A NOTE ON STEIN EQUATION FOR WEIGHTED SUMS OF INDEPENDENT $\chi^2$ DISTRIBUTIONS

XIAOHUI CHEN AND PARTHA DEY

Abstract. This note provides the Stein equation for weighted sums of independent $\chi^2$ distributions.

1. Introduction

Stein’s method, first introduced in [6], is a powerful tool to bound the distance between a probability distribution and the Gaussian distribution. Over the past decades, Stein’s method has been extended to other distributions including Poisson [2], exponential [1], $\chi^2$ [3], and gamma [4]. At the core of Stein’s method is a differential operator $\mathcal{T}$, which generates a differential equation known as the Stein equation characterizing the target distribution $\pi$:

$$\mathcal{T}f(x) = h(x) - \mathbb{E}_{Z \sim \pi}[h(Z)]$$

for a collection of functions $h \in \mathcal{H}$. On one hand, the Stein equation (1) satisfies $\mathbb{E}_{Z \sim \pi}[\mathcal{T}f(Z)] = 0$. On the other hand, if $\pi$ has an absolutely continuous density with suitable regularity conditions, then the Stein equation (1) has a unique solution $f := f_h$ for any given piecewise continuous function $h$. Thus for any random variable $X \sim \nu$, if $\mathbb{E}_{X \sim \nu}[\mathcal{T}f_h(X)] = 0$ over a rich class of functions $h \in \mathcal{H}$, then $\nu = \pi$. Quantitatively, taking expectation on both sides of (1) gives

$$\mathbb{E}_{X \sim \nu}[\mathcal{T}f_h(X)] = \mathbb{E}_{X \sim \nu}[h(X)] - \mathbb{E}_{Z \sim \pi}[h(Z)].$$

In order to control the distance between $\nu$ and $\pi$, it is enough to estimate the quantity $\mathbb{E}_{X \sim \nu}[\mathcal{T}f_h(X)]$ over $h \in \mathcal{H}$.

In this note, we derive a characterizing operator and the associated Stein equation for weighted sums of independent $\chi^2$ distributions. Such distributions arise as weak limits of degenerate $U$-statistics [5], which are useful in goodness-of-fit tests for distribution functions such as the Cramér-von Mises test statistic.

2. Characterizing operator for $\chi^2$ distributions

We first start with the characterizing operator for one $\chi^2$ distribution. Then we derive the general results for weighted sums of independent $\chi^2$ distributions.

Date: First version: February 25, 2020.

Key words and phrases. Stein equation, weighted sum of independent $\chi^2$ distributions.

X. Chen’s research is supported in part by NSF CAREER Award DMS-1752614 and a Simons Fellowship.
2.1. **Characterizing operator for one $\chi^2$ distribution.** Let $Q \sim \chi^2_p$ and $\tilde{Q} = Q - p$. Then the operator

$$\tilde{T}f(x) = 2(x + p)f'(x) - xf(x)$$

is a characterizing operator for the centered $\chi^2_p$ distribution, in the sense that $\tilde{Q} \sim \chi^2_p - p$ if and only if $\mathbb{E}(\tilde{T}f(\tilde{Q})) = 0$ for all "smooth" function $f$. To prove that $\tilde{Q} \sim \chi^2_p - p$ distribution indeed satisfies $\mathbb{E}(\tilde{T}f(\tilde{Q})) = 0$ for all "smooth" function $f$, we use the following lemma.

**Lemma 1** (Integration by parts formula). If $Q \sim \chi^2_p$, then

$$\mathbb{E}((Q - p)f(Q)) = \mathbb{E}(2Qf'(Q))$$

holds for any absolutely continuous function $f : \mathbb{R} \to \mathbb{R}$ such that the expectations $\mathbb{E}|f(Q)|, \mathbb{E}|Qf(Q)|, \mathbb{E}|Qf'(Q)|$ are finite. Equivalently, if $\tilde{Q} = Q - p$, then

$$\mathbb{E}(\tilde{Q}f(\tilde{Q})) = \mathbb{E}(2(\tilde{Q} + p)f'(\tilde{Q})).$$

**Proof of Lemma 1.** Note that we can write $Q = \sum_{i=1}^{p} Z_i^2$, where $Z_1, Z_2, \ldots, Z_p$ are i.i.d. standard Gaussian random variables and the equality holds in distribution. Then by the Gaussian integration by parts, we have

$$\mathbb{E}((Q - p)f(Q)) = \sum_{i=1}^{p} \mathbb{E}((Z_i^2 - 1)f(Q)) = \sum_{i=1}^{p} \mathbb{E}\left(Z_i \frac{\partial}{\partial Z_i} f(Q)\right) = \sum_{i=1}^{p} \mathbb{E}(2Z_i^2 f'(Q)) = \mathbb{E}(2Qf'(Q)).$$

