A proposed application of Bayesian Disease Mapping to Psychosocial/Medical Assessment

Robert A. Yaffee, Ph.D.

NYU Silver
Silver School of Social Work

Presentation to the Academy of Labor and Social Relations
Federation of Trade Unions of Ukraine
Kiev, Ukraine
October 6, 2008

Importance of spatial analysis

- When psychological symptoms exhibit geographical distributions, a study of the spatial distributions may be crucial to understanding the nature of an important problem.

- Real or perceived exposure to toxic disaster may lead to a distribution of relative risk of particular psychosocial problems. The study of the spatial manifestations of these symptoms can help policy planners design mental health programs to help the people afflicted. Spatial analysis of diseases may be informative and important for epidemiologists; distribution of chronically disabled and the mentally ill for psychologists; crime hot spot analysis for criminologists; areas of substance abuse for social workers; resource deposits for miners, prospectors and investors; political affiliation or alignment for political scientists, as well as pollution concentration and post-disaster sequelae analysis for public health policy planners .

- Spatial patterns of post traumatic stress syndrome may reflect geographical proximity to an event or inter-area distance differences in variables under consideration.
- Distribution patterns may be random, regular, or clustered.
- Therefore, thematic mapping of local case counts, dependency (spatial autocorrelation) patterns or incidence clusters of symptoms may reveal important relationships that explain real and/or perceived risk of exposure to a major toxic disaster.

Objectives

- Primary Objective: We propose an examination of spatial patterns of physical and psychological symptomatology, along with their interactions. For these consideration have for the most part been ignored.

- More specifically: We intend to examine counts per unit area of post-traumatic stress, depression, anxiety, other psychological symptoms potentially related to Chornobyl radiation exposure and relevant lifestyle factors. We will endeavor to examine the counts per unit area of reported physical illnesses there as well. We will look for interactions between them. *We sincerely hope the findings of our research can substantially help all concerned.*

Objectives—cont'd.

- We expect to find patterns of spatial dependence (correlation and/or autocorrelation) between observations at different locations, derived from the relative position (distance or arrangement) of observations in geographic space (Anselin, Spatial Econometrics, p.1).
- We expect to observe meaningful clustering (perhaps even hot spots) among these patterns of dependence of psycho-social symptomatology, radiation absorption and exposure that will inform our understanding of the nature of the problem.

Outline

1. We define our key organizing concepts to examine marked point patterns of psychosocial symptomalogical incidence.
    1. Population distribution
    2. Disease count
    3. Incidence rates
    4. Expected counts
    5. Standardized morbidity/mortality rates
    6. Relative risk
2. We explain problems with classical frequentist application of the normal distribution for this analysis.
    1. Nonnormal distribution or counts or rare events

Outline – cont'd.

3. The Basic Poisson Model for counts or rare events
    1. We will determine patterns and levels of random distribution
    2. Assumptions of the model
    3. The problem of overdispersion
4. Bayesian analysis and the Negative binomial Model
    1. Essential Bayesian Analysis
    2. The negative binomial (Poisson-gamma) model
    3. The estimation process
    4. Cluster and Hot spot detection
    5. Goodness of fit
    6. Residual Analysis
    7. A thematic or choropleth map

Population distribution

- Suppose our unit of geographical subdivision is the raion.
- From the latest census, we obtain the population count for the raion.

$$population\ density = population / area = \sum_{i=1}^{N} n_i$$

where

$n =$ *population*

$i =$ *raion or unit of geographical area*

A disease or symptom count

- We know that the incidence of a disease or psychopathogical symptom is measured in the case (as distinguished from control) count of disease.
- We examine the disease count for the unit of geographical area.

$$total\ \ disease\ \ count\ \ per\ raion = \sum_{i=1}^{I} O_i$$

$$where$$

$$i = raion$$

$$O_i = observed\ cases\ of\ \ disease\ \ per\ \ raion$$

A Crude measure of disease distribution is the Incidence rate

$$Incidence\ rate\ (IR) = \frac{O_i}{n_i}$$

$where$

$O_j = observed\ cases\ in\ raion\ i$

$n_i = number\ of\ population\ in\ raion\ i$

An incident rate is a measure of absolute risk, while the ratio of two incident rates is a measure of relative risk. If we took the ratio of the incident rate of a raion and divided it by the incident rate of the oblast, we would have the relative risk.

Spatially Random disease distribution provides the expected count (e_i)

- We obtain the incidence rate for each raion.
- We obtain the average, mode, and median disease count in each raion.

$$mean(IR)\ for\ study\ area$$

$$e_i = \frac{\sum_{i=1}^{N} O_i}{\sum_{i=1}^{N} n_i}$$

where

$e_i = mean\ incidence\ rate\ of\ country$

We can form confidence intervals, if we assume normality, around these means.

Alternatively, we can obtain confidence intervals from bootstrapping the local empirical standard errors within each raion.

Other sources of Expected counts

- May come from National Centers for health statistics
- May come from National Health Departments or Centers for Disease Control databases.
- May come from actuarial tables
- May come from national counts stratifed by age and gender.

Standard Morbidity/Mortality rate (SMR)

$$\hat{\theta} = SMR_i = \frac{o_i}{e_i}$$

where

$\hat{\theta}_i$ = estimated relative risk

i = geographical area unit (raion)

e_i = expected count in raion

o_i = observed count in raion

Relative risk

- This ratio of observed to expected counts within a unit of area describes the relative risk of the geographical unit. This is the odds of being a case over the odds of the background case rate.

$$Relative\ \ Risk\ \ within\ area\ i = \hat{\theta}_i = \frac{o_i}{e_i}$$

where

$e_i = expected\ \ normal\ or\ \ background\ \ incidence\ rate$

$o_i = observed\ \ incidence\ rate\ in\ \ geographic\ unit\ \ of\ \ area$

SMR issues

- Ratio estimators can become inflated with small denominators of expected counts.
- This can give rise to noise and outliers.
- Smoothing can be done with kernel regression (Nadaraya-Watson) or generalized additive models (if covariates come in handy to facilitate smoothing—Kelsall and Diggle).

