Limit Theory for the Domination Number of Random Class Cover Catch Digraphs

Pengfei Xiang and John C. Wierman
Department of Mathematical Sciences
The Johns Hopkins University

Abstract

We discuss the limiting behavior of the domination number of random class cover catch digraphs (CCCDs). The CCCD problem is motivated by its applications in pattern classification. For the special case of uniformly distributed data in one dimension, Priebe, Marchette and Devinney found the exact distribution of the domination number of the random data-induced CCCD, and Devinney and Wierman proved the Strong Law of Large Numbers (SLLN). We will present progress toward the SLLN and the Central Limit Theorem (CLT) for general data distributions in one dimension. The ultimate goal is to establish SLLN and CLT results for higher dimensional CCCDs.

Keywords: Class Cover Catch Digraph, Domination Number, Strong Law of Large Numbers, Central Limit Theorem, Pattern Classification

1 Introduction

1.1 Class Cover Problem

The class cover problem (CCP) was introduced by Cowen and Cannon [1], motivated by statistical pattern classification [8].

The CCP is described as follows: A *dissimilarity space* is a pair (Ω, d) where Ω is a set, and d is a function: $\Omega \times \Omega \to \mathbf{R}$ such that $d(\alpha, \beta) = d(\beta, \alpha) \geq d(\alpha, \alpha) = 0$ for all $\alpha, \beta \in \Omega$, called a *dissimilarity function*. Suppose there are two classes of Ω-elements: $\{x_i : i = 1, \cdots, n\}$ and $\{y_j : j = 1, \cdots, m\}$. For each x_i, we define its *covering ball* as follows:

Definition 1.1. $B(x_i) = \{\omega \in \Omega : d(\omega, x_i) < min_j d(y_j, x_i)\}$.

A *class cover of* $\{x_i : i = 1, \cdots, n\}$ is a subset of covering balls whose union contains all x_i's. Obviously the set consisting of all covering balls is a class cover. The CCP we consider here is to find a minimum cardinality class cover.

1.2 Class Cover Catch Digraph

We can convert the CCP to a purely graph theory problem as follows:

Definition 1.2. *The class cover catch digraph (CCCD) induced by a CCP is the digraph* $D = (V, A)$ *with the vertex set* $V = \{x_i : i = 1, \cdots, n\}$ *and the edge set* A *such that there is a directed edge* (x_i, x_j) *if and only if* $x_j \in B(x_i)$.

Definition 1.3. *The set* $S \subset V$ *is a dominating set of a diagraph* $D = (V, A)$ *if and only if for all* $v \in V$, *either* $v \in S$ *or* $(s, v) \in A$ *for some* $s \in S$.

It is easy to see that the CCP is actually equivalent to finding a minimum cardinality dominating set of the corresponding CCCD. Cowen and Cannon prove that the dominating set problem is essentially a special case of the CCP, and since the dominating set problem is NP-hard, it follows that the CCP is also NP-Hard [1].

1.3 Domination Number

Definition 1.4. *The domination number of a CCCD is the cardinality of the CCCD's minimum dominating set.*

The domination number serves as a measure of success in distinguishing the classes $\mathcal{X}$ and $\mathcal{Y}$ from each other. In 1962, Ore [12] first used the name "domination number". Due to its many applications in such fields as computer networks, social sciences and computational complexity, there has been increasing interest in this topic. Haynes, Hedetniemi and Slater [6] provide a comprehensive discussion of domination in graphs, with more advanced topics covered in [7].

1.4 Randomization

To study the problem from a statistical perspective, we need to add randomness to the Ω-valued points x_i and y_j. Specifically, we replace x_i by random variable X_i, and y_j by random variable Y_j. Here we suppose $\{X_i : i = 1, \cdots, n\}$ and $\{Y_j : j = 1, \cdots, m\}$ are two sets of i.i.d. random variables taking values in Ω, with distribution functions $F_{\mathcal{X}}$ and $F_{\mathcal{Y}}$, respectively. We assume that the X_i's are independent of the Y_j's. After

such randomization, all the previous definitions and notions can still be applied, provided that all X_i's and Y_j's are distinct with probability one. In particular, we denote the domination number by $\Gamma_{n,m}(F_{\mathcal{X}}, F_{\mathcal{Y}})$, or simply by $\Gamma_{n,m}$. Obviously, $\Gamma_{n,m}$ is a random variable whose distribution depends on $n, m, F_{\mathcal{X}}$ and $F_{\mathcal{Y}}$.

1.5 Applications in Pattern Classification

Pattern classification is "the assignment of a physical object or event to one of several pre-specified categories" (See [3, page 2]). It is widely applied to real world problems such as automated speech recognition, DNA sequence identification, and fingerprint identification. For a thorough description of pattern classification, see the two classic books [4] and [2].

The abstract mathematical model of the pattern classification problem is formulated as follows [8]. For simplicity, but without loss of generality, suppose we have two classes of objects of interest, which we will call class $\mathcal{X}$ and class $\mathcal{Y}$, respectively. We assume that the objects of both classes belong to a common dissimilarity space Ω. To model the uncertainty about which class the objects we encounter belong, we assume that there are *prior* probabilities $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ for these two classes ($\sum_{c\in\{\mathcal{X},\mathcal{Y}\}} P_c = 1$). Furthermore, we assume that given the class $\mathcal{X}$ or $\mathcal{Y}$, the objects of that class are drawn according to the *class-conditional* distribution functions $F_{\mathcal{X}}(x)$ or $F_{\mathcal{Y}}(y)$. We can generate a random pair $\big(c(\Psi), \Psi\big)$ in a two-step process: first choose the random class label $c(\Psi) \in \{\mathcal{X}, \mathcal{Y}\}$ according to the prior probabilities; then based on the chosen class, select Ψ according to the corresponding class-conditional distribution function.

In a classification problem, for an observation pair $(c(\psi), \psi)$ generated as above, only the data part ψ is given while the class label part $c(\psi)$ is unknown, so the goal of a *classifier* is to correctly guess whether $c(\psi)$ is $\mathcal{X}$ or $\mathcal{Y}$. We are given a training sample of size k with known classification:

$$D_k = \Big\{ \big(c(\psi_1), \psi_1\big), \cdots, \big(c(\psi_k), \psi_k\big) \Big\}.$$

A classifier is a function $\hat{c}_k(\psi) = \hat{c}_k(\psi, D_k)$, which, based on the training data D_k, assigns a class label $\mathcal{X}$ or $\mathcal{Y}$ to any input point $\psi \in \Omega$. The performance of a classifier $\hat{c}$ can be measured by the *probability of error*, or *misclassification rate*, given by

$$E\big[P(\hat{c}_k(\Psi) \neq c(\Psi) \mid D_k)\big].$$

The CCP has been actively studied recently, since its solution can be directly used to generate classifiers competitive with the other methods. The data points $\{X_i : i = 1, \cdots, n\}$ and $\{Y_j : j = 1, \cdots, m\}$ constitute the training data from classes $\mathcal{X}$ and $\mathcal{Y}$, respectively. Thus, in the setting

of classification, the CCP is just a problem of selecting a small set of data points to be representative of a class. We want this set as small as possible, i.e. a minimal cardinality dominating set, to make the classifier less complex while keeping most of the relevant information. A simple CCCD classifier can be constructed as follows: by switching the roles of $\mathcal{X}$ and $\mathcal{Y}$, we can get a pair of dual CCP's, resulting in two solutions such as $\mathcal{B}_{\mathcal{X}} = \{B(X_i) : i \in I, I \subset \{1, \cdots, n\}\}$ and $\mathcal{B}_{\mathcal{Y}} = \{B(Y_j) : j \in J, J \subset \{1, \cdots, m\}\}$, respectively. Define $\mathcal{C}_{\mathcal{X}} = \{\omega \in \Omega : \omega \in B(X_i) \text{ s.t. } B(X_i) \in \mathcal{B}_{\mathcal{X}}\}$, $\mathcal{C}_{\mathcal{Y}} = \{\omega \in \Omega : \omega \in B(Y_i) \text{ s.t. } B(Y_i) \in \mathcal{B}_{\mathcal{Y}}\}$. We can incorporate these two solutions into a classifier $\hat{c}(\psi) : \Omega \to \{\mathcal{X}, \mathcal{Y}\}$ as follows:

$$\hat{c}(\psi) = \begin{cases} \mathcal{X} & \psi \in \mathcal{C}_{\mathcal{X}} \cap {\mathcal{C}_{\mathcal{Y}}}^c, \\ \mathcal{Y} & \psi \in \mathcal{C}_{\mathcal{Y}} \cap {\mathcal{C}_{\mathcal{X}}}^c, \\ \text{undetermined} & \text{otherwise.} \end{cases}$$

More details about the CCP's application to classification are presented in [14].

