Evaluating biological interaction (and effect measure modification) as departures from additivity: Choosing and using additive and multiplicative models

David M. Thompson
June 28, 2016

In defining “the connection between additivity and biological independence,” Rothman (2002, p. 178), asks:

“why is it that biological interaction should be evaluated as departures from additivity of effect?”

Rothman’s discussion of biological interaction implies a logical sequence wherein an investigator (1) hypothesizes a mechanism, which may involve biological interaction; (2) selects an effect measure (measure of association) that appropriately quantifies the hypothesized effect; and finally (3) selects and constructs a statistical model that is appropriate for the chosen effect measure.

Rothman establishes links

(1) between “biological independence” and additivity of effects, and
(2) between “biological interaction” and a departure from additivity of effects.

The first part of this handout explains how one assesses additivity of effects using risk differences as the measure of association. Part III illustrates an alternative formulation of an additive model that uses risk ratios or odds ratios as the measure of association. Interleaved with these presentations are demonstrations that use SAS PROC GENMOD and a single data example to show how interaction is assessed in an additive statistical models (in Part II) and in a multiplicative model (in Part IV). The data example features a dichotomous outcome, and two dichotomous independent or predictor variables.

Consider a comparison of the “risk” or prevalence of an outcome Y, among individuals who are exposed or unexposed to one or both of two “risk factors,” X and Z.

<table>
<thead>
<tr>
<th>Z=1 (“exposed to Z”)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome Y=1</td>
<td>a₁</td>
</tr>
<tr>
<td>Outcome Y=0</td>
<td>b₁</td>
</tr>
<tr>
<td>Z=0 (“not exposed to Z”)</td>
<td></td>
</tr>
<tr>
<td>Outcome Y=1</td>
<td>a₀</td>
</tr>
<tr>
<td>Outcome Y=0</td>
<td>b₀</td>
</tr>
<tr>
<td>X=1 (“exposed to X”)</td>
<td></td>
</tr>
<tr>
<td>c₁</td>
<td>d₁</td>
</tr>
<tr>
<td>X=0 (“not exposed to X”)</td>
<td></td>
</tr>
<tr>
<td>c₀</td>
<td>d₀</td>
</tr>
</tbody>
</table>

Let pₓz be a probability with subscripts that signify the risk or prevalence of the outcome Y at levels of X and Z.

where z=1

Risk (or incidence or prevalence) among those exposed to X = p₁₁ = a₁/(a₁+b₁)
Risk for those unexposed to X = p₀₁ = c₁/(c₁+d₁)

and where z=0

Risk (or incidence, or prevalence) among those exposed to X = p₁₀ = a₀/(a₀+b₀)
Risk among those unexposed to X = p₀₀ = c₀/(c₀+d₀)

Displayed in a revised table, these probabilities are:

<table>
<thead>
<tr>
<th>X=1 (“exposed to X”)</th>
<th>Z=1 (“exposed to Z”)</th>
<th>Z=0 (“not exposed to Z”)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X=1 (“exposed to X”)</td>
<td>p₁₁</td>
<td>p₁₀</td>
</tr>
<tr>
<td>X=0 (“not exposed to X”)</td>
<td>p₀₁</td>
<td>p₀₀</td>
</tr>
</tbody>
</table>
Additivity involves the equality of joint and independent effects

Rothman (2002, p.178) states that the following relationship “establishes additivity as the definition of biological independence.”

\[ p11 - p00 = (p10 - p00) + (p01 - p00) \] (1)

When two exposures (X and Z) are independent, the effect on Y of their joint and simultaneous effects \((p11 - p00)\) is equal to the sum of the separate and independent effects of X \((p10 - p00)\) and of Z \((p01 - p00)\).

