Detecting Opioid-Related Risk Factors in Cincinnati, Ohio and Tempe, Arizona

Drew Gobbi
George Washington University
Washington, D.C., USA
drewgobbi@gwmail.gwu.edu

Maddie Warndorf
George Washington University
Washington, D.C., USA
mwarndorf65@gwu.edu

Peiyu Wang
George Washington University
Washington, D.C., USA
oliviawpy@gmail.gwu.edu

Olatunji Akinbule
George Washington University
Washington, D.C., USA
akinbule_ola@gwmail.gwu.edu

***Abstract*—This paper presents an analysis plan, complete with pilot results, for a relational study of demographic and geographic covariates and their association with opioid-related EMS calls. Building on prior literature, the paper successfully replicates the findings of a Poisson model testing for these relationships in Cincinnati Ohio. Additionally, it extends this method to Tempe, Arizona and introduces a similarly specified logistic regression model for both cities of interest. In general, we find that results do not extrapolate well between cities. Hence, further research is needed to identify generalizable opioid-related risk factors at the community level.**

Keywords—Opioids, Public Health, Machine Learning

I. Introduction

Misuse of prescription and illicit opioids has become an epidemic in the United States that claims an average of 130 lives each day [1]. Recently, public health researchers have sought to leverage widely available demographic, geospatial and temporal datasets to understand the association between certain community-level factors and indicators of opioid-related drug abuse. Determining which relationships between community factors and drug use are statistically significant can help inform public policy interventions that might improve the care available to citizens.

This paper presents an analysis plan for a relational study of city-level features and opioid use. The goal of our eventual analysis will be to determine which features represent significant associations worthy of of both targeted policy responses and a more robust causal analysis of any potential interventions. The framework for selecting these features is deliberately broad, and in the final study we intend to utilize essentially all of the demographic and geospatial indicators available to us. However, for the purposes of demonstrating the viability of our research methods, we utilize only those that are shown to be statistically significant in prior research as explanatory variables in the pilot results presented throughout this paper [2]. These include the percentage individuals in the 18-24 and 25-45 age cohorts, the percentage of males, the percentage of individuals with a bachelor's degree or higher, the area median income, and access to health care facilities within an urban census tract.

Because it is an illicit activity, opioid-related abuse is difficult to observe directly. We utilize the monthly count of opioid-related emergency medical service (EMS) calls in a given census tract as an indirect proxy for this construct. Using pilot data, we then measure the association between the monthly tract-level EMS calls through two analytical methods. First, we implement a simplified version of the Poisson regression model originally proposed by Li et al. on pilot data for Cincinnati, Ohio (as in the original study) and Tempe, Arizona to attempt to replicate the author's findings and test their generalizability in a new environment [2]. Secondly, we use a logistic regression to test the relationship between our covariates in a classification framework. In doing so, we posit that any covariates that present statistically significant relationships across all three scenarios represent key results worthy of a targeted response by policy makers.

II. Data Sources

We use opioid-related EMS call data published by the Cincinnati and Tempe open data initiatives as the dependent variables in our study [3][4]. We then extract several demographic covariates from the US Census Bureau's planning database (PDB) file and geospatial covariates (medical facility data) from city level feature levels published to ESRI to serve as our explanatory variables [5][6][7][8][9][10][11][12][13]. Each of these measures is aggregated at the census tract and month-year (e.g. January, 2018, etc.) so that our unit of analysis is a census tract in a given month in a given year.

The theoretical population we wish to extrapolate to is all opioid users, but this is only accessible across the two studies we extract data from. Our sampling frame for observing opioid use in the accessible population is the EMS call data recorded for each cities because it is the available proxy we are using to detect this phenomenon. Using the EMS call sampling frame, we have 1,243 observations in Cincinnati and 440 observations in Tempe.

Because of the limited number of data points available for each city, we use non-probability sampling and include all observations in our study. This is a valid approach because our design is focused more on detecting the presence of relationships between our explanatory variables within each city rather than proving that these relationships are causal. These data and sampling limitations introduce some key threats to external validity in our study. First, Tempe and Cincinnati represent smaller cities in rather populous states. It may not be reasonable to assume that localized results found in each will extrapolate well to the larger state or to the entire opioid using population in the United States. Secondly, the ways in which opioid-related EMS calls are classified, de-identified and published between cities may further threaten our ability to generalize results. As such, we argue that any covariate found to be predictive in both cities represents a key relationship for further study because it does so in spite of these limitations that we expect to threaten generalizability.

III. Measurement, Validity, and Reliability of Constructs

The main construct in our analysis is opioid usage, as measured through the number of opioid-related EMS calls across our sample. It is nearly impossible to measure drug usage beyond incidents where users receive treatment and a medical record is created. As a result, a non-trivial percentage of drug usage is immeasurable. Opioid-related EMS represent an available, face-valid records that allow us to observe drug-related activity. Though there is some potential for “false alarm” calls where the patient is not using opioids, EMS calls still represent an upper threshold of the incoming number of patients in an area that may have been using opioids. EMS data is also content valid because it is recorded on a standardized scale by emergency responders.

A secondary construct we are interested in measuring is accessibility to medical facilities within each census tract. Medical facilities such as hospitals or pharmacies represent places where an individual could either receive treatment for opioids or be prescribed opioids themselves. Thus, certain types of facilities may have a positive or negative correlation with the number of opioid-related EMS calls in an area. We measure this construct by the count of a given medical facility type within two miles of a census tract centroid. This is a face-valid measure of each type of medical facility’s geospatial concentration, so long as we believe that the locational data we extract is accurate.

Finally, the demographic covariates we extract from the PDB, such as the percentage of males, the percentage of 18-24 and 25-45 year olds, the percentage of college educated individuals, and the area median income are all face and content valid measures of their respective constructs. The Census has standardized methodology for which it records and calculates the measures it collects, and publishes a margin of error for each of these measures that is widely accepted in academic research.

