COMPARISON OF THE POWER FUNCTIONS FOR THE TEST OF INDEPENDENCE IN 2 × 2 CONTINGENCY TABLES

By W. L. Harkness[1] and L. Katz[2]

Pennsylvania State University and Michigan State University

1. Introduction. One of the classical problems in statistical theory is that of testing for independence in 2 × 2 contingency tables. Barnard [1] delineated three distinct experimental situations, [termed the double dichotomy (DD), the 2 × 2 comparative trial (CT), and the 2 × 2 independence trial (IT)] which lead to the presentation of data in the form of 2 × 2 contingency tables. Approximate power functions for tests of independence were considered by Patnaik [10] and Sillitto [12] for the 2 × 2 CT, and Bennett and Hsu [2] have calculated exact values of power for the 2 × 2 IT and 2 × 2 CT, in each case the test used being the conditional test devised concurrently by Yates [14] and Fisher [5]. However, power calculations for the DD have not been made, and hence comparisons of the power functions in the three situations are lacking. This paper represents a contribution in this last direction.

2. Distribution theory. Abstractly, the three experimental situations as outlined by Barnard are describable in the following manner:

I. Double Dichotomy (DD). A total of n similar balls is randomly selected from an urn containing a large number of balls, each ball labeled A_1 or A_2 and also labeled B_1 or B_2. An observed result of the experiment is represented in the form of Table I, where none of the marginal totals are fixed and n_{11} is the observed number of balls labeled A_1 and B_1. It is assumed that the probabilities of occurrence of the various markings of the balls is given by Table II, together with the marginal sums.

II. 2 × 2 *Comparative Trial* (CT). Samples of sizes $n_{1\cdot}$ and $n_{2\cdot}$ are drawn from urns A_1 and A_2 respectively. The numbers of balls labeled B_1 in the two samples (i.e., n_{11} and n_{21}) constitute independent variables with binomial distributions, where the proportion of balls marked B_1 in urn A_i is p_i, $i = 1, 2$. With this type of experiment, one set of marginal totals is fixed in advance as in Table I, namely, $n_{1\cdot}$ and $n_{2\cdot}$.

III. 2 × 2 *Independence Trial* (IT). A total of n similar balls, $n_{1\cdot}$ marked A_1 and $n_{2\cdot}$ marked A_2, are placed in an urn, then withdrawn randomly in order. They are then placed in a row of n cells, of which $n_{\cdot 1}$ have been labeled B_1 and $n_{\cdot 2}$ labeled B_2. The result of the experiment is presented in Table I, where n_{11} is the observed number of balls marked A_1 in receptacles labeled B_1. Both sets of marginal totals are fixed in advanced by the conditions of the experiment.

The probability of observing the sample point $(n_{11}, n_{12}, n_{21}, n_{22})$ in the DD

Received 12 November 1963; revised 12 March 1964.

[1] Research sponsored by the Office of Naval Research.

[2] Research sponsored by the Office of Naval Research and by NSF Grant G 18976.

TABLE I

	B_1	B_2	
A_1	n_{11}	n_{12}	$n_{1\cdot}$
A_2	n_{21}	n_{22}	$n_{2\cdot}$
	$n_{\cdot 1}$	$n_{\cdot 2}$	n

TABLE II

	B_1	B_2	
A_1	π_{11}	π_{12}	$\pi_{1\cdot}$
A_2	π_{21}	π_{22}	$\pi_{2\cdot}$
	$\pi_{\cdot 1}$	$\pi_{\cdot 2}$	1

is, by the multinomial probability law

$$f(n_{11}, n_{12}, n_{21}, n_{22}) = n! \prod_{i,j=1}^{2} \pi_{ij}^{n_{ij}}/n_{ij}! . \tag{1}$$

Replacing π_{11} by $\lambda\pi_{1\cdot}\pi_{\cdot 1}$, π_{12} by $\pi_{1\cdot}(1 - \lambda\pi_{\cdot 1})$, π_{21} by $\pi_{\cdot 1}$ $(1 - \lambda\pi_{1\cdot})$ and π_{22} by $1 - \pi_{1\cdot} - \pi_{\cdot 1} + \lambda\pi_{1\cdot}\pi_{\cdot 1}$, where $\max [0, (\pi_{1\cdot} + \pi_{\cdot 1} - 1)/(\pi_{1\cdot}\pi_{\cdot 1})] < \lambda < \min [\pi_{1\cdot}^{-1}, \pi_{\cdot 1}^{-1}]$, and n_{12} by $(n_{1\cdot} - n_{11})$, n_{21} by $(n_{\cdot 1} - n_{11})$, and n_{22} by $(n - n_{1\cdot} - n_{\cdot 1} + n_{11})$, (1) may be reparametrized as

$$f(n_{11}, n_{1\cdot}, n_{\cdot 1}) = b(n_{1\cdot}; n, \pi_{1\cdot})b(n_{11}; n_{1\cdot}, p_1)b(n_{\cdot 1} - n_{11}; n_{2\cdot}, p_2) \tag{2}$$

where $p_1 = \lambda\pi_{\cdot 1}$, $p_2 = \pi_{\cdot 1}(1 - \lambda\pi_{1\cdot})\pi_{2\cdot}^{-1}$, and

$$b(x; N, p) = \binom{N}{x} p^x(1 - p)^{N-x} .$$

For the 2×2 CT, $n_{1\cdot}$ is fixed, so that the probability of observing the two numbers n_{11} and $n_{\cdot 1} = n_{11} + n_{21}$ is obtained conditionally from (2) as

$$f(n_{11}, n_{\cdot 1} \mid n_{1\cdot}) = b(n_{11}; n_{1\cdot}, p_1)b(n_{\cdot 1} - n_{11}; n_{2\cdot}, p_2). \tag{3}$$