A departure from additivity of effect, which Rothman considers evidence of “biological interaction,” is present if the joint and simultaneous effect of the two exposures differs from the sum of the effects of each exposure when considered separately. Szklo (2004, p. 186) similarly states that “interaction occurs when the observed joint effect of [X and Z] differs from that expected on the basis of their independent effects.”
Additivity involves homogeneity of effects

Returning to equation (1) which, according to Rothman (2002, p. 178), “establishes additivity as the definition of biological independence,”

\[ p_{11} - p_{00} = (p_{10} - p_{00}) + (p_{01} - p_{00}) \]  \hspace{1cm} (1)

and rearranging equation (1)’s terms

\[ p_{11} - p_{00} - (p_{01} - p_{00}) = (p_{10} - p_{00}) \]
\[ p_{11} - p_{00} - p_{01} + p_{00} = p_{10} - p_{00} \]
\[ p_{11} - p_{01} = p_{10} - p_{00} \]  \hspace{1cm} (2)

yields a definition of additivity that involves a homogeneity of effects. Equation (2) states that the effect of X on Y is the same whether Z=1 \((p_{11} - p_{01})\) or Z=0 \((p_{10} - p_{00})\). Additivity (the absence of interaction) implies that measures of association between Y and X are homogenous (do not differ) at levels of Z.

Homogeneity of effects is reciprocal. We can rearrange the four probabilities in equation (2) and express them as:

\[ p_{11} - p_{10} = p_{01} - p_{00} \]  \hspace{1cm} (3)

Equation (3) states that the effect of Z on Y is the same whether X=1 \((p_{11} - p_{10})\) or X=0 \((p_{01} - p_{00})\). When effects of X and Z are additive: the association between Y and X is homogenous at levels of Z, and the association between Z and Y is homogenous at levels of X.

Vanderweele and Knol (2014, page 35, equation (2)) state that when \(p_{11} - p_{10} > p_{01} - p_{00}\), “interaction is sometimes said to positive or ‘super-additive’” and, similarly, when \(p_{11} - p_{10} < p_{01} - p_{00}\), “the interaction is said to be negative or ‘sub-additive’.”

Szklo and Nieto (2004, p. 186) summarize this “definition based on homogeneity or heterogeneity of effects” by stating that “interaction occurs when the effect of a risk factor A [X in our example] on the risk of an outcome Y is not homogeneous in strata formed by a third variable Z.” They comment that “when this definition is used, variable Z is often referred to as an effect modifier.”
Note that any of the three equations (1, 2 or 3) can be rearranged to arrive at any of the others. Assessing homogeneity of effects, and assessing the equality of joint and independent effects, are algebraically equivalent strategies for describing additivity.

Returning to equation (1):

\[
p_{11} - p_{00} = (p_{10} - p_{00}) + (p_{01} - p_{00})
\]  

(1)

we can construct a contrast among its proportions that tests the null hypothesis that the effects of X and Z are additive. This is equivalent to the hypothesis that no interaction exists between X and Z:

\[
p_{11} - p_{00} - (p_{01} - p_{00}) - (p_{10} - p_{00}) = 0
\]

\[
p_{11} - p_{10} - p_{01} + p_{00} = 0
\]

(1)*p_{11} + (-1)*p_{10} + (-1)*p_{01} + (1)*p_{00} =0

Rothman calls this expression the “interaction contrast” (IC). We can construct an appropriate statistical model to estimate the quantity on the left side of the equation, calculate its 95% confidence interval, and judge whether the quantity truly differs from zero.

We can estimate the interaction contrast in SAS PROC GENMOD by inserting the equation’s four coefficients (1 -1 -1 1) into an ESTIMATE statement. An important caveat is that the procedure must specify a statistical model that is additive. Part II of this handout explains and demonstrates the SAS syntax below, which specifies an additive model and directly estimates the interaction contrast (IC).

```sas
proc genmod data=two descending;
  freq count;
  class x (ref=first) y (ref=first);
  model y=x z x*z / link=identity dist=bin type3 lrci;
  lsmeans x*z / cl;
  ods output lsmeans=lsmeans;
  estimate "Interaction contrast" x*z 1 -1 -1 1;
run;
```
Data example – lung cancer incidence among workers with different exposures to asbestos and smoking

A widely studied dataset (Hammond, Selikoff, and Seidman, 1979) compared the risk of dying from lung cancer among 17,800 asbestos workers in the US and also among 73,763 men who were not exposed to asbestos. The study also enumerated smoking status, so members of the cohort displayed different combinations of exposure to cigarette smoking and to asbestos.