Basic Poisson Model for Counts or Rare events

$$o_i \sim Poisson(e_i\theta_i)$$

where

$$\theta_i = relative\ risk\ \left(odds\ ratio\ of\ observed\ \ IR\ to\ expected\ \ IR\right)$$

$$\ln(\theta) = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$$

$$\theta = \exp(\beta_0 + \beta_1 x_1 + ... + \beta_p x_p)$$

$$\theta_i = \frac{o_i}{e_i}$$

Poisson probability

The

$$probability\ of\ \ observed$$

$$count\ o\ in\ i^{th}\ area\ unit$$

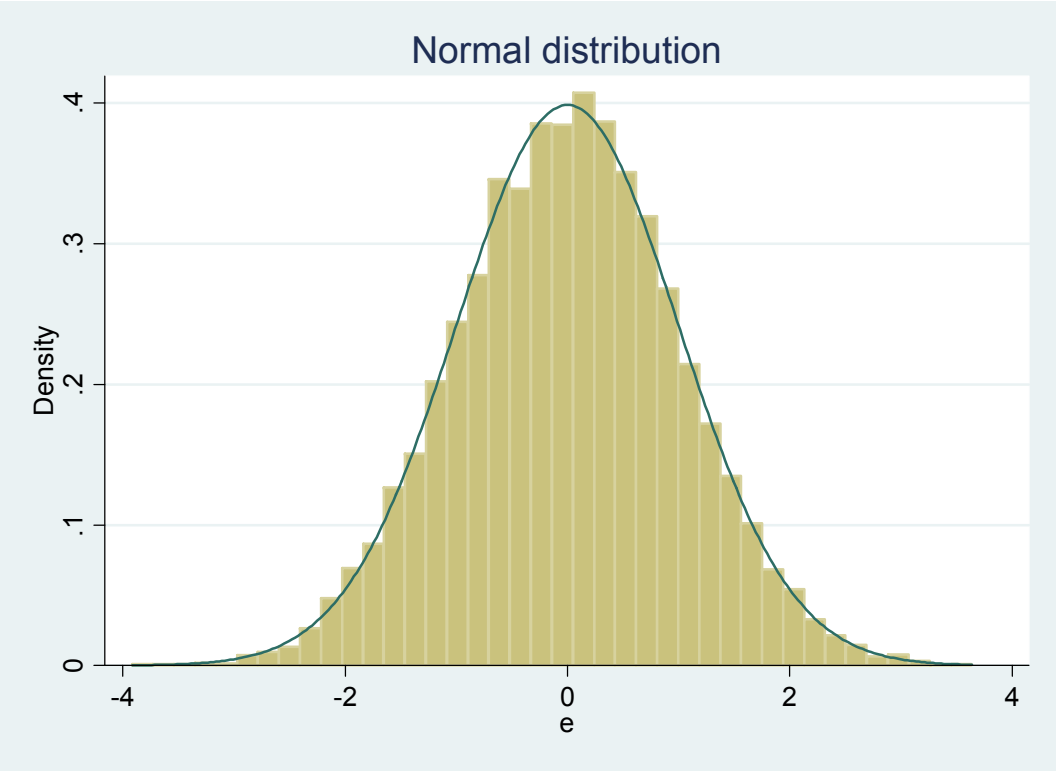
$$= \frac{e^{-\theta_i e_i}(\theta_i e_i)^o}{o!}$$

$$assuming\ that\ mean = variance,$$

$$and\ \theta(the\ relative\ risk)\ \ is\ fixed.\ By\ using\ \theta_i\ \ we\ can$$

$$allow\ \theta_i\ \ to\ \mathrm{var}\ y\ \ o\ \ between\ \ raions.$$

Inappropriateness of the Normal Distribution



But psychosocial abnormalities are not distributed as such. They tend to have long tails on the right (high positive skew). The use of such a distribution will lead to inaccurate assessment and inflated standard errors.

Model assumptions

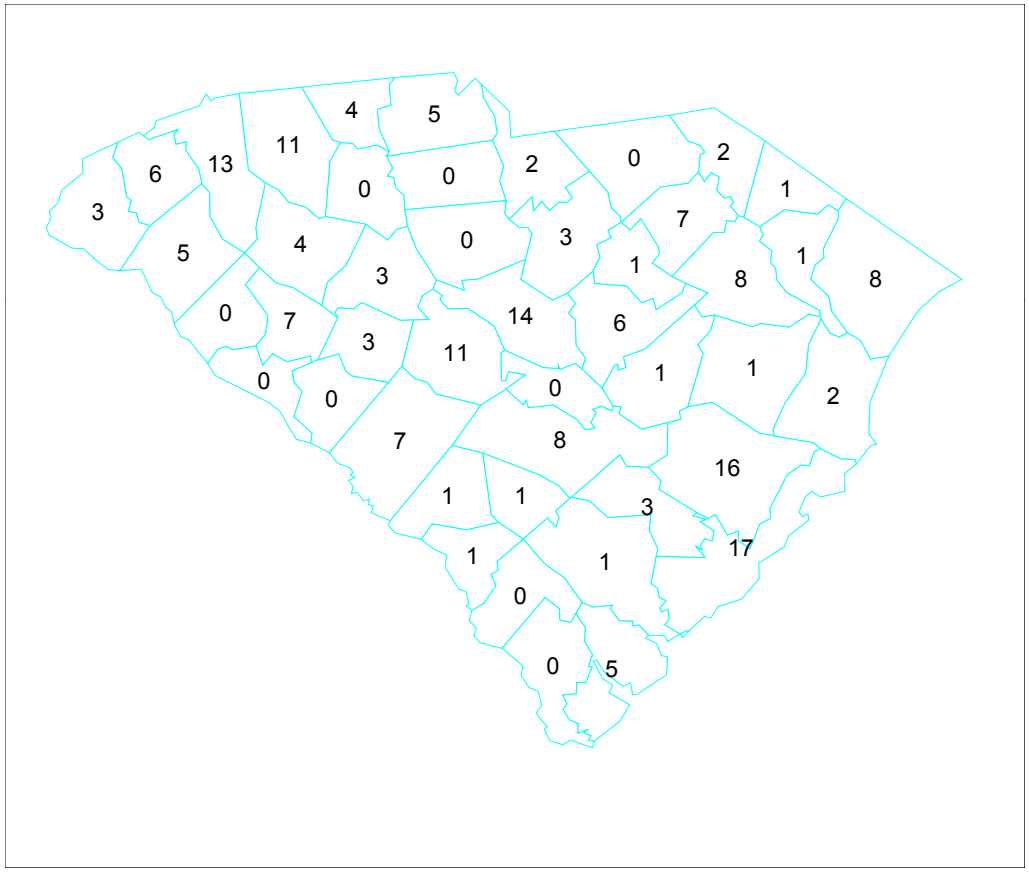
(Lawson et al, 2003).

