

# The exact binomial test between two independent proportions: A companion

Louis Laurencelle <sup>a</sup>

<sup>a</sup>Université d'Ottawa

**Abstract** ■ This note, a summary in English of a former article in French (Laurencelle, 2017), presents an outline of the theory and implementation of an exact test of the difference between two binomial variables. Essential formulas are provided, as well as two executable computer programs, one in Delphi, the other in R.

**Keywords** ■ exact test of two proportions. **Tools** ■ Delphi, R.

**Acting Editor** ■ Denis Cousineau (Université d'Ottawa)

**Reviewers**  
■ No reviewer

[louis.laurencelle@gmail.com](mailto:louis.laurencelle@gmail.com)

[10.20982/tqmp.17.2.p076](https://doi.org/10.20982/tqmp.17.2.p076)

## Introduction

Let there be two samples of given sizes  $n_1$  and  $n_2$ , and numbers  $x_1$  and  $x_2$  of “successes” in corresponding samples, where  $0 \leq x_1 \leq n_1$  and  $0 \leq x_2 \leq n_2$ . The purpose of the test is to determine whether the difference  $x_1/n_1 - x_2/n_2$  results from ordinary random variation or if it is probabilistically exceptional, given a probability threshold. Treatises and textbooks on applied statistics report a number of test procedures using some form of normal-based approximation, such as a Chi-squared test, various sorts of  $z$  (normal) and  $t$  (Student’s) tests with or without continuity correction, etc. (Howell, 2017; Siegel & Castellan, 1988; Sokal & Rohlf, 1995). The reader may also be aware of Fisher’s “exact probability test”, which is indeed an exact probability test but not a test applicable to independent proportions, the rows ( $n_1, n_2$ ) and columns ( $x_1, x_2$ ) totals being both fixed. Liddell (1978) proposes a test with an exact probability calculation, but an incomplete one that uses a single point estimate of the parametric probability for the null hypothesis of the test.

The test procedure of Liddell (1978) exploits a point estimate of  $\pi$ , the probabilistic reference parameter, which is the maximum likelihood estimator for two binomial series, i.e., here,  $\hat{\pi} = (x_1 + x_2)/(n_1 + n_2)$ . Using this single estimate and following a suggestion of Kendall and Stuart (1977), Liddell lists all possible values of the two proportions and their differences, comparing these with the ob-

served difference and then accumulating the probabilities of equal or more extreme differences.

Our proposed test, similar to that of Liddell (1978), uses the same procedure of enumerating the  $(n_1 + 1) \times (n_2 + 1)$  possible differences, this time exploiting the integral domain  $(0..1)$  of the  $\pi$  parameter in combination with an appropriate weighting function.

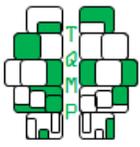
## Derivation of the exact test

Let  $R_1 = (x_1, n_1)$  and  $R_2 = (x_2, n_2)$  be the results of two binomial observations, and their observed difference  $d_O = x_1/n_1 - x_2/n_2$ , where  $R_1 \sim B(x_1|n_1, \pi_1)$ ,  $R_2 \sim B(x_2|n_2, \pi_2)$ , with  $\pi_1$  and  $\pi_2$  being the two unknown binomial probability parameters. The null hypothesis of no parametric difference stipulates that  $\pi_1 = \pi_2 = \pi_0$ . Under the null hypothesis and applying a one-sided test, what is the probability that  $d \geq d_O$ ?

The distribution of the random variable  $d = y_1/n_1 - y_2/n_2$  ( $0 \leq y_1 \leq n_1, 0 \leq y_2 \leq n_2$ ) is statistically conditioned on two factors: the size parameters  $n_1$  and  $n_2$ , which are given, and the two observed proportions  $x_1/n_1$  and  $x_2/n_2$ , with respect to which the likelihood of the underlying  $\pi_0$  value will vary.

