Published in final edited form as:

Stat Probab Lett. 2010 June 15; 80(11-12): 1015–1021. doi:10.1016/j.spl.2010.02.016.

On an asymptotic distribution of dependent random variables on a 3-dimensional lattice[☆]

Danielle J. Harvey^{a,*}, Qian Weng^b, and Laurel A. Beckett^a

^aDivision of Biostatistics, Department of Public Health Sciences, School of Medicine, University of California, Davis, CA 95616, U.S.A.

^bDepartment of Statistics, University of California, Davis, CA 95616, U.S.A.

Abstract

We define conditions under which sums of dependent spatial data will be approximately normally distributed. A theorem on the asymptotic distribution of a sum of dependent random variables defined on a 3-dimensional lattice is presented. Examples are also presented.

Keywords

Central Limit Theorem; spatial data; dependent data

1. Introduction

Imaging techniques provide non-invasive tools to track the development and progression of chronic pathologic processes. In early phase clinical trials, imaging may be used on animals to evaluate the usefulness of potential new therapeutic drugs. For example, the efficacy of a new cancer drug may be assessed by measuring the reduction in tumor size as seen on an image, negating the need to sacrifice the animals or at least extending the time until the sacrifice is conducted and allowing for multiple assessments on the same animal. Human imaging studies are also useful in the study of cancer as well as neurologic disorders such as dementia and schizophrenia where abnormalities of tissue are known to occur. Imaging enables the study of these abnormalities without the need of surgery, and is particularly useful in the context of neurodegenerative diseases where slicing into the brain is an unlikely strategy. Therefore, imaging has become a critical tool for the study of these chronic processes and the evaluation of new treatments for those conditions. However, imaging is incredibly expensive, so studies tend to be of small or moderate size.

Each image, itself, is an extremely rich data source. The images are broken down into hundreds of thousands if not millions of volume elements, or voxels, each of which contains information about the region or tissue being studied. These voxels may be thought of as data points on a 3-dimensional lattice. Due to the underlying anatomy and biology of the disease process, these

[☆]This work was supported by NIH/NIA grant # 5 P30 AG010129

© 2010 Elsevier B.V. All rights reserved.

*Corresponding author. Tel.: +1 530 752-8036; fax: +1 530-752-3239, djharvey@ucdavis.edu (Danielle J. Harvey), qweng@ucdavis.edu (Qian Weng), labeckett@ucdavis.edu (Laurel A. Beckett).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

data are likely to be highly correlated locally with the correlation decreasing as the Euclidean distance between points increases. There is, therefore, a need to perform sensible data reduction in such a way that under reasonable conditions, the reduced summary measures will yield approximately normally distributed measures for use in statistical analyses. Demonstration of the approximate normality of summary data from high-dimensional imaging would support use of standard linear model techniques for analysis even with small numbers of patients or animals, for example, use of two-sample t-tests to compare two treatments in small pre-clinical experiments.

Asymptotic distributions of sums of dependent data in one dimension have been studied extensively in the literature, using various dependence schemes. Many researchers have proved asymptotic normality for m -dependent random variables, in which random variables m units apart in sequence are assumed to be independent. Hoeffding and Robbins (1948) assumed a fixed value for m while Berk (1973) and Romano and Wolf (2000) allowed for m to change with the sample size and to grow infinitely large. Serfling (1968) examined asymptotic properties of sums of random variables under less stringent dependence structures.

Sajjan (2000) developed a central limit theorem for a stationary dependent process on the 2-dimensional lattice. For the lattice case, the idea of m -dependence is extended to (m_1, m_2) -dependence. Christofides and Mavrikiou (2003) also focused on the two-dimensional case, but relaxed the assumption of stationarity. In doing so, they made strong assumptions that were difficult to meet in practice. They also defined a dependence structure for the lattice, called ρ -radius dependence, that will be discussed below.

Use of linear combinations of data located on a lattice offers promise for the study of spatially distributed pathological processes, which may be observed through structural magnetic resonance images (MRI) or positron emission tomography (PET) scans as well as other imaging modalities. We present notation followed by a theorem describing conditions necessary for asymptotic normality of the sum of variables located on a 3-dimensional lattice. We provide the proof of the theorem as well as several examples for which the conditions of the theorem hold or the theorem might be useful.

