N is the total size of the sample. The sample variance of the sample mean is estimated at the same time by means of

$$\bar{\sigma}^2 = \frac{1}{N^2}\Sigma x_i^2 - \frac{\bar{x}^2}{N}. \qquad (2)$$

The sample standard deviation is simply the square root of this quantity. Ordinarily, if one assumes that the distribution function of x_i is Gaussian, then the probability that $\bar{x}$, the sample estimate of the mean, will differ from the true mean by more than one standard deviation in either direction is about 0.32, and that it will differ by more than various other proportional parts of a standard deviation is in accordance with the cummulative value of a standard Gaussian (normal 0,1).*

We will now proceed to examine the validity of such assumptions for a particular distribution function, namely,

$$\begin{aligned} dp &= \frac{2dx}{x^3} \qquad x > 1 \\ &= 0 \qquad x > 1. \end{aligned} \qquad (3)$$

This distribution function is a reasonably approximate representation of the distribution functions encountered in the RAND Monte Carlo transport codes as shown in the Appendix. The expected or mean value of x in Eq. (3) is obviously 2.0. On the other hand, the second moment of x, and hence the variance, is infinite. Since the variance is infinite, this immediately throws doubt on the whole idea of using the sample variance as computed by means of Eq. (2), since, if enough samples are taken, it must become infinite.

---

*The use of the standard deviation of the sample mean for establishing confidence limits on the sample mean is discussed again on p. 11 and in Ref. 2, pp. 517-519.

III. MONTE CARLO SOLUTION OF PROBLEM

A machine code was set up to obtain the solution by random sampling for the mean value of x as given by the distribution function of Eq. (3). Samples from the distribution function were obtained in one of the usual ways by setting the cummulative distribution function equal to a random number uniformly distributed between 0 and 1. Symbolically

$$P = \frac{1}{x^2} = R, \tag{4}$$

or

$$x = \frac{1}{\sqrt{R}}. \tag{5}$$

One then obtains $\bar{x}$ for a sample size N by

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{i=N} \frac{1}{\sqrt{R_i}} . \tag{6}$$

The sample variance of the sample mean is given by

$$\bar{\sigma}^2 = \frac{1}{N^2} \sum_{i=1}^{i=N} \frac{1}{R_i} - \frac{\bar{x}^2}{N} . \tag{7}$$

It should be noted that the variance given in the above equation is not the sample variance itself, which would in fact be N times the above quantity, it is rather the sample variance of the sample mean; hence presumably a measure of the statistical error in the quantity $\bar{x}$. Besides computing the sample standard deviation of $\bar{x}$, which is simply the square root of the quantity given by Eq. (7), the code also computed Student's ratio

$$T = \frac{\bar{x} - 2}{\sigma} . \qquad (8)$$

The number of samples, N, was chosen as 10. This meant that the error in $\bar{x}$ would be of the order of 25 per cent, which is representative of the errors in the actual Monte Carlo problems at their maximum range of usefulness. If we had chosen very large sample sizes, say of the order N = 1000, then the error would be of the order of only a few per cent. This is a region not particularly interesting from the standpoint of the exact interpretation of the error. When the error is, say, 2 per cent in a typical Monte Carlo problem, we are not very often interested in whether it is 2 per cent or 4 per cent, and an examination of the errors in such regions as this would therefore not be particularly rewarding for our purposes. At the other extreme, if we had taken samples of size N = 2 or 3, where the error might be in the neighborhood of 75 to 100 per cent, this too would be a rather useless region to explore since the error here would be so large that our knowing whether it was 75 per cent or 125 per cent would be of little use. A sample size of 10 was accordingly chosen so that the error would be in the range where it was most useful to have a better understanding of the exact behavior of the error.

Having chosen $N = 10$, we then ran 10,000 runs of 10 each.*
Histograms were then obtained showing the distribution functions for $\bar{x}$, $\bar{\sigma}$ and T. Figures 1, 2 and 3 give the results for these three distribution functions respectively. Tables 1, 2 and 3 give the detailed numerical results from which these figures were prepared.

* A run of 10 means that $\bar{x}$, $\bar{\sigma}$ and T were computed according to Eqs. (6), (7) and (8) for $N = 10$. This procedure was then repeated 10,000 times, independently.

