Journal of Statistics and Computer Science
Vol. 1, No. 1, 2022, pp. 1-9

URL : www.arfjournals.com
https://DOI: 10.47509 /JSCS.2022.v01i01.01

On Application of the Cramér-von Mises Distance for Equivalence Testing

Vladimir Ostrovski
ERGO Group AG, ERGO-Platz 1, 40477 Düsseldorf
Email: vladimir.ostrovski77@googlemail.com

Received April 10, 2022 • Revised June 2 & 7, 2022 • Accepted June 9, 2022 • Published July 15, 2022

Abstract: We consider an equivalence test for a fully specfied continuous distribution on ℝ. The equivalence test is a powerful tool to show that observed data are sufficiently close to a given distribution. The test under consideration is based on the time-proven Cramér-von Mises distance. We show that the test is locally asymptotically most powerful. A consistent estimator for the asymptotic variance of the test statistic is provided. The bootstrap percentile-t method is applied to improve the finite sample performance of the equivalence test. A detailed algorithm for the asymptotic and percentile-t tests is given. An extensive simulation study of the finite sample properties is performed. A practical approach to find efficient values of the tolerance parameter is provided.

MSC 2020: 62G10

Keywords: Testing equivalence, Cramér-von Mises distance, equivalence test, neighborhood-of-model validation

To cite this article

Vladimir Ostrovski. (2022). On Application of the Cramér-von Mises Distance for Equivalence Testing. *Journal of Statistics and Computer Science*. Vol. 1, No. 1, pp. 1-9. *https://DOI: 10.47509 /JSCS.2022.v01i01.01*

1 Introduction

A classic problem in test theory is to assess whether observed data are compatible with a fully specified probability measure. The common practice is to use goodness-of-fit tests for this purpose. However, goodness-of-fit tests are tailored to establish lack of fit to a given distribution, see Hodges and Lehmann (1954), Berger and Delampady (1987) and Rao and Lovric (2016) for discussions on this topic. Equivalence testing is an appropriate statistical approach to claim that observed data are suciently close to a specified distribution, see Wellek (2010) for a comprehensive review.

Let $C_1(\mathbb{R})$ be a set of continuous cumulative distribution functions (CDF's) on $\mathbb{R}$ and let $G \in C_1(\mathbb{R})$ denote a fully specified CDF. We assume that observed data follow an unknown continuous distribution with a CDF $F \in C_1(\mathbb{R})$. Baringhaus and Henze (2017) recently proposed an equivalence test based on the well-known Cramér-von Mises distance

$\Delta(F, G) = \mathbb{R}\ (F(x) = \int G(x))^2\, dG(x)$ and provided a nice probabilistic interpretation of this distance. Baringhaus, Gaigall, and Thiele (2018) consider, among other things, specific equivalence tests for uniformity that are also based on the Cramér-von Mises distance. These tests are designed for the uniform distributions on closed intervals $[a, b]$, where the parameters a and b are unknown.

The equivalence test problem is $H_0 = \{\Delta(F, G) \geq \varepsilon\}$ and $H_1 = \{\Delta(\mathrm{F}, \mathrm{G}) < \varepsilon\}$, where $\varepsilon > 0$ is a tolerance parameter. The observations $X_1, \ldots, X_n$ are independently and identically distributed according to F, where n denotes the sample size. The empirical CDF F_n of $X_1, \ldots, X_n$ can be used as a plug-in estimator of the unknown CDF F. Applying the usual normalization we obtain the test statistic $T(F_n) = \sqrt{n}(\Delta(F_n, G) - \varepsilon)$. The test statistic $T(F_n)$ has a computationally simple form because

$$n\Delta(F_n, G) = \frac{1}{12n} + \sum_{i=1}^{n}\left(U_{(i)} - \frac{2i-1}{2n}\right)^2$$

is the time-proven Cramér-von Mises statistic, where $U_i = G(X_i)$ for all $i = 1, \ldots, n$ and $U_{(1)} \leq \ldots \leq U(n)$ are the order statistics of $U_1, \ldots, U_n$.

Remark 1. Let U_n denote the empirical CDF of the transformed observations $U_1, \ldots, U_n$ and let U denote the CDF of the uniform distribution on $[0, 1]$. The equality $\Delta(Fn, G) = \Delta(U_n, U)$ can be shown using the substitution $y = G(x)$ as follows:

$\int(F_n(x) - G(x))^2\, dG(x) = \int(U_n(G(x)) - G(x))^2\, dG(x) = \int_0^1 (U_n(y) - y)^2 dy$. Therefore, testing equivalence to a continuous CDF G amounts to the testing equivalence of the transformed data to the uniform distribution on $[0, 1]$.

