Two-sided Exact Tests and Matching Confidence Intervals for Discrete Data

Michael P. Fay

National Institute of Allergy and Infectious Diseases

useR! 2010 Conference
July 21, 2010
Motivating Example 1: Fisher’s exact Test for 2×2 Table

<table>
<thead>
<tr>
<th></th>
<th>Homozygous for CCR5Δ32 mutation</th>
<th>Wild Type or Heterozygous for CCR5Δ32 mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdominal Pain</td>
<td>4 (26.7%)</td>
<td>50 (8.1%)</td>
</tr>
<tr>
<td>No Abdom. Pain</td>
<td>11 (73.3%)</td>
<td>569 (91.9%)</td>
</tr>
</tbody>
</table>

Relationship of CCR5Δ32 mutation (genetic recessive model) to Early Symptoms with West Nile Virus Infection (from Lim, et al, J Infectious Diseases, 2010, 178-185)
Step 1: Create 2 by 2 Table

> abdpain<-matrix(c(4,50,11,569),2,2,  
+       dimnames=list(c("Abdominal Pain","No Abdom. Pain"),  
+       c("Homo","WT/Hetero"))  
> abdpain

   Homo  WT/Hetero
Abdominal Pain  4      50
No Abdom. Pain  11     569
Step 2: Run test

```r
> fisher.test(abdpain)

Fisher's Exact Test for Count Data

data:  abdpain
p-value = 0.03166
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  0.9235364 14.5759712
sample estimates:
  odds ratio
    4.122741
```
Test-CI Inconsistency

Problem: Test rejects but confidence interval includes odds ratio of 1.

- Same problem in:
  - R (fisher.test), Version 2.11.1,
  - SAS (Proc Freq), Version 9.2 and
  - StatXact, (StatXact 8 Procs).

- In all 3: One and only one exact confidence for odds ratio for the 2 by 2 table is given, AND

- the confidence interval is not an **inversion** of the usual two-sided Fisher’s exact test.
  - (Test defined the same way in all 3 programs).
Example 2: One Sample Binomial Test

Observe 10 out of 100 from a simulation. Is this significantly different from a true proportion of 0.05?

> binom.test(10,100,p=0.05)

    Exact binomial test

  data:  10 and 100
number of successes = 10, number of trials = 100, p-value = 0.03411
alternative hypothesis: true probability of success is not equal to 0.05
95 percent confidence interval:
  0.04900469 0.17622260
sample estimates:
probability of success
    0.1
Example 3: Two Sample Poisson Test

If we observe rates 2/17887 (about 11.2 per 100,000) for the standard treatment and 10/20000 (50 per 100,000) for new treatment, do these two groups significantly differ by exact Poisson rate test?

```r
> poisson.test(c(10, 2), c(20000, 17877))

Comparison of Poisson rates
data:  c(10, 2) time base:  c(20000, 17877)
count1 = 10, expected count1 = 6.336, p-value = 0.04213
alternative hypothesis: true rate ratio is not equal to 1
95 percent confidence interval:
  0.952422 41.950915
sample estimates:
  rate ratio
  4.46925
```
What is happening in the examples?

- In each example, we used an exact test and an exact confidence interval, **but**, the confidence interval is **not** an inversion of the test.
What is happening in the examples?

- In each example, we used an exact test and an exact confidence interval, but,
- the confidence interval is not an inversion of the test.
- Definition: confidence interval by inversion of (a series of) tests = all parameter values that fail to reject point null hypothesis.
Definition: Inversion of Family of Tests

- Consider a series of tests, indexed by $\beta_0$
- Let $x$ be data.
- Let $p_{\beta_0}(x)$ be p-value for testing the following hypotheses:

$$
H_0 : \quad \beta = \beta_0 \\
H_1 : \quad \beta \neq \beta_0
$$

Then the inversion confidence set is

$$
C(x, 1-\alpha) = \{ \beta : p_{\beta}(x) > \alpha \}
$$

Cannot have test-confidence set inconsistency with inversion confidence set.
Figure: CCR5 data: Abdominal Pain, usual two-sided Fisher’s exact p-values
Figure: CCR5 data: Abdominal Pain, 95 % inversion confidence interval to usual two-sided Fisher’s exact
Another two-sided Fisher’s exact Test

- Define p-value as 2 times minimum of the one-sided Fisher’s exact p-values.
- Inversion of that two sided Fisher’s exact is the usual exact confidence intervals.
- Call it Central Fisher’s exact Test
Figure: CCR5 data: Abdominal Pain, gray = usual two-sided Fisher's exact p-values, red = twice minimum one-sided p-values
Figure: CCR5 data: Abdominal Pain, 95% central confidence intervals

95% central CI=(0.92,14.6)
twice one-sided p=0.063
Figure: CCR5 data: Abdominal Pain, 95 % central confidence intervals

95 % central CI=(0.92,14.6)
95 % minlike CI=(1.17,14.2)
usual two−sided p=0.032
twice one−sided p=0.063

β0
Figure: CCR5 data: Abdominal Pain, 95 % central confidence intervals

95 % central CI=(0.92,14.6)
95 % minlike CI=(1.17,14.2)

usual two−sided p=0.032

twice one−sided p=0.063

β_0
3 Ways to Calculate Two-sided p-values

**central:** 2 times minimum of one-sided p-values,

**minlike:** sum of probabilities of outcomes with likelihoods less than or equal to observed.

\[ p_m(x) = \sum_{X : f(X) \leq f(x)} f(X) \]

**blaker:** take smaller observed tail and add largest probability on the opposite tail that does not exceed observed tail.
x=number of events in treatment group

noncentral hypergeometric density

Pr[X<=4]= 0.0877
Pr[X>=10]= 0.0883

two−sided p−value= 0.176
(black+gray+green+blue)
central p−value= 0.175
2*(black+gray)

two−sided p−value= 0.116
(black+gray+blue)

odds ratio= 10.55

Figure: CCR5 data: Abdominal Pain
Solution: Use “Matching” Confidence Intervals

Smallest confidence interval that contains all parameters that fail to reject.

