Christopher S. Withers and Saralees Nadarajah

A note on bias reduction

<https://doi.org/10.5802/crmath.49>

© Académie des sciences, Paris and the authors, 2020. Some rights reserved.

This article is licensed under the Creative Commons Attribution 4.0 International License.
http://creativecommons.org/licenses/by/4.0/

Les Comptes Rendus. Mathématique sont membres du Centre Mersenne pour l’édition scientifique ouverte
wwwcentre-mersenne.org
A note on bias reduction

Christopher S. Withers\textsuperscript{a} and Saralees Nadarajah\textsuperscript{a, b}

\textsuperscript{a} Industrial Research Limited, Lower Hutt, New Zealand
\textsuperscript{b} University of Manchester, Manchester M13 9PL, UK
E-mails: kit.withers@gmail.com, mbbssn2@manchester.ac.uk

Abstract. Let \( \hat{w} \) be an unbiased estimate of an unknown \( w \in \mathbb{R} \). Given a function \( t(w) \), we show how to choose a function \( f_n(w) \) such that for \( w^* = \hat{w} + f_n(w) \), \( E t(w^*) = t(w) \). We illustrate this with \( t(w) = wa \) for some constant \( a \). For \( a = 2 \) and \( \hat{w} \) normal, this leads to the convolution equation \( c_r = c_r \otimes c_r \).

1. Introduction

Let \( \hat{w} \) be an unbiased estimate of an unknown \( w \in \mathbb{R} \). Given a function \( t(w) \), we show how to choose a function \( f_n(w) \) such that \( E t(w^*) = t(w) \) for \( w^* = \hat{w} + f_n(w) \). We illustrate this in Section 2 with \( \hat{w} \) normal and \( t(w) = wa \) for some constant \( a \). For \( a = 2 \) this gives the convolution equation \( c_r = c_r \otimes c_r \) to solve.

\( w^* \) is not an estimate since it depends on the unknown \( w \). The method extends to \( \hat{w} \) any standard estimate of an unknown \( w \in \mathbb{R} \) with respect to a given parameter \( n \). That is, \( E \hat{w} \to w \) as \( n \to \infty \) and, for \( r \geq 1 \), its \( r \)th order cumulants have magnitude \( n^{1-r} \) and can be expanded as power series in \( n^{-1} \):

\[
\kappa(\hat{w}_{i_1}, \ldots, \hat{w}_{i_r}) = \sum_{e=r-1}^{\infty} n^{-e} k_{i_1}^{e} \cdots k_{i_r}^{e},
\]

for \( 1 \leq i_1, \ldots, i_r \leq p \) and \( k_{i_0}^{0} = w_{i_0} \), where \( w_i \) is the \( i \)th component of \( w \), and the cumulant coefficients \( k_{i_1}^{e} \cdots k_{i_r}^{e} \) are bounded as \( n \to \infty \), but may depend on \( w \). For \( p = 1 \), \( \kappa(\hat{w}) = \sum_{e=r-1}^{\infty} n^{-e} k_{r}^{e} \) for \( r \geq 1 \), where \( k_{1,0} = w \). Cumulant coefficients are the building blocks of analytic methods for statistical inference. For example, methods for constructing estimates of low bias for any smooth function \( t(w) : \mathbb{R}^p \to \mathbb{R} \) were given in Mynbaev et al. [3] and Withers and Nadarajah [4–14].
Given a sequence \( a_1, a_2, \ldots \), the exponential partial Bell polynomial \( B_{i,k}(a) \) is defined by
\[
\left( \sum_{j=1}^{\infty} a_j t^j / j! \right)^k / k! = \sum_{j=k}^{\infty} B_{i,k}(a) t^j / j!
\]
for \( t \in R \) and \( k = 0, 1, \ldots \). It is tabled on p. 307–308 of Comtet [2] for \( 1 \leq r \leq 12 \). Given two sequences \( a_1, a_2, \ldots \) and \( b_1, b_2, \ldots \), their discrete convolution is defined by
\[
a_r \otimes b_r = \sum_{i=1}^{r} a_i b_{r-i}.
\]
Set \( \delta_{i,1} = I(i = 1) \) and \( (a)_j = a!/ (a-j)! \).

