AN ABSTRACT OF THE THESIS OF

KENNETH NORMAN OLSON for the M.S. in Mathematics
(Name) (Degree) (Major)

Date thesis is presented July 26, 1966

Title APPLYING PROBABILITY TO A PARTICULAR HIRING PRACTICE

Redacted for Privacy

Abstract approved

Major professor

Before we apply the laws of probability to hiring practices, a foundation of basic probability theory will be presented. In this presentation a number of theorems related to probability will be proven. These theorems are not necessarily applicable to the problem which follows; however, they are basic to probability theory and any discussion of probability would be incomplete without them.

The particular problem which follows the discussion of probability theory is only one of many that could have been chosen. The purpose of the chapter is to show that probability can be applied to hiring practices of an employer who uses a method similar to the one mentioned. Its practicality is also shown by arriving at a function which in turn can be programmed for a computer.

As the reader continues into Chapter I, he should keep in mind that the use of only basic probability theory was intended; therefore, the greater part of the discussion will be devoted to the science of

counting and, although the solution is stated in terms of probability, the solution will be arrived at mainly through the science of counting.

APPLYING PROBABILITY TO A PARTICULAR HIRING PRACTICE

by

KENNETH NORMAN OLSON

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

June 1967

APPROVED:

Redacted for Privacy

Professor of Mathematics

In Charge of Major

Redacted for Privacy

Chairman of Department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented July 26, 1966

Typed by Carol Baker

TABLE OF CONTENTS

APPLYING PROBABILITY TO A PARTICULAR HIRING PRACTICE

CHAPTER I

THE THEORY OF COUNTING

One problem arises more often than any other in our study of counting. That is the problem of determining the size, number of elements, of a set of n-tuples of objects of some nature. Because of its frequent occurrence and its applicability to other ideas of counting it is regarded as the basic principle of counting and is developed as follows: If X is a set of ordered n-tuples of objects, we determine the size of X by first determining the number, m_1, of objects that may be used as the first component of an n-tuple. Next determine the number, m_2, of objects that may be used as the second component of an n-tuple where the first component has already been chosen. Then determine the number, m_3, of objects that may be used as the third component where the first two components have already been chosen. We continue in this manner until we have determined the number, m_n, of objects that may be used as the n^{th} component of an n-tuple where the first $(n-1)$ components have already been chosen. The size of the set X of n-tuples is the product of the numbers $m_1, m_2, m_3, \cdots, m_n$.

A permutation is one form of the n-tuple. It is an arrangement of n different objects in a given order. In determining the number of permutations or n-tuples there would be n choices for the first component, n-1 for the second, n-2 for the third, continuing until all n spaces are filled. The number of permutations will then be the product $n(n-1)(n-2)\cdots 1$. This number is named $n!$ and reads "n factorial."

An important application of the above is to the problem of finding the number of subsets of a set. Suppose we are to find how many possible subsets can be formed from a set X if the set contains n elements. We may first find the number of subsets of size k that can be found where $k = 1, 2, \cdots, n$. The sum of these will be the total number of subsets. Let x_k be the number of subsets of size k. Each subset of size k would have $k!$ permutations; therefore, $x_k \cdot k!$ would be the number of n-tuples of size k resulting from our original set X. By our basic principle, $n(n-1)(n-2)\cdots(n-k+1)$ is also the number of n-tuples which can be drawn from set X. Therefore,

$$x_k \cdot k! = n(n-1)(n-2)\cdots(n-k+1)$$

$$x_k = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

$$x_k = \frac{n(n-1)(n-2)\cdots(n-k+1)(n-k)!}{k!\,(n-k)!}$$

$$x_k = \frac{n!}{k!\,(n-k)!} \quad .$$

The expression $\frac{n!}{k!\,(n-k)!}$ occurs so frequently that we replace it with a symbol, namely $\binom{n}{k}$, and is the number of subsets of size k that may be formed from a set of size n.

Another problem of counting is finding the number of partitions of a set. A <u>partition</u> of a set X is a subdivision of the set into disjoint subsets, called cells of the partition, which exhaust the whole set. That is, we need to know the number of ways in which one can partition a set of size n into r cells so that the first cell has size n_1, the second cell has size n_2 and so on, where $n_1 + n_2 + n_3 + \cdots + n_r = n$.

For the first cell of n_1 items there are n items available; therefore, there are $\binom{n}{n_1}$ ways in which the elements in the first cell can be selected. There are now $n-n_1$ items available from which to select the n_2 items that go into the second cell, so there are $\binom{n-n_1}{n_2}$ ways in which to select the elements for the cell of n_2 items. Continuing in this manner, we determine that the elements in the r^{th} cell containing n_r items can be selected in $\binom{n-n_1-n_2-n_3-\cdots-n_{r-1}}{n_r}$ ways. The product of these expressions is the number of ways in which a set of size n can be partitioned into r cells and would be

$$\binom{n}{n_1} \binom{n-n_1}{n_2} \binom{n-n_1-n_2}{n_3} \cdots \binom{n-n_1-n_2-\cdots-n_{r-1}}{n_r}.$$

