The Inductive Method, Unbounded Couplings and Stein's Method

Jay Bartoff, Larry Goldstein
University of Southern California
[arxiv:1005.4390] [arxiv:1202.0909]

July 11^{th}, 2012

Stein's Method

Stein's method is a powerful technique for obtaining bounds of optimal rates in distributional approximations.

Works for many target distributions: Normal, Poisson, ...

Works in the presence of dependence.

Applications include: molecular sequence analysis, permutation tests, statistical physics ...

Some applications of the method require a coupling of Y_n to a Y_n' possessing certain inconvenient properties, e.g. boundedness.

Stein's Method

Stein's method is a powerful technique for obtaining bounds of optimal rates in distributional approximations.

Works for many target distributions: Normal, Poisson, ...

Works in the presence of dependence.

Applications include: molecular sequence analysis, permutation tests, statistical physics ...

Some applications of the method require a coupling of Y_n to a Y_n' possessing certain inconvenient properties, e.g. boundedness.

Stein's Method

Stein's method is a powerful technique for obtaining bounds of optimal rates in distributional approximations.

Works for many target distributions: Normal, Poisson, . . .

Works in the presence of dependence.

Applications include: molecular sequence analysis, permutation tests, statistical physics . . .

Some applications of the method require a coupling of Y_n to a Y_n' possessing certain inconvenient properties, e.g. boundedness.

Stein's Method

Stein's method is a powerful technique for obtaining bounds of optimal rates in distributional approximations.

Works for many target distributions: Normal, Poisson, . . .

Works in the presence of dependence.

Applications include: molecular sequence analysis, permutation tests, statistical physics . . .

Some applications of the method require a coupling of Y_n to a Y_n' possessing certain inconvenient properties, e.g. boundedness.

Stein's Method

Stein's method is a powerful technique for obtaining bounds of optimal rates in distributional approximations.

Works for many target distributions: Normal, Poisson, . . .

Works in the presence of dependence.

Applications include: molecular sequence analysis, permutation tests, statistical physics . . .

Some applications of the method require a coupling of Y_n to a Y_n' possessing certain inconvenient properties, e.g. boundedness.

The Nutshell

For the normal approximation of Y_n, situations where bounded couplings exist are the lucky ones.

Some unbounded couplings can be handled by showing that a certain configuration V_n reduced from the one that gives Y_n, conditioned on J_n, is the same as the original problem on a slightly smaller scale:

$$\mathcal{L}_\theta(V_n|J_n) = \mathcal{L}_{\psi_{n,\theta}}(Y_{n-L_n})$$

Based on ideas from Bolthausen (1985) for the Combinatorial Central Limit Theorem, $A \in \mathbb{R}^{n \times n}, \pi \in \mathcal{S}_n$,

$$Y = \sum_{i=1}^{n} a_{i,\pi(i)}$$

The Nutshell

For the normal approximation of Y_n, situations where bounded couplings exist are the lucky ones.

Some unbounded couplings can be handled by showing that a certain configuration V_n reduced from the one that gives Y_n, conditioned on J_n, is the same as the original problem on a slightly smaller scale:

$$\mathcal{L}_\theta(V_n|J_n) = \mathcal{L}_{\psi_{n,\theta}}(Y_{n-L_n})$$

Based on ideas from Bolthausen (1985) for the Combinatorial Central Limit Theorem, $A \in \mathbb{R}^{n\times n}, \pi \in \mathcal{S}_n$,

$$Y = \sum_{i=1}^{n} a_{i,\pi(i)}$$

The Nutshell

For the normal approximation of Y_n, situations where bounded couplings exist are the lucky ones.

Some unbounded couplings can be handled by showing that a certain configuration V_n reduced from the one that gives Y_n, conditioned on J_n, is the same as the original problem on a slightly smaller scale:

$$\mathcal{L}_\theta(V_n|J_n) = \mathcal{L}_{\psi_{n,\theta}}(Y_{n-L_n})$$

Based on ideas from Bolthausen (1985) for the Combinatorial Central Limit Theorem, $A \in \mathbb{R}^{n\times n}, \pi \in \mathcal{S}_n$,

$$Y = \sum_{i=1}^{n} a_{i,\pi(i)}$$

Examples: Graph and Occupancy Model

Erdős-Rényi graph: Let Y_n be the number of degree d vertices in the random graph with n vertices, where for each distinct pair of vertices u, v the indicators $E_{\{u,v\}}$ of the existence of an edge between u and v are independent Bernoulli random variables with success probability $\theta/(n-1)$.