> library(exact2x2)
Loading required package: exactci
> fisher.exact(abdpain)

Two-sided Fisher's Exact Test (usual method using minimum likelihood)

data:  abdpain
p-value = 0.03166
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  1.1734 14.1659
sample estimates:
  odds ratio
  4.122741
Solution: Use “Matching” Confidence Intervals

> fisher.exact(abdpain, tsmethod="central")

    Central Fisher's Exact Test

    data:  abdpain
    p-value = 0.06332
    alternative hypothesis: true odds ratio is not equal to 1
    95 percent confidence interval:
    0.9235364 14.5759712
    sample estimates:
    odds ratio
    4.122741
Solution: Use “Matching” Confidence Intervals

> blaker.exact(abdpain)

Blaker's Exact Test

data:  abdpain
p-value = 0.03166
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  1.1734 14.2183
sample estimates:
  odds ratio
  4.122741
Example 2: One Sample Binomial

\begin{verbatim}
> library(exactci)
> binom.exact(10,100,p=0.05)

    Exact two-sided binomial test (central method)

data:  10 and 100
number of successes = 10, number of trials = 100, p-value = 0.05638
alternative hypothesis: true probability of success is not equal to 0.05
95 percent confidence interval:
 0.04900469 0.17622260
sample estimates:
probability of success
   0.1
\end{verbatim}
Example 2: One Sample Binomial

> binom.exact(10,100,p=0.05,tmethod="minlike")

        Exact two-sided binomial test (sum of minimum likelihood method)

data:  10 and 100
number of successes = 10, number of trials = 100, p-value = 0.03411
alternative hypothesis: true probability of success is not equal to 0.05
95 percent confidence interval:
  0.0534 0.1740
sample estimates:
  probability of success
        0.1
Example 2: One Sample Binomial

> binom.exact(10, 100, p=0.05, tsmethod="blaker")

    Exact two-sided binomial test (Blaker's method)

data:  10 and 100
number of successes = 10, number of trials = 100, p-value = 0.03411
alternative hypothesis: true probability of success is not equal to 0.05
95 percent confidence interval:
  0.0513  0.1723
sample estimates:
  probability of success
                 0.1
Example 3: Two Sample Poisson

\[
\text{poisson.exact(c(10,2),c(20000,17877))}
\]

**Exact two-sided Poisson test (central method)**

data:  c(10, 2) time base:  c(20000, 17877)
count1 = 10, expected count1 = 6.336, p-value = 0.06056
alternative hypothesis: true rate ratio is not equal to 1
95 percent confidence interval:
  0.952422 41.950915
sample estimates:
  rate ratio
    4.46925
Example 3: Two Sample Poisson

```
> poisson.exact(c(10,2),c(20000,17877),tsmethod="minlike")

Exact two-sided Poisson test (sum of minimum likelihood method)

data:  c(10, 2) time base:  c(20000, 17877)
count1 = 10, expected count1 = 6.336, p-value = 0.04213
alternative hypothesis: true rate ratio is not equal to 1
95 percent confidence interval:
   1.061630 28.412707
sample estimates:
  rate ratio
   4.46925
```
Example 3: Two Sample Poisson

```r
> poisson.exact(c(10,2),c(20000,17877),tsmethod="blaker")

Exact two-sided Poisson test (Blaker's method)

data:  c(10, 2) time base:  c(20000, 17877)
count1 = 10, expected count1 = 6.336, p-value = 0.04213
alternative hypothesis: true rate ratio is not equal to 1
95 percent confidence interval:
  1.068068 28.412707
sample estimates:
rate ratio
  4.46925
```
An Anomaly: Unavoidable Test-CI Inconsistency

Made-Up Example:

<table>
<thead>
<tr>
<th></th>
<th>Group A</th>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>7 (2.67 %)</td>
<td>30 (6.07 %)</td>
</tr>
<tr>
<td>No Event</td>
<td>255 (97.33 %)</td>
<td>464 (93.93 %)</td>
</tr>
</tbody>
</table>

- usual two-sided Fisher’s exact test \( p = 0.04996 \)
- 95% inversion confidence set:
  \[
  \{ \beta : \beta \in (0.177, 0.993) \text{ or } \beta \in (1.006, 1.014) \}
  \]

Matching CI defined as smallest interval that contains all elements of inversion confidence set:

\[
(0.177, 1.014)
\]

Unavoidable test-CI inconsistency!
Figure: Made-up example, gray=usual two-sided Fisher’s exact, blue=Blaker’s exact p-values, red=twice minimum one-sided p-values
References

- Fay (2010) Biostatistics 373-374
- R package: exact2x2
- R package: exactci