

Approximating the Distribution of Functions In Poisson Binomial by the Chen-Stein Method

Mia Pang Rey¹

Received: February, 2007; Revised: February, 2008

ABSTRACT

The paper gives a proof of a Poisson approximation theorem for an unbounded function in the Poisson binomial random variable W using the Chen-Stein method. The bound for the Radon-Nikodym derivative involving W and the Poisson random variable Z is used to determine a bound for the modified variation distance between the two variables. The mean of Z is defined to be exactly the mean of W , and may be large.

Key words and phrases: Chen-Stein method, Poisson distribution, total variation distance

I. INTRODUCTION

Let X_1, X_2, \dots, X_n be independent Bernoulli random variables such that $P(X_i = 1) = 1 - P(X_i = 0) = p_i$, for $0 \leq p_i \leq 1$, and $\lambda = \sum_{i=1}^n p_i$. Let $W = \sum_{i=1}^n X_i$, $W^{(i)} = W - X_i$, and Z be a Poisson distributed random variable with mean λ . We say that W has Poisson binomial distribution with $Eh(W) = \lambda$. We give a proof of a Poisson approximation theorem for $Eh(W)$ involving higher orders of an unbounded real valued function h and with λ unrestricted. Throughout the paper, we denote I_A to be the indicator function for A where $A \subseteq \mathcal{Z}^*$. If $A = \{r\}$, $r \in \mathcal{Z}^*$, we use I_r instead $I_{(r)}$. \mathcal{R} denotes the set of real numbers.

II. THE CHEN-STEIN METHOD

In 1972, Charles Stein developed a technique in approximating probability distributions of random variables. His method relies mainly on differential operators and exchangeability of random variables which looks very simple yet gives very significant results. It was initially applied in normal approximation. Chen gave the Poisson version of the technique in his 1975 paper. We give a brief overview of Chen's results and some works stemming from them.

A characterization of the Poisson random variable is given by the following lemma.

¹ Assistant Professor of the Department of Accounting and Finance, College of Business Administration, University of the Philippines, Diliman, Quezon City; email address: miapangrey@yahoo.com

Lemma 2.1. (Stein, 1986)

Let Z be a random variable taking values from the set of positive-integers. Then, Z has a Poisson distribution with mean λ if and only if for all bounded functions $f: \mathbb{Z}^+ \rightarrow \mathbb{R}$

$$EZf(Z) = E\lambda f(Z + 1).$$

Because of above lemma, it is natural to say that if we want $L(W) \cong L(Z)$, then

$$EWf(W) - \lambda Ef(W + 1) \cong 0.$$

Now we consider the difference equation

$$\lambda f(w + 1) - wf(w) = h(w) - Eh(Z) \quad (1)$$

where $w \in \mathbb{Z}^+$ and $h: \mathbb{Z}^+ \rightarrow \mathbb{R}$.

Applying W to (1) and getting the expectation of both sides, we get the following expression

$$E\lambda f(W + 1) - EWf(W) = Eh(W) - Eh(Z).$$

Hence, to get the bound of the total variation distance between the Poisson Binomial distribution and the Poisson Distribution, it is sufficient to work on the bound of

$$E\lambda f(W + 1) - EWf(W).$$

(1) is called the **Stein equation for the Poisson approximation** or the **Chen-Stein equation**.

The solution of the Chen-Stein equation is given by the next theorem.

Theorem 2. (Stein, 1986)

The Chen-Stein equation,

$$\lambda f(w + 1) - wf(w) = h(w) - Eh(Z),$$

has a solution $U_\lambda h$ for every function h defined on \mathbb{Z}^+ such that $E|h(Z)| < \infty$. It is given by

$$U_\lambda h(w) = \frac{E(h(Z) - Eh(Z))I(Z \geq w)}{\lambda P(Z = w - 1)}$$

The solution is unique except when $w = 0$ in which it can be chosen arbitrarily. Also, when h is bounded, f is also bounded.