Occupancy: Let Y_n be the number of bins having d balls when distributing n balls over m urns. Species counting problem, and author attribution literature: how many words did Shakespeare know?

In both cases, Y_n is the sum of (globally) dependent variables.

Examples: Graph and Occupancy Model

Erdős-Rényi graph: Let Y_n be the number of degree d vertices in the random graph with n vertices, where for each distinct pair of vertices u, v the indicators $E_{\{u,v\}}$ of the existence of an edge between u and v are independent Bernoulli random variables with success probability $\theta/(n-1)$.

Occupancy: Let Y_n be the number of bins having d balls when distributing n balls over m urns. Species counting problem, and author attribution literature: how many words did Shakespeare know?

In both cases, Y_n is the sum of (globally) dependent variables.

Examples: Graph and Occupancy Model

Erdős-Rényi graph: Let Y_n be the number of degree d vertices in the random graph with n vertices, where for each distinct pair of vertices u, v the indicators $E_{\{u,v\}}$ of the existence of an edge between u and v are independent Bernoulli random variables with success probability $\theta/(n-1)$.

Occupancy: Let Y_n be the number of bins having d balls when distributing n balls over m urns. Species counting problem, and author attribution literature: how many words did Shakespeare know?

In both cases, Y_n is the sum of (globally) dependent variables.

The Nutshell

Removing a randomly chosen vertex I from the Erdős-Rényi random graph leaves a graph of the same type on a vertex set one size smaller.

Removing a randomly chosen bin in the uniform occupancy problem leaves a uniform occupancy problem with one fewer urn, and a Binomial number fewer of balls.

The Nutshell

Removing a randomly chosen vertex I from the Erdős-Rényi random graph leaves a graph of the same type on a vertex set one size smaller.

Removing a randomly chosen bin in the uniform occupancy problem leaves a uniform occupancy problem with one fewer urn, and a Binomial number fewer of balls.

Stein Equation for Normal

The random variable Z is $\mathcal{N}(0,1)$ if and only if

$$E[Zf(Z)] = E[f'(Z)]$$

for all absolutely continuous functions f for which these expectations exist.

If W has a distribution close to $\mathcal{N}(0,1)$ then

$$E[f'(W) - Wf(W)] \quad \text{will be close to zero.}$$

Stein equation for test function h,

$$f'(w) - wf(w) = h(w) - Eh(Z).$$

Stein Equation for Normal

The random variable Z is $\mathcal{N}(0,1)$ if and only if

$$E[Zf(Z)] = E[f'(Z)]$$

for all absolutely continuous functions f for which these expectations exist.
If W has a distribution close to $\mathcal{N}(0,1)$ then

$$E[f'(W) - Wf(W)] \quad \text{will be close to zero.}$$

Stein equation for test function h,

$$f'(w) - wf(w) = h(w) - Eh(Z).$$

Stein Equation for Normal

The random variable Z is $\mathcal{N}(0,1)$ if and only if

$$E[Zf(Z)] = E[f'(Z)]$$

for all absolutely continuous functions f for which these expectations exist.
If W has a distribution close to $\mathcal{N}(0,1)$ then

$$E[f'(W) - Wf(W)] \quad \text{will be close to zero.}$$

Stein equation for test function h,

$$f'(w) - wf(w) = h(w) - Eh(Z).$$

Stein Equation

Given h, solve for f in

$$f'(w) - wf(w) = h(w) - Eh(Z).$$

Substitute W for w and take expectation to yield

$$E[f'(W) - Wf(W)] = Eh(W) - Eh(Z).$$

Note expectation on the left involves only one random variable.

Stein Equation

Given h, solve for f in

$$f'(w) - wf(w) = h(w) - Eh(Z).$$

Substitute W for w and take expectation to yield

$$E[f'(W) - Wf(W)] = Eh(W) - Eh(Z).$$

Note expectation on the left involves only one random variable.

Stein Equation

Given h, solve for f in

$$f'(w) - wf(w) = h(w) - Eh(Z).$$

Substitute W for w and take expectation to yield

$$E[f'(W) - Wf(W)] = Eh(W) - Eh(Z).$$

Note expectation on the left involves only one random variable.