However,

$$\binom{n}{n_1} = \frac{n!}{n_1!\,(n-n_1)!}, \quad \binom{n-n_1}{n_2} = \frac{(n-n_1)!}{n_2!\,(n-n_1-n_2)!}, \quad \binom{n-n_1-n_2}{n_3} = \frac{(n-n_1-n_2)!}{n_3!\,(n-n_1-n_2-n_3)!}.$$

Therefore,

$$\binom{n}{n_1}\binom{n-n_1}{n_2} = \frac{n!\,(n-n_1)!}{n_1!\,n_2!\,(n-n_1)!\,(n-n_1-n_2)!} = \frac{n!}{n_1!\,n_2!\,(n-n_1-n_2)!}$$

and

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3} = \frac{n!\,(n-n_1-n_2)!}{n_1!\,n_2!\,n_3!\,(n-n_1-n_2)!\,(n-n_1-n_2-n_3)!}$$

$$= \frac{n!}{n_1!\,n_2!\,n_3!\,(n-n_1-n_2-n_3)!} \quad .$$

Continuing in this manner, we get for our final expression

$\dfrac{n!}{n_1!\,n_2!\,n_3!\cdots n_r!}$, which we denote by the symbol $\binom{n}{n_1, n_2, n_3, \cdots, n_r}$.

The above partition is an ordered partition. An ordered partition is the case where two partitions with identical cells are considered to be two different partitions if the order in which the cells appear is changed. In determining the number of unordered partitions, the case where two partitions as described above are considered the same, we first determine in how many ways one partition could be written by merely changing the order of the cells.

Suppose a partition of a set X contains r cells. Let x_r be the number of unordered partitions of that set. Then $r!$ would be the number of ways each of the unordered partitions could be written by merely permuting the cells. Therefore, $x_r \cdot r!$ would be the number of ordered partitions. It follows that

$$x_r \cdot r! = \binom{n}{n_1, n_2, n_3, \cdots, n_r}$$

and

$$x_r = \binom{n}{n_1, n_2, n_3, \cdots, n_r}\frac{1}{r!}$$

which is the number of ways in which a set of size n can be partitioned into r unordered cells.

CHAPTER II

BASIC PROBABILITY THEORY

Probability is defined as follows: First determine U, the possibility set, that is the set of all possible events. To each element of U assign a positive number for a measure (weight), $m(x)$, so that the sum of the weights assigned is one. Then the measure of a set, $m(X)$, is the sum of the weights of its elements. Find the truth set of the statement, p, under consideration and the measure of this set is the probability of statement p and is denoted $Pr[p]$.

Before considering any particular probability problems we have to develop a few properties of probability measure. Since the probability of a statement is obtained directly from the measure, $m(X)$, of its truth set, we may develop properties of $m(X)$ and translate to statements about probability.

1. $m(X) = 1$ if and only if $X = U$.

   By definition we assigned a positive measure to each element of U such that the sum of the elements is 1. Since U contains all the elements, it has measure 1. Assume $X \neq U$. Then X must be a proper subset of U; therefore, U contains elements that are not contained in X. Then

$m(X) = m(U)$ minus the weight of those elements which is positive by definition; therefore, $m(X) = 1$ minus some positive number and thus is less than one.

2. $m(X) = 0$ if and only if $X = \phi$.

If $X = \phi$, then X has no elements and therefore $m(X) = 0$. Assume that $X \neq \phi$. Then X has at least one element and by definition this element has a positive measure and $m(X)$ is positive.

A consequence of the proofs of the above two properties is the following property.

3. $0 \leq m(X) \leq 1$ for any set X.

4. For two sets X and Y, $m(X \cup Y) = m(X)+m(Y)$ if and only if X and Y are disjoint.

$m(X) + m(Y)$ is the sum of weights of the elements of X added to the sum of the weights of the elements of Y. If X and Y are disjoint, then the weight of every element of $X \cup Y$ is added once and only once and $m(X) + m(Y) = m(X \cup Y)$.

If X and Y are not disjoint, then the weight of every element contained in both X and Y is added twice;

that is, every element of $X \cap Y$ has its weight added twice in the sum $m(X) + m(Y)$. Thus this sum is greater than $m(X \cup Y)$ by the amount $m(X \cap Y)$. By properties 1 and 2, if $X \cap Y$ is not empty then $m(X \cap Y) > 0$. Hence $m(X) + m(Y) > m(X \cup Y)$. Our proof also shows that in general we have the following property.

5. For any two sets X and Y, $m(X \cup Y) = m(X) + m(Y) - m(X \cap Y)$.

6. $m(\overline{X}) = 1 - m(X)$.

In the statement of property 5, let $Y = \overline{X}$, the complement of X.

$$m(X \cup \overline{X}) = m(X) + m(\overline{X}) - m(X \cap \overline{X})$$

$$m(U) = m(X) + m(\overline{X}) - m(\phi)$$

$$1 = m(X) + m(\overline{X}) - 0$$

$$m(\overline{X}) = 1 - m(X) .$$

Translating the above properties to properties of probabilities we arrive at the following:

1. $Pr[p] = 1$ if and only if p is logically true.

2. $Pr[p] = 0$ if and only if p is logically false.

