Optimal Data Analysis
Vol. 2, Release 2 (December 1, 2013), 222-227

2155-0182/10/$3.00

Confirmatory Analysis for an Ordered Series of a Dichotomous Attribute: Airborne Radiation and Congenital Hypothyroidism of California Newborns

Paul R. Yarnold, Ph.D., and Robert C. Soltysik, M.S.
Optimal Data Analysis, LLC

Ordered series involving a dichotomous (binary) variable are widely used to describe changes in phenomena which occur across time. Examples of such series include the percentage of a sample or population each year (or other unit of time) that marries, dies, or is arrested. This article demonstrates how UniODA is used for such designs to test a confirmatory (*a priori*), omnibus (overall), optimal (maximum-accuracy) hypothesis, subsequently disentangled by a confirmatory (if hypotheses are composed) or otherwise by an exploratory (*post hoc*) optimal range test. This methodology is demonstrated with an example assessing the effect of airborne beta nuclear radiation emanating from the Fukushima nuclear meltdown on the risk of congenital hypothyroidism (CH) for newborns in California in the years 2011-2012.

Data for exposition are extracted from recent research investigating the effect of increased airborne radiation originating from the Fukushima nuclear meltdown on the risk of congenital hypothyroidism (CH) for newborns in California in 2011-2012.[1] Table 1 gives the data and omnibus *a priori* hypothesis developed for borderline plus confirmed cases of CH, and Table 2 gives the data and the identical *a priori* hypothesis for only the confirmed cases.[1]

Table 1: Data and *A Priori* Hypothesis for *Borderline* Plus *Confirmed* CH of Newborns in California

Dates	1/1-3/16/2011	3/17-6/30/2011	7/1-12/31/2011	1/1-3/16/2012	3/17-6/30/2012	7/1-12/31/2012
Date Code	1	2	3	4	5	6
Exposure	Not Exposed	Exposed	Exposed	Not Exposed	Not Exposed	Not Exposed
Class Code	Class=**0**	Class=**1**	Class=**1**	Class=**0**	Class=**0**	Class=**0**
n Class=1	500	777	1,360	450	504	1,079
Total *n*	99,953	142,592	252,874	99,122	138,529	259,056
% Class=1	0.500	0.545	0.538	0.454	0.364	0.417

2155-0182/10/$3.00

Table 2: Data and *A Priori* Hypothesis for *Confirmed* CH of Newborns in California

Dates	1/1-3/16/2011	3/17-6/30/2011	7/1-12/31/2011	1/1-3/16/2012	3/17-6/30/2012	7/1-12/31/2012
Date Code	1	2	3	4	5	6
Exposure	Not Exposed	Exposed	Exposed	Not Exposed	Not Exposed	Not Exposed
Class Code	Class=**0**	Class=**1**	Class=**1**	Class=**0**	Class=**0**	Class=**0**
n Class=1	69	94	200	63	88	144
Total *n*	99,953	142,592	252,874	99,122	138,529	259,056
% Class=1	0.0690	0.0659	0.0791	0.0636	0.0635	0.0556

Table 3: Data and *A Priori* Hypothesis for *Borderline* CH of Newborns in California

Dates	1/1-3/16/2011	3/17-6/30/2011	7/1-12/31/2011	1/1-3/16/2012	3/17-6/30/2012	7/1-12/31/2012
Date Code	1	2	3	4	5	6
Exposure	Not Exposed	Exposed	Exposed	Not Exposed	Not Exposed	Not Exposed
Class Code	Class=**0**	Class=**1**	Class=**1**	Class=**0**	Class=**0**	Class=**0**
n Class=1	431	683	1,160	387	416	935
Total *n*	99,884	142,498	252,674	99,059	138,441	258,912
% Class=1	0.432	0.479	0.459	0.391	0.300	0.361

As seen, Table 1 combines newborns with borderline and confirmed CH. In Table 3 the number of confirmed cases (see Table 2) was subtracted from *n* Class=1 and Total *n* in Table 1, yielding data and *a priori* hypothesis for only the *borderline* newborns.