IV. RESULTS

Many interesting and unexpected conclusions can be drawn by examining Figs. 1, 2 and 3. Turning our attention first of all to Fig. 1 and Table 1, which give the distribution function for $\bar{x}$, one notices that its expected value is extremely close to the actual expected value of x, or $\bar{x}$, namely 2.0. However, the median and most probable values are considerably less than the true mean of 2.0. In the Monte Carlo results obtained from the RAND transport codes, it is customary to draw by eye a smooth curve through the various values of a histogram. Even though one has drawn hundreds of such curves, it is not obvious which ones of the above statistics are really incorporated in such a process. In other words, does one try to draw the smooth curve through the expected value, the median value, the most probable value, or some combination of these statistics? My guess is that something like the median or the most probable value is more heavily weighted than the expected value. If this is true, then one can see by referring again to the distribution shown in Fig. 1 that a biased estimate of the mean is being obtained, which in this particular case can underestimate the true mean by around 10 per cent.

Turning our attention now to Fig. 2 and Table 2, which give the distribution of the sample standard deviation of the sample mean, it can be seen at a glance that although this statistic has an expected value of infinity, a very reasonable sort of distribution function is nevertheless obtained. One can see, however, that its expectancy tends to be large because some very large values are cropping up in the tail of the curve. It is of interest to note that the semi-interquartile range of $\bar{x}$, 0.290 is very close to the median of $\bar{\sigma}$, 0.273, but quite different from the mean of $\bar{\sigma}$, 0.435. A possible interpretation of these facts is that the error in $\bar{x}$ will be greater than $\bar{\sigma}$ about half of the time and less than $\bar{\sigma}$ half of the time. Rather than examine the distribution function for $\bar{\sigma}$ in more detail, we will now look at the results for Student's ratio,* which is an error-measuring statistic of more value in this case than the standard deviation itself. Referring now to Fig. 3 and Table 3, we see first of all that the distribution function for T is extremely skew. The reason for this lies in the fact that there is a correlation between $\bar{x}$ and $\bar{\sigma}$. As is well known, the only distribution function where the sample mean and the sample variance are distributed independently is Gaussian. Our original distribution function of Eq. (3) is very far from Gaussian; hence we expect a sizeable correlation between the sample mean and the sample variance. If the sample mean is in error on the low side, then the sample standard deviation will probably be a large underestimate. Student's ratio will then be a large negative number; hence the extreme skewness to the left of Fig. 3.

*Student's ratio is a term usually applied to samples from a normal parent population. However, here we use the term in the more general sense, and apply it to samples from the distribution of Eq. (3).

In the distribution function of Student's ratio for $N \to \infty$, and an original normal distribution one expects to find about 68 per cent of the area of the distribution curve between -1 and 1. In the Student's distribution for $N = 10$ and an original normal distribution one finds some 65 per cent of the area between -1 and 1. In the distribution curve of Student's ratio, for our example, there is only 48 per cent of the area between -1 and 1. In order to contain 68 per cent of the area of Fig. 3 one must encompass the range -1.8 to +1.8. Therefore in interpreting the meaning of the sample standard deviation in Monte Carlo results whence the probability density function is similar to Eq. (3), one should be guided by the nature of the distribution functions obtained in this paper rather than mechanically using the usual Gaussian assumption. A rather complete discussion of the use of Student's ratio for the establishment of a confidence interval for the mean is given on pp. 517-519 of Ref. 2.[2]

It seems likely that if the assumed distribution function of Eq. (3) were modified by increasing the exponent of x from 3 to $3 + \varepsilon$, the same sort of results would be obtained. In other words, even though the variance would now be finite, the distribution functions for $\bar{\sigma}$ and Student's ratio would remain essentially the same, and therefore the interpretation of the confidence interval would be unchanged.

V. LIMITATIONS

One must keep in mind the assumptions which were made in the problem solved above. The flux was assumed to be isotropic and of one velocity (or energy). Also only the case N = 10 was treated. Thus one cannot use the results of this problem to make quantitative predictions of confidence limits in actual Monte Carlo results. However the qualitative aspects of the solution given here will apply to Monte Carlo results which are obtained by the sampling scheme given in the Appendix.

Values of N other than 10 could easily be run with the present code and more information could thus be obtained. In view of the approximate nature of the assumed distribution function, this added effort did not seem justified.

VI. CONCLUSIONS

By means of a simple example we have shown that the sample variance in a Monte Carlo problem may be used to estimate the error of the sample mean even though the variance of the distribution function itself is infinite, provided that adequate precautions are taken.