Coupling

To handle the term $E[Wf(W)]$:

Coupling

To handle the term $E[Wf(W)]$:

1. Exchangeable pairs

Coupling

To handle the term $E[Wf(W)]$:

1. Exchangeable pairs
2. Size Bias coupling

Coupling

To handle the term $E[Wf(W)]$:

1. Exchangeable pairs
2. Size Bias coupling
3. Zero Bias Coupling

Coupling

To handle the term $E[Wf(W)]$:

1. Exchangeable pairs
2. Size Bias coupling
3. Zero Bias Coupling
4. G coupling

Size Bias Coupling

For a nonnegative random variable Y with finite, nonzero mean μ, we say that Y^s has the Y-size bias distribution if

$$E[Yf(Y)] = \mu E[f(Y^s)]$$

for all functions f for which these expectations exist.

When $Y = \sum_{i=1}^n X_i$, a sum of nonnegative random variables, we may construct Y^s by choosing a summand proportional to its mean, i.e. $P(I = i) = EX_i/\mu$, replacing X_I by X_I^s, and adjusting the remaining variables to have their conditional distribution given the new value.

If X is a non-trivial Bernoulli random variable, $X^s = 1$.

Size Bias Coupling

For a nonnegative random variable Y with finite, nonzero mean μ, we say that Y^s has the Y-size bias distribution if

$$E[Yf(Y)] = \mu E[f(Y^s)]$$

for all functions f for which these expectations exist.

When $Y = \sum_{i=1}^{n} X_i$, a sum of nonnegative random variables, we may construct Y^s by choosing a summand proportional to its mean, i.e. $P(I = i) = EX_i/\mu$, replacing X_I by X_I^s, and adjusting the remaining variables to have their conditional distribution given the new value.

If X is a non-trivial Bernoulli random variable, $X^s = 1$.

Size Bias Coupling

For a nonnegative random variable Y with finite, nonzero mean μ, we say that Y^s has the Y-size bias distribution if

$$E[Yf(Y)] = \mu E[f(Y^s)]$$

for all functions f for which these expectations exist.

When $Y = \sum_{i=1}^{n} X_i$, a sum of nonnegative random variables, we may construct Y^s by choosing a summand proportional to its mean, i.e. $P(I = i) = EX_i/\mu$, replacing X_I by X_I^s, and adjusting the remaining variables to have their conditional distribution given the new value.

If X is a non-trivial Bernoulli random variable, $X^s = 1$.

Distances

Distance $d(\mathcal{L}(W), \mathcal{L}(Z))$ obtained by Stein's method depends on class of functions to which test function h belongs.

1. Wasserstein distance

$$d_1(\mathcal{L}(W), \mathcal{L}(Z)) = \int_{-\infty}^{\infty} |F(x) - G(x)| dx,$$

consider h Lipschitz, satisfying $|h(y) - h(x)| \leq |y - x|$.

2. Kolmogorov distance

$$d_\infty(\mathcal{L}(W), \mathcal{L}(Z)) = \sup_{z \in \mathbb{R}} |F(x) - G(x)|,$$

consider $h(x) = \mathbf{1}(x \leq z)$ indicator functions.

Distances

Distance $d(\mathcal{L}(W), \mathcal{L}(Z))$ obtained by Stein's method depends on class of functions to which test function h belongs.

1. Wasserstein distance

$$d_1(\mathcal{L}(W), \mathcal{L}(Z)) = \int_{-\infty}^{\infty} |F(x) - G(x)| dx,$$

consider h Lipschitz, satisfying $|h(y) - h(x)| \leq |y - x|$.

2. Kolmogorov distance

$$d_\infty(\mathcal{L}(W), \mathcal{L}(Z)) = \sup_{z \in \mathbb{R}} |F(x) - G(x)|,$$

consider $h(x) = \mathbf{1}(x \leq z)$ indicator functions.

Distances

Distance $d(\mathcal{L}(W), \mathcal{L}(Z))$ obtained by Stein's method depends on class of functions to which test function h belongs.

1. Wasserstein distance

$$d_1(\mathcal{L}(W), \mathcal{L}(Z)) = \int_{-\infty}^{\infty} |F(x) - G(x)| dx,$$

   consider h Lipschitz, satisfying $|h(y) - h(x)| \leq |y - x|$.