2. Asymptotic Optimality

In this section, we show that the asymptotic α-level equivalence tests based on the Cramér-von Mises statistic are locally asymptotically most powerful (LAMP), see van der Vaart (1998, Chapter 25) for an introduction to the asymptotic semi-parametric theory. Let $F_0 \in C_1(\mathbb{R})$ be a fixed CDF such that $\Delta(F_0, G) = \varepsilon$. Let $[-\delta, \delta] \to C_1(\mathbb{R})$, $t \to F_t$ be a parametric sub-model for $\delta > 0$, such that $\Delta(F_t, G) < \varepsilon$ for $t < 0$ and $\Delta(F_t, G) > \varepsilon$ for $t > 0$. We consider local alternatives $F_{t/\sqrt{n}}$ for $t > 0$ and $n \to \infty$. If the curve $t \to F_t$ is mean square dierentiable at F_0 then there exists an asymptotic upper bound for the power of asymptotic α-level tests at the local alternatives, see van der Vaart (1998, p.384, Theorem 25.44). An asymptotic α-level test is called LAMP if its power function attains this upper bound asymptotically for all sub-models, which are mean square dierentiable at F_0. First we derive the efficient influence function of the statistical functional $F \to \Delta(F, G)$.

Proposition 2. The statistical functional $\kappa : \mathrm{F} \to \Delta(F, G)$ is differentiable at F_0 with the efficient influence function

$$\tilde{\kappa}(x) = 2\int (F_0(s) - G(s))\,(1_{(-\infty,\,s]}(x) - F_0(s)dG(s).$$

Proof. Let $[-\delta, \delta] \to C_1(\mathbb{R})$, $t \to F_t$, be a parametric sub-model for $\delta > 0$, which is mean square dierentiable at F_0 with tangent h. Applying Bickel, Klaassen, Ritov, and Wellner (1993, p. 457, Proposition 2), we compute the derivative $\frac{\partial}{\partial t} F_t(s) = \frac{\partial}{\partial t}\int 1_{(-\infty,\,s]}(x) dF_t(x) = \int h(x) 1_{(-\infty,\,s]}(x) dF_0(x)$ at $t = 0$. Using differentiation under the integral, we obtain

$$\frac{\partial}{\partial t}\Delta(F_t, G) = 2\int (F_0(s) - G(s))\,[\int h(x)\,1_{(-\infty,\,s]}(x)\,dF_0(x)\,dG(s) \text{ at } t = 0.$$

By Fubini's theorem, we conclude:

$$\frac{\partial}{\partial t}\Delta(F_t, G) = \int h(x)[2\int (F_0(s) - G(s))\,1_{(-\infty,\,s]}(x) dF(x)\,dF_0(s) \text{ at } t = 0.$$

Therefore, the function $\phi(x) = 2\int (F_0(s) - G(s))\,1_{(-\infty,\,s]}(x)\,dG(s)$ is an influence function of κ at F_0. The function $\phi(x)$ is continuous because

$$\begin{aligned}|\phi(x + \in) - \phi(x)| &= |2\int (F_0(s) - G(s))\,(1_{(-\infty,\,s]}(x + \in) - 1_{(-\infty,\,s]}(x) dG(s)| \\ &\le 2\int |F_0(s) - G(s))|\,|(1_{(-\infty,\,s]}(x + \in) - 1_{(-\infty,\,s]}(x)| dG(s) \\ &\le 2\int (1_{(-\infty,\,s]}(x + \in) - 1_{(-\infty,\,s]}(x) dG(s) \\ &= 2G(x + \in) - G(x))\end{aligned}$$

for any $\in > 0$. It is easy to show that $\int \phi^2 dF_0 < \infty$. We do not impose any additional constraints on the CDF's $F_t \in C_1(\mathbb{R})$. Hence, the tangent space of F_0 contains all continuous functions g such that $\int g dF_0 = 0$ and $\int g^2 dF_0 < \infty$ by van der Vaart (1998, p. 364, Example 25.16). Then the effecient influence function of κ at F_0 is $\tilde{\kappa}(x) = \phi(x) - \int \phi(x)\,dF_0(x)$ because $\tilde{\kappa}(x)$ is the projection of $\phi(x)$ into the tangent space of F_0. Applying Fubini's theorem, we obtain $\int \phi(x)\,dF_0(x) = 2\int (F_0(s) - G(s))\,F_0(s)\,dG(s)$, which concludes the proof.