$U_\lambda h$ may also be written as follows:

$$U_\lambda h(w) = - \sum_{k=w}^{\infty} \frac{(w-1)!}{k!} \lambda^{k-w} \left\{ h(k) - \sum_{i=0}^{\infty} h(i) \frac{e^{-\lambda} \lambda^i}{i!} \right\} \quad (2)$$

$$U_\lambda h(w) = \sum_{k=0}^{w-1} \frac{(w-1)!}{k!} \lambda^{k-w} \left\{ h(k) - \sum_{i=0}^{\infty} h(i) \frac{e^{-\lambda} \lambda^i}{i!} \right\} \quad (3)$$

Below are some properties of $U_\lambda h$.

Lemma 2.2. (Stein, 1986)

For any real-valued h defined on \mathbb{Z}^+ where $E | h(Z) | < \infty$,

$$U_\lambda h(w) = \sum_{r=0}^{\infty} h(r) U_\lambda I_r(w)$$

Lemma 2.3. (Chen, Choi (1992))

For any function $h : \mathbb{Z}^+ \rightarrow \mathbb{R}$, $l \in \mathbb{Z}^+$ and Z , a Poisson random variable with mean λ , such that $E | Z^l h(Z) | < \infty$,

$$E | U_\lambda h(Z+1) | = \frac{1}{l} \{ E | h(Z) | + E | h(Z+1) | \}. \tag{4}$$

III. THE RADON-NIKODYM DERIVATIVE

The following lemmas give us a bound for the Radon-Nikodym derivative which is significant to the proof of the main theorem.

Lemma 3.1. (Borisov, Ruzankin (2002))

For p_i not necessarily equal for $i = 1, 2, \dots, n$ where $\tilde{p} = \max(p_1, p_2, \dots, p_n)$, we have the following inequality:

$$\sup_j \frac{P(W = j)}{P(Z = j)} \leq \frac{1}{(1 - \tilde{p})^2}.$$

Lemma 3.2. (Barbour, Chen, Choi (1995))

Let

$$\tilde{C}(p_1, p_2, \dots, p_n) = \sup \left\{ \frac{P(W = r)}{P(Z = r)} : r \geq 0 \right\}$$

and

$$C^*(p_1, p_2, \dots, p_n) = \sup \left\{ \frac{P(W^{(1)} = r)}{P(Z = r)} : r \geq 0 \right\}.$$

We have

$$(1 - p)C^* \leq \tilde{C}.$$

IV. THE MAIN RESULT

Theorem 4.

With C^* as defined in Lemma 3.2, let h be a real-valued function defined on \mathbb{Z}^+ such that $EZ^6|h(Z)| < \infty$. We have

$$\left| Eh(W) - Eh(Z) + \frac{1}{2} \sum_{i=1}^n p_i^2 E\Delta^2 h(Z) - \frac{1}{3} \sum_{i=1}^n p_i^3 E\Delta^3 h(Z) - \frac{1}{8} \left(\sum_{i=1}^n p_i^2 \right)^2 E\Delta^4 h(Z) \right| \\ \leq C^* \left(\left(\sum_{i=1}^n p_i^2 \right)^3 Q_1 + 2 \sum_{i=1}^n p_i^2 \sum_{i=1}^n p_i^3 Q_2 + \sum_{i=1}^n p_i^4 Q_3 \right)$$

where

$$Q_1 = 80(1 \wedge \lambda^{-3}) E|h(Z+3)| \\ + \frac{1}{6} (1 \wedge \lambda^{-2}) \{51E|h(Z+4)| - 86E|h(Z+3)| + 24E|h(Z+2)| + 6E|h(Z+1)| + 5E|h(Z)|\} \\ + \frac{1}{120} (1 \wedge \lambda^{-1}) \{38E|h(Z+5)| - 145E|h(Z+4)| + 200E|h(Z+3)| - 110E|h(Z+2)| \\ + 10E|h(Z+1)| + 7|h(Z)|\} \\ + \frac{1}{48} \{E|h(Z+6)| - 6E|h(Z+5)| + 15E|h(Z+4)| - 20E|h(Z+3)| + 15E|h(Z+2)| \\ - 6E|h(Z+1)| + E|h(Z)|\}$$

$$Q_2 = 120(1 \wedge \lambda^{-2}) \{E|h(Z+3)| + E|h(Z+2)|\} \\ + \frac{1}{6} (1 \wedge \lambda^{-1}) \{7E|h(Z+4)| - 18E|h(Z+2)| + 8E|h(Z+1)| + 3E|h(Z)|\} \\ + \frac{1}{120} \{8E|h(Z+5)| - 25E|h(Z+4)| + 20E|h(Z+3)| - 100E|h(Z+2)| \\ + 80E|h(Z+1)| + 17|h(Z)|\}$$

$$Q_3 = 2(1 \wedge \lambda^{-1}) \{E|h(Z+3)| + 2E|h(Z+2)| + E|h(Z+1)|\} \\ + \frac{1}{12} \{3E|h(Z+4)| + 4E|h(Z+3)| - 6E|h(Z+2)| - 12E|h(Z+1)| + 11E|h(Z)|\}$$