The SAS program below creates a dataset that approximates the published risk proportions, which are reported as lung cancer deaths per 100,000. To obtain a dataset whose properties parallel those of the published one, a smoking prevalence of about 0.28 was assigned to both the asbestos workers and to the comparison group.

```
data two;
input asbestos smk lungcadeath count;
cards;
0 0 0 53103
0 0 1 6
0 1 0 20628
0 1 1 25
1 0 0 12809
1 0 1 7
1 1 0 4954
1 1 1 30;
/*summary statistics*/
proc freq data=two;
weight count;
tables asbestos*smk*lungcadeath / nocol nopct outpct out=three;
run;
data four;
set three;
perhunthou=pct_row*1000;
run;
proc format;
  value smkf 1="Smokers" 0="Non-smokers";
  value gpf 1="Asbestos Workers (n= 17800)"
  0="Comparison Group (n=73763)";
run;
/*two by two table like that reported in Gordis and other sources.
Deaths from lung cancer (per 100,000) among individuals
with and without exposures to cigarette smoking and
to working with asbestos*/
proc report nowd data=four;
  where lungcadeath=1;
columns smk asbestos, perhunthou;
define smk / group "Cigarette smoking" format=smkf. order=internal;
define asbestos / across "Asbestos Exposure" format=gpf. order=internal;
define perhunthou / analysis '' format=5.1;
run;
```
The PROC REPORT step produces, with some editing, this table:

<table>
<thead>
<tr>
<th>Cigarette smoking</th>
<th>Asbestos Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Comparison Group</td>
</tr>
<tr>
<td></td>
<td>(n=73763)</td>
</tr>
<tr>
<td></td>
<td>Asbestos Workers</td>
</tr>
<tr>
<td></td>
<td>(n=17800)</td>
</tr>
<tr>
<td>Non-smokers</td>
<td>p00 = 11.3</td>
</tr>
<tr>
<td>Smokers</td>
<td>p10 = 121.0</td>
</tr>
</tbody>
</table>

Following the notation introduced in Part I, where X denotes cigarette smoking (1=smokers and 0=nonsmokers) and z denotes the strata for asbestos exposure, the risk probabilities $p_{xz}$ are:

- $p00 = 11.3$ deaths per 100,000 in those with neither exposure
- $p01 - p00 = 54.6 - 11.3 = 43.3$ deaths per 100,000. These represent excess deaths attributable to asbestos exposure.
- $p10 - p00 = 121.0 - 11.3 = 109.7$ deaths per 100,000. These represent excess deaths attributable to smoking.
- $p11 - p00 = 601.9 - 11.3 = 590.6$ excess deaths per 100,000 among those with joint exposure.

**The data example illustrates non-additivity of effects**

In the absence of interaction, we expect that the effect of experiencing both exposures would equal the sum of their separate effects.

$$p11 - p00 = (p01 - p00) + (p10 - p00)$$

$$590.6 \neq (43.3 + 109.7)$$

$$590.6 - (43.3 + 109.7) = 437.6$$

The risk of lung cancer death where both exposures are present is greater (by about 437.6 deaths per 100,000) than the sum of the risks from the two individual exposures. This inequality may be evidence of a synergy or positive interaction between the exposures.
Part II – Assessing departure from additivity in an additive model that uses risk differences as its effect measure

Recall that in part I, we expressed additivity in terms of independence,

\[ p_{11} - p_{00} = (p_{10} - p_{00}) + (p_{01} - p_{00}) \]  