Proposition 3. Let c_α denote the lower α-quantile of the standard normal distribution. Let $c_{\alpha,n}$ be a sequence of critical values depending on $X_1, ..., X_n$ such that $c_{\alpha,n} \to c$ in probability for $n \to \infty$. Let $\sigma_n^2 = \sigma_n^2(X_1, ..., X_n)$ be a consistent sequence of estimators for the asymptotic variance of the test statistic $T(F_n)$. Then the test, which rejects H_0 if $T(F_n)/\sigma_n \le c_{\alpha,n}$, is locally asymptotically most powerful (LAMP).

Proof. By van der Vaart (1998, p. 384, Theorem 25.44 and Lemma 25.45), it is sufficient to show that the asymptotic distributions of $T(F_n)$ and $S_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\tilde{\kappa}(X_i)$ coincide under F_0. The test statistic $T(F_n)$ converges weakly to $2\int (F_0 - G)B(F_0)\,dG$ for $n \to \infty$ by Shorack and Wellner (2009, p. 178, Theorem 4.4.2), where B denotes the standard Brownian bridge. The statistic S_n can be rewritten as $S_n = 2\int (F_0 - G)[\sqrt{n}(F_n - F_0)]dG$. By Donsker's theorem, $\sqrt{n}(F_n - F_0)$ converges weakly to $B(F_0)$ for $n \to \infty$. By the continuous mapping theorem, we conclude $S_n \to 2\int(F_0 - G)B(F_0)\, dG$ for $n \to \infty$.

3. Asymptotic and Percentile-*t* tests

In this section, we provide a detailed algorithm for the asymptotic and percentile-t equivalence tests. First, a consistent estimator for the asymptotic variance of the test statistic $T(F_n)$ will be derived. The asymptotic variance of $T(F_n)$ is $\sigma^2(F) = 4\iint(F(t) - G(t))\,(F(s) - G(s))\,(F(\min(s, t)) - F(s)\,F(t))\, dG(s)\, dG(t)$ by Shorack and Wellner (2009, p.42, Proposition 2.2.1). An estimator of $\sigma^2(F)$ is given in the next proposition in terms of $U_{(1)}, \ldots, U_{(n)}$.

Proposition 4. Set $U_{(0)} = 0$ and $U_{(n+1)} = 1$. Then $\sigma^2(F_n) = 4\sum_{k=0}^{n}\sum_{l=0}^{n} s_{kl}$ is a consistent estimator of $\sigma^2(F)$, where $s_{kl} = \left(\frac{\min(k,l)}{n} - \frac{k}{n}\frac{l}{n}\right) f_k f_l$ and $f_k = \frac{k}{n}\left(U_{(k+1)} - U_{(k)}\right) - \frac{1}{2}\left(U_{(k+1)}^2 - U_{(k)}^2\right)$.

Proof. By the Glivenko Cantelli theorem, $F_n \to F$ almost surely for $n \to \infty$. Consequently, $\sigma^2(F_n) \to \sigma^2(F)$ almost surely for $n \to \infty$ by the continuous mapping theorem. Using the definition of the empirical CDF we obtain:

$$\sigma^2(F_n) = 4\sum_{k=0}^{n}\sum_{l=0}^{n}\int_{X_{(k)}}^{X_{(k+1)}}\int_{X_{(l)}}^{X_{(l+1)}}\left(\frac{k}{n} - G(s)\right)\left(\frac{l}{n} - G(t)\right)\left(\frac{\min(k,l)}{n} - \frac{k}{n}\frac{l}{n}\right) dG(s)\, dG(t).$$

Let s_{kl} denote a single summand. Fubini's theorem implies

$$s_{kl} = \left(\frac{\min(k,l)}{n} - \frac{k}{n}\frac{l}{n}\right)\left(\int_{X_{(x)}}^{X_{(k+1)}}\left(\frac{k}{n} - G(s)\right) dG(s)\right)\left(\int_{X_{(1)}}^{X_{(l+1)}}\left(\frac{l}{n} - G(t)\right) dG(t)\right)$$