Finally, conditional on fixed values of $n_{1\cdot}$ and $n_{\cdot 1}$, the probability of observing n_{11} in the 2×2 IT, is easily seen to be

$$f(n_{11} \mid n_{1\cdot}, n_{\cdot 1}) = k(n_{1\cdot}, n_{\cdot 1}, t) \binom{n_{1\cdot}}{n_{11}} \binom{n_{2\cdot}}{n_{\cdot 1} - n_{11}} t^{n_{11}} \tag{4}$$

where

$$k(n_{1\cdot}, n_{\cdot 1}, t) = \left[\sum_j \binom{n_{1\cdot}}{j} \binom{n_{2\cdot}}{n_{\cdot 1} - j} t^j\right]^{-1}$$

and

$$t = (\pi_{11}\pi_{22})/(\pi_{12}\pi_{21}) = (p_1q_2)/(p_2q_1)$$
$$= [\lambda(1 - \pi_{1\cdot} - \pi_{\cdot 1} + \lambda\pi_{1\cdot}\pi_{\cdot 1})]/[(1 - \lambda\pi_{1\cdot})(1 - \lambda\pi_{\cdot 1})], \quad q_i = 1 - p_i, \; i = 1, 2,$$

with $0 < t < +\infty$.

3. Tests of independence and exact power functions. In each of the three cases described above, we may be interested in testing for independence of the two classifications. For the DD, the hypothesis of interest is that $\pi_{11} = \pi_{1\cdot}\pi_{\cdot 1}$; for the 2 × 2 CT we wish to test the equality of the two binomial proportions p_1 and p_2; while for the 2 × 2 IT, the null hypothesis is that the markings A_1 or A_2 are independent of the labelings B_1 or B_2. If the hypothesis is correct, the conditional distribution of n_{11}, given $n_{\cdot 1}$ and $n_{1\cdot}$, is the hypergeometric distribution, corresponding to putting $t = 1$ in (4). Since $\pi_{11} = \pi_{1\cdot}\pi_{\cdot 1}$, $p_1 = p_2$, and $t = 1$ if and only if $\lambda = 1$, we may specify the null hypothesis by $H_0 : \lambda = 1$, for each case. Any alternative hypothesis may be expressed as $H_1 : \lambda \neq 1$, for any of the three cases, so that H_1 is composite. In terms of λ, H_0 is simple. The nuisance parameters $\pi_{1\cdot}$ and $\pi_{\cdot 1}$ make it composite, also.

We proceed to test the independence hypothesis in each case using the uniformly most powerful unbiased size α test (UMPUT), as described, for example, by Lehmann [8]. This test is a conditional test based on the two tails, the test statistic $\varphi_{n_{1\cdot}, n_{\cdot 1}}(x)$ for given values of $n_{1\cdot}$ and $n_{\cdot 1}$ being given by

$$\begin{aligned} \varphi_{n_{1\cdot}, n_{\cdot 1}}(x) &= 1 && \text{if } x < c_1(n_{1\cdot}, n_{\cdot 1}) \text{ or } x > c_2(n_{1\cdot}, n_{\cdot 1}) \\ &= \epsilon_i && \text{if } x = c_i(n_{1\cdot}, n_{\cdot 1}), \ i = 1, 2 \\ &= 0 && \text{if } c_1(n_{1\cdot}, n_{\cdot 1}) < x < c_2(n_{1\cdot}, n_{\cdot 1}) \end{aligned}$$

where c_i and ϵ_i are constants uniquely determined by the two equations

$$\text{(i)} \quad E[\varphi_{n_{1\cdot}, n_{\cdot 1}}(x)] = \alpha, \qquad \text{(ii)} \quad E[x\varphi_{n_{1\cdot}, n_{\cdot 1}}(x)] = \alpha E[x]$$

and where the expectations are taken with respect to the hypergeometric distribution, that is, (4) with $t = 1$. The first equation (i) reflects the fact that the test is *similar* (on the boundary of H_0 and H_1) whereas the condition that the test be unbiased leads to the second equation (ii). The randomization feature of the UMPUT arises directly from these considerations. As a consequence, the type I error is exactly equal to α, whereas in the Fisher-Yates test, the size of the test is usually considerably less than the nominal α-value, at least for small samples.