We have also shown that in the particular example used a biased estimate of the mean value may be obtained which will lead to an underestimate of the mean by about 10 per cent.

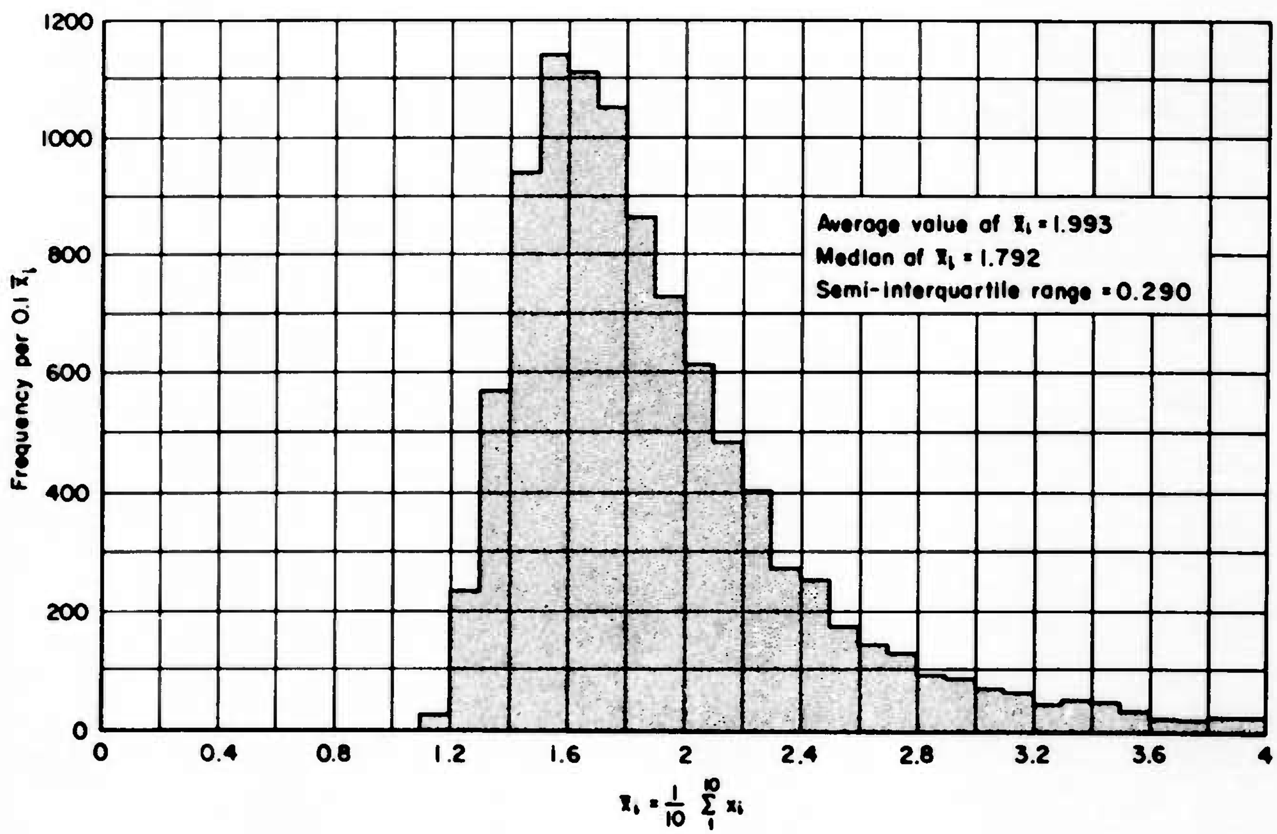

