L^1 Bounds in Normal Approximation

Larry Goldstein

University of Southern California

Zero Bias Transformation

Motivated by Stein's Lemma

$$E[Zf(Z)] = \sigma^2 Ef'(Z) \quad \text{if and only if } Z \sim \mathcal{N}(0, \sigma^2).$$

For every Y with mean zero and variance σ^2, there exists a unique law for Y^* such that

$$E[Yf(Y)] = \sigma^2 Ef'(Y^*)$$

for all smooth f. (Goldstein and Reinert, 1997)

Stein's Lemma becomes: The distributional transformation $Y \to Y^*$ has $\mathcal{N}(0, \sigma^2)$ as its unique fixed point.

L^1 Bound

Principle: If $\mathcal{L}(Y)$ and $\mathcal{L}(Y^*)$ are close, then $\mathcal{L}(Y)$ is close to being a fixed point of the transformation, so is close to the unique fixed point, the normal.

In L^1 (Wasserstein, Dudley, Fortet-Mourier or Kantarovich) distance, this principle is evidenced by

$$||Y - Z|| \leq 2||Y^* - Y||.$$

The right hand side can sometimes be conveniently computed by coupling.

Examples and their Couplings

- Independent Sums
  -
- Combinatorial Central Limit Theorem
  -
- Cone Measure on the Sphere
  -

Examples and their Couplings

- Independent Sums
  - Replace One
- Combinatorial Central Limit Theorem
  -
- Cone Measure on the Sphere
  -

Examples and their Couplings

- Independent Sums
  - Replace One
- Combinatorial Central Limit Theorem
  - Exchangeable Pair
- Cone Measure on the Sphere
  -

Examples and their Couplings

- Independent Sums
  - Replace One
- Combinatorial Central Limit Theorem
  - Exchangeable Pair
- Cone Measure on the Sphere
  - Square Biasing Under Symmetry

Zero and Size Biasing

Let Y be nonnegative with finite mean μ. Recall Y^s has the Y-sized biased distribution if

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

for all smooth f.

Parallel to the zero bias transform: mean is replaced by variance, f is replaced by f'.

Both are special cases of 'distributional biasing' of the form

$$E[P(Y)f(Y)] = \alpha Ef^{(m)}(Y^{(P)}).$$

(Goldstein and Reinert, 2005)

Zero (Size) Biasing an Independent Sum

To zero (size) bias a sum

$$Y = \sum_{i=1}^{n} X_i$$

of mean zero (non-negative) independent variables, pick one proportional to its variance (mean) and replace with biased version.

Answers intuitively the basic question of when a sum Y is close to normal.

Zero Bias CLT Rationale

If Y is the sum of comparable, independent, mean zero variables then Y^* differs from Y by only one summand.

Hence Y^* is close to Y, so Y is nearly a fixed point of the zero bias transformation, and hence close to normal.

Sums of i.i.d. variables

When $W = n^{-1/2} \sum X_i$ for $X, X_1, \ldots, X_n$ i.i.d. with mean zero, variance 1, and distribution function G,

$$W^* - W = n^{-1/2}(X_I^* - X_I).$$

Hence by previous bound

$$||F - \Phi||_1 \leq \frac{2}{\sqrt{n}} ||G^* - G||.$$

The distribution function G^* of X^* is given explicitly by

$$G^*(x) = E[X(X - x)\mathbf{1}(X \leq x)].$$

Distribution Specific Constant

Bernoulli* is Uniform, and

$$||F - \Phi|| \leq \frac{E|X_1|^3}{\sqrt{n}}, \quad \text{so } c = 1.$$

For $\mathcal{U}[-\sqrt{3}, \sqrt{3}]$,

$$G^*(x) = -\frac{\sqrt{3}x^3}{36} + \frac{\sqrt{3}x}{4} + \frac{1}{2} \quad \text{for} \quad x \in [-\sqrt{3}, \sqrt{3}],$$

yields

$$||F - \Phi|| \leq \frac{\sqrt{3}}{4\sqrt{n}} = \frac{E|X_1|^3}{3\sqrt{n}}, \quad \text{so } c = 1/3$$

Combinatorial Central Limit Theorem

Given $n \times n$ real matrix A, obtain bounds to the normal approximation for

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

where π is uniform on $\mathcal{S}_n$.

(Hoeffding, Chen and Ho, Bolthausen, von Bahr)

Uniform distribution $\mathcal{U}(\mathcal{S}_n)$

1. Simple random sampling is a special case of the special case where $a_{ij} = c_i d_j$.

2. Measure of uniformity of π, letting $a_{ij} = |i - j|$ gives $a_{i,\pi(i)} = |i - \pi(i)|$. $Y = 0$ on id, how far from zero is Y when π is uniform?

3. Distribution of permutation test statistics.

Exchangeable Pair

Given Y, construct Y' such that (Y, Y') is exchangeable, and

$$E(Y'|Y) = (1 - \lambda)Y, \quad \lambda \in (0, 1).$$

Should we exchange it?

Computation of certain bounds when applying the exchangeable pair may require the (sometimes difficult) calculation of quantities such as

$$\sqrt{\mathsf{Var}\{E[(Y' - Y)^2|Y]\}}.$$

But such is not required for the computation of $E|Y^* - Y|$ for L^1 bounds.

Nevertheless, one method of obtaining a zero bias coupling involves the pair.

Zero Bias from Exchangeable Pair

Let (Y', Y'') be an exchangeable pair with joint distribution F satisfying $E(Y''|Y') = (1-\lambda)Y'$, $\mathsf{Var}(Y') = \sigma^2$. Let

$$dF^\dagger(y^\dagger, y^\ddagger) = \frac{(y^\dagger - y^\ddagger)^2}{2\lambda\sigma^2} dF(y^\dagger, y^\ddagger),$$

and $U \sim \mathcal{U}[0,1]$ independent. Then when $(Y^\dagger, Y^\ddagger) \sim F^\dagger$,

$$Y^* = UY^\dagger + (1-U)Y^\ddagger$$

has the Y^*-zero bias distribution.

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

Given $\pi' \sim \mathcal{U}(\mathcal{S}_n)$, let τ_{IJ} be the transposition of I and J, chosen distinct and uniformly, and let

$$\pi'' = \pi' \tau_{I,J}.$$

Then with Y'' formed using π'', the pair Y', Y'' is exchangeable,

$$E(Y''|Y') = (1 - \frac{2}{n-1})Y',$$

and

$$Y'' - Y' = a_{I,\pi'(J)} + a_{J,\pi'(I)} - (a_{I,\pi'(I)} + a_{J,\pi'(J)}).$$

Square Difference Bias and Coupling

To form $\pi^\dagger, \pi^\ddagger$, consider $I^\dagger, J^\dagger, K^\dagger, L^\dagger$ with distribution

$$p(i,j,k,l) = \frac{\left[(a_{ik}+a_{jl})-(a_{il}+a_{jk})\right]^2}{4n^2(n-1)\sigma^2}.$$

Now (letting $\pi = \pi'$) set

$$\pi^\dagger = \begin{cases} \pi\tau_{\pi^{-1}(K^\dagger),J^\dagger} & \text{if } L^\dagger = \pi(I^\dagger), K^\dagger \neq \pi(J^\dagger) \\ \pi\tau_{\pi^{-1}(L^\dagger),I^\dagger} & \text{if } L^\dagger \neq \pi(I^\dagger), K^\dagger = \pi(J^\dagger) \\ \pi\tau_{\pi^{-1}(K^\dagger),I^\dagger}\tau_{\pi^{-1}(L^\dagger),J^\dagger} & \text{otherwise,} \end{cases}$$

and $\pi^\ddagger = \pi^\dagger\tau_{I^\dagger,J^\dagger}$.

Aside: Non Uniform π Distribution

May consider distribution constant on conjugacy classes, for example, π uniform over all involutions $\pi^2 = \text{id}$.

In general, we may construct an exchangeable pair for such a π by letting

$$\pi'' = \tau_{IJ} \pi \tau_{IJ},$$

when A is symmetric and the probability of fixed points is zero.

Bound by computing $E|Y^* - Y'|$

Letting

$$a_{\bullet j} = \frac{1}{n}\sum_{i=1}^{n} a_{ij} \quad a_{i\bullet} = \frac{1}{n}\sum_{j=1}^{n} a_{ij} \quad a_{\bullet\bullet} = \frac{1}{n^2}\sum_{i,j=1}^{n} a_{ij}$$

$$a_3 = \sum_{i,j=1}^{n} |a_{ij} - a_{i\bullet} - a_{\bullet j} + a_{\bullet\bullet}|^3,$$

with $\sigma^2 = \mathrm{Var}(Y')$,

$$||F - \Phi||_1 \leq \frac{a_3}{(n-1)\sigma^3}\left(16 + \frac{56}{n-1} + \frac{8}{(n-1)^2}\right).$$

Cone Measure $\mathcal{C}_p^n$ in $\mathbb{R}^n$

$$S(\ell_p^n) = \{\mathbf{x} : \sum_{i=1}^n |x_i|^p = 1\}, \quad B(\ell_p^n) = \{\mathbf{x} : \sum_{i=1}^n |x_i|^p \le 1\}$$

With μ^n Lebesgue measure in $\mathbf{R}^n$, for $A \subset S(\ell_p^n)$ and

$$[0,1]A = \{ta : a \in A, 0 \le t \le 1\}$$

let

$$\mathcal{C}_p^n(A) = \frac{\mu^n([0,1]A)}{\mu^n(B(\ell_p^n))}.$$

Cone Measure $\mathcal{C}_p^n$ in $\mathbb{R}^n$

Special cases

1. $p = 1$: Uniform distribution over the simplex

$$\sum_{i=1}^{n} |x_i| = 1.$$

2. $p = 2$: Uniform distribution over the sphere

$$\sum_{i=1}^{n} x_i^2 = 1.$$

Projection

For $\mathbf{X} \sim \mathcal{C}_p^n$ for some $p > 0$ and $\boldsymbol{\theta} \in \mathbb{R}^n$ a unit vector, consider the projection

$$Y = \boldsymbol{\theta} \cdot \mathbf{X}.$$

When $\boldsymbol{\theta} = n^{-1/2}(1, 1, \ldots, 1)$, then $Y = n^{-1/2} \sum_{i=1}^n X_i$.

Diaconis and Freedman: for $p = 2$ considered total variation bounds

Meckes and Meckes: for random vectors with symmetries in general, considered supremum and total variation bounds

Zero Bias Construction for Coordinate Symmetric Vectors

$$(Y_1, \ldots, Y_n) =_d (e_1 Y_1, \ldots, e_n Y_n), \quad \forall e_i \in \{-1, 1\}.$$

Since $Y_i =_d -Y_i$ and $(Y_i, Y_j) =_d (Y_i, -Y_j)$ when $i \neq j$, when second moments exist we have

$$EY_i = 0 \quad \text{and} \quad \text{Cov}(Y_i, Y_j) = 0.$$

With $\sigma_i^2 = \text{Var}(Y_i)$, the construction depends on the *square bias* distributions in direction i,

$$EY_i f(\mathbf{Y}) = \sigma_i^2 E f(\mathbf{Y}^i) \quad \text{or} \quad dF^i(\mathbf{y}) = \frac{y_i^2}{\sigma_i^2} dF(\mathbf{y}).$$

Square Bias Construction

Let $Y = \sum_{i=1}^{n} Y_i$, I an independent random index with distribution $P(I = i) = \frac{\sigma_i^2}{\sum_{j=1}^{n} \sigma_j^2}$ and $U \sim \mathcal{U}[-1, 1]$ independent of all other variables. Then

$$Y^* = UY_I^I + \sum_{j \neq I} Y_j^I.$$

Generalizes the 'replace one' construction for independent variables given earlier.

For coupling under dependence, pick i according to I, generate y_i^i, then 'adjust' $Y_j, j \neq i$ according to the conditional distribution given $Y_i = y_i^i$.

Coupling for Cone Measure

If $\{G_j, \epsilon_j, j = 1, \ldots, n\}$ are independent variables with $G_j \sim \Gamma(1/p, 1)$, $G_{1,n} = \sum_{i=1}^n G_i$ and $\epsilon_j \in \{-1, 1\}$ equally likely, then

$$\mathbf{X} = \left(\epsilon_1 (\frac{G_1}{G_{1,n}})^{1/p}, \ldots, \epsilon_n (\frac{G_n}{G_{1,n}})^{1/p} \right) \sim \mathcal{C}_p^n.$$

Square Bias in given Direction

With $G'_j \sim \Gamma(2/p, 1)$ independent,

$$X_i^i = \epsilon_i \left(\frac{G_i + G'_i}{G_{1,n} + G'_i} \right)^{1/p}$$

has the X_i square bias distribution, and the vector with components

$$\left\{ \begin{array}{cc} \left(\frac{1-|X_i^i|^p}{1-|X_i|^p} \right)^{1/p} X_j & j \neq i \\ X_i^i & j = i \end{array} \right.$$

has the $\mathbf{X}$ distribution square biased in direction i.

L^1 Cone Measure Bound to Normal

Let $\mathbf{X}$ have cone measure $\mathcal{C}_p^n$ on the sphere $S(\ell_p^n)$ for some $p > 0$, and let

$$Y = \sum_{i=1}^{n} \theta_i X_i$$

be the one dimensional projection of $\mathbf{X}$ along the direction $\boldsymbol{\theta} \in \mathbf{R}^n$ with $||\boldsymbol{\theta}|| = 1$. Then with $\sigma_{n,p}^2 = \mathsf{Var}(X_1)$ and $m_{n,p} = E|X_1^1|$,

$$||F - \Phi|| \leq 3\left(\frac{m_{n,p}}{\sigma_{n,p}}\right)\sum_{i=1}^{n}|\theta_i|^3 + \left(\frac{1}{p} \vee 1\right)\frac{4}{n+2}.$$

Special Cases

$p = 1$

$$||F - \Phi|| \leq \frac{9}{\sqrt{2}} \sum_{i=1}^{n} |\theta_i|^3 + \frac{4}{n+2}$$

$p = 2$

$$||F - \Phi|| \leq \frac{9}{\sqrt{3}} \sum_{i=1}^{n} |\theta_i|^3 + \frac{4}{n+2}$$

Summary

- Independent Sums
  - Replace One
- Combinatorial Central Limit Theorem
  - Exchangeable Pair
- Cone Measure on the Sphere
  - Square Biasing Under Symmetry

Extensions

- Independent Sums
  - Replace One
- Combinatorial Central Limit Theorem
  - Exchangeable Pair
- Cone Measure on the Sphere
  - Square Biasing Under Symmetry
- More...
  - ?