The exact power functions for the unconditional tests of independence in the DD, 2 × 2 CT and 2 × 2 IT are then seen to be, respectively

$$\begin{aligned} P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1}) &= \sum_{n_{1\cdot}=0}^{n} \sum_{n_{\cdot 1}=0}^{n} \sum_{x} \varphi_{n_{1\cdot}, n_{\cdot 1}}(x) f(x, n_{1\cdot}, n_{\cdot 1}) \\ Q_n(p_1, p_2 \mid n_{1\cdot}) &= \sum_{n_{\cdot 1}=0}^{n} \sum_{x} \varphi_{n_{1\cdot}, n_{\cdot 1}}(x) f(x, n_{\cdot 1} \mid n_{1\cdot}) \qquad (5) \\ R_n(t \mid n_{1\cdot}, n_{\cdot 1}) &= \sum_{x} \varphi_{n_{1\cdot}, n_{\cdot 1}}(x) f(x \mid n_{1\cdot}, n_{\cdot 1}). \end{aligned}$$

Harkness [7] has computed (for the UMPUT of independence) exact values of $P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1})$ for $n = 10$, 20, and 30 and $(\pi_{1\cdot}, \pi_{\cdot 1}) = (.1, .1), \cdots, (.5, .5)$, with a wide range of values for λ. Also, $Q_n(p_1, p_2 \mid n_1)$ was computed for $n = 10$, 20, and 30 for $n_{1\cdot} = 2, \cdots, n/2$ and $(p_1, p_2) = (.1, .1), \cdots, (.9, .9)$. An extensive set of values of $P_n(t \mid n_{1\cdot}, n_{\cdot 1})$ were also computed for $n = 10, 20$, and 30 and selected values of $n_{1\cdot}$, $n_{\cdot 1}$, and t. The results of some of these calculations are given in Tables A, B_1, B_2, and B_3. A brief discussion of the implications of these calculations follow our next remarks on an investigation of power in large samples.

4. Asymptotic power. Our main tool in investigating the large sample behavior of the various power functions is the following theorem due to Hannan and Harkness [6].

THEOREM 1. *Let P_i, Q_i $(0 < P_i < 1, Q_i = 1 - P_i)$ be the unique solutions of the equations*

$$t = (P_1Q_2)/(P_2Q_1), \qquad n_{1\cdot}P_1 + n_{2\cdot}P_2 = n_{\cdot 1}. \tag{6}$$

Let $H_i^{-2} = n_{i\cdot}P_iQ_i$ $(i = 1, 2)$, $H^2 = H_1^2 + H_2^2$, and $X_{n_{11}} = H(n_{11} - n_{1\cdot}P_1)$. Then

(i) $$f(n_{11} \mid n_{1\cdot}, n_{\cdot 1}) \sim H\varphi(X_{n_{11}}) \quad as \quad H, HX_{n_{11}}^3 \to 0$$

(ii) $$\sum_{n_{11}=a}^{b} f(n_{11} \mid n_{1\cdot}, n_{\cdot 1}) \sim \Phi(X_{b+\frac{1}{2}}) - \Phi(X_{a-\frac{1}{2}}) \quad as \quad H, HX_a^3, HX_b^3 \to 0$$

(iii) $$\sum_{n_{11}=a}^{\min(n_{1\cdot}, n_{\cdot 1})} f(n_{11} \mid n_{1\cdot}, n_{\cdot 1}) \sim 1 - \Phi(X_{a-\frac{1}{2}}) \quad as \quad H, HX_a^3 \to 0$$

where $\varphi(x) = (2\pi)^{-\frac{1}{2}} \exp(-x^2/2)$, $\Phi(u) = \int_{-\infty}^{u} \varphi(x)\, dx$, and "$\sim$" means the ratio of the two sides tends to one.

In general, it may be rather difficult to determine the values of n_{11} for which $HX_{n_{11}}^3 \to 0$, but for the important special case when $(n_{1\cdot}/n) \to \theta_1$, and $(n_{\cdot 1}/n) \to \theta_2$, it can be shown that $HX_{n_{11}}^3 \to 0$ if and only if $n^{-2}(n_{11} - n_{1\cdot}P_1)^3 \to 0$, or equivalently, $n^{-\frac{1}{2}}X_{n_{11}}^3 \to 0$, which is Feller's [3] condition for the validity of his normal approximation theorem to the binomial.

Solving the equations in (6) for P_1 and P_2, noting, according to Theorem 1, that the mean and variance of the distribution given in (4) are asymptotically given by $\mu = n_{1\cdot}P_1$ and $\sigma^2 = H^{-2}$ respectively, it is found that

$$\mu = n_{1\cdot}\, P_1 = \lambda^*(n_{1\cdot}\, n_{\cdot 1}/n), \qquad \sigma^2 = \left[\sum_{i,j=1}^{2} (1/\pi_{ij}^*)\right]^{-1}$$

where $P_1 = \{-d + [d^2 + 4n_{1\cdot}n_{\cdot 1}t(1 - t)]^{\frac{1}{2}}\}/\{2(1 - t)n_{1\cdot}\}$, $d = n - (n_{1\cdot} + n_{\cdot 1})(1 - t)$, $\lambda^* = (nP_1)/(n_{\cdot 1})$, and the π_{ij}^* are the π_{ij} in Table II expressed in terms of λ, $\pi_{1\cdot}$, $\pi_{\cdot 1}$, with λ^*, $(n_{1\cdot}/n)$, $(n_{\cdot 1}/n)$, replacing λ, $\pi_{1\cdot}$, $\pi_{\cdot 1}$. (The equations in (6) lead to a quadratic equation which P_1 must satisfy —one root is discarded since it leads to impossible values for P_1.) If $t = 1$, then $\mu = (n_{1\cdot}n_{\cdot 1})/n$ and $\sigma^2 = [n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}]/[n^2(n - 1)] = h^2$.