2. Theoretical Considerations

Christofides and Mavrikiou (2003) defined the following concept of local dependence appropriate for the issue considered here:

Definition: For a positive integer r let $\mathbf{N}^r$ denote the r -dimensional positive integer lattice and let $\{X_{\mathbf{i}}, \mathbf{i} \in \mathbf{N}^r\}$ be an array of random variables defined on a common probability space $(\Omega, \mathcal{A}, P)$. Let $\rho \geq 0$. The random variables $\{X_{\mathbf{i}}, \mathbf{i} \in \mathbf{N}^r\}$ are said to be **ρ -radius dependent** if $X_{\mathbf{i}_1}$ and $X_{\mathbf{i}_2}$ are independent whenever $d(\mathbf{i}_1, \mathbf{i}_2) > \rho$, where $d(\mathbf{i}_1, \mathbf{i}_2)$ is the Euclidean distance between $\mathbf{i}_1$ and $\mathbf{i}_2$.

Let $\{X_{\mathbf{i}}, \mathbf{i} = (i_1, i_2, i_3) \leq (n_1, n_2, n_3)\}$ be an array of ρ -radius dependent three-dimensionally indexed random variables. $X_{\mathbf{i}}$ represent either individual values or weighted values, and we are interested in asymptotic properties of their sum. n_1 defines the vertical dimension (back to front), n_2 defines the horizontal dimension (left to right) and n_3 defines the spatial dimension (bottom to top).

Let $\rho^* = \lceil \rho \rceil$, the smallest integer greater than or equal to ρ .

Let v_n be a positive integer greater than ρ^* , which is allowed to change with $n = (n_1, n_2, n_3)$.

Let $T_{i_1, i_2, i_3} = \{(j_1, j_2, j_3) : i_k - v_n \leq j_k \leq i_k + v_n, k = 1, 2, 3\}$ so that T_{i_1, i_2, i_3} is the $(2v_n + 1) \times (2v_n + 1) \times (2v_n + 1)$ cube whose center is the point (i_1, i_2, i_3) .

Let

$$S_{i_1, i_2, i_3}^{(T)} = \sum_{(j_1, j_2, j_3) \in T_{i_1, i_2, i_3}} X_{j_1, j_2, j_3},$$

the total (or weighted total) in the cube.

The lattice may be divided into a set of independent cubes and borders separating the cubes, beginning in the back left corner of the bottom layer. The first cube is centered at (k_n, k_n, k_n) where $k_n = v_n + 1$. Cube centers will be regularly spaced at intervals of $\lambda_n = 2v_n + \rho^* + 1$, where λ_n represents the width of the cube and the adjacent border. Because the variables X_{i_1, i_2, i_3} are ρ -radius dependent and the cubes are separated by a border of width $\rho^* > \rho$, the cube sums are independent. We assume that $n_i + \rho^*$ may be evenly divided by λ_n , so that there are no partial cubes in the i^{th} dimension ($i = 1$: back to front, $i = 2$: left to right, $i = 3$: bottom to top). The cubes are centered at $(i_1, i_2, i_3) = (k_n + j_1\lambda_n, k_n + j_2\lambda_n, k_n + j_3\lambda_n)$, where $0 \leq j_i \leq D_i - 1$ and D_i is the number of cubes in each dimension.

The regions of points that do not belong to any of the cubes may be divided into seven zones for each cube (Figure 1). Three regions are the “flat” rectangular areas adjacent to the right, front, and top of a cube, but not symbolically extending past the edges of the cube. Next we identify three strips adjacent to the edges but not extending past the corners of the cube. Finally, the remaining region is the small cube adjacent to the top right front corner.

For each cube, T_k where $k = (i_1, i_2, i_3)$, define $S_k^{(B)}$ to be the sum of the points in the seven boundary regions surrounding it.

The proof of the theorem relies on taking limits in such a fashion that the normalized sum of the sums of the random variables in the cubes is asymptotically normal, by standard central limit theorem arguments, while the normalized sum of the random variables in the boundary regions goes to zero in probability.

2.1. Theorem

The main theorem for the asymptotic distribution of the sum of random variables located on a spatial lattice is given below.

