A BOOTSTRAP TEST FOR SYMMETRY BASED ON RANKED SET SAMPLES

REZA DRIKVANDI[1], REZA MODARRES[2] AND ABDULLAH H. JALILIAN[1]

[1] *Department of Statistics, Shahid Beheshti University, 1983963113 G.C, Tehran, Iran*
[2] *Department of Statistics, The George Washington University, Washington, DC, USA*

Abstract. To test the hypothesis of symmetry about an unknown median we propose the maximum of a partial sum process based on ranked set samples. We discuss the properties of the test statistic and investigate a modified bootstrap ranked set sample bootstrap procedure to obtain its sampling distribution. The power of the new test statistic is compared with two existing tests in a simulation study.

Keywords: *Ranked Set Sampling; Symmetry; Bootstrap; Simulation.*

1. **Introduction**

Ranked set sampling (RSS) is a sampling protocol that can often improve the cost efficiency of an experiment. It is appropriate for situations where quantification of sampling units is costly or difficult, but ranking the units in a small set is easy and inexpensive.

In RSS one first draws k^2 units at random from the population and partitions them into k sets of k units. The k units in each set are ranked without making actual measurements. Ranking can be performed based on concomitant variables, expert judgment, visual inspection or any means that does not involve actually quantifying the observations. For example, in environmental sampling, the level of a soil contaminant can be costly to measure, however; we might use soil texture, color or PH level to rank the experimental units. In the rth set the observation with the rth judgment rank is quantified for $r = 1, \ldots, k$. The entire process is repeated m cycles and yields the ranked set sample $\left\{X_{(r)j}, r = 1, ..., k;\ j = 1, ..., m\right\}$. For an excellent treatment of the RSS literature, see Chen et al. (2004).

Important reasons for general interest in testing symmetry can be found in Antille et al. (1982). Based on simple random samples (SRS), many authors, including Gupta (1967), Butler (1969), Gastwirth (1971), Rothman and Woodroofe (1972), Boos (1982), Modarres and Gastwirth (1996), have proposed statistics for detecting asymmetry. Based on ranked set samples, Ozturk (2001) constructed an RSS analog of the Rothman-Woodroofe (1972) statistic for testing symmetry and Hui and Modarres (2007) proposed a sign test for the hypothesis of symmetry when the median of the distribution is assumed known.

In this paper, we assume that a ranked set sample $\left\{X_{(r)j}, r = 1, ..., k;\ j = 1, ..., m\right\}$ is drawn from a continuous differentiable distribution F with density function f and unknown median θ. Our interest is to test the hypothesis that F is symmetric about θ against asymmetric alternatives.

When both F and θ are unknown, one approach is to estimate θ by a consistent estimator $\hat{\theta}$ and use it in a test statistic for symmetry about a known θ. The problem with this approach is that the resulting tests are not distribution-free. We sidestep this problem by using a modified bootstrap procedure to estimate the associated critical and/or p-values.

In Section 2, we discuss the existing statistics and define a new statistic for testing symmetry. In Section 3, we describe a modified bootstrap procedure to obtain the sampling distribution of the proposed test. In Section 4, we compare the new test with those of Ozturk (2001) and Hui and Modarres (2007) via a simulation study and make recommendations. We conclude the article in Section 5.

2. **The Test Statistic**

For a known center of symmetry $\theta \in \mathbb{R}$, let

$\mathcal{P}^\theta$: set of all continuous distributions with median θ;
$\mathcal{P}_S^\theta$: set of all symmetric continuous distributions with median θ.

Based on a sample from F, $F \in \mathcal{P}^\theta$, the testing problem is

$$\begin{cases} H_0: & F \in \mathcal{P}_S^\theta \\ H_1: & F \in \mathcal{P}^\theta - \mathcal{P}_S^\theta \end{cases}$$

which is equivalent to (see Theorem 2.1)

$$\begin{cases} H_0: & F(x) + F(2\theta - x) = 1 \ ; \ \forall x \in \mathbb{R} \\ H_1: & F(x_0) + F(2\theta - x_0) \neq 1 \ ; \ \exists x_0 \in \mathbb{R}. \end{cases}$$

Existing tests of symmetry that use a simple random sample (SRS) $\{X_j : j = 1, \ldots, n\}$ include the the sign test $S(\theta) = \sum_{j=1}^n I(X_j > \theta)$, and its variants (Gastwirth, 1971), the modified runs test (Modarres and Gastwirth, 1996) and a modified sign test by Cheng and Balakrishnan (2004).