Corollary 4.

$$\sup_{|h(Z)| \leq 1} \left| Eh(W) - Eh(Z) + \frac{1}{2} \sum_{i=1}^n p_i^2 E\Delta^2 h(Z) - \frac{1}{3} \sum_{i=1}^n p_i^3 E\Delta^3 h(Z) - \frac{1}{8} \left(\sum_{i=1}^n p_i^2 \right)^2 E\Delta^4 h(Z) \right|$$

$$\leq 8C^* \left(10 \left(\sum_{i=1}^n p_i^2 \right)^3 (1 \wedge \lambda^{-3}) + 6 \sum_{i=1}^n p_i^2 \sum_{i=1}^n p_i^3 (1 \wedge \lambda^{-2}) + \sum_{i=1}^n p_i^4 (1 \wedge \lambda^{-1}) \right).$$

The proof of Theorem 4 needs the following details.

Let $V_\lambda h(w) = \Delta U_\lambda h(w+1)$. Below are some facts about $V_\lambda h(w)$.

$$V_\lambda I_r(w) > 0 \Leftrightarrow w = r - 1 \tag{5}$$

$$V_\lambda I_r = \sum_{i=1}^{\infty} V_\lambda I_r(s) I_s \tag{6}$$

$$V_\lambda I_{r+1}(r) \leq (1 \wedge \lambda^{-1}) \tag{7}$$

Proofs of (5) and (6) are found in Stein (1986) while (7) follows from Lemma 2.2.

Lemma 4.1. (Barbour, Chen, Choi (1995))

For $k \geq 0$ and any real-valued function h such that we have $EZ^{k+2}|h(Z)| < \infty$, we have

$$EV_\lambda h(Z+k) = \frac{-1}{(k+1)(k+2)} \{ (k+1)Eh(Z+K+2) - (k+2)Eh(Z+K+1) + Eh(Z) \} \tag{8}$$

Lemma 4.2. (Barbour, Chen, Choi (1995))

Let k, m be nonnegative integers and h a nonnegative function such that $EZ^{k+m+2}|h(Z)| < \infty$. Then we have

$$\begin{aligned} \sum_{r=0}^{\infty} h(r+m)E|V_\lambda I_r(Z+k)| &\leq 2(1 \wedge \lambda^{-1})Eh(Z+k+m+1) + \frac{1}{k+2}Eh(Z+k+m+2) \\ &\quad - \frac{1}{k+1}Eh(Z+k+m+1) + \frac{1}{(k+1)(k+2)}Eh(Z+m) \end{aligned}$$

Lemma 4.3. (Barbour, Chen, Choi (1995))

Let k be nonnegative integers and h a nonnegative function such that $EZ^{k+2}|h(Z)| < \infty$. Then we have

$$\begin{aligned} E|V_\lambda h(Z+k)| &\leq 2(1 \wedge \lambda^{-1})E|h(Z+k+1)| + \frac{1}{k+2}E|h(Z+k+2)| \\ &\quad - \frac{1}{k+1}E|h(Z+k+1)| + \frac{1}{(k+1)(k+2)}E|h(Z)| \end{aligned}$$

Lemma 4.4. Let m be a positive integer and h be real-valued function such that

$$EZ^{2m}|h(Z)| < \infty.$$

Then we have

$$EV_\lambda^m h(Z) = (-1)^m \frac{1}{2^m} \frac{1}{m!} E\Delta^{2m} h(Z)$$

Proof:

We prove the lemma by induction. Observe that when $m = 1$, we have

$$EV_\lambda h(Z) = -\frac{1}{2} E\Delta^2 h(Z)$$

which is true by (8). When $m = 2$, again by (8),

$$EV_\lambda^2 h(Z) = EV_\lambda (V_\lambda h(Z)) = -\frac{1}{2} \{EV_\lambda h(Z+2) - 2EV_\lambda h(Z+1) + EV_\lambda h(Z)\}$$

Reapplying (8) to the three terms above, we arrive at the expression

$$EV_\lambda^2 h(Z) = \frac{1}{8} E\Delta^4 h(Z) = (-1)^2 \frac{1}{2^2} \frac{1}{2!} E\Delta^4 h(Z)$$

Suppose it holds for $m = k$, that is,

$$EV_\lambda^k h(Z) = (-1)^k \frac{1}{2^k} \frac{1}{k!} E\Delta^{2k} h(Z).$$

We show it holds for $m = k + 1$.

$$\text{Now, } EV_\lambda^{k+1} h(Z) = EV_\lambda^k (V_\lambda \Delta^2 h(Z)) = (-1)^k \frac{1}{2^k} \frac{1}{k!} E\Delta^{2k} h(Z)$$

But by (8)

$$\begin{aligned} E\Delta^{2k} V_\lambda h(Z) &= \sum_{j=0}^{2k} (-1)^j \binom{2k}{j} EV_\lambda h(Z+j) \\ &= \sum_{j=0}^{2k} (-1)^j \binom{2k}{j} \left\{ -\frac{1}{j+2} Eh(Z+j+2) + \frac{1}{j+1} Eh(Z+j+1) \right. \\ &\quad \left. - \frac{1}{(j+1)(j+2)} Eh(Z) \right\} \\ &= -\frac{1}{2k+2} Eh(Z+2k+2) + \sum_{j=2}^{2k+1} (-1)^{j-1} \left\{ \binom{2k}{j-1} + \binom{2k}{j-2} \right\} Eh(Z+j) \\ &\quad + Eh(Z+1) + \sum_{j=0}^{2k} (-1)^{j+1} \binom{2k}{j} \frac{1}{(j+1)(j+2)} Eh(Z) \end{aligned}$$

Now

$$\binom{2k}{j-1} + \binom{2k}{j-2} = \frac{j}{k+2} \binom{2k+2}{j}$$

and

$$\sum_{j=0}^{2k} (-1)^{j+1} \binom{2k}{j} \frac{1}{(j+1)(j+2)} = -\frac{1}{2k+2}.$$

Hence,

$$\begin{aligned} E\Delta^{2k} V_\lambda h(Z) &= -\frac{1}{2k+2} \left\{ Eh(Z+2k+2) + \sum_{j=1}^{2k+1} (-1)^j j \binom{2k+2}{j} Eh(Z+j) + Eh(Z) \right\} \\ &= -\frac{1}{2k+2} E\Delta^{2k+2} h(Z). \end{aligned}$$

Therefore,

$$EV_\lambda^{k+1} h(Z) = (-1)^{k+1} \frac{1}{2^k} \frac{1}{k!} \frac{1}{2k+2!} E\Delta^{2k+2} h(Z) = (-1)^{k+1} \frac{1}{2^{k+1}} \frac{1}{k+1!} E\Delta^{2(k+1)} h(Z) \quad \blacksquare$$

Lemma 4.5.

Let h be a real-valued function such that $EZ^6 |h(Z)| < \infty$. Then

$$\begin{aligned} E|V_\lambda^3 h(Z)| &\leq 80(1 \wedge \lambda^{-3}) E|h(Z+3)| \\ &+ \frac{1}{6} (1 \wedge \lambda^{-2}) \{51E|h(Z+4)| - 86E|h(Z+3)| + 24E|h(Z+2)| + 6E|h(Z+1)| + 5E|h(Z)|\} \\ &+ \frac{1}{120} (1 \wedge \lambda^{-1}) \{38E|h(Z+5)| - 145E|h(Z+4)| + 200E|h(Z+3)| - 110E|h(Z+2)| \\ &\quad + 10E|h(Z+1)| + 7|h(Z)|\} \\ &+ \frac{1}{48} \{E|h(Z+6)| - 6E|h(Z+5)| + 15E|h(Z+4)| - 20E|h(Z+3)| + 15E|h(Z+2)| \\ &\quad - 6E|h(Z+1)| + E|h(Z)|\} \end{aligned}$$

An outline of the proof is as follows:

By identities (6) and (7), observe that

$$\begin{aligned} V_\lambda^3 I_r &= V_\lambda^2 (V_\lambda I_r) = V_\lambda^2 \left(\sum_{s=0}^{\infty} V_\lambda I_r(s) I_r \right) = \sum_{s=0}^{\infty} V_\lambda I_r(s) V_\lambda^2 I_r \\ &= \sum_{s=0}^{\infty} \sum_{t=0}^{\infty} \sum_{k=0}^{\infty} V_\lambda I_r(s) V_\lambda I_s(t) V_\lambda I_t(k) I_k \\ &= V_\lambda I_r(r-1) \sum_{t=0}^{\infty} \sum_{k=0}^{\infty} V_\lambda I_{r-1}(t) V_\lambda I_t(k) I_k \\ &\quad + \sum_{s \neq r-1} \sum_{t=0}^{\infty} \sum_{k=0}^{\infty} V_\lambda I_r(s) V_\lambda I_s(t) V_\lambda I_t(k) I_k \end{aligned}$$

Using (5),

$$|V_\lambda^3 h| = \left| V_\lambda^3 \sum_{r=0}^{\infty} h(r) I_r \right| = \left| \sum_{r=0}^{\infty} h(r) V_\lambda^3 I_r \right| = \sum_{r=0}^{\infty} |h(r)| V_\lambda^3 I_r = A + B + C + D + E$$

where

$$\begin{aligned} A &= 2 \sum_{r=3}^{\infty} |h(r)| V_\lambda I_r (r-1) V_\lambda I_{r-1} (r-2) V_\lambda I_{r-2} (r-3) \\ B &= 2 \sum_{r=1}^{\infty} |h(r)| V_\lambda I_r (r-1) \sum_{t \neq (r-2)} \sum_{k \neq (t-1)} V_\lambda I_{r-1} (t) V_\lambda I_t (k) I_k \\ C &= 2 \sum_{r=0}^{\infty} |h(r)| \sum_{s \neq (r-1)} \sum_{k \neq (s-2)} V_\lambda I_r (s) V_\lambda I_s (s-1) V_\lambda I_{s-1} (k) I_k \\ D &= 2 \sum_{r=0}^{\infty} |h(r)| \sum_{s \neq (r-1)} \sum_{t \neq (s-1)} V_\lambda I_r (s) V_\lambda I_s (t) V_\lambda I_t (t-1) I_{t-1} \\ E &= \sum_{r=0}^{\infty} |h(r)| V_\lambda^3 I_r = V_\lambda^3 |h| \end{aligned}$$

Determine an upper bound for each of A,B,C,D and E by using (7). This leads to

$$\begin{aligned} E|V_\lambda^3 h(Z)| &\leq 8(1 \wedge \lambda^{-3}) E|h(Z+3)| \\ &\quad + (1 \wedge \lambda^{-2}) \left\{ 8E|V_\lambda |h|Z+2| + 4E|V_\lambda^2 |h|Z+1| \right\} \\ &\quad + (1 \wedge \lambda^{-1}) \left\{ 4E|V_\lambda^2 |h|Z+1| + 2 \sum_{r=0}^{\infty} E|V_\lambda^2 I_{r(Z)}| \right\} \\ &\quad + EV_\lambda^3 |h|(Z). \end{aligned}$$

Use Lemmas 4.2, 4.3 and 4.4 to complete the proof. ■

Proof of Theorem 4.

Using the identity

$$Eh(W) - Eh(Z) = \sum_{i=1}^n p_i^2 EV_\lambda h(W^{(i)})$$

where $W^{(i)} = W - X_i$ from Stein (1986) and iterating twice, we get the following expression

$$\begin{aligned} Eh(W) - Eh(Z) &- \sum_{i=1}^n p_i^2 EV_\lambda h(Z) - \sum_{i=1}^n (p_i^2)^2 EV_\lambda^2 h(Z) + \sum_{i=1}^n p_i^3 E\Delta V_\lambda h(Z) \\ &= \sum_{i=1}^n (p_i^2)^2 \sum_{j=1}^n p_j^2 EV_\lambda^3 h(W^{(j)}) - \sum_{i=1}^n p_i^2 \sum_{j=1}^n p_j^3 E\Delta V_\lambda^2 h(W^{(j)}) \\ &\quad - \sum_{i=1}^n p_i^3 \sum_{j=1}^n (p_j^2)^2 E\Delta V_\lambda^2 h(W^{(j)}) - \sum_{i=1}^n p_i^4 E\Delta^2 V_\lambda h(W^{(i)}) \end{aligned}$$