Theorem 1. Let $\{X_{i_1, i_2, i_3}, (i_1, i_2, i_3) \leq (n_1, n_2, n_3)\}$ be an array of ρ -radius dependent three dimensionally indexed random variables. Without loss of generality, assume that these variables have mean zero. Let v_n be a positive integer greater than $\rho^* = \lceil \rho \rceil$, as described above in the notation, and $\lambda_n = 2v_n + \rho^* + 1$. Let $n_c = D_1 D_2 D_3$ be the total number of cubes. Let

σ_{i_1, i_2, i_3}^2 be the variance of $S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)}$ and let also $r_{i_1, i_2, i_3}^3 = E \left[S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)} \right]^3$. Assume the finiteness of these quantities, and define

$$r_{n_1, n_2, n_3}^3 = \sum_{i_1=0}^{D_1-1} \sum_{i_2=0}^{D_2-1} \sum_{i_3=0}^{D_3-1} r_{i_1, i_2, i_3}^3,$$

$$\sigma_{n_1, n_2, n_3}^2 = \sum_{i_1=0}^{D_1-1} \sum_{i_2=0}^{D_2-1} \sum_{i_3=0}^{D_3-1} \sigma_{i_1, i_2, i_3}^2.$$

Assume that $n_1, n_2, n_3 \rightarrow \infty$ monotonically, that $n_c \rightarrow \infty$, and that $v_n \rightarrow \infty$ at a rate slower than n_1, n_2, n_3 , such that

$$\frac{\sum_{k=1}^{n_c} \text{Var}(S_k^{(B)})}{n_c \sigma_{n_1, n_2, n_3}^2} \rightarrow 0 \quad (1)$$

$$\frac{r_{n_1, n_2, n_3}}{\sigma_{n_1, n_2, n_3}} \rightarrow 0. \quad (2)$$

Let

$$\bar{X}_{n_1, n_2, n_3} = \frac{1}{n_1 n_2 n_3} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \sum_{i_3=1}^{n_3} X_{i_1, i_2, i_3}.$$

Then

$$\frac{\sqrt{n_s}(\bar{X}_{n_1, n_2, n_3})}{\sigma_{n_1, n_2, n_3}} \xrightarrow{\mathcal{L}} N(0, 1)$$

$$\text{where } n_s = \frac{(n_1 n_2 n_3)^2}{n_c}.$$

PROOF. The proof of this theorem uses a method first introduced by Bernstein (1927). We first note that we can decompose the sum of the components of the lattice as follows:

$$\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \sum_{i_3=1}^{n_3} X_{i_1, i_2, i_3} = \sum_{i_1=0}^{D_1-1} \sum_{i_2=0}^{D_2-1} \sum_{i_3=0}^{D_3-1} S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)} + \sum_{k=1}^{n_c} S_k^{(B)}$$

so that

$$\bar{X}_{n_1, n_2, n_3} = \frac{1}{n_1 n_2 n_3} \sum_{i_1=0}^{D_1-1} \sum_{i_2=0}^{D_2-1} \sum_{i_3=0}^{D_3-1} S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)} + \frac{1}{n_1 n_2 n_3} \sum_{k=1}^{n_c} S_k^{(B)}.$$

We prove the Theorem using Slutsky's Theorem, which requires showing that

$$\frac{\sqrt{n_s}}{n_1 n_2 n_3 \sigma_{n_1, n_2, n_3}} \sum_{i_1=0}^{D_1-1} \sum_{i_2=0}^{D_2-1} \sum_{i_3=0}^{D_3-1} S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)} \xrightarrow{\mathcal{L}} N(0, 1)$$

and that

$$\frac{\sqrt{n_s} \sum_{k=1}^{n_c} S_k^{(B)}}{n_1 n_2 n_3 \sigma_{n_1, n_2, n_3}} \xrightarrow{p} 0$$

$$\text{where } n_s = \frac{(n_1 n_2 n_3)^2}{n_c}.$$

We begin by showing the first part. As mentioned above, the cubes defined by

$$\left\{ S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)}, (i_1, i_2, i_3) \leq (D_1 - 1, D_2 - 1, D_3 - 1) \right\}$$

are independent. They have mean zero, and a finite variance, σ_{i_1, i_2, i_3}^2 . Lyapounov's condition (Billingsley (1995)) is satisfied for $\delta = 1$, since

$$\lim_{(n_1, n_2, n_3) \rightarrow \infty} \frac{r_{n_1, n_2, n_3}}{\sigma_{n_1, n_2, n_3}} = 0 \Rightarrow \lim_{(n_1, n_2, n_3) \rightarrow \infty} \frac{r_{n_1, n_2, n_3}^3}{\sigma_{n_1, n_2, n_3}^3} = 0.$$