Fig. 1—Frequency distribution of sample mean for 10,000 runs of 10 each

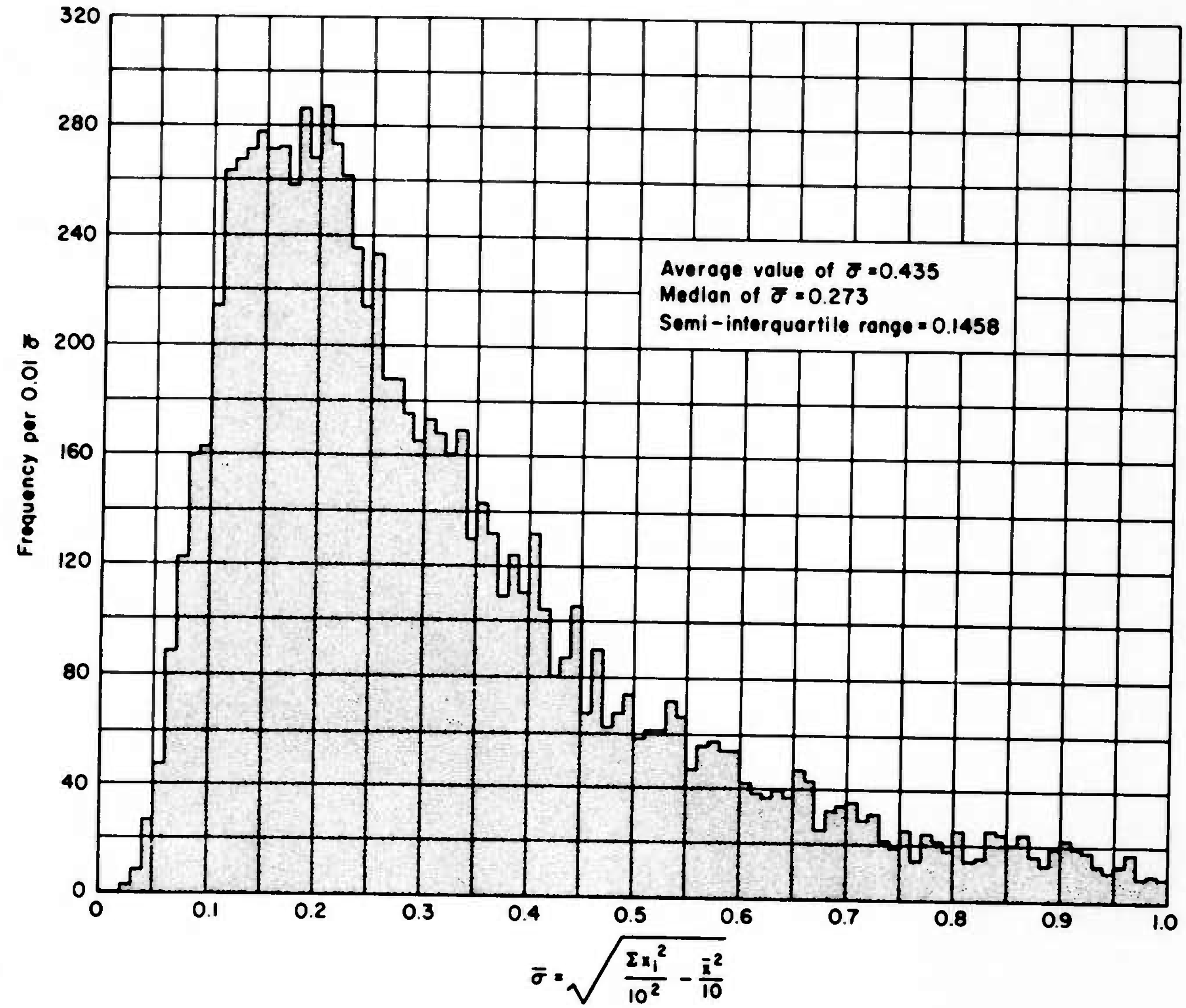

Fig. 2—Frequency distribution for sample standard deviation of sample mean for 10,000 runs of 10 each

