PharmaSUG2011 - Paper SP03

Formulas Calculating Risk Estimates and Testing for Effect Modification and Confounding

By Manojkumar B Agravat

Tampa, Florida

Abstract

The author is demonstrating formulas for point estimates, effect modification algorithms, and confounding inferences that may have bias and inferences of causality with new formulas, that involves PROC LOGISTIC, PROC IML, and PROC MIXED (SAS TM) algorithms. The death penalty and a sample randomized clinical trial is given for bivariate and multivariate applications of the formulas. The Cochran Mantel Haenszel Test is compared with new Agravat's formulas using novel codes of "R" from the author for strength of association. Logistic regression and Cochran Mantel Haenszel test are compared for confounding. In addition, an algorithm for hyper-geometric distribution is shown for myocardial infarction that did not exist before for effect modification. A new probability formula, hazard ratio, and logits are shown to explain head neck cancer (INHANCE) for one exposure no drinking, along with Agravat series which is hyper-geometric. A new algorithm and distribution is introduced to deal with the probability of events for discussion called Agravat's distribution and probability algorithm. Agravat series is displayed and involves features for special relativity of Einstein, Poincaire` conjecture which is given graphically, Bose's work, and Dirac's comments are discussed.

Keywords: logit, point estimates, effect modification, confounding, hyper-geometric, randomized clinical trials.

1 INTRODUCTION

The problem of point estimates when controlling for both confounding and explanatory variables is not possible for some statistics with the standard method (Cochran Mantel Haenszel Test) for data with conflicting risks. The risk may be greater or lesser than the risk whose variable is in question compared to the confounding statistics alone. The author proposes a method that calculates a risk statistic which can be utilized to decide if there are significant risks important for confounding and effect modification based on significant P-values or significant difference that has causal inference potential. For the effect modification method, a data transformation step is done to modify the count data and making a new variable "fit", "zxy", and "xzy" whose beta estimates are obtained from the use of R and a survival function survreg. Y represents the outcome variable. X represents the explanatory variable. Z represents the confounding variable. In R YX may be described as the interaction between Y and X; likewise ZY means the interaction between Z and Y. The author's new equations are for bivariate explanatory covariates and adjust for interaction. The data on death penalty is very biased according to results; hence lognormal distribution is used with the scale values for variables. Normally, the beta estimates used are directly from the output of results. As in the paper "A New Effect Modification P Value Test Demonstrated" [Agravat 2009] [Agravat 2008] the survreg function is used for beta survival estimates which is parametric because distributions may be chosen. The expectation is to utilize a distribution with the better P-value for scientific purpose of the study data. The traditional method uses PROC FREQ Cochran Mantel Haenszel test for the point estimates or PROC LOGISTIC (SAS TM) for odds ratios and Breslow Day test. PROC IML will give avenues to calculate effect modification with a program of the author that is shown to work. Subsequently, the "aem" statistic is shown to follow an asymptotic chi-square and F-statistic distribution using PROC MIXED (SAS TM) that is multivariate. The odds ratio equation is shown with the proof of the Cochran Mantel Haenszel Test and assumption. A new algorithm and distribution is shown to exist for predicting probabilities related to quantum statistics.

2 METHODS

Method for Point Estimates The author will present a new method for calculating point estimates for odds ratio's and relative risks 1 and 2 when controlling for the confounder and explanatory variable. The object is to use new type of formulas similar to logits which incorporates vector math and complex numbers and geometrical coordinate system. The formulas use R and the survreg function which ordinarily uses time for censored events to obtain betas. You use the lognormal distribution when using survreg in R because the data is requiring this distribution for cases where there may be some extreme bias or not found not shown in analysis in studies using other methods to avoid large beta estimates and avoidable error in the case of death penalty. The new method incorporates the concept of calculating area under the curve of vectors. The formulas allow the use of complex numbers that result in more area being calculating for the point estimates. The author will present formulas and outputs for bivariate covariates in this article. The datasets include death penalty from Florida Law Review by authors Radelet and Pierce from prisoners who have committed multiple murders in Florida [Radelet and Pierce]. The lognormal distribution is needed for this data because of extreme bias and log scale is used for beta estimates are required due to likely extreme bias. The risk estimates are greater if adjusted for over-dispersion.

Table 1: Death Penalty Verdict By Defendant's Race/Victim's Race Regarding Multiple Murders in Florida

Defendant's Race	Victim's Race	Death Penalty(Yes)	Death Penalty(No)	Percent (Yes)
White	White	53	414	11.3
	Black	0	16	0.0
Black	White	11	37	22.9
	Black	4	139	2.8

2.1 Formulas for Point Estimates for Bivariate Levels and Non-normal Data in "R"

```
deathpenaltyb<-c(0,1,0,1,0,1,0,1)
defrb<-c(1,0,1,0,1,0,1,0)
vicrb<-c(1,0,1,0,1,0,1,0)
countb<-c(16,139,1,4,37,11,53,414)


penaltyb<-data.frame(deathpenaltyb,defrb,vicrb,countb)

penaltyb

bz<-.61
bzx<-.626
byz<-.61
by<-.442
bxz<-.61
sey<-.354
sez<-.344
sezx<-.344
byx<-.923
bx<-.626
seyz<-.344
sez<-.344
stderrz<-.523
n<-8


OR1<-(1-exp(bz+bxz))/(exp(-(bz)))*sqrt(sez)
```

exp{OR1}

RR1n<--(exp((bz+byx))/(1-exp(bz+byx)))*sqrt(stderrz)

RR1nb<--(exp(bz+byx+by)/(1-exp(bz+byx+by)))*sqrt(stderrz)

(RR1n+RR1nb)/2

exp{RR1}

RR2<-(exp(by+bxz)/(exp(by)))

exp{RR2}

Controlling for Explanatory Variable

OR2<-(1-exp(by+byz))/-(exp(-(by)))*sqrt(sez)

exp{OR2}

RR1<-(exp(-by)/(1-exp(by)))*sqrt(sey)

exp{RR1}

RR2<-(exp(by+byz)/(1-exp(by)))

exp{RR2}

The formulas here are from the author and based on the new use of the exponentiation of natural logarithm to calculate the area under the curve based on vectors that may be added, subtracted, or multiplied in an equation. The odds ratios here reflect that there is no interaction based on residuals (figure 2) with parallel residuals. These new equations for bivarate variables are similar to logits to reflect risks in 3-D of the X, Y, and Z plane.

2.2 Agravat's Novel "R" Codes and Outputs for survreg

legalyl<-survreg(Surv(count,vicr)~deathpenalty, data=penaltyb, dist="lognormal")

summary(legalyl)

Call:

survreg(formula = Surv(count, vicr) ~ deathpenalty, data = penalty,

dist = "lognormal")

Value Std. Error

(Intercept) 3.686 0.935 3.94324 8.04e-05

deathpenalty 11.223 5486.343 0.00205 9.98e-01 z p

Log(scale) 0.626 0.354 1.76981 7.68e-02

Scale= 1.87

Log Normal distribution

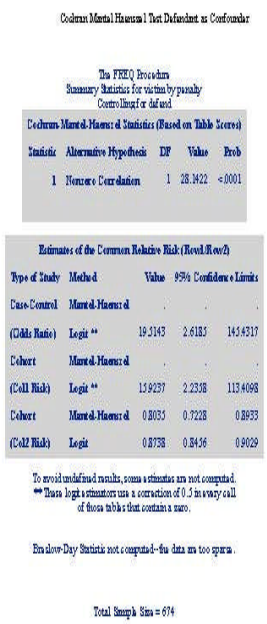
Loglik(model)= -22.9 Loglik(intercept only)= -24.4

Chisq= 2.9 on 1 degrees of freedom, p= 0.089

Number of Newton-Raphson Iterations: 17

n= 8

Figure 1: PROC FREQ (SAS) CMH Test Outputs Defendant / Victim



Cochran Mantel Haenszel Test Defendant as Confounder

The FREQ Procedure
Summary Statistics for victim by penalty
Controlling for defend

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	28.3422	<.0001

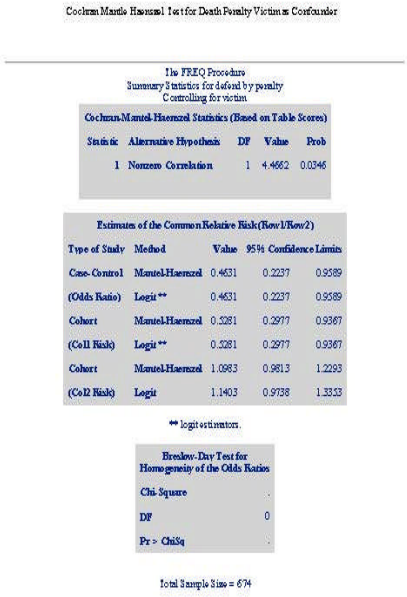
Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Limits	
Case-Control	Mantel-Haenszel	.	.	.
(Odds Ratio)	Logit **	19.5143	2.6183	145.4317
Cohort	Mantel-Haenszel	.	.	.
(Col1 Risk)	Logit **	1.9237	2.2338	1.134098
Cohort	Mantel-Haenszel	0.8035	0.7228	0.8933
(Col2 Risk)	Logit	0.8738	0.8456	0.9029

To avoid undefined results, some estimates are not computed.
** These logit estimators use a correction of 0.5 in every cell of those tables that contain a zero.

Breslow-Day Statistic not computed--the data are too sparse.

Total Sample Size = 674



Cochran Mantle Haenszel Test for Death Penalty Victim as Confounder

The FREQ Procedure
Summary Statistics for defend by penalty
Controlling for victim

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	4.4662	0.0346

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Limits	
Case-Control	Mantel-Haenszel	0.4631	0.2237	0.9589
(Odds Ratio)	Logit **	0.4631	0.2237	0.9589
Cohort	Mantel-Haenszel	0.5281	0.2977	0.9367
(Col1 Risk)	Logit **	0.5281	0.2977	0.9367
Cohort	Mantel-Haenszel	1.0983	0.9813	1.2293
(Col2 Risk)	Logit	1.1403	0.9738	1.3353

** logit estimators.

Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square	.
DF	0
Pr > ChiSq	.

Total Sample Size = 674

The codes are using a survival analysis function, survreg from “R” that is novel, and a new use in this application. Since this is a survival function, the full effect of risks due to censored events may be incorporated. The count is placed first and then the bivariate exposure next (Agravat 2008, Agravat 2009). P-values are used to pick the distribution used unless a distribution is already chosen by the study designer.

2.3 Randomized Drug Trial (SAS.com/PROC GENMOD)

```
data drug;
    input drug$ x r n @@;
    datalines;
  A  .1   1  10   A  .23  2  12   A  .67  1   9
  B  .2   3  13   B  .3   4  15   B  .45  5  16
  C  .04  0  10   C  .15  0  11   C  .56  1  12
  D  .34  5  10   D  .6   5   9   D  .7   8  10
  E  .2  12  20   E  .34 15  20   E  .56 13  15
  ;
```

2.3.1 Agravat’s R Code for Randomized Clinical Trial Sample Multivariate Data

```
drug<-c(1,2,3,4,5,1,2,3,4,5,1,2,3,4,5)

x<-c(1,1,1,0,1,1,0,1,0,0,0,0,0,0,0)

r<-c(1,1,1,0,0,1,1,0,0,0,1,0,0,0,0)

count<-c(10,13,10,10,20,12,15,11,9,20,9,16,12,10,15)

ex29_1a<-data.frame(drug,x,r,count)

ex29_1a



ex29b<-survreg(Surv(count,x)~drug, data=ex29_1a, dist="weibull")

summary(ex29b)

ex29d<-survreg(Surv(count,x)~drug:r, data=ex29_1a, dist="weibull")

summary(ex29d)

ex29e<-survreg(Surv(count,x)~r, data=ex29_1a, dist="weibull")

summary(ex29e)
```

2.3.2 Output for Randomized Clinical Trial (Multivariate)

```
Call:

survreg(formula = Surv(count, x) ~ drug, data = ex29_1a, dist = "weibull")

             Value Std. Error     z        p

(Intercept)  2.407     0.1547 15.55 1.53e-54

drug         0.139     0.0552  2.52 1.17e-02

Log(scale)  -1.704     0.3157 -5.40 6.80e-08
```

Scale= 0.182

Weibull distribution

Loglik(model)= -19 Loglik(intercept only)= -21.8

Chisq= 5.74 on 1 degrees of freedom, p= 0.017

Number of Newton-Raphson Iterations: 6

n= 15

Call:

survreg(formula = Surv(count, x) ~ drug:r, data = ex29_1a, dist = "weibull")

	Value	Std. Error	z	p
(Intercept)	3.009	0.1275	23.60	4.12e-123
drug:r	-0.200	0.0774	-2.59	9.70e-03
Log(scale)	-1.587	0.3162	-5.02	5.24e-07

Scale= 0.205

Weibull distribution

Loglik(model)= -19.4 Loglik(intercept only)= -21.8

Chisq= 4.9 on 1 degrees of freedom, p= 0.027

Number of Newton-Raphson Iterations: 6

n= 15

Call:

survreg(formula = Surv(count, x) ~ r, data = ex29_1a, dist = "weibull")

	Value	Std. Error	z	p
(Intercept)	3.049	0.137	22.18	4.85e-109
r	-0.465	0.165	-2.81	4.96e-03
Log(scale)	-1.690	0.312	-5.41	6.32e-08