(1)

and then derived Rothman’s “interaction contrast,” which amounts to a test of additivity:

\[ p_{11} - p_{00} - (p_{01} - p_{00}) - (p_{10} - p_{00}) = 0 \]

\[ p_{11} - p_{10} - p_{01} + p_{00} = 0 \]

H0: \((1)p_{11} + (-1)p_{10} + (-1)p_{01} + (1)p_{00} = 0\)

Statistical models exist that can test for additivity using the risk difference as an effect measure (or measure of association). These include the “binomial model for the risk difference” (Spiegelman and Hertzmark, 2005), also called the “binomial regression model” (Cheung 2007).

The following SAS PROC GENMOD step constructs an additive model that considers an outcome that follows a binomial distribution (dist=bin in the MODEL statement) and links that outcome to risk differences (link=identify in the MODEL statement). The ESTIMATE statement directly tests the null hypothesis of additivity. Evidence against this null hypothesis suggests “non-additivity.”

Note that the MODEL statement uses an identity link to create an additive model. The ESTIMATE statement lists four coefficients that reproduce exactly the ones in the expression for the interaction contrast.

```sas
proc genmod data=two descending;
  freq count;
  class smk (ref=first) asbestos (ref=first) ;
  model lungcadcemortality = smk asbestos smk*asbestos / link=identity dist=bin type3 lrci;
  lsmeans smk*asbestos / cl;
  ods output lsmeans=lsmeans estimates=estimates;
  estimate "IC" smk*asbestos 1 -1 -1 1;
run;
```

The next steps produce a report and a graph that show how the joint effects of cigarette smoking and exposure to asbestos are non-additive with respect to an outcome, lung cancer mortality (predicted deaths per 100,000). More specifically, the effects are “super-additive” or synergistic.
data mortality;
set lsmeans;
mortality=estimate*100000;
ucl=upper*100000;
lcl=lower*100000;
run;
proc print noobs data=mortality;
  var smk asbestos estimate mortality lcl ucl;
run;

<table>
<thead>
<tr>
<th>asbestos</th>
<th>smk</th>
<th>Estimate</th>
<th>Deaths per 100,000</th>
<th>95% CI on estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.006019</td>
<td>601.926</td>
<td>387.183 816.669</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.000546</td>
<td>54.619</td>
<td>14.169 95.070</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0.001210</td>
<td>121.048</td>
<td>73.627 168.469</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.000113</td>
<td>11.298</td>
<td>2.258 20.337</td>
</tr>
</tbody>
</table>

proc print noobs data=estimates;
  format meanestimate meanlowercl meanuppercl 12.8;
run;

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Estimate</th>
<th>Lower CL</th>
<th>Upper CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC</td>
<td>0.00437557</td>
<td>0.00213768</td>
<td>0.00661345</td>
</tr>
</tbody>
</table>

The estimated contrast (0.004376 or 437.6 per 100,000) matches our earlier calculation. The confidence interval on the estimate, and its associated p value, which is less than 0.0001, indicate that the true contrast is not equal to zero. We reject the null hypothesis that the effects are additive; the model detects evidence of a departure from additivity.

A plot of the model’s predictions illustrates the departure from additivity.

proc sgplot data=mortality;
  series y=mortality x=smk / group=asbestos name="one"
    groupdisplay=cluster clusterwidth=0.05
    markers markerattrs=(symbol=squarefilled size=10);
  highlow x=smk high=ucl low=lcl / group=asbestos
    groupdisplay=cluster clusterwidth=0.05;
  xaxis values=(0 1) label="" valueattrs=(size=14 weight=bold);
  yaxis label="Lung cancer deaths per 100,000"
    labelattrs=(size=14 weight=bold)
    valueattrs=(size=14 weight=bold);
  format smk smkf. asbestos gpf.;
  keylegend "one" / title="" location=inside down=2 position=topleft
    valueattrs=(size=12 weight=bold);
run;
Part III. Using ratios, instead of risk differences, to assess departure from an additive model.