- Individuals within the study population behave independently with regard to propensity to disease, after controlling for relative and confounding variables.
- The at-risk background intensity is has a continuous and random spatial distribution within a specified background.
- Case-events are occur as singly spatially separate phenomena.

A Marked Point Process

- Each case is geo-coded, in that a latitude and longitude, for the location of residence, is recorded in the dataset as a marked point on a map. In this way, the different symptomatology can be graphed on a map.
- The count of cases is a measure of a count. Counts are generally approximated with the use of a poisson distribution.
- Poisson processes can also approximate rare case counts-- when the area is large and the case counts are small.

Congenital abnormality deaths in counties of South Carolina in1990



Andrew Lawson lecture Polygon shape file provided by A. Lawson.

The Poisson process for counts and rare events

- The number of cases of a disease can be modeled as poisson process.

$$o_i = Poisson(\theta_i e_i)$$

$$where$$

$$o_i = observed\ \ count\ \ for\ \ area\ \ unit\ \ i$$

$$\theta_i = relative\ risk(ratio\ of\ incidence\ rates)$$

$$e_i = expected\ count\ for\ area\ unit\ i$$

Poisson Fixed effects model

$$\theta_i = \exp(b_0 + b_1 x_{1i} + b_2 x_{2i} + e_i)$$
$$= \exp(X_i B + e_i)$$
$$\ln(\theta_i) = b_0 + b_1 x_{1i} + b_2 x_{2i} + e_i$$

- Where
- Θ_i= relative risk of geographical area i
- b_0 = common level of rate over study region from i=1 to m (maximum number of geographical area units under study)
- x_{1i}=easting (longitude)of area unit centroid
- x_{2i}=northing (latitude) of area unit centroid
- X_i= mxp matrix of covariates
- B= px1 parameter estimate vector

Assumptions of the Poisson model

- The mean= variance
- That all effects included in the model determine the excess risk surface.

The Problem of Overdispersion

- When the variance > mean, overdispersion occurs.
- Overdispersion may come from clustering of the counts at a particular scale.
- This may come from sparse data and low expectation per unit of area.
- This may come from variations in frailty of the individuals—that is, susceptibility to the disease or symptom under consideration.
- Violation of the dispersion assumption invalidates use of the Poisson model.
- We relax the assumption of the mean=variance with Bayesian analysis.

Why Bayesian Analysis?

- Bayesian analysis tells us how to update prior beliefs in light of evidence (Jackman, S. Yale, 4/23/04)
- Any model estimable by maximum likelihood can be estimated by Bayesian simulation .
- Bayesian simulation permits us to handle models and data too difficult for maximum likelihood.
- Data sets with nonnormal data (using other distributions to model them)
- We get more accurate estimates with smaller confidence intervals (called credible intervals by Bayesians).

Bayes's Rule of Conditional Probability

Consider parameter θ *and data* y.

$$P(\theta \mid y) = \frac{P(\theta, y)}{P(y)}$$

Therefore, $P(\theta \mid y)P(y) = P(\theta, y)$

This can be re $-$ exp *ressed as*

$$P(\theta, y) = p(\theta)\, p(y \mid \theta)$$

where $p(\theta) =$ *prior distribution*

and $p(y \mid \theta) =$ *sampling* (*data*) *distribution*

The sampling distribution is called the likelihood

$$So \quad P(\theta \mid y) = \frac{P(\theta \mid y)P(y)}{P(y)}$$

$$Similarly, \quad p(y \mid \theta) \propto p(\theta)\, p(y \mid \theta)$$

$$f(y \mid \theta) \propto f(\theta)\, f(y \mid \theta)$$

Bayesian Analysis

- A posterior (predictive distribution) is proportional to a prior distribution times a likelihood

$$Posterior \propto Prior * Likelihood$$

$$p(\theta | data) \propto p(\theta) * f(data | \theta)$$

How do we combine the prior with the data

- We pool them by forming a weighted average
- Suppose the posterior can be re-arranged to yield

$$\theta \mid y \sim N(\mu, \tau^2)$$

where

$$\mu = \frac{\frac{1}{\tau_0^2}\mu + \frac{n}{\sigma^2}\bar{y}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}} \quad \textit{is a weighted average}$$

The averaging is weighted by the precision of the samples

$$\tau^{-2}\, total\ \ precision$$

$$= \ \tau_0^{-2}\ \ prior\ \ precision + \frac{n}{\sigma^2}\ data\ \ precision$$

$$because \quad Var(\theta \mid y) = \frac{1}{\frac{1}{\sigma^2} + \frac{1}{\tau_0^2}}$$

Guido Imben's 2007 Summer Institute Lecture on Bayesian Inference at the National Bureau of Economic Research in the "What's new in Econometrics" series.

Bayesian combination of normal prior and normal likelihood

$$\theta \sim N(\mu, \tau^2) \quad prior$$

$$Y \mid \theta \sim N(\theta, \sigma^2) \quad likelihood(data \mid prior)$$

$$The \ \ posterior \ \ distribution \ \ has \ \ mean \ \ and \ \ variance:$$

$$E(\theta \mid Y) = \alpha\mu + (1-\alpha)Y$$

$$Var(\theta \mid Y) = (1-\alpha)\sigma^2$$

$$where$$

$$\alpha = \frac{\sigma^2}{\sigma^2 + \tau^2}$$

Bayesian shrinkage

- We postulate that $0 \leq \alpha \leq 1$.
- The posterior mean is a weighted average of the prior mean, μ, and the direct estimate Y; the posterior estimate is shrunk toward the posterior mean.
- The α weight of the prior mean, μ, depends on the relative variability of the prior distribution and likelihood.
- If σ^2 is large relative to the prior variance, τ^2, (our prior knowledge is more precise than the data variance), then B is close to 1 producing substantial shrinkage.
- If σ^2 is small relative to prior variance, τ^2, then B is close to 0 and the estimate is moved very little toward the prior mean.