Scale= 0.185

Weibull distribution

Loglik(model)= -18.4 Loglik(intercept only)= -21.8

Chisq= 6.8 on 1 degrees of freedom, p= 0.0091

Number of Newton-Raphson Iterations: 6

n= 15

2.3.3 Results and New Formulas for Sample RCT:

“X” is coded 0 for concentrations above .2 for the drugs. R is 0 for “R” 5 and greater. R corresponds to

Randomization and N is the count variable.

```
> bz<-.139
> byz<--.200
> by<--.465
> bxz<--.0641
> bxy<--.567
> bx<--6.59
> sez<-.1547
> seyz<-.1275
> sey<-.137
> sexz<-.2140
> seyx<-.101
> sez<-.12
```

```
OR1<-(exp(bz+bxz+by)/(1-exp(bz+bxz+by)))*sqrt(sez)

RR1at<-(exp(-(bz))/(1-exp((bz+bxz+by))))*sqrt(sez)

RR1bt<-(-(exp(bz)))*(exp(-(bz)))*sqrt(sez)

RR1t<-(RR1at+RR1bt)*.5

RR2<-(exp(by+bxz)/(exp(by)))
```

Percent_attributable_risk_z<-(RR1-RR2)/RR1=-224%

Attributable_risk_z<-RR1-RR2=-65%

Efficacy_z<-(RR2-RR1)/RR2=69.1 %

Control for x

```
OR2<-((exp(bz+bxz+by))/(exp(bz)))

RR1a<-(exp(-(by))/(1-exp(by)))*sqrt(sey)

RR1b<-exp(-(by))*(-(exp(by)))*sqrt(sey)

RR1x<-(RR1a+RR1b)*.5
```

RR2a<-(exp(by+byz)/(1-exp(by)))

RR2b<-(exp(by+byz))*(exp(-(by+byz)))*sqrt(sexz)

RR2x<-(RR2a+RR2b)*.5

Percent_attributable_risk_x<-(RR1-RR2)/RR1=-126.2%

Attributable_risk_x<-RR1-RR2=-77%

Efficacy_x<-(RR2-RR1)/RR2=55.8%

The odds ratio when controlling for drugs or a possible interaction term for this sample study is .72 which means that the odds for outcome of a randomization due to no exposures of concentrations over .2 is .72 times based on drugs with statistically significant 95 % confidence intervals (.77,.67). The relative risk 1 is .29 which shows there is a 71 % protective effect due to the drugs being given for outcome that is also statistically significant (.31,.27). Relative risk 2 is not significant for not getting the outcome. Controlling for the explanatory variable is also showing statistically significant risk statistics for odds ratio, and relative risks as shown in table 2.

The efficacy when controlling for confounder or effect modifier is 69.1 %. When controlling for the explanatory variable, the efficacy is significant 69.1 % and the odds ratio is also significant .59 with 95 % confidence interval of .63 and .55. This may mean that the concentration has a significant effect, and that concentration above .2, relatively speaking with regards to the study and units, shows significant difference for the outcome with the randomization due to 95 % confidence interval and relative risk 1: .61, with 95 % CI of .62 and .56 there is a 39 % less risk of outcome due to exposure. Based on the side effects, this drug may or may be useful for the potential problem because both the presence and concentration are statistically significant. The drug itself may have an effect worth noting due to its protective effect of 71 % and 39 % due to its concentration showing consistency a causal criterion of Hill. Expect more requirements for inferences to be demonstrated further and supported in the future for this hypothetical randomized clinical trial to know what dose and side effects are shown and necessary for possible side effects.

Table 2: Randomized Clinical Trial Sample Agravat's Statistics Overall Risks for Multivariate Risk Estimates

Risk Statistic	Risk-z	95 % C. I.	Risk-x	95 % C. I.
Odds Ratio	.72	(.77, .67)	.59	(.63, .55)
Rel. Risk 1	.29	(.31, .27)	.61	(.62 .56)
Rel. Risk 2	.94	(1.01,.87)	1.38	(1.49,1.28)

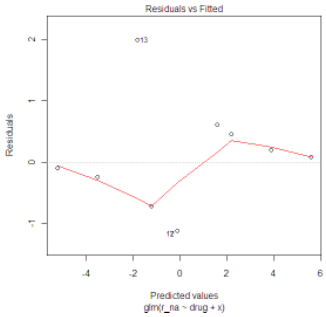


Figure 2: Residuals of Randomized Clinical Trial Plot

The plot of the residuals above using glm in "R" shows that the residuals are not interacting because the residuals are parallel. Interaction does not need to be taken into count here to calculate the risk statistics.

Risk	Defendant's Race Confounder	95 CI	Victim's Race	95 CI
Odds Ratio	.076	(.088, .066)	5.48	(6.54, 4.59)
Relative Risk 1	2.41	(2,88, 2.02)	.50	(.60 .42)
Relative Risk 2	6.30	(7.52,5.28)	.006	(.007, .005)
Risk	Victim's Race Confounder	95 CI	Defendant's Race	95 CI
Odds Ratio	.40	(.46, .35)	1.67	(1.99, 1.39)
Relative Risk 1	5.14	(6.14, 4.29)	.63	(.76.,.53)
Relative Risk 2	6.30	(7.53,5.27)	.06	(.07,.05)

Table 3: Outputs Using Agravat's Method in "R" and New Formulas
Black victims whose victim s were black had odds ratio of 5.48.

2.4 Proc Logistic and Death Penalty Outputs

The existing method for confounding and point estimates utilizes the Cochran Mantel Haenszel Method (CMH) and logistic regression. The point estimates come from PROC FREQ CMH option (SAS TM) CMH test which do not work well for point estimates when data has conflicting risks. The non-zero correlation P-value method is the test for independence for the CMH test giving P<.0001 and P <.034 when controlling for defendant and victim's race. Often the CMH test does not work for other distributions when conflicting risks exist such as in non-normal data. PROC LOGISTIC is often also used to check for confounding by determining if the beta estimates with and without confounder have a greater than 10 percent change but in this case.

Assumptions for Non-normal Distributions The assumptions of the death penalty dataset, shows that the Shapiro-Wilks P- value test, represents non-normal data with P <.0009. The Durbin Watson Statistic DW =2.28 indicating data is not auto correlated.

2.4 Results from Agravat's Method for Point Estimates for Death Penalty

The distribution chosen here is lognormal for this parametric method. Black was coded"1" (controlling for defendant's race) first in table 3. While whites were code as"0" and death penalty was coded as"1" for the first half of the analysis and reversed for defendant and victim's race in the second half. For this special situation and use of lognormal distribution and case of extreme bias, the odds are obtained by exponentiation of the output from the formulas for lognormal. The odds ratio for controlling for defendants race is .076 which means that for black defendants whose victims were black the odds of death penalty is .076 for victim's race being black [KCobb]. The 95 percent confidence interval is significant (.088,.066). hence there is evidence for bias due to significance and increased risks .076, especially since the odds ratio and relative risks are different .076 vs. 2.41 for defendant's race as confounder (see table 3) and .40 vs 5.14 when controlling for the other variable The risk for death penalty controlling for defendant's race is 2.41, meaning black defendants had a 2.41 times increased risk for death penalty vs. white defendants. The RR2 is 6.3 meaning the risk of not getting the death penalty for white defendants is 6.3. The RR2 is .006 for victims race controlled so the risk of not getting the death penalty if victim is black is 99.4% less.

There is evidence for interaction between defendant's race and victims' race for death penalty. The risk for not getting death penalty, when black vs. whites is .006 (R.R. 2) when controlling for victims race with defendant's race as confounder (see table 3). The new statistics show that the odds ratio when controlling for defendant's race and victims race, they differ by 81% (.076 vs .40) respectively and both statistically significant. The explanatory variable when controlled for victim's race and defendant's race is also showing a greater than 69.5 % (5.48 vs. 1.67) difference. This shows bias or is it double jeopardy based on race or another personal factor? (This new risk is not available with the PROC FREQ CMH option.) These risks were reported without adjustment for over-dispersion.

SAS Code 1a. Standard CMH SAS code for Death Penalty

```
data dpl;
input victim $ defend $ penalty $ count ;
cards;
 1 1 1  53
 1 0 1  11
 0 1 1   0
 1 1 1   4
 1 1 0  414
 1 0 0   37
 0 1 0   16
 0 0 0  139
;
```

```
Run;
Proc freq data=dpl;
weight count;
tables victim*defend*penalty/cmh;
Run;
```

SAS Code 1b. Death Penalty PROC LOGISTIC

```
data dpl;
input victim $ defend $ penalty $ count;
cards;
 1 1 1  53
 1 0 1  11
 0 1 1   0
 1 1 1   4
 1 1 0  414
 1 0 0   37
 0 1 0   16
 0 0 0  139
;
proc logistic data=dpl descending;
class defend victim penalty;
weight count;
model penalty=defend victim;
run;
proc logistic data=dpl descending;
class defend penalty;
weight count;
model penalty=defend ;
run;
proc logistic data=dpl descending;
class  victim penalty;
weight count;
model penalty= victim;
run;
```

PROC LOGISTIC (SAS) Outputs for Death Penalty

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-10.1285	384.0	0.0007	0.9790
defend	**0**	1	0.3849	0.1857	4.2973	0.0382
victim	**0**	1	-8.5306	384.0	0.0005	0.9823

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-2.3967	0.1706	197.2866	<.0001
defend	**0**	1	-0.3759	0.1706	4.8537	0.0276

Analysis of Maximum Likelihood Estimates						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-10.1245	389.0	0.0007	0.9792
victim	**0**	1	-8.2326	389.0	0.0004	0.9831

The new odds ratio when controlling for victim is .076 meaning that black victims whose defendants are black the odds for death penalty is .076 for black defendants (see table 3). The risk for death penalty when victim's race is black is 2.4 with black coded as 1 and 1.67 when black is coded as 0 for defendant. So when controlling for explanatory variable, victims who were white had a .40 odds ratio meaning that white defendants whose victims are white had a .40 chance of death penalty. White defendants whose victims were black had a 6.30 increased risk for not getting death penalty.

The Cochran Mantel Haenszel test did not work well for this death penalty dataset for point estimates when defendant was seen as confounder. The distribution assumption may not have been easily compatible with PROC FREQ and CMH test or there may have been conflicting risks (see figure 1).

Only the non-zero correlation showed that P <.0001 and P <.034 (see figure 1), hence there was independence or confounding was suspected when controlling for defendant's race and victim's race. The PROC LOGISTIC shows that for victim's race there is no confounding for victim's race because the beta does not change by 10 percent or more (3.6 %). The defendant's race shows there is confounding because the betas change over 10 percent. The signs are reversed for the betas and estimates nearly same. The standard errors are high though same when individual parameters are measured alone. Logistic regression requires equal variance which is not met for the full model because standard errors differ.

2.5 Hyper geometric Distribution Birth Control and Myocardial Infarction

Hyper geometric Distribution Code Agravat's method for hyper-geometric distribution works well for the typical groups of data that involves time and age. Hyper-geometric distribution is involved with the number of successes of drawing from a population without replacement. Often the studies involve the term total, referring to total number in sample size. One category is count and may involve the number in time or age stratum like this example of myocardial infarction with exposures including birth control. The grouping of data to analyze variables still involves outcome variables first include the frequency of that outcome. The outcome variable for this code with PROC MIXED involves bivariate outcomes for grouped data that can be set by the individual and used to test per selected category for effect modification with "lAEM" variable log of "aem". Fit variable is still used with the same procedure described for lung cancer data (Agravat 2009). There is also a fit variable called "fitrawz" where the raw count is kept for the effect modifier level group. There may be a variable for count of number of cases with exposure "-z" or effect modifier

variable and number with exposure including age or time strata. When the weight is 'n' the result is that there is a significant difference in risk for outcome of age category for exposure birth control use which includes the number in level of birth control adjusted for cases involving that category because P <.0001 (at alpha = .05) that is asymptotic chi- square with "1" for age being a categorical variable which is selected by the study designer is between 35-49 years old.In this type of study the log may be used as shown in the SAS code showing that "laem" or the logo of "aem" follows an asymptotic chi square distribution for large sample size Other methods do not exist or give results for hyper-geometric data (Zelterman: (samples/a57496)). The "AEM" variable is calculated by PROC IML and placed in each row for the data calculated based on the O statistic as shown in example 4a and 4b for the separate data in section 2.8 give P<0 which is statistically significant for chi square test value of 12.80 and one may reject the null of homogeneous odds for the null of no association.

The use of birth control and risk for myocardial infarction for older ages over 35 is supported by (womenshealth.gov). The Mayo clinic suggests that birth control use such as patches, vaginal rings, and pills may increase risk of high blood pressure (Mayo Clinic). Side effects increase for women who smoke and take birth control and these side effects are: heart failure, diabetes, stroke, clots, high blood pressure, liver tumors, gallstones and jaundice (ehow.com) as they are older. Women who are obese and have other risk factors for cardiovascular disease with birth control are more strongly advised not to smoke and take birth control (womenshealth.gov). For the age specific question, the criteria of total population versus age category variable, the significant results may show Agravat's method as having better causal inference potentials because the time variable can be included in terms of age categorized in bivariate terms. Cases of cardiovascular heart disease (CHD) have significant relationships with respect to using birth control for older ages, 35 and older, coded "1". Ages vary from 25-49 in 5 categories originally but coded "0" for ages less than 35. Further studies may be needed to test if birth control is conditionally a risk for CHD and if at all under what terms and which risk factors. The PROC IML code is shown to give the "AEM" values from the data transformation method of the author. AEM is calculated from the O statistic which is then put into the PROC IML code of the author, and then the AEM or the O statistic is placed into the SAS code for PROC MIXED. The observed minus observed mean determinants squared divided by observed mean is divided to calculate "aem" through the O statistic.