Original Approximate Statistical Analysis

The original analysis used a Mantel Haenszel chi-square test and reported: "…the CH rate for the cohort born March 17 to December 31, 2011, or those exposed *in utero* to environmental radioactivity from Fukushima was significantly elevated from those born in other periods in 2011 and 2012, for confirmed cases (p<0.013). Adding borderline cases to the confirmed cases resulted in a much higher level of statistical significance (p< 0.00000001)." At this point the original analysis was completed.

This analysis process reflects the modus operandi in applied research using this method.[2] However, when viewed from the perspective of the ODA paradigm some issues are unresolved.[3] Specifically, the analysis did *not* (a) determine if combining the borderline and confirmed cases resulted in Simpson's paradox[4]; (b) provide an index of strength of effect[3] (only of Type I error rate); (c) evaluate the assumptions which are required for the Mantel Haenszel chi-square test to be a valid, although inherently two-tailed[2] approximate statistic; (d) disentangle omnibus effects; (e) analyze data in Table 3; or (f) correct for multiple tests of statistical hypotheses to control experimentwise Type I error rate (this didn't affect the outcome of analysis presently).[3]

Exact Optimal (Maximum-Accuracy) Confirmatory Statistical Analysis

The *a priori* hypothesis that exposure to airborne beta-radiation arising from the nuclear meltdown is tested using MegaODA software because the sample is too large for UniODA software.[5-8] Data are ordered over time but they can't be ipsatively standardized because the data in each time period are presented as a scalar, each time for a different population.[1,9-12]

An ASCII data set (nuclear.dat) was constructed in which every observation formed a separate row.[13] Data included the date code called *code* (1-6); the class variable called *exposure* (if the new-born was exposed then exposure=1, and if the newborn was not

2155-0182/10/$3.00

exposed then exposure=0); and three binary attributes. The first attribute, *confirm*, was coded as 1 if thyroid stimulating hormone (TSH)>29 µIL/ml, and as 0 otherwise (Table 2).[1] The second attribute, *border*, was coded as missing (-9) if TSH>29 µIL/ml, as 1 if TSH> 19 µIL/ml, and as 0 otherwise (Table 3). The third attribute, *both*, was coded as 1 if TSH>19 µIL/ml, and as 0 otherwise (Table 1).

The planned analysis unfolds in three separate stages, with each stage considering one of the three attributes.

Confirmed Congenital Hypothyroidism

For expository purposes the first analysis tests the omnibus hypothesis for the confirmed cases of CH using the special-purpose TABLE command in MegaODA (and UniODA) software.[3] The numerators (94+200) and then the denominators (142,492+ 252,874) of the class=1 cells in Table 2 are integrated (added); the same procedure is performed for class=0 observations (364 and 586,660, respectively); and the results are used to create tabular input as shown for the present application for MegaODA software (the data table indicates the denominator minus the numerator as entries in the first column).[3]

```
OPEN DATA;
OUTPUT nuclear.out;
CAT ON;
TABLE 2;
CLASS ROW;
DIR < 1 2;
DATA;
586296 364
395072 294
END DATA;
GO;
```

Monte Carlo (MC) simulation is *not* used to estimate exact *p*, because for binary problems the randomization algorithm in ODA software and Fisher's exact test are isomorphic, and the exact *p* is provided in program output).[3] The resulting test of the *a priori* hypothesis is statistically significant at the experimentwise criterion[3] (for the initial test this is per-comparison $p<0.05$): exact per-comparison $p<0.00664$. However, the accuracy of the model is minute (ESS=0.01, where 0=the accuracy expected by chance, and 100=perfect accuracy), and the predictive value (PV) of the model is very weak (ESP= 4.83, where 0=the PV expected by chance, and 100=perfect PV).[3,14] The finding of the omnibus test is symbolically represented as:

(2,3) > (1,4,5,6),

where commas within parentheses indicate relationships which must be disentangled—that is, replaced by equality or inequality signs on the basis of an optimal range test.[15,16]

The left-hand parentheses of the symbolic representation is disentangled by replacing the data table in the code above with the following data table; deleting the DIR command (because no *a priori* hypothesis concerning a comparison of these two groups was made); and then re-running the analysis (Sidak criterion for experimentwise $p<0.05$ for two tests of statistical hypotheses is per-comparison $p< 0.0253$).[3]

```
142498 94
252674 200
```

For this analysis exact $p<0.162$ (ESS= 0.01; ESP=4.1), and the symbolic representation is updated:

(2=3) > (1,4,5,6).