Additive models that explore risk differences must have a suitable estimate for $p_{00}$, which is the risk, incidence or prevalence of the outcome among those who are unexposed to either of the risk factors $X$ or $Z$. However, an estimate of $p_{00}$ may not be available, for example when observational data are obtained in a case control design. This limitation motivates the use of an alternative measure of association to assess departures from additivity.

Recall that we defined additivity on the basis of risk differences:
\[
p_{11} - p_{00} = (p_{10} - p_{00}) + (p_{01} - p_{00})
\]
(1)

and then manipulated equation (1) to arrive at Rothman’s “interaction contrast:
\[
p_{11} - p_{10} - p_{01} + p_{00} = 0
\]

Dividing every term by $p_{00}$ yields.
\[
\frac{p_{11}}{p_{00}} - \frac{p_{01}}{p_{00}} - \frac{p_{10}}{p_{00}} + 1 = 0
\]

Recognizing that these ratios of probabilities are relative risks (RR), we obtain:
\[
RR_{11} - RR_{01} - RR_{10} + 1 = 0
\]
(4)

Rothman (1986) refers to the quantity on the left side of this equation as the “Relative Excess Risk due to Interaction” (RERI). Rothman and Greenland also call it an “interaction contrast ratio” (ICR). An equivalent expression, obtained in a few algebraic steps, is:
\[
RR_{11} - 1 = (RR_{10} - 1) + (RR_{01} - 1)
\]
(5)

The left side of equation (5) reflects the relative risk of the outcome that is associated with the joint or simultaneous effects of both the exposure variable $X$ and the stratum variable $Z$. The right side reflects the sum of the individual effects of either variable by itself.

If the equality holds, then we consider the risk ratios to conform on an additive scale. In this case, we say that there is no interaction or effect measure modification on an additive scale or in an additive model.

However, evidence against the equality is evidence of interaction or departure of additivity between these effect measures.
Analogous equations apply to odds ratios when we use them as effect measures:

\[
\begin{align*}
OR_{11} - OR_{01} - OR_{10} + 1 &= 0 \\
OR_{11} - 1 &= (OR_{10} - 1) + (OR_{01} - 1)
\end{align*}
\]

Hosmer and Lemeshow (1992) show how to obtain estimates (and confidence intervals) for the RERI using logistic regression. For reasons founded in statistical theory (logistic regression uses the logit link, which is the canonical link for outcomes that follow a binomial distribution), logistic regression is a reliable choice for multivariable models that assess a dichotomous outcome.

As an alternative to logistic regression, Cheung (2007) suggests using a generalized linear model that retains the identity link described earlier, but that treats the outcome as a Poisson random variable. Because this approach necessarily mis-specifies the variance, Cheung suggests modifying the approach so that it calculates standard errors that are robust in spite of this mis-specification.
Part IV. Assessing departure from multiplicativity in multiplicative models that use relative risks or odds ratios as their effect measures

Parts I, II, and III of this handout explore additivity of effects. Multiplicativity of effects can be defined in an analogous manner. If two exposures’ (X and Z) joint effect on an outcome (Y) is equal to the product of their separate and independent effects, we can predict that:

\[
p(\text{y}=1 \mid x=1, z=1) = \frac{p(\text{y}=1 \mid x=1, z=0) \cdot p(\text{y}=1 \mid z=1, x=0)}{p(\text{y}=1 \mid x=0, z=0) \cdot p(\text{y}=1 \mid z=0, x=0)}
\]

\[
\text{RR}_{xz} = \text{RR}_x \cdot \text{RR}_z
\]

If we compare the joint and independent effects of two exposures on the odds of an outcome:

\[
\text{OR}_{xz} = \text{OR}_x \cdot \text{OR}_z
\]

These equations amount to tests of the null hypothesis that there is no departure from multiplicativity. Evidence against these equalities suggest a departure from multiplicativity.