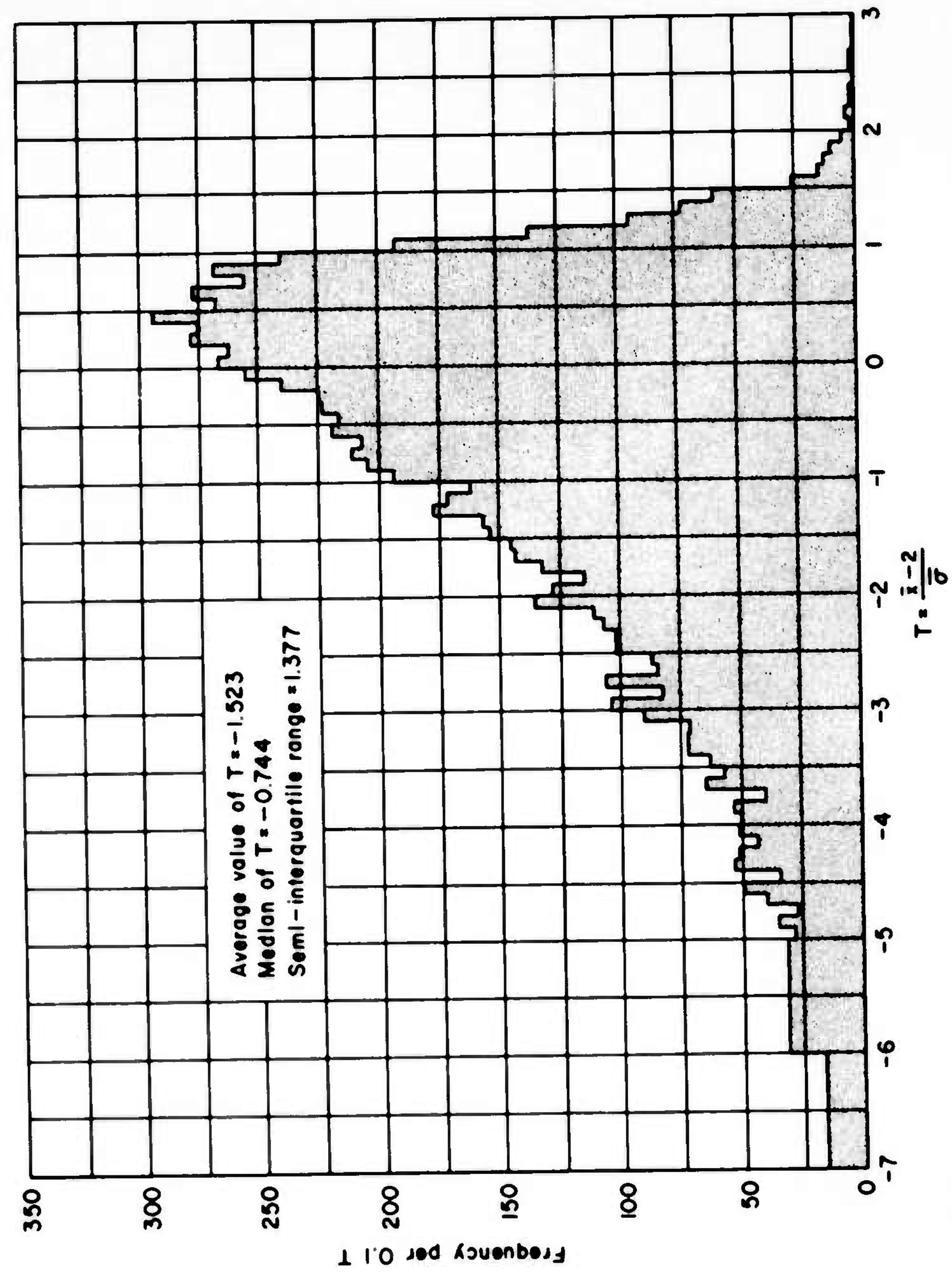

Fig. 3—Frequency distribution of student's ratio for 10,000 runs of 10 each

Table 1

FREQUENCY AND CUMULATIVE DISTRIBUTIONS OF $\bar{x}$ FOR 10,000 SAMPLES OF 10 EACH

Interval for $\bar{x}$		Frequency Distribution	Cumulative Distribution
0	.1	0	0
.1	.2	0	0
.2	.3	0	0
.3	.4	0	0
.4	.5	0	0
.5	.6	0	0
.6	.7	0	0
.7	.8	0	0
.8	.9	0	0
.9	1.0	0	0
1.0	1.1	0	0
1.1	1.2	26	26
1.2	1.3	237	263
1.3	1.4	569	832
1.4	1.5	942	1774
1.5	1.6	1141	2915
1.6	1.7	1112	4027
1.7	1.8	1052	5079
1.8	1.9	865	5944
1.9	2.0	732	6676
2.0	2.1	614	7290
2.1	2.2	488	7778
2.2	2.3	407	8185
2.3	2.4	277	8462
2.4	2.5	257	8719
2.5	2.6	177	8896
2.6	2.7	148	9044
2.7	2.8	132	9176
2.8	2.9	95	9271
2.9	3.0	89	9360
3.0	3.1	72	9432
3.1	3.2	66	9498
3.2	3.3	45	9543
3.3	3.4	53	9596
3.4	3.5	48	9644
3.5	3.6	44	9688
3.6	3.7	23	9711
3.7	3.8	21	9732
3.8	3.9	24	9756
3.9	4.0	24	9780
4.0	4.1	20	9800

$$\bar{x} = \frac{1}{10} \sum_{i=1}^{i=10} x_i$$

Mean = 1.99

Median = 1.79

Semi-interquartile Range = 0.290

Table 1
(Continued)