By Lemmas 4.1 and 4.4, we have

$$EV_\lambda h(Z) = -\frac{1}{2}E\Delta^2 h(Z)$$

$$EV_\lambda^2 h(Z) = \frac{1}{8}E\Delta^4 h(Z)$$

$$E\Delta V_\lambda h(Z) = EV_\lambda h(Z+1) - EV_\lambda h(Z) = -\frac{1}{3}E\Delta^3 h(Z)$$

From Lemmas 3.1 and 3.2, $P(W^{(i)} = r) \leq C^* P(Z = r)$ for $r > 0$ and $0 \leq i \leq n$. With

$$\lambda_k = \sum_{i=1}^k p_i,$$

we have

$$\begin{aligned} \left| \lambda_2^2 \sum_{i=1}^n p_i^2 EV_\lambda^3 h(W^{(j)}) \right| &\leq \lambda_2^2 \sum_{i=1}^n p_i^2 E|V_\lambda^3 h(W^{(j)})| = \lambda_2^2 \sum_{i=1}^n p_i^2 \sum_{r=0}^\infty |V_\lambda^3 h(r)| P(W^{(i)} = r) \\ &= \lambda_2^2 \sum_{i=1}^n p_i^2 \sum_{r=0}^\infty |V_\lambda^3 h(r)| \frac{P(W^{(i)} = r)}{P(Z = r)} P(Z = r) \\ &\leq C^* \lambda_2^3 \sum_{r=0}^\infty |V_\lambda^3 h(r)| P(Z = r) \\ &= C^* \lambda_2^3 E|V_\lambda^3 h(Z)| \end{aligned}$$

Similarly, it can be shown that

$$\begin{aligned} \left| \lambda_2 \sum_{i=1}^n p_i^3 E\Delta V_\lambda^2 h(W^{(j)}) \right| &\leq C^* \lambda_2 \lambda_3 E|\Delta V_\lambda^2 h(Z)| \\ \left| \lambda_3 \sum_{i=1}^n p_i^2 E\Delta V_\lambda^2 h(W^{(j)}) \right| &\leq C^* \lambda_3 \lambda_2 E|\Delta V_\lambda^2 h(Z)| \\ \left| \sum_{i=1}^n p_i^4 E\Delta^2 V_\lambda h(W^{(j)}) \right| &\leq C^* \lambda_4 E|\Delta^2 V_\lambda h(Z)| \end{aligned}$$

By Lemmas 4.2, 4.3 and 4.5, we get the desired result. ■

The corollary is a direct consequence where the supremum is achieved by a function h which satisfies $|h| = 1$ for all positive integers.

The main theorem gives a Poisson approximation for the distribution of an unbounded function in the Poisson binomial random variable. It involves higher orders of an unbounded function h which extends the results of Barbour, Chen and Choi in their 1995 paper.

Acknowledgement

This is part of my dissertation which I did under the supervision of Prof. Louis Chen of the Institute for Mathematical Science, National University of Singapore and Prof. Polly W. Sy of the Institute of Mathematics, University of the Philippines, Diliman Campus. I am greatly indebted to them.

References

- BARBOUR, A.D. (1987) *Asymptotic expansions in the Poisson Limit Theorem*. Ann. Prob. 15, 748–766
- BARBOUR A.D., CHEN L. H.Y. and CHOI K.P. (1995) *Poisson Approximation for Unbounded Functions, I: Independent Summands*. Statistica Sinica 5, 746–766.
- BORISOV I. S. and RUZANKIN. P.S. (2002) *Poisson Approximation for Expectation of Unbounded Function of Independent Random Variables*. Ann. Prob. 30, 1657–1680
- STEIN C. (1986). *Approximate Computation of Expectations*. IMS, Hayward, California