All of the measures we list above are ratio-measures because zero represents the absence of any of these phenomena in our sample. We intend to demonstrate the convergent-discriminant validity between these measures using principle component analysis (PCA) of our demographic and geospatial predictors. This will reduce each set of explanatory variables to only a set of components that are uncorrelated with one another and explain a large proportion of the overall variance. We have not implemented a PCA at this time because the four PDB demographic variables we include in our pilot analysis are all largely uncorrelated with one another with coefficients less than 0.5. We will implement a full dimensionality reduction of all covariates, including the entire PDB dataset when we implement our final analysis plan.

Much like the validity, we are unable to demonstrate the reliability of each constructs because we don't have multiple measures for each construct. We know that the EMS calls are not as reliable as possible other measurements of opioid usage, but we are hoping to implement another measurement in the future.

IV. Causal Theory, Proposed Quasi-Experimental Design, and Limitiations

Our causal theory is that some features of a community, such as median income or age could be associated with higher or lower incidences of opioid-related EMS calls. The presence of a positive correlation in this sense could imply that certain features are "risk factors" for a community, while a negative correlation could imply some element of drug use deterrence in an area. This theory is necessarily broad, because we do not know for sure what could be associated with higher incidences of opioid-related EMS calls.

Our pilot study uses a more explicit specification of this hypothesis and assumes that the results of the study conducted by Li et al. are correct in Cincinnati and that they extrapolate well to other cities. So, we posit that the covariates Li finds to be statistically significant and positively associated with opioid-related EMS calls, the percent of males and percent of individuals age 25-45, will also be positively correlated in our replication in Cincinnati and in our extension to Tempe. Pearson correlation tests of the relationship between these covariates and the monthly count of opioid-related EMS calls provide little evidence to support this theory, with most coefficients falling below 0.3. Since, the monthly count number of EMS calls is power law distributed, we expect that the linear Pearson correlations to underfit to the data. Accordingly, expect some of these regressors to have a more pronounced effect when we fit our Poisson regression model [2].

It is impossible for us to conduct an experimental design in this study, because opioid use is self-selected behavior amongst the population that we can merely observe amongst the data available. Even if its random assignment were possible here, it would hardly be ethical. Instead, we utilize a quasi-experimental, patched up, mixed-materials design to test our findings. For each unit of analysis in Tempe and Cincinnati, we can observe a breakdown of our selected demographic and geographic covariates and determine which appear to be more prevalent in areas with a higher number of opioid-related EMS calls. We can then compare these results in each city to our control, which is the results found in the Li. et al. paper. Covariates that are found to be statistically significant and of the same magnitude and direction across all three scenarios are said to generalize well and be reliable indicators of potential opioid use.

Our design faces some self-selection related threats to internal validity. First, individuals in each census tract choose to live in that area. This could bias our estimates because the non-random assignment of subjects across space throughout our sample could be related to the prevalence of EMS calls, without being a function of the demographic makeup of an area. Similarly, because individuals can move in and out of an area throughout our study period, we could face a history confound that further biases our results. For example, if individuals with a higher propensity to dial 911 migrate into a high opioid using tract, we could expect the number of EMS calls to go up regardless of the demographic makeup of that area. Secondly, we are unable to determine the difference between prescribed and non-prescribed opioid-related calls since we are only looking at tract level monthly totals. Our EMS call measures of drug use represent an upper bound that could overestimate drug-usage. Third, limited sample sizes and potential data mislabeling can limit the overall generalizability between each of our three groups. Due to these constraints, we rely only on the standard 0.05 Type-1 error rate in our pilot analysis and do not perform any corrections due to our experimental design.

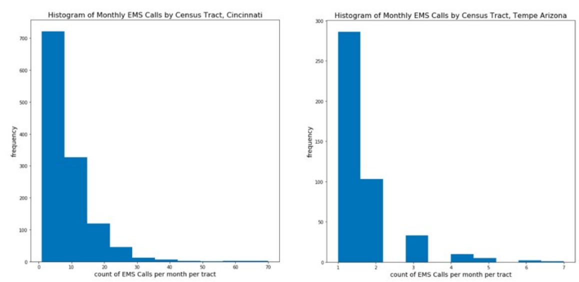
The main statistical method we will use to test our design is a simplified version of the Poisson regression used by Li. et al [2]. We choose the Poisson distribution family for our generalized linear model because we observe the monthly count of EMS calls to be power law distributed, as shown in the histograms below. As such, the Li. et al. model follows the form:

$$y_{it} \sim Poissosn(\lambda_{it})$$

$$\log(\lambda_{it}) = \beta X_{it} + \alpha_i + \varphi_t + \delta_{it}$$

where it is assumed that the number of EMS calls per month in each tract is a Poisson random variable determined by an intensity parameter lambda. This intensity parameter is then said to be a function of each of our covariates, contained in the vector Xit, fixed effects alpha and phi that represent the time and tract dependent heterogeneity created by each data point, and a stochastic error term delta.