Test statistics that are based on measures of discrepancy include Butler (1969) statistic

$$h_n(\theta) = \sup_{x \in \mathbb{R}} \left| \hat{F}_n(x) + \hat{F}_n(2\theta - x) - 1 \right|,$$

and the Rothman-Woodroofe (1972) statistic

$$h_n(\theta) = \int_{\mathbb{R}} \left(\hat{F}_n(x) + \hat{F}_n(2\theta - x) - 1 \right)^2 dx, \tag{2.1}$$

where

$$\hat{F}_n(x) = \frac{1}{n} \sum_{j=1}^n I(X_j \leq x) \ ; \ x \in \mathbb{R},$$

is the empirical cdf based on the SRS.

An RSS analog of the SRS sign test (Hettmansperger, 1995) is based on the statistic

$$S_m(\theta) = \sum_{r=1}^k \sum_{j=1}^m I(X_{(r)j} > \theta)$$

where

$$P_{H_0}\{S_m = y\} = \sum_{i_1 + \cdots + i_k = y} \prod_{r=1}^k \binom{m}{i_j} p_r^{i_j} (1 - p_r)^{m - i_j} \ ; \ y = 0, \ldots, mk$$

where

$$p_r = \text{Beta}(r, k-r+1, \frac{1}{2}).$$

This test does not use the rank information and has a lower power than other test statistics.

In order to take into account the rank information available in RSS, Ozturk (2001) and Hui and Modaress (2007) considered the fact that for all $F \in \mathcal{P}_S^\theta$,

$$F_{(r)}(x) = 1 - F_{(k-r+1)}(2\theta - x), \quad \forall r = 1, 2, ..., k, \ \forall x \in \mathbb{R}, \tag{2.2}$$

where $F_{(r)}$ is the distribution function of the r-th order statistic based on a SRS from F with size k. Assuming θ is known, Hui and Modaress (2007) introduced a modified sign test with test statistic

$$B_m(\theta) = \begin{cases} \sum_{r=1}^{\frac{k}{2}} \sum_{j=1}^{m} I\big(X_{(r)j} + X_{(k-r+1)j} > 2\theta\big), k \text{ is even} \\ \sum_{r=1}^{\frac{k-1}{2}} \sum_{j=1}^{m} I\big(X_{(r)j} + X_{(k-r+1)j} > 2\theta\big) + \sum_{j=1}^{m} 1\big(X_{(\frac{k+1}{2})j} > \theta\big), k \text{ is odd} \end{cases}$$

which, under H_0, is distributed as

$$B_m \sim \text{Bin}\big(m\Big[\frac{k+1}{2}\Big], \frac{1}{2}\big).$$

Since for all $F \in \mathcal{P}_S^\theta$,

$$\begin{aligned} F(x) + F(2\theta - x) - 1 &= \frac{1}{k}\sum_{r=1}^{k} F_{(r)}(x) + \frac{1}{k}\sum_{r=1}^{k} F_{(k-r+1)}(2\theta - x) - 1 \\ &= \frac{1}{k}\sum_{r=1}^{k}\Big[F_{(r)}(x) + F_{(k-r+1)}(2\theta - x) - 1\Big]; \ \forall x \in \mathbb{R}, \end{aligned}$$

we will show that (2.2) is the sufficient and necessary condition for symmetry of underlying distribution F.

Theorem 2.1. *Suppose ranking in RSS is perfect. $F \in \mathcal{P}_S^\theta$ if and only if*

$$\forall r = 1, 2, ..., k, \ \forall x \in \mathbb{R}: \ F_{(r)}(x) = 1 - F_{(k-r+1)}(2\theta - x).$$

Proof: See Appendix.

Ozturk (2001) proposed the following test statistic

$$T_m\left(\hat{\theta}\right) = mk\left\{-1 + \frac{\sum_{j=1}^{m}\sum_{r=1}^{k}\sum_{i=1}^{m}\sum_{r'=1}^{k}\left|X_{(r)j} + X_{(r')i} - 2\hat{\theta}\right|}{\sum_{j=1}^{m}\sum_{r=1}^{k}\sum_{i=1}^{m}\sum_{r'=1}^{k}\left|X_{(r)j} - X_{(r')i}\right|}\right\},$$

where $\hat{\theta}$ is a suitable estimator for θ. He used a bootstrap procedure for testing H_0 based on T_m.