Note that

$$\begin{aligned} & \frac{\sqrt{n_s}}{n_1 n_2 n_3 \sigma_{n_1, n_2, n_3}} \sum_{i_1=0}^{D_1-1} \sum_{i_2=0}^{D_2-1} \sum_{i_3=0}^{D_3-1} S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)} \\ &= \frac{n_c^{-1/2}}{\sigma_{n_1, n_2, n_3}} \sum_{i_1=0}^{D_1-1} \sum_{i_2=0}^{D_2-1} \sum_{i_3=0}^{D_3-1} S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)} \\ &= \frac{\sqrt{n_c}}{\sigma_{n_1, n_2, n_3}} \bar{S} \end{aligned}$$

where $\bar{S} = \frac{1}{n_c} \sum_{i_1=0}^{D_1-1} \sum_{i_2=0}^{D_2-1} \sum_{i_3=0}^{D_3-1} S_{(k_n + \lambda_n i_1, k_n + \lambda_n i_2, k_n + \lambda_n i_3)}^{(T)}$. By the Lyapounov version of the Central Limit Theorem,

$$\frac{\sqrt{n_c}}{\sigma_{n_1, n_2, n_3}} \bar{S} \xrightarrow{\mathcal{L}} N(0, 1).$$

We now show the second part:

$$\frac{\sqrt{n_s} \sum_{k=1}^{n_c} S_k^{(B)}}{n_1 n_2 n_3 \sigma_{n_1, n_2, n_3}} = \frac{\sum_{k=1}^{n_c} S_k^{(B)}}{\sqrt{n_c} \sigma_{n_1, n_2, n_3}}.$$

By Chebyshev's inequality, if we can show that

$$E \left(\frac{\sum_{k=1}^{n_c} S_k^{(B)}}{\sqrt{n_c} \sigma_{n_1, n_2, n_3}} - 0 \right)^2 \rightarrow 0$$

then

$$\frac{\sum_{k=1}^{n_c} S_k^{(B)}}{\sqrt{n_c} \sigma_{n_1, n_2, n_3}} \xrightarrow{p} 0.$$

$$\begin{aligned} E \left(\frac{\sum_{k=1}^{n_c} S_k^{(B)}}{\sqrt{n_c} \sigma_{n_1, n_2, n_3}} - 0 \right)^2 &= \frac{1}{n_c \sigma_{n_1, n_2, n_3}^2} E \left(\sum_{k=1}^{n_c} S_k^{(B)} \right)^2 \\ &= \frac{1}{n_c \sigma_{n_1, n_2, n_3}^2} \text{Var} \left(\sum_{k=1}^{n_c} S_k^{(B)} \right) \\ &\leq \frac{27}{n_c \sigma_{n_1, n_2, n_3}^2} \sum_{k=1}^{n_c} \text{Var} (S_k^{(B)}) \end{aligned} \quad (3)$$

$$\rightarrow 0 \quad (4)$$

where (3) follows because each cube and its border regions only touch borders of the 26 adjacent cubes that share faces, edges, or corners. Also, $\text{Var}(A + B) \leq \text{Var}A + \text{Var}B + 2\sigma_A\sigma_B$ and $2\sigma_A\sigma_B \leq \sigma_A^2 + \sigma_B^2$. Line (4) follows from the assumptions and the proof is complete.

2.2. Notes

The assumptions of the theorem are intuitive and are likely to be met by measures of interest related to image data. They are also similar to assumptions used in other versions of the Central Limit Theorem. In particular, Assumption (2) is the standard Central Limit Theorem assumption that the distribution does not have extremely long tails. Assumption (1) guarantees that the lattice may be decomposed into independent blocks that contribute the dominant share of information (variance) when we sum over the lattice, while the share of the border regions can be made arbitrarily small by choosing large enough $n = n_1 n_2 n_3$. One way to allow for the blocks to provide the majority of the information in the lattice is to allow the block size to grow as the overall dimension of the lattice grows (something that had not been considered by Christofides and Mavrikiou (2003)).