While the TABLE command may be used to efficiently disentangle binary effects, larger problems must be disentangled using the data file created earlier in combination with the following MegaODA software commands (indicated in red). The Monte Carlo simulator is set for a third test of a statistical hypothesis with experimentwise $p<0.05$.

2155-0182/10/$3.00

```
OPEN nuclear.dat;
OUTPUT nuclear.out;
VARS code exposure confirm border both;
CLASS exposure;
ATTR confirm;
CAT confirm border both;
MISSING ALL (-9);
EX code=2;
EX code=3;
MC ITER 25000 TARGET .05 SIDAK 3
    STOP 99.9 STOPUP 99.9;
GO;
```

The resulting model was not statistically significant: estimated $p<0.33$, confidence for $p>0.10$ is >99.99%, ESS=3.9, ESP=0.01). The symbolic representation is thus completed:

(2=3) > (1=4=5=6).

These results indicate that for confirmed cases the CH rate for the cohort born March 17 to December 31, 2011, or those exposed *in utero* to environmental radioactivity from Fukushima (which were statistically comparable), were significantly elevated from those born in other periods in 2011 and 2012 (which were statistically comparable). Although statistically significant, the effect is very weak if considered from an ecological perspective.

Borderline Congenital Hypothyroidism

Analyses conducted for the confirmed CH data were repeated for the borderline data. The test of the omnibus *a priori* hypothesis is statistically significant at the experimentwise criterion (exact per-comparison $p<0.29\text{x}10^{-14}$). As for analysis of confirmed data, the accuracy of the model for borderline data is minute (ESS =0.10) and the PV of the model is very weak (ESP= 6.10).The finding of the omnibus *a priori* test is symbolically represented as before:

(2,3) > (1,4,5,6).

The test for the left-hand parentheses was not statistically significant (exact $p<0.37$; ESS=0.02; ESP=1.0), and symbolic representation is updated:

(2=3) > (1,4,5,6).

In contrast to the findings for confirmed cases, for borderline cases the test for the right-hand parentheses was statistically significant at the experimentwise criterion (estimated $p=0$, confidence for $p<0.01696$ is >99.9%; ESS=4.4; ESP=0.07). The ODA model was: if class=1 or class=4 then predict proportion of borderline CH cases is greater than for class=5 or class=6. The symbolic representation is updated:

(2=3) > (1,4) > (5,6).

The test comparing classes 1 and 4 was not statistically significant (exact $p<0.17$; ESS= 2.5; ESP=0.04), and the symbolic representation is refreshed:

(2=3) > (1=4) > (5,6).

Finally, the test comparing classes 5 and 6 was statistically significant at the experimentwise criterion (exact $p<0.0017$; ESS=4.4; ESP=0.07), and the symbolic representation is completed:

(2=3) > (1=4) > 6 > 5.

These results indicate that for boderline cases the CH rate for the cohort born March 17 to December 31, 2011, or those exposed *in utero* to environmental radioactivity from Fukushima (which were statistically comparable), were significantly elevated from those born before the Fukushima catastrophe, or born in the first three months of 2012 (which were statistically comparable). Compared to these four time periods the borderline CH rate was significantly lower for the most recent six-month period, and was significantly lowest between March 17 and June 30 of 2012.

2155-0182/10/$3.00

Summary of Findings

For confirmed CH cases there were essentially two rates: moderate and high, and the exposed newborns had a high rate while unexposed newborns had a moderate rate.