With any given value of  $\pi_0$ , the probability sought for is  $P\{d \geq d_O\} = p(H_0)$  is :

$$p(\pi_0) = \sum_{y_1, y_2} b(y_1|n_1, \pi_0) \times b(y_2|n_2, \pi_0), \quad (1)$$



where  $b(y|n, \pi)$  is the binomial probability given by  $\binom{n}{y} \pi^y (1 - \pi)^{n-y}$ . The sum is computed using every value of  $y_1$  and  $y_2$ ,  $0 \leq y_1 \leq n_1$ ,  $0 \leq y_2 \leq n_2$ , for which

$$y_1/n_1 - y_2/n_2 \geq x_1/n_1 - x_2/n_2 \quad (2)$$

(for a right-side difference test). However, as pointed out above, any value of  $\pi_0$  is not as likely as another, considering the observed, real values of binomial variables  $x_1$  and  $x_2$ . Indeed, the likelihood of a given  $\pi_0$  value with respect to the observed data is simply

$$L(\pi_0|x_1, n_1; x_2, n_2) = \pi_0^{x_1+x_2} \pi_0^{n_1+n_2-x_1-x_2}. \quad (3)$$

This function, akin to the Beta function (Johnson, Kotz, & Balakrishnan, 1995), has here its maximum value at  $\hat{\pi} = (x_1 + x_2)/(n_1 + n_2)$ , the so-called maximum likelihood estimate used in Liddell (1978) test procedure.

Our procedure, which encompasses the whole domain of  $\pi_0$ , consists in weighting the summation (1) by the likelihood function (3), resulting in the following formula (see Laurencelle, 2017, for more details):

$$P\{d_{random} \geq d_O\} = \frac{(N + 1)!n_1!n_2!}{(2N + 1)!X!(N - X)!} \sum_{y_1, y_2}^* \frac{(X + y_1 + y_2)!(2N - X - y_1 - y_2)!}{y_1!(n_1 - y_1)!y_2!(n_2 - y_2)!} \quad (4)$$

where  $N = n_1 + n_2$ ,  $X = x_1 + x_2$ , and the right-hand summation  $\sum^*$  of (4) is restricted to  $(y_1, y_2)$  pairs such that  $y_1/n_1 - y_2/n_2 \geq x_1/n_1 - x_2/n_2$  for a right-side difference test, and vice-versa for a left-side test.

In Listing 1 and Listing 2 are provided two computer codes, one for Pascal/Delphi language, the other for R language, which perform the computation of  $p\{d_{random} \geq d_O\}$ . For a two-sided test, state the probability  $\min(2 \times p, 1)$ ,  $p$  begin likely to overflow  $1/2$  due to the discrete nature of the distribution of  $d$ .

### Two examples

Below are two data sets illustrating the results of formula (4) for the difference between two independent proportions, together with the results of some other test procedures. The classical Chi-squared and  $z$ -tests (without continuity correction) need not be presented. As for the Anscombe (1948), Chanter (1975) and Freeman and Tukey (1950) tests, here F-T, they are normal  $z$  tests using Fisher's arcsine normalizing transformation of proportions, each offering an original variant. A full-scale study of these and other variants is coming soon ("The statistical treatment of proportions including analysis of variance, with examples", in French in this journal).

### Authors' note

The author wishes to thank the Editor for his helpful and clarifying suggestions.

### References

- Anscombe, F. J. (1948). The transformation of poisson, binomial and negative-binomial data. *Biometrika*, 35, 246–254. doi:10.1093/biomet/35.3-4.246
- Chanter, D. O. (1975). Modifications of the angular transformation. *Journal of the Royal Statistical Society Series B (Applied Statistics)*, 24, 354–359. doi:10.2307/2347101
- Freeman, M. F., & Tukey, J. W. (1950). Transformations related to the angular and the square root. *Annals of Mathematical Statistics*, 21, 607–611. doi:10.1214/aoms/117729756
- Howell, D. C. (2017). *Fundamental statistics for the behavioral sciences (9th ed.)* Washington: Cengage Learning.
- Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions (vol. 2) (2nd ed.)*. New York: Wiley.
- Kendall, M. G., & Stuart, A. (1977). *The advanced theory of statistics, vol. 1*. New York.
- Laurencelle, L. (2017). Le test binomial exact de la difference entre deux proportions et ses approximations / the exact binomial test between two proportions and its approximations. *The Quantitative Methods for Psychology*, 13(2), 141–147. doi:10.20982/tqmp.13.2.p141
- Liddell, D. (1978). Practical tests of 2x2 contingency tables. *The Statistician*, 25, 295–304. doi:10.2307/2988087
- Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- Sokal, R. R., & Rohlf, F. J. (1995). *Biometry (3rd ed.)* New York: Freeman.