Researchers may question the assumption of zero mean in the theorem, since an image made up of zero-mean voxels would not be very interesting. However, if all of the voxels had the same mean, the theorem would still hold. Segmentation strategies for structural images often assume that intensities of particular tissue type, such as grey matter in the brain, come from a mixture of Gaussian distributions (Ashburner and Friston (2005)). Therefore, it is assumed that voxels belonging to a particular class of tissue all have the same mean. If voxels have different means, subtracting a three-dimensional array of the voxel means from the image would result in an array of zero-mean variables in which the correlation structure has been preserved. Therefore, as long as the conditions of the theorem still hold on this transformed array, the linear combinations of the voxel-level data will be approximately normally distributed (see Example 2.2). Finally, patterns in the image may be explained by demographic or clinical information. If linear regression models are used to predict voxel-level data, the array consisting of residuals from these models would have zero mean at each position. Summaries

based on this residual array would be helpful for residual diagnostics to determine if there is any additional structure remaining in the data. Therefore, there are many contexts in which the image data do not have zero mean in which the theorem may still be applied.

It is important to note that in this imaging setting, $n = n_1 n_2 n_3$ represents the size of a single image that generates the desired spatial process, not the number of individuals imaged. Taking the limit as the size of the lattice grows infinitely large may be thought of as looking at larger and larger images or regions. Therefore, the theorem states that under suitable conditions, linear summaries of imaging data will be approximately normal if the image had enough voxels, rather than requiring a large number of subjects (as is the usual interpretation of $n \rightarrow \infty$ in other Central Limit Theorems.) This theorem supports the use of standard normal-theory approaches for analysis of studies with even fairly modest numbers of people using a variety of common-sense linear combinations of voxel-level data. In addition, it provides a rationale for use of standard linear model techniques with normality assumptions to examine the role of clinical or environmental variables in explaining variation in image data

However, the theorem says that summary measures derived by adding combinations of the voxel-level data will be approximately normally distributed, since in practice, the number of voxels will always be finite. Given a fixed number of voxels, if the spatial dependence is too strong, this approximation may not be appropriate, so it is important to understand the properties of the underlying data. Some preliminary work we have done with structural images suggest that the spatial correlation of error terms drops off drastically as the distance between voxels increases (data not shown), but that cannot be guaranteed for all data generated from images. Researchers should also be careful if summaries are based on small regions of the brain, since again, the approximation may not be appropriate. However, many brain regions of interest, even though anatomically small, may still contain relatively large numbers of voxels. For example, a recent publication by Schuff et al. (2009) presents average volumes of the hippocampus, a small region of interest in aging studies, that range from about 1600 mm³ to over 2000 mm³. Since voxels in MRI are often 1 mm³, these volumes translate to well over 1000 voxels on which the summaries are based. In situations of strong spatial dependence or small regions, more robust statistical analytic techniques that do not require the assumption of exact normality, such as resampling or permutation methods, should be considered.

2.3. Examples

Example 2.1 In the simplest case of independent, identically distributed random variables (ρ – radius dependent with $\rho = 0$), by definition, the border regions are non-existent, since the width of the border region is the smallest integer greater than or equal to ρ (0). Therefore, the lattice may be decomposed entirely into independent cubes and it is easy to show that the conditions of the theorem are met as long as the distribution has finite mean, variance, and absolute third moment.

Example 2.2 Now consider a slightly more difficult case, which incorporates non-normal data as well as dependence. Many uses of imaging involve the identification of abnormalities (e.g. abnormal tissue or function). Therefore, it would be useful to identify conditions under which the sum of a set of dependent Bernoulli random variables is approximately normal. Let $\{X_{i_1, i_2, i_3}, (i_1, i_2, i_3) \leq (n_1, n_2, n_3)\}$ be an array of ρ -radius dependent three dimensionally indexed Bernoulli (p_{i_1, i_2, i_3}) random variables with p_{i_1, i_2, i_3} bounded away from zero and 1 ($0 < p_{\min} \leq p_{i_1, i_2, i_3} \leq p_{\max} < 1$ for all (i_1, i_2, i_3)), where $p_{\min}$ and $p_{\max}$ are the smallest and largest probabilities of a success across the voxels. Assume that the overall proportion of successes (abnormalities) in the entire lattice remains constant as the lattice grows and that the random variables are positively correlated. The assumption of positive correlation is reasonable, since damage is thought to spread locally, so that if one voxel is damaged, surrounding voxels are more likely to also be damaged. All conditions of the theorem are met in this case (see the

sketch of the proof below), so that summaries consisting of simple sums of the voxel-level data over sufficiently large images will be approximately normal.

