ON THE NUMBER OF SUCCESSES IN INDEPENDENT TRIALS[1]

BY S. M. SAMUELS

Purdue University

1. Summary. The distribution of the number of successes in independent trials is shown to be bell-shaped of every order. The most likely number of successes is "almost uniquely" determined from the mean number and from the mean plus the largest and smallest probability of success on any trial. Bounds on the distribution function of the number of successes are obtained and extended to an infinite number of trials, including the Poisson distribution.

2. The shape of the distribution. We consider n independent trials with probabilities p_i of success on the ith trial. We shall always assume $0 < p_i < 1$; the generalization to $p_i = 0$ or 1 being immediate, though sometimes, (e.g. in (5) and Theorem 1), burdensome to formulate. We let

$$f(k) = \text{probability of } k \text{ successes}$$

$$f_i(k) \doteq \text{probability of } k \text{ successes in all but the } i\text{th trial.}$$

Then

$$f(k) = p_i f_i(k-1) + (1-p_i) f_i(k), \tag{1}$$

$$\textstyle\sum_{k=0}^{n} f(k) z^k = \prod_{i=1}^{n} (1 - p_i + p_i z). \tag{2}$$

An inequality of Newton, proved in [3], p. 104, states that if $a_1, \cdots, a_n$ are any non-zero real numbers (positive or negative) and if $b_0, b_1, \cdots, b_n$ are defined by

$$\textstyle\sum_{k=0}^{n} \binom{n}{k} b_k z^k = \prod_{i=1}^{n} (1 + a_i z), \tag{3}$$

then,

$$b_k^2 > b_{k-1} b_{k+1} \qquad \text{for } k = 1, \cdots, n-1, \tag{4}$$

unless all the a_i's are equal, in which case equality holds.

From (2), it follows that

$$(f(k)/\tbinom{n}{k})^2 > [f(k-1)/\tbinom{n}{k-1}][f(k+1)/\tbinom{n}{k+1}] \quad \text{for } k = 1, \cdots, n-1, \tag{5}$$

unless all the p_i's are equal, in which case equality holds. This inequality has also been derived in [5], p. 88. From (5), we have the weaker inequality

$$f^2(k) > f(k-1) f(k+1) \qquad \text{for } k = 1, \cdots, n-1, \tag{6}$$

Received 4 November 1964; revised 6 February 1965.

[1] This paper is based on the author's Ph.D. thesis at Stanford University. Research supported by the Office of Naval Research under contract No. Nonr-225(28) at Stanford University and by the Aerospace Research Laboratories under Contract No. AF 33(657-11737) at Purdue University.

(i.e. f is log concave). Hence f is unimodal, first increasing, then decreasing, and the mode is either unique or shared by two adjacent integers.

This result can be extended, as follows: Let

$$D_0(k) = f(k)$$
$$D_r(k) = D_{r-1}(k) - D_{r-1}(k-1) \qquad \text{for } r = 1, 2, \cdots.$$

Then

$$\sum_{k=0}^{n+r} D_r(k)z^k = (1-z)^r \prod_{i=1}^{n} (1 - p_i + p_i z) \qquad \text{for } r = 0, 1, \cdots.$$

Hence, by (4)

$$(D_r(k)/\tbinom{n+r}{k})^2 > [D_r(k-1)/\tbinom{n+r}{k-1}][D_r(k+1)/\tbinom{n+r}{k+1}]$$
$$\text{for } r = 0, 1, \cdots, \text{ and } k = 1, \cdots, n+r-1.$$

The weaker versions of these inequalities, namely, $D_r^2(k) > D_r(k-1)D_r(k+1)$, imply that D_{r+1} can have at most one strict sign change between successive strict sign changes of D_r; hence D_r has at most r strict sign changes (i.e. f is "bell-shaped of order r" for every r).

This result, for $r = 1, 2$, was given by Darroch ([1], Theorems 1 and 2).

3. The most likely number of successes. If $p_1 = \cdots = p_n = p$, then the most likely number of successes is well-known to be the integer k (or pair of adjacent integers) such that

$$k/(n+1) \leqq p \leqq (k+1)/(n+1).$$

Equivalently, if for some k, $k \leqq np \leqq k+1$, then the mode is $k(k+1)$ if np is $\leqq (\geqq)\ k + (n-k)/(n+1)$. An analogous result for unequal p_i's follows immediately from two simple lemmas. The first, given in [2], states that

$$(7) \quad f(k+1)/f(k) \quad \text{is strictly increasing in each } p_i \text{ for } k = 0, 1, \cdots, n-1,$$

since

$$(d/dp_i)(f(k+1)/f(k)) = [f_i^2(k) - f_i(k-1)f_i(k+1)]/f^2(k)$$

which is positive by (6). The second states that

$$(8) \qquad kf(k) = \sum_{i=1}^{n} p_i f_i(k-1) \quad \text{for any } k,$$

which follows by differentiating (2) with respect to z and identifying coefficients.