For borderline CH cases there were four rates: high, moderate, low, and very low. The exposed newborns had a high rate; newborns in the period immediately before the nuclear event, or immediately after the elevated radioactivity diminished had a moderate rate; newborns in the subsequent three months had a very low rate; and newborns in the most recent six-month period had a low rate of borderline CH.

Clearly these data should not be combined, as they had different underlying structure, and combining different structures often generates spurious findings attributable to paradoxical confounding.[4] In addition, as was stated earlier, these data represent an ordered series: ipsative transformation is therefore necessary to circumvent paradoxical confounding if the measured variables (e.g., radiation level, CH rate) differ over time. Preliminary research in the ODA laboratory suggests that both of these measures have changed over time, and thus double ipsative transformations are needed: this is the subject of a working manuscript.

It is also interesting to note that in the case of borderline cases, an exploratory analysis would likely identify the same model that was identified testing the *a priori* hypothesis. However, for confirmed data, it appears that substantially different result would accrue: the data in 2011 have high, homogeneous CH rates; the CH rate in the latter half of 2012 is lowest; and the data for the first half of 2012 are homogeneous and intermediate.

References

[1]Mangano J, Sherman J, Busby C (2013). Changes in confirmed plus borderline cases of congenital hypothyroidism in California as a function of environmental fallout from the Fukushima nuclear meltdown. *Open Journal of Pediatrics*, *3*, 370-376.

[2]Grimm LG, Yarnold PR (Eds., 2000). *Reading and understanding more multivariate statistics*. Washington, DC: APA Books.

[3]Yarnold PR, Soltysik RC (2005). *Optimal data analysis: Guidebook with software for Windows*. Washington, D.C.: APA Books.

[4]Yarnold PR (1996). Characterizing and circumventing Simpson's paradox for ordered bivariate data. *Educational and Psychological Measurement*, *56*, 430-442.

[5]Yarnold PR, Feinglass J, Martin GJ, McCarthy WJ (1999). Comparing three pre-processing strategies for longitudinal data for individual patients: An example in functional outcomes research. *Evaluation in the Health Professions*, *22*, 254-277.

[6]Soltysik RC, Yarnold PR (2013). MegaODA large sample and BIG DATA time trials: Separating the chaff. *Optimal Data Analysis*, *2*, 194-197.

[7]Soltysik RC, Yarnold PR (2013). MegaODA large sample and BIG DATA time trials: Harvesting the Wheat. *Optimal Data Analysis*, *2*, 202-205.

[8]Yarnold PR, Soltysik RC (2013). MegaODA large sample and BIG DATA time trials: Maximum velocity analysis. *Optimal Data Analysis*, *2*, 220-221.

[9]Yarnold PR, Soltysik RC (2013). Ipsative transformations are *essential* in the analysis of serial data. *Optimal Data Analysis*, 2, 94-97.

[10]Yarnold, P.R. (2013). Comparing attributes measured with "identical" Likert-type scales in single-case designs with UniODA. *Optimal Data Analysis*, 2, 148-153.

2155-0182/10/$3.00

[11]Yarnold, P.R. (2013). Comparing responses to dichotomous attributes in single-case designs. *Optimal Data Analysis*, 2, 154-156.

[12]Yarnold PR (2013). Ascertaining an individual patient's *symptom dominance hierarchy*: Analysis of raw longitudinal data induces Simpson's Paradox. *Optimal Data Analysis*, 2, 159-171.

[13]Bryant FB, Harrison PR (2013). How to create an ASCII input data file for UniODA and CTA software. *Optimal Data Analysis*, *2*, 2-6.

[14]Yarnold PR (2013). Standards for reporting UniODA findings expanded to include ESP and all possible aggregated confusion tables. *Optimal Data Analysis*, 2, 106-119.

[15]Yarnold PR, Brofft GC (2013). ODA range test *vs*. one-way analysis of variance: Comparing strength of alternative line connections. *Optimal Data Analysis*, 2, 198-201.

[16]Yarnold PR (2013). ODA range test *vs*. one-way analysis of variance: Patient race and lab results. *Optimal Data Analysis*, 2, 206-210.

Author Notes

ODA Blog: odajournal.com