SAS CODE 2: Agravat's Method for Myocardial Infarction for Hyper-geometric Data

```
proc iml;
  * Read data into IML ;
  use micasesHo;
  read all ;

 * combine x0 x1 x2 into a matrix X ;
  var =  zxy || xzy||n;
 var2 = age || fit ;
  newvar=var*var2;


varC1={
 1  1  6,
 1  1  21};

varC2={
 598874  11363  37,
 898311  5601  71,
 1       1     99};

V1={.3  .3  1.5,
   .8  .8  .7};

O1={1  1  6,
    1  1  21};


deter1=(V1);

print deter1;
```

```
deterO1=(O1);

D1=(deterO1-deter1)#(deterO1-deter1)#1/deterO1;

print D1;
D1new={
  0.49  0.49  3.375
  0.04  0.04  0.845};

 ars=D1new[,+];

print ars;

print D1;

aem1=ars[+,];

print aem1;(5.28)


V2={
247541  15,
531279  42.2,
.01  .01};


O2={
610274  37,
894043  71,
101  99};

deter2=(V2);

print deter2;

deterO2=(O2);

D2=(deterO2-deter2)#(deterO2-deter2)#1/deterO2;

print D2;
D2new={

 215600.25  13.081081
 147193.95  11.682254
 100.98  98.980001};

 ars=D2new[,+];

print ars;

print D2;

aem2=ars[+,];

print aem2; (363024.2)

Ot=sum(aem1+aem2);

print Ot;

/***3130534
```

```
*/
P=PROBCHI(363042.98
, 5, 0)=.95 ;
print P; /****P<0  **/
```

```
print p;


lOt=log(Ot);
print lOt;

P=PROBCHI(12.80
, 5, 0)=.95 ;
print P; '(P<0)'
```

SAS CODE 2b: Hyper-geometric Code

```
data micasesHoaemnewD;
input n age fit fitrawz lcwocz locusex total laem ;
label
fitrawz = 'fit with the raw count for confounder z'
cwoc = 'cases with oc use'
lcwocz = log of 'cases with oc use adjuste for confounder'
locusex =log of '# in age stratum using oc''s adjusted for explanatory'
n = '# of cases in age stratum'
ltotal =log of 'sample size in this age stratum'
laem='log(aem)';
;
datalines;
6  0 1 2 1 1 292  .72
21 0 1 9 1 1 444  1
37 1 0 4 5.77 4.06 393  1
71 1 0 6 5.95 3.75 442   1
99 1 1 6 1 1 405 5.56
;
run;
proc mixed data=micasesHoaemnewD;
weight total; class  lcwocz  ;
model age=  lcwocz  laem /solution ddfm=satterth covb chisq ;
run;
```

Fit Statistics	
-2 Res Log Likelihood	-18E307
AIC (smaller is better)	-18E307
AICC (smaller is better)	-18E307
BIC (smaller is better)	-18E307

Solution for Fixed Effects						
Effect	log of cases with oc use adjuste for confounder	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		0.7864	0.03923	1	20.05	0.0317
lcwocz	1	-0.9754	0.04740	1	-20.58	0.0309
lcwocz	5.77	-132E-18	0.05509	1	-0.00	1.0000
lcwocz	5.95	0	.	.	.	.
laem		0.2136	0.01051	1	20.32	0.0313

Type 3 Tests of Fixed Effects						
Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
lcwocz	2	1	604.47	302.23	<.0001	0.0406
laem	1	1	412.74	412.74	<.0001	0.0313

SAS Code 2c with AEM and PROC MIXED for Myocardial Infarction

```
proc mixed data=micasesHoaemnewA;
weight n; class aem cwocz ;
model age=  cwocz  aem /solution ddfm=satterth covb chisq ;
run;
```

Fit Statistics	
-2 Res Log Likelihood	-18E307
AIC (smaller is better)	-18E307
AICC (smaller is better)	-18E307
BIC (smaller is better)	-18E307

Solution for Fixed Effects							
Effect	aem	cases with oc use adjuste for confounder	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			1.0000	0	1	Infty	<.0001
cwocz		1	1.39E-16	.	.	.	.
cwocz		598874	1.05E-16	.	.	.	.
cwocz		898311	0	.	.	.	.
aem	24.058333		-1.0000	0	1	-Infty	<.0001
aem	363018.92		0	.	.	.	.

Type 3 Tests of Fixed Effects						
Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
cwocz	2	1	0.00	0.00	1.0000	1.0000
aem	1	1	2.99E15	2.99E15	<.0001	<.0001

Fit Statistics	
-2 Res Log Likelihood	-18E307
AIC (smaller is better)	-18E307
AICC (smaller is better)	-18E307
BIC (smaller is better)	-18E307

2.6 New Method for Effect Modification of Head/Neck Cancer Data: Results of Head Neck Cancer Study INHANCE and the category of never drinkers vs. never smokers were previously too small to be valid statistically according to Mia Hashibe. Thus a larger category of this pool was added to study head neck cancer which is normally casually linked to cigarette smoking and drinking alcohol 75 percent [Mia Hashibe, et. al., 2007]. According to the check on assumptions for data distribution of head cancer from nondrinkers/nonsmokers (comprising 15.6 and 26.6 percent respectively of cases and controls for non-drinkers versus 10.5 and 37.9 percent of cases and controls for non-smokers from the pool from the study) and according to race coded: 1 for non-Hispanic, 2 for black, and 3 for Hispanic, the chi-square is P <.0001 meaning that a random variable is likely (zxy for race) [Mia Hashibe et. al., 2007]. The non-drinkers were mostly below the age of 40 years or older than 75 years and were female. Likewise for the non-smokers, subjects were more educated. Shapiro-Wilk's test is indicating that a non-normal distribution with P <.0003 from PROC UNIVARIATE (SAS) exists. The PROC AUTOREG (SAS) shows that the Durbin Watson Statistic is 3.23 indicating negative autocorrelation, hence the data is non-normal and involves random effects model. Heterogeneity is expected in this study, and regression may not be fixed or linear as a result. One is coded never drinkers versus 0 never smokers for xzy. Effect modification is expected to exist for the outcome head/neck cancer for the level of race from the exposure of never drinking/never smoking based on P <.0001 statistic. The risks for head/neck cancer for the three races (non-Hispanic, black, and Hispanic) vary by more than 10 percent for never drinkers vs. never smokers. The C statistic is .799 indicating a very good confidence for the results. The data is fairly large, over 11,500, and the Likelihood assumption, Score test, and Wald test are all valid for the global null so the results may be valid for a large population approximation. The algorithm for effect modification program converged as indicated. The P <.0001 hence there is reason to believe that there is statistical significance for interaction of the risk of head/ neck cancer from never drinking alcohol and never smoking according to races being non-Hispanic, black, and Hispanic to be different by more than 10 percent in this study of the International Head and Neck Cancer Epidemiology (INHANCE), a multinational study conducted by France, in Europe, United States, India, and around the world. The effect modification test shows that there is a difference for head/neck cancer by race and the exposure in question of no drinking/ no smoking (INHANCE) due to P<.0001. The C statistic is .788 which gives a good deal of confidence in the test.

SAS CODE 3 Head Neck Cancer INHANCE Study

```
data smtobAEM;
input cases fit zxy xzy count;
datalines;
1 1 1   1   795
0 1 1   1   2586
1 0 332  569  763
0 0 1913 3279 4397
1 1 1   1   111
0 1 1   1   233
1 0 27   46   62
0 0 104   178  238
1 1 1   1   40
0 1 1   1   152
1 0 20    34   45
```

```
0 0 74   127  170
;
run;
proc logistic data=smtobAEM descending;
freq count;
class cases fit;
model cases= xzy zxy /lackfit rsq ;
run;
```

c 0.856

Partition for the Hosmer and Lemeshow Test

		cases = 1		cases = 0	
Group	Total	Observed	Expected	Observed	Expected
1	4397	0	42.84	4397	4354.16
2	3917	946	975.58	2971	2941.42
3	1278	870	797.59	408	480.41

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
61.9428	1	<.0001

The R square for PROC MIXED AEM OUTPUT

R-Square	0.5070	**Max-rescaled R-Square**	0.8163

PROC LOGISTIC OUTPUT for Standard Count Data

R-Square	0.2536	**Max-rescaled R-Square**	0.4083

SAS Code 3b: PROC IML and PROC MIXED SAS CODES of AEM Chi-square for Effect Modification

```
proc iml;
  * Read data into IML ;
  use smtobAEM;
  read all ;
 * combine x0 x1 x2 into a matrix X ;
  var =  zxy || xzy||n;
 var2 = cases || fit ;
  newvar=var*var2;
  varB1={
1  1,
```

```
 0  1};
varB2={
 1  0,
 0  0};
varB3={
 1  1,
 0  1};
varB4={
 1  0,
 0  0};
varB5={
 1  1,
 0  1};
varB6={
 1  0,
 0  0};
varC1={
  1  1  795,
  1  1  2586};
varC2={
  332  569  763,
  1913  3279  4397};
varC3={
 1  1  111,
 1  1  233};
varC4={
  27  46  62,
  99  178  238};
varC5={
 1  1  40,
 1  1  152};
varC6={
  20  34  45,
  74  127  170};
varA1=varB1*varC1;
varA2=varB2*varC2;
varA3=varB3*varC3;
varA4=varB4*varC4;
varA5=varB5*varC5;
varA6=varB6*varC6;
print varA1;
print varA2;
print varA3;
print varA4;
print varA5;
print varA6;
V1={
 1.0  1.0  1916,
 0.4  0.4  1120};
O1={
 2  2  3385,
 1  1  2586};
deter1=(V1);
print deter1;
deterO1=(O1);
D1=(deterO1-deter1)#(deterO1-deter1)#1/deterO1;
print D1;
D1new={
    0.50  0.50  637.50,
    0.36  0.36  831.07};
 ars=D1new[,+];
print ars;
```

```
print D1;
aem1=ars[+,];
print aem1;
/******/
print varA2;
V2={
316  569  763,
0  0  0};
O2={
332  569  763,
0  0  0};
deter2=(V2);
print deter2;
deterO2=(O2);
D2=(deterO2-deter2)#(deterO2-deter2)#1/deterO2;
print D2;
D2new={
 0  0  0,
0  0  0};
 ars=D2new[,+];
print ars;
print D2;
aem2=ars[+,];
print aem2;
/**********/
print varA3;
V3={
1.2  1.2  205,
.4  .4  93.9};
O3={
2  2  344,
1  1  233};
deter3=(V3);
print deter3;
deterO3=(O3);
D3=(deterO3-deter3)#(deterO3-deter3)#1/deterO3;
D3new={
0.32  0.32  56.165698,
0.36  0.36  83.042103};
 ars=D3new[,+];
print ars;
print D3;
aem3=ars[+,];
print aem3;
/***************/
print varA4;
V4={
27  46  62,
0  0  0};
O4={
27  46  62,
0  0  0};
deter4=(V4);
print deter4;
deterO4=(O4);
D4=(deterO4-deter4)#(deterO4-deter4)#1/deterO4;
print D4;
D4new={
0  0  0,
0.  0.  0.};
 ars=D4new[,+];
print ars;
```

```
print D4;
aem4=ars[+,];
print aem4;
V5={
1.1  1.1  108.1,
.4  .4  65.5};
O5={
2  2  192,
1  1  150};
deter5=(V5);
print deter5;
deterO5=(O3);
D5=(deterO5-deter5)#(deterO5-deter5)#1/deterO5;
print D5;
D5new={
 0.41  0.405  161.7698,
 0.36  0.36  120.41309};
 ars=D5new[,+];
print ars;
print D3;
aem5=ars[+,];
print aem5;
/***************/
print varA6;
V6={
20  34  45,
0  0  0};
O6={
20  34  45,
0  0  0};
deter6=(V6);
print deter6;
deterO6=(O6);
D6=(deterO6-deter6)#(deterO6-deter6)#1/deterO6;
print D6;
D6new={
0  0  0,
0.  0.  0.};
 ars=D6new[,+];
print ars;
print D4;
aem6=ars[+,];
print aem6;
Ot=sum(aem1+aem2 +aem3+aem4+aem5+aem6);
print Ot;
P=PROBCHI(1894.5757, 12, 0)=.95 ;
print P;/******P<0 ******/
```

SAS Code 3c. AEM CODE with PROC MIXED

```
data smtobAEMmixed;
input cases fit zxy xzy aem count;
datalines;
1 1 1    1    1470.29    795
0 1 1    1       1       2586
1 0 332  569     1       763
0 0 1913 3279    1       4397
1 1 1    1    140.5678  111
0 1 1    1       1       233
1 0 27   46      1       62
```

```
0 0 104   178      1        238
1 1 1     1     283.71789   40
0 1 1     1        1        152
1 0 20    34       1        45
0 0 74    127      1        170
;
run;
```

```
proc mixed data=smtobAEMmixed;
weight count;
class  zxy ;
model cases= zxy aem /solution ddfm=satterth covb chisq ;
run;
```

Fit Statistics	
-2 Res Log Likelihood	20.6
AIC (smaller is better)	22.6
AICC (smaller is better)	24.6
BIC (smaller is better)	22.0