Under the conditions of Theorem 1, asymptotically $H_0{:}\lambda = 1$ is rejected if $|n_{11} - (n_{1\cdot}n_{\cdot 1}/n)| \geq hu_{\alpha/2}$, where $u_{\alpha/2}$ satisfies $\Phi(u_{\alpha/2}) = 1 - \alpha/2$. Thus, as an immediate consequence of Theorem 1, we have

THEOREM 2. *If* $\sigma^{-4}[hu_{\alpha/2} + (1 - \lambda^*)(n_{1}.n_{.1}/n)] \to 0$, *then*

$$R_n(t \mid n_{1.}, n_{.1}) \sim \Phi[-u_{\alpha/2}(h/\sigma) + (1 - \lambda^*)(n_{1}.n_{.1}/n\sigma)] + \Phi[-u_{\alpha/2}(h/\sigma) - (1 - \lambda^*)(n_{1}.n_{.1}/n\sigma)]. \tag{7}$$

If t is kept fixed as $n_{1.}, n_{.1}, n \to \infty$, then the power of the test for independence tends to one (meaning, of course, that the test is consistent). In order to examine the situation in which power is not close to one in large samples, we must either let the significance probability decrease to zero, or consider a sequence of alternative hypotheses converging to the null hypothesis. We discuss the second case. In the following "$\to$" always means "as the relevant variables tend to $+\infty$."

THEOREM 3. *If* $|n_{1.} - n\eta_1| = o(n^{2/3})$ *and* $|n_{.1} - n\eta_2| = o(n^{2/3})$, $0 < \eta_i < 1$, $i = 1, 2$, *then*

$$R_n(t \mid n_{1.}, n_{.1})]_{t_n} \to \Phi(-u_{\alpha/2} + \delta) + \Phi(-u_{\alpha/2} - \delta)$$

where $\delta^2 = \eta_1(1 - \eta_1)\eta_2(1 - \eta_2)c$ *and* $t_n = 1 - n^{-\frac{1}{2}}c$.

PROOF. The hypotheses imply $(n_{1.}/n) \to \eta_1$, $(n_{.1}/n) \to \eta_2$. Since t is positive, the choice of P_1 and P_2 in Theorem 1 ensure that $\sigma^2 \to +\infty$. Hence we need only to show that

$$(u_{\alpha/2}h + n^{-1}n_{1}.n_{.1} - n_{1}.P_1)\sigma^{-1} \to u_{\alpha/2} + \delta \tag{8}$$

as $n_{1.}, n_{.1}, n$ tend to ∞, since then the hypothesis of Theorem 2 will be obviously satisfied. Noting that P_1, P_2 satisfy (6), with $t_n = 1 - n^{-\frac{1}{2}}c$, and n sufficiently large, it is easily established that P_1 and P_2 both converge to η_2 (note that P_1 and P_2 depend on n, but we shall not indicate this explicitly). Since $P_2 = (1 - cQ_1n^{\frac{1}{2}})^{-1}$,

$$n^{\frac{1}{2}}(P_2 - P_1) = (n^{\frac{1}{2}} - cQ_1)^{-1}n^{\frac{1}{2}}cP_1Q_1 \to \eta_2(1 - \eta_2)c.$$

Similarly, $\sigma^2/n \to \eta_1(1 - \eta_1)\eta_2(1 - \eta_2)$, so that

$$hu_{\alpha/2}/\sigma = u_{\alpha/2}\left[\frac{n_{1.}\, n_{2.}\, n_{.1}\, n_{.2}}{n^3(n - 1)}\frac{n}{\sigma^2}\right]^{\frac{1}{2}} \to u_{\alpha/2}.$$

Finally,

$$[n^{-1}n_{1}.n_{.1} - n_{1}.P_1]\sigma^{-1} = n_{1}.[n_{1}.P_1 + n_{2}.P_2 - nP_1](n\sigma)^{-1}$$
$$= (n_{1.}/n)(1 - [n_{1.}/n])(\sigma^2/n)^{-\frac{1}{2}}n^{\frac{1}{2}}(P_2 - P_1) \to [\eta_1(1 - \eta_1)\eta_2(1 - \eta_2)c]^{\frac{1}{2}}.$$

Thus, (8) is established, so that the theorem follows.

Now consider a sequence of 2×2 tables with fixed marginal probabilities $\pi_{1.}$ and $\pi_{.1}$. Applying Theorem 3 with $\eta_1 = \pi_{1.}$, $\eta_2 = \pi_{.1}$, and

$$t = \lambda(1 - \pi_{1.} - \pi_{.1} + \lambda\pi_{1.}\pi_{.1})/(1 - \lambda\pi_{1.})(1 - \lambda\pi_{.1}),$$

with $\lambda = 1 - n^{-\frac{1}{2}}\gamma \equiv \lambda_n$ it is readily verified that $n^{\frac{1}{2}}(1 - t) \to \gamma(\pi_{2.}\pi_{.2})^{-1}$. Thus,

$$(9) \qquad R_n(t \mid n_{1\cdot}, n_{\cdot 1})]_{\lambda_n} \to \Phi(-u_{\alpha/2} + \delta) + \Phi(-u_{\alpha/2} - \delta)$$

where $\delta^2 = \gamma^2\pi_{1\cdot}\pi_{\cdot 1}/\pi_{2\cdot}\pi_{\cdot 2}$.