Interval for $\bar{x}$		Frequency Distribution	Cumulative Distribution
4.1	4.2	13	9813
4.2	4.3	17	9830
4.3	4.4	13	9843
4.4	4.5	7	9850
4.5	4.6	16	9866
4.6	4.7	12	9878
4.7	4.8	13	9891
4.8	4.9	5	9896
4.9	5.0	6	9902
5.0	6.0	37	9939
6.0	7.0	26	9965
7.0	8.0	8	9973
8.0	9.0	7	9980
9.0	10.0	5	9985
10.0	11.0	4	9989
11.0	12.0	3	9992
12.0	13.0	1	9993
13.0	14.0	0	9993
14.0	15.0	1	9994
15.0	16.0	2	9996
16.0	17.0	1	9997
17.0	18.0	1	9998
18.0	19.0	1	9999
19.0	20.0	0	9999
20.0	21.0	0	9999
21.0	22.0	0	9999
22.0	23.0	0	9999
23.0	24.0	0	9999
24.0	25.0	0	9999
25.0	100.0	1	10000

Table 2

FREQUENCY AND CUMULATIVE DISTRIBUTIONS OF $\bar{\sigma}$ FOR 10,000 SAMPLES OF 10 EACH

Interval for $\bar{\sigma}$		Frequency Distribution	Cumulative Distribution
0	.01	0	0
.01	.02	0	0
.02	.03	3	3
.03	.04	8	11
.04	.05	26	37
.05	.06	47	84
.06	.07	88	172
.07	.08	122	294
.08	.09	159	453
.09	.10	162	615
.10	.11	214	829
.11	.12	263	1092
.12	.13	267	1359
.13	.14	271	1630
.14	.15	277	1907
.15	.16	271	2178
.16	.17	272	2450
.17	.18	258	2708
.18	.19	286	2994
.19	.20	263	3257
.20	.21	287	3544
.21	.22	273	3817
.22	.23	261	4078
.23	.24	235	4313
.24	.25	214	4527
.25	.26	233	4760
.26	.27	188	4948
.27	.28	188	5136
.28	.29	175	5311
.29	.30	165	5476
.30	.31	173	5649
.31	.32	168	5817
.32	.33	160	5977
.33	.34	169	6146
.34	.35	130	6276
.35	.36	143	6419
.36	.37	132	6551
.37	.38	109	6660
.38	.39	124	6784
.39	.40	110	6894
.40	.41	132	7026

$$\bar{\sigma} = \sqrt{\frac{1}{100}\sum_{i=1}^{i=10} x_i^2 - \frac{\bar{x}^2}{10}}$$

Mean = 0.435

Median = .273

Semi-interquartile Range = .146

Table 2
(Continued)

Interval for $\bar{\sigma}$		Frequency Distribution	Cumulative Distribution
.41	.42	104	7130
.42	.43	80	7210
.43	.44	87	7297
.44	.45	105	7402
.45	.46	67	7469
.46	.47	90	7559
.47	.48	62	7621
.48	.49	67	7688
.49	.50	74	7762
.50	.51	58	7820
.51	.52	61	7881
.52	.53	61	7942
.53	.54	72	8014
.54	.55	66	8080
.55	.56	47	8127
.56	.57	56	8183
.57	.58	57	8240
.58	.59	54	8294
.59	.60	54	8348
.60	.61	42	8390
.61	.62	38	8428
.62	.63	37	8465
.63	.64	40	8505
.64	.65	37	8542
.65	.66	47	8589
.66	.67	43	8632
.67	.68	25	8657
.68	.69	32	8689
.69	.70	34	8723
.70	.71	35	8758
.71	.72	29	8787
.72	.73	31	8818
.73	.74	21	8839
.74	.75	19	8858
.75	.76	25	8883
.76	.77	14	8897
.77	.78	24	8921
.78	.79	21	8942
.79	.80	18	8960
.80	.81	25	8985
.81	.82	14	8999
.82	.83	15	9014
.83	.84	25	9039
.84	.85	24	9063
.85	.86	20	9083

Table 2
(Continued)