Solution for Fixed Effects

Effect	zxy	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-0.00067	0.07990	4	-0.01	0.9937
zxy	1	0.03779	0.1244	4	0.30	0.7764
zxy	20	1.0000	0.7938	4	1.26	0.2763
zxy	27	1.0000	0.6776	4	1.48	0.2140
zxy	74	4.48E-17	0.4141	4	0.00	1.0000
zxy	104	-429E-19	0.3526	4	-0.00	1.0000
zxy	332	1.0000	0.2078	4	4.81	0.0086
zxy	1913	0	.	.	.	.
aem		0.000668	0.000144	4	4.64	0.0097

Covariance Matrix for Fixed Effects

Row	Effect	zxy	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9
1	Intercept		0.006384	-0.00638	-0.00638	-0.00638	-0.00638	-0.00638	-0.00638		-2.07E-8
2	zxy	1	-0.00638	0.01547	0.006384	0.006384	0.006384	0.006384	0.006384		-6.31E-6
3	zxy	20	-0.00638	0.006384	0.6302	0.006384	0.006384	0.006384	0.006384		
4	zxy	27	-0.00638	0.006384	0.006384	0.4591	0.006384	0.006384	0.006384		
5	zxy	74	-0.00638	0.006384	0.006384	0.006384	0.1715	0.006384	0.006384		
6	zxy	104	-0.00638	0.006384	0.006384	0.006384	0.006384	0.1243	0.006384		
7	zxy	332	-0.00638	0.006384	0.006384	0.006384	0.006384	0.006384	0.04317		
8	zxy	1913									
9	aem		-2.07E-8	-6.31E-6							2.068E-8

Type 3 Tests of Fixed Effects						
Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
zxy	6	4	27.19	4.53	0.0001	0.0826
aem	1	4	21.56	21.56	<.0001	0.0097

The PROC MIXED and PROC IML code shows evidence again for effect modification by the significant chi-square for "AEM". Also the F-statistic, F <21.56, and Chi-square, X^2 <21.56, are identical with different P-values. The t value has P 0.0097 for "aem". PROC IML yields P< 0 for the data set indicating statistically significant evidence to reject the null of homogeneous odds for head/neck cancer due to no drinking for level of race.

2. 7 Probability Equation from the Author for z=3 and Hazard Ratio

The probability of three events for confounder/effect modifier can be described by (Agravat 2009 (unpublished), Agravat 2009)):

$$\frac{odds\ (z)}{odds\ (y)} = \frac{1}{odds\ (y) + P(z) - odds\ (y) * P(z)} =$$

$$odds\ (z)[\ odds\ (y) + P(z) - odds\ (y) * P(z)] = 1$$

$$= [\ odds\ (y) + P(z) - odds\ (y) * P(z)] = \frac{1}{odds\ (z)}$$

$$P(z)(1 - odds\ (y)) = \frac{1}{odds\ (z)} - odds\ (y) =$$

$$\hat{P}(z)_{new} = \frac{\frac{1}{odds\ (z)} - odds\ (y)}{1 - odds\ (y)}$$

$$P(z)_{new} = \frac{1 - odds\ (\hat{y})}{1 - odds\ (\hat{y})} = 1; if \rightarrow \beta_z = 0; \beta_y = 0$$

$$\rightarrow undefined \quad (In - Linear - regression)$$

The $P(z)_{new}$ equation shows that the probability follows that CMH test assumption if Beta=0 or not!

This probability equation from the author gives much more insights into probability not possible before when the level of z=3. In the case of head/neck cancer, the beta values are: By=-1.527; Bz -1.548 giving:

$P(z)_{new}$ = 4.74-.2171/1-.2171=5.726 There is no assumption of Beta=0 or no harm as is in logistic regression or linear regression for this new method by the author. The author is expecting non linear regression or interaction to be expected for the sample data. The author's formulas however follow the Cochran Mantel Haenszel test assumption for common conditional odds equal to one. The formulas for probability and hazard ratio used were from the author are:

$$\frac{1}{odds(y) + P(z) - odds(y) * P(z)}$$, which is 1/OR1 (Agravat 2009 (unpublished)). When solved for odds ratio the new odds (z) is:

$$(z)_{new} = \text{odds}(z)_{actual} = \frac{\frac{1}{P(z)_{new}} - \text{odds}(y)}{1 - \text{odds}(y)}$$

If Bz and By =0, then the Cochran Mantle Haenszel test assumption holds for the reciprocal of odds ratio equation presented by the author. Utilizing the new probability equation, the odds (z) is described as follows which is:

1) If Bz and By both are ≠ to 0 then $P(z)_{new}$ exists.

2) If Bz and By both are ≠ to 0 then odds $(z)_{new}$ exists.

3) If By =0 and Bz ≠ to 0 then (z) $_{new}$ approaches odds $(z)_{new} = \frac{\frac{(1+z)}{z} - y}{1-y} = \infty$

4) If Bz =0 and By ≠ to 0 then (z)$_{new}$ odds $(z)_{new} = \frac{2-y}{1-y}$

5) If Bz and By both are =to 0 then P (z) $_{new}$ does not exist

6) If Bz and By both are = to 0 then (z) $_{new}$ does not exist

7) However, if both Bz and By are = 0 then the CMH assumption holds but no probability exists.

The (z) $_{new}$ statistic in terms of odds ratio new is based on the inverse of the probability P (z) $_{new}$ equation solved (Agravat 2007 (unpublished)).

$$P(z)new = \frac{\frac{1}{odds(z)} - odds(y)}{1 - odds(y)} = P(z)new * (1 - odds(y)) + odds(y) =$$

$$\frac{1}{odds(z)} = P(z)new * (1 - odds(y)) + odds(y);$$

$$HR(z)_p = odds(z)new = \frac{1}{P(z)new * (1 - \hat{odds}(y)) + \hat{odds}(y))}$$

Hazard Ratio new or HR $(z)_p$ equals .213 and shows increased risk by .213 times for outcome of death due to the group in question by 1 unit of time without assuming b_z =0 and b_y =0 or no harm because. The normal hazard ratio of "z" is exp (z)=.213. The $(z)_{new}$ has ramifications: 1) inverse and 2) negative of the inverse from (z) $_{new}$ equation does not prove to be correct using standard definitions of probability and odds because then 0=1 which is not true if implicit derivatives of the new equation (z) $_{new}$ $\frac{\partial y}{\partial z}$ or $\frac{\partial z}{\partial y}$ are done to the equation. The moral of this lesson (3) is never assume the outcome has no harm and vary the other covariate! Hazard ratios may also utilize a logit similar to logistic regression however the assumptions differ.

New Probability and P-value for this algorithm of probability by the author for the head neck cancer problem (INHANCE) shows:

$$(\frac{y}{1-y}) = P(z)_{new} - \frac{\frac{1}{z}}{1-y} = -.154 \Rightarrow$$

$$\exp(\frac{y}{1-y}) = .857$$

$$\frac{\exp(\frac{y}{1-y})^2}{n_r} = .0612$$

$$P < .0612$$

This P-value for the odds of outcome for this odds is not significant P<.0612.

$\frac{y}{1-y} = \frac{.2171}{1-.2171} = .277$ Is new from the author for odds of outcome of head neck cancer from data (INHANCE) with exposure no drinking / no smoking based on race with three levels using the novel code from R. The new log odds is: $\log(.277) = -.5571$. The anti-logit involving non-normal data without assuming Beta=0 from binomial or logistic

distribution is $\log(\frac{\exp(-.5571)}{1+\exp(-.5571)}) = \log(\frac{.573}{1+.573}) = -.439$ and has a log odds that can be described.

$\log(\frac{y}{1-y}) = \log(\frac{.2171}{1-.2171}) = -.5568$ .The new logit is -.5568 for head neck cancer from the model (Agresti, 1996) from the author.

The interpretation of the **hazard ratio risk statistic** is that the risk of death is .213 times less for no drinking as exposure which is without assuming other factors in consideration. One expects a less harmful prognosis due to this hazard ratio for head neck cancer from no drinking. The exposure for the head neck cancer has no drinking as exposure and no smoking as referent for exposure according to the 1) nonhispanic, 2) black, 3) Hispanic race.

2.8 The $\hat{O}$ Stat Method and AEM Chi-square for Effect Modification

The $\hat{O}$ Stat Method works by utilizing the data transformed values that come from the data transformation shown in the SAS code and making matrices from "A New Effect Modification P Value Test Demonstrated" by the author. The first two rows and 2x2 matrixes are multiplied by the 2x3 matrix for the next columns. This yields an observed table 3a and expected table is calculated by the O statistic. Next, calculate the expected matrix by first estimating the expected 2x3 table. The observed is obtained by getting the first two rows and all columns of data with variables "cases", "fit", "zxy", and "xzy". Then the 2x3 table is multiplied by the 2x2 table giving a 2x3 table. The new expected table is a type of mean estimate of observed values calculated by multiplying the observed value by row total, then divided by sum total. The {O} statistic is calculated obtaining through matrices. Sum the {O} statistic of each set of these products of matrices for first step $\hat{O}$ has DF 2. The above matrix gives a 2x3 matrix. The same procedure is repeated for the next set and this is shown in tables for rows 3 and 4 of the data transformation. This step yields several 0 values for the observed and the expected values for the 2x3 table for both the observed and the expected hence this may indicate characteristics of singularity. This procedure is repeated in SAS/IML (TM).The method of the "AEM" or the O stat comes from the formula of:

$\hat{O} = \frac{(Obs - \overline{Obs})^2}{(Obs)}$ The matrix calculation is depicted below. This pattern continues in pairs of rows of the data transformational method's procedure.

Observed O stat Matrix

$$\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} * \begin{vmatrix} 1 & 1 & 795 \\ 1 & 1 & 2586 \end{vmatrix}$$

Observed O stat Matrix: $\begin{vmatrix} 2 & 2 & 3381 \\ 2 & 2 & 2586 \end{vmatrix}$

Expected O stat Matrix

$$\begin{vmatrix} 1.1 & 1.1 & 1916 \\ .4 & .4 & 1120 \end{vmatrix}$$

The above matrices are from the head/neck cancer example (INHANCE study).The properties of matrices include commutative properties regarding rings and for multiplication. If E1 is nonsingular, then k*E1 is non-singular (invertible matrix (Wikipedia)). For a square matrix, which is also possible, for the matrices the commutative property of multiplication shows that if E1 or E2 are singular when 0 is a possible determinant when dealing with the O statistics determinants of certain elements, then if E1 and E2 are singular then concentric rings are possible which may be related to how the matrices shown demonstrates the involvement of complex and real numbers which shows properties of being both nonsingular and singular because of nonzero values and being square and having 0 values. Thus the allowance of complex numbers allows more calculations with than without them.

Table 4a: O Stat Matrix Values Observed

2x2 Observed $\hat{O}$	**Col1**	**Col2**	**Col3**	**Total**
	2	2	3381	3385
	1	1	2586	2588
Total	3	3	6967	6973

2x2 Mean $\hat{O}$	**Col1**	**Col2**	**Col3**	**Total**
	1.1	1.1	1916	1918.2
	.4	.4	1120	1120.8
Total	1.5	1.5	3036	3039

Table 7b: O Stat Matrix Values Mean

The columns of the **O** stat are column totals from the PROC IML code. The sum of the O stat from the columns of Each matrix multiplication is summed for each pair of rows of the data transformational code and summed with a P-value outputted by the SAS code in PROC IML and the **O** stat total can be calculated to give the chi square P-value of the “aem” which is the from the **O** stat calculated by the formula above. The mean **O** stat is calculated by the value times the row total divided by overall total. The alternating matrix calculations give O statistics that are used as “aem” for the PROC MIXED calculations. Each set of calculations will give a number value followed by a 0 for the “0” level fit variable which were adjusted for by beta estimates. The number 1 is put for the column of first O statistic and “aem” value that includes the “0” fit variable continued throughout. Deter01 represents observed values and deter1 represents observed mean in the PROC IML SAS code. AEM is another name for the O stat for that matrix calculation. The zeros are ignored in the matrix calculation.

The sum of the “aem” or the O stats after all the matrix calculations give is 1894.5757 with 12 degrees of freedom for the data with a P-value from PROC IML of 0 and P<.0001 for chi-square for aem from PROC MIXED for the same variable “aem” that are statistically significant. You may then reject the null of homogeneous odds. You may then reject the null of homogeneous odds for head/neck cancer and no drinking as exposure and the level of races (Non-Hispanic, black, and Hispanic) of the INHANCE study population. You may generalize that there may be a difference by race for head/neck cancer for no drinking as exposure. The interaction may be less due to protection which is possible. The hazard ratio is .213 which indicates less harm for the outcome head/neck cancer and exposure no drinking based on level of race. The PROC MIXED SAS code shows that the corresponding AEM values in the program created by the author is significant for chi-square and P-value of 21.56 and P<.0001. The “aem” is also significant by the F-statistic, which is for multivariate analysis, rejecting the difference for all groups 21.56 and P<.0001. Hence the SAS code for the data transformational method follows an asymptotic chi square statistic. The zxy” variable has a significant P<0.0001 that is chi-square. ZXY is the effect modified transformed variable. You may conclude therefore that, the effect modifier variable is giving significant evidence to reject effect modification null of homogeneous odds (Agravat 2010 unpublished).