Figure 1: Histogram of Monthly EMS Calls in Cincinnati (L) and Tempe (R)

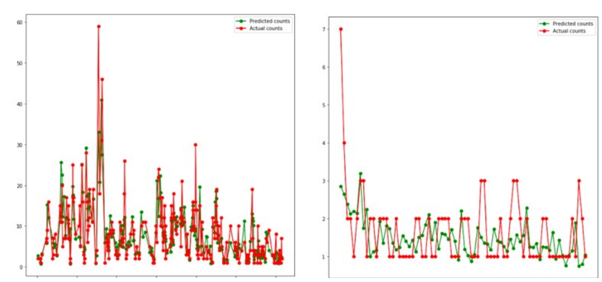


As expected, the Poisson model is largely successful at replicating Li et. al's findings. We find that the percentage of males and the percentage of those who are 25-45 have a positive association with EMS calls. Likewise, we find that the percentage of individuals with a college education, the percentage of individuals who are 18-24, area median income, and the concentration of certain types of healthcare facilities such as mental health centers, naloxone distribution centers, urgent care, and prescription drug drop offs have a negative association. The exponential of nearly these coefficients is approximately 1. For example, this means that for something like a college education, a one unit increase in the proportion of college educated individuals in a given census tract could be associated with one fewer opioid-related EMS call per month. Conversely, a one unit increase in the proportion of males in a census tract in Cincinnati could be associated with one more opioid-related EMS calls per month in that tract.

Our results differ quite a bit in Tempe. Here, we find that the area median income and the concentration of drug drop off sites are positively associated with opioid-related EMS calls while only the proportion of college educated individuals is negatively associated. Nearly all other covariates are statistically insignificant. This could be an effect size issue, where the smaller dataset for Tempe is unable to detect these relationships while the larger Cincinnati dataset is. However, a rudimentary power analysis, shown in the preceding section appears to show that this should not be a problem. These differing results could also be due to the limitations of our design in terms of external generalizability between cities.

Model predictions on a test set appear to demonstrate a good fit to the data. One potential concern to note is that the model deviance of the Cincinnati regression slightly exceeds the critical X2 value, which could imply that the monthly count of EMS calls is not Poisson distributed. Indeed, the mean number of EMS calls is not equal to its variance (although it is approximately equal to standard deviation) as we would expect in a Poisson distributed variable. While there are similar problems in the Tempe data, model deviance is below the critical value, implying some goodness of fit.

Figure 2: Predicted (green) vs. Actual (red) EMS Calls in Cincinnati (L) and Tempe (R)



We also analyze the relationships between the EMS calls and our covariates in a classification framework using logistic regression. To do this, we convert the monthly EMS calls per census tracts into a binary variable equal to one if the number of EMS calls is above the sample median and 0 otherwise. Results appear to be entirely driven by our geospatial predictors (e.g. the locations of certain types of medical facilities) rather than our demographic predictors. In Cincinnati, we find that hospitals, surgical centers and pain management facilities have positively associated odds ratios with the above median EMS call class. This suggests that an increase in the concentration of these types of facilities in a given census tract would lead to an increase in the likelihood of having above median opioid-related EMS calls in that area. Conversely, we find that facilities like drug drop offs, psychiatric hospitals, and children's hospitals would decrease this same likelihood. In Tempe, our model fails to converge entirely. Models in both cities produce pseudo-R2 metrics below .25, suggesting that they explain less than a quarter of the variation across our sample. This means that our classifier can stand to improve significantly. In our full-analysis, we also intend to test additional classifiers.

Finally, we run power analyses in G*Power for each model to determine if the pilot data we are using is adequately sized to detect reasonable effect sizes. To detect a small effect size of 0.25 with a type-1 error rate of 0.05 and 80% power, we find that the model degrees of freedom for both Cincinnati and Tempe imply that we have an adequate sample size. Using the same key statistics, for our logistic regressions, we find that only Cincinnati has enough data for detection. The intensity parameter of one in Tempe would require about five times as much data as currently available. Full results of these analyses are presented in the paper's technical appendix.

V. Discussion of Results and Conclusion

Using a patched-up quasi-experimental design, we propose to test the reliability of tract-level demographic features (extracted from the Census) and medical facility concentration data on predicting higher incidences of opioid related EMS calls in Cincinnati, Ohio and Tempe, Arizona. We intend to use both a Poisson regression model and a classification framework to detect these relationships. Moreover, we leverage the results of Li. et al. to test a pilot hypothesis that the proportion of males and middle aged individuals may increase the propensity of these EMS calls [2].

Initial results of our pilot study open several key insights into our analysis plan. While we find that the Li. et al. Poisson model replicates well in Cincinnati, we find a relatively low goodness of fit and poor generalizability when we extend the same results to Tempe [2]. Similarly, we find that a logistic regression using the same covariates as our main model has poor explanatory power and keys in on a different set of predictors entirely. These divergences could arise from numerous factors beyond the limitations of our design. For example, Ohio and Arizona receive similar levels of funding on a per capita basis, but have very different ways in which they address the opioid epidemic. Ohio tends to be patient centered and focuses on helping users, while Arizona's policies are more focused on improving the effectiveness of emergency treatment. Accounting for these differences within the context of our model will be a key focus going forward.

References

[1] "Opioid Basics", Opioid Overdose, 2019. [Online]. Available: https://www.cdc.gov/drugoverdose/epidemic/index.html. [Accessed: 10- Dec- 2019].

[2] Z. Li et al., "Suspected heroin-related overdoses incidents in Cincinnati, Ohio: A spatiotemporal analysis", PLOS Medicine, vol. 16, no. 11, 2019. Available: https://journals.plos.org/plosmedicine/article/file?id=10.1371/journal.pmed.1002956&type=printable. [Accessed 10 November 2019].

[3] "Heroin", City of Cincinnati, 2019. [Online]. Available: https://data.cincinnati-oh.gov/Safety/Heroin/7mtn-nnb5. [Accessed: 10-Oct-2019]

[4] "Layer: Opioid Calls (ID:0)", ArcGIS REST Services Directory, 2019. [Online]. Available: https://services.arcgis.com/lQySeXwbBg53XWDi/ArcGIS/rest/services/Opioid_Calls/FeatureServer/0. [Accessed: 10- Oct- 2019].