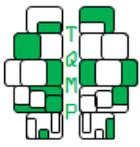


Table 1 ■ Two examples of computations

$x_1/n_1$	$x_2/n_2$	$p_1$	$p_2$	$d_O$	$p_{Exact}$	$p_{\chi^2}, p_z$	$p_{Anscombe}$	$p_{Chanter}$	$p_{F-T}$
13/20	15/29	0.650	0.517	0.133	0.183	0.178	0.183	0.183	0.186
13/112	89/473	0.188*	0.116*	0.072	0.032	0.035	0.031	0.030	0.033

Note. \* The programs being coded for right-tail extreme probability evaluation, the  $(x_1/n_1)$  and  $(x_2/n_2)$  data are automatically swapped so that  $p_1 \geq p_2$ .

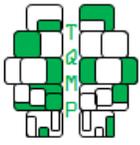
Appendix: The programs

Listing 1: Original Pascal / Delphi program

```

program difference_between_two_independent_proportions;
{$APPTYPE CONSOLE}
Uses SysUtils, Math;
Const nmax=10000; {>= 2 *(n1 + n2)}
var x1, x2, n1, n2, N, X, R, i, y1, y2: integer;
    lf: array[0..10000] of real;
    p1, p2, dif, lcons, lprob, sum: real;
begin
  writeln('L. Laurencelle, July 2005, 2017');
  writeln('Exact right-tail probability of the difference between two proportions');
  {to avoid overflow, precompute log of the factorials}
  lf[0]:=0;
  for i:=1 to nmax do lf[i]:=lf[i-1]+ln(i);
  repeat writeln;
    write('x1_n1: '); readln(x1, n1);
    write('x2_n2: '); readln(x2, n2);
    if x1/n1 < x2/n2 then begin y1:=x1; x1:=x2; x2:=y1; y1:=n1; n1:=n2; n2:=y1 end;
    p1:=x1/n1; p2:=x2/n2; dif:=p1-p2;
    write('p1=', p1:6:3, ' p2=', p2:6:3, ' dif=', dif:6:3);
    if abs(p1-p2)<=0 then writeln(' pExact=0.5') else
      begin
        N:=n1+n2; X:=x1+x2; R:=N-X;
        {Selective summation by enumeration}
        sum:=0;
        lcons:=lf[N+1]+lf[n1]+lf[n2]-lf[2*N+1]-lf[X]-lf[N-X];
        for y1:=0 to n1 do for y2:=0 to n2 do
          if y1/n1-y2/n2>=dif then
            begin
              lprob:=lf[X+y1+y2]+lf[2*N-X-y1-y2]-lf[y1]-lf[n1-y1]-lf[y2]- lf[n2-y2];
              sum:=sum+exp(lcons+lprob)
            end;
          write(' pEXACT=', sum:8:4);
        end
      until false
    end.

```

**Listing 2: R implementation**

```
# pExact computes the exact probability of a difference larger than
# the observed difference between the two proportions x1 over n1 vs. x2 over n2.
# L. Laurencelle, July 2005, 2017
pExact <- function(x1, n1, x2, n2) {
  if ((x1>n1)|(x2>n2)) stop("Invalid input: x1 (x2) must be smaller than n1 (n2)")

  # all computations are in log to avoid overflow
  lf = lfactorial # log(factorial(x)) = ln(x!)

  if ((x1/n1)<(x2/n2)) {# swap the values
    temp <- x1; x1 <- x2; x2 <- temp
    temp <- n1; n1 <- n2; n2 <- temp
  }

  # if equal proportions, then p is 0.5
  if (x1/n1 == x2/n2) {
    return(0.5)
  } else {
    N = n1+n2
    X = x1+x2
    total = 0
    lcons = lf(N+1)+lf(n1)+lf(n2) -lf(2*N+1)-lf(X)-lf(N-X)
    for (y1 in 0:n1) for (y2 in 0:n2)
      if (round( (y1/n1 - y2/n2) - (x1/n1 - x2/n2), 10)>=0) {
        lprob = lf(X+y1+y2)+lf(2*N-X-y1-y2)-lf(y1)-lf(n1-y1)-lf(y2)-lf(n2-y2)
        total = total + exp(lcons + lprob)
      }
    return(total)
  }
}
```

**Citation**

Laurencelle, L. (2021). The exact binomial test between two independent proportions: A companion. *The Quantitative Methods for Psychology*, 17(2), 76–79. doi:[10.20982/tqmp.17.2.p076](https://doi.org/10.20982/tqmp.17.2.p076)

Copyright © 2021, Laurencelle. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 07/05/2021 ~ Accepted: 14/05/2021