Using (9), Theorems 1 and 3, and applying Tchebycheff's inequality, we see that if $|n_{1\cdot} - n\pi_{1\cdot}| = o(n^{2/3})$, then

$$(10) \qquad Q_n(p_1, p_2 \mid n_{1\cdot})]_{\lambda_n} \to \Phi(-u_{\alpha/2} + \delta) + \Phi(-u_{\alpha/2} - \delta)$$

where $p_1 = \lambda\pi_{\cdot 1}$, $p_2 = [\pi_{\cdot 1}(1 - \lambda\pi_{1\cdot})]/\pi_{2\cdot}$, and also

$$(11) \qquad P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1})]_{\lambda_n} \to \Phi(-u_{\alpha/2} + \delta) + \Phi(-u_{\alpha/2} - \delta).$$

Thus, (9), (10), and (11) assert that the three power functions have the same limit, when evaluated at $\lambda = 1 - n^{-\frac{1}{2}}\gamma = \lambda_n$.

Finally, we consider two special cases of Theorem 2. First, putting $n_{\cdot 1} = n_{1\cdot}p_1 + n_{2\cdot}p_2 = \nu$ in (7) and $t = p_1q_2/p_2q_1$, we obtain, after some simplification,

$$(12) \qquad R_n(t \mid n_{1\cdot}, \nu) \sim \Phi(-h'u_{\alpha/2}\sigma' + \theta') + \Phi(-h'u_{\alpha/2}\sigma' - \theta')$$

where $h' = [n_{1\cdot}n_{2\cdot}\nu(n - \nu)/n^2(n - 1)]^{\frac{1}{2}}$, $\sigma' = (\sigma_1^{-2} + \sigma_2^{-2})^{-\frac{1}{2}}$, $\theta' = [(n_{1\cdot}\nu/n) - n_{1\cdot}p_1][\sigma']^{-1}$ and $\sigma_i^2 = n_{i\cdot}p_iq_i$, $i = 1, 2$. (Note that p_1, p_2 are now solutions of (6), and that $\lambda^* = nP_1/n_{\cdot 1}$).

Secondly, if $t = \lambda(1 - \pi_1 - \pi_{\cdot 1} + \lambda\pi_{1\cdot}\pi_{\cdot 1})/(1 - \lambda\pi_{1\cdot})(1 - \lambda\pi_{\cdot 1})$, replacing $n_{1\cdot}$ and $n_{\cdot 1}$ by $n\pi_{1\cdot}$ and $n\pi_{\cdot 1}$, respectively, yields

$$(13) \qquad R_n(t \mid n\pi_{1\cdot}, n\pi_{\cdot 1}) \sim \Phi(-h''u_{\alpha/2}\sigma_\pi^{-1} + \theta) + \Phi(-h''u_{\alpha/2}\sigma_\pi^{-1} - \theta)$$

where $h'' = [n^2\pi_{1\cdot}\pi_{2\cdot}\pi_{\cdot 1}\pi_{\cdot 2}/(n - 1)]^{\frac{1}{2}}$, $\sigma_\pi^2 = n[\sum_{i,j=1}^{2} \pi_{ij}^{-1}]^{-1}$, $\theta = n(1 - \lambda)\pi_{1\cdot}\pi_{\cdot 1}\sigma_\pi^{-1}$, and the π_{ij}, $i, j = 1, 2$, are the cell entries in Table II expressed in terms of λ, $\pi_{1\cdot}$, and $\pi_{\cdot 1}$, as suggested following (1).

5. Remarks. The power functions $Q_n(p_1, p_2 \mid n_{1\cdot})$ and $P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1})$ are expressible as weighted averages of the power function $R_n(t \mid n_{1\cdot}, n_{\cdot 1})$. Explicitly,

$$(14) \qquad P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1}) = E_{(n_{1\cdot}, n_{\cdot 1})}[R_n(t \mid n_{1\cdot}, n_{\cdot 1})]$$

and

$$(15) \qquad Q_n(p_1, p_2 \mid n_{1\cdot}) = E_{(n_{\cdot 1} \mid n_{1\cdot})}[R_n(t \mid n_{1\cdot}, n_{\cdot 1})].$$

In (14), $t = [\lambda(1 - \pi_{1\cdot} - \pi_{\cdot 1} + \lambda\pi_{1\cdot}\pi_{\cdot 1})]/[(1 - \lambda\pi_{1\cdot})(1 - \lambda\pi_{\cdot 1})]$ and $E_{(n_{1\cdot}, n_{\cdot 1})}[R_n(t \mid n_{1\cdot}, n_{\cdot 1})]$ denotes the expected value of $[R_n(t \mid n_{1\cdot}, n_{\cdot 1})]$, with respect to the joint distribution of $n_{1\cdot}$ and $n_{\cdot 1}$. In (15), $t = p_1q_2/p_2q_1$ and the expectation is taken with respect to the distribution of $n_{\cdot 1}$, conditional on fixed values of $(n_{1\cdot})$. Thus, $R_n(t \mid n_{1\cdot}, n_{\cdot 1})$ is an unbiased estimator of $P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1})$ and $Q_n(p_1, p_2 \mid n_{\cdot 1})$, in the sense of Lehmann ([8], p. 140). Since $\nu = n_{1\cdot}p_1 + n_{2\cdot}p_2$ is the mean value of the distribution of $n_{\cdot 1}$ in the 2×2 CT, and $n\pi_{1\cdot}$, $n\pi_{\cdot 1}$ are the mean values of $n_{1\cdot}$ and $n_{\cdot 1}$ respectively, in the DD, the motivation for considering the approximating (12) and (13) becomes clearer.