Sketch of proof: We first transform the variables so that they have zero mean: $Y_{i_1, i_2, i_3} = X_{i_1, i_2, i_3} - p_{i_1, i_2, i_3}$. Each cube has dimension $(2v_n + 1) \times (2v_n + 1) \times (2v_n + 1)$ and consists of a total of $(2v_n + 1)^3$ random variables (voxels). The border regions between cubes ($S_k^{(B)}$) consist of a total of $3(2v_n + 1)^2 \rho^* + 3(\rho^*)^2(2v_n + 1) + (\rho^*)^3$ random variables. It is then easy to show that for a given cube size, σ_{i_1, i_2, i_3}^2 and r_{i_1, i_2, i_3}^3 are both finite.

To show Assumption (1),

$$\begin{aligned} \text{Var}(S_k^{(B)}) &\leq C_1(2v_n + 1)^4 \\ \Rightarrow \sum_{k=1}^{n_c} \text{Var}(S_k^{(B)}) &\leq n_c C_1(2v_n + 1)^4 \end{aligned} \quad (5)$$

$$\begin{aligned} \sigma_{i_1, i_2, i_3}^2 &\geq \sum_{i_1} \sum_{i_2} \sum_{i_3} p_{i_1, i_2, i_3} (1 - p_{i_1, i_2, i_3}) \\ &\geq \sum_{i_1} \sum_{i_2} \sum_{i_3} p_{\min} (1 - p_{\min}) = C_2(2v_n + 1)^3 \\ \Rightarrow \sigma_{n_1, n_2, n_3}^2 &\geq n_c C_2(2v_n + 1)^3. \end{aligned} \quad (6)$$

$$\begin{aligned} \text{Therefore, } \frac{\sum_{k=1}^{n_c} \text{Var}(S_k^{(B)})}{n_c \sigma_{n_1, n_2, n_3}^2} &\leq \frac{n_c C_1(2v_n + 1)^4}{n_c^2 C_2(2v_n + 1)^3} \\ &= \frac{2v_n + 1}{n_c} \left(\frac{C_1}{C_2} \right) \rightarrow 0. \end{aligned} \quad (7)$$

In the above, C_1 and C_2 are constants that do not depend on n_1 , n_2 , or n_3 . To show (5), use the definition of the variance of a sum of random variables and note that

$\text{Cov}(X_{i_1, i_2, i_3}, X_{i_1', i_2', i_3'}) \leq P(X_{i_1, i_2, i_3} = 1, X_{i_1', i_2', i_3'} = 1) \leq 1$. (6) follows from the definition of the variance of a sum of random variables and from the assumption of positively correlated random variables. Finally, (7) follows from the assumption that v_n grows at a rate slower than n_1 , n_2 , n_3 , and, therefore, slower than n_c .

To show Assumption (2),

$$\begin{aligned} r_{i_1, i_2, i_3}^3 &\leq C_3(2v_n + 1)^9 \\ \Rightarrow r_{n_1, n_2, n_3}^3 &\leq n_c C_3(2v_n + 1)^9 \\ \Rightarrow r_{n_1, n_2, n_3} &\leq n_c^{1/3} C_3^{1/3} (2v_n + 1)^3 \\ \sigma_{n_1, n_2, n_3}^2 &\geq n_c C_2(2v_n + 1)^3 \\ \Rightarrow \sigma_{n_1, n_2, n_3} &\geq n_c^{1/2} C_2^{1/2} (2v_n + 1)^{3/2}. \end{aligned} \quad (8)$$

$$\begin{aligned}
 \text{Therefore, } \frac{r_{n_1, n_2, n_3}}{\sigma_{n_1, n_2, n_3}} &\leq \frac{n_c^{1/3} C_3^{1/3} (2v_n + 1)^3}{n_c^{1/2} C_2^{1/2} (2v_n + 1)^{3/2}} \\
 &= \frac{(2v_n + 1)^{3/2}}{n_c^{1/6}} \left(\frac{C_3^{1/3}}{C_2^{1/2}} \right) \rightarrow 0.
 \end{aligned} \tag{9}$$

Once again, C_3 is a constant that does not depend on n_1 , n_2 , or n_3 . By using the triangle inequality and then expanding the cubic polynomial in r_{i_1, i_2, i_3}^3 (8) can be shown. Finally, because we assume that v_n grows at a rate slower than n_1 , n_2 , n_3 (and therefore, n_c), (9) follows and the assumption is met.