Interval for $\bar{\sigma}$		Frequency Distribution	Cumulative Distribution
.86	.87	24	9107
.87	.88	16	9123
.88	.89	13	9136
.89	.90	19	9155
.90	.91	22	9177
.91	.92	19	9196
.92	.93	18	9214
.93	.94	13	9227
.94	.95	11	9238
.95	.96	13	9251
.96	.97	17	9268
.97	.98	9	9277
.98	.99	10	9287
.99	1.00	9	9296
1.00	1.1	100	9396
1.1	1.2	90	9486
1.2	1.3	69	9550
1.3	1.4	59	9614
1.4	1.5	42	9656
1.5	1.6	31	9687
1.6	1.7	45	9732
1.7	1.8	27	9759
1.8	1.9	22	9781
1.9	2.0	22	9803
2.0	2.1	19	9822
2.1	2.2	11	9833
2.2	2.3	13	9846
2.3	2.4	6	9852
2.4	2.5	11	9863
2.5	2.6	11	9874
2.6	2.7	12	9886
2.7	2.8	3	9889
2.8	2.9	9	9898
2.9	3.0	11	9909
3.0	3.1	6	9915
3.1	3.2	5	9920
3.2	3.3	6	9926
3.3	3.4	1	9927
3.4	3.5	4	9931
3.5	3.6	5	9936
3.6	3.7	3	9939
3.7	3.8	3	9942
3.8	3.9	3	9945
3.9	4.0	1	9946

Table 2
(Continued)

Interval for $\bar{\sigma}$		Frequency Distribution	Cumulative Distribution
4.0	4.1	3	9949
4.1	4.2	1	9950
4.2	4.3	1	9951
4.3	4.4	1	9952
4.4	4.5	3	9955
4.5	4.6	4	9959
4.6	4.7	1	9960
4.7	4.8	2	9962
4.8	4.9	1	9963
4.9	5.0	4	9967
5.0	5.5	6	9973
5.5	6.0	8	9981
6.0	6.5	2	9983
6.5	7.0	1	9984
7.0	7.5	1	9985
7.5	8.0	0	9985
8.0	8.5	3	9988
8.5	9.0	1	9989
9.0	9.5	1	9990
9.5	10.0	3	9993
10.0	15.0	5	9998
15.0	20.0	1	9999
20.0	25.0	1	10000

Table 3

FREQUENCY AND CUMULATIVE DISTRIBUTIONS OF STUDENT'S RATIO, T, FOR 10,000 SAMPLES OF 10 EACH

Interval for T		Frequency Distribution	Cumulative Distribution
-100	-25	4	4
-25	-24	1	5
-24	-23	0	5
-23	-22	2	7
-22	-21	2	9
-21	-20	2	11
-20	-19	2	13
-19	-18	3	16
-18	-17	6	22
-17	-16	5	27
-16	-15	5	32
-15	-14	14	46
-14	-13	19	65
-13	-12	27	92
-12	-11	29	121
-11	-10	50	171
-10	-9	68	239
-9	-8	101	340
-8	-7	112	452
-7	-6	171	623
-6	-5	316	939
-5	-4.9	28	967
-4.9	-4.8	35	1002
-4.8	-4.7	27	1029
-4.7	-4.6	40	1069
-4.6	-4.5	50	1119
-4.5	-4.4	34	1153
-4.4	-4.3	53	1206
-4.3	-4.2	51	1257
-4.2	-4.1	43	1300
-4.1	-4.0	51	1351
-4.0	-3.9	50	1401
-3.9	-3.8	53	1454
-3.8	-3.7	40	1494
-3.7	-3.6	65	1559
-3.6	-3.5	57	1616
-3.5	-3.4	63	1679
-3.4	-3.3	72	1751
-3.3	-3.2	72	1823
-3.2	-3.1	71	1894
-3.1	-3.0	90	1984
-3.0	-2.9	104	2088

$$T = \frac{\bar{x} - 2}{\sigma}$$

Mean = - 1.52

Median = - 0.744

Semi-interquartile Range = 1.37

Table 3
(Continued)