2. Adding bias to \( \tilde{w} \) to reduce the bias of \( t(\tilde{w}) \)

Suppose that \( \tilde{w} \sim \mathcal{N}(w, v(w)/n) \), a normal estimate. Set
\[
v = v(w), \quad f_n(w) = \sum_{i=1}^{\infty} k_i n^{-i}, \quad (2)
\]
where \( k_i = b_i / i! \) may depend on \( w \). Theorems 1 and 3 show how to choose \( k_i \) or \( b_i \) so that for any given \( a \),
\[
E(\tilde{w}^*)^a = w^a, \quad \tilde{w}^* = \tilde{w} + f_n(w).
\]
Then \( E(\tilde{w}^*)^2 = w^2 \).

**Theorem 1.** Take \( w^* = \tilde{w} + f_n(w) \) with
\[
k_r = -c_r v^r / D^{2r-1},
\]
where \( D = 2w \) and, for \( r \geq 2 \),
\[
c_1 = 1, \quad c_r = c_r \otimes c_r. \quad (5)
\]

**Proof.** For \( w^* \) of (3),
\[
E(\tilde{w}^*)^2 = m_n^2 + v/n = w^2 + 2w f_n(w) + f_n(w)^2 + v/n = w^2 + \sum_{i=1}^{\infty} T_i n^{-i},
\]
where \( T_i = D k_i + s_i + v \delta_{i,1}, \quad s_1 = 0, \) and for \( i \geq 2 \),
\[
s_i = k_i \otimes k_i = \sum_{j=1}^{i-1} k_j k_{i-j}.
\]
So, \( T_i = 0 \) if we take \( k_i = - (s_i + v \delta_{i,1}) / D \) for \( i \geq 1 \). This gives the result.

**Corollary 2.** For \( r \geq 2 \),
\[
c_r = - \left( \frac{1/2}{r} \right) (-4c_1)^r / 2 = 2^{-r-1} 1 \cdot 3 \cdot (2r-3) / r!.
\]

**Proof.** Set
\[
C(t) = \sum_{r=1}^{\infty} c_r t^r,
\]
where \( c_1 \) is now arbitrary. By (5), \( C(t) = c_1 t + C(t)^2 \). Since \( C(0) = 0 \), this gives
\[
C(t) = \left[ 1 - (1 - 4c_1 t)^{1/2} \right] / 2 = - \sum_{r=1}^{\infty} \left( \frac{1/2}{r} \right) (-4c_1 t)^r / 2
\]
which implies the result.

**Theorem 3.** Take \( w^* = \hat{w} + f_n(w) \) with

\[
f_n(w) = w \sum_{j=0}^{\infty} A_j \left( c / n^3 \right)^j / j! + wu \sum_{j=0}^{\infty} B_j \left( c / n^3 \right)^j / j! - w,
\]

where \( c = 4\nu^3 \), \( a_j = (1/2)^j/2 \), \( u = -\nu n^{-1} \), \( B_k = a_{k+1} / a_k (k+1) \) and

\[
A_j = \sum_{k=0}^{j} \left( (1/3)_k B_{j,k}(a) \right).
\]

Then \( E(w^*)^3 = w^3 \).

**Proof.** For \( w^* \) of (3),

\[
E(w^*)^3 = m_n^3 + 3m_n \nu / n = w^3
\]

if \( m_n^3 + 3m_n \nu / n - w^3 = 0 \). Set \( \gamma = (\nu / n)^3 + w^6 / 4 \). Since \( \gamma > 0 \), this cubic has one real root given by Equation (3.8.2) of Abramowitz and Stegun [1]:

\[
m_n = S_1^{1/3} + S_2^{1/3}
\]

where \( S_j = w^3 / 2 \pm \gamma^{1/2} \). Suppose that \( w > 0 \). (If not, replace \( w \) by \( |w| \).) Then for \( \nu \) of (4),

\[
\gamma^{1/2} = w^3 (1 + d)^{1/2} / 2,
\]

where \( d = cn^{-3} \). Furthermore,

\[
(1 + d)^{1/2} = \sum_{j=0}^{\infty} \left( 1/2 \right)_j d^j / j!, \quad S_1 = 1 + D
\]

for

\[
D = \sum_{j=1}^{\infty} \left( 1/2 \right)_j d^j / j^2 = \sum_{j=1}^{\infty} a_j d^j / j!.
\]