[5] "2019 Planning Database Tract Data", Planning Database with 2010 Census and 2013 – 2017 American Community SurveyData, 2019. [Online]. Available: https://www.census.gov/topics/research/guidance/planning-databases.2019.html. [Accessed: 10- Oct- 2019].

[6] "Health Care Facility Directories", Health Care Facilities and Regulations, 2019. [Online]. Available: https://chfs.ky.gov/agencies/os/oig/dhc/Pages/hcf.aspx. [Accessed: 10- Oct- 2019].

[7] J. Lindemann, "National Prescription Drop Box Layer", ArcGIS, 2018. [Online]. Available: https://www.arcgis.com/home/item.html?id=7726ad6095c543b2a75ffae24b27e419#overview. [Accessed: 10- Oct- 2019].

[8] T. Peacock, "Walgreens Drug Kiosk Locations", ArcGIS, 2016. [Online]. Available: https://opioidepidemic.maps.arcgis.com/home/item.html?id=835b14782ffb4c259c3c2913e68ca0ef#overview. [Accessed: 10- Oct- 2019].

[9] "Ohio Pharmacies Dispensing Naloxone Without a Prescription", Naloxone Pharmacy, 2019. [Online]. Available: https://www.pharmacy.ohio.gov/Licensing/NaloxonePharmacy.aspx. [Accessed: 10- Oct- 2019].

[10] "Ohio OneSource", Ohio Public Health Information Warehouse, 2019. [Online]. Available: http://publicapps.odh.ohio.gov/EDW/DataBrowser/Browse/OhioOneSource. [Accessed: 10- Oct- 2019].

[11] "DrugDropOffLocations (FeatureServer)", ArcGIS REST Services Directory, 2019. [Online]. Available: https://services1.arcgis.com/mpVYz37anSdrK4d8/ArcGIS/rest/services/DrugDropOffLocations/FeatureServer. [Accessed: 10- Oct- 2019].

[12] "Pharmacies (FeatureServer)", ArcGIS REST Services Directory, 2019. [Online]. Available: https://services1.arcgis.com/mpVYz37anSdrK4d8/ArcGIS/rest/services/Pharmacies/FeatureServer. [Accessed: 10- Oct- 2019].

[13] "AZLicensedFacilities (FeatureServer)", ArcGIS REST Services Directory, 2019. [Online]. Available: https://services1.arcgis.com/mpVYz37anSdrK4d8/ArcGIS/rest/services/AZLicensedFacilities/FeatureServer. [Accessed: 10- Oct- 2019].

[14] G. Hoagland et al., Tracking Federal Funding to Combat the Opioid Crisis. Washington, D.C.: Bipartisan Policy Center, 2019.

TECHNICAL APPENDIX

1. Pearson Correlations between EMS Calls in Cincinnati (L) and TEMPE (R):

Cincinnati, OH	ems_call
HospitalsGeneralCount2mi	0.137245
HospitalsPsychCount2mi	-0.029317
SurgicalCentersCount2mi	-0.062300
UrgentCareCount2mi	-0.118978
WomensClinicsCount2mi	0.172335
PainManagementCount2mi	0.257200
PhysicalTherapyCount2mi	-0.001588
MentalFacilitiesCount2mi	0.099455
SubstanceAbuseCount2mi	0.102888
OtherMedicalFacilitesCount2mi	0.210124
NursingAssistedLiveCount2mi	-0.047322
HospiceCount2mi	0.124001
HospitalsChildCount2mi	-0.027375
ChildFacilitiesCount2mi	0.271944
PharmCount2mi	0.166354
DrugDropCount2mi	0.039115
NaloxoneDistribCount2mi	0.056832
pct_College_ACS_13_17	-0.178175
pct_Males_ACS_13_17	0.207492
pct_Pop_18_24_ACS_13_17	-0.070584
pct_Pop_25_44_ACS_13_17	0.025740

Tempe, AZ	ems_call
pdb2018trv4_us_csv_Med_HHD_Inc_	-0.079807
pdb2018trv4_us_csv_pct_Males_AC	0.018494
pdb2018trv4_us_csv_pct_Pop_18_2	0.134348
pdb2018trv4_us_csv_pct_Pop_25_4	-0.054822
pdb2018trv4_us_csv_pct_College_	-0.055565
ems_call	1.000000
HospitalsGeneralCount2mi	-0.070949
HospitalsPsychCount2mi	-0.074100
SurgicalCentersCount2mi	-0.186461
UrgentCareCount2mi	-0.065176
WomensClinicsCount2mi	-0.048554
PainManagementCount2mi	0.138430
PhysicalTherapyCount2mi	-0.100956
MentalFacilitiesCount2mi	0.017484
SubstanceAbuseCount2mi	-0.029791
OtherMedicalFacilitesCount2mi	-0.067443
NursingAssistedLiveCount2mi	-0.208814
HospiceCount2mi	-0.097138
ChildFacilitiesCount2mi	-0.170895
PharmCount2mi	-0.165553
DrugDropCount2mi	0.248664
NaloxoneDistribCount2mi	-0.026190

2. Poisson Regression Results (excluding fixed effects) for Cincinnati

Dep. Variable:	ems_call	**No. Observations:**	1018
Model:	GLM	**Df Residuals:**	893
Model Family:	Poisson	**Df Model:**	124
Link Function:	log	**Scale:**	1.0000
Method:	IRLS	**Log-Likelihood:**	-2489.9
Date:	Wed, 11 Dec 2019	**Deviance:**	1338.8
Time:	22:19:56	**Pearson chi2:**	1.37e+03
No. Iterations:	6	**Covariance Type:**	nonrobust