Log binomial models, which estimate relative risks as the measure of association, can test the first null hypothesis. Logistic regression models, which estimate odds ratios, can test the second. In both statistical models, the inclusion of product terms provides a direct test of the null hypothesis that there is no departure from multiplicativity.

In the log binomial model:

\[
\ln[p(\text{y}=1)] = \beta_0 + \beta_1 \cdot X + \beta_2 \cdot Z + \beta_3 \cdot X \cdot Z
\]

\[
p(\text{y}=1) = \exp(\beta_0 + \beta_1 \cdot X + \beta_2 \cdot Z + \beta_3 \cdot X \cdot Z)
\]

The model yields expressions for relative risks or, depending on the sampling scheme, a prevalence proportion ratio. Either of these is a ratio of two exponentiated linear functions:

\[
\text{RR}_{xz} = \exp(\beta_0 + \beta_1 \cdot X + \beta_2 \cdot Z + \beta_3 \cdot X \cdot Z) / \exp(\beta_0) = \exp(\beta_1 \cdot X + \beta_2 \cdot Z + \beta_3 \cdot X \cdot Z)
\]

\[
\text{RR}_x = \exp(\beta_0 + \beta_1 \cdot X) / \exp(\beta_0) = \exp(\beta_1 \cdot X)
\]

\[
\text{RR}_z = \exp(\beta_0 + \beta_2 \cdot Z) / \exp(\beta_0) = \exp(\beta_2 \cdot Z)
\]

\[
\text{RR}_x \cdot \text{RR}_z = \exp(\beta_1 \cdot X) \cdot \exp(\beta_2 \cdot Z) = \exp(\beta_1 \cdot X + \beta_2 \cdot Z)
\]
If there is no departure from multiplicativity among relative risks:

\[ RR_{xz} = RR_x \times RR_z \]

\[ \exp(\beta_1 X + \beta_2 Z + \beta_3 XZ) = \exp(\beta_1 X + \beta_2 Z) \]

Note that the equality holds if \( \beta_3 = 0 \), where \( \beta_3 \) is the regression coefficient associated with the product term. The log binomial model produces a test of the hypothesis \( H_0: \beta_3 = 0 \). Evidence that \( \beta_3 \neq 0 \) leads us to reject \( H_0 \) and suspect a departure from multiplicativity among the relative risks.

Logistic regression models apply similar logic and similar algebra to assess multiplicativity among odds ratios:

Recall that the log binomial model is of the form:

\[ \ln[p(y = 1)] = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ \]

so that

\[ p(y = 1) = \exp(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ) = e^{\beta_0} e^{\beta_1 X} e^{\beta_2 Z} e^{\beta_3 XZ} \]

The logistic regression model is of the form:

\[ \ln \frac{p(y = 1)}{p(y = 0)} = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ \]

so that odds are expressed:

\[ \frac{p(y = 1)}{p(y = 0)} = \exp(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ) = e^{\beta_0} e^{\beta_1 X} e^{\beta_2 Z} e^{\beta_3 XZ} \]

With respect to risks, probabilities and odds, logistic regression and log binomial models are inherently multiplicative. Their estimated coefficients (after exponentiation) relate to “\( e^{\beta} \)-fold difference” in the outcome for a “one-unit” difference in the predictor variable.

If there is no departure from multiplicativity among odds ratios, then:

\[ OR_{xz} = OR_x \times OR_z \]

\[ \frac{\exp(\beta_0 + \beta_1 \times 1 + \beta_2 \times 1 + \beta_3 \times 1 \times 1)}{\exp(\beta_0 + \beta_1 \times 0 + \beta_2 \times 0 + \beta_3 \times 0 \times 0)} = \frac{\exp(\beta_0 + \beta_1 \times 1 + \beta_2 \times 0 + \beta_3 \times 1 \times 0)}{\exp(\beta_0 + \beta_1 \times 0 + \beta_2 \times 0 + \beta_3 \times 0 \times 0)} \]