3. $0 \leq Pr[p] \leq 1$ for any statement p.

4. $Pr[p \vee q] = Pr[p] + Pr[q]$ if and only if p and q are inconsistent.

5. $Pr[p \vee q] = Pr[p] + Pr[q] - Pr[p \wedge q]$ for any two statements p and q.

6. $Pr[\sim p] = 1 - Pr[p]$.

In our definition of probability the method of assigning weights to the possibility set was left open. The assigning of these weights depends upon the likelihood of each of the possibilities which varies from one situation to another.

One instance where the assigning of weights can be formulated is the situation where all possibilities are equally likely and we have what we refer to as the case of the equiprobable measure. In this case, all possibilities will be assigned the same weight.

Consider a situation where U has n elements and all are equally likely. Since the sum of the n measures is 1, each element has measure $\frac{1}{n}$. Therefore, for any subset X of U, $m(X)$ is $\frac{r}{n}$ where r is the number of elements in X.

CHAPTER III

A PARTICULAR HIRING PRACTICE

Consider now an employer who in his hiring practices prefers to have the applicants interviewed and rated by a number of people on his staff. This could be done by having all interviewers interview all applicants and having each interviewer rank the applicant he feels is most qualified first and to continue to rank them from there as he feels they are qualified. Of the many requirements the employer could now select, to decide which of the applicants he would hire, let us suppose he chooses the following one.

An applicant will be hired if he is ranked first by at least two of the interviewers and at least second by a third interviewer. It seems quite obvious, however, that with a small number of applicants and a large number of interviewers a number of applicants could meet the requirements even though the interviewers were to rank the applicants at random. In fact, it will be shown that with five applicants and nine interviewers someone must receive at least two firsts and a second.

It is also obvious that the same problem would result if the employer should change his requirements to three firsts, one first and two seconds or to some other requirement. Therefore, the employer is interested in determining how many interviewers are needed with different numbers of applicants so that the probability of someone receiving two firsts and at least a second when ranked at random is kept reasonably low. This will assure him that a person meeting the requirements is considered qualified.

The question now arises as to what this probability should be. To help determine this let us suppose that this practice is conducted four times a year. Over a period of ten years it will be used 40 times. Therefore, if we choose a number of interviewers such that the probability is $1/40$ it is likely that someone may be hired during the ten years even though they may have been rated at random. Let us suppose he decides that a probability of 0.01 would be sufficiently small. This will not guarantee that he will avoid a selection by chance alone, without regard to qualifications, but it reduces the probability to one for every 100 times used or one in every 25 years.

Our problem is then to determine, with n applicants and m interviewers, the probability that someone may be rated first by two interviewers and at least second by a third when the ratings are done at random.

Our statement p is "Someone will receive two firsts and at least one second." This problem will then be solved by determining $Pr[\sim p]$ and by the property $Pr[p] = 1 - Pr[\sim p]$.

Consider first the specific case of three interviewers and n applicants.

Step i. Determine U. U is the set of all possible sets of ratings. An element of U is then a set containing one order of ratings from each interviewer.

Step ii. Determine the measure of each element of U. Since all sets of ratings are equally likely this is the case of the equiprobable measure. We therefore determine the size of U. The total number of sets of reports, elements of U, possible is $(n!)^3$. This was found by realizing that after the first interviewer rated the applicants in $n!$ ways each succeeding interviewer could also rate in $n!$ ways. For three interviewers we then arrive at $(n!)^3$ sets of ways. The measure of each element is then $\frac{1}{(n!)^3}$.

Step iii. Determine the truth set for $\sim p$.

First we have this happening when no one receives two firsts. How many ways can this happen ?

First interviewer has n choices for first.

Second interviewer has $n-1$ choices for first.

Third interviewer has $n-2$ choices for first.

After which each has $(n-1)!$ ways of rating the rest of the applicants. Therefore, $n(n-1)(n-2)[(n-1)!]^3$ is the number of ways in which no one receives two firsts.

Second, someone may receive two firsts but not another first or second from the third interviewer. To determine in how many ways this can happen let A, B and C be the three interviewers.

	A	B	C
Choices for first	n	1	$n-1$
Choices for second	$n-1$	$n-1$	$n-2$
Choices for third	$n-2$	$n-2$	$n-2$
Choices for fourth	$n-3$	$n-3$	$n-3$
.	.	.	.
.	.	.	.
.	.	.	.