2. Kolmogorov distance

$$d_\infty(\mathcal{L}(W), \mathcal{L}(Z)) = \sup_{z \in \mathbb{R}} |F(x) - G(x)|,$$

   consider $h(x) = \mathbf{1}(x \leq z)$ indicator functions.

$$f'(w) - wf(w) = h(w) - Eh(Z)$$

Wasserstein d_1 distance easier to handle than Kolmogorov d_∞. For instance, for d_1 we have

$$|f'(w) - f'(w+t)| \leq t||f''|| \quad \text{and} \quad ||f''|| \leq 2||h'||.$$

For d_∞ test functions the solution f does not posses two derivatives, and h must be smoothed. Bothausen's inequality:

$$|f'(w) - f'(w+t)| \leq |t| \left(1 + |w| + \frac{1}{\lambda} \int_0^1 1_{[z,z+\lambda]}(w+ut)du\right)$$

where $\lambda > 0$ is amount of smoothing performed on $h(w) = \mathbf{1}(w \leq z)$.

Term in integral yields probability of small interval, not of 'moment' type, but easy to bound when coupling is bounded.

$$f'(w) - wf(w) = h(w) - Eh(Z)$$

Wasserstein d_1 distance easier to handle than Kolmogorov d_∞. For instance, for d_1 we have

$$|f'(w) - f'(w+t)| \le t||f''|| \quad \text{and} \quad ||f''|| \le 2||h'||.$$

For d_∞ test functions the solution f does not posses two derivatives, and h must be smoothed. Bothausen's inequality:

$$|f'(w) - f'(w+t)| \le |t| \left(1 + |w| + \frac{1}{\lambda}\int_0^1 1_{[z,z+\lambda]}(w+ut)du\right)$$

where $\lambda > 0$ is amount of smoothing performed on $h(w) = \mathbf{1}(w \le z)$.

Term in integral yields probability of small interval, not of 'moment' type, but easy to bound when coupling is bounded.

$$f'(w) - wf(w) = h(w) - Eh(Z)$$

Wasserstein d_1 distance easier to handle than Kolmogorov d_∞. For instance, for d_1 we have

$$|f'(w) - f'(w+t)| \leq t||f''|| \quad \text{and} \quad ||f''|| \leq 2||h'||.$$

For d_∞ test functions the solution f does not posses two derivatives, and h must be smoothed. Bothausen's inequality:

$$|f'(w) - f'(w+t)| \leq |t| \left(1 + |w| + \frac{1}{\lambda}\int_0^1 1_{[z,z+\lambda]}(w+ut)du\right)$$

where $\lambda > 0$ is amount of smoothing performed on $h(w) = \mathbf{1}(w \leq z)$.

Term in integral yields probability of small interval, not of 'moment' type, but easy to bound when coupling is bounded.

Unbounded Coupling: Erdős-Rényi Graph G_n

To size bias the number Y_n of vertices of degree d in G_n, select a vertex, say I, uniformly, and add or subtract edges uniformly, as needed so that I has degree d. The number Y_n^s of degree d vertices in the graph so formed has the Y_n-size biased distribution.

The coupling of Y_n and Y_n^s is not bounded. It is possible that the chosen vertex I had many edges, and all but d of them were then removed in order for I to attain degree d.

V_n: Removing I and all its incident edges leaves an Erdős-Rényi random graph on $n-1$ vertices.

J_n:The vertex I and its degree $D(I)$ are independent of the reduced graph, yet

$$|Y_n^s - Y_n| \leq 1 + |d - D(I)|.$$

Unbounded Coupling: Erdős-Rényi Graph G_n

To size bias the number Y_n of vertices of degree d in G_n, select a vertex, say I, uniformly, and add or subtract edges uniformly, as needed so that I has degree d. The number Y_n^s of degree d vertices in the graph so formed has the Y_n-size biased distribution.

The coupling of Y_n and Y_n^s is not bounded. It is possible that the chosen vertex I had many edges, and all but d of them were then removed in order for I to attain degree d.

V_n: Removing I and all its incident edges leaves an Erdős-Rényi random graph on $n-1$ vertices.

J_n:The vertex I and its degree $D(I)$ are independent of the reduced graph, yet

$$|Y_n^s - Y_n| \leq 1 + |d - D(I)|.$$

Unbounded Coupling: Erdős-Rényi Graph G_n

To size bias the number Y_n of vertices of degree d in G_n, select a vertex, say I, uniformly, and add or subtract edges uniformly, as needed so that I has degree d. The number Y_n^s of degree d vertices in the graph so formed has the Y_n-size biased distribution.