To explain the rationale for the new test statistic proposed in this paper, first consider the case where θ is known and note the difference process

$$d_{rj} = \frac{X_{(r)j} + X_{(k-r+1)j}}{2} - \theta \qquad r = 1, ..., k, \quad j = 1, 2, ...m.$$

Each component in the above process has zero mean under H_0. To measure deviations of the difference process from 0 we propose to use the maximum of the partial sum process as the test statistic

$$D_m(\theta) = \frac{1}{\sqrt{km}} \max_{i=1}^{k} \left| \sum_{r=1}^{i} \sum_{j=1}^{m} d_{rj} \right|, \tag{2.3}$$

This is similar in spirit to Lim et al. (2008), who used the maximum residual process to test constancy of mean in repeated measures data.

Now, we consider the case in which θ is unknown. When θ is unknown, we estimate it by $\hat{\theta}$ and compute $D_m(\hat{\theta})$ to test H_0. Hence,

$$D_m(\hat{\theta}) = \frac{1}{\sqrt{km}} \max_{i=1}^{k} \left| \sum_{r=1}^{i} \sum_{j=1}^{m} (\frac{X_{(r)j} + X_{(k-r+1)j}}{2} - \hat{\theta}) \right|,$$

where $\hat{\theta}$ is an unbiased and consistent estimator of θ under H_0. We use $\bar{X}_{\mathrm{RSS}}$ (the RSS sample mean) to estimate θ.

Under H_0, the asymptotic distribution of the test statistic (D_m) is in the form of maxima of the Brownian motion when (i) $m \to \infty$ with a fixed k and (ii) $k \to \infty$ with a fixed m. For details, see Lim et al. (2008).

3. **Bootstrap Test For Symmetry**

In this section, we discuss a bootstrap test for symmetry based on a ranked set sample. Since finding the sampling distributions of statistics under RSS is often difficult, inference is typically drawn using asymptotic results. Moreover, depending on the statistic of interest, asymptotic results may not be readily available and, when they are available, they may not be valid for the finite sized sample at hand. Chen et al. (2004) suggest a natural method to obtain bootstrap samples from each row of RSS. Modarres et al. (2006) introduced two other methods that are designed to obtain more stratified resamples from the given sample. We use the bootstrap RSS (BRSS) algorithm of Modarres et al. (2006) to estimate the distribution of D_m.

Suppose H_0 is true and we want to estimate the distribution of D_m. Let $\hat{F}_{n,RSS}$ be the empirical cdf based on the ranked set sample. Since $\hat{F}_{n,RSS}$ is not symmetric, bootstrap samples should not be taken from $\hat{F}_{n,RSS}$, but rather from a symmetric distribution which is close to $\hat{F}_{n,RSS}$. We study bootstrap samples taken from the closest symmetric distribution to $\hat{F}_{n,RSS}$ in integrated square error (or L^2-norm) sense. Schuster and Barker (1987) call this procedure the symmetric bootstrap.

Theorem 3.1. *Given an RSS sample from $F \in \mathcal{P}^{\theta}$, the closest symmetric distribution function to*

$$\hat{F}_{n,RSS}(x) = \frac{1}{km} \sum_{r=1}^{k} \sum_{j=1}^{m} I\big(X_{(r)j} \leq x\big)$$

is

$$G_n(x) = \frac{1}{2}\Big(\hat{F}_{n,RSS}(x) + 1 - \hat{F}_{n,RSS}(2\theta - x)\Big)$$

Proof: See Appendix.

Note that G_n is empirical distribution function of a random sample

$$W = \{X_{(1)1}, X_{(1)2}, \ldots, X_{(k)m}, 2\theta - X_{(1)1}, 2\theta - X_{(1)2} \ldots, 2\theta - X_{(k)m}\}$$

with size $2km$. Thus, generating a random sample of size n from G_n is equivalent to sampling n times with replacement from W. For an RSS sample, to decide whether to reject H_0 or accept H_0 based on test statistic D_m, we propose this algorithm:

(1) Compute observed D_m for original RSS sample.
(2) Generate $b = 1000$ RSS samples with the same k and m values as original sample from the closest symmetric distribution function G_n,
(3) Compute D_m for each of these bootstrap samples,
(4) Compute percentage of bootstrap samples with D_m values greater than observed D_m,
(5) If this percentage is greater than $\alpha = 0.05$, accept H_0, otherwise reject H_0.

In next section, Monte Carlo studies are performed to compute and compare the power of B, T and D tests.