Assuming A and B had rated the same person first, B had only one choice for first after A had chosen. If no one was to receive three firsts C had $n-1$ choices for first and if the first choice of A and B was not to receive a second, C had $n-2$ choices for second. Therefore, when A and B had the same person first we had $n(n-1)(n-1)(n-1)(n-2)[(n-2)!]^3$ sets of ratings satisfying $\sim p$. However, A and C could have rated the same man first, or likewise, B and C; therefore, three times the above number of ratings actually exist where someone receives two firsts but not at least another second. Adding the number of ways no one receives two firsts and the number of ways someone receives two firsts but not at least another second we have $n(n-1)(n-2)[(n-1)!]^3 + 3n(n-1)^3(n-2)[(n-2)!]^3$ elements in our truth set for $\sim p$. Since the measure of each element was $\frac{1}{(n!)^3}$, the measure of the truth set will be

$$\frac{n(n-1)(n-2)[(n-1)!]^3 + 3n(n-1)^3(n-2)[(n-2)!]^3}{(n!)^3}$$

which is the probability of $\sim p$. Therefore,

$$\begin{aligned}
\Pr[p] &= 1 - \frac{[(n-2)!]^3[n(n-1)^4(n-2) + 3n(n-1)^3(n-2)]}{n^3(n-1)^3[(n-2)!]^3} \\
&= 1 - \frac{n^2-4}{n^2} \\
&= \frac{4}{n^2} .
\end{aligned}$$

Example 1. Let $n = 3$

$$\Pr[p] = \frac{4}{3^2} = \frac{4}{9}$$

Example 2. Let $n = 10$

$$\Pr[p] = \frac{4}{10^2} = \frac{1}{25}$$

Consider now, the general case of m interviewers and n applicants.

Step i. Determine U. U, as before, is the set of all sets of ratings.

Step ii. Determine the measure of each element of U. Since each of the m interviewers can rank the applicants in $n!$ ways the total number of sets of ratings is $(n!)^m$. The measure of each element is then $\frac{1}{(n!)^m}$.

Step iii. Determine the truth set for $\sim p$. The statement $\sim p$ is "No one receives two firsts and at least one second." This can happen in a number of ways.

First, this can happen when no one receives two firsts, as follows: Let $A_1, A_2, A_3, \cdots, A_m$ be the m interviewers. As before, the number of choices for each place, when no one receives two firsts, is shown below.

	A_1	A_2	A_3	A_4	$A_5 \cdots$	A_m
first choice	n	$n-1$	$n-2$	$n-3$	$n-4$	$n-(m-1)$
second choice	$n-1$	$n-1$	$n-1$	$n-1$	$n-1$	$n-1$
third choice	$n-2$	$n-2$	$n-2$	$n-2$	$n-2$	$n-2$
fourth choice	$n-3$	$n-3$	$n-3$	$n-3$	$n-3$	$n-3$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

Therefore, no one will receive two firsts in

$n(n-1)(n-2)\cdots[(n-(m-1)][(n-1)!]^m$ ways.

Second, this could happen when one person receives two firsts but not another first or second. Below we have assumed the person received his two firsts from A_1 and A_2 and have determined the number of possible choices just as we did earlier in our specific example with three interviewers.

	A_1	A_2	A_3	A_4	$A_5 \cdots$	A_m
first choice	n	1	n-1	n-2	n-3	n-(m-2)
second choice	n-1	n-1	n-2	n-2	n-2	n-2
third choice	n-2	n-2	n-2	n-2	n-2	n-2
fourth choice	n-3	n-3	n-3	n-3	n-3	n-3
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

However, we assumed A_1 and A_2 rated the same man first. It could have been any pair of interviewers so we must multiply the product of our above expressions, which is the number of ratings possible for each pair, times the number of possible pairs. To find the number of possible pairs we determine the number of ways the m raters can be partitioned into a group of two and another group of $m-2$ which is $\binom{m}{2,\, m-2}$. Thus we arrive at the total

$$\frac{m!}{2!\,(m-2)!}\, n(n-1)(n-2)\cdots[n-(m-2)]\,(n-1)^2(n-2)^{m-2}[(n-2)!]^m \; .$$

Third, this can happen when two persons each receive two firsts but neither receives another first or second. The number of choices for each place by each rater making this possible is shown below.

	A_1	A_2	A_3	A_4	A_5	A_6	$A_7 \cdots$	A_m
first choice	n	1	n-1	1	n-2	n-3	n-4	n-(m-3)
second choice	n-2	n-2	n-2	n-2	n-3	n-3	n-3	n-3
third choice	n-2	n-2	n-2	n-2	n-2	n-2	n-2	n-2
fourth choice	n-3	n-3	n-3	n-3	n-3	n-3	n-3	n-3
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

The product of these is the number of ratings when A_1 and A_2 rate the same person first and A_3 and A_4 rate the same person first. We must now multiply this product by the number of ways we can select two pairs out of m things. The order of the two pairs is of no consequence to us so we now determine the number of unordered partitions of m things into three cells of sizes 2, 2, and $m-4$ which is $\binom{m}{2,2,m-4}\frac{1}{2!}$. Therefore,

$$\frac{m!}{2!\,2!\,(m-4)!\,2!}\,n(n-1)(n-2)\cdots[n-(m-3)]\,(n-2)^4(n-3)^{m-4}[(n-2)!]^m$$ is the number of ways two persons might each get two firsts but not another first or second.

Continuing in this manner, three persons receiving two firsts but not another first or second can happen in

$$\frac{m!}{2!\,2!\,2!\,(m-6)!\,3!}\,n(n-1)(n-2)\cdots[n-(m-4)]\,(n-3)^6(n-4)^{m-6}[(n-2)!]^m$$

ways.