The second part is an immediate consequence of the first part. □

2.2. **Some combinatorial results.** Given a sequence of distinct non-zero real numbers $\lambda_1, \lambda_2, \ldots, \lambda_r$, we define

$$\Lambda_{k,i} := \sum_{S \subseteq [r] \setminus \{i\}, |S| = k} \prod_{j \in S} \lambda_j$$

and

$$\Lambda_k := \sum_{S \subseteq [r], |S| = k} \prod_{j \in S} \lambda_j$$

for $i, k \in [r]$, where $[r] = \{1, \ldots, r\}$. Define $\Lambda_0 \equiv \Lambda_{0,i} \equiv 1$. Clearly $\Lambda_{r,i} = 0$ for all $i \in [r]$.

**Lemma 2.** For $i, k \in [r]$, we have

$$\Lambda_k - \Lambda_{k-1,i}\lambda_i = \Lambda_{k,i} \quad \text{and} \quad \sum_{i=1}^{r} \Lambda_{k,i}\lambda_i = (k + 1)\Lambda_{k+1}.$$

**Proof of Lemma 2.** The first claim follows from the definitions of $\Lambda_{k,i}$ and $\Lambda_k$. Note that $\Lambda_k$ involves the summation of the product terms of $k$ distinct $\lambda_i$’s, which implies
2.3. Stein equation for weighted sums of independent \( \chi^2 \) distributions. Let \( Q_i \sim \chi_{m_i}^2, i = 1, 2, \ldots, r \) be independent \( \chi^2 \) random variables and \( \lambda_1, \lambda_2, \ldots, \lambda_r \) be a sequence of distinct non-zero real numbers. We consider the random variable

\[
U = \sum_{i=1}^{r} \lambda_i Q_i,
\]

which is a weighted sum of independent \( \chi^2 \) random variables. Define \( \mu = \mathbb{E}(U) = \sum_{i=1}^{r} \lambda_i m_i, \tilde{U} := U - \mu, \tilde{Q}_i := Q_i - m_i, i = 1, 2, \ldots, r. \) We also define

\[
\mu_k := \sum_{i=1}^{r} \lambda_i^2 \Lambda_{k-1,i} m_i \quad \text{for} \quad k \geq 1
\]

and \( \mu_0 = 0. \) The main result of this note is the following Stein equation for \( \tilde{U}. \)

*Theorem 3* (Stein equation for \( \tilde{U}. \)). Let \( f : \mathbb{R} \to \mathbb{R} \) be an \( r \)-th differentiable function such that \( \mathbb{E} |f^{(k)}(U)| \) and \( \mathbb{E} |U f^{(k)}(U)|, k = 0, 1, \ldots, r \) are finite. Then we have \( \mathbb{E} \tilde{T} f(\tilde{U}) = 0, \) where

\[
\tilde{T} f(x) = \sum_{k=0}^{r} (-2)^k \left( \mu_k + \Lambda_k x \right) f^{(k)}(x). \tag{2}
\]

**Remark 4.** Stein equation for the non-centered weighted sum \( U \) of independent \( \chi^2 \) distributions is given by \( \mathbb{E} T f(U) = 0, \) where

\[
T f(x) = \sum_{k=0}^{r} (-2)^k \left( \mu_k + \Lambda_k x - \Lambda_k \mu \right) f^{(k)}(x). \tag{3}
\]

**Proof of Theorem 3.** Take a smooth function \( f. \) By the integration by parts formula for \( \chi^2 \) distribution in Lemma 1, we have

\[
\mathbb{E} \left( \tilde{U} f(\tilde{U}) \right) = 2 \sum_{i=1}^{r} \lambda_i^2 \mathbb{E} \left( Q_i f^{(1)}(\tilde{U}) \right)
\]

\[
= 2 \sum_{i=1}^{r} \lambda_i^2 m_i \mathbb{E} \left( f^{(1)}(\tilde{U}) \right) + 2 \sum_{i=1}^{r} \lambda_i^2 \mathbb{E} \left( \tilde{Q}_i f^{(1)}(\tilde{U}) \right)
\]

\[
= 2 \mu_1 \mathbb{E} \left( f^{(1)}(\tilde{U}) \right) + 2 \Lambda_1 \mathbb{E} \left( \tilde{U} f^{(1)}(\tilde{U}) \right) - 2 \sum_{i=1}^{r} \lambda_i \Lambda_{1,i} \mathbb{E} \left( \tilde{Q}_i f^{(1)}(\tilde{U}) \right)
\]