\[ \frac{\exp(\beta_0 + \beta_1 + \beta_2 + \beta_3)}{\exp(\beta_0)} = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} \times \frac{\exp(\beta_0 + \beta_2)}{\exp(\beta_0)} \]

\[ \exp(\beta_1 + \beta_2 + \beta_3) = \exp(\beta_1) \times \exp(\beta_2) \]

\[ \exp(\beta_1 + \beta_2 + \beta_3) = \exp(\beta_1 + \beta_2) \]
Note that the equality holds only if $\beta_3=0$, where $\beta_3$ is the regression coefficient associated with the product term. The logistic regression model produces a test of the hypothesis $H_0: \beta_3=0$. Evidence that $\beta_3\neq0$ leads us to reject $H_0$ and suspect a departure from multiplicativity among the odds ratios.

The worked SAS example below explores the same lung cancer mortality data (Hammond and Selikoff, 1979) that we examined earlier.

A log binomial finds no evidence of departure from multiplicativity.

/*The log binomial model is inherently multiplicative because it employs a log transformation of $Y$*/

```sas
proc genmod data=two descending;
  freq count;
  class smk (ref=first) asbestos (ref=first) ;
  model lungcadeneth = smk asbestos smk*asbestos
   / link=log dist=bin type3 ;
  lsmeans smk*asbestos / cl;
  ods output lsmeans=lsmeans ;
run;
```

```sas
data mortality;
  set lsmeans;
  mortality=exp(estimate)*100000;
  ucl=exp(upper)*100000;
  lcl=exp(lower)*100000;
run;
```

This model’s mortality estimates are identical to those derived from the additive model. The confidence limits differ some, because the models assume different variance structures.

<table>
<thead>
<tr>
<th>asbestos</th>
<th>smk</th>
<th>Estimate</th>
<th>Deaths per 100,000</th>
<th>lcl</th>
<th>ucl</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>-9.0883</td>
<td>11.298</td>
<td>5.076</td>
<td>25.146</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>-6.7167</td>
<td>121.048</td>
<td>81.812</td>
<td>179.099</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>-7.5125</td>
<td>54.619</td>
<td>26.044</td>
<td>114.546</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-5.1128</td>
<td>601.926</td>
<td>421.312</td>
<td>859.968</td>
</tr>
</tbody>
</table>

The absence of statistical interaction in this multiplicative model indicates there is no departure from multiplicativity.
Indeed, a graph of the predicted log probabilities illustrates the lack of departure from multiplicativity on the measurement scale that the log binomial model employs.

```sas
proc sgplot data=lsmeans;
  series y=estimate x=smk / group=asbestos name="one"
       groupdisplay=cluster clusterwidth=0.05
       markers markerattrs=(symbol=squarefilled size=10);
  xaxis values=(0 1) label=" " valueattrs=(size=14 weight=bold); 
  yaxis label="Ln (p of Lung cancer)"
       labelattrs=(size=14 weight=bold)
       valueattrs=(size=14 weight=bold); 
  format smk smkf. asbestos gpf.;
  keylegend "one" / title="" location=inside down=2 position=topleft 
       valueattrs=(size=12 weight=bold) ;
run;
```
The statistical model that we applied earlier to these same data, which used risk differences as its effect measure, found evidence of interaction. A multiplicative model that uses risk ratios as its effect measure, finds no evidence of interaction. The lack of interaction in the multiplicative model is NOT evidence of a lack of biological interaction. It is, however, a reminder of the importance of using constructing a statistical model that uses an effect measure that is matched to the scientific question.

This comparison of statistical models underscores Rothman’s assertion that biological interaction should be assessed as a departure from an additive model. Multiplicative models are not well suited for evaluating biological interaction. Rothman concludes (2002, p. 180) that “statistical evaluation of interaction using [multiplicative] models will not give an appropriate assessment of biological interaction.”

References:


