DVORETZKY–KIEFER–WOLFOWITZ INEQUALITIES FOR THE TWO-SAMPLE CASE

DVORETZKY–KIEFER–WOLFOWITZ INEQUALITIES FOR THE TWO-SAMPLE CASE

FAN WEI AND R. M. DUDLEY

ABSTRACT. The Dvoretzky–Kiefer–Wolfowitz (DKW) inequality says that if F_n is an empirical distribution function for variables i.i.d. with a distribution function F, and K_n is the Kolmogorov statistic $\sqrt{n}\sup_x |(F_n - F)(x)|$, then there is a finite constant C such that for any $M > 0$, $\Pr(K_n > M) \leq C\exp(-2M^2)$. Massart proved that one can take $C = 2$ (DKWM inequality) which is sharp for F continuous. We consider the analogous Kolmogorov–Smirnov statistic $KS_{m,n}$ for the two-sample case and show that for $m = n$, the DKW inequality holds with $C = 2$ if and only if $n \geq 458$. For $n_0 \leq n < 458$ it holds for some $C > 2$ depending on n_0.

For $m \neq n$, the DKWM inequality fails for the three pairs (m, n) with $1 \leq m < n \leq 3$. We found by computer search that for $n \geq 4$, the DKWM inequality always holds for $1 \leq m < n \leq 200$, and further that it holds for $n = 2m$ with $101 \leq m \leq 300$. We conjecture that the DKWM inequality holds for pairs $m \leq n$ with the $457 + 3 = 460$ exceptions mentioned.

1. Introduction

This paper is a long version, giving many more details, of our shorter paper [16]. Let F_n be the empirical distribution function based on an i.i.d. sample from a distribution function F, let

$$D_n := \sup_x |(F_n - F)(x)|,$$

and let K_n be the Kolmogorov statistic $\sqrt{n}D_n$. Dvoretzky, Kiefer, and Wolfowitz in 1956 [7] proved that there is a finite constant C such that for all n and all $M > 0$,

$$\Pr(K_n \geq M) \leq C\exp(-2M^2). \tag{1}$$

We call this the DKW inequality. Massart in 1990 [12] proved (1) with the sharp constant $C = 2$, which we will call the DKWM inequality. In this paper we consider possible extensions of these inequalities to the two-sample case, as follows. For $1 \leq m \leq n$, the null hypothesis H_0 is that F_m and G_n are independent empirical distribution functions from a continuous distribution function F, based altogether on $m + n$ samples i.i.d. (F). Consider the Kolmogorov–Smirnov statistics

$$D_{m,n} = \sup_x \mid (F_m - G_n)(x) \mid, \quad KS_{m,n} = \sqrt{\frac{mn}{m+n}}D_{m,n}. \tag{2}$$

All probabilities to be considered are under H_0.

For given m and n let $L = L_{m,n}$ be their least common multiple. Then the possible values of $D_{m,n}$ are included in the set of all k/L for $k = 1, \ldots, L$. If $n = m$

Date: January 6, 2012.
1991 *Mathematics Subject Classification.* 2008 MSC: 62G10, 62G30.
Key words and phrases. Kolmogorov–Smirnov test, empirical distribution functions.

then all these values are possible. The possible values of $KS_{m,n}$ are thus of the form

$$M = \sqrt{(mn)/(m+n)}k/L_{m,n}. \tag{3}$$

We will say that the DKW (resp. DKWM) inequality holds in the two-sample case for given m, n, and C (resp. $C = 2$) if for all $M > 0$, the following holds:

$$P_{m,n,M} := \Pr(KS_{m,n} \geq M) \leq C \exp(-2M^2). \tag{4}$$

It is well known that as $m \to +\infty$ and $n \to +\infty$, for any $M > 0$,

$$P_{m,n,M} \;\to\; \beta(M) := \Pr(\sup_{0\leq t\leq 1} |B_t| > M) = 2\sum_{j=1}^{\infty}(-1)^{j-1}\exp(-2j^2M^2), \tag{5}$$

where B_t is the Brownian bridge process.

Remark. For M large enough so that H_0 can be rejected according to the asymptotic distribution given in (5) at level $\alpha \leq 0.05$, the series in (5) is very close in value to its first term $2\exp(-2M^2)$, which is the DKWM bound (when it holds). Take M_α such that $2\exp(-2M_\alpha^2) = \alpha$, then for example we will have

$$\beta(M_{.05}) \doteq 0.04999922, \quad \beta(M_{.01}) \doteq 0.009999999. \tag{6}$$

Let $r_{\max} = r_{\max}(m, n)$ be the largest ratio $P_{m,n,M}/(2\exp(-2M^2))$ over all possible values of M for the given m and n. We summarize our main findings in Theorem 1 and Facts 2, 3, and 4.

1. **Theorem.** *For $m = n$ in the two-sample case:*
   - (*a*) *The DKW inequality always holds with $C = e \doteq 2.71828$.*
   - (*b*) *For $m = n \geq 4$, the smallest n such that H_0 can be rejected at level 0.05, the DKW inequality holds with $C = 2.16863$.*
   - (*c*) *The DKWM inequality holds for all $m = n \geq 458$, i.e., for all $M > 0$,*

$$P_{n,n,M} = Pr(KS_{n,n} \geq M) \leq 2e^{-2M^2}. \tag{7}$$

   - (*d*) *For each $m = n < 458$, the DKWM inequality fails for some M given by (3).*
   - (*e*) *For each $m = n < 458$, the DKW inequality holds for $C = 2(1 + \delta_n)$ for some $\delta_n > 0$, where for $12 \leq n \leq 457$,*

$$\delta_n < -\frac{0.07}{n} + \frac{40}{n^2} - \frac{400}{n^3}. \tag{8}$$

Remark. The bound on the right side of (8) is larger than $2\delta_n$ for $n = 16,\ 40, 70,$ 440, and 445 for example, but is less than $1.5\delta_n$ for $125 \leq n \leq 415$. It is less than $1.1\delta_n$ for $n = 285,\ 325,\ 345$.

Theorem 1 (a), (b), and (c) are proved in Section 2. Parts (d) and (e), and also parts (a) through (c) for $n < 6395$, were found by computation.

For $m \neq n$ we have no general or theoretical proofs but report on computed values. The methods of computation are summarized in Subsection 3.2. Detailed results in support of the following three facts are given in Subsection 3.3 and Appendix B.

2. **Fact.** *Let $1 \leq m < n \leq 200$. Then:*
   - (*a*) *For $n \geq 4$, the DKWM inequality holds.*

(*b*) *For each* (m, n) *with* $1 \le m < n \le 3$*, the DKWM inequality fails, in the case of* $\Pr(D_{m,n} \ge 1)$*.*
(*c*) *For* $3 \le m \le 100$*, the* n *with* $m < n \le 200$ *having largest* $r_{\max}$ *is always* $n = 2m$*.*
(*d*) *For* $102 \le m \le 132$ *and* m *even, the largest* $r_{\max}$ *is always found for* $n = 3m/2$ *and is increasing in* m*.*
(*e*) *For* $169 \le m \le 199$ *and* $m < n \le 200$*, the largest* $r_{\max}$ *occurs for* $n = m+1$*.*
(*f*) *For* $m = 1$ *and* $4 \le n \le 200$*, the largest* $r_{\max} = 0.990606$ *occurs for* $n = 4$ *and* $d = 1$*. For* $m = 2$ *and* $4 \le n \le 200$*, the largest* $r_{\max} = 0.959461$ *occurs for* $n = 4$ *and* $d = 1$*.*

In light of Fact 2(c) we further found:

3. **Fact.** *For* $n = 2m$*:*

(*a*) *For* $3 \le m \le 300$*, the DKWM inequality holds;* $r_{\max}(m, 2m)$ *has relative minima at* $m = 6,\ 10,$ *and* 16 *but is increasing for* $m \ge 16$*, up to* 0.9830 *at* $m = 300$*.*
(*b*) *The* p*-values forming the numerators of* $r_{\max}$ *for* $100 \le m \le 300$ *are largest for* $m = 103$ *where* $p \doteq 0.3019$ *and smallest at* $m = 294$ *where* $p \doteq 0.2189$*.*
(*c*) *For* $101 \le m \le 199$*, the smallest* $r_{\max}$ *for* $n = 2m$*, namely* $r_{\max}(101, 202) \doteq 0.97334$*, is larger than every* $r_{\max}(m', n')$ *for* $101 \le m' < n' \le 200$*, all of which are less than* 0.95*, the largest being* $r_{\max}(132, 198) \doteq 0.9496$*.*
(*d*) *For* $3 \le m \le 300$*,* $r_{\max}$ *is attained at* $d_{\max} = k_{\max}/n$ *which is decreasing in* n *when* $k_{\max}$ *is constant but jumps upward when* $k_{\max}$ *does;* $k_{\max}$ *is nondecreasing in* m*.*

The next fact shows that for a wide range of pairs (m, n), but not including any with $n = m$ or $n = 2m$, the correct p-value $P_{m,n,M}$ is substantially less than its upper bound $2\exp(-2M^2)$ and in cases of possible significance at the 0.05 level or less, likewise less than the asymptotic p-value $\beta(M)$:

4. **Fact.** *Let* $100 < m < n \le 200$*. Then:*

(*a*) *The ratio* $2\exp(-2M^2)/P_{m,n,M}$ *is always at least* 1.05 *for all possible values of* M *in* (3)*. The same is true if the numerator is replaced by the asymptotic probability* $\beta(M)$ *and* $\beta(M) \le 0.05$*.*
(*b*) *If in addition* $m = 101,\ 103,\ 107,\ 109,$ *or* 113*, then part* (*a*) *holds with* 1.05 *replaced by* 1.09*.*

Remark. For small p-values, for example of order 10^{-15}, with $d \le 1/2$, few or no significant digits can be computed by the method we used. But, one can compute accurately an upper bound for such p-values, which we used in ranges $d_0(m, n) \le D_{m,n} \le 1/2$ where p-values are less than 10^{-8}, to verify Facts 2, 3, and 4 for those ranges. We give details in Section 3 and Appendix B, Tables 6, 7, and 8.

We have in the numerator of $r_{\max}$ the p-values of 0.2189 (corresponding to $m = 294$) or more in Fact 3(b) (Table 8), and similarly p-values of 0.26 or more in Table 6 and 0.27 or more in Table 7, which together support Fact 2. These substantial p-values suggest, although they of course do not prove, that more generally, large $r_{\max}$ do not tend to occur at small p-values.

2. Proof of Theorem 1

B. V. Gnedenko and V. S. Korolyuk in 1952 [9] gave an explicit formula for $P_{n,n,M}$, and M. Dwass (1967) [8] gave another proof. The technique is older: the reflection principle dates back to André [1]. Bachelier in 1901 [2, pp. 189-190] is the earliest reference we could find for the method of repeated reflections, applied to symmetric random walk. He emphasized that the formula there is rigorous ("rigoureusement exacte"). Expositions in several later books we have seen, e.g. in 1939 [4, p. 32], are not so rigorous, assuming a normal approximation and thus treating repeated reflections of Brownian motion. According to J. Blackman [5, p. 515] the null distribution of $\sup|F_n - G_n|$ had in effect "been treated extensively by Bachelier" in 1912, [3] "in connection with certain gamblers'-ruin problems."

The formula is given in the following proposition.

5. **Proposition** (Gnedenko and Korolyuk)**.** *If $M = k/\sqrt{2n}$, where $1 \le k \le n$ is an integer, then*

$$Pr(KS_{n,n} \ge M) = \frac{2}{\binom{2n}{n}} \left(\sum_{i=1}^{\lfloor n/k \rfloor} (-1)^{i-1} \binom{2n}{n+ik} \right).$$

Since the probability $P_{n,n,M} = \Pr(KS_{n,n} \ge M)$ is clearly not greater than 1, we just need to consider the M such that

$$2e^{-2M^2} \le 1,$$

i.e., we just need to consider the integer pairs (n,k) where

$$k \ge \sqrt{n \ln 2}. \tag{9}$$

The exact formula for $P_{n,n,M}$ is complicated. Thus we want to determine upper bounds for $P_{n,n,M}$ which are of simpler forms. We prove the main theorem by two steps: we first find two such upper bounds for $P_{n,n,M}$ as in Lemma 6 and 14 and then show (7) holds when $P_{n,n,M}$ is replaced by the two upper bounds for two ranges of pairs (k,n) respectively, as will be stated in Propositions 13 and 16.

6. **Lemma.** *An upper bound for $P_{n,n,M}$ can be given by $2\binom{2n}{n+k}/\binom{2n}{n}$.*

Proof. This is clear from Proposition 5, since the summands alternate in signs and decrease in magnitude. Therefore we must have

$$\sum_{i=2}^{\lfloor n/k \rfloor} (-1)^{i-1} \binom{2n}{n+ik} \le 0.$$

□

As a consequence of Lemma 6, to prove (7) for a pair (n,k), it will suffice to show that

$$2\binom{2n}{n+k} \Big/ \binom{2n}{n} < 2\exp\left(-k^2/n\right). \tag{10}$$

We first define some auxiliary functions.

7. **Notation.** For all $n, k \in \mathbb{R}$ such that $1 \le k \le n$, define

$$PH(n,k) := \ln \binom{2n}{n+k} - \ln \binom{2n}{n} + \frac{k^2}{n},$$

where for $n_1 \geq n_2$,

$$\binom{n_1}{n_2} = \frac{\Gamma(n_1+1)}{\Gamma(n_1-n_2+1)\Gamma(n_2+1)},$$

and $\Gamma(x)$ is the Gamma function, defined for $x > 0$ by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

It satisfies the well-known recurrence $\Gamma(x+1) \equiv x\Gamma(x)$.

It is clear that $PH(n,k) \leq 0$ if and only if (10) holds.

8. **Notation.** For all $n, k \in \mathbb{R}$ such that $1 \leq k \leq n$, define

$$\begin{aligned} DPH(n,k) &:= PH(n,k) - PH(n,k-1) \\ &= \ln\left(\frac{n-k+1}{n+k}\right) + \frac{2k-1}{n}. \end{aligned} \tag{11}$$

9. **Lemma.** *When $n \geq 19$, $DPH(n,k)$ is decreasing in k when $k \geq \sqrt{n \ln 2}$.*

Proof. Clearly $DPH(n,k)$ is differentiable with respect to k on the domain $n, k \in \mathbb{R}$ such that $n > 0$ and $0 < k < n + 1/2$, with partial derivative given by

$$\frac{\partial}{\partial k} DPH(n,k) = \frac{-2k^2+2k+n}{n\left(-k^2+k+n^2+n\right)}. \tag{12}$$

It is easy to check that the denominator is positive on the given domain. Thus (12) is greater than 0 if and only if $-2k^2 + 2k + n > 0$, which is equivalent to

$$\frac{1}{2}\left(1-\sqrt{2n+1}\right) < k < \frac{1}{2}\left(1+\sqrt{2n+1}\right).$$

Since we have that when $n \geq 19$,

$$\sqrt{n \ln 2} > \frac{1}{2}\left(1+\sqrt{2n+1}\right),$$

$DPH(n,k)$ is decreasing in k whenever $n \geq 19$. □

10. **Lemma.** *(a) For $0 < \alpha < 2/\sqrt{\ln 2}$ and all $n \geq 1$,*

$$n - \alpha\sqrt{n}\sqrt{\ln 2} + 1 > 0. \tag{13}$$

(b) For $\sqrt{3/(2\ln 2)} < \alpha < 2/\sqrt{\ln 2}$ and n large enough,

$$\frac{d}{dn} DPH(n, \alpha\sqrt{n \ln 2}) > 0.$$

(c) For $n \geq 3$, $DPH(n, \sqrt{3n})$ is increasing in n.
(d) $DPH(n, \sqrt{3n}) \to 0$ as $n \to \infty$.
(e) For all $n \geq 3$, $DPH(n, \sqrt{3n}) < 0$.

Proof. Part (a) holds because the left side of (13), as a quadratic in $\sqrt{n}$, has the leading term $n = \sqrt{n}^2 > 0$ and discriminant $\Delta = \alpha^2 \ln 2 - 4 < 0$ under the assumption.

For part (b), by plugging $k = \alpha\sqrt{n \ln 2}$ into $DPH(n,k)$, we have

$$DPH(n, \alpha\sqrt{n \ln 2}) = \frac{2\alpha\sqrt{n \ln 2} - 1}{n} + \ln\left(\frac{-\alpha\sqrt{n \ln 2} + n + 1}{\alpha\sqrt{n \ln 2} + n}\right), \tag{14}$$

which is well-defined by part (a). It is differentiable with respect to n with derivative given by

$$\frac{d}{dn}DPH(n,\alpha\sqrt{n\ln 2}) \tag{15}$$
$$=\frac{n\left(2\alpha^3\ln^{\frac{3}{2}}(2)-3\alpha\sqrt{\ln 2}\right)+\sqrt{n}\left(2-4\alpha^2\ln 2\right)+2\alpha\sqrt{\ln 2}}{2n^2\left(\alpha\sqrt{\ln 2}+\sqrt{n}\right)\left(-\alpha\sqrt{n}\sqrt{\ln 2}+n+1\right)}.$$

By part (a), the denominator

$$2n^2\left(\alpha\sqrt{\ln 2}+\sqrt{n}\right)\left(-\alpha\sqrt{n}\sqrt{\ln 2}+n+1\right)$$

is positive. The numerator will be positive for n large enough, since the coefficient of its leading term,

$$2\alpha^3\ln^{3/2}(2)-3\alpha\sqrt{\ln 2},$$

is positive by the assumption $\alpha>\sqrt{3/(2\ln 2)}$ in this part. So part (b) is proved.

For part (c), when $\alpha=\sqrt{3}/\sqrt{\ln 2}$, we have

$$\frac{d}{dn}DPH(n,\sqrt{3n})=\frac{3\sqrt{3}n-10\sqrt{n}+2\sqrt{3}}{2\left(\sqrt{n}+\sqrt{3}\right)\left(n-\sqrt{3}\sqrt{n}+1\right)n^2}.$$

This is clearly positive when $3\sqrt{3}n-10\sqrt{n}+2\sqrt{3}\geq 0$, which always holds when $n\geq 3$. This proves part (c).

For part (d), plugging $\alpha=\sqrt{3/\ln 2}$ into (14), we have

$$\begin{aligned}&\lim_{n\to\infty}DPH(n,\sqrt{3n})\\ =\ &\lim_{n\to\infty}\left(\frac{2\sqrt{3n}-1}{n}+\ln\left(\frac{n-\sqrt{3n}+1}{n+\sqrt{3n}}\right)\right)\\ =\ &0,\end{aligned}$$

proving part (d). Part (e) then follows from parts (c) and (d). □

11. **Lemma.** *For $n\geq 1$,*

$$DPH(n,\sqrt{n\ln 2})>0.$$

Proof. By (15) for $\alpha<2/\sqrt{\ln 2}$, in this case $\alpha=1$, we have that

$$\frac{d}{dn}DPH(n,\sqrt{n\ln 2})=\frac{n\left(2\ln^{3/2}(2)-3\sqrt{\ln 2}\right)+\sqrt{n}(2-4\ln 2)+2\sqrt{\ln 2}}{2n^2\left(\sqrt{n}+\sqrt{\ln 2}\right)\left(n-\sqrt{n}\sqrt{\ln 2}+1\right)}.$$

The denominator is always positive for $n\geq 1$ by (13). The numerator as a quadratic in $\sqrt{n}$ has leading coefficient $2\ln^{3/2}(2)-3\sqrt{\ln 2}<0$. This quadratic also has a negative discriminant, so the numerator is always negative when $n\geq 1$.

Similarly, we have

$$\begin{aligned}&\lim_{n\to\infty}DPH(n,\sqrt{n\ln 2})\\ =\ &\lim_{n\to\infty}\left(\frac{2\sqrt{n\ln 2}-1}{n}+\ln\left(\frac{n-\sqrt{n\ln 2}+1}{n+\sqrt{n\ln 2}}\right)\right)\\ =\ &0.\end{aligned}$$

Therefore $DPH(n, \sqrt{n \ln 2}) > 0$ for all $n \geq 1$. $\square$

Summarizing Lemmas 9, 10, and 11, we have the following corollary:

12. **Corollary.** *For any fixed $n \geq 19$, $DPH(n,k)$ is decreasing in k when $k \geq \sqrt{n \ln 2}$. Furthermore,*

$$DPH\left(n, \sqrt{n \ln 2}\right) > 0, \ DPH\left(n, \sqrt{3n}\right) < 0.$$

13. **Proposition.** *The inequality (7) holds for all integers n, k such that $n \geq 108$ and $\sqrt{3n} \leq k \leq n$.*

Proof. By Lemma 6, the probability $P_{n,n,M}$ is bounded above by $2\binom{2n}{n+k}/\binom{2n}{n}$. We here prove this proposition by showing that (10) holds for all integers n, k such that $\sqrt{3n} \leq k \leq n$ and $n \geq 108$.

To prove (10) is equivalent to proving

$$\ln \binom{2n}{n+k} - \ln \binom{2n}{n} + \frac{k^2}{n} < 0 \tag{16}$$

for $k = t\sqrt{n}$ where $t \geq \sqrt{3}$, by Notation 7.

Rewriting (16), we need to show that for $k \geq \sqrt{3n}$,

$$\ln \left(\frac{n!n!}{(n+k)!(n-k)!} \right) + \frac{k^2}{n} < 0. \tag{17}$$

We will use Stirling's formula with error bounds. Recall that one form of such bounds [13] states that

$$\sqrt{2\pi} \exp\left(\frac{1}{12s} - \frac{1}{360s^3} - s\right) s^{s+1/2} \leq s! \leq \sqrt{2\pi} \exp\left(\frac{1}{12s} - s\right) s^{s+1/2}$$

for any positive integer s. We plug the bounds for $s!$ into $\dfrac{n!n!}{(n+k)!(n-k)!}$, getting

$$\frac{n!n!}{(n+k)!(n-k)!} \leq \frac{n^{2n+1}(n+k)^{-n-k-\frac{1}{2}}(n-k)^{k-n-\frac{1}{2}} \exp\left(\frac{1}{6n}\right)}{\exp\left(\frac{1}{12}\left[\frac{1}{n+k} + \frac{1}{n-k}\right] - \frac{1}{360}\left[\frac{1}{(n+k)^3} + \frac{1}{(n-k)^3}\right]\right)}.$$

By taking logarithms of both sides of the preceding inequality, we have

$$\begin{aligned} \text{LHS of (17)} \leq & \frac{k^2}{n} + \frac{1}{6n} - \frac{1}{12}\left(\frac{1}{n+k} + \frac{1}{n-k}\right) + \frac{1}{360}\left(\frac{1}{(n+k)^3} + \frac{1}{(n-k)^3}\right) \\ & - \left(n+k+\frac{1}{2}\right) \ln\left(1+\frac{k}{n}\right) - \left(n-k+\frac{1}{2}\right) \ln\left(1-\frac{k}{n}\right). \end{aligned} \tag{18}$$

Plugging $k = t\sqrt{n}$ into the RHS of (18), we can write the result as $I_1 + I_2 + I_3$, where

$$I_1 = -\,n\left(\left(1-\frac{t}{\sqrt{n}}\right)\ln\left(1-\frac{t}{\sqrt{n}}\right)+\left(\frac{t}{\sqrt{n}}+1\right)\ln\left(\frac{t}{\sqrt{n}}+1\right)\right),$$
$$I_2 = -\,\frac{1}{2}\left(\ln\left(1-\frac{t}{\sqrt{n}}\right)+\ln\left(\frac{t}{\sqrt{n}}+1\right)\right),$$
$$I_3 = -\,\frac{1}{12\left(n-\sqrt{n}t\right)}-\frac{1}{12\left(\sqrt{n}t+n\right)}+\frac{1}{360\left(n-\sqrt{n}t\right)^3}+\frac{1}{360\left(\sqrt{n}t+n\right)^3}$$
$$+\frac{1}{6n}+t^2.$$

Then we want to prove that for n large enough,

$$I_1 + I_2 + I_3 < 0. \tag{19}$$

Then as a consequence, (17) will hold.

By Corollary 12 and the fact that $PH(n,k)$ is decreasing in k for n,k integers and $k \geq t\sqrt{n}$ where $t \geq \sqrt{3}$, if we can show that (19) holds for the smallest integer k such that $\sqrt{3n} \leq k \leq n$, then (16) will hold for all integers $\sqrt{3n} \leq k \leq n$. Notice that if k is the smallest integer not smaller than $\sqrt{3n}$, then $\sqrt{3n} \leq k < \sqrt{3n}+1$. It is equivalent to say that $\sqrt{3} \leq t \leq \left(\sqrt{3n}+1\right)/\sqrt{n}$, and the RHS is smaller than 2 for all $n \geq 14$. So our goal now is to prove (19) holds for all $n \geq 108$, as assumed in the proposition, and $\sqrt{3} \leq t < 2$.

By Taylor's expansion of $(1+x)\ln(1+x)+(1-x)\ln(1-x)$ around $x=0$, we find an upper bound for I_1, given by

$$\begin{aligned} I_1 &= -n\left(\sum_{i=1}^{\infty}\frac{t^{2i}}{n^i i(2i-1)}\right) \\ &< -t^2-\frac{t^4}{6n}-\frac{t^6}{15n^2}-\frac{t^8}{28n^3}. \end{aligned} \tag{20}$$

For I_2, by using Taylor's expansion again, we have

$$\begin{aligned} I_2 = -\frac{1}{2}\left(\ln\left(1-\frac{t^2}{n}\right)\right) &= \sum_{j=1}^{\infty}\frac{1}{2j}\left(\frac{t^2}{n}\right)^j \\ &\leq \frac{t^2}{2n}+\frac{t^4}{4n^2}+\frac{1}{2}R_3, \end{aligned} \tag{21}$$

where $R_3 = \sum_{j=3}^{\infty}\frac{1}{j}\left(\frac{t^2}{n}\right)^j < \frac{1}{3}\sum_{j=3}^{\infty}\left(\frac{t^2}{n}\right)^j = t^6/\left[3n^3\left(1-\frac{t^2}{n}\right)\right]$.

We only need to show (19) holds for all $\sqrt{3} \leq t < 2$, and thus want to bound $t^6/\left[3n^3\left(1-\frac{t^2}{n}\right)\right]$ by a sharp upper bound. This means we want $\frac{t}{\sqrt{n}}$ to be small. We have $n \geq 64$, which implies $\frac{t}{\sqrt{n}} < \frac{1}{4}$. Then we have an upper bound for R_3:

$$R_3 \leq \frac{1}{3}\frac{t^6}{(15n^3/16)}.$$

It follows that

$$I_2 \leq \frac{t^2}{2n} + \frac{t^4}{4n^2} + \frac{8t^6}{45n^3}. \tag{22}$$

We now bound I_3 by studying two summands separately. For the first part of I_3, we have

$$\begin{aligned} -\frac{1}{12\left(n-\sqrt{n}t\right)} - \frac{1}{12\left(\sqrt{n}t+n\right)} &= -\frac{1}{12n}\left(\frac{1}{1-t/\sqrt{n}} + \frac{1}{1+t/\sqrt{n}}\right) \\ &= -\frac{1}{6n}\left(1+\left(\frac{t}{\sqrt{n}}\right)^2 + \left(\frac{t}{\sqrt{n}}\right)^4 + \dots\right) \\ &< -\frac{1}{6n} - \frac{t^2}{6n^2}. \end{aligned}$$

For the second part of I_3, we have that when $t/\sqrt{n} \leq 1/4$,

$$\begin{aligned} \frac{1}{\left(\sqrt{n}t+n\right)^3} + \frac{1}{\left(n-\sqrt{n}t\right)^3} &= \frac{1}{n^3}\left(\frac{1}{\left(1+t/\sqrt{n}\right)^3} + \frac{1}{\left(1-t/\sqrt{n}\right)^3}\right) \\ &< \frac{1}{n^3}\left(\frac{1}{\left(5/4\right)^3} + \frac{1}{\left(3/4\right)^3}\right) \\ &< 3/n^3. \end{aligned}$$

Therefore we have

$$I_3 < -\frac{t^2}{6n^2} + \frac{3}{n^3} + t^2.$$

Summing I_1 through I_3, we have

$$\begin{aligned} I_1 + I_2 + I_3 &< t^2 - \frac{t^8}{28n^3} - \frac{t^6}{15n^2} - \frac{t^4}{6n} - t^2 + \frac{t^2}{2n} + \frac{t^4}{4n^2} + \frac{8t^6}{45n^3} - \frac{t^2}{6n^2} + \frac{3}{n^3} \\ &< \frac{1}{n}\left(\frac{t^2}{2} - \frac{t^4}{6}\right) + \frac{1}{n^2}\left(-\frac{t^2}{6} + \frac{t^4}{4} - \frac{t^6}{15}\right) + \frac{1}{n^3}\left(3 - \frac{t^8}{28} + \frac{8t^6}{45}\right) \end{aligned} \tag{23}$$

when $\dfrac{t}{\sqrt{n}} < \dfrac{1}{4}$, i.e., $n \geq 16t^2$.

We now want to show that $I_1 + I_2 + I_3 < 0$ for all $n \geq 108$ and $\sqrt{3} \leq t < 2$. We will consider the coefficients of $\frac{1}{n}, \frac{1}{n^2}, \frac{1}{n^3}$ in (23). The coefficient of $\frac{1}{n}$ is $\frac{t^2}{2} - \frac{t^4}{6}$, which is decreasing in t when $\sqrt{3} \leq t < 2$; thus by plugging in $t = \sqrt{3}$, we have

$$\frac{t^2}{2} - \frac{t^4}{6} \leq 0.$$

The coefficient of $\frac{1}{n^2}$ is $-\frac{t^6}{15} + \frac{t^4}{4} - \frac{t^2}{6}$, which is also decreasing in t when $\sqrt{3} \leq t < 2$. Thus by plugging in $t = \sqrt{3}$, we have

$$-\frac{t^6}{15} + \frac{t^4}{4} - \frac{t^2}{6} \leq -\frac{1}{20}.$$

The coefficient of $\frac{1}{n^3}$ is $-\frac{t^8}{28} + \frac{8t^6}{45} + 3$. By calculation, we have that when $\sqrt{3} \leq t < 2$,

$$-\frac{t^8}{28} + \frac{8t^6}{45} + 3 < 5.4.$$

Thus when $n \geq 108 > 64$ and $\sqrt{3} \leq t < 2$, we have

$$I_1 + I_2 + I_3 < \frac{5.4}{n^3} - \frac{1}{20n^2}. \tag{24}$$

Therefore if we can show that for some n,

$$\frac{5.4}{n^3} - \frac{1}{20n^2} \leq 0, \tag{25}$$

then $I_1 + I_2 + I_3 < 0$ for those n. Solving (25), we obtain $n \geq 108$. □

Remark. The coefficient of $\frac{1}{n}$ in (23) is the same as the coefficient of $\frac{1}{n}$ in the Taylor expansion of $I_1 + I_2 + I_3$. So when the leading coefficient $\frac{t^2}{2} - \frac{t^4}{6}$ is positive, i.e., $t < \sqrt{3}$, the upper bound $2\binom{2n}{n+k}/\binom{2n}{n}$ from Lemma 6 will tend to be larger than $e^{-k^2/n}$.

Now we want to show that (7) holds for all integer pairs $(n, t\sqrt{n})$ with $\sqrt{\ln 2} < t < \sqrt{3}$ and n greater than some fixed value. By the argument in the remark, we need to choose another upper bound for $P_{n,n,M}$.

14. **Lemma.** *We have* $P_{n,n,M} \leq \dfrac{2\binom{2n}{n+k} - \binom{2n}{n+2k}}{\binom{2n}{n}}$, *where* $M = k/\sqrt{2n}$, $k = 1, \ldots, n$.

Proof. Let A be the event that $\sup \sqrt{n}(F_n - G_n) \geq M$ and B the event that $\inf \sqrt{n}(F_n - G_n) \leq -M$. We want an upper bound for $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$. Let S_j be the value after j steps of a simple, symmetric random walk on the integers starting at 0. Then

$$\Pr(S_{2n} = 2m) = \frac{1}{4^n}\binom{2n}{n+m}$$

for $m = -n, -n+1, \cdots, n-1, n$. By a well-known reflection principle we have nice exact expressions for $\Pr(A)$ and $\Pr(B)$,

$$\Pr(A) = \Pr(B) = \frac{\Pr(S_{2n} = 2k)}{\Pr(S_{2n} = 0)} = \frac{\binom{2n}{n+k}}{\binom{2n}{n}}.$$

Therefore we want a lower bound for $\Pr(A \cap B)$. Let C be the event that for some $s < t$, $\sqrt{n}(F_n - G_n)(s) \geq M$ and $\sqrt{n}(F_n - G_n)(t) \leq -M$. Then we can exactly evaluate $\Pr(C)$ by two reflections, e.g. [9], specifically,

$$\Pr(C) = \frac{\Pr(S_{2n} = 4k)}{\Pr(S_{2n} = 0)} = \frac{\binom{2n}{n+2k}}{\binom{2n}{n}},$$

and $C \subset A \cap B$, so the bound holds. □

15. **Lemma.** *Let n, k be positive integers, $n \geq 372$, and $\sqrt{2n} < k = t\sqrt{n} \leq \sqrt{3n}$. Then*

$$\binom{2n}{n+2k} > \binom{2n}{n+k} e^{-3t^2 - 0.05}.$$

Proof. By Stirling's formula with error bounds, we have

$$\ln\left(\frac{\binom{2n}{n+2k}}{\binom{2n}{n+k}}\right) = \ln\left(\frac{(n+k)!(n-k)!}{(n+2k)!(n-2k)!}\right) > \ln(A_n)$$

where A_n is defined as

$$\frac{(n-k)^{n-k+\frac{1}{2}}(k+n)^{k+n+\frac{1}{2}}\exp\left(\frac{1}{12}\left[\frac{1}{k+n}+\frac{1}{n-k}\right]-\frac{1}{360}\left[\frac{1}{(k+n)^3}+\frac{1}{(n-k)^3}\right]\right)}{\exp\left(\frac{1}{12(2k+n)}+\frac{1}{12(n-2k)}\right)(n-2k)^{-2k+n+1/2}(2k+n)^{2k+n+1/2}},$$

and so

$$\begin{aligned}\ln(A_n) &= -\frac{1}{12(2k+n)}-\frac{1}{12(n-2k)}+\frac{1}{12(n-k)}+\frac{1}{12(k+n)}\\ &\quad -\frac{1}{360(n-k)^3}-\frac{1}{360(k+n)^3}-\left(-2k+n+\frac{1}{2}\right)\ln(n-2k)\\ &\quad +\left(-k+n+\frac{1}{2}\right)\ln(n-k)+\left(k+n+\frac{1}{2}\right)\ln(k+n)\\ &\quad -\left(2k+n+\frac{1}{2}\right)\ln(2k+n)\\ &= I_4+I_5, \qquad (26)\end{aligned}$$

where

$$\begin{aligned}I_4 &= -\left(-2k+n+\frac{1}{2}\right)\ln(n-2k)+\left(-k+n+\frac{1}{2}\right)\ln(n-k)\\ &\quad +\left(k+n+\frac{1}{2}\right)\ln(k+n)-\left(2k+n+\frac{1}{2}\right)\ln(2k+n),\\ I_5 &= \frac{1}{12(n-k)}+\frac{1}{12(k+n)}-\frac{1}{12(2k+n)}-\frac{1}{12(n-2k)}\\ &\quad -\frac{1}{360(n-k)^3}-\frac{1}{360(k+n)^3}.\end{aligned}$$

Using again (20) and (21), we have for $|x|<1$,

$$\begin{aligned}x^2+\frac{x^4}{6} &< (1-x)\ln(1-x)+(x+1)\ln(x+1)\\ &< x^2+\frac{x^4}{6}+\frac{1}{15}\sum_{i=3}^{\infty}x^{2i} = x^2+\frac{x^4}{6}+\frac{x^6}{15(1-x^2)},\end{aligned}$$

and also

$$\begin{aligned}-x^2 &> \ln(1-x)+\ln(x+1)\\ &> -x^2-\frac{1}{2}\sum_{i=2}^{\infty}x^{2i} = -x^2-\frac{1}{2}\frac{x^4}{(1-x^2)}.\end{aligned}$$

So by plugging in $k = t\sqrt{n}$, we have that for $\dfrac{t}{\sqrt{n}} < \dfrac{1}{4}$,

$$\begin{aligned}
I_4 = {} & n\left(\left(1-\frac{k}{n}\right)\ln\left(1-\frac{k}{n}\right)+\left(\frac{k}{n}+1\right)\ln\left(\frac{k}{n}+1\right)\right) \\
& + \frac{1}{2}\left(\ln\left(1-\frac{k}{n}\right)+\ln\left(\frac{k}{n}+1\right)\right) \\
& - n\left(\left(1-\frac{2k}{n}\right)\ln\left(1-\frac{2k}{n}\right)+\left(\frac{2k}{n}+1\right)\ln\left(\frac{2k}{n}+1\right)\right) \\
& - \frac{1}{2}\left(\ln\left(1-\frac{2k}{n}\right)+\ln\left(\frac{2k}{n}+1\right)\right) \\
> {} & n\left(\left(\frac{t}{\sqrt{n}}\right)^2+\frac{1}{6}\left(\frac{t}{\sqrt{n}}\right)^4\right)-\frac{1}{2}\left(\left(\frac{t}{\sqrt{n}}\right)^2+\frac{8}{15}\left(\frac{t}{\sqrt{n}}\right)^4\right) \\
& - n\left(\left(\frac{2t}{\sqrt{n}}\right)^2+\frac{1}{6}\left(\frac{2t}{\sqrt{n}}\right)^4+\frac{4}{45}\left(\frac{2t}{\sqrt{n}}\right)^6\right)+\frac{1}{2}\left(\frac{2t}{\sqrt{n}}\right)^2 \\
= {} & t^2+\frac{t^4}{6n}-\frac{t^2}{2n}-\frac{4t^4}{15n^2}-4t^2-\frac{8t^4}{3n}-\frac{256t^6}{45n^2}+\frac{2t^2}{n} \\
= {} & -\frac{1}{n^2}\left(\frac{256t^6}{45}+\frac{4t^4}{15}\right)+\frac{1}{n}\left(\frac{3t^2}{2}-\frac{5t^4}{2}\right)-3t^2.
\end{aligned}$$

Now we proceed to find a lower bound for I_5. For all $k \le n/8$, in other words $t := k/\sqrt{n}$ such that $8t \le \sqrt{n}$,

$$\begin{aligned}
I_5 = {} & \frac{1}{12}\left(\frac{1}{n-k}+\frac{1}{k+n}-\frac{1}{(2k+n)}-\frac{1}{(n-2k)}\right) \\
& - \frac{1}{360}\left(\frac{1}{(k+n)^3}+\frac{1}{(n-k)^3}\right) \\
= {} & \frac{1}{12}\left(\frac{1}{\sqrt{n}t+n}+\frac{1}{n-\sqrt{n}t}-\frac{1}{2\sqrt{n}t+n}-\frac{1}{n-2\sqrt{n}t}\right) \\
& - \frac{1}{360}\left(\frac{1}{(\sqrt{n}t+n)^3}+\frac{1}{\left(n-\sqrt{n}t\right)^3}\right) \\
= {} & \frac{1}{6(n-t^2)}-\frac{1}{6\left(n-4t^2\right)}-\frac{n+3t^2}{180n\left(n-t^2\right)^3} \\
> {} & \frac{1}{6n}-\frac{1}{3n}-\frac{n+3t^2}{90n^4} \\
= {} & -\frac{1}{6n}-\frac{1}{90n^3}-\frac{t^2}{30n^4}.
\end{aligned}$$

Since $t \le \sqrt{3}$, we know that as long as $n \ge 192$, the condition $8t \le \sqrt{n}$ will hold.

Adding our lower bounds for I_4 and I_5, we have that when $n \ge 192$ and $\sqrt{\ln 2} \le t \le \sqrt{3}$,

$$\begin{aligned}
I_4 + I_5 \quad > \quad & -\frac{t^2}{30n^4}-\frac{1}{90n^3}-\frac{1}{n^2}\left(\frac{256t^6}{45}+\frac{4}{15}t^4\right)-\frac{1}{n}\left(\frac{5t^4}{2}-\frac{3t^2}{2}+\frac{1}{6}\right)-3t^2 \\
> \quad & -3t^2-\gamma, \qquad (27)
\end{aligned}$$

for some γ. When $\gamma = 0.05$, we want to show that for n large enough, (27) always holds. In other words, we need

$$0.05 > \frac{t^2}{30n^4} + \frac{1}{90n^3} + \frac{1}{n^2}\left(\frac{256t^6}{45} + \frac{4}{15}t^4\right) + \frac{1}{n}\left(\frac{5t^4}{2} - \frac{3t^2}{2} + \frac{1}{6}\right). \tag{28}$$

Notice that when $\sqrt{\ln 2} < t < \sqrt{3}$, the coefficient $\frac{5t^4}{2} - \frac{3t^2}{2} + \frac{1}{6}$ is positive and is increasing in t; the RHS of (28) is increasing in t and decreasing in n. Thus we just need to make sure the inequality holds for $t = \sqrt{3}$. Therefore we need

$$0.05 > \frac{1}{10n^4} + \frac{1}{90n^3} + \frac{156}{n^2} + \frac{109}{6n}. \tag{29}$$

Solving (29) numerically, we find that it holds for $n \geq 372$.

Therefore, by (26) and (27), we have shown that when $n \geq 372$,

$$\ln\left[\binom{2n}{n+2k}\Big/\binom{2n}{n+k}\right] > -3t^2 - 0.05,$$

for $k = t\sqrt{n}$ and $\sqrt{\ln 2} < t < \sqrt{3}$, proving Lemma 15. □

16. **Proposition.** *Let $k = t\sqrt{n}$, where $\sqrt{\ln 2} < t < \sqrt{3}$, and k, n integers. Then the inequality*

$$\left[2\binom{2n}{n+k} - \binom{2n}{n+2k}\right]\Big/\binom{2n}{n} < 2\exp\left(-k^2/n\right)$$

holds for $n \geq 6395$.

Proof. By Lemma 15, it will suffice to show that for $n \geq 6395 > 372$,

$$\binom{2n}{n+k}\left(1 - e^{-3t^2-0.05}/2\right)\Big/\binom{2n}{n} < \exp(-k^2/n). \tag{30}$$

Rewriting (30) by taking logarithms of both sides, we just need to show

$$\ln\binom{2n}{n+k} - \ln\binom{2n}{n} + \frac{k^2}{n} + \ln\left(1 - e^{-3t^2-0.05}/2\right) < 0.$$

By (17), (18), and (23), we have that

$$\ln\binom{2n}{n+k} - \ln\binom{2n}{n} + \frac{k^2}{n} < \frac{3 - \frac{t^8}{28} + \frac{4t^6}{45}}{n^3} + \frac{-\frac{t^2}{6} + \frac{t^4}{4} - \frac{t^6}{15}}{n^2} + \frac{\frac{t^2}{2} - \frac{t^4}{6}}{n}$$

for $n > 16t^2$. So now we just need

$$\frac{3 - \frac{t^8}{28} + \frac{4t^6}{45}}{n^3} + \frac{-\frac{t^2}{6} + \frac{t^4}{4} - \frac{t^6}{15}}{n^2} + \frac{\frac{t^2}{2} - \frac{t^4}{6}}{n} + \ln\left(1 - e^{-3t^2-0.05}/2\right) < 0. \tag{31}$$

When $\sqrt{\ln 2} < t < \sqrt{3}$, the coefficient $\dfrac{t^2}{2} - \dfrac{t^4}{6} > 0$. Next, using $t < \sqrt{3}$,

$$\begin{aligned}
&\frac{1}{n^3}\left(3 - \frac{t^8}{28} + \frac{4t^6}{45}\right) + \frac{1}{n^2}\left(-\frac{t^2}{6} + \frac{t^4}{4} - \frac{t^6}{15}\right) + \frac{1}{n}\left(\frac{t^2}{2} - \frac{t^4}{6}\right) \\
< \;& \frac{1}{n}\left(\frac{t^2}{2} - \frac{t^4}{6}\right) + \frac{t^4}{4n^2} + \frac{1}{n^3}\left(3 + \frac{4t^6}{45}\right) \\
< \;& \frac{1}{n}\left(\frac{t^2}{2} - \frac{t^4}{6}\right) + \frac{9}{4n^2} + \frac{27}{5n^3}.
\end{aligned}$$

Clearly, the maximum value of $\ln\left(1-e^{-3t^2-0.05}/2\right)$ for $\sqrt{\ln 2} \le t \le \sqrt{3}$ is achieved when $t=\sqrt{3}$. Plugging in $t=\sqrt{3}$ into $\ln\left(1-e^{-3t^2-0.05}/2\right)$, we have

$$\ln\left(1-e^{-3t^2-0.05}/2\right) \le -0.0000586972.$$

Now we find the maximum value of $\frac{t^2}{2}-\frac{t^4}{6}$ for $\sqrt{\ln 2} \le t \le \sqrt{3}$. The derivative with respect to t is $t-\dfrac{2t^3}{3}$, which equals zero when $t=\sqrt{1.5}$. This critical point corresponds to the maximum value of $\frac{t^2}{2}-\frac{t^4}{6}$ for $\sqrt{\ln 2}<t<\sqrt{3}$, and this maximum value is 0.375.

Accordingly, when $\sqrt{\ln 2}<t<\sqrt{3}$,

$$\text{LHS of (31) } < -0.0000586972+\frac{9}{4n^2}+\frac{39}{5n^3}+\frac{3}{8n}.$$

We just need

$$-0.0000586972+\frac{9}{4n^2}+\frac{39}{5n^3}+\frac{3}{8n}<0. \tag{32}$$

The LHS of (32) is decreasing in $n>0$. By numerically solving the inequality in n we have that $n \ge 6395$. Therefore we have proved that when $n>6395$, the original inequality (7) holds for all positive integer pairs (k,n) such that $\sqrt{n\ln 2}<k<\sqrt{3n}$ and $k \le n$. $\square$

Recall that by (9), the inequality (7) holds for all $k \le \sqrt{n\ln 2}$. Combining Propositions 13 and 16, we have the following conclusion.

17. **Theorem.** *(a) When $n \ge 6395$, (7) holds for all (n,k) such that $0 \le k \le n$.*
(b) When $6395>n\ge 372$, (7) holds for all integer pairs (n,k) such that $0\le k\le \sqrt{n\ln 2}$ and $\sqrt{3n}<k\le n$.

Then by computer searching for the rest of the integer pairs (n,k), namely, $1\le k\le n$ when $1\le n\le 371$ and $\sqrt{n\ln 2}<k\le\sqrt{3n}$ when $372\le n<6395$, we are able to find the finitely many counterexamples to the inequality (7), and thus prove Theorem 1.

3. Treatment of $m \neq n$

3.1. **One- and two-sided probabilities.** For given positive integers $1\le m\le n$ and d with $0<d\le 1$, let pv_{os} be the one-sided probability

$$pv_{os}(m,n,d)=\Pr(\sup_x (F_m-G_n)(x)\ge d)=\Pr(\inf_x (F_m-G_n)(x)\le -d), \tag{33}$$

where the equality holds by symmetry (reversing the order of the observations in the combined sample). Let the two-sided probability (p-value) be

$$P(m,n;d):=\Pr(\sup_x |(F_m-G_n)(x)|\ge d).$$

The following is well known, e.g. for part (b), [10, p. 472], and easy to check:

18. **Theorem.** *For any positive integers m and n and any d with $0<d\le 1$ we have*
(a) $pv_{os}(m,n,d)\le P(m,n;d)\le pv_{ub}(m,n,d):=2pv_{os}(m,n,d)$.
(b) If $d>1/2$, $P(m,n;d)=pv_{ub}(m,n,d)$.

3.2. **Computational methods.** To compute p-values $P(m, n; d)$ for the 2-sample test for $d \leq 1/2$ we used the Hodges (1957) "inside" algorithm, for which Kim and Jennrich [11] gave a Fortran program and tables computed with it for $m \leq n \leq 100$. We further adapted the program to double precision. The method seems to work reasonably well for $m \leq n \leq 100$; for $n = 2m$ with $m \leq 94$ and $d = (m+1)/n$ it still gives one or two correct significant digits, see Table 1. The inside method finds p-values $\Pr(D_{m,n} \geq d)$ as $1 - \Pr(D_{m,n} < d)$. When p-values are very small, e.g. of order 10^{-15}, the subtraction can lead to substantial or even total loss of significant digits, due to subtracting numbers very close to 1 from 1 (again see Table 1).

The one-sided probabilities $pv_{os}(m, n, d)$ and thus $P(m, n; d)$ for $d > 1/2$ by Theorem 18(b) can be computed by an analogous "outside" method with only additions and multiplications (no subtractions), so it can compute much smaller probabilities very accurately. The smallest probability needed for computing the results of the paper is $\Pr(D_{300,600} \geq 1)$ which was evaluated by the outside program as $1.147212371856 \cdot 10^{-247}$, confirmed to the given number (13) of significant digits by evaluating $2/\binom{900}{300}$. Moreover the ratio of this to $2\exp(-2M^2)$ is about $3 \cdot 10^{-74}$, so great accuracy in the p-value is not needed to see that the ratio is small. For $m = n$ we can compare results of the outside method to those found from the Gnedenko–Korolyuk formula in Proposition 5. For $\Pr(D_{500,500} \geq 0.502)$ the outside method needs to add a substantial number of terms. It gives $1.87970906825 \cdot 10^{-57}$ which agrees with the Gnedenko–Korolyuk result to the given accuracy.

For large enough m, n there will be an interval of values of d,

$$d_0(m, n) \leq d \leq 1/2, \tag{34}$$

in which the p-values pvi as computed by the inside method are small, $pvi \leq p_0$, and so may have fewer reliable significant digits than is desirable. We can verify the DKWM inequality in these ranges using Theorem 18(a) by showing that

$$pv_{ub}(m, n, d) \leq 2\exp(-2M^2) \tag{35}$$

where as usual $M = \sqrt{mn/(m+n)}d$, and did so computationally for $3 \leq m < n \leq 200$ and $202 \leq n = 2m \leq 600$ as shown by ratios less than 1 in the last columns of Tables 6 and 7, and less than $r_{\max}$ in Table 8, respectively. These moreover show that $r_{\max}$ is not attained in the range (34) for any (m, n) covered by the three tables.

Originally we had taken $p_0 = 10^{-14}$ or 10^{-13} and actually found no spurious values of $r_{\max}$ (as we had with $p_0 = 10^{-15}$), but then pvi might have had few reliable significant digits. In the present version, we have taken $p_0 = 10^{-8}$, giving 5 or 6 additional reliable significant digits to pvi where it is computed.

With either the inside or outside method, evaluation of an individual probability takes $O(mn)$ computational steps, which is more (slower) than for $m = n$. For mn large, rounding errors accumulate, which especially affect the inside method. Moreover, to find the p-values for all possible values of D_{mn}, in the general case that m and n are relatively prime, as in a study like the present one, gives another factor of mn and so takes $O(m^2n^2)$ computational steps.

The algorithm does not require storage of $m \times n$ matrices. Four vectors of length n, and various individual variables, are stored at any one time in the computation.

For $n = 2m$, the smallest possible $d > 1/2$ is $d = (m+1)/n$. Let pvo be the p-value $\Pr(D_{m,n} \geq d)$ as computed by the outside method. Let the relative error of pvi as an approximation to the more accurate pvo be $reler = \left|\frac{pvi}{pvo} - 1\right|$. For $n = 2m$, $m = 1, \ldots, 120$, and $d = (m+1)/n$, the following $m = m_{\max}$ give larger $reler$ than for any $m < m_{\max}$, with the given pvo.

TABLE 1. p-values for $n = 2m$, $d = (m+1)/n$

$m_{\max}$	$reler$	pvo
10	$5.55 \cdot 10^{-15}$	0.0290
20	$7.88 \cdot 10^{-13}$	$8.94 \cdot 10^{-4}$
28	$2.04 \cdot 10^{-12}$	$5.48 \cdot 10^{-5}$
40	$1.32 \cdot 10^{-9}$	$8.29 \cdot 10^{-7}$
49	$6.51 \cdot 10^{-9}$	$3.58 \cdot 10^{-8}$
60	$1.01 \cdot 10^{-6}$	$7.66 \cdot 10^{-10}$
70	$4.76 \cdot 10^{-5}$	$2.32 \cdot 10^{-11}$
80	$2.19 \cdot 10^{-3}$	$7.07 \cdot 10^{-13}$
93	0.063	$7.52 \cdot 10^{-15}$
95	0.109	$3.74 \cdot 10^{-15}$
98	0.525	$1.31 \cdot 10^{-15}$
100	1.045	$6.52 \cdot 10^{-16}$
105	9.758	$1.14 \cdot 10^{-16}$
120	2032.4	$6.01 \cdot 10^{-19}$

The small relative errors for $m \leq 10, 20$, or 40, indicate that the inside and outside programs algebraically confirm one another. As m increases, pvo becomes smaller and $reler$ tends to increase until for $m = 100$, pvi has no accurate significant digits. For $m = 105$, pvi is off by an order of magnitude and for $m = 120$ by three orders. For $m = 122$, $n = 244$, and $d = 123/244$, for which $pvo = 2.99 \cdot 10^{-19}$, pvi is negative, $-4.44 \cdot 10^{-16}$. That is, the inside computation gave $\Pr(D_{122,244} < 123/244) \doteq 1 + 4.44 \cdot 10^{-16}$, which is accurate to 15 decimal places but useless.

Of course, p-values of order 10^{-15} are not needed for applications of the Kolmogorov–Smirnov test even to, say, tens of thousands of simultaneous hypotheses as in genetics, but in this paper we are concerned with the theoretical issue of validity of the DKWM bound.

3.3. **Details related to Facts 2, 3, and 4.** Fact 2(b) states that for $1 \leq m < n \leq 3$ the DKWM inequality fails. The following lists $r_{\max}(m, n) > 1$ for each of the three pairs and the $d_{\max}$, equal to 1 in these cases, for which $r_{\max}$ is attained.

m	n	$r_{\max}$	$d_{\max}$
1	2	1.264556	1
1	3	1.120422	1
2	3	1.102318	1

Fact 2(a) states that if $1 \leq m < n \leq 200$ and $n \geq 4$, the DKWM inequality holds. Searching through the specified n for each m, we got the following.

For $m = 1, 2$, the results of Fact 2(f) as stated were found.

For $3 \leq m \leq 199$ and $m < n \leq 200$ we searched over n for each m, finding $r_{\max}(m, n)$ for each n and the $n = n_{\max}$ giving the largest $r_{\max}$. Tables 6 and 7 in Appendix B show that all $r_{\max} < 1$, completing the evidence for Fact 2(a), and were always found at $n_{\max} = 2m$ for $m \leq 100$, as Fact 2(c) states.

For Fact 2 (d) and (e) and Fact 3, the results stated can be seen in Tables 7 and 8.

Fact 3(a) in regard to relative minima of $r_{\max}$ is seen to hold in Table 6. Increasing $r_{\max}$ for $16 \le m \le 300$ is seen in Tables 6 and 8. Fact 3(b) is seen in Table 8.

In Fact 3(c), the minimal $r_{\max}(m, 2m)$ for $m \ge 101$ is at $m = 101$ by part (a) with value 0.973341 in Table 8. The largest $r_{\max}$ in Table 7 is $0.949565 < 0.973341$ as seen with the aid of Fact 2(d). For Fact 3(d), one sees that $k_{\max}$ is nondecreasing in m in Tables 6 and 8.

Regarding Fact 4, the relative error of the DKWM bound as an approximation of a p-value, namely

$$reler(dkwm, m, n, d) := \frac{2\exp(-2M^2)}{P_{m,n,M}} - 1, \tag{36}$$

where M is as in (3) with $d = k/L_{m,n}$, is bounded below for any possible d by

$$reler(dkwm, m, n, d) \ge \frac{1}{r_{\max}(m, n)} - 1. \tag{37}$$

From our results, over the given ranges, the relative error has the best chance to be small when $n = m$ and the next-best chance when $n = 2m$. On the other hand, in Table 7 in Appendix B, where $rmaxx = rmaxx(m) = \max_{m<n\le 200} r_{\max}(m, n)$, we have for each m, n with $100 < m < n \le 200$ and possible d that

$$reler(dkwm, m, n, d) \ge \frac{1}{rmaxx(m)} - 1. \tag{38}$$

Thus Fact 4(a) holds by Fact 3(c) and the near-equality of $\beta(M)$ and $2\exp(-2M^2)$ if either is ≤ 0.05 (6). Fact 4(b) holds similarly by inspection of Table 7.

3.4. **Conservative and approximate p-values.** Whenever the DKWM inequality holds, the DKWM bound $2\exp(-2M^2)$ provides simple, conservative p-values. The asymptotic p-value $\beta(M)$ given in (5) is very close to the DKWM bound in case of significance level ≤ 0.05 or less (6).

In general, by Fact 4 for example, using the DKWM bound as an approximation can give overly conservative p-values. We looked at $m = 20$, $n = 500$. For $\alpha = 0.05$ the correct critical value for $d = k/500$ is $k = 151$ whereas the approximation would give $k = 155$; for $\alpha = 0.01$ the correct critical value is $k = 180$ but the approximation would give $k = 186$. For $180 \le k \le 186$ the ratio of the true p-value to its DKWM approximation decreases from 0.731 down to 0.712.

Stephens [15] proposed that *in the one-sample case*, letting $N_e := n$ and

$$F := \sqrt{N_e} + 0.12 + 0.11/\sqrt{N_e}, \tag{39}$$

one can approximate p-values by $\Pr(D_n \ge d) \sim \beta(Fd)$ for $0 < d \le 1$, with β from (5). Stephens gave evidence that the approximation works rather well. In the one-sample case the distributions of the statistics D_n and K_n are continuous for fixed n and vary rather smoothly with n.

Some other sources, e.g. [14, pp. 617-619], propose in the two-sample case setting $N_e = mn/(m + n)$, defining $F := F_{m,n}$ by (39), and approximating $\Pr(D_{m,n} \ge d)$ by $S_{\text{pli}} := \beta(Fd)$ ["Stephens approximation plugged into" two-sample]. Since F in (39) is always larger than $\sqrt{N_e}$, S_{pli} is always less than the asymptotic probability $\beta(M)$ for $M = \sqrt{N_e}d$ which, in turn, is always less than the DKWM approximation

$2\exp(-2M^2)$. The approximation $S_{\rm pli}$ is said in at least two sources we have seen (neither a journal article) to be already quite good for $N_e \geq 4$. That may well be true in the one-sample case. In the two-sample case it may be true when $1 < m \ll n$ but not when $n \sim m$. Table 2 compares the two approximations $dkwm = 2\exp(-2M^2)$ and $S_{\rm pli}$ to critical p-values for some pairs (m, n). For $m = n$, and to a lesser extent when $n = 2m$, it seems that $dkwm$ is preferable. For other pairs, $S_{\rm pli}$ is. For the six pairs (m, n) with $L_{m,n} = n$ or $2n$, $S_{\rm pli} < pv$. For the other two (relatively prime) pairs, $pv < S_{\rm pli}$. For $m = 39, n = 40$, $S_{\rm pli}$ has rather large errors, but those of $dkwm$ are much larger.

In Table 2, $d = k/L_{m,n}$ and pv is the true p-value. After each of the two approximations, $dkwm$ and $S_{\rm pli}$, is its relative error $reler$ in approximating pv.

Table 2. Comparing two approximations to p-values

m	n	N_e	k	d	pv	$dkwm$	$reler$	$S_{\rm pli}$	$reler$
40	40	20	12	.3	.05414	.05465	.0094	.04313	.2033
40	40	20	13	.325	.02860	.02925	.0226	.02216	.2253
40	40	20	14	.35	.014302	.01489	.0413	.01079	.2453
40	40	20	15	.375	.006761	.00721	.0669	.00498	.2628
200	200	100	27	.135	.05214	.05224	.0020	.04745	.0899
200	200	100	28	.14	.03956	.03968	.0030	.03578	.0955
200	200	100	32	.16	.011843	.01195	.0092	.01044	.1183
200	200	100	33	.165	.008539	.00864	.0113	.00748	.1240
25	50	16.67	16	.32	.06066	.06586	.0858	.05129	.1545
25	50	16.67	17	.34	.03847	.04242	.1025	.03198	.1687
25	50	16.67	19	.38	.014149	.01624	.1479	.01141	.1933
25	50	16.67	20	.4	.008195	.00966	.1783	.00653	.2029
39	40	19.75	456	.2923	.05145	.06847	.3309	.05476	.0644
39	40	19.75	457	.2929	.04968	.06746	.3579	.05390	.0850
39	40	19.75	541	.3468	.010159	.01731	.7036	.01264	.2439
39	40	19.75	542	.3474	.009849	.01701	.7267	.01240	.2593
20	500	19.23	150	.3	.05059	.06276	.2406	.04973	.0171
20	500	19.23	151	.302	.04817	.05992	.2439	.04733	.0175
20	500	19.23	179	.358	.010608	.01446	.3634	.01038	.0214
20	500	19.23	180	.36	.009998	.01368	.3688	.009787	.0211
21	500	20.15	3074	.29276	.050052	.06319	.2626	.050410	.0072
21	500	20.15	3076*	.29295	.049882	.06291	.2612	.050170	.0058
21	500	20.15	3686	.35105	.010040	.01392	.3869	.010062	.0022
21	500	20.15	3687	.35114	.009979	.01389	.3917	.010033	.0054
100	500	83.33	73	.146	.0534470	.0572963	.07202	.051661	.03343
100	500	83.33	74	.148	.0483882	.0519476	.07356	.0467046	.03479
100	500	83.33	88	.176	.0104170	.0114528	.09943	.0098532	.05413
100	500	83.33	89	.178	.0092390	.010178	.1016	.0087264	.05548
400	600	240	104	.08667	.0521403	.0543568	.04251	.051221	.01763
400	600	240	105	.0875	.0486074	.0506988	.04303	.047719	.01827
400	600	240	125	.10417	.0103748	.0109416	.05463	.0100418	.03210
400	600	240	126	.105	.0095362	.0100634	.05528	.0092231	.03283

(* For $(m, n) = (21, 500)$, the value $k = 3075$ is not possible.)

The pair $(400, 600)$ was included in Table 2 because, according to Fact 2(d), the ratio $n/m = 3/2$ seemed to come next after $1/1$ and $2/1$ in producing large $r_{\max}$, and so possibly small relative error for *dkwm* as an approximation to *pv*, and $r_{\max}$ was increasing in the range computed for this ratio, $m = 102, 104, ..., 132$. Still, the relative errors of S_{pli} in Table 2 are smaller than for *dkwm*.

It is a question for further research whether the usefulness of S_{pli}, which we found for $m = 20$ or 21 and $n = 500$, extends more generally to cases where m is only moderately large and $m \ll n$.

3.5. **Obstacles to asymptotic expansions.** This is to recall an argument of Hodges [10]. Let

$$Z^+ := Z^+_{m,n} := \sqrt{\frac{mn}{m+n}} \sup_x (F_m - G_n)(x),$$

a one-sided two-sample Smirnov statistic. There is the well-known limit theorem that for any $z > 0$, if $m, n \to \infty$ and $z_{m,n} \to z$, then $\Pr(Z^+_{m,n} \geq z_{m,n}) \to \exp(-2z^2)$. Suppose further that $m/n \to 1$ as $n \to \infty$. Then $\sqrt{mn/(m+n)} \sim \sqrt{n/2}$. A question then is whether there exists a function $g(z)$ such that

$$\Pr\left(Z^+_{m,n} \geq z_{m,n}\right) = \exp\left(-2z^2 - \frac{g(z)}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right)\right). \tag{40}$$

Hodges [10, pp. 475-476,481] shows that for an "average" function $g(z) = \sqrt{2}z$, rather than a $o(1/\sqrt{n})$ error, there is an "oscillatory" term which is only $O(1/\sqrt{n})$. Hodges considers $n = m + 2$ (with our convention that $n \geq m$).

If $m = n$, successive possible values of $F_m - G_n$ differ by $1/n$, and values of $Z^+_{m,n}$ (or our M) by $1/\sqrt{2n}$. Thus for fixed z, which are of interest in finding critical values, $z_{n,n}$ can only converge to z at a $O(1/\sqrt{n})$ rate. Then, rather than (40), one might consider it with $z_{m,n}$ substituted for z on the right. For $n = m + 2$, successive possible values of $F_m - G_n$ typically (although not always) differ by at most $2/(n(n-2))$, and possible values of $Z^+_{m,n}$ by $O(n^{-3/2})$, so $z_{m,n}$ can converge to z at that rate. Then (40) is more plausible and it is of interest that Hodges showed it fails.

Here are numerical examples for $m = n - 1$, so $L_{m,n} = n(n-1)$, and for $D_{m,n}$ rather than $Z^+_{m,n}$. We focus on critical values k and $d = k/(n(n-1))$ at the 0.05 level, having p-values *pv* a little less than 0.05. Let *reler* be the relative error of *dkwm* as an approximation to *pv*. By analogy with (40), let us see how $\sqrt{n} \cdot \mathit{reler}$ behaves.

TABLE 3. Behavior of the relative error of *dkwm* for $m = n - 1$

n	k	pv	$reler$	$\sqrt{n} \cdot reler$	n	k	pv	$reler$	$\sqrt{n} \cdot reler$
40	457	.04968	.3579	2.264	400	15066	.049986	.1379	2.758
100	1850	.049985	.2395	2.395	500	21216	.049983	.08052	1.800
200	5302	.049885	.1627	2.301	600	27889	.049984	.08250	2.021
300	9771	.049995	.1448	2.507					

Here the numbers $\sqrt{n} \cdot \mathit{reler}$ also seem "oscillatory" rather than tending to a constant.

Hodges' argument suggests that the approximation S_{pli}, or any approximation implying an asymptotic expansion, cannot improve on the $O(1/\sqrt{n})$ order of the relative error of the simple asymptotic approximation $\beta(M)$; it may often (but not always, e.g. for $m = n$) give smaller multiples of $1/\sqrt{n}$, but not $o(1/\sqrt{n})$.

Appendix A. Details for $m = n \leq 458$

Here we give details on δ_n as in Theorem 1(e), giving data to show by how much (7) fails when $n \leq 457$.

Recall that for $m = n$, we define $M = k/\sqrt{2n}$. For each $1 \leq n \leq 457$, we define $k_{\max}$ to be the k such that $1 \leq k \leq n$ and $\frac{P_{n,n,M}}{2e^{-2M^2}}$ is the largest. Since (7) fails for $n \leq 457$, when plugging in $k = k_{\max}$, we must have

$$\frac{P_{n,n,M}}{2e^{-2M^2}} > 1.$$

Define

$$\delta_n := \frac{P_{n,n,M}}{2e^{-2M^2}} - 1,$$

where $M = k_{\max}/\sqrt{2n}$. Then for any fixed $n \leq 457$ and $M > 0$,

$$P_{n,n,M} = \Pr(KS_{n,n} \geq M) \leq 2(1+\delta_n)e^{-2M^2}.$$

When n increases, the general trend of δ_n is to decrease, but δ_n is not strictly decreasing, e.g. from $n = 7$ to $n = 8$ (Table 5). For $N \leq 457$, we define

$$\Delta_N = \max\{\delta_n : N \leq n \leq 457\}.$$

Then it is clear that for all $n \geq N$ and $M > 0$,

$$P_{n,n,M} = \Pr(KS_{n,n} \geq M) \leq 2(1+\Delta_N)e^{-2M^2}. \tag{41}$$

In Table 4 we list some pairs (N, Δ_N) for $1 \leq N \leq 455$. The values of δ_n and Δ_N were originally output by Mathematica rounded to 5 decimal places. We added .00001 to the rounded numbers to assure getting upper bounds.

Table 4. Selected Pairs (N, Δ_N)

N	Δ_N	N	Δ_N	N	Δ_N
1	0.35915	75	0.00276	215	0.00045
2	0.23152	80	0.00234	225	0.00041
3	0.13811	85	0.00229	230	0.00039
4	0.08432	90	0.00203	235	0.00036
5	0.08030	95	0.00192	240	0.00034
6	0.06223	100	0.00177	250	0.00032
7	0.04287	105	0.00160	255	0.00028
9	0.04048	110	0.00155	265	0.00028
10	0.03401	115	0.00136	270	0.00026
11	0.02629	120	0.00133	275	0.00024
13	0.02603	125	0.00124	285	0.00023
14	0.02376	130	0.00112	290	0.00020
15	0.02065	135	0.00111	305	0.00018
16	0.01773	140	0.00101	310	0.00016
18	0.01755	145	0.00095	325	0.00015
20	0.01511	150	0.00092	330	0.00013
24	0.01237	155	0.00083	345	0.00012
28	0.00923	160	0.00080	350	0.00011
32	0.00865	165	0.00078	355	0.00010
36	0.00707	170	0.00070	365	0.00009
40	0.00645	175	0.00068	370	0.00008

Continued on next page

N	Δ_N	N	Δ_N	N	Δ_N
44	0.00549	180	0.00066	375	0.00007
48	0.00509	185	0.00060	390	0.00006
52	0.00433	190	0.00058	395	0.00005
56	0.00415	195	0.00056	415	0.00004
60	0.00348	200	0.00052	420	0.00003
65	0.00338	205	0.00048	440	0.00002
70	0.00280	210	0.00048	455	0.00001

For $451 \leq N \leq 458$, values of Δ_N which are more precise than those Mathematica displays (it gives just 5 decimal places) are as follows. In all these cases $k = 35$. For $N = 458$, $k = 36$ would give a still more negative value. Theorem 1(c) shows that no k would give $\Delta_N > 0$ for any $N \geq 458$.