Set $f_k = \int_{X_{(k)}}^{X_{(k+1)}}\left(\frac{k}{n} - G(s)\right) dG(s)$. By applying the substitution $u = G(s)$, we get

$$f_k = \int_{U_{(k)}}^{U_{(k+1)}}\left(\frac{k}{n} - u\right) du = \frac{k}{n}\left(U_{(k+1)} - U_{(k)}\right) - \frac{1}{2}\left(U_{(k+1)}^2 - U_{(k)}^2\right).$$

If $\Delta(F, G) = \varepsilon$ then the asymptotic distribution of the test statistic $T(F_n)$ is normal with mean zero and variance $\sigma^2(F)$, see Shorack and Wellner (2009, p.42, Proposition 2.2.1 and p. 178, Theorem 4.4.2) for details. Therefore, the asymptotic test rejects H_0 if $T(F_n) \le c_\alpha \sigma(F_n)$, where c_α is the lower α-quantile of the standard normal distribution. The asymptotic test is LAMP by Proposition 3. The minimum value of the tolerance parameter ε, for which the asymptotic test can reject H_0, equals

$$\varepsilon_{\min}(F_n) = \Delta(F, G) - n^{-\frac{1}{2}} c_\alpha \sigma(F_n) \tag{3.1}$$

The asymptotic test can be carried out as follows:

1. Given are the observed data $X_1, \ldots, X_n$, the tolerance parameter ε and the significance level ε.
2. Transform the variables $U_i = G(X_i)$ for all $i = 1, \ldots, n$.
3. Compute the order statistics $U_{(1)}, \ldots, U(n)$. Set $U_{(0)} = 0$ and $U_{(n+1)} = 1$.
4. Compute the estimator $\sigma^2(F_n)$ for the asymptotic variance of the test statistic $T(F_n)$, see Proposition 4 for the closed-form formula.
5. Compute $\varepsilon_{\min}(F_n)$, see (3.1) for the formula.
6. Reject H_0 if $\varepsilon_{\min}(F_n) \le \varepsilon$.

In order to improve the finite sample performance of the considered equivalence test, we use the well known percentile-*t* method, see van der Vaart (1998, Chapter 23). The bootstrap samples $\hat{X}_1, \ldots, \hat{X}_n$ should be generated from original data by using sampling with replacement. Let $\hat{F}_n$ denote the empirical CDF, which is based on the bootstrap sample $\hat{X}_1, \ldots, \hat{X}_n$. Let the data sample $X_1, \ldots, X_n$ be fixed and the bootstrap samples $\hat{X}_1, \ldots, \hat{X}_n$ be random. The percentile-*t* method uses the bootstrap distribution of the normalized statistic $(T(\hat{F}_n) - T(F_n)) / \sigma(\hat{F}_n)$ to estimate the distribution of the test statistic $T(F_n)$. Let $c_\alpha(\hat{F}_n)$ denote the empirical lower α-quantile of the normalized test statistic $(T(\hat{F}_n) - T(F_n)) / \sigma(\hat{F}_n)$. The quantile $c_\alpha(\hat{F}_n)$ can be computed by means of the simulation to any degree of accuracy. The percentile-*t* test rejects H_0 if $T(F_n) \le \sigma(F_n)\, c_\alpha(\hat{F}_n)$. The minimum value of the tolerance parameter ε, for which the percentile-*t* test can reject H_0, is

$$\varepsilon_{\min}(\hat{F}_n) = \Delta(F, G) - n^{-\frac{1}{2}} c_\alpha(\hat{F}_n) \sigma(F_n). \tag{3.2}$$

The percentile-*t* test is consistent by van der Vaart (1998, p. 330, Theorem 23.4 and p. 331, Theorem 23.5). Consequently, the percentile-*t* test is also LAMP by Proposition 3. The percentile-*t* test can be performed similarly to the asymptotic test, with the following additional steps to calculate $c_\alpha(\hat{F}_n)$:

1. Generate m bootstrap samples $\hat{X}_1, \ldots, \hat{X}_n$ from the original data using sampling with replacement, where *m* is a given number of bootstrap samples.
2. Compute the normalized test statistic $(T(\hat{F}_n) - T(F_n)) / \sigma(\hat{F}_n)$ for each bootstrap sample.
3. Compute the empirical lower α-quantile $c_\alpha(\hat{F}_n)$ of the normalized test statistic.

The percentile-*t* test is computationally intensive due to the calculation of the variance estimator $\sigma(\hat{F}_n)$.