	Coefficient	Absolute values, e^coefficient	P-value
HospitalsGeneralCount2mi	0.183403	1.201298	7.335100e-09
HospitalsPsychCount2mi	0.085906	1.089704	2.121867e-01
SurgicalCentersCount2mi	0.160123	1.173655	3.134214e-08
UrgentCareCount2mi	-0.568132	0.566583	4.024145e-36
WomensClinicsCount2mi	-0.245960	0.781953	7.610889e-08
PainManagementCount2mi	0.456629	1.578742	4.829786e-12
PhysicalTherapyCount2mi	-0.172056	0.841932	1.679064e-03
MentalFacilitiesCount2mi	-0.084763	0.918730	1.537189e-02
SubstanceAbuseCount2mi	0.014819	1.014929	6.845912e-01
OtherMedicalFacilitesCount2mi	0.023969	1.024258	3.129950e-03
NursingAssistedLiveCount2mi	0.064082	1.066180	9.666082e-07
HospiceCount2mi	-0.203770	0.815650	1.133279e-03
HospitalsChildCount2mi	-0.212292	0.808728	1.193597e-16
ChildFacilitiesCount2mi	-0.361914	0.696342	1.258922e-04
PharmCount2mi	0.053491	1.054947	4.208775e-10
DrugDropCount2mi	-0.177135	0.837666	1.931585e-05
NaloxoneDistribCount2mi	0.016717	1.016857	1.686287e-02
pct_College_ACS_13_17	-0.004788	0.995223	4.726036e-02
pct_Males_ACS_13_17	0.007364	1.007391	7.250650e-02
pct_Pop_18_24_ACS_13_17	-0.007660	0.992369	2.173639e-03
pct_Pop_25_44_ACS_13_17	0.013479	1.013570	1.377294e-02
mhi	0.000006	1.000006	3.767413e-02

3. Poisson Regression Results (excluding fixed effects) for Tempe

Dep. Variable:	ems_call	**No. Observations:**	341
Model:	GLM	**Df Residuals:**	274
Model Family:	Poisson	**Df Model:**	66
Link Function:	log	**Scale:**	1.0000
Method:	IRLS	**Log-Likelihood:**	-433.99
Date:	Wed, 11 Dec 2019	**Deviance:**	85.453
Time:	22:34:24	**Pearson chi2:**	88.8
No. Iterations:	4	**Covariance Type:**	nonrobust

	Coefficient	Absolute values, e^coefficient	P-value
pdb2018trv4_us_csv_Med_HHD_Inc_	0.000012	1.000012	0.070188
pdb2018trv4_us_csv_pct_Males_AC	0.006599	1.006621	0.706949
pdb2018trv4_us_csv_pct_Pop_18_2	0.008771	1.008810	0.311391
pdb2018trv4_us_csv_pct_Pop_25_4	-0.001459	0.998542	0.937167
pdb2018trv4_us_csv_pct_College_	-0.025139	0.975174	0.005870
HospitalsGeneralCount2mi	-0.050671	0.950591	0.707555
HospitalsPsychCount2mi	0.069240	1.071694	0.745290
SurgicalCentersCount2mi	-0.015776	0.984347	0.925388
UrgentCareCount2mi	-0.098587	0.906117	0.541792
WomensClinicsCount2mi	-0.229649	0.794813	0.376651
PainManagementCount2mi	-0.060735	0.941073	0.761052
PhysicalTherapyCount2mi	0.087568	1.091516	0.586368
MentalFacilitiesCount2mi	-0.074678	0.928042	0.655207
SubstanceAbuseCount2mi	0.067339	1.069658	0.614647
OtherMedicalFacilitesCount2mi	-0.048747	0.952422	0.271095
NursingAssistedLiveCount2mi	-0.003253	0.996752	0.938702
HospiceCount2mi	0.132250	1.141393	0.454148
ChildFacilitiesCount2mi	-0.036940	0.963734	0.446834
PharmCount2mi	0.044927	1.045951	0.244065
DrugDropCount2mi	0.278071	1.320580	0.001419
NaloxoneDistribCount2mi	-0.090861	0.913145	0.209548