3 CONCLUSIONS

Agravat's method for point estimates of this type of non-normal data with these new risk formulas, indicate that for the outcome of death penalty based on defendant's race, there may be bias due to difference of risks which is over 81 % difference when controlling for potential confounder as black vs white being coded "1". The risk is different by 81 percent for categories when one stratum may be different or lie outside the other. However, white defendants whose victims were black had a 6.30 times increased risk for not getting death penalty which shows tremendous bias. What also is true is that, all other strata when controlling for victim's race with defendant as confounder, were significant. This indicates that there is gross "injustice" or unbalanced scales in the cases of death penalty for black defendants by race in Florida according to this dataset of Radelet and Pierce from the Florida Law Review even without adjusting for over-dispersion. It is not necessarily the size of one strata being huge, but since so many strata's are significant in the side of bias and there is obvious confounding in cases of black defendants where they are statistically significant from confidence intervals where there may be injustice. The justice system may be overly punitive in Florida. The Cochran Mantel Haenszel test shows fewer point estimates because there may be a problem with strata's that have conflict with each other in terms of size of risk differences or are not in the same direction. There is indication of confounding for defendants race by non-zero correlation P-value but no risks produced for this data from the standard method.

The odds ratio of .46 and relative risk of .53 might indicate small difference hence no confounding may be present for controlling from the PROC FREQ CMH test and PROC LOGISTIC. Point estimates alone based on risk the standard risk statistic are not enough, but the difference with Agravat's method is larger scope of risk estimates for visible inferences, e.g. .076 to .066 for odds ratio when controlling for defendant' race and victim's race first which may indicate that there is bias while the standard test does not offer odds ratios due to conflicting risks with PROC FREQ. The other advantage for this new method is that when trying to calculate defendant's race as confounder (all the point estimates are seen in table 2) most of the categories have significant confidence intervals. Also this method is parametric, using "R" for the many choices of distributions available that may be scientifically chosen. For the Cochran Mantel Haenszel Test, the CMH estimates are not available when controlling for the defendant's race which is the question of the analysis. As for inferences, in this case if defendant's race is the confounder and victim's race is the exposure, there are no point estimates with CMH test. For sake of inferences, the Agravat's method for point estimates allow a newer viewpoint showing that when controlling for the explanatory variable, risk is 5.48 and statistically significant not limited by conflicting risks. The advantage of having both point estimates is important and needed for statistical inferences for the future. If this set of conclusions were true consistently over other data sets, you may conclude infer that bias is statistically significant and likely. The results indicate a greater difference for bias in the case for defendant's race being black. The author believes that while most people say "where there is a will there is a way", the author states " where society has a will there's a way", in a responsible society.

Distributions such as Rayleigh and Weibull are available based on the type of data or study given when bias may not be so strong and data follows different patterns. The standard chi-square of R.A. Fisher is defined by:

$$\chi^2 = \sum_i \sum_j \sum_k \frac{(O_{ijk} - E_{ijk})^2}{E_{ijk}}$$. However the author's equation also allows asymptotic chi-square values

with this new technique and the new formula is: $\hat{O} = \frac{(Obs - \overline{Obs})^2}{(Obs)}$.

PROC LOGISTIC shows that for confounding, the chance for confounding is small for victim's race and probable for defendant's race though only the defendant's race has significant P-values P<.038 and <.0276. The author's formulas show significant risks for both the defendant's race and victim's race (see table 3) based on confidence intervals. Confounding must have the outcome dependent on both the explanatory variable and possible confounder. For victim's race there is no significant evidence for the given data from PROC LOGISTIC (SAS) but by the author's formulas proof is given for confounding.

Effect Modification Results The results for risk of head neck cancer for races is different showing that there are statistically significant results P <.0001 for never drinkers may be generalizable for this study of random effects modeled data for the outcome by both the Lackfit option that is chi-square and "AEM" variable proving to be asymptotic chi-square. There may be interaction P <.0001, however based on hazard ratio .213, the effect may be protective and less for the INHANCE study. Despite the exposure never drinkers and never smokers from that pool of data, the risks for head neck cancer are still different by more than 10 percent indicating further investigations are needed because the risk may be explained by indicating that heterogeneity exists for this data and may be helpful in addition to other statistics such as likelihood ratio tests for randomness. The possible interaction as represented by

difference in risks across strata gives evidence that without ignoring this question of data correlation by covariates toward the outcome and without fixing the intercept to 0, there is evidence of significant risk differences for head/neck cancer for race, the effect modifier, that the exposure not smoking and not drinking, is different and statistically significant. This generalization may be a topic of inquiry that may not be ignored for further studies of head neck cancer and race. This is evidence that provides the answer for inferences and without this statistic of risk interaction the question of difference of head neck cancer with regards to race may not be available for consideration. This new effect modification test's P-value allows one to have a statistic to compare between subject heterogeneity. Many tests allow inferences of effect for estimates for fixed effects in non-normal models in longitudinal data analysis where random and fixed effects are incorporated, but this comparison test of risk statistic for P-value being different across strata is useful for many types of data including head neck cancer and race. The new method is known to work on non-normal tests and potentially works on models with low correlations as well in contrast to the correlated data often involved with random effects models. The numbers of strata for potential effect modifier can be multivariate and large in sample size too as seen with the head/neck cancer data. Baseline data may need to be examined further in a newer study of risks and other factors that are also causal.

Effect Modification of Lung Cancer

The author wants to amend the results for lung cancer and passive smoke exposure for countries for "A New Effect Modification P Value Test Demonstrated." The algorithm works but the method used for significant results is given by testing more possibilities for significant p values:

```
proc logistic data=passlungexp5d descending;
Weight count;
class xzy  fit cases;
model cases= fit zxy  xzy / rsq lackfit;
run;
```

Hosmer and Lemeshow Goodness-of-Fit Test SAS ®

Chi-Square	DF	Pr > ChiSq
8.6778	5	0.0003

Therefore one concludes that passive smoke exposure or spouse smoking risks are different for lung cancer based on level of countries United States, Great Britain, and Japan. Hence, effect modification is exists for lung cancer from spouse smoking and counties of United States, Great Britain, and Japan and the null of homogeneous odds is rejected. The author's research for the same data give odds ratios of 1.38 and 1.10 when controlling for confounder and explanatory variable countries and spouse smoking (Agravat 2011) for lung cancer in the same dataset.

Discussion of Randomized Clinical Trial "X" is coded 0 for concentrations above .2 for the drugs. R is 0 for "R" 5 and greater. R corresponds to randomization number. N is the count variable. OR=.72 for "drug" means that the ratio of odds for no positive response by exposure "x" controlling for 'drugs' is .72 times odds of positive response without "x" controlling for "drugs". The relative risk 1 for non-positive response states that the proportion of non-positive response to drugs is 71 % lower for those taking drugs controlling for "z" and 39 % lower due to "x". When controlling for "x", the OR is .59 meaning ratio of odds for no positive response for 'drugs' controlling for "x" is .59 times odds of positive response without "drugs" controlling for "x"'. The relative risk 2 is 1.38 for those with exposure "x" when controlling for "x". One is not able to report attributable risk due to non-causality inferences for the drug as confounder. The efficacy when controlling for drug is 69.1%. This efficacy explains that the no positive response that can be removed due to this drug is 69.1 %. The amount of no positive response that can be removed if the exposure "x" is removed is 55.8 % when controlling for "x" and is statistically significant based on confidence intervals (1.28, 1.49) because causality inference may have support in this trial for the drugs' concentration here due to statistical significance. Levels of toxicity, safety, and effectiveness have to be measured and met according to criteria for a drug to be allowed for public utilization.

Agravat's probability algorithm and Agravat's distribution is included for discussion because it allows you to deal with the probability of events as in the case of the data taken 1 at a time for three variables for example for the probability of outcome.

$$P(n \mid i) = \frac{\binom{(n-i)}{\quad}\binom{(n-i)}{\quad}}{(n-i)^n}$$

$$P(n \mid i=0) = \frac{\binom{(n-i)}{\quad}\binom{(n-i)}{\quad}}{(n-i)^n} = \frac{\binom{(3-0)}{\quad}\binom{(3-0)}{\quad}}{3^3} = \frac{1}{3}$$

$$P(n \mid i=1) = \frac{\binom{(3-1)}{\quad}\binom{3-1)}{\quad}}{2^3} = \frac{1}{2}$$

$$* P(n \mid i=2) = \frac{\binom{(3-2)}{\quad}\binom{(3-2)}{\quad}}{(1)^3} = 1$$

For the last condition,

$$P(n=3 \mid i=3) = \frac{1}{P(n,i)} = \frac{1}{\frac{n!}{(n-i)!}} = \frac{(n-i)!}{n!} = \frac{1}{6}$$

1) For n=3 and i=0:

The derivation below explains the checking of the answers for the above algorithm that addresses a general question that if there are three possibilities to start, and there exist i different possibilities, what is the probability of outcome? In the case of P (3|2) inspections leads one to conclude the answer is 0, but the answer is supported by the derivation below. The probability of P (3|3) is given below as well. This algorithm may work to calculate essential or important probabilities more than demonstrated.

Ln(P(n|i=0))=

Ln(n-i)+ Ln(n-i)- nLn(n-i)-ln1/3+

Ln(n-i)+ Ln(n-i)- nLn(n-i)-ln1/2+

Ln(n-i)+ Ln(n-i)- nLn(n-i)-ln1+

Ln(n-i)+ Ln(n-i)- nLn(n-i)-ln1/6+

=8ln (n-i)-4n*ln (n-i)-ln2

8/4n*2=1/n>>n=3 so P(n|i=0) =1/3

*>>Ln(P(n|i=2))=

Ln(n-i)+ Ln(n-i)- nLn(n-i)-ln1+

Ln(n-i)!+ Ln(n!)- ln1/6=0

$$P(n=3\,|\,i=3)=$$

$$\frac{(n-i)(n-i)}{(n-i)^n}$$

$$\frac{(n-i)^n}{(n-i)*(n-i)} \sim \frac{n\ln(n-i)}{(n-i)(n-i)} =$$

$$\frac{n}{(n-i)} \to \frac{n!}{(n-i)!} \sim P_i{}^n$$

$$P(n=3\,|\,i=3)=\frac{1}{P_i{}^n}$$

$$\frac{(n-i)(n-i)}{(n-i)^n} =$$

$$2\log(n-i) - n\log(n-i)$$

$$2\log n - 2\log i - nlohn + n\log i$$

$$=$$

$$2\log n - i\log 2 - n\log n + i\log n$$

$$2\log n - n\log n - i\log 2 =$$

$$2\log n + n\log\frac{1}{n} - i\log 2 =$$

$$n*n(\frac{1}{n}-2)^{n-i} =$$

$$\pi = 2n$$

$$n*n(1-2n)^{n-i}$$

$$n*n(1-\pi)^{n-i}$$

$$\frac{\pi}{2}*\frac{\pi}{2}(1-\pi)^{n-i}$$

$$(\frac{\pi}{2})^2*(1-\pi)^{n-i}$$

Agravat's – Distribution

$$f(i)=c(i,\phi)+(\frac{\pi}{2})^2*(1-\pi)^{n-i}$$

$$\log\ f(i)=\log\binom{n}{i}+2\log(\frac{\pi}{2})+(n-i)\log(1-\pi)$$

$$\log\ f(i)=\log\binom{n}{i}+2i\log(\frac{\pi}{1-\pi})-\frac{n}{4}\log(1-\pi)$$

Canonical – function

$$\theta=\log(\frac{\pi}{1-\pi})$$

$$\pi=\frac{\exp(\theta)}{1+\exp(\theta)}$$

$$1-\pi=\frac{1}{1+\exp(\theta)}$$

Cumulant – function:

$$b(\theta)=(n-i)\log(1-\pi)$$

$$Mean: b`(\theta)=(\frac{n-i}{1-\pi})$$

$$Variance: b``(\theta)=\frac{(n-i)}{(1-\pi)^2}$$

$Deviance$:

$$D(y,\hat{u}) = 2\sum_{i=1}^{n} y(\tilde{\theta} - \hat{\theta}) - b(\tilde{\theta}) + b(\hat{\theta})$$

$$\theta = u$$

$$b(\theta) = \frac{1}{2}(\theta^2) = \frac{1}{2}y^2$$

$$D(y,\hat{u}) = 2[y^2 - yu - \frac{1}{2}y^2 + \frac{1}{2}u^2]$$

$$= 2(\frac{1}{2}y^2 - yu + \frac{1}{2}u^2)$$

$$Deviance - Agravat - Distribution$$

$$D(y,\hat{u}) = (y - \hat{u})^2$$

$$Maximum - Log - Likelihood$$

$$-2\log\lambda = \frac{D(\omega_1) - D(\omega_2)}{\phi}$$

(data.Princeton). The distribution is similar to the binomial or exponential distributon with mean and variance like the gamma distribution due to proportions.

The denominator of $1-\pi$ in Agravat distribution is similar to problems in quantum statistics in the denominator of $\int_0^\infty \frac{x^n}{e^x \pm 1} dx$ solved with the Riemann Zeta function as well because it also depends on a binomial distribution like probability. The probability algorithm when tested for autocorrelation for the resultant value from Agravat's distribution shows that the intercept and 1-prob are equal but have opposite sign values and has equal standard errors for the two variables intercept and "omp" (1 –Pr) for sample data. The value n if plotted is linear (figure 4a) with the value for Agravat distribution. Figure 4b shows that probability is increasing with "i" for sample data. N and f(i) are independent with the chi-square test with P<4E-7 to P<.000457. Probability is not independent of n P<.2270 or f(i) and they have identical PROC FREQ values for chi-square test. A PROC MIXED test of REML and COVTEST shows that intercept is equal to i=0 value except opposite sign for the "smoktobdev" dataset of sample data (Verbeke and Molenberghs) with residual equal to 0. The subsequent probability shows that probability calculated from Agravat's algorithm follows the t, F, and Chi square distribution as can be seen below from the sample dataset with converging values and are statically significant. The Durbin Watson statistic of 2.44 indicates no autocorrelation for the Algorithm probability and distribution value of f(i). In figure 4 the distribution shows ability to calculate the important electron orbital levels for example for level 9 and 6 orbits equaling to 84-54=30 levels that are important.