2 Previous Results

There have been several research results on the probabilistic properties of $\Gamma_{n,m}$ in the case of $\Omega = \mathbf{R}$. In this one dimensional situation, we denote $Y_{(j)}$ as the jth order statistic of $Y_0 = 0, Y_1, \cdots, Y_m, Y_{m+1} = 1$, and let the random variable $\alpha_{j,m}$ be the minimum number of covering balls needed to cover the $N_{j,m}$ $\mathcal{X}$-class points located between $Y_{(j)}$ and $Y_{(j+1)}$. We refer to $\alpha_{j,m}$ $(j = 0, m)$ as external components, and $\alpha_{j,m}$ $(j = 1, \cdots, m-1)$ as internal components. It should be noted that $\Gamma_{n,m} = \sum_{j=0}^{m} \alpha_{j,m}$. This way we are able to decompose the problem into $m+1$ sub-problems of finding the domination number $\alpha_{j,m}$ in the interval $[Y_{(j)}, Y_{(j+1)}]$.

It is obvious that $\alpha_{j,m} = 0$ if and only if $N_j = 0$. We see that $\alpha_{j,m}$ can be at most 2, because all X_i's in $[Y_{(j)}, Y_{(j+1)}]$ are contained in the covering balls of the two $\mathcal{X}$ points that are closest to midpoint of this interval on the right and left. Since the domination number is a non-negative integer, $\alpha_{j,m}$ can only be $0, 1$ or 2.

With careful analysis, the probability of each of these values was determined exactly by Priebe, Devinney and Marchette [13]. They found the conditional distribution of $\alpha_{j,m}$ given $N_{j,m}$ for the special case of $\Omega = \mathbf{R}$ and $F_{\mathcal{X}} = F_{\mathcal{Y}} = U[0,1]$, where $U[0,1]$ is the uniform distribution on the interval $[0,1]$:

Theorem 2.1. *Suppose $\Omega = \mathbf{R}$. If $F_{\mathcal{X}} = F_{\mathcal{Y}} = U[0,1]$, then the following are true:*

- *For $j \in \{0, 1, \cdots, m\}$, if $N_{j,m} = 0$ then $\alpha_{j,m} = 0$.*

- *For $j \in \{0, m\}$, if $N_{j,m} > 0$ then $\alpha_{j,m} = 1$.*
- *For $j \in \{1, 2, \cdots, m-1\}$, if $N_{j,m} = n_{j,m} > 0$ then*

$$\begin{aligned} P(\alpha_{j,m} = 1 \mid N_{j,m} = n_{j,m}) &= 1 - P(\alpha_{j,m} = 2 \mid N_{j,m} = n_{j,m}) \\ &= \frac{5}{9} + \frac{4}{9}\frac{1}{4^{n_{j,m}-1}}. \end{aligned}$$

Note that the interval components $\{\alpha_{j,m} : 1 \leq j \leq m-1\}$ are identically distributed, and the external components $\alpha_{0,m}$ and $\alpha_{m,m}$ are identically distributed.

From the above theorem, we know that for $j \in \{1, 2, \cdots, m-1\}$, given $N_{j,m} = n_{j,m} > 0$, the conditional probability of $\alpha_{j,m} = 2$ is an increasing function of $n_{j,m}$, which just means that $\alpha_{j,m}$ tends to become larger as the number of $\mathcal{X}$ points increases, for fixed m.

Under the same assumptions as Theorem 1.1, Devinney and Wierman proved a strong law of large numbers for $\Gamma_{n,m}$ [5]:

Theorem 2.2. *Suppose $\Omega = \mathbf{R}$, and $F_{\mathcal{X}} = F_{\mathcal{Y}} = U[0,1]$. If $m = \lfloor rn \rfloor, r \in (0, \infty)$, then*

$$\lim_{n \to +\infty} \frac{\Gamma_{n,m}}{m} = g(r) \quad a.s.$$

where

$$g(r) \equiv \frac{12r + 13}{3(r+1)(4r+3)}.$$

As suggested in Figure 1, $g(r) \to 0$ when $r \to \infty$, which is justified by the fact that asymptotically the interval between $Y_{(j)}$ and $Y_{(j+1)}$ contains no $\mathcal{X}$ point almost surely. Moreover, $g(r) \to \frac{13}{9}$ as $r \to 0$. This corresponds to the situation where each interval between $Y_{(j)}$ and $Y_{(j+1)}$ contains very large number of $\mathcal{X}$ points. According to Theorem 1.1, the probability of $\alpha_{j,m} = 1$ is approximately $\frac{5}{9}$, while the probability of $\alpha_{j,m} = 2$ is approximately $\frac{4}{9}$, so $\frac{13}{9} = \frac{5}{9} \cdot 1 + \frac{4}{9} \cdot 2$ can be just viewed as an expectation value of $\alpha_{j,m}$.

In their proof in [5], DeVinney and Wierman first prove the special case of $r = 1$. They construct two related Poisson processes A and B, with common rate $\lambda \in (0, \infty)$. Points of A play the role of $\mathcal{X}$ points, and points of B play the role of $\mathcal{Y}$ points. The classical SLLN can be applied to a CCP induced from these A and B points, then the result is transferred back to the original setting using the conditional uniformity property of the Poisson processes.

For the $r \neq 1$ case, the proof is easily extended by letting process A having rate $r\lambda$ and process B having rate λ.

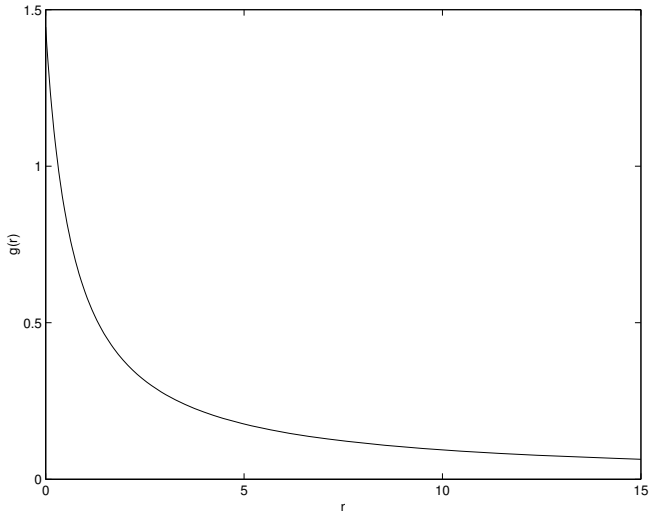

Figure 1: A graph of the limit function $g(r)$, plotted using MATLAB.

Remark: We have found an alternative proof to Theorem 2.2 by using an existing SLLN theorem for *quadrant dependent* random variables [10]. The concept of quadrant dependence was first introduced by E.L. Lehmann in [9], and the limiting theory for quadrant dependent random variables is comprehensively discussed in [11].

3 Strong Law of Large Numbers (SLLN)

In Theorem 2.2, the classes $\mathcal{X}$ and $\mathcal{Y}$ both have uniform distributions. But in real world applications, they usually have different non-uniform distributions. In fact, the principal motive is classifying objects belonging to different classes. Our research has proved an extension to Theorem 2.2 for more general distribution functions in the one dimensional case:

Theorem 3.1. *Suppose $\Omega = \mathbf{R}$. Assume the densities $f_{\mathcal{X}}(x)$ and $f_{\mathcal{Y}}(y)$ are bounded functions with a finite number of discontinuities. If $m/n \to r$, then*

$$\lim_{n\to+\infty} \frac{\Gamma_{n,m}}{m} = \int g\Big(r\frac{f_{\mathcal{Y}}(u)}{f_{\mathcal{X}}(u)}\Big) f_{\mathcal{Y}}(u)\,du \quad a.s. \tag{1}$$

Proof Sketch. Our proof is conducted in two phases: We first consider piece-

wise constant densities $f_{\mathcal{X}}$ and $f_{\mathcal{Y}}$, i.e.