4. Simulation Study

In this section, we study the finite sample performance of the equivalence tests by means of simulation. It is sufficient to consider the case where G is the CDF of the uniform distribution on [0, 1], see Remark 1. The equivalence tests are implemented in *R* and the source code is freely available under https://github.com/TestingEquivalence/EquivalenceCM. All simulations are performed in *R*-Studio on a scientific workstation.

4.1 Appropriate values of the tolerance parameter ε

In order to shed some light on the appropriate values of the tolerance parameter ε, the power of the test is computed at the uniform distribution U on [0, 1] for the different sample sizes n, see Table 1. The values of the power of the test 0.9, 0.8 and 0.7 are column names. The asymptotic test is abbreviated as AT and the percentile-*t* test is abbreviated as PT. All tests are carried out at the nominal level 0.05. The number of simulations is 1000 for each experiment and the number of bootstrap samples is 1000. The power of the asymptotic test is higher than the power of the percentile-*t* test for sample sizes $n \in \{50, 100, 200, 500\}$. The power of both tests is almost equal for $n = 1000$. For a given sample size n, the value of the tolerance parameter ε can be considered appropriate if the test power at U is sufficiently high. We would set the value of ε for a given n so that the test power at U is at least 0.9.

Table 1: Tolerance parameter ε as a function of the power of the test

	0.9		*0.8*		*0.7*	
n	*AT*	*PT*	*AT*	*PT*	*AT*	*PT*
50	0.0168	0.0214	0.0128	0.0131	0.0104	0.0087
100	0.0089	0.0119	0.0066	0.0070	0.0054	0.0047
200	0.0043	0.0059	0.0034	0.0037	0.0027	0.0024
500	0.0019	0.0025	0.0014	0.0015	0.0011	0.0010
1000	0.0009	0.0012	0.0007	0.0007	0.005	0.0005

4.2 Type I error rates

The type I error rates of the equivalence tests are studied in this section. The boundary points of H_0 are based on alternatives that are often considered in the literature on goodness-of-fit tests, see Blinov and Lemeshko (2014), Lemeshko, Blinov, and Lemeshko (2016), Marhuenda, Morales, and Pardo (2005), Rayner and Rayner (2001), Zhang (2002) and Vexler and Gurevich (2010) among others. For this purpose, we use the beta distributions Beta (p, q) with different parameters p, q and Stephens alternatives. The Stephens alternatives A_k, B_k and C_k have the CDFs:

$A_k(x) = 1 - (1 - x)^k$ for $x \in [0, 1]$,

$B_k(x) = 2^{k-1}x^k$ for $x \in [0, 0.5]$ and $B_k(x) = 1 - 2^{k-1} (1 - x)^k$ for $(0.5, 1]$;

$C_k(x) = \frac{1}{2} - 2^{k-1}\left(\frac{1}{2} - x\right)^k$ for $x \in [0, 0.5]$ and $C_k(x) = \frac{1}{2} + 2^{k-1}\left(x - \frac{1}{2}\right)^k$ for (0.5, 1], see Stephens (1974) for details. The parameter $k > 0$ controls the shape of the CDF.

The construction of the boundary points of H_0 is similar to Baringhaus and Henze (2017). Let F be a CDF so that $\Delta(F, U) > \varepsilon$. Then the CDF of the corresponding boundary point is $wF + (1 - w)U$, where the parameter w is optimized so that $\Delta(wF + (1 - w)U, U) = \varepsilon$. The parameter w can be found using any line search method. Table 2 summarizes the computed test power at the boundary points for the sample size $n = 200$ and the tolerance parameter $\varepsilon = 0.006$. The boundary points are based on the distributions in the column Alternative. The Cramér-von Mises distance between an alternative and U is in the column $\Delta(F, U)$. The asymptotic test is abbreviated as AT and the percentile-t test is abbreviated as PT. All tests are carried out at the nominal level 0.05. The number of simulations is 1000 for each experiment and the number of bootstrap samples is 1000.

The value of the tolerance parameter ε is chosen so that the power of the both tests is larger than 0.9 for the sample size $n = 200$. The power of the asymptotic test varies considerably from point to point. The asymptotic test is not conservative at some boundary points. The power of the percentile-t test is much closer to the nominal level 0.05 compared to the asymptotic test. Similar results were observed for sample sizes $n \in \{50, 100, 500\}$. Overall, the percentile-t test performs significantly better than the asymptotic test.

Table 2. The power of the tests at boundary points of H_0.

