Optimal Data Analysis
Vol. 5 (May 29, 2016), 31-36

2155-0182/10/$3.00

Ascertaining Intervention Efficacy

Paul R. Yarnold, Ph.D.
Optimal Data Analysis, LLC

UniODA[1,2] is used to study the effect of interruptions on the course of behavior normally seen in longitudinal (temporal) series, for cases or for groups, in applications such as modeling of mortality rates after exposure to environmental toxins[3-7], evaluating symptom reduction in chronic disease after human[8-10] or artificial[11-13] therapy, or assessing the validity of efficacy claims (regarding an outcome) associated with change in public policy.[14,15] This paper uses UniODA to assess the immediate short-term longevity of efficacy (if any) of back-to-back interventions (advertisements published on a hobby shop webpage) with respect to two serial outcomes each assessed as counts: the daily number of webpage visitors, and of page views.

Data were obtained for exposition as a convenience sample for an interrupted time-series design.[12] Daily number of visitors and of views for the webpage were recorded beginning 22 days prior to implementation of a two-stage advertising intervention. On Day 1 of the intervention the initial ad (first intervention, class=0) was posted, and the second ad (second intervention, class=1) was posted on Day 13. Each additional successive day of post-intervention follow-up is another one-unit increment in a two-stage prospective validity study that is unfolding in real time. The study terminated once 22 total days of post-intervention follow-up data were recorded. Table 1 gives raw data (ipsative standardization yielded identical findings[2,16,17]).

Two *a priori* hypotheses are tested for the number of visitors, and also for the number of views (considered to be separate experiments since they address different aspects of marketing success). First, it is hypothesized that there will increase in the number of (visitors, views) after the first intervention (coded as “0”) compared to the pre-intervention (baseline) period (“1”). It is also hypothesized that there will increase in the number of (visitors, views) after the second (“2”) versus the first intervention period.

First Intervention

The first analysis tested the *a priori* hypothesis that the daily number of visitors pre-intervention (Day -22 to -1) is lower versus after (Day 1 to 12) the first advertisement. Analysis unfolded sequentially: the first analysis compared Day 1 versus pre-intervention data; the second analysis compared Days 1 and 2 versus pre-intervention data, and so forth up to a maximum of 12 comparisons, the last contrasting the pre-intervention data versus Days 1 through 12. In each step the criterion to achieve statistical significance is adjusted for the number of tests of statistical hypotheses being made vis-à-vis a sequentially rejective Sidak Bonferroni-type multiple comparisons procedure.[1,2]

2155-0182/10/$3.00

Table 1: Raw Pre- and Post-Intervention Data (Counts)

Pre-Intervention Data (Class=0)			Post-Intervention Data (Class=1)		
Day	Views	Visitors	Day	Views	Visitors
-22	31	7	1	63	27
-21	34	9	2	53	27
-20	26	12	3	38	20
-19	21	13	4	40	18
-18	24	14	5	57	23
-17	26	15	6	39	16
-16	27	15	7	15	13
-15	6	5	8	35	10
-14	12	6	9	34	17
-13	12	9	10	33	15
-12	26	14	11	25	12
-11	15	11	12	31	16
-10	42	19	13	65	29
-9	32	23	14	91	17
-8	39	16	15	58	12
-7	16	10	16	56	14
-6	21	13	17	63	20
-5	25	13	18	56	13
-4	21	13	19	20	17
-3	32	18	20	18	11
-2	23	13	21	19	10
-1	14	8	22	31	11

Analysis in the first step was conducted using UniODA[1] and MegaODA[18-20] command syntax given below (all data from the second intervention, and from the first intervention except for Day 1, are excluded):

```
OUTPUT example.out;
OPEN example.dat;
VARS day interven visitors views;
CLASS interven;
ATTR visitors;
EX interven=3;
EX day>1;
DIR < 0 1;
MC ITER 25000;
GO;
```

In the second step the EX syntax for day was modified:

```
EX day>2;
```

Table 1 summaries the findings of the analysis, that concluded with the 7th step of the procedure: the finding in the 8th step wasn't statistically reliable because model p > Sidak criterion for experimentwise $p<0.05$.

2155-0182/10/$3.00

Table 1: Results for Number of Visitors: Pre-Intervention versus First Intervention

Day	*p* < Sidak	*p* < Model	Optimal Threshold	ESS
1	0.05	0.043	25	100.0
2	0.0254	0.0029	25	100.0
3	0.0170	0.0019	19.5	95.4
4	0.0128	0.0026	17	86.4
5	0.0103	0.0008	17	86.4
6	0.0086	0.0004	15.5	81.8
7	0.0074	0.0042	15.5	67.5
8	0.0064	0.011	15.5	56.8

In the Tables all *p*-values are rounded up and reported using sufficient precision to make a numerical comparison versus the associated Sidak criterion (ipsative standardization of view and visitor counts didn't alter findings herein).[2]

As hypothesized the first advertisement significantly increased (experimentwise $p<0.05$) the number of viewers compared to pre-advertisement levels, for seven days. The effect was perfect (ESS=100) for Days 1 and 2; very strong for Day 3; strong for Days 4 through 6; and relatively strong for Days 7 and 8.

