# Statistica Sinica Preprint No: SS-2015-0204.R3


| | |
|---|---|
| **Title** | Convex Mixtures Imputation and Applications |
| **Manuscript ID** | SS-2015-0204.R3 |
| **URL** | http://www.stat.sinica.edu.tw/statistica/ |
| **DOI** | 10.5705/ss.202015.0204 |
| **Complete List of Authors** | Jianhui Ning<br>Michelle Liou and<br>Philip E. Cheng |
| **Corresponding Author** | Michelle Liou |
| **E-mail** | mliou@stat.sinica.edu.tw |
| Notice: Accepted version subject to English editing. | |

Statistica Sinica

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)



where

$$\hat{m}_{CM}(X_i) = w_i\hat{m}_{KR}(X_i) + (1 - w_i)\hat{m}_{kNN}(X_i), \tag{2.2}$$

and $w_i$ is given in (1.8). By (1.5) and (1.8), the first summand of the convex estimate (2.2) is a local kernel regression estimate based on the observed responses, and the second summand furnishes the $k$-NN regression estimate using the non-observation (missing) weight. Thus, the convex mixtures (CM) estimator (2.1) is defined using the local convex regression estimate (2.2) to balance the trade-off effect between the sample bias and variance given by the two regression estimates. Similar to reducing the bias (hence the MSWE) of the KR (1.4) by using the IPW (1.9), the IPW version of the CM estimator is defined as

$$\hat{\mu}_{CMIPW} = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{m}_{CM}(X_i) + \frac{\delta_i\left[Y_i - \hat{m}_{CM}(X_i)\right]}{w_i}\right]. \tag{2.3}$$

By analogy with the asymptotic equivalence between KR and IPW of Lemma 1.1, the same asymptotic normality is acquired by both CM and CMIPW as stated below, and proved in the Appendix.

**Theorem 2.1** *Assume the regularity conditions (H), (S) and (W) in the Appendix. The imputation estimators $\hat{\mu}_{CM}$ of (2.1) and $\hat{\mu}_{CMIPW}$ of (2.3) approximates the same normal distribution*

$$\sqrt{n}(\hat{\mu}_{CM} - \mu) \to N(0, \sigma^2_{CM}),$$

*where*

$$\begin{aligned}\sigma^2_{CM} &= Var[m(X)] + E\left[\frac{\sigma^2(X)}{p(X)}\right] \\ &\quad + \frac{1}{k}E\left[\sigma^2(X)\{1 - p(X)\}^3\left(1 + \frac{1}{k}\right)\right],\end{aligned} \tag{2.4}$$

*and the first two terms on the right-hand side of (2.4) yields $\sigma^2_{KR}$ of (1.10).*

The asymptotic variances of the imputation estimators, the $k$-NN (1.6), IPW (1.9) (or, the KR and HT), CM (2.1) and CMIPW (2.3) can be compared as follows. It was known from Ning and Cheng (2012, Theorem 1) that

$$\sigma^2_{kNN} = \sigma^2_{KR} + \frac{1}{k}E\left[\sigma^2(X)\{1 - p(X)\}\right]. \tag{2.5}$$

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)



The asymptotic variances of (1.10) and (2.4) are related by the equation

$$\sigma^2_{CM} = \sigma^2_{KR} + \frac{1}{k}E\left[\sigma^2(X)\{1 - p(X)\}^3(1 + \frac{1}{k})\right]. \tag{2.6}$$

Like $k$-NN, under regularity conditions, the CM and CMIPW yield larger variances than KR and IPW, hence larger MSE, because these estimators are all asymptotically unbiased under regularity conditions. However, under non-regular conditions, KR and IPW may yield larger bias and MSE as mentioned after Lemma 1.1. In view of the two IPW versions, the IPW (of the KR) and CMIPW, it is possible to form a third IPW version using a combination of these two. This new version is termed a convex regression (CR) imputation estimator for the mean $\mu$, defined as

$$\hat{\mu}_{CR} = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{m}_{CM}(X_i) + \frac{\delta_i\{Y_i - \hat{m}_{KR}(X_i)\}}{w_i}\right]. \tag{2.7}$$

Under regularity conditions, the CR yields a different asymptotic normality from the previous estimators.

**Theorem 2.2** *Under the same conditions of Theorem 2.1, the CR imputation estimator $\hat{\mu}_{CR}$ approximates the normal distribution*

$$\sqrt{n}(\hat{\mu}_{CR} - \mu) \to N(0, \sigma^2_{CR})$$

*where*

$$\sigma^2_{CR} = \sigma^2_{KR} + \frac{1}{k}E\left[\sigma^2(X)\{1 - p(X)\}^2\right] \tag{2.8}$$

Like $\sigma^2_{CM}$, it is seen from (2.8) that $\sigma^2_{CR}$ is larger than $\sigma^2_{KR}$, but $\sigma^2_{CR}$ is smaller than $\sigma^2_{kNN}$,

$$\sigma^2_{kNN} = \sigma^2_{CR} + \frac{1}{k}E\left[p(X)\{1 - p(X)\}\sigma^2(X)\right]. \tag{2.9}$$

It is understood that valid theoretical results can only be acquired under regularity conditions, that is, no theory can be derived under non-regularity conditions, because sample bias and MSE can vary widely when regularity conditions are violated. As mentioned, typical non-regular conditions include the case having jump discontinuites of the missing pattern function $p(x)$ or the conditional variance function $\sigma^2(x)$, and the case where $p(x)$ may decrease towards zero over an interval within the domain of $X$. To