Figure 4. Sample Probability Dataset

```
data  SMOKTOBDEV ;
input  pr ni n i omp fi;

datalines;

0.00000021   1     9     0     0.99999970   0.00000029
0.00000048   9     9     1     0.99999950   -11.68000000
0.00000121   36    9     2     0.99999870   -22.43
0.00000357   84    9     3     0.99999600   -24.759
0.00001280   126   9     4     0.99993800   -37
0.00006100   125   9     5     0.99993800   -40.04
0.00045700   84    9     6     0.99954000   -38.157
```

```
;
run;

proc autoreg data=smoktobdev;
model fi=pr n i omp/dw=2 dwprob;
run;
```

Durbin-Watson
Statistics

Order	DW
1	2.1620
2	2.4403

n = 8.99999 * Intercept + 5.4E-6 * pr - 301E-13 * i - 5.73E-6 * omp

Regress R-Square 0.9802

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	B	-55867	75103	-0.74	0.5109
pr	B	83751	71443	1.17	0.3257
n	0	0	.	.	.
i	B	-8.0119	0.8701	-9.21	0.0027
omp	B	55865	75103	0.74	0.5109

Figure 4a. Plot of "n" with Pr from Agravat's distribution and Agravat's Probability Algorithm

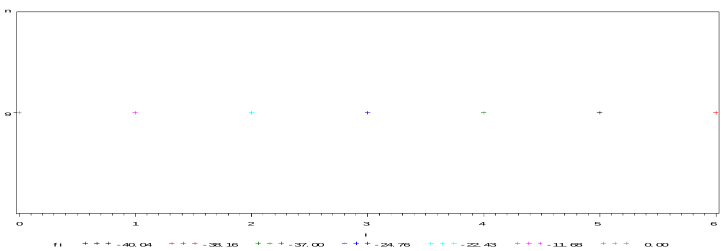


Figure 4b. Plot of Probability with "i" in P(n|i)

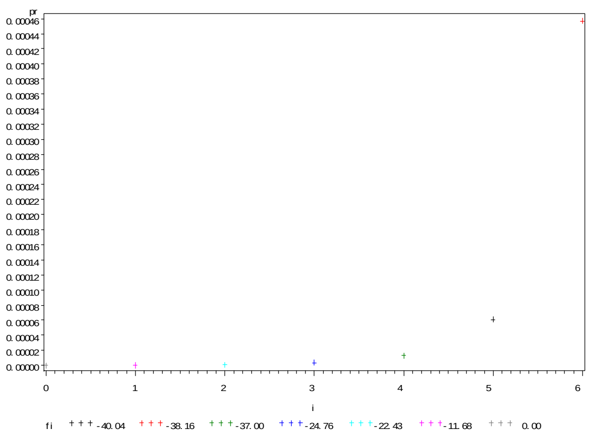


Figure 4c. Statistics of Agravat distribution *f(i)* and Probability form Agravat's Algortihm and Other Variables

The SAS System

The Mixed Procedure

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
pr	31122	9339.31	4	3.33	0.0290

Covariance Matrix for Fixed Effects

Row	Effect	Col1	Col2	Col3	Col4
1	Intercept	4.3063		-1.2456	7363.85
2	n				
3	i	-1.2456		0.5348	-4682.00
4	pr	7363.85		-4682.00	87222797

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
n	0	.	.	.	.	.
i	1	4	128.98	128.98	<.0001	0.0003
pr	1	4	11.10	11.10	0.0009	0.0290

```
proc mixed data=smoktobdev method=reml covtest;
model fi=n i pr/ solution ddfm=satterth covb chisq;
run;
```

Comparison of the analysis of Agravat's distribution and Agravat algorithm $\frac{(n-i)(n-i)}{(n-i)^n}$ shows similar values to the PROC MIXED codes for effect modification and AEM variable using O statistic $\hat{O} = \frac{(Obs - \overline{Obs})^2}{(\overline{Obs})}$ which are proportional to concept of n^n which is a fixed value. On a logarithmic scale the base is same and exponent different. As denominator of F-statistic goes to infinity, i and probability converge to chi-square statistic of the Agravat

distribution and algorithm for probability as in O statistics which use mean values. The probability distribution and

$$t_B \sim \frac{\hat{B} - B_0}{s.e.(\hat{B})}$$

$$t_B * s.e.(\hat{B}) \sim \hat{B} - B_0$$

$$t_B * s.e.(\hat{B}) \sim Estimate$$

algorithm shows that significance for the t and Chi-square distribution (Basu's Theorem): *and* .

$$\frac{(x - \bar{u})^2}{\sigma^2} \sim \chi^2$$

SIGNIFICANCE of PROC IML and PROC MIXED both yield significant results with a new method that is showing how that data transformation method of Agravat's algorithm works. The procedure may work for any produced transformation of the data in PROC IML. PROC MIXED also gives accurate results that include the "AEM" variable for the O stat calculated by the method shown from the author when introduced into the author's algorithm that is asymptotic chi-square. The O stat is a distinct type of formula from the standard statistic for effect modification because it has observed value minus observed mean value squared divided by observed value. Normally, chi-square is observed minus expected squared divided by expected. According to the author, the data transformational method of the author is asymptotic chi-square and the statistics of the P-values follows the asymptotic chi–square distribution. PROC MIXED and "AEM" is asympototic chi-square and the PROC IML algorithm in SAS can be used to calculate the P-value for effect modification as an alternative and log of the ""aem" can be used to simplify calculations.

Since statistic and F-statistic squared which is also the ratio off two chi-squares squared, one may safely conclude that one measure is going towards infinity. At the same time, the lcwocz variable for effect modifier/confounder is P < .0001. This statistic has the range of - ∞ to + ∞ as in intercept and AEM variable in the hyper-geometric distribution data example. For hyper-geometric distributions for n values that are fixed, while N approaches positive infinity (Wackerly, Mendenhall, and Scheaffer). However for the data on myocardial infarction, the values for t-distribution are *positive* or *negative* infinity for intercept in "aem" with estimates of 1 and -1 respectively. There is an anomaly here because t~ $(F)^2$ which is normally positive however t can be negative. The estimate of the intercept is 1 and value is infinity. Since the F test can have high variance in numerator that is negative for between effects which is fixed the value can approach infinity as long as the numerator is small. The denominator may have values of within variance that is small while accounting for randomness. Still the t value may approach negative infinity. There may be sometimes mixed effects for when the within variance is small and numerator or between variance is negative and tends toward negative infinity which may also produce negative infinity values involving the t -distribution.

"Generalized Linear Models", (Nelder and McCullough 1989), stated that in large samples give the approximate distribution for χ2. With normality, there may be exact results. As n approaches ∞ , the degrees of freedom, in the denominator approaches infinity, and the F-statistic is equivalent to χ2. However, the author's method is able to produce this convergence with non-normal data and small sample size, as in hyper-geometric data set regarding myocardial infarction, as well as in large samples as in the case of head/neck cancer. The author's methods with the matrices, algorithm, aem, and PROC IML with PROC MIXED produces, asymptotic chi-squares in the case of effect modification.

In pharmacokinetics, the question of steady state chemistry as well as the physics of energy does become a factor in the metabolism of drugs. Drugs being broken down related to their patho-physiology of the condition may give more insights. This question concerns thermodynamics and thermo-kinetics as well and "first pass effect" in metabolism. Drugs with problems with steady state may present a problem in the human body and may have risks! Many drugs are known to have side effects, but the doctor must choose which drug is best and related to the condition of the patient and according to factors for prognosis due to the stage of the disease for complicated cases.

The analysis and study of drugs and different phases may utilize these formulas for inferences that include causality criteria. In phase I, when the drugs are tested for pharmacokinetics, dosage is a key issue at what concentration the effect is seen. For this trial, after dose .2 units, there is a significant effect statistically with randomization. In phase II, toxicity is important at that stage and safety. If a drug has significant risks then there will be a higher risk that can be seen with the formulas of the author. The therapeutics can be analyzed for long term efficacy in phase III supported for when controlling for the concentration. Finally, phase IV will involve the study of side effects before a drug is found to be safe and goes to the market.

The probability equation given by the author further derives the Agravat series when the third, fourth, and fifth implicit derivative is done to the odds ratio equation when the reciprocal of OR 1 equation is equal to z/y and

solved. These equations are hyper-geometric because they are non-factorable and negative. This equation has z with three levels. The odds ratio equation also has three levels.

Agravat's Proof of Cochran Mantel Haenszel Test Assumption of Θ=1

$$\frac{1}{odds(y) + P(z) - odds(y) * P(z)} = \frac{odds(z)}{odds(y)}$$

$$(odds(z)(odds(y) + P(z) * odds(z) - odds(y) * (Pz) * odds(z)) = odds(y)$$

$$odds(y) - odds(y) * odds(z) + odds(z) * odds(y) * P(z) = P(z) * odds(z)$$

$$odds(y) = odds(y) * odds(z) + P(z) * odds(z) - odds(z) * odds(y) * P(z)$$
$$odds(y) - odds(y) * odds(z) + odds(z) * odds(y) * P(z) = P(z) * odds(z)$$
$$odds(y) - odds(y) * odds(z) = 0$$
$$odds(z) = 1 >>>>> \; z = 1 >> when \ldots . z = 0$$
$$odds(y) - odds(y) * odds(z) + odds(z) * odds(y) * P(z) = P(z) * odds(z)$$
$$>> \; y = 0, odds(y) = 1$$

The author just proved above that using the author's background information for odds ratio, and setting the reciprocal of the odds ratio to z/y, the author is able to prove that the common conditional odds ratio is 1. The assumption is Bz is = 0 and By is shown to be 0. Hence, z/y=1/1 or that the common conditional odds ratio equals 1 as suggested by the COCHRAN MANTEL HAENSZEL test assumption and this author's proof as support for the assumption.

The matrix math, described by observed, mean observed tables, coming from the O statistic and the code and the PROC IML code, give statistics that may yield equal values for the F-statistic and the chi-square for the AEM variable. The t statistic normally requires the normal distribution; however the data is fitting the hyper-geometric distribution and gives negative and positive ∞. For the non-normal data, these equations give more area that can be quantified and explain graphing the three body problem of Poincaire´ (figure 6). The Poincaire´ conjecture (Figure 6) has implications in algebraic topology where f(x, y, z) = (0) can exist and have complex numbers as coordinates as shown in the 3- d sphere at the vertex. The author's hyper-geometric series, Agravat series, allows complex numbers that the author believes may be a factor in accounting for time at least partly. The beta estimates and the use of the formulas from the author allow the vectors and multidimensional mathematics utilized in calculating risks.

Agravat Series The proof behind the Agravat series is as follows:

$$\frac{z}{y} = y + \frac{z}{1+z} - y * (\frac{z}{1+z}) = y^2 + y * (\frac{z}{1+z}) - y^2 * (\frac{z}{1+z}) = z = y^2 + y(1+\frac{1}{z}) - y^2(1+\frac{1}{z})$$

$$= \frac{y^2}{z} + y + \frac{y}{z} - y^2 * (1+\frac{1}{z}) = \frac{y^2}{z} + y + \frac{y}{z} - y^2 - \frac{y^2}{z} = y + \frac{y}{z} - y^2 - \frac{y^2}{z}$$

$$\frac{\partial z`}{\partial y}(1+(y*\frac{\partial(z)^{-1}}{\partial y}+\frac{1}{z})-2y-(y^2*\frac{\partial(z)^{-1}}{\partial y}+\frac{1}{z})=1+(y*\partial\frac{(z)^{-1}}{\partial y}+\frac{1}{z})-2y+\frac{y^2}{z^2}-\frac{1}{z}=$$

$$=1+\frac{2}{z}-\frac{y}{z^2}+\frac{y^2}{z^2}-2y$$

$$\frac{\partial z``}{\partial y}(1+\frac{2}{z}-\frac{y}{z^2}+\frac{y^2}{z^2}-2y)=-\frac{2}{z^2}+(-y*\frac{\partial(z^{-2})}{\partial y}+\frac{1}{z^2})+(y^2*\frac{\partial(z^{-2})}{\partial y}+\frac{1}{z^2})=$$

$$=\frac{2y}{z^2}-\frac{2y^2}{z^2}$$

$$\frac{\partial z```}{\partial y}(\frac{2y}{z^2}-\frac{2y^2}{z^2})=(2y*\frac{\partial(z^{-2})}{\partial y}+\frac{1}{z^2})-(2y^2*\frac{\partial(z^{-2})}{\partial y}+\frac{1}{z^2})=$$

$$\frac{\partial z``}{\partial y}(-\frac{4yz}{z^3}+\frac{1}{z^2}+\frac{4y^2z}{z^3}+\frac{1}{z^2})=\frac{4y}{z^2}-\frac{4y^2}{z^2}+\frac{2}{z^2}$$

$$\frac{\partial z````}{\partial y}(\frac{4y}{z^2}-\frac{4y^2}{z^2}+\frac{2}{z^2})=(4y*\frac{\partial(z^{-2})}{\partial y}+\frac{1}{z^2})+(-4y^2*\frac{\partial(z^{-2})}{\partial y}+\frac{1}{z^2})+\frac{4}{z^2}=$$

$$=-\frac{8y}{z^2}+\frac{6}{z^2}+\frac{8y^2}{z^2}$$

$$\frac{\partial z`````}{\partial y}(\frac{8y}{z^2}+\frac{6}{z^2}+\frac{8y^2}{z^2})=\frac{16y}{z^2}-\frac{12}{z^2}-\frac{16y^2}{z^2}$$