We also observe that an application of the results of Mitra [9] shows that if the parameters $\pi_{1\cdot}$ and $\pi_{\cdot 1}$ are assumed to be unknown in the DD, and are estimated by $n_{1\cdot}/n$ and $n_{\cdot 1}/n$ respectively, then the limiting power function of the asymptotically equivalent frequency chi-square test is given by the non-central chi-square distribution with one degree of freedom and non-centrality parameter δ^2.

6. Comparison of exact and approximate power. The test for independence in each of the three distinct experimental situations outlined in the introduction is carried out in terms of the same conditional test, so that the exact power functions in the 2 × 2 CT and DD (as previously noted) are weighted averages of the power function $R_n(t \mid n_{1\cdot}, n_{\cdot 1})$ in the 2 × 2 IT. It is therefore of some interest to compare the values of these power functions.

In the 2 × 2 CT, we put $p_1 = \lambda\pi_{\cdot 1}$, $p_2 = [\pi_{\cdot 1}(1 - \lambda\pi_{1\cdot})]/\pi_{2\cdot}$ and $n_{1\cdot} = n\pi_{1\cdot}$, and in the 2 × 2 IT we let

$$t = \frac{\lambda(1 - \pi_{1\cdot} - \pi_{\cdot 1} + \lambda\pi_{1\cdot}\,\pi_{\cdot 1})}{(1 - \lambda\pi_{1\cdot})(1 - \lambda\pi_{\cdot 1})}, \qquad n_{1\cdot} = n\pi_{1\cdot}, \quad \text{and} \quad n_{\cdot 1} = n\pi_{\cdot 1},$$

where λ, $\pi_{1\cdot}$, and $\pi_{\cdot 1}$ are the parameters in the DD. Then for large samples the three power functions $P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1})$, $Q_n(p_1, p_2 \mid n\pi_{1\cdot})$ and $R_n(t \mid n\pi_{1\cdot}, n\pi_{\cdot 1})$ should be very nearly equal, in view of (9), (10), and (11). In order to examine the rapidity with which these power functions converge together, some exact values of these functions are given in Tables B_1, B_2, and B_3.

On examining these tables, it is seen that in general values of power in the 2 × 2 CT and 2 × 2 IT are greater than those for the DD, with power greatest in the 2 × 2 CT for small λ, while for large values of λ, power is greatest in the 2 × 2 IT. For $n = 10$, there are very substantial differences in power between the three cases, but for $n = 30$, these differences tend to be negligible. It can also be seen from these tables that for $n = 30$ there is an ordering in the values of power, with $R_n(t \mid n\pi_{1\cdot}, n\pi_{\cdot 1}) > Q_n(p_1, p_2 \mid n\pi_{1\cdot}) > P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1})$. The level of significance $\alpha = .05$ was used in all computations of power in the tables which follow. We note that the approximations (9), (12), and (13) coincide for the

TABLE III

Comparison of exact and approximate values of power in the 2 × 2 CT

p_1	p_2	Patnaik* (2nd Approx.)	Sillitto* (Approx.)	(12)† (Approx.)	$Q_{30}(p_1, p_2 \mid 15)$ (Exact)
.3	.1	.316	.293	.241	.254
.6	.1	.925	.871	.876	.852
.7	.2	.860	.824	.812	.802
.8	.2	.967	.941	.947	.932
.7	.3	.634	.617	.585	.587
.6	.4	.199	.197	.180	.186

* Calculation of values of Patnaik's and Sillitto's approximations given here were carried out by the present authors.

† Values computed using the normal approximation given in Equation (12).

particular choice of parameters in Tables B_1, B_2, and B_3. For $n = 30$, their common value has been calculated for several values of λ and tabulated alongside the exact values of the three power functions which it simultaneously approximates.

In the 2×2 CT, Patnaik [10] and Sillitto [12] have also supplied normal approximations for the power of the two-sided test for equality of the proportions p_1 and p_2. In Table III we compare their approximations to that given by (12), with $n = 30$, and $n_{1\cdot} = 15$. Bennett and Hsu [2] have made a similar comparison for one-sided tests with "exact" values of power based on the Fisher-Yates test using Finney's [4] tables. Whereas in the UMPUT the size of the test is exactly α, for the test based on Finney's tables the effective size is almost always considerably less than the nominal α value. Consequently, the "exact" values of power given by Bennett and Hsu are, in general, much less than those given by the approximations of Patnaik and Sillitto. An examination of Table III shows that this phenomenon is not present when the UMPUT is used.

REFERENCES

[1] Barnard, G. A. (1947). Significance tests for 2×2 tables. *Biometrika* **34** 123–138.

[2] Bennett, B. M. and Hsu, P. (1960). On the power function of the exact test for the 2×2 contingency table. *Biometrika* **47** 393–398.

[3] Feller, William (1957). *An Introduction to Probability Theory and Its Applications*, 1(2nd ed.). Wiley, New York.

[4] Finney, D. J. (1948). The Fisher-Yates test of significance in 2×2 contingency tables. *Biometrika* **35** 145–156.

[5] Fisher, R. A. (1935). The logic of inductive inference. *J. Roy. Statist. Soc.* **98** 39–54.

[6] Hannan, J. and Harkness, W. (1963). Normal approximation to the distribution of two independent binomials, conditional on fixed sum. *Ann. Math. Statist.* **34** 1593–1595.