On Days 1 and 2 of the intervention the optimal threshold value was 25, and this fell to 15.5 in the final three days (Table 1). As this threshold value (cutpoint) decreases ESS also decreases because model sensitivity (for post-intervention data) falls as values above the cutpoint become increasingly prevalent in pre-intervention data.

Even while maintaining strict control over the experimentwise Type I error rate this methodology is highly sensitive to incremental changes in the data. For example, two sequential days (7 and 8) with mediocre counts of new visitors were sufficient to terminate the experiment. However, had the count for Day 9 occurred after Day 6 (i.e., the "17" come the day before the "13"), then in Table 1, row 8 would instead have been:

8	0.0064	0.0002	15.5	81.8

If this had happened then the intervention would have been determined to have supported the *a priori* hypothesis for a total of 10 days, rather than the 7 day boost that actually materialized (the reader may wish to verify this conclusion).

The second analysis tested the *a priori* hypothesis that the daily number of views pre-intervention (Day -22 to -1) is lower versus after (Day 1 to 12) the first advertisement: the first comparison in made using UniODA/MegaODA command syntax modified as indicated below (data from the second intervention, and from the first intervention except Day 1, are excluded):

```
ATTR views;
EX day>1;
```

Table 2 summaries the findings of the analysis, that concluded with the 7th step of the procedure: the finding in the 8th step wasn't statistically reliable because model *p* > Sidak criterion for experimentwise $p<0.05$.

Table 2: Results for Number of Views: Pre-Intervention versus First Intervention

Day	*p* < Sidak	*p* < Model	Optimal Threshold	ESS
1	0.05	0.044	52.5	100.0
2	0.0254	0.0041	47.5	100.0
3	0.0170	0.0049	36	90.9
4	0.0128	0.0012	36	90.9
5	0.0103	0.0002	36	90.9
6	0.0086	0.0001	36	90.9
7	0.0074	0.0008	36	76.6
8	0.0064	0.0003	34.5	78.4
9	0.0057	0.0002	33	75.2
10	0.0052	0.0001	32.5	76.4
11	0.0047	0.0007	32.5	68.2
12	0.0043	0.0012	32.5	61.4

As hypothesized the first advertisement significantly increased (experimentwise p<0.05) the number of views compared to pre-advertisement levels, for twelve days. The effect was perfect (ESS=100) for Days 1 and 2; very strong for Days 3-6; strong for Days 7-9; and relatively strong for Days 10-12.

Second Intervention

The third analysis tested the *a priori* hypothesis that the daily number of visitors after the first advertisement (Day 1 to 12) is lower versus after the second advertisement (Day 13 to 22). The analysis unfolded sequentially: the first analysis compared Days 13 and 14 versus Days 1-12; the second analysis compared Days 13-15 versus Days 1-12, and so forth up to a maximum of nine comparisons, the last contrasting the Days 13-22 versus Days 1-12. In each step the criterion to achieve statistical significance is adjusted for the number of tests of statistical hypotheses being made using a sequentially rejective Sidak Bonferroni-type multiple comparisons procedure.

Unlike the initial pair of analyses, the comparisons involving the second intervention didn't begin using only the first day of the second intervention due to insufficient statistical power: at least 20 days of data are needed in the first intervention period to obtain p<0.05 for ESS=100 with only one observation in the second intervention period.[2] Instead, adding one more observation to the second intervention, for ESS=100, p<0.011. The first axiom of novometric theory mandates adequate statistical power, and it behooves researchers to comply.[2]

Analysis in the first step was conducted using the UniODA/MegaODA command syntax modified as indicated below (pre-intervention data, and all data from the second intervention except Days 13 and 14, are excluded):

```
ATTR visitors;
DIR < 1 2;
EX interven=0;
EX day>14;
```

The results of the analysis failed to reject the null hypothesis: p<0.38, ESS=50.0. It is thus concluded that the second advertisement didn't increase the number of visitors relative to the first intervention.

The final analysis tested the *a priori* hypothesis that the daily number of views after the first advertisement (Day 1-12) is lower versus after the second advertisement (Day 13-22). The comparison in made vis-à-vis the UniODA/MegaODA command syntax given below (pre-intervention data, and all data from the second intervention except Days 13 and 14, are excluded):

```
ATTR views;
EX day>14;
```

Table 3 summaries the findings of the analysis, that concluded with the 5th step (Day 18) of the procedure: the finding in the 6th step (Day 19) wasn't statistically reliable because model p > Sidak criterion for experimentwise p<0.05.