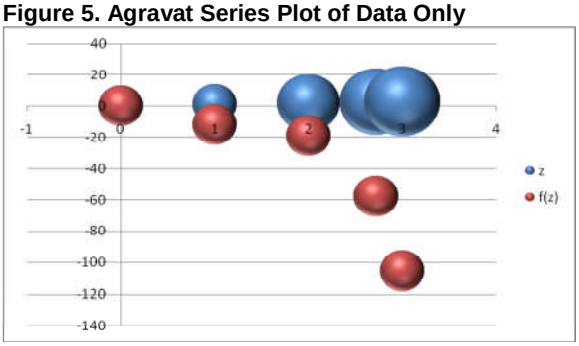
$$f(z_{3rd})=\sum_{n=1}^{n=\infty}-\frac{2}{z^2}+\frac{4y^n}{z^2}-\frac{4y^{n+1}}{z^2}.....Agravat\ –Series-Expression-1$$

$$f(z_{4th})=\sum_{n=1}^{n=\infty}-\frac{6}{z^2}+\frac{8y^n}{z^2}-\frac{8y^{n+1}}{z^2}.....Agravat\ –Series-Expression-2$$

$$f(z_{5th})=\sum_{n=1}^{n=\infty}-\frac{12}{z^2}+\frac{16y^n}{z^2}-\frac{16y^{n+1}}{z^2}.....Agravat-Series-Expression-3$$

(Expressions 2 and 3 come from the implicit derivative of the previous expression).

Figure 5. Agravat Series Plot of Data Only



Agravat series shows a similar relation to the problem of Poincair`e, three body problem, dealing with the sun, moon, and earth's space. The description of the matter of particles may be better described by the hyper-geometric Agravat series for steady state. According to Paul Dirac "Some time before the discovery of quantum mechanics people realized that the connection between light waves and photons must be of a statistical character" (Dirac, 1958).The author believes this link may be hyper-geometric. The 3-D spheres as shown by the author's Agravat series involve the polarization of light (Baym). In addition, the probability of a photon is described by Maxwell's equations. In addition, the author states that the order of these equations may be related to special relativity of Einstein but you may not ignore Heisenberg's uncertainty principle! Heisenberg states in matrix math that his matrices for position and displacement were not commutative (Werner Heisenberg: (wikipedia)),which is important for the steady state but Agravat's O statistic and matrices are commutative (Agravat, 2010 (unpublished)) and also agree that time must be a

factor in complex numbers which can be described by the Agravat series. For $e = mc^2$, with f (z) =c and m=y, there is less destruction or steady state that is associated versus the expanding space with m=z. The author believes the transmutability of mass and energy due to the restriction imposed by the infinite power of destruction possible or may be limited to other particles by the possibility of mass being described in a way that does not take into the account but rather of mass in terms of the outcome and mass in terms of another "variable" that can involve a hyper-geometric function or characteristic to involve transmutability (Theory of Relativity :(wikipedia)) in the.

Figure 6. Radar Plot $e = mc^2$, f (z) =c, **and Agravat Series Transformation , and Delta Function**

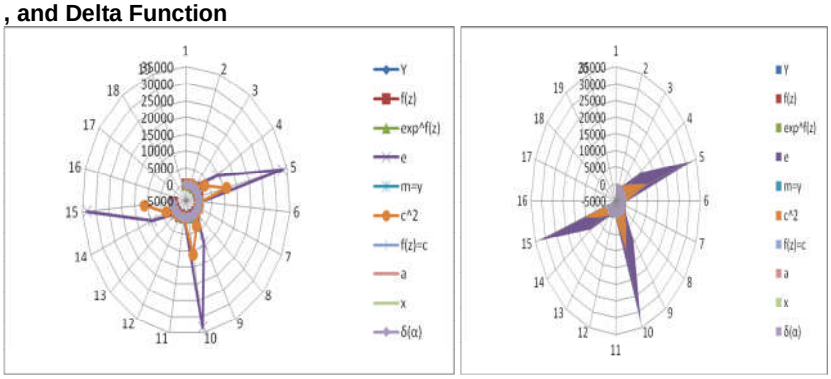
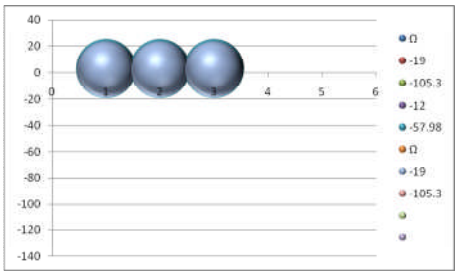


Figure 6 shows the relationship between $e = mc^2$, and Agravat Series and using f (z) =c where energy is going beyond the speed of light squared and mass is central in the plot. Einstein states that mass is a manifestation of energy and a small amount of ass may converted to more energy as shown in figure 6.



$$f(z) = \sum_{n=1}^{n=\infty} -\frac{12}{z^2} + \frac{16y^n}{z^2} - \frac{16(y)}{z^2}^{n+1}$$

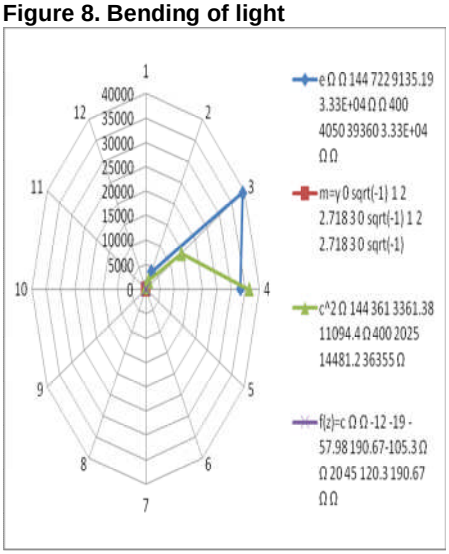
Figure 7. Poincaire` Conjecture and Agravat Series

The use of Agravat series better describes the cohesion of particles for that may be having hyper-geometric properties that can involve the author's Agravat series meant for level of three variables which can be subatomic to hold together in steady state not described by $e = mc^2$ while speed of light ~ to the f (z) function in the third expression of Agravat series. According to Einstein and Bose the temperature and particle energy condensate and behave differently (Bose-Einstein: (hyper physics)). Bose also believed similarly however, the author believes that Bose assumed that when describing photons behavior, he did not assume independence. The author's algorithm and distribution does not assume independence either but the probability algorithm and distribution produce statistically significant values (4c). The author's algorithm's last term is similar to Maxwell-Boltzman distribution which is:

$\frac{1}{P_i^n}$ versus $N!\prod_{i=1}^{k}\frac{1}{N_i!}$ .The Maxwell-Boltzman calculation which can be used to calculate number of ways N atoms can be arranged in an atom $N!\prod_{i=1}^{k}\frac{g^{Ni}}{N_i!}$ in g boxes (Maxwell-Boltzman Statistics:(Wikipedia)). However the author has a distribution as well to supplement the probability algorithm. The probability algorithm of the author gives better answers for larger numbers of variables such as P(5|3)=.125 for the author versus if Maxwell Boltzman is used:1/20. The more precise inferences of the author's algorithm may be more useful for calculations dealing with the chemistry of atoms and hence the pharmacokinetics in the long run. For the behavior of condensates, Bose's assumption may

be correct, and the author's algorithm and distribution does work for probability according to Agravat's distribution. Bose had been told he made an error, but may have been correct when saying that for behavior of photons the behavior may not be independent and had a 1/3 chance for P(3|0) for photons. Thus certain particles do condensate or cluster while others fermions don't at the subatomic level. The author believes that regarding mutability of matter, this factor is a characteristic of subatomic behavior that may share hyper-geometric properties. Mass may be proportional to the y the variable of outcome. If the mass is described by the variable "z", this may result in destruction or decreases of energy levels in the system in steady state. Delta function states that the position of a particle from start is described by this function, however the author believes only when slightly perturbed. The less perturbed perhaps the motion may follow a hyper-geometric distribution or other type of function in part and in rest also an extreme stimulation towards destruction with mass dependent on other variables without control.

Figure 8. Bending of light



The figure 8 shows that there may be bending of light's speed from straight to tangential and reflecting inwards of energy in terms of direction through the Agravat series transformation involved in special relativity not seen before by other methods.

Included is a discussion on the relationship of the Agravat probability and Agravat distribution. To begin with:

$$1)\sigma^2 \sim \frac{z^2}{n} >> z^2 \sim n >> z \sim \sqrt{n}$$

$$2)(s.e.)^2 \sim \frac{s.d.^2}{n} \sim \frac{\sigma}{n}$$

$$3)\sigma \sim (s.e.)^2 * n >> \sigma^2 \sim (s.e.)^4 * n^2$$

$$4)\sigma^2 * n \sim (x - \bar{u})^2 \sim \frac{z^2 * \sigma^2}{n-1}$$

$$5)z \sim \sqrt{n(n-1)}$$

$$6)(s.e.)^4 * n^2 * n \sim (x - \bar{u})^2 > D \sim (s.e.)^4 * n^3$$

The sas code below summarizes the data of new probabilities and distribution values form the author's work that includes value D.

```
data Sve;
input id  se n  i  Pr  omp  logfi  D;
datalines;
1       0.645           3       0       0.3333  0.667       0.131         4.67
2       0.645           3       1       0.5000  0.5         0.702         4.67
3       0.645           3       2       1.0000  0           1/0             4.67
4       0.645           3       3       0.1667  0.8333    0.33      4.67
5       0.763           5       0       0.0080  0.992       0.00436       42.36
6       0.763           5       1       0.0156  0.9844    -1.31         42.36
7       0.763           5       2       0.0370  0.963      -5.47          42.36
8       0.763           5       3       0.1250  0.875      -3.27          42.36
9       0.763           5       4       1.0000  0          1/0           42.36
10      0.763           5       5       0.0083  0.991700   -17.9             42.36
11      0.95700000      9       0       0.00000021      0.9999997 0.00000029    611.47000000
12      0.95700000      9       1       0.00000048      0.9999995       -11.6800    611.47000000
13      0.95700000      9       2       0.00000121      0.9999987       -22.43      611.47000000
14      0.95700000      9       3       0.00000357      0.9999960       -24.759     611.47000000
15      0.95700000      9       4       0.00001280      0.9999380       -37           611.47000000
16      0.95700000      9       5       0.00006100      0.99993800      -40.04      611.47000000
17      0.95700000      9       6       0.00045700      0.99954000      -38.157     611.47000000
18      0.95700000      9       7       0.00078100      0.99210000      -41.89      611.47000000
19      0.95700000      9       8       1.00000000      0.00000000      1/0           611.47000000
20      0.95700000      9       9       0.00000255      0.99999700      0.00000429 611.47000000

;
run;

proc genmod data=Sve descending;
        model logfi= n pr D i/ link=log dist=gamma type3;
ods output type3=te;run;
```

The SAS System
multivariate data

Obs	Source	DF	ChiSq	Prob ChiSq	test	power
1	n	1	9.3056	0.0023	3.84146	0.86226
2	Pr	1	9.4812	0.0021	3.84146	0.86847
3	D	1	9.4824	0.0021	3.84146	0.86851
4	i	1	16.127	<.0001	3.84146	0.98010