[7] Harkness, William Leonard (1959). An investigation of the power function for the test of independence in 2×2 contingency tables. Ph.D. Thesis, Michigan State Univ. Library.

[8] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[9] Mitra, S. K. (1958). On the limiting power function of the frequency chi-square test. *Ann. Math. Statist.* **29** 1221–1233.

[10] Patnaik, P. B. (1948). The power function of the test for the difference between two proportions in a 2×2 table. *Biometrika* **35** 157–175.

[11] Pearson, E. S. (1947). The choice of statistical tests illustrated on the interpretation of data classed in a 2×2 table. *Biometrika* **34** 139–167.

[12] Sillitto, G. P. (1949). Note on approximation to the power of the 2×2 comparative trial. *Biometrika* **36** 347–352.

[13] Stevens, W. L. (1951). Mean and variance of an entry in a contingency table. *Biometrika* **38** 468–470.

[14] Yates, F. (1934). Contingency tables involving small numbers and the χ^2-test. *Suppl., J. Roy. Statist. Soc.* **1** 217–235.

TABLE A

Values of exact power for UMPUT test of independence in double dichotomy

$\pi_{\cdot 1}$	.5	.4	.3	.2	.1
λ		$P_{10}(\lambda, .5, \pi_{\cdot 1})$			
.1	.801	.527	.297	.146	.071
.2	.647	.416	.238	.124	.066
.3	.497	.319	.189	.106	.062
.4	.365	.240	.149	.090	.059
.5	.258	.176	.116	.077	.056
.6	.176	.127	.091	.067	.054
.7	.118	.092	.073	.060	.052
.8	.079	.068	.060	.054	.051
.9	.057	.054	.052	.051	.050
1.0	.050	.050	.050	.050	.050
λ		$P_{20}(\lambda, .5, \pi_{\cdot 1})$			
.1	.998	.945	.758	.434	.144
.2	.975	.849	.624	.344	.122
.3	.899	.714	.491	.268	.104
.4	.758	.558	.369	.205	.089
.5	.577	.406	.266	.154	.077
.6	.392	.274	.184	.114	.067
.7	.239	.172	.123	.085	.059
.8	.131	.102	.081	.065	.054
.9	.069	.063	.058	.054	.051
1.0	.050	.050	.050	.050	.050
λ		$P_{30}(\lambda, .5, \pi_{\cdot 1})$			
.1	1.000	.996	.940	.705	.263
.2	.999	.970	.844	.573	.212
.3	.985	.897	.711	.448	.170
.4	.923	.766	.558	.336	.135
.5	.781	.591	.407	.243	.108
.6	.571	.407	.275	.169	.086
.7	.352	.249	.173	.115	.070
.8	.181	.135	.103	.078	.059
.9	.081	.071	.063	.057	.052
1.0	.050	.050	.050	.050	.050

$P_n(1 - \lambda, .5, \pi_{\cdot 1}) = P_n(1 + \lambda, .5, \pi_{\cdot 1})$ by symmetry.

TABLE B_1

Comparison of the three exact power functions P_n, Q_n, R_n and normal approximation*

$\pi_{1\cdot} = \pi_{\cdot 1} = .5$

λ	n = 10			n = 20		
	P_n	Q_n	R_n	P_n	Q_n	R_n
.1	.801	.933		.998	.999	1.000
.2	.647	.786	.805	.975	.983	.991
.3	.497	.617	.620	.889	.916	.939
.4	.365	.457	.447	.758	.783	.812
.5	.258	.321	.308	.577	.602	.626
.6	.176	.215	.206	.392	.412	.426
.7	.118	.139	.134	.239	.250	.256
.8	.079	.088	.086	.131	.136	.137
.9	.057	.059	.069	.069	.071	.071
1.0	.050	.050	.050	.050	.050	.050

λ	n = 30			
	P_n	Q_n	R_n	N.A.
.1	1.000	1.000	1.000	1.000
.2	.999	.999	.999	1.000
.3	.985	.988	.991	.995
.4	.923	.932	.942	.947
.5	.781	.796	.809	.805
.6	.571	.587	.598	.585
.7	.352	.363	.368	.357
.8	.181	.186	.200	.180
.9	.081	.083	.090	.078
1.0	.050	.050	.050	.050

* Here and in Tables B_2 and B_3, $P_n = P_n(\lambda, \pi_{1\cdot}, \pi_{\cdot 1})$, $Q_n = Q_n(p_1, p_2 \mid n\pi_{1\cdot})$; $R_n = R_n(t \mid n\pi_{1\cdot}, n\pi_{\cdot 1})$, $p_1 = \lambda\pi_{\cdot 1}$, $p_2 = \pi_{\cdot 1}(1 - \lambda\pi_{1\cdot})/\pi_{\cdot 2}$;

$$t = \lambda(1 - \pi_{1\cdot} - \pi_{\cdot 1} + \lambda\pi_{1\cdot}\pi_{\cdot 1})/(1 - \lambda\pi_{1\cdot})(1 - \lambda\pi_{\cdot 1}),$$

and N. A. is the normal approximation given by the equivalent expressions (9), (12), or (13) for the particular parameters used here.