Example 2.3 Many biologically driven summaries of image data will be linear combinations of the voxel-level data with scalar multipliers different from one (the case in Example 2.2). However, a minor constraint on the multipliers guarantees that the conditions of the theorem are met. Let $\{X_{i_1, i_2, i_3}, (i_1, i_2, i_3) \leq (n_1, n_2, n_3)\}$ be an array of ρ -radius dependent three dimensionally indexed Bernoulli (p_{i_1, i_2, i_3}) random variables, as described in the previous example. Consider the new array $\{C_{i_1, i_2, i_3} X_{i_1, i_2, i_3}, (i_1, i_2, i_3) \leq (n_1, n_2, n_3)\}$, where $|C_{i_1, i_2, i_3}| < M$ for some constant M . This new array satisfies the conditions of the theorem. Therefore, as long as the multipliers used in the linear combination are bounded (and the conditions of the theorem are met for the original voxel-level data), these summaries will also be approximately normal.

Example 2.4 As a practical example in neuroimaging research, consider a set of positron emission tomography (PET) scans, measuring the glucose metabolism of the brain. Suppose the voxel values have been standardized to a normal, healthy population by centering at the normal mean and dividing by the standard deviation of the normal population. The hypothesis of interest is that the scans of these subjects show reduced glucose metabolism relative to the healthy population. Under the null hypothesis of metabolism similar to the healthy individuals, the standardized voxel values (X_{i_1, i_2, i_3}) have zero mean. If we further assume that $E|X_{i_1, i_2, i_3}|^3$ is finite and that only those voxels that are close by (for example, sharing a side) are positively correlated, the conditions of the theorem hold and the sum of the voxel values will be approximately normally distributed. The proof for the assumptions follows similar to that given above for Example 2.2.

3. Conclusion

We present an asymptotic result for spatially correlated data and provide examples for which the theorem applies. The theorem suggests that linear combinations of voxel-level data over a large enough region (obtained through an image) may be used as outcomes in a regression setting. We have shown that linear combinations generated from images of people or animals with similar underlying characteristics will be approximately normally distributed even for small to moderate numbers of individuals, provided the images themselves cover a large enough region with a high enough resolution (large number of voxels). Such linear combinations include simple sums of the image data or weighted sums with weights defined according to an underlying biological hypothesis. Applications where this theorem might be useful include understanding patterns of tissue or metabolic abnormalities seen on structural MRI or PET scans or activation patterns observed through functional MRI, all imaging approaches featuring high-resolution images, provided the region of interest is substantially larger than the local correlation radius.

Acknowledgments

The authors would like to thank the two referees for their extremely helpful comments, which greatly improved the manuscript.

References

- Ashburner J, Friston KJ. Unified segmentation. *Neuroimage* 2005;26:839–851. [PubMed: 15955494]
- Berk KN. A central limit theorem for m -dependent random variables with unbounded m . *The Annals of Probability* 1973;1:352–354.
- Bernstein S. Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Mathematische Annalen* 1927;97:1–59.
- Billingsley, P. *Probability and Measure*. New York: John Wiley & Sons; 1995.
- Christofides TC, Mavrikiou PM. Central limit theorem for dependent multidimensionally indexed random variables. *Statistics & Probability Letters* 2003;63:67–78.
- Hoeffding W, Robbins H. The central limit theorem for dependent random variables. *Duke Mathematical Journal* 1948;15:773–780.
- Romano JP, Wolf M. A more general central limit theorem for m -dependent random variables with unbounded m . *Statistics & Probability Letters* 2000;47:115–124.
- Sajjan SG. A note on central limit theorems for lattice models. *Journal of Statistical Planning and Inference* 2000;83:283–290.
- Schuff N, Woerner N, Boreta L, Kornfield T, Shaw LM, Trojanowski JQ, Thompson PM, Jack CR Jr, Weiner MW. The Alzheimer's Disease Neuroimaging Initiative. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain* 2009;132:1067–1077. [PubMed: 19251758]
- Serfling RJ. Contributions to central limit theory for dependent variables. *The Annals of Mathematical Statistics* 1968;39:1158–1175.

Figure 1.

The different components of $S_k^{(B)}$; “A” = above, “R” = right, “F” = front, “RF” = right and front, “AR” = above and right, “AF” = above and front, “RFA” = right, front, and above.