Combining the Prior and the data to get the Posterior

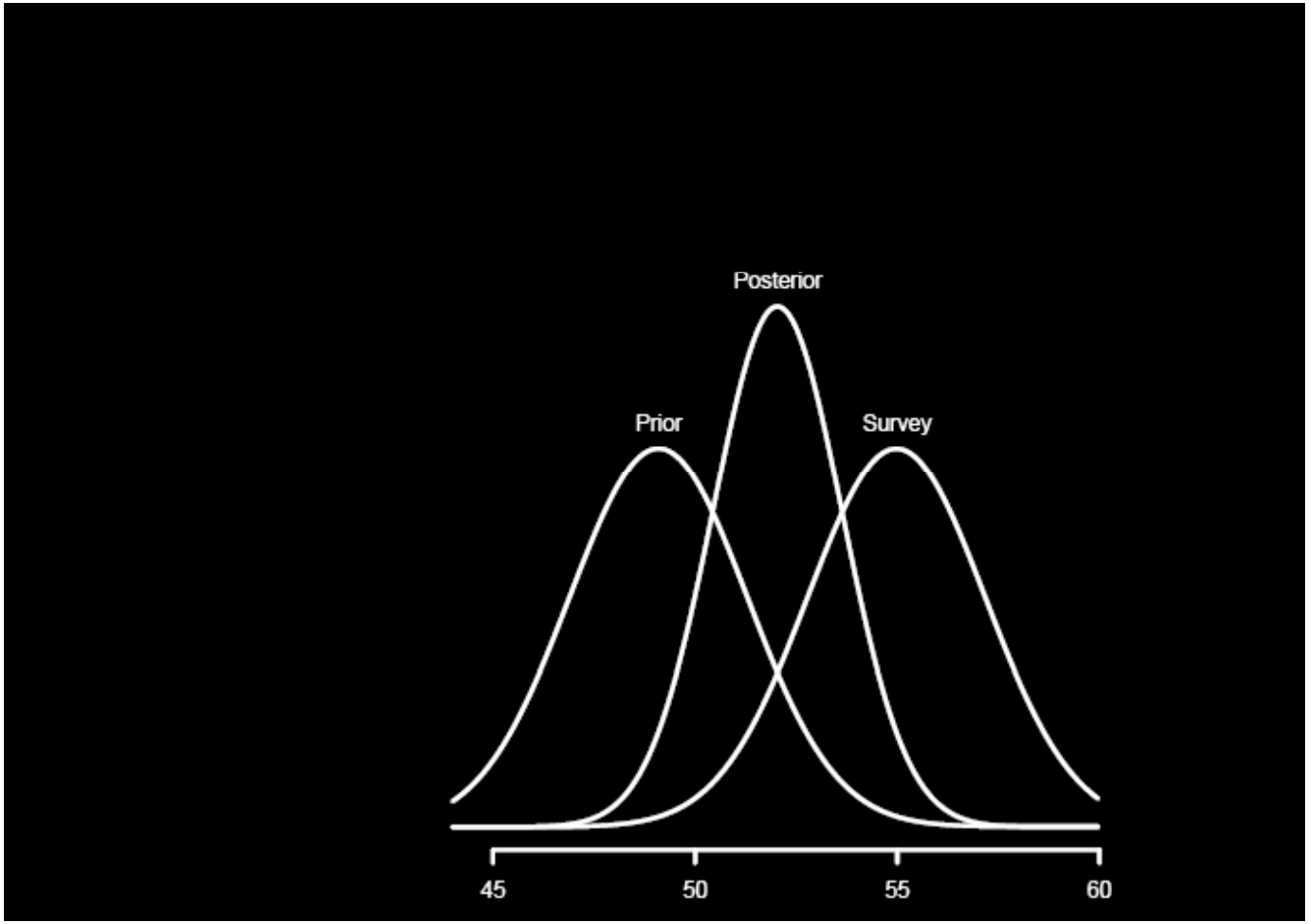


Figure from Simon Jackman, Yale Lecture, Day 1.

Bayesian Analysis

- As priors become vague, they become noninformative, so

$$p(\theta) -> c(a\ constant)$$

We say that the priors become diffuse, flat, or noninformative. As this happens Bayesian analysis yields the same results as maximum likelihood analysis, except that Bayesian analysis can provide answers when ml cannot.

Inappropriateness of the Closed-form Normal Distribution for abnormal symptoms

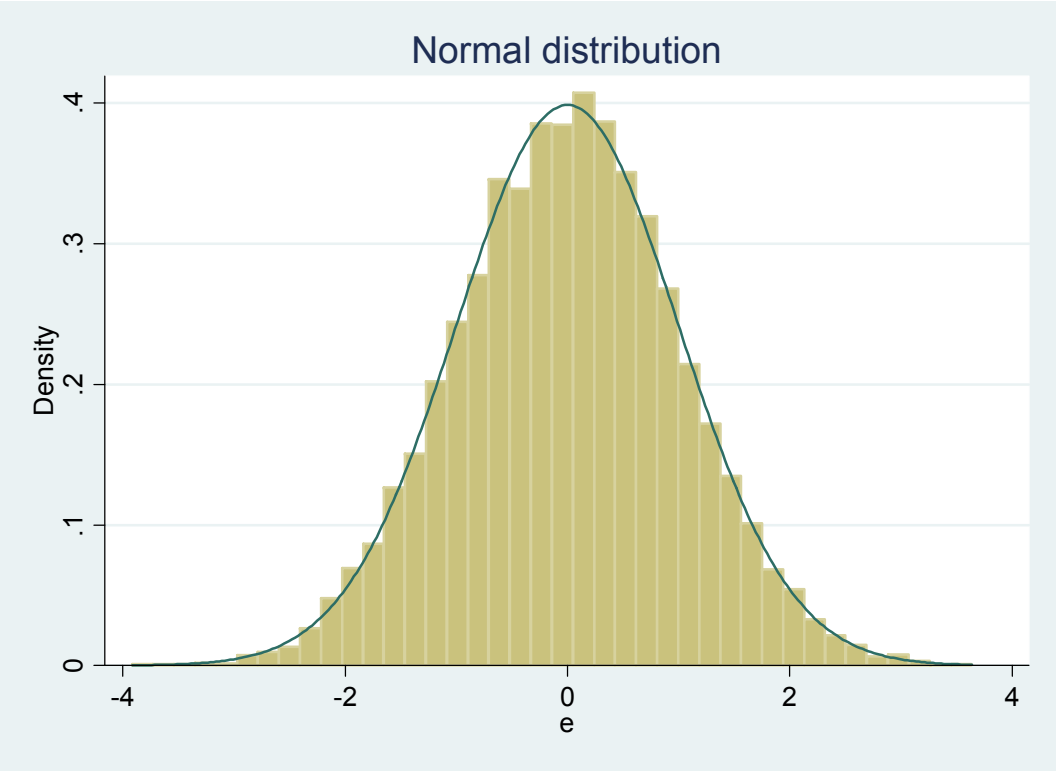
- We specify a normal likelihood as

$$f(y \mid \theta, \sigma^2) = \prod_{i=1}^{n} \phi(y_i \mid \theta, \sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i-\theta)^2}{2\sigma^2}\right)$$

$$\propto \exp\left(\frac{-\sum_{i=1}^{n}(y_i-\theta)^2}{2\sigma^2}\right)$$

A normal distribution does not reflect the presence of psychopathology



We use a gamma(Γ) distribution as a prior distribution

$$Gamma(\alpha,\beta)\ \ distribution:$$

$$p(\theta)=\frac{\beta^{\alpha}\theta^{\alpha-1}e^{-\theta\beta}}{\Gamma(\alpha)},\quad \theta>0,\alpha>0,\beta>0$$

$$where$$

$$E(\theta)=\alpha/\beta$$

$$\mathrm{var}(\theta)=\alpha/\beta^{2}$$

$$\Gamma=(n-1)!$$

By multiplying the gamma by the Poisson likelihood, we obtain a Negative Binomial.

- Use of the negative binomial distribution allows us to relax the assumption that the mean must equal the variance, which is an unrealistic assumption.
- With a negative binomial distribution, we have a more flexible model.

Marginalized Log-Likelihood of the Negative Binomial model

- From the marginalized LL :

$$L(\alpha,\beta)=\sum_i\left[\ln\frac{\Gamma(o_i+\alpha)}{\Gamma(\alpha)}+\beta\ln(\alpha)-(o_i+\alpha)\ln(e_i+\alpha)\right]$$

Clayton and Kaldor(1987) Lawson(2008)

$$E_{poisson}(O_i) = \mu_i = e_i\theta_i$$

and

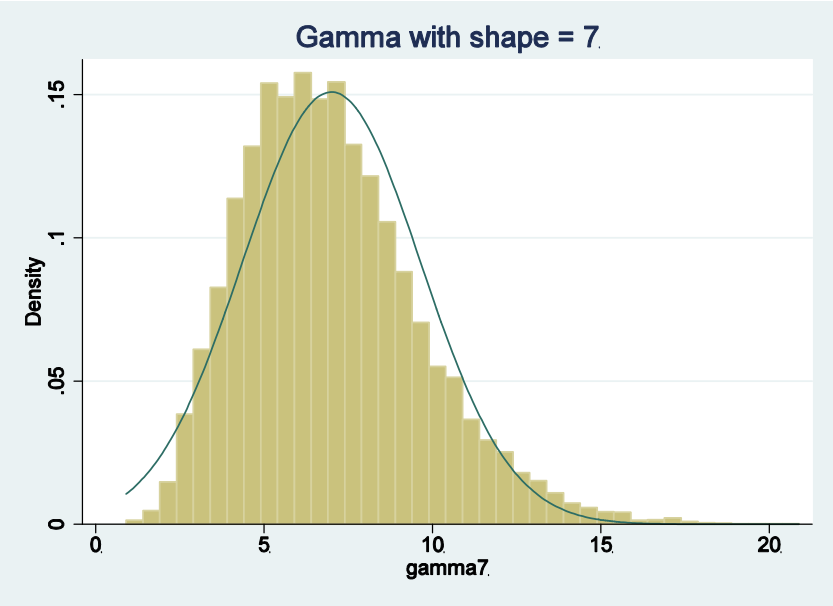
$$\theta_i \sim \Gamma(\alpha, \beta), \ then$$