TABLE B_2

Comparison of the exact power functions P_n, Q_n, R_n and the normal approximation

$\pi_{1\cdot} = \pi_{\cdot 1} = .4$

λ	n = 10			n = 20		
	P_n	Q_n	R_n	P_n	Q_n	R_n
.1	.333	.410	.309	.806	.851	.849
.2	.266	.323	.266	.669	.713	.695
.3	.209	.251	.222	.528	.565	.547
.4	.163	.192	.181	.397	.426	.412
.5	.126	.145	.142	.285	.305	.298
.6	.097	.109	.110	.195	.208	.205
.7	.076	.083	.084	.129	.135	.135
.8	.061	.064	.065	.084	.087	.087
.9	.053	.054	.054	.058	.059	.059
1.0	.050	.050	.050	.050	.050	.050
1.1	.053	.054	.054	.058	.059	.059
1.2	.062	.065	.066	.084	.086	.089
1.3	.077	.083	.086	.128	.134	.140
1.4	.099	.111	.115	.192	.203	.216
1.5	.128	.148	.154	.277	.294	.317
1.6	.167	.195	.206	.380	.404	.439
1.7	.214	.254	.271	.497	.527	.573
1.8	.272	.325	.353	.620	.654	.707
1.9	.341	.407	.451	.737	.772	.824
2.0	.420	.501	.565	.838	.870	.913
2.1	.508	.604	.689	.914	.940	.967
2.2	.603	.712	.810	.963	.980	.992
2.3	.702	.821	.912	.988	.996	.999
2.4	.798	.921	.978	.997	1.000	—

λ	n = 30			
	P_n	Q_n	R_n	N.A.
.1	.959	.969	.975	.996
.2	.874	.890	.895	.923
.3	.744	.763	.769	—
.4	.588	.606	.613	.594
.5	.430	.444	.452	—
.6	.290	.299	.307	.282
.7	.181	.186	.191	—
.8	.106	.108	.111	.105

TABLE B_2—*Continued*

λ	$n = 30$			
	P_n	Q_n	R_n	N.A.
.9	.064	.064	.065	—
1.0	.050	.050	.050	.050
1.1	.064	.064	.065	—
1.2	.105	.108	.112	.106
1.3	.178	.184	.196	—
1.4	.282	.293	.313	.293
1.5	.413	.429	.459	—
1.6	.559	.579	.617	.584
1.7	.702	.724	.763	—
1.8	.823	.844	.877	.859
1.9	.912	.927	.949	—
2.0	.964	.974	.985	.985
2.1	.989	.994	.997	—
2.2	.998	.999	1.000	1.000
2.3	1.000	1.000	1.000	—
2.4	1.000	1.000	1.000	1.000

TABLE B_3

Comparison of the three exact power functions P_n, Q_n and R_n for $\pi_{\cdot 1} = \pi_{1\cdot} = .3$

λ	n = 10			n = 20			n = 30		
	P_n	Q_n	R_n	P_n	Q_n	R_n	P_n	Q_n	R_n
.1	.122	.125	.101	.343	.367	.306	.602	.642	.779
.2	.106	.109	.093	.272	.290	.254	.473	.501	.573
.3	.093	.095	.085	.214	.226	.207	.362	.381	.439
.4	.081	.083	.077	.166	.175	.166	.270	.282	.318
.5	.072	.073	.069	.128	.134	.131	.196	.204	.225
.6	.064	.065	.063	.099	.102	.102	.140	.145	.156
.7	.058	.058	.058	.077	.079	.080	.099	.101	.107
.8	.053	.054	.054	.062	.063	.063	.071	.072	.074
.9	.051	.051	.051	.053	.053	.053	.055	.055	.056
1.0	.050	.050	.050	.050	.050	.050	.050	.050	.050
1.1	.051	.051	.051	.053	.053	.053	.055	.055	.056
1.2	.054	.054	.055	.062	.063	.064	.071	.072	.073
1.3	.061	.059	.058	.077	.079	.082	.097	.100	.101
1.4	.065	.067	.071	.099	.103	.108	.136	.140	.142
1.5	.073	.077	.084	.128	.134	.142	.186	.193	.195
1.6	.084	.089	.101	.164	.174	.184	.247	.257	.261
1.7	.098	.105	.124	.208	.221	.234	.319	.333	.337
1.8	.114	.124	.152	.259	.276	.292	.400	.417	.422
1.9	.133	.146	.186	.317	.339	.356	.485	.507	.513
2.0	.155	.171	.227	.380	.408	.426	.573	.598	.606
2.1	.180	.201	.275	.448	.482	.500	.658	.687	.696
2.2	.208	.234	.330	.520	.558	.576	.737	.768	.778
2.3	.240	.272	.393	.591	.636	.651	.806	.838	.848
2.4	.275	.315	.464	.661	.711	.724	.864	.894	.904
2.5	.313	.363	.539	.727	.781	.792	.909	.937	.946
2.6	.355	.416	.619	.787	.843	.853	.943	.966	.973
2.7	.400	.475	.700	.840	.896	.904	.966	.984	.988
2.8	.448	.540	.778	.885	.938	.944	.981	.994	.996
2.9	.499	.611	.849	.920	.968	.972	.990	.998	—
3.0	.552	.689	.910	.947	.986	.989	.995	1.000	—
3.1	.605	.774	.956	.966	.996	.997	.997	1.000	—
3.2	.659	.866	—	.979	.999	—	.999	1.000	—
3.3	.712	.965	—	.989	1.000	—	.999	1.000	—