Furthermore,

\[
D^k / k! = \sum_{j=k}^{\infty} B_{j,k}(a) d^j / j!
\]

implies

\[
S_1^{1/3} = \sum_{k=0}^{\infty} \left( 1/3 \right)_k D^k = \sum_{j=0}^{\infty} A_j d^j / j!.
\]

Also

\[
S_2 = 1/2 - (1 + d)^{1/2} / 2 = - \sum_{j=1}^{\infty} a_j d^j / j! = -a_1 d (1 + U)
\]

for

\[
U = \sum_{k=1}^{\infty} B_k d^k / k!
\]

and

\[
U^j / j! = \sum_{k=j}^{\infty} B_{k,j}(B) d^k / k!.
\]

Then

\[
S_2^{1/3} = u (1 + U)^{1/3} = u \sum_{j=0}^{\infty} \left( 1/3 \right)_j U^j = u \sum_{j=0}^{\infty} \left( 1/3 \right)_j U^j / j! = \sum_{k=0}^{\infty} C_k d^k / k!,
\]

where

\[
C_k = \sum_{j=0}^{k} \left( (1/3)_j B_{j,k}(B) \right).
\]

Hence, for the choice of \( f_n(w) \), \( E(w^*)^3 = w^3 \).
The method of Theorems 1 and 3 will not work for \( t(w) = w^5 \) since there is no explicit solution to a quintic. However, we now show how to obtain an unbiased or bias-reduced estimate of \( w^a \) for any \( a > 0 \). Set \( \Delta = \hat{w} - w = w^* - m_n \). Then

\[
E\left[ (w^*)^a \right] = E\left( (m_n + \Delta)^a \right) = \sum_{j=0}^{\infty} \left( \begin{array}{c} a \\ j \end{array} \right) m_n^{a-j} \mu_j (\hat{w}) = \sum_{j=0}^{\infty} \left( \begin{array}{c} a \\ 2j \end{array} \right) m_n^{a-2j} \mu_{2j} (\hat{w}),
\]

where

\[
\mu_{2j} (\hat{w}) = N_j v^j, \quad N_0 = 1, \quad N_j = 1 \cdots (2j - 1)
\]

for \( j \geq 1 \). By (2),

\[
f_n(w)^k / k! = \sum_{i=k}^{\infty} B_{i,k}(b)n^{-i} / i!.
\]

So,

\[
(m_n / w)^a = \left[ 1 + f_n(w) / w \right]^{a} = \sum_{k=0}^{\infty} \left( \begin{array}{c} a \\ k \end{array} \right) w^{-k} f_n(w)^k / k! = \sum_{l=0}^{\infty} D_{a,i} n^{-i} / i!
\]

for

\[
D_{a,i} = \sum_{k=0}^{i} \left( \begin{array}{c} a \\ k \end{array} \right) w^{-k} B_{i,k}(b).
\]

This implies

\[
E\left[ (w^*)^a / w^a \right] = \sum_{k=0}^{\infty} n^{-k} E_k
\]

for

\[
E_k = \sum_{i+j=k} \left( \begin{array}{c} a \\ 2j \end{array} \right) N_j v^j D_{a-2,i,j} / i!.
\]

So,

\[
b_1 = -(a - 1) w / 2, \quad E_1 = 0, \quad E\left[ (w^*)^a \right] = w^a + O\left( n^{-2} \right),
\]

\[
b_2 = w(a - 1) \left[ -(a - 1)^2 + (a - 1) w v - (a - 2)^2 v^2 \right], \quad E_2 = 0, \quad E\left[ (w^*)^a \right] = w^a + O\left( n^{-3} \right).
\]

In this way, we can construct \( f_n(w) \) so that for any given \( a > 0 \) and \( k \geq 1 \),

\[
E\left[ \hat{w} + f_n(w) \right]^a = w^a + O\left( n^{-k} \right).
\]

References