Similarly, for four persons, the number of possibilities will be

$$\frac{m!}{2!\,2!\,2!\,2!\,(m-8)!\,4!}\,n(n-1)(n-2)\cdots[n-(m-5)]\,(n-4)^8(n-5)^{m-8}[(n-2)!]^m .$$

For five persons, the number of possibilities will be

$$\frac{m!}{2!\,2!\,2!\,2!\,2!\,(m-10)!\,5!}\,n(n-1)(n-2)\cdots[n-(m-6)]\,(n-5)^{10}(n-6)^{m-10}[(n-2)!]^m .$$

Consider now the general case where exactly r applicants receive two firsts but none of these r applicants receive at least another second. Setting a chart up as before, $2r$ raters are needed in order that r people get two firsts. A_{2r} would have rated first the same applicant as A_{2r-1}; therefore, A_{2r} has only one choice possible for first. A_{2r-1} would have $r-1$ of the n choices unavailable to him for first since $r-1$ pairs were formed through rater A_{2r-2}. A_{2r+1} would have one less first place choice than A_{2r-1} and each succeeding rater would have one less choice than the previous rater. The last rater, A_m, would not be able to choose any of the r applicants who have two firsts. He also would not be able to choose any of the first choices of the raters between A_{2r}

and himself which is $m-2r-1$ more choices not available. He therefore has $n-[m-(r+1)]$ choices available for first.

In the second position, raters $A_1, A_2, \cdots, A_{2r}$ have $n-r$ choices available if none of the r applicants with two firsts are to receive a second. The remaining raters have the same r choices and their own first choice unavailable leaving $n-(r+1)$ choices for second place.

The remaining places have no bearing on the hiring; therefore, the only choices unavailable are the ones each rater has chosen previously.

	A_1	A_2	$A_3 \cdots$	A_{2r-1}	A_{2r}	A_{2r+1}	$A_{2r+2} \cdots$	A_m
first choice	n	1	$n-1$	$n-(r-1)$	1	$n-r$	$n-(r+1)$	$n-[m-(r+1)]$
second choice	$n-r$	$n-r$	$n-r$	$n-r$	$n-r$	$n-(r+1)$	$n-(r+1)$	$n-(r+1)$
third choice	$n-2$	$n-2$	$n-2$	$n-2$	$n-2$	$n-2$	$n-2$	$n-2$
fourth choice	$n-3$	$n-3$	$n-3$	$n-3$	$n-3$	$n-3$	$n-3$	$n-3$
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

The product of these numbers of choices is the number of ratings possible for this pairing of raters. We now determine in how many ways we can select r pairs out of the m things if the order of the r pairs is inconsequential. We have r cells of size two and a remaining cell of size $m-2r$; thus we arrive at $\frac{m!}{(2!)^r(m-2r)!\,r!}$ ways of pairing the m raters. We can now determine the number of ways r applicants can each receive two firsts while none receive at least another second with m raters. This number will be

$$\frac{m!}{(2!)^r(m-2r)!\,r!}n(n-1)\cdots[n-[m-(r+1)]]\,(n-r)^{2r}[n-(r+1)]^{m-2r}[(n-2!)]^m .$$

The largest number of persons that could receive two firsts when there are m interviewers is the greatest integer part of $\frac{m}{2}$. This greatest integer part is to be noted by the symbol $\left[\frac{m}{2}\right]$. If we continue to count possibilities beyond the case where five persons receive two firsts but not another first or second we will eventually arrive at the case where $\left[\frac{m}{2}\right]$ persons receive two firsts. Following the pattern, the number of possibilities will be

$$\frac{m!}{(2!)^{\left[\frac{m}{2}\right]}(m-\left[\frac{m}{2}\right]2)!\left[\frac{m}{2}\right]!}n(n-1)(n-2)\cdots[n-\{m-(\left[\tfrac{m}{2}\right]+1)\}](n-\left[\tfrac{m}{2}\right])^{2\left[\frac{m}{2}\right]}$$

$$\times\ [n-(\left[\tfrac{m}{2}\right]+1)]^{m-2\left[\frac{m}{2}\right]}\ [(n-2)!]^m .$$

The sum of all these numbers is the number of possibilities in the truth set of $\sim p$. Remembering that each element has measure $\frac{1}{(n!)^m}$, the measure of the truth set of $\sim p$, which is $Pr[\sim p]$, is given by the following expression.

$$\frac{1}{(n!)^m}\Bigg(n(n-1)(n-2)\cdots[n-(m-1)][(n-1)!] +$$

$$\frac{m!}{2!(m-2)!}n(n-1)(n-2)\cdots[n-(m-2)](n-1)^2(n-2)^{m-2}[(n-2)!]^m +$$

$$\frac{m!}{2!2!2!(m-4)!}n(n-1)(n-2)\cdots[n-(m-3)](n-2)^4(n-3)^{m-4}[(n-2)!]^m + \cdots$$

$$\cdots + \frac{m!}{(2!)^r(m-2r)!r!}n(n-1)(n-2)\cdots[n-\{m-(r+1)\}](n-r)^{2r}[n-(r+1)]^{m-2r}[(n-2)!]^m + \cdots$$

$$\cdots + \frac{m!}{(2!)^{\left[\frac{m}{2}\right]}(m-\left[\frac{m}{2}\right]2)!\left[\frac{m}{2}\right]!}n(n-1)(n-2)\cdots\left[n-\left\{m-\left(\left[\frac{m}{2}\right]+1\right)\right\}\right]$$

$$\times\ \left(n-\left[\frac{m}{2}\right]\right)^{2\left[\frac{m}{2}\right]}\left[n-\left(\left[\frac{m}{2}\right]+1\right)\right]^{m-2\left[\frac{m}{2}\right]}[(n-2)!]^m\Bigg).$$