$$\theta_i = \frac{1}{m}\sum_i \frac{o_i + \alpha_i}{e_i + \beta}$$

where

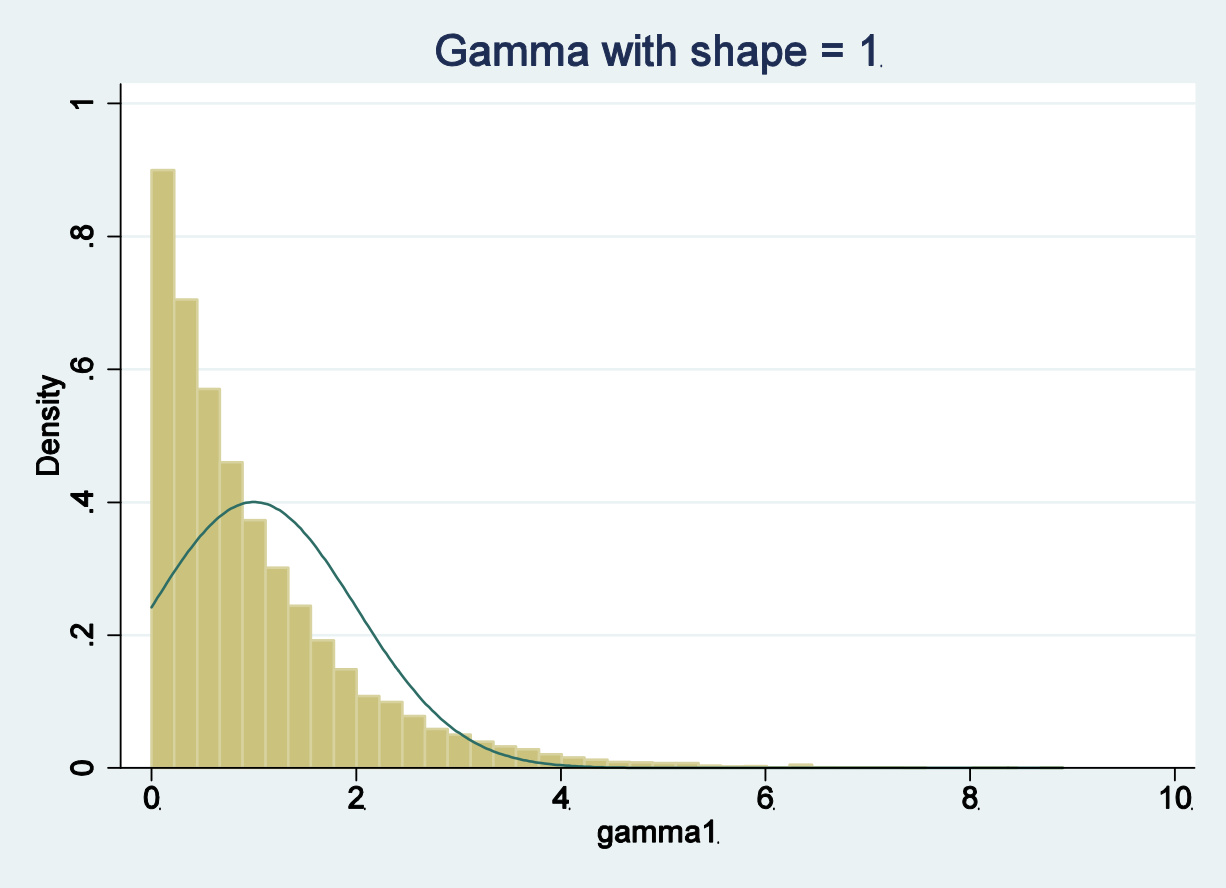
$$m = sample\ size$$

Flexible Prior
which we can tune by changing a shape or location parameter



We can parameterize the Gamma distribution as a 3 parameter distribution a = shape, b = location, and g = scale
Gamma(a, (x – b)/g) Merely, by varying the a, we can control the shape of the Gamma distribution.

We would like a flexible prior distribution



Conjugacy

Prior	***Likelihood***	***Posterior***
Normal	Normal	Normal
Beta	Binomial	Beta
Gamma	Poisson	Poisson
Wishart	Multivariate normal	Wishart
Exponential	Poisson	Poisson

Bayesian Analysis of a Negative Binomial process

Random effects have a prior distribution –such as a ($\Gamma(\alpha,\beta)$) distribution. The α and β parameters are parameters of the prior distribution. Then there are the data. By a mixture of the prior distribution with the distribution of the sample (which is a multiplication if parameters are logged, as they are in a Poisson process), then we obtain a posterior distribution

Noninformative Prior distributions

A noninformative prior is may be a uniform or flat prior.

A Jeffrey's prior is distributed as a Beta(.5,.5) and may be used as a noninformative prior.

It does not favor one parameter value over any other.

Even if such a prior is used, there will still be shrinkage of the confidence intervals.

Hence, Bayesian analysis generally yields smaller credible intervals than classical frequentist confidence intervals.

Sufficient statistics are computed

Sufficient statistics are those which together can define a distribution. For a normal distribution, sufficient statistics are the mean and the variance.

From the mean and variance, which are the same in the Poisson distribution, we can generate the distribution.

Estimation of the Bayesian Posterior

This is done by repeatedly resampling this combination to obtain he (probability mass function of) posterior distribution. By repeatedly resampling a large number of times (which is easy with fast computers) using a Gibbs sampler (Gelfand and Smith, 1990), we obtain a measure of the area under the posterior curve.

From this posterior simulation, we can easily obtain a mean and variance as well as any other sufficient statistics.

Resampling the area under the curve using GIBBS sampling with Winbugs.
Winbugs (Windows Bayesian estimation using Gibbs Sampling) by Spiegelhalter et al.(Imperial College of London) was used to generate the maps of congenital birth defects in South Carolina shown earlier.

Markov Chain Monte Carlo Algorithm for Bayesian Estimation

Gibbs sampler
(Gelfand and Smith, 1990)

"Sequentially sampling parameter values from a Markov chain whose stationary distribution is exactly the desired joint posterior distribution of interest(Carlin and Lewis, 120)."

Gibbs Sampler

Carlin and Lewis, 121-122.

Let our model contains k parameters $\theta = (\theta_1, ..., \theta_k)$

"Assume that samples can be generated from each of the complete conditional distributions"

$$\left\{ p(\theta_i \mid \theta_{j \neq i,} y) \quad for\ i = 1, ..., k \right\}\ in\ the\ model.$$

"Under mild [stationary] conditions, the collection of full conditional distributions uniquely determines the joint posterior distribution …"

$$p(\theta \mid y),\ and\ hence,$$

$$all\ marginal\ posterior$$

$$distributions,\ \ p(\theta_i \mid y_i),\ i = 1, ...k\ .$$

Gibbs Sampler -- cont'd

"Given an arbitrary set of starting values" $\left\{\theta_2^0,\ldots,\theta_k^0\right\},$

The algorithm proceeds such that

For (t=1,…, T), repeat:

$$Step\ 1: Draw\ \theta_1^t\ \ from\ \ p(\theta_1 \mid \theta_2^{t-1},\theta_3^{t-1},\ldots,\theta_k^{t-1},y)$$
$$Step\ 2: Draw\ \theta_2^t\ \ from\ \ p(\theta_2 \mid \theta_1^{t-1},\theta_3^{t-1},\ldots,\theta_k^{t-1},y)$$
$$\vdots$$
$$Step\ k: Draw\ \theta_k^t\ \ from\ \ p(\theta_k \mid \theta_1^{t-1},\theta_3^{t-1},\ldots,\theta_{k-1}^{t-1},y)$$

Gibbs Sampler-cont'd

Carlin and Lewis, 121-122

What is obtained converges in distribution to a draw of the true joint posterior distribution.

A histogram provides a simulated-consistent estimator of the marginal distribution.

In practice we run 3 to 5 chains of these processes simultaneously to assess sampler convergence, discarding all samples from the burn in period.

$$\hat{E}(\theta_i \mid y) = \frac{1}{m(T - t_{bi})} \sum_{j=1}^{T} \sum_{t=t_{bi+1}}^{T} \theta_{i,j}^{t}$$

$$where$$

$$m = number\ of\ chains$$

$$T = total\ number\ of\ obs$$

$$t_{bi} = number\ of\ obs\ in\ burn-in\ period$$

This reduces the autocorrelation bias in the estimation of the posterior mean and variance.

Bayesian updating

The process of obtaining a posterior distribution from a prior and a likelihood is called an Bayesian updating.

One updating may be for a former updating. This is called Sequential Updating.

In this case, the earlier distribution is called a hyper distribution with hyperparameters.