4. **Monte Carlo Power Estimation**

Let $\pi_{k,m}(F)$ denote the power of D-test at F. Given k, m, $F \in \mathcal{P}^\theta$, and the sample $\left\{X_{(r)j} : r = 1, ..., k;\ j = 1, ..., m\right\}$ from F, $\pi_{k,m}(F)$ is the probability of rejecting H_0. By controlling Type I error at $\alpha = 0.05$, we expect that

$$\pi_{k,m}(F) \le 0.05\ ;\ \forall F \in \mathcal{P}_S^\theta,$$

and as a powerful test, we expect that for all $F \in \mathcal{P}^\theta - \mathcal{P}_S^\theta$, $\pi_{k,m}(F)$ be as large as possible.

We selected the standard Normal, standard Cauchy and standard Logistic distributions from $\mathcal{P}_S^\theta$. To represent asymmetric distributions we selected Gamma distribution with parameter $(\alpha, \beta) = (5, 7)$, and $(\alpha, \beta) = (3, 5)$, Fisher-Snedecor or F distribution with parameter $(n_1, n_2) = (8, 9)$ and Log-Normal distribution with parameter $(\mu, \sigma^2) = (0, 1)$ from $\mathcal{P}^\theta - \mathcal{P}_S^\theta$ were selected. The skewness values of these distributions are 0.89, 1.15,4.24 and 6.18, respectively. For each distribution, we obtained 2000 ranked set samples with $k = 4, 5$ and $m = 5, 10, 15, 20$. We obtained the rejection frequency of B, and T and D tests. A bootstrap procedure with $b = 1000$ was used to obtain the sampling distribution of $T$and D tests. Test statistic B was based on an center.

Table 1 and Table 2 show the performance of the three test statistics for symmetry under different distributions and sample sizes.

Table 1 shows that for symmetric distributions T and D tests perform well and have Type I error rate close to the nominal 0.05 level. B-test has inflated levels under a logistic distribution. Table 2, for asymmetric distributions, shows that when the skewness of the distribution is low (e.g, gamma), D-test is more powerful than B-test and T-test. But, for extremely skewed distributions (e.g., lognormal) T-test has better performance. It should be noted that increasing m leads to more growth in power for D-test in comparison to T-test. Since detecting the asymmetry of underling distribution is difficult when the skewness is small, we recommend to use D-test for data with small skewness.

Table 1. Percent rejections at the 5% level for B , T and D tests (the first, second and the third entry of each cell). Symmetric distributions are Normal, Cauchy and Logistic.

Symmetric Distributions Density Function	k	Power Estimates m			
$N(0,1)$		5	10	15	20
	4	6.4 4.8 5.0	6.5 4.7 4.9	6.1 5.4 4.5	6.0 4.6 4.6
	5	7.9 4.0 5.0	6.4 4.6 4.4	5.4 5.3 4.9	3.9 4.2 4.3
$C(0,1)$		5	10	15	20
	4	6.7 2.8 2.6	8.1 2.7 2.2	7.0 2.9 2.4	6.4 3.0 2.0
	5	6.7 2.6 2.4	7.5 3.4 2.3	5.5 3.7 2.2	4.8 3.5 2.1
$L(0,1)$		5	10	15	20
	4	7.3 5.3 4.9	8.1 6.1 5.7	7.2 5.5 5.6	6.8 5.0 5.1
	5	7.1 4.6 4.7	6.6 4.3 4.5	6.1 4.9 5.5	5.5 4.4 4.5

5. **Conclusion**

For testing symmetry based on RSS samples, we proposed a distribution-free test based on the maximum of a partial sum process. Although the asymptotic distribution of the test statistic is in the form of maxima of the Brownian motion, in practice the size of the data set is not always large enough to apply the asymptotic results. To overcome this problem, we used a symmetric bootstrap procedure to approximate the finite-sample distribution of the test statistic. We have demonstrated, via simulation, that the proposed test has Type I error rate close to the nominal level. In addition, the results of the simulations indicate that the proposed test is more powerful than the test proposed by Hui and Modarres (2007). Moreover, for the data with small skewness, the proposed test is more powerful than the test proposed by Ozturk (2001). A power advantage of the proposed test statistic with respect to the test statistic proposed by Ozturk (2001) is that the asymptotic distribution of the proposed test statistic can be explicitly derived.

Table 2. Percent rejections at the 5% level for B , T and D tests (the first, second and the third entry of each cell). Asymmetric distributions are Gamma(5,7), Gamma(3,5), F(8,9) and Log-Normal with skewness values 0.89, 1.15,4.24 and 6.18, respectively.