THEOREM 1. *If k is an integer such that*

$$k \leqq \sum_{i=1}^{n} p_i \leqq k+1,$$

and if $p_1 = \min(p_1, \cdots, p_n)$, $p_n = \max(p_1, \cdots, p_n)$, *then*

$$(9) \qquad f_1(k) > f_1(k-1),$$

$$(10) \qquad f_n(k) > f_n(k+1),$$

unless $p_1 = p_n = k/n$ *or* $(k + 1)/n$ *in which cases equality holds in* (9) *or* (10) *respectively.*

PROOF. By hypothesis and by (8), there are indices r and s such that

$$f_r(k-1) \leqq f(k), \tag{11}$$

$$f_s(k) \geqq f(k+1). \tag{12}$$

Hence, by (1),

$$f_r(k-1) \leqq f_r(k), \tag{13}$$

$$f_s(k) \geqq f_s(k+1). \tag{14}$$

If $p_1 + \cdots + p_n > k$, then strict inequality can be obtained in (11), hence in (13) and therefore, by (7), for $r = 1$. If $p_1 + \cdots + p_n = k$, $p_1 < p_n$, then (7) again implies strict inequality in (13) for $r = 1$. If $p_1 = p_n = k/n$, then

$$\begin{aligned} f_1(k) &= \tbinom{n-1}{k}(k/n)^k(1-(k/n))^{n-1-k} \\ &= \tbinom{n-1}{k-1}(k/n)^{k-1}(1-(k/n))^{n-k} = f_1(k-1). \end{aligned}$$

A similar argument at the other end-point completes the proof.

As immediate corollaries to the theorem, we have:

$$p_1 + \textstyle\sum_{i=1}^n p_i > k \Rightarrow f(k) > f(k-1), \tag{15}$$

$$p_n + \textstyle\sum_{i=1}^n p_i < k + 1 \Rightarrow f(k) > f(k+1). \tag{16}$$

Hence, if the hypotheses of both (15) and (16) hold for some k, then k is the unique mode of f. Moreover,

$$k \leqq \textstyle\sum_{i=1}^n p_i \leqq k+1 \Rightarrow f(k-1) < f(k); \qquad f(k+1) > f(k+2), \tag{17}$$

so if the mean number of successes is between k and $k + 1$, then the most likely number of successes is k or $k + 1$.

An improvement on (17)—in fact a complete characterization of the mode in terms of the mean—was obtained by Darroch ([1], Theorem 4) by using the following result: If g is any function defined on $0, 1, \cdots, n$, and if $p_1 + \cdots + p_n = np$, then

$$\begin{aligned} \min_{\substack{r=0,1,\cdots,[n(1-p)] \\ s=0,1,\cdots,[np]}} \textstyle\sum_{k=0}^n g(k)f(k \mid np, r, s) \leqq Eg &= \textstyle\sum_{k=0}^n g(k)f(k) \\ &\leqq \max_{\substack{r=0,1,\cdots,[n(1-p)] \\ s=0,1,\cdots,[np]}} \textstyle\sum_{k=0}^n g(k)f(k \mid np, r, s), \end{aligned} \tag{18}$$

where

$$\begin{aligned} f(k \mid np, r, s) \\ = \tbinom{n-r-s}{k-s}((np-s)/(n-r-s))^{k-s}((n-r-np)/(n-r-s))^{n-r-k} \end{aligned} \tag{19}$$

and $[x]$ denotes the integer part of x. This result was obtained by Tchebicheff [7] for a special function, g, but without making use of the special form of that g. It was also obtained by Hoeffding [4] from a more detailed theorem.