Interval for T		Frequency Distribution	Cumulative Distribution
-2.9	-2.8	82	2170
-2.8	-2.7	106	2276
-2.7	-2.6	84	2360
-2.6	-2.5	86	2446
-2.5	-2.4	101	2547
-2.4	-2.3	101	2648
-2.3	-2.2	107	2755
-2.2	-2.1	111	2866
-2.1	-2.0	135	3001
-2.0	-1.9	128	3129
-1.9	-1.8	115	3244
-1.8	-1.7	132	3376
-1.7	-1.6	143	3519
-1.6	-1.5	145	3664
-1.5	-1.4	154	3818
-1.4	-1.3	156	3974
-1.3	-1.2	177	4151
-1.2	-1.1	171	4322
-1.1	-1.0	161	4483
-1.0	-0.9	193	4676
-0.9	-0.8	205	4881
-0.8	-0.7	211	5092
-0.7	-0.6	207	5299
-0.6	-0.5	219	5518
-0.5	-0.4	216	5734
-0.4	-0.3	223	5957
-0.3	-0.2	224	6181
-0.2	-0.1	240	6421
-0.1	0	255	6676
0	0.1	266	6942
0.1	0.2	262	7204
0.2	0.3	278	7482
0.3	0.4	275	7757
0.4	0.5	294	8051
0.5	0.6	267	8318
0.6	0.7	277	8595
0.7	0.8	255	8850
0.8	0.9	268	9118
0.9	1.0	240	9358
1.0	1.1	192	9550
1.1	1.2	136	9686
1.2	1.3	94	9780
1.3	1.4	72	9852

Table 3
(Continued)

Interval for T		Frequency Distribution	Cumulative Distribution
1.4	1.5	59	9911
1.5	1.6	38	9949
1.6	1.7	15	9964
1.7	1.8	13	9977
1.8	1.9	10	9987
1.9	2.0	5	9992
2.0	2.1	1	9993
2.1	2.2	3	9996
2.2	2.3	1	9997
2.3	2.4	1	9998
2.4	2.5	0	9998
2.5	2.6	1	9999
2.6	2.7	1	10000

Appendix

DERIVATION OF BASIC PROBABILITY DENSITY FUNCTION FOR THE FLUX

There are several ways of recording flux or some function of the flux, such as dosage, in a Monte Carlo code. One common scheme used in the RAND neutron and gamma ray transport codes, as well as in codes developed by Los Alamos and Sandia Corporation,[3] is to record certain functions of a particle as it passes through a given plane in the sample space. The parameters which need be known are particularly the statistical weight of the particle, its energy, and the angle of crossing with respect to a normal to the plane.

Generally speaking, particles will cross any one point on the plane with many values of statistical weight, energy, and angle. Here we will consider the simplest case which approximates these complex conditions. In particular, we assume that all particles have the same weight, the same energy, and that their directions are distributed isotropically in space.

Let θ be the angle of a particle with respect to the normal to the recording plane. The probability density function for a particle at angle θ is simply

$$dp = \tfrac{1}{2} \sin\theta \, d\theta \quad 0 < \theta < \pi. \tag{A1}$$

If we consider particles which cross a fixed area of the recording plane, then the probability density function for a particle crossing this area of the plane at angle θ is

$$dp = \sin\theta \, |\cos\theta| \, d\theta \quad 0 < \theta < \pi. \qquad (A2)$$

In the Monte Carlo code, when a particle actually crosses the recording plane at an angle θ, the flux or dosage contributed by that particle is then given

$$\varphi = \frac{1}{|\cos\theta|}, \qquad (A3)$$

aside from a constant of proportionality, which we may neglect. If we eliminate θ from Eqs. (2) and (3), we obtain the probability density function for the flux per particle crossing, which is

$$dp = \frac{2d\varphi}{\varphi^3} \qquad \varphi = 1 \text{ to } \infty \qquad (A4)$$

$$= 0 \qquad \varphi = 0 \text{ to } 1.$$

It is this equation which the whole substance of this memorandum is based upon. Note that the expected value of the flux in Eq. (4) is 2. This simply means that the average flux of particles crossing a plane in an isotropic field is twice the value which would be obtained by having the same number of particles cross at normal incidence, and this is in fact a well-known result.

ADDENDUM ON RUNNING TIME

The running time, including printing, for the 10,000 samples of 10 each was $4\frac{1}{2}$ minutes on the 7090.

REFERENCES

1. Kahn, Herman, "Random Sampling (Monte Carlo) Techniques in Neutron Attenuation Problems--II", *Nucleonics*, May 1950, p. 27, Part II, *Nucleonics*, June 1950, p. 60.

2. Cramer, Harold, *Mathematical Methods of Statistics*, Princeton University Press, 1946.

3. Marcum, J. I., *Neutron Fluxes in Air: A Comparison of Monte Carlo Computations by RAND, Los Alamos and Sandia*, The RAND Corporation, RM-2556-PR, July 1, 1960.