Table 3: Results for Number of Views: First versus Second Intervention

Day	p < Sidak	p < Model	Optimal Threshold	ESS
13-14	0.05	0.011	64	100.0
15	0.0254	0.011	57.5	91.7
16	0.0170	0.0089	54.5	83.3
17	0.0128	0.0032	54.5	83.3
18	0.0103	0.0020	54.5	83.3
19	0.0086	0.0087	54.5	69.1

As hypothesized the second advertisement significantly increased (experimentwise p<0.05) the number of views compared to first-advertisement levels, for six days. The effect was perfect (ESS=100) for the first two days;

2155-0182/10/$3.00

very strong for the third day; and strong for the final three days of the second intervention.

Disciplined adherence to statistical rules may not mesh perfectly with all corporate cultures. Fortunately, stopping rules governing the experimental design are very flexible. For example, the experimental protocol could call for eliminating data from weekends and from holidays if there is a defensible reason for the exclusion and inclusion criteria (this practice is common in medical research). This method can also be used to investigate time-lagged efficacy. For example, one may track time-to-onset of the first significant boost for slower-accelerating series. Or one may identify regions of significant gains or losses over the domain of the series to identify underlying patterns over time.

References

[1]Yarnold PR, Soltysik RC. *Optimal data analysis: Guidebook with software for Windows*. Washington, DC: APA Books, 2005.

[2]Yarnold PR, Soltysik RC (2016). *Maximizing predictive accuracy*. Chicago, IL: ODA Books.

[3]Yarnold PR (2013). Statistically significant increases in crude mortality rate of North Dakota counties occurring after massive environmental usage of toxic chemicals and biocides began there in 1998: An optimal *static* statistical map. *Optimal Data Analysis*, *2*, 98-105.

[4]Yarnold PR (2013). The most recent, earliest, and *K*th significant changes in an ordered series: Traveling backwards in time to assess when annual crude mortality rate most recently began increasing in McLean County, North Dakota. *Optimal Data Analysis*, *2*, 143-147.

[5]Yarnold PR (2013). Determining when annual crude mortality rate most recently began increasing in North Dakota counties, I: Backward-stepping little jiffy. *Optimal Data Analysis*, *2*, 217-219.

[6]Yarnold PR, Soltysik RC (2013). Confirmatory analysis for an ordered series of a dichotomous attribute: Airborne radiation and congenital hypothyroidism of California newborns. *Optimal Data Analysis*, *2*, 222-227.

[7]Yarnold PR, Soltysik RC (2013). Exploratory analysis for an ordered series of a dichotomous attribute: Airborne radiation and congenital hypothyroidism of California newborns. *Optimal Data Analysis*, *2*, 228-231.

[8]Collinge W, Kahn J, Walton T, Kozak L, Bauer-Wu S, Fletcher K, Yarnold PR, Soltysik RC (2013). Touch, Caring, and Cancer: Randomized controlled trial of a multimedia caregiver education program. *Supportive Care in Cancer*, *21*, 1405-1414.

[9]Collinge W, Yarnold PR (2001). Transformational breath work in medical illness: Clinical applications and evidence of immunoenhancement. *Subtle Energies & Energy Medicine*, *12*, 139-156.

[10]Collinge W, Yarnold PR, Raskin E (1998). Use of mind/body self-healing practice predicts positive health transition in chronic fatigue syndrome: A controlled study. *Subtle Energies & Energy Medicine*, *9*, 171-190.

[11]Collinge WC, Soltysik RC, Yarnold PR (2010). An internet-based intervention for fibromyalgia self-management: Initial design and alpha test. *Optimal Data Analysis*, *1*, 163-175.

[12]Yarnold PR, Soltysik RC, Collinge W (2013). Modeling individual reactivity in serial designs: An example involving changes in weather and physical symptoms in fibromyalgia. *Optimal Data Analysis*, *2*, 37-42.

2155-0182/10/$3.00

[13]Collinge W, Yarnold PR, Soltysik RC (2013). Fibromyalgia symptom reduction by online behavioral self-monitoring, longitudinal single subject analysis and automated delivery of individualized guidance. *North American Journal of Medical Sciences, 5,* 546-553.

[14]Linden A, Yarnold PR (2016). Using machine learning to identify structural breaks in single-group interrupted time series designs. *Journal of Evaluation in Clinical Practice*. DOI: 10.1111/jep.12544

[15]Yarnold BM, Yarnold PR (2010). Maximizing the accuracy of Probit models via UniODA. *Optimal Data Analysis*, *1*, 41-42.

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

[17]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.

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

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

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

Author Notes

The study analyzed de-individuated data and was exempt from Institutional Review Board review. No conflict of interest was reported.

Mail: Optimal Data Analysis, LLC
6348 N. Milwaukee Ave., #163
Chicago, IL 60646
USA