The coupling of Y_n and Y_n^s is not bounded. It is possible that the chosen vertex I had many edges, and all but d of them were then removed in order for I to attain degree d.

V_n: Removing I and all its incident edges leaves an Erdős-Rényi random graph on $n-1$ vertices.

J_n:The vertex I and its degree $D(I)$ are independent of the reduced graph, yet

$$|Y_n^s - Y_n| \leq 1 + |d - D(I)|.$$

Unbounded Coupling: Erdős-Rényi Graph G_n

To size bias the number Y_n of vertices of degree d in G_n, select a vertex, say I, uniformly, and add or subtract edges uniformly, as needed so that I has degree d. The number Y_n^s of degree d vertices in the graph so formed has the Y_n-size biased distribution.

The coupling of Y_n and Y_n^s is not bounded. It is possible that the chosen vertex I had many edges, and all but d of them were then removed in order for I to attain degree d.

V_n: Removing I and all its incident edges leaves an Erdős-Rényi random graph on $n-1$ vertices.

J_n:The vertex I and its degree $D(I)$ are independent of the reduced graph, yet

$$|Y_n^s - Y_n| \leq 1 + |d - D(I)|.$$

Unbounded Coupling: Multinomial Occupancy $\mathbf{M}_n$

Similarly, form a configuration where a randomly selected urn I contains exactly d balls by removing or adding balls uniformly from all other urns. This coupling has the same unboundedness property as the previous.

Inductive step: Removing urn I and it contents of $M(I)$ balls leaves occupancy problem with $n - M(I)$ balls and $m - 1$ urns.

The multinomial occupancy problem is substantially more difficult than graph degree. The number of balls in the one, unlike the number of edges in the other, must remain fixed: compare removing a vertex connected to all other vertices, and removing an urn containing all balls.

Also has a substantially more difficult 'variance calculation.'

Unbounded Coupling: Multinomial Occupancy $\mathbf{M}_n$

Similarly, form a configuration where a randomly selected urn I contains exactly d balls by removing or adding balls uniformly from all other urns. This coupling has the same unboundedness property as the previous.

Inductive step: Removing urn I and it contents of $M(I)$ balls leaves occupancy problem with $n - M(I)$ balls and $m - 1$ urns.

The multinomial occupancy problem is substantially more difficult than graph degree. The number of balls in the one, unlike the number of edges in the other, must remain fixed: compare removing a vertex connected to all other vertices, and removing an urn containing all balls.

Also has a substantially more difficult 'variance calculation.'

Unbounded Coupling: Multinomial Occupancy $\mathbf{M}_n$

Similarly, form a configuration where a randomly selected urn I contains exactly d balls by removing or adding balls uniformly from all other urns. This coupling has the same unboundedness property as the previous.

Inductive step: Removing urn I and it contents of $M(I)$ balls leaves occupancy problem with $n - M(I)$ balls and $m - 1$ urns.

The multinomial occupancy problem is substantially more difficult than graph degree. The number of balls in the one, unlike the number of edges in the other, must remain fixed: compare removing a vertex connected to all other vertices, and removing an urn containing all balls.

Also has a substantially more difficult 'variance calculation.'

Unbounded Coupling: Multinomial Occupancy $\mathbf{M}_n$

Similarly, form a configuration where a randomly selected urn I contains exactly d balls by removing or adding balls uniformly from all other urns. This coupling has the same unboundedness property as the previous.

Inductive step: Removing urn I and it contents of $M(I)$ balls leaves occupancy problem with $n - M(I)$ balls and $m - 1$ urns.

The multinomial occupancy problem is substantially more difficult than graph degree. The number of balls in the one, unlike the number of edges in the other, must remain fixed: compare removing a vertex connected to all other vertices, and removing an urn containing all balls.

Also has a substantially more difficult 'variance calculation.'

Inductive Step: Main Idea

Conditioning on J_n leaves smaller problem; need that the bound K_n on $|Y_n^s - Y_n|$ is function of J_n.

Can take expectation by first conditioning. Conditional expectation pulls past bound K_n, conditional expectation of term with W_n, standardized value of Y_n, can be expressed by distance to normal for smaller problem.

$$E\left(K_n \int_{-K_n/\sigma_n}^{K_n/\sigma_n} \int_0^1 \mathbf{1}_{[z,z+\lambda]}(W_n + ut)dudt\right).$$

Obtain recursion for bound to normality.