If $k < np < k + 1$ and

$$g(x) = 1 \quad \text{if} \quad x = k + 1$$
$$= -1 \quad \text{if} \quad x = k$$
$$= 0 \quad \text{otherwise,}$$

then $Eg = f(k + 1) - f(k)$, which, by (18), is always (i.e., no matter how the p_i's are chosen, subject to the constraint $p_1 + \cdots + p_n = np$) negative if $k < np < k + 1/(k + 2)$, always positive if $k + 1 - 1/(n - k + 1) < np < k + 1$, and can be of either sign if $k + 1/(k + 2) < np < k + 1 - 1/(n - k + 1)$. Since this holds for all k, we have the following:

THEOREM 2 (Darroch). *If, for some integer k,*

$$k - 1/(n - k + 2) < \sum_{i=1}^{n} p_i < k + 1/(k + 2),$$

then k is the (unique) most likely number of successes. If, on the other hand,

$$k + 1/(k + 2) < \sum_{i=1}^{n} p_i < k + 1 - 1/(n - k + 1),$$

for some k, then the mode may be either k or $k + 1$ depending on how the p_i's are chosen.

We should remark that, while some improvement of these results may be possible using other simple functions of the p_i's, there is no rational function of the p_i's which uniquely determines the mode. This is so even when $n = 2$.

4. Bounds on the distribution function. We define $F(k)$ to be the probability of at most k successes in the n independent trials:

$$F(k) = \sum_{j=0}^{k} f(j)$$
$$F(k \mid np, r, s) = \sum_{j=0}^{k} f(j \mid np, r, s)$$
$$= \sum_{j=0}^{k} \binom{n-r-s}{j-s} ((np - s)/(n - r - s))^{j-s} \cdot ((n - r - np)/(n - r - s))^{n-r-j}.$$

If we let

$$g(x) = 1 \quad \text{if } x \leqq k$$
$$= 0 \quad \text{if } x > k,$$

then $Eg = F(k)$ so that (18) gives upper and lower bounds on this distribution function in terms of the mean. Hoeffding [4] has shown that these bounds can be simplified as follows:

LEMMA 1 (Hoeffding).

$$\min_{p_1+\cdots+p_n=np} F(k) = F(k \mid, np, 0, 0) \qquad \textit{if } np \leqq k$$
$$= \min_{r=0,1,\cdots,[n(1-p)]} F(k \mid np, r, 0) \qquad \textit{if } k < np < k + 1$$
$$= 0 \qquad \textit{if } np \geqq k + 1$$
$$\max_{p_1+\cdots+p_n=np} F(k) = F(k \mid np, 0, 0) \qquad \textit{if } np \geqq k + 1$$
$$= \max_{s=0,1,\cdots,[np]} F(k \mid np, 0, s) \qquad \textit{if } k < np < k + 1$$
$$= 1 \qquad \textit{if } np \leqq k.$$

Similarly, we can obtain a lower bound on the probability of less than k successes minus the probability of more than k successes, in terms of the mean:

LEMMA 2.

$$\min_{p_1+\cdots+p_n=np}\{F(k-1)+F(k)\}$$

$$= F(k-1 \mid np, 0, 0) + F(k \mid np, 0, 0) \qquad \textit{if } np \leqq a(k)$$

$$= \min_{r=0,1,\cdots,[n(1-p)]}\{F(k-1 \mid np, r, 0) + F(k \mid np, r, 0)\} \quad \textit{if } a(k) < np \leqq k$$

$$= \min_{\substack{r=0,1,\cdots,[n(1-p)]\\ s=0,1,\cdots,[np]}}\{F(k-1 \mid np, r, s) + F(k \mid np, r, s)\} \quad \textit{if } k < np < k+1$$

$$= 0 \qquad \textit{if } np \geqq k+1,$$

where

$$a(k) = (k(k-1))^{1/2}\cdot(n-1)/\{(k(k-1))^{1/2} + ((n-k)(n-k-1))^{1/2}\}.$$

The method of proof is essentially the same as that used by Hoeffding in deriving Lemma 1.

Lemma 2 complements the Simmons inequality, derived in [6], which states that

$$(20) \qquad F(k-1 \mid np, 0, 0) > 1 - F(k \mid np, 0, 0) \qquad \text{if} \quad p \leqq k/n \leqq \tfrac{1}{2},$$

unless $p = k/n = \frac{1}{2}$, in which case equality holds.