The UNIVARIATE Procedure
Variable: i

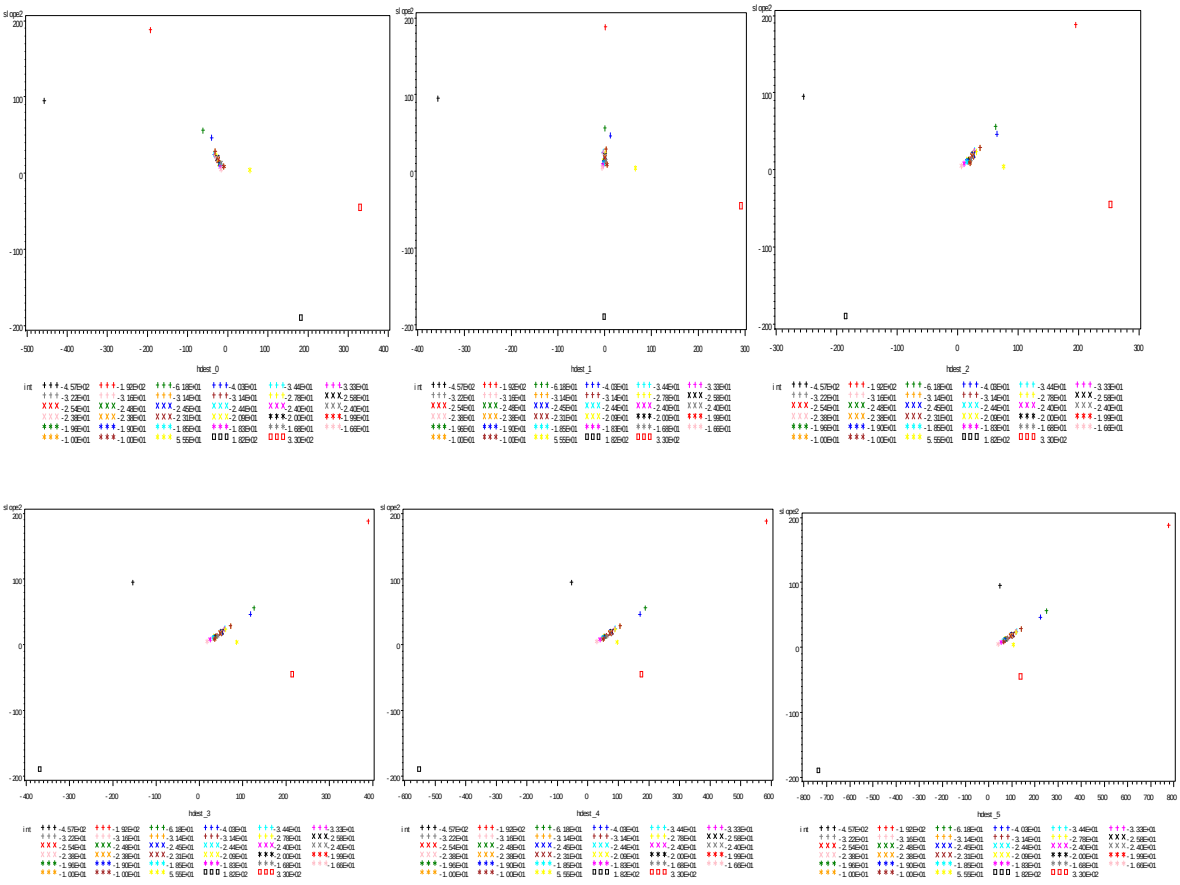
Tests for Normality

Test	--Statistic---		-----p Value------	
Shapiro-Wilk	W	0.931246	Pr < W	0.1632
Kolmogorov-Smirnov	D	0.144608	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.066713	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.442284	Pr > A-Sq	>0.2500

```
axis1 symbol1 color=black interpol=join value=dot  w=.4 l=1;
axis2 symbol2 color=blue interpol=rl value=triangle  w=.4 l=2;

proc gplot data=randest3;
plot slope2*(hdest_0-hdest_10)=int/ haxis=axis1 vaxis=axis2;
plot slope1*(hdest_0-hdest_10)=int/ haxis=axis1 vaxis=axis2;
run;
```

The plots below show a few of the resulting motions and collisions actually predicted by the probability and Agravat's distribution values may be applicable to subatomic motion and with increasing slopes. In this case, slope2 corresponds to n=5. The algorithm and distribution can work for a great many number of possibilities to further knowledge in quantum statistics. In addition, Shapiro Wilk's test shows that "i" is normal P<.1632.



```
PROC MIXED  data=Sve METHOD=ml;
  CLASS N;
  MODEL logfi = n /SOLUTION;
  RANDOM INTERCEPT n /SUB=id TYPE=UN G S gcorr solution;
ODS LISTING exclude SOLUTIONR; ODS OUTPUT SOLUTIONR=randestd;
run;
```

The PROC MIXED code with another code is used to predict slopes and intercepts for the plot of n at 3, 5, and 9. While

Table 5 Spearman's Rho Test from Agravat's Algorithm and Distribution in SPSS (TM)

Correlations								
		D stat	n	i	Pr	1-Pr	log(f(i))	standard error
D stat	Correlation Coefficient	1.000	1.000**	.449*	-.743**	.744**	-.737**	1.000**
	Sig. (2-tailed)	.	.	.047	.000	.000	.001	.
	N	20	20	20	20	20	17	20
n	Correlation Coefficient	1.000**	1.000	.449*	-.743**	.744**	-.737**	1.000**
	Sig. (2-tailed)	.	.	.047	.000	.000	.001	.
	N	20	20	20	20	20	17	20
i	Correlation Coefficient	.449*	.449*	1.000	.020	-.019	-.596*	.449*
	Sig. (2-tailed)	.047	.047	.	.932	.937	.012	.047
	N	20	20	20	20	20	17	20
Pr	Correlation Coefficient	-.743**	-.743**	.020	1.000	-1.000**	.471	-.743**
	Sig. (2-tailed)	.000	.000	.932	.	.000	.057	.000
	N	20	20	20	20	20	17	20
1-Pr	Correlation Coefficient	.744**	.744**	-.019	-1.000**	1.000	-.473	.744**
	Sig. (2-tailed)	.000	.000	.937	.000	.	.055	.000
	N	20	20	20	20	20	17	20
log(f(i))	Correlation Coefficient	-.737**	-.737**	-.596*	.471	-.473	1.000	-.737**
	Sig. (2-tailed)	.001	.001	.012	.057	.055	.	.001
	N	17	17	17	17	17	17	17
standard error	Correlation Coefficient	1.000**	1.000**	.449*	-.743**	.744**	-.737**	1.000
	Sig. (2-tailed)	.	.	.047	.000	.000	.001	.
	N	20	20	20	20	20	17	20

Interpretations of the table of Spearman's rho shows that D value is statsitcally significant for Pr, 1-Pr, log(f(i)) with negative correlation of -.737 for log(f(i)) and -.743 correlation with Probability. "I" is significantly correlated but with factor of.449. "I" has power of 98 % with PROC GENMOD for the Agravat's distribution.

Appendix of data for relativity figures 7.

Y	f(z)	exp^f(z)	e	m=y	c^2	f(z)=c	a	x	δ(α)
0	Ω	Ω	Ω	0	Ω	Ω	0	0	-0.651
1	-12	6.14E-06	144	1	144	-12	1	1	0.001679
2	-19	5.60E-09	722	2	361	-19	2	2	0.326
2.718	-57.98	6.60E-26	9135.19	2.718	3361	-57.98	2.718	2.718	0.2339
3	-105.3	1.86E-46	3.33E+04	3	11094	-105.3	3	3	0.217
0	Ω	Ω	Ω	0	Ω	Ω	0	0	-0.651
1	-12	6.14E-06	144	1	144	-12	1	1	0.001679
2	-19	5.60E-09	722	2	361	-19	2	2	0.326
2.718	-57.98	6.60E-26	9135.19	2.718	3361	-57.98	2.718	2.718	0.2339
3	-105.3	1.86E-46	3.33E+04	3	11094	-105.3	3	3	0.217
0	Ω	Ω	Ω	0	Ω	Ω	0	0	-0.651
1	-12	6.14E-06	144	1	144	-12	1	1	0.001679
2	-19	5.60E-09	722	2	361	-19	2	2	0.326
2.718	-57.98	6.60E-26	9135.19	2.718	3361	-57.98	2.718	2.718	0.2339

```
3       -       1.86E-46  3.33E+04  3       11094   -105.3  3       3       0.217
        105.3
```

References

Agravat 2007 “Test for Calculating the Common Conditional Odds Ratio, Relative Risk, Confounding, Effect Modification, and Matched Pair Odds Ratio.”(Unpublished)

Agravat 2008 “Method for Calculating Odds Ratios Relative Risks, And Interaction”, at:

http://health.usf.edu/medicine/research/**ABSTRACT** BOOK .pdf .

Agravat 2009 “New Method for Calculating Odds Ratios, Relative Risks, and Confidence Intervals after Controlling for Confounders with Three Levels” (unpublished submission: American Journal of Epidemiology).

Agravat 2009 “A New Effect Modification P Value Test Demonstrated” SESUG 2009, Statistics and Data Analysis at: http://analytics.ncsu.edu/sesug/2009/SD018.Agravat.pdf .

Agravat 2010 “Computational Method for Effect Modification with R, SAS, SPSS, and Graphics”, unpublished submission: www.jstatsoft.org).

Agravat 2011 “New Methods to Evaluate Effect Modification with ‘N’ Statistics, and Point Estimates, and Probability for Analysis of Lung Cancer and Passive Smoke Exposure for United States, Great Britain, and Japan” (possible submission: American Journal of Epidemiology).

Agresti, A. “An Introduction to categorical Data Analysis”. NY: John Wiley and Sons; 1996.

Baym, G. “Lectures on Quantum Mechanics”, W.A Benajmin; 1969.

Babette Brumback, Arthur Berg (2008). “Statistics in Medicine” On Effect-Measure Modification: Relationships Among Changes in the Relative Risk, Odds Ratio, and Risk Difference.

Blot W, Fraumeni J (1986), “Passive smoking and lung cancer”, Journal of National Cancer Institute.Vol.77 No.5:993-1000.

Cochran W (1950). “The Comparison of percentages in matched samples.” Biometrika Trust. Vol. (37) No.3/4:256-266.

Paul Dirac. “The Principles of Quantum Mechanics”, Fourth Edition. Oxford : (1958).

Mia Hashibe ET. al. (2007) “Alcohol Drinking in Never Users of Tobacco, Cigarette Smoking in Never Drinkers, and the Risk of Head and Neck Cancer: Pooled Analysis in the International Head and Neck Cancer Epidemiology Consortium”. Journal of National Cancer Institute. Vol: (99) No.10/777-7789.

Lindsey, J.”Applying Generalized Linear Models”. Springer-Verlag:1997.

Nelder and McCullagh. “Generalized Linear Models”. Chapman and Hall/CRC: 1989.

Thomas, G Jr.. “Calculus and Analytic Geometry”. Addison-Wesley: 1972.

Szklo, M., & Nieto, F. “Epidemiology Beyond the Basics”. Jones and Bartlett; 2007.

Verbeke, G, Molenberghs, G.: “Linear Mixed Models for Longitudinal Data”. Springer Series in Statistics, 2000.

Dennis Wackerly, William Mendenhall, Richard Scheaffer, 2002. “Mathematical Statistics with Applications”.

Duxbury Advanced Series.

Mayo Clinic “Medications and supplements that can raise your blood pressure”, at: http://www.mayoclinic.org/mcitems/mc5700-mc5799/mc5751-16.pdf .

R Development Core Team (2004). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3 http://www.R-project.org .

SAS Institute Inc (2003). The SAS System, Learning Edition 4.2. Cary, NC. http:// www.sas.com/ .

Basu’s theorem (Accessed 4/02/2011) http://en.wikipedia.org/wiki/Basu%E2%80%99s_theorem .

“Birth Control Pills for Women Who Smoke” at red http://www.ehow.com/about 5572248 birthcontrolpills-women-smoke.html .

“Elementary Calculus” (Accessed 3/2011)

http://www.math.wisc.edu/~keisler/calc.html .

Cox-Model (accessed 2/02/2011)

http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/COX_MODEL.pdf .

“The GENMOD Procedure: Logistic Regression” (accessed 01/2007) http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_genmod_sect053.htm .

“Generalized Linear Model Theory”

http://data.princeton.edu/wws509/notes/a2.pdf

SAS Institute Inc (2003). The SAS System, Version 9.1. Cary, NC. URL http://www.sas.com/ .

“Introduction to SAS. UCLA: Academic Technology Services, Statistical Consulting Group”. Available at: http://www.ats.ucla.edu/stat/sas/notes2/ .

“Heart Disease” at http://www.womenshealth.gov/faq/heart-disease.cfm .

”Homogeneous Association” available at: http://onlinecourses.science.psu.edu/stat504/?q=book/export/html/114 .

“Invertible Matrix” (accessed 4/2010) http://en.wikipedia.org/wiki/Invertible_matrix .

Lab Objectives (accessed January, 2008)

www.stanford.edu/~kcobb/hrp261/lab2.doc .

Normal distribution (accessed 2/20/2011) http://www.answers.com/topic/normal-distribution .

Kenneth Portier at:
http://ifasstat.ifas.ufl.edu/SAS/New%20Folder/lesson14/lesson14.htm .

"Bose-Einstein Distribution" (accessed 3/14/2011)

http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/disbe.html .

" The Poincaire´ conjecture" (accessed 1/31/11)

http://www.claymath.org/poincare/ .

"Theory of Special Relativity" (Accessed 2/25/2011)
http://en.wikipedia.org/wiki/Theory_of_relativity .

"Werner Heisenberg"
http://en.wikipedia.org/wiki/Werner_Heisenberg .

Daniel Zelterman. "Advanced Log-linear" at http://ftp.sas.com/samples/A57496 .

R Development Core Team (2004). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3- 900051-00-3, http://www.R-project.org.

PASW 18 Gradpack SPSS ™ IBM Inc. (2010). Windows User's Guide. SPSS Inc., Chicago,IL.

Acknowledgements
The author wants to thank his daddy, Dr. Bansidas M. Agravat, for telling me never to quit and giving support (including financial), mom Gauri B. Agravat, and wife Kavita M Agravat, son Gopal M Agravat, for support, brothers Jaydeep and Sanjay for computer time, late grandparents, cousins, and few friends and faculty that that added to my inspiration to help people. I include mathematics Phd instructors I knew, and the study of medicine which gives understanding. I thank Pt. Muley who encouraged me to be patient. Also, personal metaphysics inspires me further to understand man's limits. The tunes of Hendrix, plus more sends me back making me happy to think of more, and especially God thanks! Sometimes inspiration comes from the "hunger" for it! Inspiration may come to anyone, whether while sitting under a tree or a sudden flash of brilliance, but one has to have the hunger and the hard work to follow up with it.

Manoj B. Agravat MPH
20108 Bluff Oak Blvd
Tampa, Florida
E-mail:m_agravat@yahoo.com
(813)-846-0903