So that we may write the expression within the parentheses as a summation, we rewrite the first term of the summation as

$$\frac{m!}{(2!)^0 0!(m-0)!} n(n-1)(n-2)\cdots[n-(m-1)](n-0)^0(n-1)^m[(n-2)!]^m .$$

Now factoring out $m![(n-2)!]^m$, expressing the remainder as a summation, reducing our coefficient of the summation and solving for $Pr[p]$ we arrive at the following equation.

$$Pr[p] = 1 - \frac{m!}{n^m(n-1)^m} \sum_{j=1}^{[\frac{m}{2}]+1} \frac{n(n-1)(n-2)\cdots[n-(m-j)][n-(j-1)]^{2(j-1)}(n-j)^{m-2(j-1)}}{2^{j-1}[m-2(j-1)]!(j-1)!} .$$

CHAPTER IV

PROGRAMMING THE PROBABILITY FUNCTION

From the expression derived in Chapter III it is still difficult to draw any conclusions about the probability that, with a given number of applicants and a given number of interviewers rating the applicants at random, an applicant might receive two firsts and at least another second. To do this we will have to compute the Pr[p] for various numbers of applicants and interviewers. This can most easily be done through the use of a computer.

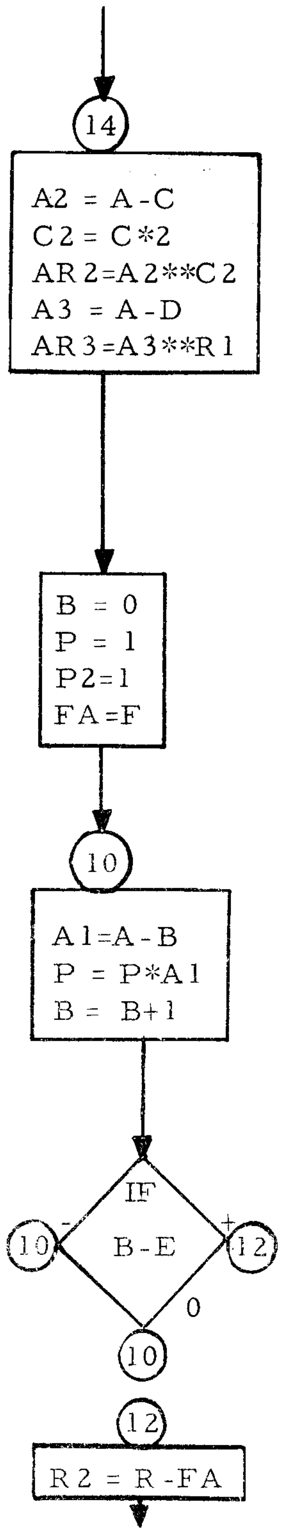
The basic strategy is to give the computer an initial number of applicants, A, and an initial number of interviewers (raters), RA, and have it determine Pr[p]. After printing these values of A, RA, and Pr[p] it will increment RA by one, determine a new value for Pr[p] and print the new values of A, RA and Pr[p]. It will continue incrementing RA and printing data until some designated final value for RA is reached. The computer will then return RA to its initial value, increment A by one, determine Pr[p] and then return to incrementing RA by one until the final value of RA is reached for A+1 applicants. It will continue in this manner, incrementing through all values of RA for each value of A.

The initial and final values of A and RA are given to the computer on a data card following the program enabling us to use the program again later with any new values of A and RA.

The program which follows is written in FORTRAN language. Variables different from those in the expression of Chapter III are used making it necessary to explain the program in detail. To do this the flow chart will be drawn vertically along the left side with the change in variables and a necessary explanation along the right.

Flow Chart	Change of Variables	Explanation
(1) Read nn, A, RA, AF, RF	A = initial value of n RA = initial value of m AF = final value of n RF = final value of m	Instructs computer to read data card. The numbers appearing in the circles are used only as labels so that we can later instruct the computer to return to that point. These are not followed sequentially.
R = RA		This enables us to save our initial RA so that we may return to it after incrementing A each time.
(6) $I = \frac{R}{2}$ S = 0 D = 1 R1 = R F = 0 G = 0 E = R - 1 C = 0	$I = [\frac{m}{2}]$ D = j R1 = m - 2(j - 1) F = 2(j - 1) G = j - 1 E = m - j C = j - 1	When evaluating the summation we begin with j=1. The values given at this time are correct for j=1; however, when we move to j=2, each will have to be incremented in its own way as is done later in the program.

S represents the entire summation and changes as we increase j. We begin with S = 0 and add to it the value found for each j.