How do we identify psycho-social symptom clusters

- We can develop and intensity function of the incidence rate (per capita) in each raion.
- When the disease density exceeds the confidence intervals around the average in each raion, which would indicate random normal spatial heterogeneity, we are able to plot this and measure the disease clustering.
- We may plot ranges of the intensity function on a thematic or Choropleth (a shaded or color coded) map indicating gradations in the quantile of the spatially distributed variable under consideration.

Intensity Function

Intensity $\lambda(s) = number\ of\ events\ expected\ within\ a\ unit\ area\ s$

$\lambda(s)$ *integrates to an overall mean number of events per unit area*

$$\lambda(s) = \sum_{||s-s_i|| \le b} \frac{3}{\pi b^2}\left(1 - \frac{|| s - s_i ||^2}{b^2}\right)^2$$

where

$b = bandwidth\ size\ (distance)$

$s = s\ vector\ of\ grid\ of\ locations\ defining\ the\ study\ area$

Cluster detection

Any bounded area of elevated risk is a spatial cluster.

Any area displaying unusual or excess risk is called a "hot spot."

Only adjacent areas can be together called a cluster.

Sometimes correlated heterogeneity may be called a cluster.

In a $\ln(\Theta) = a + e(i)$, the a is the smooth part of the model, whereas the e(i) is the residual. Residuals may be decomposed into correlated (clustered) and uncorrelated. Analysis of the residuals may indicate spatial patterns of elevated correlation and risk, that are useful.

If the relative risk of the posterior distribution exceeds a critical threshold in a particular area, that could be deemed a cluster. (Lawson, 2008, Chapter 6).

Kriging

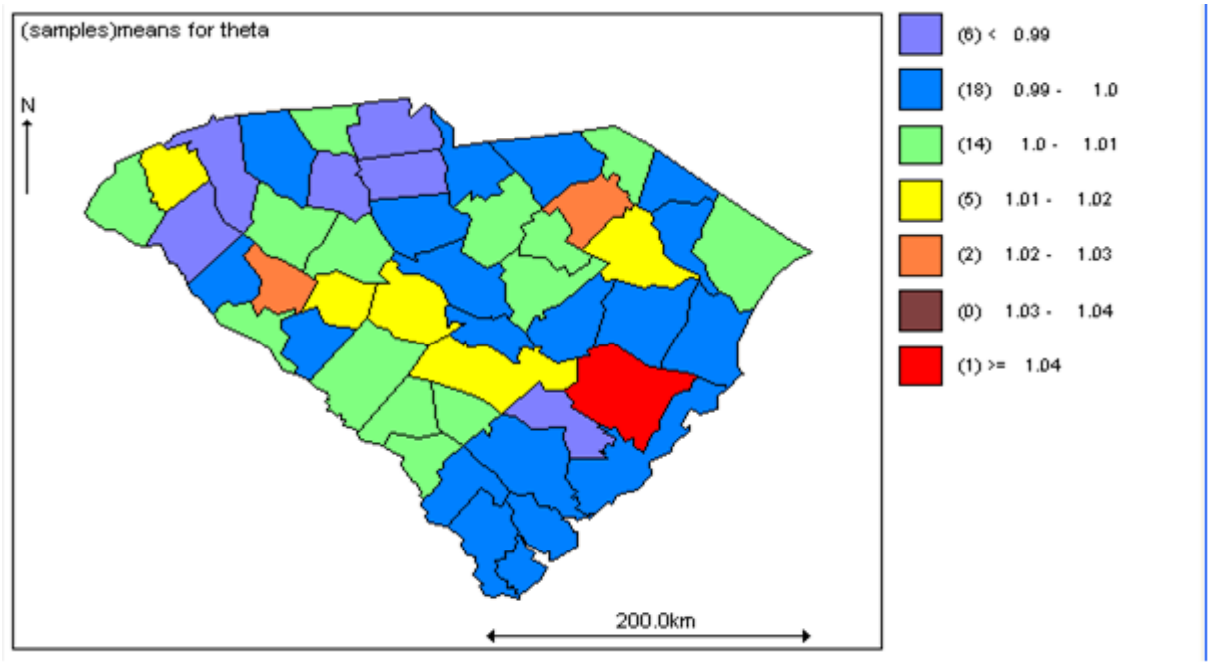
To map the area between the points, ordinary kriging, a form of spatial interpolation is performed.

Without going into a lot of detail, ordinary kriging is a weighted average of neighboring data to generate a linear prediction of an unknown data value.