Alternative	*Δ(F, U)*	*AT*	*PT*
Beta(0.5, 1.0)	0.0333	0.083	0.045
Beta(0.5,1.5)	0.0732	0.091	0.046
Beta(0.5, 2.0)	0.1065	0.067	0.030
Beta(1, 1.5)	0.0119	0.084	0.033
Beta(1, 2)	0.0333	0.084	0.028
Beta(1.5, 2)	0.0089	0.069	0.061
A0.25	0.1111	0.092	0.060
A0.5	0.0333	0.089	0.053
A1.5	0.0119	0.065	0.035
A2	0.0333	0.082	0.039
A2.5	0.0556	0.086	0.046
A3	0.0762	0.100	0.048
B0.25	0.0278	0.017	0.042
B0.5	0.0083	0.034	0.052
B2	0.0083	0.038	0.062
B2.5	0.0139	0.034	0.056
B3	0.0190	0.037	0.061
C0.25	0.0278	0.032	0.056
C0.5	0.0083	0.043	0.065
C2	0.0083	0.028	0.051
C2.5	0.0139	0.023	0.049
C3	0.0190	0.024	0.050

Conclusion

The equivalence tests under consideration can be successfully applied to assess whether observed data are close to a fully specified continuous distribution on $\mathbb{R}$. The percentile-*t* test has shown good finite sample performance and should be preferred. The asymptotic test is not conservative and should only be used when computational resources are scarce. For a given sample size *n*, the efficient value of the tolerance parameter ε can be found by simulation of the test power at the uniform distribution on [0, 1].

Acknowledgments

The author would like to thank the anonymous referees for the useful comments and suggestions leading to a significant improvement of this paper.

References

Baringhaus, L., Gaigall, D., and Thiele, J. P. (2018). Statistical inference for L2-distances to uniformity. *Computational Statistics*, 33 (4), 18631896. doi. 10.1007/s00180-018-0820-0

Baringhaus, L., and Henze, N. (2017). Cramér-von Mises distance. Probabilistic interpretation, confidence intervals, and neighbourhood-of-model validation. *Journal of Nonparametric Statistics*, 29 (2), 167-188. doi. 10.1080/10485252.2017.1285029

Berger, J. O., and Delampady, M. (1987). Testing precise hypotheses. *Statistical Science*, 2 (3), 317 335. doi. 10.1214/ss/1177013238

Bickel, P., Klaassen, C., Ritov, Y., and Wellner, J. (1993). *Efficient and adaptive estimation for semiparametric models.* Johns Hopkins University Press, Baltimore.

Blinov, P., and Lemeshko, B. (2014). A review of the properties of tests for uniformity. In *12th international conference on actual problems of electronic instrument engineering.* doi. 10.1109/APEIE.2014.7040743

Hodges, J., and Lehmann, E. (1954). Testing the approximate validity of statistical hypotheses. *J. Roy. Statist. Soc. Ser. B*, 16 (2), 261268.

Lemeshko, B.Y., Blinov, P. Y., and Lemeshko, S. B. (2016). Goodness-of-fit tests for uniformity of probability distribution law. *Optoelectronics, Instrumentation and Data Processing*, 52 (2), 128140. doi. 10.3103/S8756699016020047

Marhuenda, Y., Morales, D., and Pardo, M. (2005). A comparison of uniformity tests. *Statistics*, 39 , 315-327. doi. 10.1080/02331880500178562

Rao, C. R., and Lovric, M. M. (2016). Testing point null hypothesis of a normal mean and the truth. 21st century perspective. *Journal of Modern Applied Statistical Methods*, 15 . doi. 10.22237/jmasm/1478001660

Rayner, G., and Rayner, J. (2001). Power of the Neyman smooth tests for the uniform distribution. *Journal of Applied Mathematics and Decision Sciences*, 5 . doi. 10.1155/S117391260100013X

Shorack, G., and Wellner, J. (2009). *Empirical processes with applications to statistics.* Society for Industrial and Applied Mathematics.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69 (347), 730737. doi.10.2307/2286009

van der Vaart, A. (1998). Asymptotic statistics. Cambridge University.

Vexler, A., and Gurevich, G. (2010). Empirical likelihood ratios applied to goodness-of-fit tests based on sample entropy. *Computational Statistics and Data Analysis*, 54 (2), 531-545.

Wellek, S. (2010). Testing statistical hypotheses of equivalence and noninferiority. CRC Press.

Zhang, J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society*. Series B (Statistical Methodology), 64 (2), 281-294.