$A2 = n-(j-1)$
$C2 = 2(j-1)$
$AR2 = (n-(j-1))^{2(j-1)}$
$A3 = n-j$
$AR3 = (n-j)^{m-2(j-1)}$

AR2 and AR3 are two expressions in the numerator which are first evaluated for j=1. After all expressions have been evaluated for j=1 we will return and evaluate for j=2, and so on.

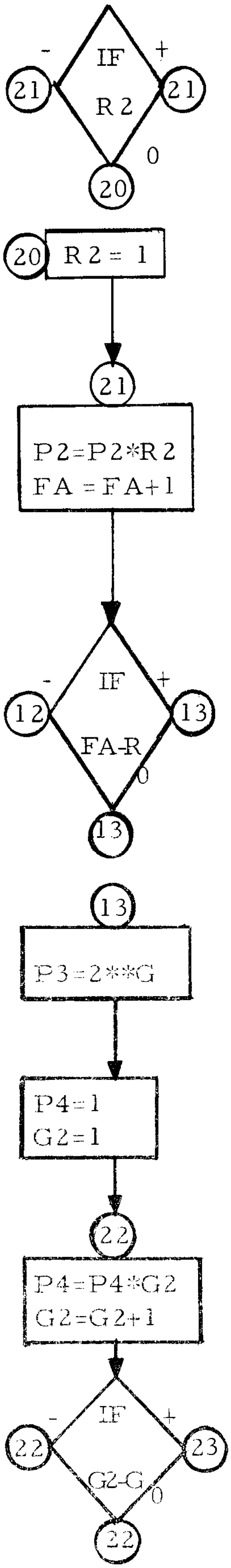
$FA = 2(j-1)$

B, P and P2 will be used to increment and determine $n(n-1)\cdots(n-(m-j))$

$A1 = n, n-1, \cdots, n-(m-j)$
$P = n(n-1)\cdots(n-(m-j))$

This cycle enables us to determine P, the final part of the numerator.

$R2 = m-2(j-1)$



Before we proceed to determine $(m-2(j-1))!$ for values of $m-2(j-1)$ other than zero we must instruct the computer to replace 0! by 1.

$P2=(m-2(j-1))!$

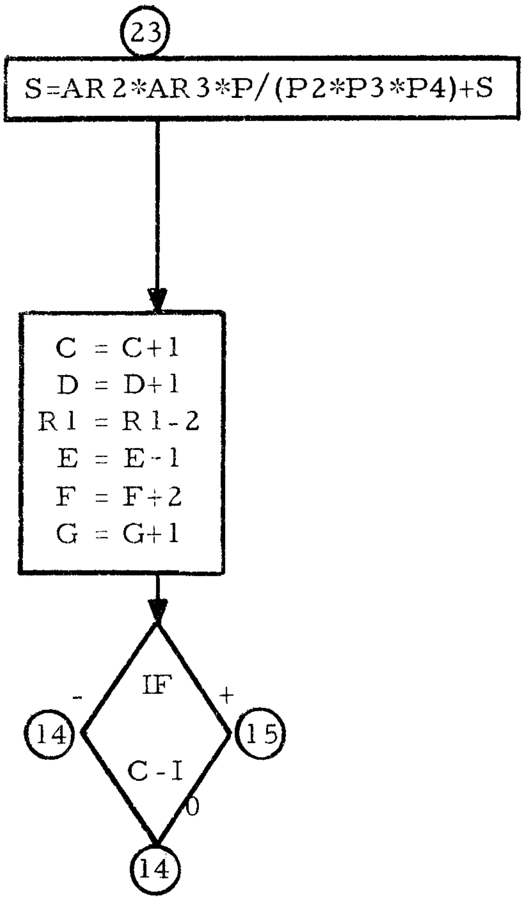
This cycle determines P2, a part of the denominator, for values other than zero.

$P3=2^{j-1}$

Enables us to increment so that we may determine $(j-1)!$

$P4=(j-1)!$

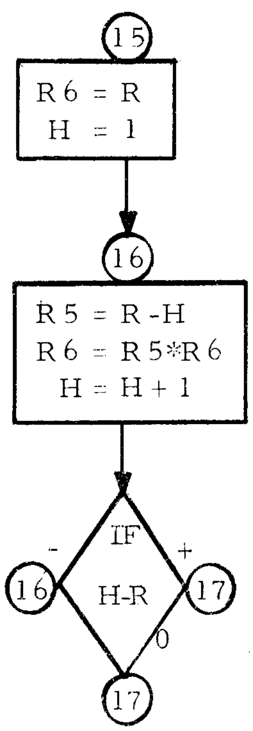
This cycle evaluates P4, which completes the denominator.



When the computer first reaches this point, it has evaluated the summation for j=1.

These are the increments necessary to proceed to the next value of j.