$$f_{\mathcal{X}}(x) = \sum_{l=1}^{k} a_l I_{[c_{l-1},c_l)}(x),$$

and

$$f_{\mathcal{Y}}(y) = \sum_{l=1}^{k} b_l I_{[c_{l-1},c_l)}(y),$$

where $a = c_0 < c_1 < \cdots < c_k = b$. To prove (1) for this type of density function, we divide the CCP into sub-CCP's with conditional uniform distributions for data points in the intervals $[c_{l-1}, c_l]$. In each interval, the ratio between the number of $\mathcal{Y}$ points and $\mathcal{X}$ points is asymptotically $r\frac{f_{\mathcal{Y}}(u)}{f_{\mathcal{X}}(u)}$. From Theorem 2.2, we know that

$$\frac{\text{domination number in } [c_{l-1}, c_l]}{\text{number of } \mathcal{Y} \text{ points in} [c_{l-1}, c_l]}$$

is asymptotically $g\big(r\frac{f_{\mathcal{Y}}(u)}{f_{\mathcal{X}}(u)}\big), u \in [c_{l-1}, c_l]$. By adding up the domination numbers for all the intervals, we get an approximation Γ' to $\Gamma_{n,m}$. We can prove that equation (1) holds if $\frac{\Gamma_{n,m}}{m}$ is replaced by $\frac{\Gamma'}{m}$. Since the difference between Γ' and $\Gamma_{n,m}$ is bounded by $2k$, we conclude that equation (1) is also true.

For the general continuous case, we construct a sequence of piece-wise constant density functions $f_{\mathcal{X},k}$ and $f_{\mathcal{Y},k}$ converging to $f_{\mathcal{X}}$ and $f_{\mathcal{Y}}$, respectively. Based on X_i and Y_j, we define two new sequences of random variables $X_{i,k}$ and $Y_{j,k}$, which are respectively distributed according to $F_{\mathcal{X},k}$ and $F_{\mathcal{Y},k}$. From the first step in our proof, we know that the SLLN is true for the domination number of the CCCD induced by the newly defined points $X_{i,k}$ and $Y_{j,k}$. By using the relation between X_i and $X_{i,k}$, and between Y_i and $Y_{i,k}$, we can argue that the SLLN still holds for the original densities $F_{\mathcal{X}}$ and $F_{\mathcal{Y}}$. □

Intuitively, when class $\mathcal{X}$ and class $\mathcal{Y}$ both have the same distribution pattern, it means that their objects tend to be interspersed. Hence a larger dominating set is needed to distinguish $\mathcal{X}$ from $\mathcal{Y}$ than when they have different distributions. In the following theorem, we have actually proved that equal $f_{\mathcal{X}}$ and $f_{\mathcal{Y}}$ give the maximum limit in the SLLN. This result could be used to construct an asymptotically distribution-free statistical test for equality of densities.

Theorem 3.2. *Under the same assumptions as in Theorem 3.1,*

$$\int g\Big(r\frac{f_{\mathcal{Y}}(u)}{f_{\mathcal{X}}(u)}\Big) f_{\mathcal{Y}}(u)\,du \le g(r)$$

where the equality holds if and only if $f_{\mathcal{X}}(u) = f_{\mathcal{Y}}(u) \quad a.s.$

Proof. Note that $g(r)$ is a convex function, hence $g^*(r) = g(\frac{1}{r})$ is a concave function. Expressing g in terms of g^*, and then applying Jensen's inequality, we have:

$$\begin{aligned}\int g\Big(r\frac{f_{\mathcal{Y}}(u)}{f_{\mathcal{X}}(u)}\Big)f_{\mathcal{Y}}(u)du &= \int g^*\Big(\frac{1}{r}\frac{f_{\mathcal{X}}(u)}{f_{\mathcal{Y}}(u)}\Big)f_{\mathcal{Y}}(u)du \\ &\leq g^*\Big(\int \frac{1}{r}\frac{f_{\mathcal{X}}(u)}{f_{\mathcal{Y}}(u)}f_{\mathcal{Y}}(u)du\Big) \\ &= g^*(\frac{1}{r}) = g(r) \quad \square\end{aligned}$$

4 Asymptotic Variance

Our ultimate goal is to prove the CLT for $\Gamma_{n,m}$. To achieve this, an important first step is to calculate the limiting variance:

Theorem 4.1. *Suppose* $\Omega = \mathbf{R}$, *and* $F_{\mathcal{X}} = F_{\mathcal{Y}} = U[0,1]$. *If* $m/n \to r$, *then*

$$\lim_{n\to\infty} \frac{Var(\Gamma_{n,m})}{m} = v(r) \tag{2}$$

where

$$v(r) \equiv \frac{1536r^5 + 6848r^4 + 11536r^3 + 8836r^2 + 2793r + 180}{9(r+1)^3(4r+3)^4}.$$

Proof Sketch. By decomposing $\Gamma_{n,m}$ into internal and external components, we can write

$$Var(\Gamma_{n,m}) = Var(\alpha_{0,m} + \sum_{j=1}^{m-1} \alpha_{j,m} + \alpha_{m,m}),$$

which can be expressed as a sum of variances and covariances of the components. Thus in order to get (2), we just need to calculate the limiting values of the $Var(\alpha_{j,m})$ and $Cov(\alpha_{j_1,m}, \alpha_{j_2,m})$.

We compute $Var(\alpha_{j,m}) = E({\alpha_{j,m}}^2) - (E(\alpha_{j,m}))^2$ using the identities $E({\alpha_{j,m}}^k) = E[E({\alpha_{j,m}}^k \mid N_{j,m})], k = 1, 2$. Using the conditional probabilities given in Theorem 1.1, $E({\alpha_{j,m}}^k \mid N_{j,m})$ can be expressed in terms of $4^{N_{j,m}}$. Thus we only need to calculate $E(4^{N_{j,m}})$, which can be further written as $E[E(4^{N_{j,m}}|L_{j,m})]$ where $L_{j,m} = Y_{(j+1)} - Y_{(j)}$. This iterated

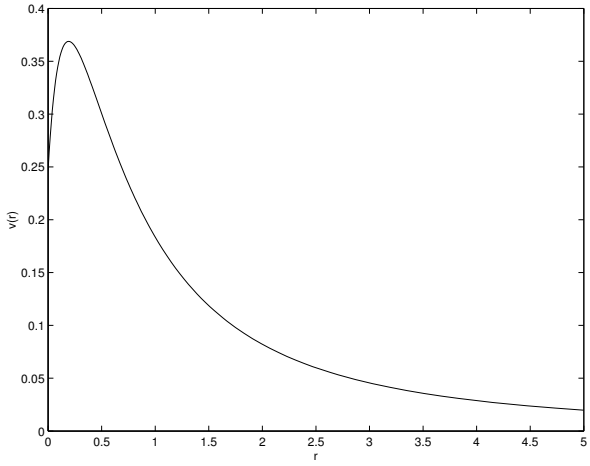

Figure 2: A graph of the limit function $v(r)$, plotted using MATLAB.

expectation can be calculated by using the fact that $N_{j,m}$ is binomial distributed given $L_{j,m}$ and using knowledge of the joint distribution of the order statistics $Y_{(j)}$. Extensive calculations yield

$$Var(\alpha_{j,m}) = \frac{144r^3 + 360r^2 + 237r + 20}{9(r+1)^2(4r+3)^2} + o(1) \quad \text{for } j \in \{1, \cdots, m-1\}.$$

For $Cov(\alpha_{j_1,m}, \alpha_{j_2,m})$, similarly we first compute the conditional expectation $E(\alpha_{j_1,m}\alpha_{j_2,m} \mid N_{j_1,m}, N_{j_2,m})$, then calculate $E(4^{N_{j_1,m}+N_{j_2,m}})$ by conditioning on $L_{j_1,m}$ and $L_{j_2,m}$. A much more intricate analysis is needed to prove that the covariances are of order $O(\frac{1}{m})$. We obtain

$$\begin{aligned} Cov(\alpha_{j_1,m}, \alpha_{j_2,m}) &= \frac{-r^2(2304r^4 + 9984r^3 + 16096r^2 + 11440r + 3025)}{9(r+1)^3(4r+3)^4 m} \\ &\quad + O(\frac{1}{m^2}) \quad \text{for } j_1, j_2 \in \{1, \cdots, m-1\}. \end{aligned}$$

The fact that $\alpha_{j,m}$'s are weakly dependent in the sense that the covariances tend to 0 in the order of $O(\frac{1}{m})$ may be helpful in proving the Central Limit Theorem. □