4. Logistic Regression Results for Cincinnati (Tempe does not converge)

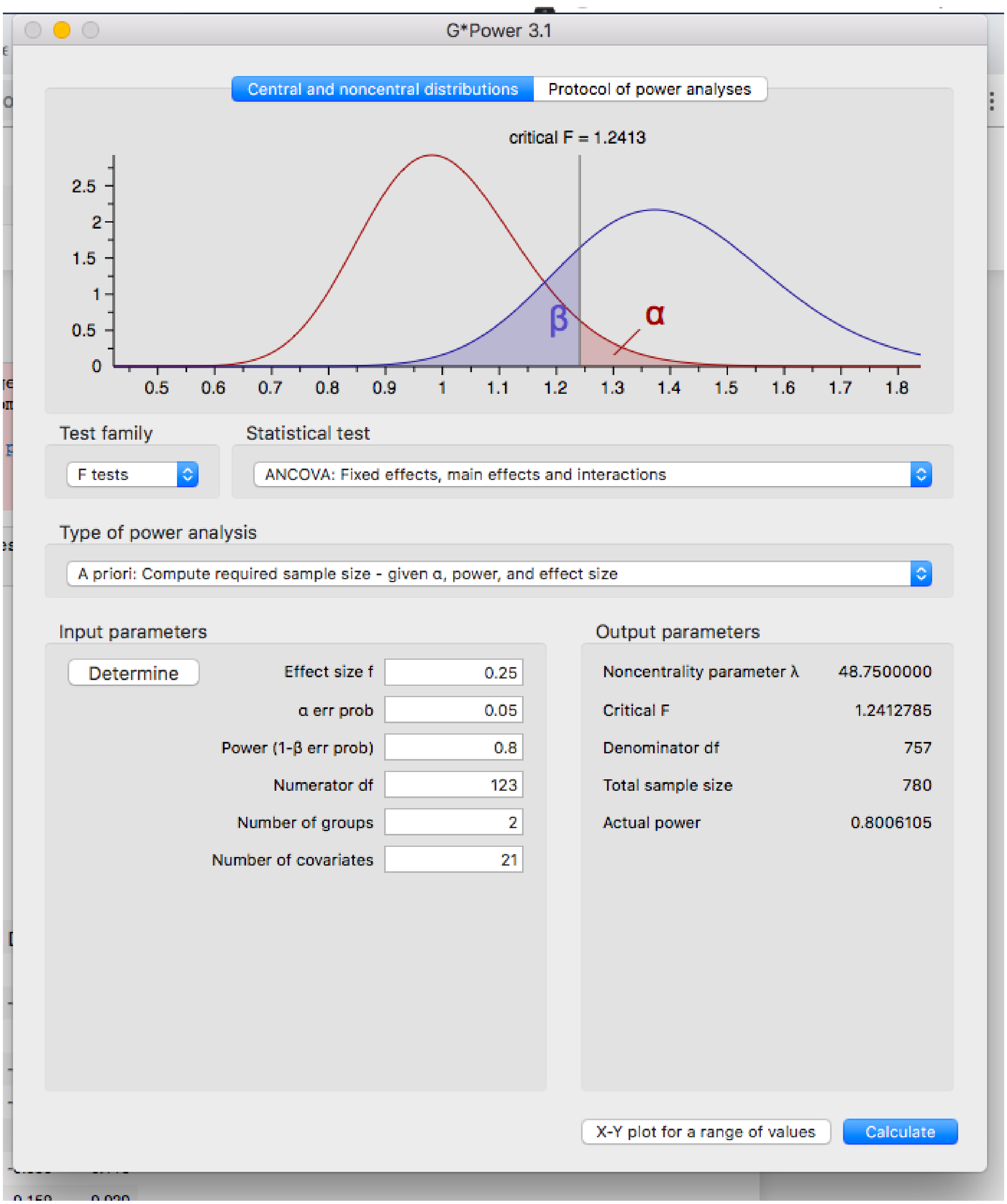
Logit Regression Results

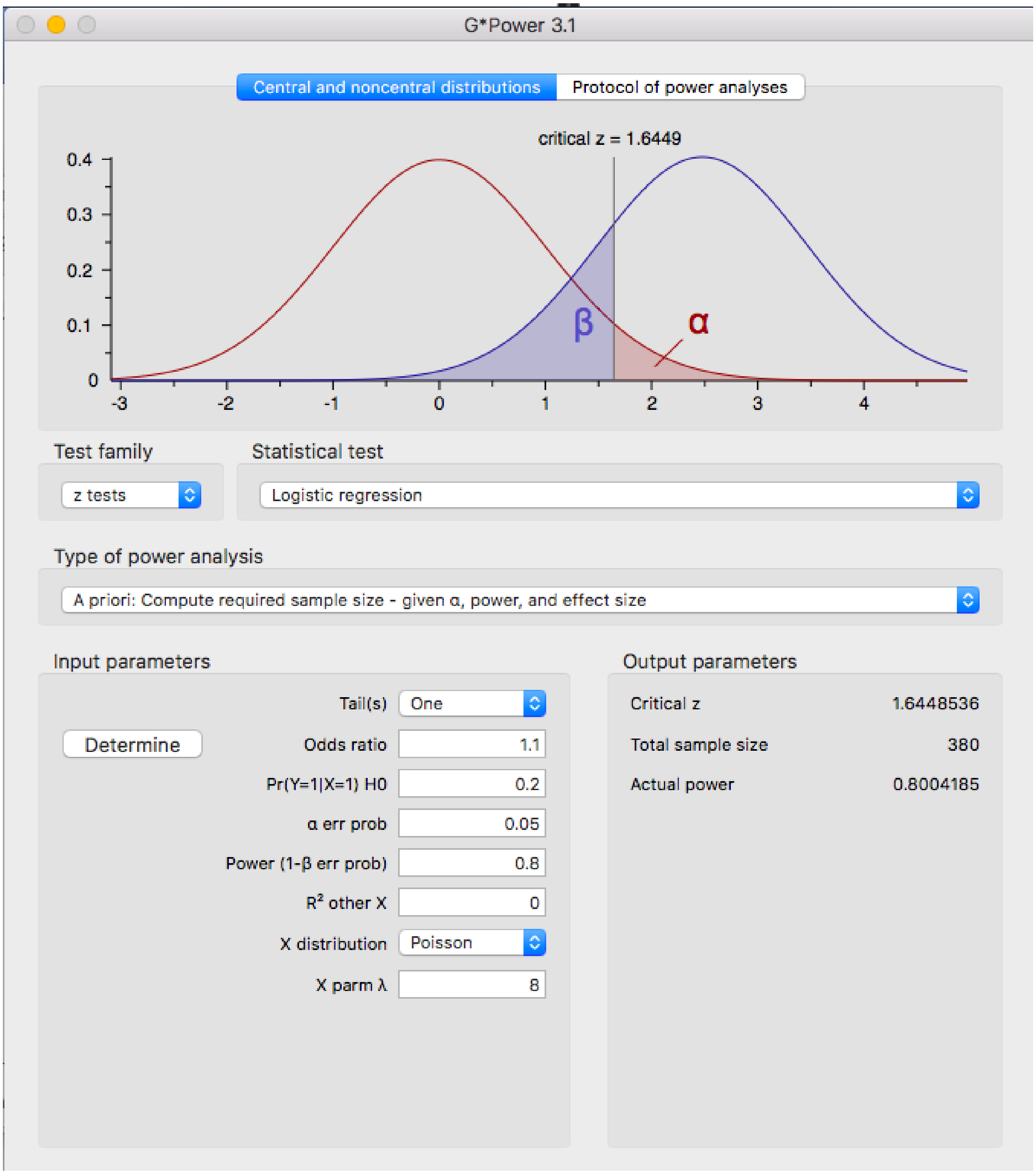
Dep. Variable:	ems_flag	**No. Observations:**	1238
Model:	Logit	**Df Residuals:**	1219
Method:	MLE	**Df Model:**	18
Date:	Wed, 11 Dec 2019	**Pseudo R-squ.:**	0.1823
Time:	22:47:07	**Log-Likelihood:**	-697.36
converged:	True	**LL-Null:**	-852.86
		LLR p-value:	2.624e-55