Asymmetric Distributions Density Function	k	m = 5	10	15	20
$G(5,7)$		5	10	15	20
	4	6 7 11	11 18 24	15 29 35	18 42 48
	5	9 13 18	13 29 35	16 48 55	18 65 67
$G(3,5)$		5	10	15	20
	4	6 14 16	14 34 39	22 56 59	26 71 72
	5	11 23 28	18 52 55	25 75 77	31 87 88
$F(8,9)$		5	10	15	20
	4	9 34 33	26 77 66	37 94 92	46 99 97
	5	14 55 50	30 91 87	39 98 96	48 100 100
$LN(0,1)$		5	10	15	20
	4	12 59 46	41 97 85	58 100 97	69 100 100
	5	19 79 63	43 99 95	61 100 99	76 100 100

References

Antille, A., Kersting, G. and Zucchini, W. (1982). Testing symmetry. *Journal of the American Statistical Association.* 77, 639-646.

Boos, D. D. (1982). A test for asymmetry associated with the Hodges-Lehmann estimator. *Journal of the American Statistical Association.* 77, 647-651.

Butler, C. C. (1969). A test for symmetry using the sample distribution function. *the Annals of Mathematical Statistics.* 40, 2209-2210.

Chen, Z., Bai, Z. and Sinha, B. (2004). Ranked set sampling: Theory and application. *New York: Springer-Verlag.*

Cheng, W. H. and N. Balakrishnan, (2004). A modified sign test for symmetry. *Communications in Statistics - Simulation and Computation.* 33(3), 703-709.

Gastwirth, J. L. (1971). On the sign test of symmetry. *Journal of the American Statistical Association.* 66, 821-823.

Gupta, M. K. (1967). An asymptotically nonparametric test of symmetry. *the Annals of Mathematical Statistics.* 4, 263-270.

Hettmansperger, T. P. (1995). The ranked-set sample sign test. *Journal of Nonparametric Statistics.* 4, 263-270.

Hui, T. and Modarres, R. (2007). A test for symmetry using ranked set samples. *Technical Report, Department of Statistics. George Washington University.*

Lim, J., Wang, X., Lee, S. and Jung, S. H., (2008). A distribution-free test of constant mean in linear mixed effects models. *Statistics in Medicine.* 27, 3833-3846.

Modarres, R. and Gastwirth, J. L. (1996). A modified runs test for symmetry. *Statistics and Probability Letters.* 31, 107-112.

Modarres, R., Hui, T. and Zheng, G. (2006). Resampling methods for ranked set samples. *Computational Statistics and Data Analysis.* 51, 1039-1050.

Ozturk, O. (2001). A nonparametric test of symmetry versus asymmetry for ranked-set samples. *Communications in Statistics - Theory and Methods.* 30(10), 2117-2133.

Rothman, E. D. and Woodroofe, M. (1972). A Cramer-von Mises type statistic for testing symmetry. *the Annals of Mathematical Statistics.* 43, 2035-2038.

Appendix A. **Proof of Theorems**

A.1. **Proof of Theorem 2.1.** Sufficiency: If f is symmetric about $\theta \in \mathbb{R}$, then for all $r = 1, \dots, k$ and for all $x \in \mathbb{R}$,

$$\begin{aligned} F_{(r)}(x) \;=\; \mathrm{Beta}\big(r, k-r+1, F(x)\big) &= k\binom{k-1}{r-1}\int_0^{F(x)} y^{r-1}(1-y)^{k-r}dy \\ &= k\binom{k-1}{r-1}\int_{1-F(x)}^{1}(1-u)^{r-1}u^{k-r}du \\ &= 1-\mathrm{Beta}\big(k-r+1, r, 1-F\,(x)\big) \\ &= 1-\mathrm{Beta}\big(k-r+1, r, F(2\theta-x)\big) \\ &= 1-F_{(k-r+1)}(2\theta-x). \end{aligned}$$

Necessity: We know that

$$F(x) = \frac{1}{k}\sum_{r=1}^{k} F_{(r)}(x)\ ;\ \forall x \in \mathbb{R}.$$

Thus, for all $x \in \mathbb{R}$, when k is even,

$$\begin{aligned} F(x) &= \frac{1}{k}\sum_{r=1}^{\frac{k}{2}}\Big(F_{(r)}(x) + F_{(k-r+1)}(x)\Big) \\ &= \frac{1}{k}\sum_{r=1}^{\frac{k}{2}}\Big(1 - F_{(k-r+1)}(2\theta-x) + F_{(k-r+1)}(x)\Big) \\ &= \frac{1}{k}\sum_{r=1}^{\frac{k}{2}}\Big(2 - F_{(k-r+1)}(2\theta-x) - F_{(r)}(2\theta-x)\Big) \\ &= \frac{1}{k}(2\frac{k}{2}) - \frac{1}{k}\left[\sum_{r=1}^{\frac{k}{2}} F_{(r)}(2\theta-x) + \sum_{r=\frac{k}{2}}^{k} F_{(r)}(2\theta-x)\right] \\ &= 1 - \frac{1}{k}\sum_{r=1}^{k} F_{(r)}(2\theta-x) = 1 - F(2\theta-x), \end{aligned}$$