Inductive Step: Main Idea

Conditioning on J_n leaves smaller problem; need that the bound K_n on $|Y_n^s - Y_n|$ is function of J_n.

Can take expectation by first conditioning. Conditional expectation pulls past bound K_n, conditional expectation of term with W_n, standardized value of Y_n, can be expressed by distance to normal for smaller problem.

$$E\left(K_n \int_{-K_n/\sigma_n}^{K_n/\sigma_n} \int_0^1 \mathbf{1}_{[z,z+\lambda]}(W_n + ut)dudt\right).$$

Obtain recursion for bound to normality.

Inductive Step: Main Idea

Conditioning on J_n leaves smaller problem; need that the bound K_n on $|Y_n^s - Y_n|$ is function of J_n.

Can take expectation by first conditioning. Conditional expectation pulls past bound K_n, conditional expectation of term with W_n, standardized value of Y_n, can be expressed by distance to normal for smaller problem.

$$E\left(K_n \int_{-K_n/\sigma_n}^{K_n/\sigma_n} \int_0^1 \mathbf{1}_{[z,z+\lambda]}(W_n + ut)dudt\right).$$

Obtain recursion for bound to normality.

Graph Degree and Occupancy: L^∞ History

Graph Degree: Kordecki (1990) case $d = 0$, Neammanee and Suntadkarn (2009) (within ϵ) .

Graph Degree and Occupancy: L^∞ History

Graph Degree: Kordecki (1990) case $d = 0$, Neammanee and Suntadkarn (2009) (within ϵ) .

Occupancy: Englund (1981) $d = 0$, Penrose (2009) $d = 1$.

Graph Degree and Occupancy: L^∞ History

Graph Degree: Kordecki (1990) case $d = 0$, Neammanee and Suntadkarn (2009) (within ϵ) .

Occupancy: Englund (1981) $d = 0$, Penrose (2009) $d = 1$.

Chen and Rollin (2010) (quite general situations), within poly log.

Graph Degree and Occupancy: L^∞ History

Graph Degree: Kordecki (1990) case $d = 0$, Neammanee and Suntadkarn (2009) (within ϵ) .

Occupancy: Englund (1981) $d = 0$, Penrose (2009) $d = 1$.

Chen and Rollin (2010) (quite general situations), within poly log.

Inductive method: Applied for the graph and occupancy situations obtains the correct rate for all $d \in \{0, 1, \ldots, \}$.

Cast of Main Characters: Occupancy

1. Y_n: How many of the m urns contain d balls when n balls are uniformly distributed.
2. Y_n^s: Number of m urns with d balls after adding or removing balls from urn I_n, randomly selected, so that it has occupancy d.
3. J_n: Identity I_n of selected urn and its occupancy $M_n(I_n)$.
4. K_n: Bound $1 + |M_n(I_n) - d|$ the absolute number $|Y_n^s - Y_n|$ of urns whose occupancy is affected by adding or removing balls when forming Y_n^s.
5. V_n: Number of urns other than I_n with occupancy d.

$$\mathcal{L}_m(V_n|J_n) = \mathcal{L}_{m-1}(Y_{n-M_n(I_n)})$$

Further Directions

Presently only the size bias coupling is handled. The inductive method can also be used for exchangeable pairs, and the more general G-coupling framework of Chen and Roellin, as well as with the zero bias coupling.

Should yield results on counts more general than number of urns with exactly d balls, eg. numbers of urns with more than d balls, or number of balls in excess of d summed over all urns.

Also can be applied to more general statistics on the Erdős-Rényi graph.

Further Directions

Presently only the size bias coupling is handled. The inductive method can also be used for exchangeable pairs, and the more general G-coupling framework of Chen and Roellin, as well as with the zero bias coupling.

Should yield results on counts more general than number of urns with exactly d balls, eg. numbers of urns with more than d balls, or number of balls in excess of d summed over all urns.

Also can be applied to more general statistics on the Erdős-Rényi graph.

Further Directions

Presently only the size bias coupling is handled. The inductive method can also be used for exchangeable pairs, and the more general G-coupling framework of Chen and Roellin, as well as with the zero bias coupling.

Should yield results on counts more general than number of urns with exactly d balls, eg. numbers of urns with more than d balls, or number of balls in excess of d summed over all urns.

Also can be applied to more general statistics on the Erdős-Rényi graph.