5 Future Research Directions

We plan to continue investigating the limiting behavior of the domination number, namely, the strong law of large numbers and central limit theorem

for $\Gamma_{n,m}$. This research will continue in two directions: one is to prove the SLLN for the higher dimensional space $\Omega = \mathbf{R}^d, d \geq 2$; the other is to prove the CLT for $\Gamma_{n,m}$ for the case of $\Omega = \mathbf{R}$ and $F_{\mathcal{X}} = F_{\mathcal{Y}} = U[0,1]$, then extend it to more general distribution functions, and finally to higher dimensional spaces. In this effort, we are investigating the use of subadditive process methods for the SLLN, and characteristic function methods, linear quadrant dependence, Stein's method, and methods of Yukich and collaborators [15] for the CLT.

It should be noted that the one dimensional problem is mainly a testing ground for identifying approaches that might be useful in higher dimensions. The real goals are the SLLN and CLT in higher dimensional CCCD problems. One difficulty we encounter in higher dimension situations is how to divide the whole sample space into regions, as we divided the $[0,1]$ into intervals $(Y_{(j)}, Y_{(j+1)})$ in the one dimensional case. Therefore most likely we will not have such a simple identity as $\Gamma_{n,m} = \sum_{j=1}^{m} \alpha_{j,m}$.

In addition to the CLT and SLLN for $\Gamma_{n,m}$, we would like to apply the developing methods to other similar functions of the CCCD besides the domination number. One example is the size of greedy algorithm approximation to the minimum dominating set.

Acknowledgements

This work is partially supported by the Acheson J. Duncan Fund for the Advancement of Research in Statistics.

References

[1] A. Cannon and L. Cowen (2000) Approximation algorithms for the class cover problem, *6th International Symposium on Artificial Intelligence and Mathematics.*

[2] Luc Devroye, Laszlo Gyorfi, and Gabor Lugosi (1996) *A Probabilistic Theory of Pattern Recognition*, Springer.

[3] Richard O. Duda and Peter E. Hart (1973) *Pattern Classification and Scene Analysis*, Wiley-Interscience.

[4] Richard O. Duda, Peter E. Hart, and David G. Stork (2001) *Pattern Classification*, Second edition, Wiley-Interscience.

[5] Jason DeVinney and John C. Wierman (2002) A SLLN for a one-dimensional class cover problem, *Statistics and Probability Letters 59*, 425-435.

[6] Teresa W. Haynes, Stephen T. Hedetniemi, and Peter J. Slater (1998) *Domination in Graphs, Fundamentals*, Marcel Dekker, Inc., New York.

[7] Teresa W. Haynes, Stephen T. Hedetniemi, and Peter J. Slater (1998) *Domination in Graphs, Advanced Topics*, Marcel Dekker, Inc., New York.

[8] S.R. Kulkarni, G. Lugosi, and S.S. Venkatesh (1998) Learning pattern classification – a survey, *IEEE Transactions on Information Theory 44*, 2178-2206.

[9] E.L. Lehmann (1966) Some concepts of dependence, *Annals of Mathematical Statistics 37*,1087-1437.

[10] Przemyslaw Matula (1992) A note on the almost sure convergence of sums of negatively dependent random variables, *Statistics and Probability Letters 15*, 209-213.

[11] Charles M. Newman (1984) Asymptotic independence and limit theorems for positively and negatively dependent random variables, *Inequalities in Statistics and Probability*, IMS Lecture Notes-Monograph Series, Vol. 5, 127-140.

[12] O. Ore (1962) *Theory of Graphs*, Amer. Math. Soc. Publ., Providence, RI.

[13] Carey Priebe, Jason DeVinney, and David Marchette (2001) On the distribution of the domination number for random class cover catch digraphs, *Statistics and Probability Letters 55*, 239-246.

[14] Carey Priebe, David Marchette, Jason DeVinney, and Diego Socolinsky (2002) Classification using class cover catch digraphs, Technical Report #628, Department of Mathematical Sciences, Johns Hopkins University.

[15] Mathew D. Penrose and J.E. Yukich (2001) Central limit theorem for some graphs in computational geometry, *Annals of Applied Probability 11*, 1005-1041.