and when k is odd, similarly,

$$\begin{aligned} F(x) &= \frac{1}{k}\left(\sum_{r=1}^{\frac{k-1}{2}}\Big(F_{(r)}(x) + F_{(k-r+1)}(x)\Big) + F_{(\frac{k+1}{2})}(x)\right) \\ &= \frac{1}{k}\left(\sum_{r=1}^{\frac{k-1}{2}}\Big(2 - F_{(k-r+1)}(2\theta-x) - F_{(k-r+1)}(2\theta-x)\Big) + \big(1 - F_{(\frac{k+1}{2})}(2\theta-x)\big)\right) \\ &= \frac{1}{k}(2\frac{k-1}{2}+1) - \frac{1}{k}\left[\sum_{r=1}^{\frac{k-1}{2}} F_{(r)}(2\theta-x) + \sum_{r=\frac{k+1}{2}}^{k} F_{(r)}(2\theta-x) + F_{(k-r+1)}(2\theta-x)\right] \\ &= 1 - \frac{1}{k}\sum_{r=1}^{k} F_{(r)}(2\theta-x) = 1 - F(2\theta-x). \end{aligned}$$

A.2. **Proof of Theorem 3.1.** For any $f \in L^2$, the L^2-norm is defined by
$\|f\| = \left[\int |f(x)|^2 dx\right]^{\frac{1}{2}}$.
Now, under L^2-norm, for symmetric (around θ) distribution G, we have

$$
\begin{aligned}
d\left(\hat{F}_{n,RSS}, G_n\right) &= \left\|\hat{F}_{n,RSS} - G_n\right\| = \left[\int \left|\hat{F}_{n,RSS}(x) - G_n(x)\right|^2 dx\right]^{\frac{1}{2}} \\
&= \left[\int \left|\hat{F}_{n,RSS}(x) - \frac{1}{2}\left[\hat{F}_{n,RSS}(x) + 1 - \hat{F}_{n,RSS}(2\theta - x)\right]\right|^2 dx\right]^{\frac{1}{2}} \\
&= \frac{1}{2}\left[\int \left|\hat{F}_{n,RSS}(x) - 1 + \hat{F}_{n,RSS}(2\theta - x)\right|^2 dx\right]^{\frac{1}{2}} \\
&= \frac{1}{2}\left[\int \left|\hat{F}_{n,RSS}(x) - G(x) + G(x) - 1 + \hat{F}_{n,RSS}(2\theta - x)\right|^2 dx\right]^{\frac{1}{2}} \\
&\leq \frac{1}{2}\left(\left[\int \left|\hat{F}_{n,RSS}(x) - G(x)\right|^2 dx\right]^{\frac{1}{2}} + \left[\int \left|1 - G(x) - \hat{F}_{n,RSS}(2\theta - x)\right|^2 dx\right]^{\frac{1}{2}}\right) \\
&\leq \frac{1}{2}\left(\left[\int \left|\hat{F}_{n,RSS}(x) - G(x)\right|^2 dx\right]^{\frac{1}{2}} + \left[\int \left|G(2\theta - x) - \hat{F}_{n,RSS}(2\theta - x)\right|^2 dx\right]^{\frac{1}{2}}\right) \\
&\leq \frac{1}{2}\left(2\left[\int \left|\hat{F}_{n,RSS}(x) - G(x)\right|^2 dx\right]^{\frac{1}{2}}\right) \\
&= \left[\int \left|\hat{F}_{n,RSS}(x) - G(x)\right|^2 dx\right]^{\frac{1}{2}} = d\left(\hat{F}_{n,RSS}, G\right),
\end{aligned}
$$

where first inequality is derived by Minkowsky inequality.
Thus, we conclude that

$$
d\left(\hat{F}_{n,RSS}, G_n\right) = \inf_G \left\{d\left(\hat{F}_{n,RSS}, G\right)\right\}.
$$