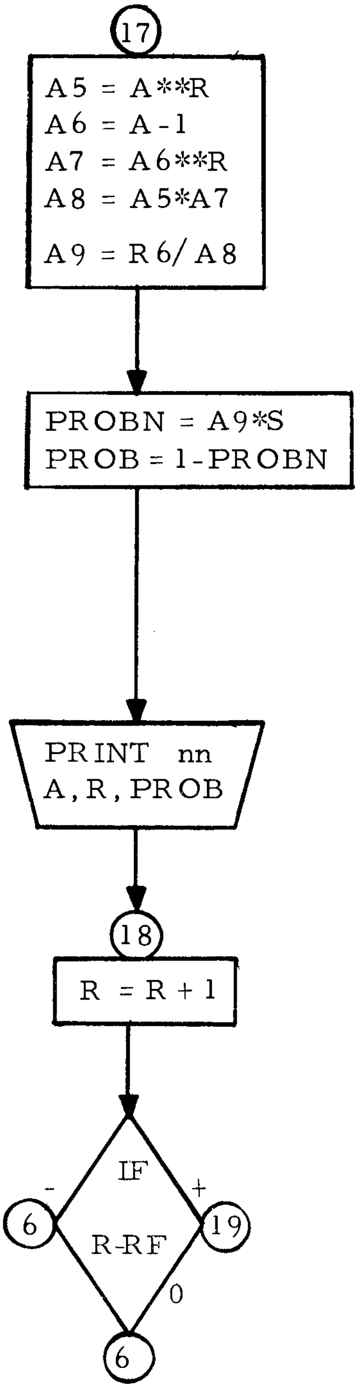
Tests on the value of j. If $j \leq [\frac{m}{2}] + 1$ go back to 14 and solve for S. If $j > [\frac{m}{2}] + 1$ move out of the cycle.



R6 = m

R6 = m!

This cycle determines R6, the numerator of the coefficient of the summation.



$A5 = n^m$
$A6 = n-1$
$A7 = (n-1)^m$
$A8 = n^m(n-1)^m$
$A9 = \dfrac{m!}{n^m(n-1)^m}$

A8 is the denominator of the coefficient of the summation.

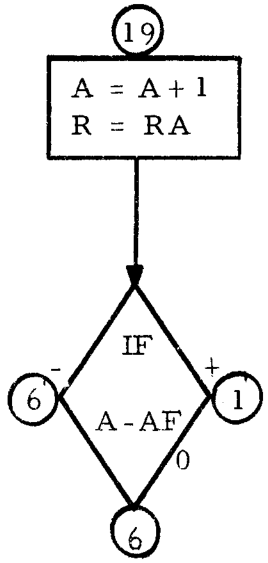
A9 is the coefficient of the summation.

PROBN = Pr[~p]
PROB = Pr[p]

When the computer reaches this point for the first time it determines Pr[p] for initial A and initial RA.

Increments number of raters.

Tests to see if we have reached the final value of R. If not, the computer again evaluates using a new value of R. If we have reached final value of R we move on.



We now increment the number of applicants and restore the number of raters back to initial value.

This is a test to see if we have passed the final number of applicants. If not, we go back and continue. If the final value has been passed we proceed to 1 which stops the computer.

CHAPTER V

CONCLUSION

After determining the probability that having n applicants for a job and m interviewers rating them at random someone might receive at least two firsts and another second when $n = 3, 4, 5, 6, \cdots, 100$ and $m = 3, 4, 5, \cdots, 10$ we have a table of probabilities containing 784 pairs of numbers. This being too large a table to include in its entirety, only the more interesting results are included in the brief table below.

APPLICANTS	INTERVIEWERS	PROBABILITY
3	3	0.44444445
10	3	0.04000000
5	9	1.00000000
19	3	0.01108034
20	3	0.01000000
35	4	0.01256344
40	4	0.00966718
62	5	0.00996066
88	6	0.00986202
100	6	0.00768118
100	7	0.01324901

The first two results have been included because these were the specific examples chosen in the early part of Chapter III. The third was chosen because it was stated in the first part of Chapter III that with five applicants and nine interviewers someone must satisfy the requirements. The probability of one means that it is true that someone will receive two firsts and at least one second.

The remaining eight were chosen because this particular employer decided that a probability of 0.01 would be sufficiently low. It can be seen that with 19 applicants and three interviewers the probability is greater than 0.01. However, with 20 applicants, three interviewers appear to be satisfactory. To keep the probability less than or equal to 0.01, we cannot increase to four interviewers until we have 40 applicants for the job. The rest of these data show when we can increase to five and also to six, but with 100 or less applicants we would never need seven interviewers.

By using the same program and changing only the data card we could determine the probability for any combination of applicants and interviewers. This will enable us to choose a number of interviewers for any number of applicants so that the probability of someone receiving two firsts and at least a second, by chance alone, will be below any number we wish.

BIBLIOGRAPHY

1. Kemeny, John G., J. Laurie Snell and Gerald L. Thompson. Introduction to finite mathematics. Englewood Cliffs, N. J., Prentice Hall, 1956. 372 p.

2. Parzen, Emanuel. Modern probability theory and its applications. New York, John Wiley and Sons, 1960. 464 p.

3. Ryser, Herbert John. Combinatorial mathematics. Buffalo, N. Y., Mathematical Association of America; distr. by Wiley, New York, 1963. 154 p. (The Carus Mathematical Monographs, no. 14)

4. Thompson, Bruce. For-train. Minneapolis, Minnesota. Control Data Corp., 1966. 177 p.