77]:

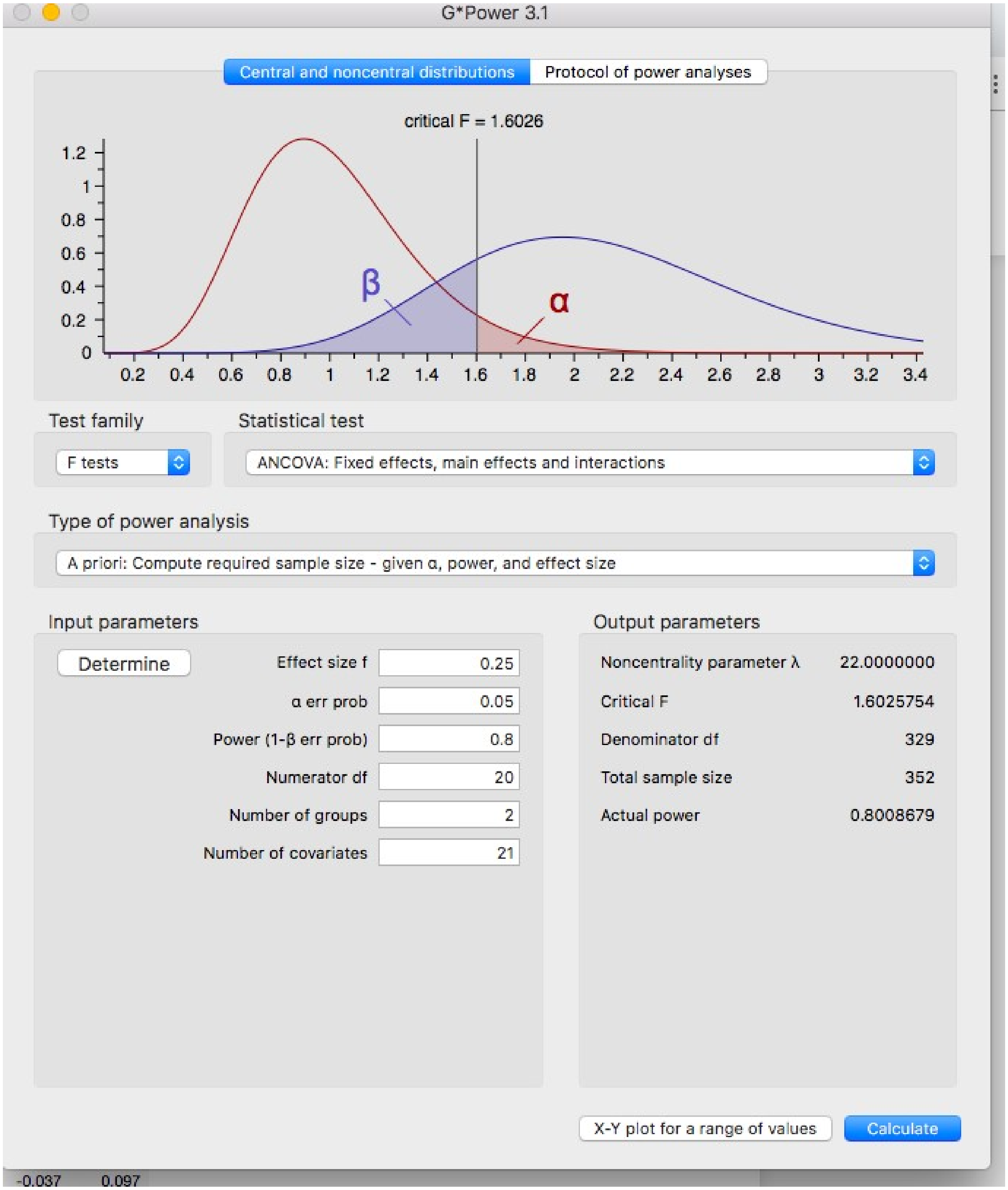
	Coefficient	Absolute values, e^coefficient	P-value
HospitalsGeneralCount2mi	0.601669	1.825163	2.255290e-06
HospitalsPsychCount2mi	-0.544081	0.580375	2.942064e-02
SurgicalCentersCount2mi	0.211552	1.235594	4.171637e-03
UrgentCareCount2mi	-0.917611	0.399472	4.665740e-14
WomensClinicsCount2mi	-0.241997	0.785059	1.677478e-01
PainManagementCount2mi	1.157386	3.181606	1.709394e-08
SubstanceAbuseCount2mi	-0.334848	0.715447	5.700931e-03
OtherMedicalFacilitesCount2mi	0.038051	1.038784	2.785218e-01
NursingAssistedLiveCount2mi	0.168246	1.183228	4.062118e-07
HospiceCount2mi	-0.151569	0.859359	3.350343e-01
HospitalsChildCount2mi	-0.458197	0.632423	5.919144e-03
ChildFacilitiesCount2mi	0.269813	1.309719	2.548730e-01
PharmCount2mi	0.056220	1.057830	3.476478e-02
DrugDropCount2mi	-0.570767	0.565092	4.464126e-06
NaloxoneDistribCount2mi	0.018019	1.018182	4.242165e-01
pct_College_ACS_13_17	-0.018148	0.982016	1.269494e-04
pct_Males_ACS_13_17	0.010914	1.010974	1.367741e-01
pct_Pop_18_24_ACS_13_17	-0.029436	0.970993	1.248845e-04
pct_Pop_25_44_ACS_13_17	-0.003563	0.996443	7.554182e-01