A bound on F which complements Lemma 1 is presented by the following lemma:

LEMMA 3.

$$(21) \qquad F(k \mid np, 0, 0) \geqq [1 - p/(k+1-kp)]^{n-k} \qquad \text{if} \quad np \leqq k+1.$$

PROOF. Equality holds for $k = 0, n$; $p = 0, 1$. We assume $1 \leqq k \leqq n-1$. Let

$$b_k(p) = F(k \mid np, 0, 0) - [1 - p/(k+1-kp)]^{n-k}.$$

Then

$$(d/dp)b_k(p) = -n\binom{n-1}{k}p^k(1-p)^{n-1-k} + \{(n-k)(k+1)/(k+1-kp)^2\}[1 - p/(k+1-kp)]^{n-1-k},$$

which has the same sign as

$$(n-k)(k+1)^{n-k}/n\binom{n-1}{k} - p^k(k+1-kp)^{n+1-k} = A(k) - B(k; p),$$

where

$$A(k) > 1 \quad \text{for} \quad 1 \leqq k < n-1,$$

$$A(n-1) = 1,$$

$$B(k; 0) = 0,$$

$$B(k; 1) = 1,$$

$$(d/dp)B(k;p) > 0 \quad \text{for} \quad p < (k+1)/(n+1),$$
$$< 0 \quad \text{for} \quad p > (k+1)/(n+1).$$

Hence (21) is true for $k = n - 1$ and for $1 \leqq k < n - 1$ it is sufficient to verify (21) for $np = k + 1$,

We do so by induction. Letting the superscripts below denote the number of trials, we have, for $n \geqq k$,

$$(22) \quad b_k^{(n)}((k+1)/n) - b_{k-1}^{(n-1)}(k/(n-1)) = F^{(n)}(k \mid k+1, 0, 0) - F^{(n-1)}(k-1 \mid k, 0, 0),$$

since, for $p_1 = (k+1)/n$, $p_2 = k/(n-1)$,

$$[1 - p_1/(k+1-kp_1)]^{n-k} = [1 - p_2/(k-(k-1)p_2)]^{(n-1)-(k-1)} = [1 - 1/(n-k)]^{n-k}.$$

Now

$$F^{(n-1)}(k-1 \mid k, 0, 0) = F^{(n)}(k \mid k+1, 0, 1);$$

hence, by Lemma 1, (22) is non-negative. Since (21) holds for $k = 0$, the lemma is proved.

Combining (20) and (21) with Lemmas 1 and 2, we have, in particular,

$$(23) \qquad F(k) \geqq [1 - p/(k+1-kp)]^{n-k} \quad \text{if} \quad p_1 + \cdots + p_n = np \leqq k,$$

$$(24) \quad F(k-1) \geqq 1 - F(k) \qquad \text{if} \quad p_1 + \cdots + p_n = np \leqq a(k);\ k/n \leqq \tfrac{1}{2}.$$

The Inequalities (20) and (21) can be immediately extended to the Poisson distribution, in which case we have:

$$(25) \qquad \sum_{j=0}^{k-1} (\lambda^j/j!)e^{-\lambda} > \sum_{j=k+1}^{\infty} (\lambda^j/j!)e^{-\lambda} \qquad \text{if} \quad \lambda \leqq k,$$

$$(26) \qquad \sum_{j=0}^{k} (\lambda^j/j!)e^{-\lambda} \geqq e^{-\lambda/(k+1)} \qquad \text{if} \quad \lambda \leqq k+1,$$

which imply, in particular,

$$(27) \qquad \sum_{j=0}^{k} (k^j/j!)e^{-k} > \tfrac{1}{2} \qquad (k \text{ an integer}),$$

$$(28) \qquad \sum_{j=0}^{k} (\lambda^j/j!)e^{-\lambda} > e^{-1} \qquad \text{if} \quad \lambda < k+1,$$

as shown by Teicher [8].

The generalization of (23) and (24) to an infinite number of trials is likewise immediate and similar.

REFERENCES

[1] DARROCH, J. N. (1964). On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* **35** 1317–1321.

[2] GHURYE, S. G. and WALLACE, D. L. (1959). A convolution class of monotone likelihood ratio families. *Ann. Math. Statist.* **30** 1158–1164.

[3] HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G. (1959). *Inequalities*. Cambridge Univ. Press.

[4] HOEFFDING, W. (1956). On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* **27** 713–721.

[5] LÉVY, P. (1937). *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, Paris.

[6] RÉNYI, A. (1962). *Wahrscheinlichkeitsrechnung*. VEB Deutscher Verlag der Wissenschaften, Berlin.

[7] TCHEBICHEFF, P. (1846). Démonstration élémentaire d'une proposition générale de la théorie des probabilités. *Jour. f.d. reine u. angew. Math.* **33** 259–267.

[8] TEICHER, HENRY (1955). An inequality on Poisson probabilities. *Ann. Math. Statist.* **26** 147–149.