\[
= 2 \mathbb{E} \left( (\mu_1 + \Lambda_1 \tilde{U}) f^{(1)}(\tilde{U}) \right) - 2^2 \sum_{i=1}^{r} \lambda_i^2 \Lambda_{1,i} \mathbb{E} \left( Q_i f^{(2)}(\tilde{U}) \right),
\]
where the third equality follows from Lemma 2. Expanding the last term we have
\[
\sum_{i=1}^{r} \lambda_i^2 \Lambda_{1,i} E \left( Q_i f^{(2)}(\tilde{U}) \right) = \sum_{i=1}^{r} \lambda_i^2 \Lambda_{1,i} m_i E \left( f^{(2)}(\tilde{U}) \right) + \sum_{i=1}^{r} \lambda_i^2 \Lambda_{2,i} E \left( \tilde{Q}_i f^{(2)}(\tilde{U}) \right) \]
\[
= \mu_2 E \left( f^{(2)}(\tilde{U}) \right) + \Lambda_2 E \left( \tilde{U} f^{(2)}(\tilde{U}) \right) - 2 \sum_{i=1}^{r} \lambda_i \Lambda_{2,i} E \left( \tilde{Q}_i f^{(2)}(\tilde{U}) \right) \]
\[
= \mu_2 E \left( (\mu_2 + \Lambda_2 \tilde{U}) f^{(2)}(\tilde{U}) \right) - 2 \sum_{i=1}^{r} \lambda_i^2 \Lambda_{2,i} E \left( Q_i f^{(3)}(\tilde{U}) \right). \]

Using induction we finally get
\[
E \left( \tilde{U} f(\tilde{U}) \right) = \sum_{k=1}^{r-1} (-1)^{k-1} 2^k E \left( \mu_k f^{(k)}(\tilde{U}) + \lambda_k \tilde{U} f^{(k)}(\tilde{U}) \right) + (-1)^{r-1} 2^r \sum_{i=1}^{r} \lambda_i^2 \Lambda_{r-1,i} E \left( Q_i f^{(r)}(\tilde{U}) \right) \]
\[
= \sum_{k=1}^{r-1} (-1)^{k-1} 2^k E \left( (\mu_k + \Lambda_k \tilde{U}) f^{(k)}(\tilde{U}) \right) + (-1)^{r-1} 2^{r} \Lambda_r E \left( (\mu + \tilde{U}) f^{(r)}(\tilde{U}) \right), \]

where the last step follows from Lemma 2 and \( \Lambda_{r,i} = 0 \). Thus \( \tilde{U} \) satisfies the relation
\[
E \tilde{F} f(\tilde{U}) = 0, \quad \text{where} \quad \tilde{F} f(x) = xf(x) + \sum_{k=1}^{r-1} (-2)^k (\mu_k + \Lambda_k x) f^{(k)}(x) + (-2)^r \Lambda_r (\mu + x) f^{(r)}(x). \]

Then (2) follows from the last identity together with \( \mu_0 = 0, \Lambda_0 = 1 \), and \( \mu_r = \Lambda_r \mu \). ■

References

[1] Sourav Chatterjee, Jason Fulman, and Adrian Röllin. Exponential approximation by stein’s method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.*, 2011.
[2] Louis H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probab.*, 3(3):534–545, 06 1975.
[3] Robert E. Gaunt, Alastair Pickett, and Gesine Reinert. Chi-square approximation by stein’s method with application to pearson’s statistic. *Annals of Applied Probability* (arXiv:1507.01707), 2016.
[4] Ho Ming Luk. *Stein’s Method for the Gamma Distribution and Related Statistical Applications*. PhD thesis, University of Southern California, 1994.
[5] Robert J. Serfling. *Approximation theorems of mathematical statistics*. Wiley series in probability and mathematical statistics : Probability and mathematical statistics. Wiley, 1980.
[6] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, pages 583–602, Berkeley, Calif., 1972. University of California Press.
STEIN EQUATION FOR WEIGHTED SUMS OF INDEPENDENT $\chi^2$ DISTRIBUTIONS

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
S. WRIGHT STREET, CHAMPAIGN, IL 61820
E-mail: xhchen@illinois.edu
URL: http://publish.illinois.edu/xiaohuichen/

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
S. WRIGHT STREET, CHAMPAIGN, IL 61820
E-mail: psdey@illinois.edu
URL: https://faculty.math.illinois.edu/psdey/