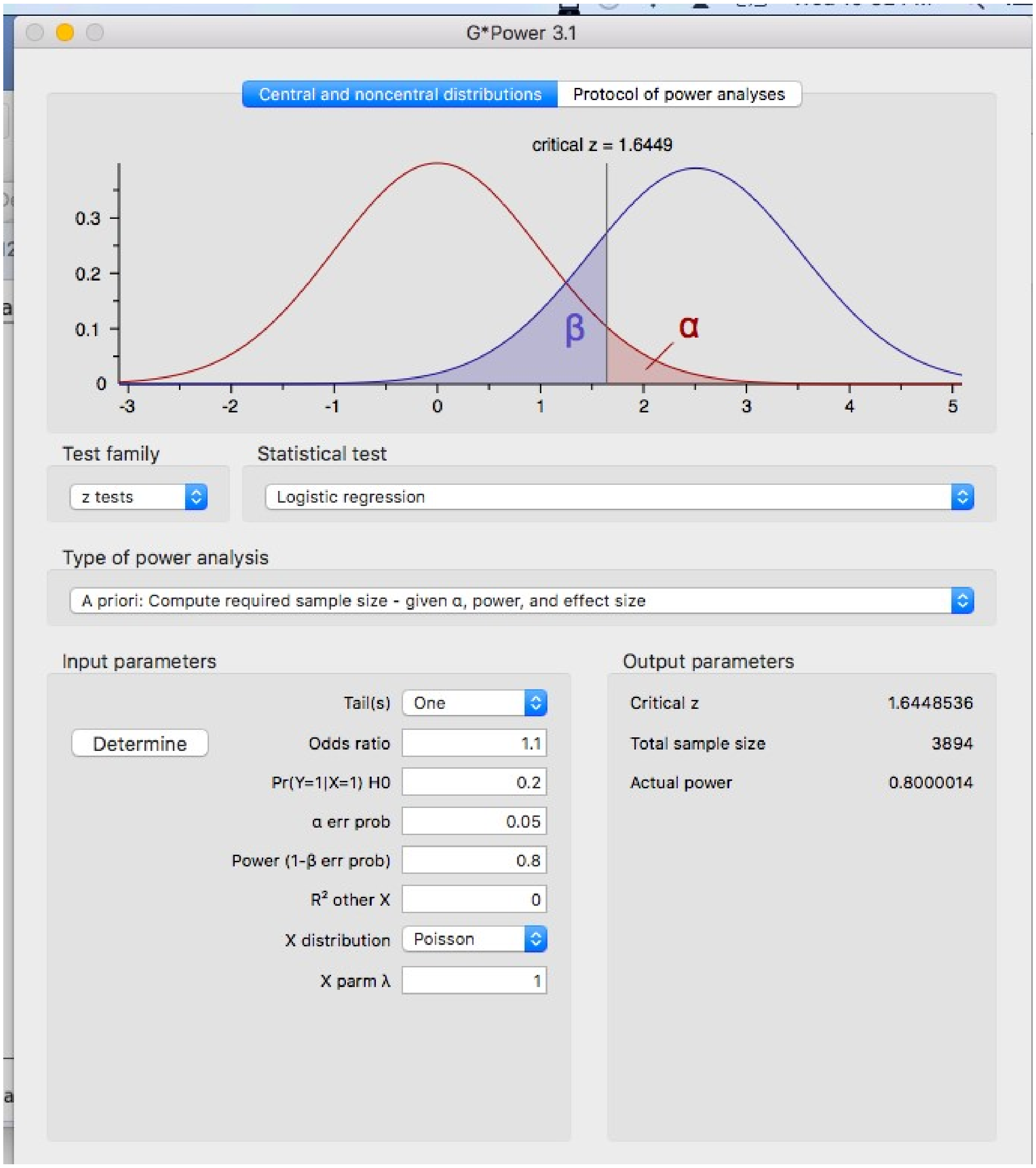
5. Power Analyses for Cincinnati – Poisson Regression (Top) and Logistic Regression (Bottom)




G*Power 3.1
Central and noncentral distributions
Protocol of power analyses
critical z = 1.6449
0.4
0.3
0.2
0.1
0
-3
-2
-1
0
1
2
3
4
β
α
Test family
z tests
Statistical test
Logistic regression
Type of power analysis
A priori: Compute required sample size - given α, power, and effect size
Input parameters
Tail(s) One
Determine
Odds ratio 1.1
Pr(Y=1|X=1) H0 0.2
α err prob 0.05
Power (1-β err prob) 0.8
R² other X 0
X distribution Poisson
X parm λ 8
Output parameters
Critical z 1.6448536
Total sample size 380
Actual power 0.8004185

6. Power Analyses for Tempe – Poisson Regression (Top) and Logistic Regression (Bottom)




G*Power 3.1
Central and noncentral distributions
Protocol of power analyses
critical z = 1.6449
0.3
0.2
0.1
0
-3
-2
-1
0
1
2
3
4
5
β
α
Test family
z tests
Statistical test
Logistic regression
Type of power analysis
A priori: Compute required sample size - given α, power, and effect size
Input parameters
Tail(s) One
Determine
Odds ratio 1.1
Pr(Y=1|X=1) H0 0.2
α err prob 0.05
Power (1-β err prob) 0.8
R² other X 0
X distribution Poisson
X parm λ 1
Output parameters
Critical z 1.6448536
Total sample size 3894
Actual power 0.8000014