Pooled location and scale estimators

Chanseok Park* and Min Wang†

December 2022

Abstract

In this note, we provide a brief summary of pooled location and scale estimators, which are provided in the `pooledEstimator` function of the robust quality control chart (rQCC) R package. The `rcc` function is based on the pooled location and scale estimators provided here.

1 Location Estimation

The pooled location estimator has three types, denoted by A, B, C. Type A estimator is obtained by the unweighted average of the unbiased estimators. Type B estimator is obtained by the weighted average with the weight proportional to the sample size. Type C estimator is obtained by the optimally weighted average in a sense of the BLUE.

Suppose that there are m samples and $\hat{\mu}_i$ is an unbiased estimator of μ from the ith sample of size n_i. The `pooledEstimator` calculates the pooled estimator for the sample mean (default), the sample median, and the Hodges-Lehmann [1] estimator based on one of the following three pooling methods.

- Type A:

$$\bar{\hat{\mu}}_A = \sum_{i=1}^{m} w_i \hat{\mu}_i,$$

  where $w_i = 1/m$.

- Type B:

$$\bar{\hat{\mu}}_B = \sum_{i=1}^{m} w_i \hat{\mu}_i,$$

  where $w_i = n_i / \sum_{j=1}^{m} n_j$.

*Applied Statistics Laboratory, Department of Industrial Engineering, Pusan National University, Busan 46241, Korea. His work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2022R1A2C1091319).

†Department of Management Science and Statistics, The University of Texas at San Antonio, San Antonio, TX 78249, USA.

- Type C:
$$\bar{\hat{\mu}}_C = \sum_{i=1}^{m} w_i \hat{\mu}_i,$$
where w_i is given by the BLUE in [2].

It should be noted that $\bar{\hat{\mu}}_A$, $\bar{\hat{\mu}}_B$, and $\bar{\hat{\mu}}_C$ are all unbiased for μ.

2 Scale Estimation

The `pooledEstimator` also calculates the pooled sclae estimator for the standard deviation (sd), range, median absolute deviation, and Shamos estimators based on one of the following three pooling methods.

Suppose that there are m samples and $\hat{\sigma}_i$ is an estimator of σ from the ith sample of size n_i. Let C_i is an unbiasing factor so that $E(\hat{\sigma}_i)/C_i = \sigma$. Type A estimator is obtained by the unweighted average of the unbiased estimators. Type B estimator is obtained by the weighted average with the weight proportional to the unbiasing factors. Type C estimator is obtained by the optimally weighted average in a sense of the BLUE.

- Type A:
$$\bar{\hat{\sigma}}_A = \sum_{i=1}^{m} w_i \cdot \frac{\hat{\sigma}_i}{C_i},$$
where $w_i = 1/m$.
- Type B:
$$\bar{\hat{\sigma}}_B = \sum_{i=1}^{m} w_i \cdot \frac{\hat{\sigma}_i}{C_i}$$
where $w_i = C_i / \sum_{j=1}^{m} C_j$.
- Type C:
$$\bar{\hat{\sigma}}_C = \sum_{i=1}^{m} w_i \frac{\hat{\sigma}_i}{C_i},$$
where w_i is given by the BLUE in [2].

It should be noted that $\bar{\hat{\sigma}}_A$, $\bar{\hat{\sigma}}_B$, and $\bar{\hat{\sigma}}_C$ are all unbiased for σ.

3 R Usages

```
> # Type A with mean (default)
> x1 = c(1,2)
> x2 = c(2,5,9)
> data = list(x1, x2)
> pooledEstimator(data)
```

```
> # Type C with HL1 (Hodges-Lehmann)
> pooledEstimator(data, estimator="HL1", poolType="C")
```

```
> # Type A with standard deviation (sd)
> pooledEstimator(data, estimator="sd")
```

```
> # Type C with Shamos
> pooledEstimator(data, estimator="shamos", poolType="C")
```

References

[1] J. L. Hodges and E. L. Lehmann. Estimates of location based on rank tests. *Annals of Mathematical Statistics*, 34:598–611, 1963.

[2] C. Park, L. Ouyang, and M. Wang. Development of robust x-bar charts with unequal sample sizes. `http://dx.doi.org/10.48550/arXiv.2212.10731`, 2022. ArXiv e-prints.