Thematic (Choropleth) Map

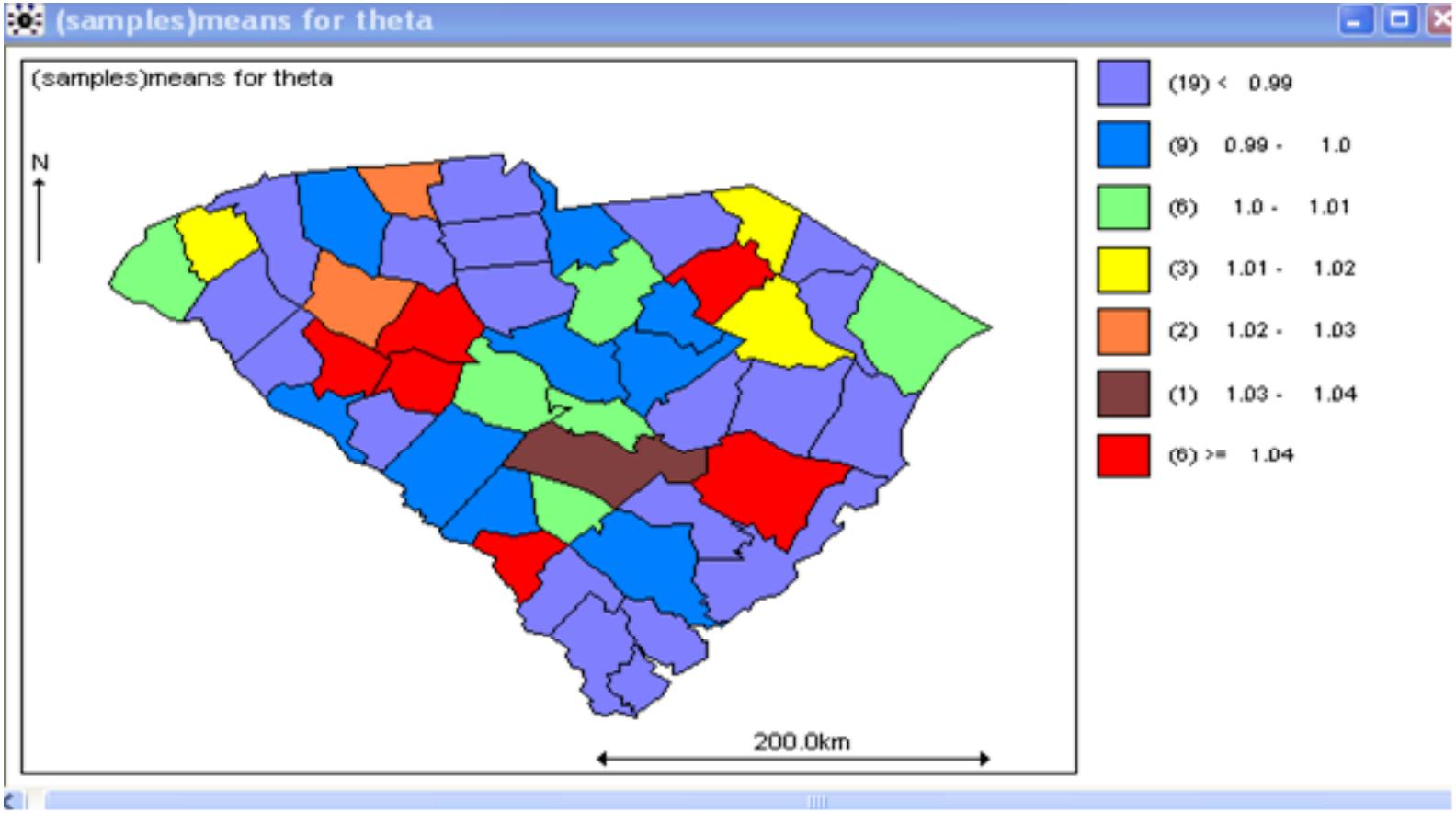
- These maps will reveal patterns of diffusion or spatial relationships. The intensity function may be divided according to quantiles and the different disease densities may be shaded or colored to represent them.
- Choropleth maps are thematic maps that display these gradations.
- See next two pages for an example.

Thematic Map
generated from a Normal Prior



There is no clear pattern here—only a spatial cluster in one county.

Choropleth map depicting levels of congenital abnormality deaths in terms of SMR in 1990. Map generated from a Gamma Prior.



Here we can see a pattern of higher risk in the central counties as well as some in the west .

Other spatial autoregression models

SAR: simultaneous autoregressive models

CAR: conditional autoregressive models

$$Y(s_i) = x(s_i)'\beta + \sum_{j=1}^{N} b_{ij}\varepsilon(s_j) + v(s_j)$$

$$b_{ij} = spatial\ \ dependence\ \ parameters$$

These models express the SMR as a function of a set of fixed covariates and the error correlation among neighbors

Waller and Gotway, 363.

Space time analysis suggested by Bernardinelli et al.(1995)

$$\ln(\theta) = a + u_i + v_i + \beta t_k + \delta_i * t_k$$

where

$\beta = parameter\ estimate\ of\ \ time\ trend\ \ t_k$

$\delta_i * t_k = interaction\ random\ effect\ between\ space\ and\ \ time$

Goodness of Fit

Bayesian Information Criterion

$$BIC = -2\ln Likelihood + 2p\ln(T)$$

$$p = number\ of\ parameters$$

$$T = number\ of\ observations$$

Avg LL for G posterior estimates

$$ALL = \frac{1}{G}\sum_{i=1}^{G} LL_i$$

Residual Analysis

Lawson et al. (2003)(notes that we can analyze residuals for lack of systematic pattern to diagnose our model

general residual

$$r_{1i} = o_{1i} - \hat{o}_{1i}$$

standardized residual

$$r_{si} = \frac{o_{1i} - \hat{o}_{1i}}{\sqrt{\text{var}(o_{1i} - \hat{o}_{1i})}}$$

Bayesian residual(*Carlin and Lewis*,1996)

$$r_b = \frac{1}{G}\sum_{i=1}^{G} E(o_i \mid \theta_i^g)$$

where

$G = number\ of\ posterior\ samples$

Recapitulation

We have argued that Bayesian methods may be more amenable to analyzing nonnormal samples than classical methods.

We have shown that Bayesian methods, combined with the new computer power, permit estimation of sample distributions without closed form that conventional methods cannot.

We have shown that by using these methods, we can relax over constricting assumptions that lead to faulty assessment.

We have used these methods to analyze counts and possibly rare events with greater precision than conventional methods.

We use these methods to map psychopathological symptomatology that are generally nonnormally distributed.

Implications

When working with skewed or nonnormal distributions, we are better able to assess these populations with a Bayesian approach.

We obtain more accurate assessments if we have an idea of what the prior distribution looks like and we know what conjugate distributions may be used as that prior.

If we don't know what the prior distribution appears to be we can use a diffuse prior and still get a good estimate.

We do this by shrinking our confidence intervals and getting more accurate estimates.

This approach may be helpful for disease or symptomalogical mapping to help us appreciate the nature of perceived risk to a toxic catastrophe.

References

- Anselin, L. (2006). Spatial Econometrics, in (Mills, T. and Patterson, K, eds), Palgrave Handbook of Econometrics, (1),902-970.
- Carlin, B. and Lewis, T. (2008) Bayesian Methods for Data Analysis. 3rd Ed., Chapman and Hall, 121.
- Chandraskaran, K. and Arivarignan, G. (2006). Disease mapping using Mixture Distributions, Indian Journal of Medical Res. Vol. 123, 788-798.
- Jackman, S. (2004). Bayesian Modeling, Yale University (March 23, 2004).
- Lawson, A. B., Browne, W. J., Videl Rodiero, C. L. (2003). Bayesian Disease Mapping with Winbugs and Mlwin. New York: Wiley, 1-75.
- Lawson, A. B. (2008). Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. Chapman Hall, 117-188, chapter 6.
- Ma, H.,Virnig, B.A., and Carlin, B.P. (2005) Spatial Methods in Geographic Administrative Data Analysis.
- Messner, S. and Anselin, L. "General Methodological Issues with the Analysis of areal data, section three of Spatial analysis of homicide rates with areal data.