N	Δ_N
451	$5.116 \cdot 10^{-6}$
452	$4.707 \cdot 10^{-6}$
453	$4.156 \cdot 10^{-6}$
454	$3.462 \cdot 10^{-6}$
455	$2.627 \cdot 10^{-6}$
456	$1.649 \cdot 10^{-6}$
457	$5.309 \cdot 10^{-7}$
458	$-7.284 \cdot 10^{-7}$

Recall that for $n \geq 458$, we have $\delta_n \leq 0$. As stated in Theorem 1(e) we have that for $12 \leq n \leq 457$,

$$\delta_n < -\frac{0.07}{n} + \frac{40}{n^2} - \frac{400}{n^3}. \tag{42}$$

(More precisely, (42) should be read as: the Mathematica output δ_n plus 0.00001 is smaller than the right hand side of (42) when $11 < n < 458$.) The formula was found by regression and experimentation. In Table 5, we provide the values of δ_n when $1 \leq n \leq 11$.

TABLE 5. δ_n for $n \leq 11$

n	δ_n[1]	n	δ_n[1]
1	0.35914	7	0.04286
2	0.23151	8	0.04434
3	0.1381	9	0.04047
4	0.08431	10	0.034
5	0.08029	11	0.02628
6	0.06222		

[1]The data shown in Table 5 are the Mathematica output without adding 0.00001.

Appendix B. Tables for $m < n$

Tables 6 and 7 were both found by a program which performed the following searches. For each m with $1 \leq m < 200$, there was a search over all n with $m < n \leq 200$. For each such n, $r_{\max}(m, n)$ was found, itself requiring a search over the possible values of $d = k/L_{m,n}$ for $k = 1, 2, \ldots, L_{m,n}$, recalling that $L_{m,n}$ is the least common multiple of m and n. Then for given m, there is a largest such $r_{\max}$, called *rmaxx* in the two tables, attained at $n = n_{\max}$ and for that n, at $d = \text{dmaxx} = k_{\max}/L_{m,n_{\max}}$, and with a p-value "pvatmax" in the numerator of *rmaxx*. There are columns in the tables for each of these.

By separate calculations, we had found that for $m = 1$, $n_{\max} = 2$ with an $r_{\max} > 1$, and for $m = 2$, $n_{\max} = 3$, also with $r_{\max} > 1$ (the values of $r_{\max}$ were given in Subsection 3.3). We are concerned here with cases where $r_{\max} < 1$, so we begin with $m = 3$.

For $3 \leq m < n \leq 200$ and each possible value d of $D_{m,n}$ in the range (34) where the p-value *pvi* by the inside method was found to be less than 10^{-8}, we evaluated instead the upper bound $pv_{ub}(m, n, d)$ as in Theorem 18(a) and took the ratio

$$r_{ub}(m, n, d) = pv_{ub}(m, n, d)/\left(2\exp(-2M^2)\right) \tag{43}$$

where as usual $M = \sqrt{mn/(m + n)}d$. We took the maximum of these for the possible values of d and the ratio of that maximum to $r_{\max}(m, n)$ as evaluated for all other possible values of d. Then we took in turn the maximum of all such ratios for fixed m, over n with $m < n \leq 200$, giving *mrmr* ("maximum ratio of maximum ratios") in the last column of the tables. As all these are less than 1 (in fact, less than 0.63 in Table 6, less than 0.75 in Table 7), we confirm that $r_{\max}(m, n)$ is not attained in the range (34) for $3 \leq m < n \leq 200$ and so the given values of $n_{\max}$ and *rmaxx* are confirmed.

First, we give Table 6 for $3 \leq m \leq 100$ and $m < n \leq 200$. In this range $n = 2m$ is among the possible values of n and in all these cases $n_{\max} = 2m$. For $m \leq 43$, no values of $pvi < 10^{-8}$ were computed for any $n \leq 200$ and $d \leq 1/2$, so the ranges (34) were empty, $d_0(m, 200)$ has an artificial initialized value of 2.0, and *mrmr* is set equal to 0.

Table 6. $3 \leq m \leq 100$, $m < n \leq 200$

m	$n_{\max}$	*rmaxx*	$k_{\max}$	pvatmax	*dmaxx*	$d_0(m, 200)$	*mrmr*
3	6	0.986116	4	0.678571	0.666667	2.000000	0
4	8	0.973325	4	0.513131	0.5	2.000000	0
5	10	0.951143	4	0.654679	0.4	2.000000	0
6	12	0.938437	5	0.468003	0.416667	2.000000	0
7	14	0.947585	6	0.341305	0.428571	2.000000	0
8	16	0.950533	6	0.424185	0.375	2.000000	0
9	18	0.949182	6	0.500403	0.333333	2.000000	0
10	20	0.944748	6	0.569105	0.3	2.000000	0
11	22	0.946271	7	0.42873	0.318182	2.000000	0
12	24	0.946955	8	0.320096	0.333333	2.000000	0
13	26	0.949675	8	0.368058	0.307692	2.000000	0
14	28	0.950815	8	0.414328	0.285714	2.000000	0
15	30	0.950668	8	0.458559	0.266667	2.000000	0
16	32	0.950333	9	0.351588	0.28125	2.000000	0
17	34	0.951642	9	0.388814	0.264706	2.000000	0
18	36	0.952087	9	0.424878	0.25	2.000000	0
19	38	0.9527	10	0.32966	0.263158	2.000000	0
20	40	0.953956	10	0.360358	0.25	2.000000	0
21	42	0.954631	10	0.390399	0.238095	2.000000	0

Continued on next page

m	$n_{\max}$	$rmaxx$	$k_{\max}$	pvatmax	$dmaxx$	$d_0(m,200)$	$mrmr$
22	44	0.954788	10	0.419677	0.227273	2.000000	0
23	46	0.95505	11	0.330725	0.23913	2.000000	0
24	48	0.955966	11	0.356137	0.229167	2.000000	0
25	50	0.956499	11	0.381112	0.22	2.000000	0
26	52	0.956683	11	0.405588	0.211538	2.000000	0
27	54	0.957278	12	0.323585	0.222222	2.000000	0
28	56	0.958022	12	0.345065	0.214286	2.000000	0
29	58	0.958501	12	0.366261	0.206897	2.000000	0
30	60	0.958735	12	0.387131	0.2	2.000000	0
31	62	0.958918	13	0.311609	0.209677	2.000000	0
32	64	0.959602	13	0.330051	0.203125	2.000000	0
33	66	0.960091	13	0.348314	0.19697	2.000000	0
34	68	0.960399	13	0.366366	0.191176	2.000000	0
35	70	0.960536	13	0.384182	0.185714	2.000000	0
36	72	0.961028	14	0.313042	0.194444	2.000000	0
37	74	0.961533	14	0.328951	0.189189	2.000000	0
38	76	0.9619	14	0.344729	0.184211	2.000000	0
39	78	0.962136	14	0.360355	0.179487	2.000000	0
40	80	0.962249	14	0.375811	0.175	2.000000	0
41	82	0.962708	15	0.309089	0.182927	2.000000	0
42	84	0.963123	15	0.322988	0.178571	2.000000	0
43	86	0.963437	15	0.336793	0.174419	2.000000	0
44	88	0.963654	15	0.350491	0.170455	0.498182	0.348133
45	90	0.963776	15	0.364068	0.166667	0.494444	0.349028
46	92	0.964152	16	0.301667	0.173913	0.489565	0.379506
47	94	0.964521	16	0.313932	0.170213	0.485957	0.387321
48	96	0.964812	16	0.326132	0.166667	0.482500	0.395004
49	98	0.965027	16	0.338257	0.163265	0.477653	0.402108
50	100	0.965171	16	0.350299	0.16	0.480000	0.411874
51	102	0.965387	17	0.29201	0.166667	0.470588	0.414091
52	104	0.965731	17	0.30292	0.163462	0.467692	0.421264
53	106	0.966015	17	0.313788	0.160377	0.464151	0.423325
54	108	0.966239	17	0.324605	0.157407	0.460926	0.439284
55	110	0.966407	17	0.335364	0.154545	0.458182	0.440581
56	112	0.966519	17	0.346059	0.151786	0.455000	0.44585
57	114	0.966794	18	0.29073	0.157895	0.451930	0.450127
58	116	0.967076	18	0.300472	0.155172	0.448793	0.46219
59	118	0.967311	18	0.310182	0.152542	0.445847	0.462783
60	120	0.9675	18	0.319853	0.15	0.445000	0.466195
61	122	0.967645	18	0.329482	0.147541	0.440574	0.481306
62	124	0.967746	18	0.339061	0.145161	0.438226	0.481427
63	126	0.968	19	0.286669	0.150794	0.435476	0.484516
64	128	0.968245	19	0.295428	0.148438	0.433750	0.49833
65	130	0.968453	19	0.304163	0.146154	0.430769	0.498012
66	132	0.968624	19	0.312871	0.143939	0.427879	0.499076
67	134	0.96876	19	0.321547	0.141791	0.426269	0.503188

Continued on next page

m	$n_{\max}$	$rmaxx$	$k_{\max}$	pvatmax	$dmaxx$	$d_0(m, 200)$	$mrmr$
68	136	0.968862	19	0.330188	0.139706	0.424118	0.499238
69	138	0.969058	20	0.280649	0.144928	0.421232	0.511505
70	140	0.96928	20	0.28857	0.142857	0.420000	0.506893
71	142	0.969473	20	0.296476	0.140845	0.417042	0.519134
72	144	0.969636	20	0.304361	0.138889	0.415556	0.514055
73	146	0.96977	20	0.312224	0.136986	0.413288	0.526253
74	148	0.969876	20	0.320062	0.135135	0.411351	0.520679
75	150	0.969993	21	0.273263	0.14	0.410000	0.532889
76	152	0.970201	21	0.280462	0.138158	0.407632	0.546669
77	154	0.970385	21	0.287651	0.136364	0.405390	0.53993
78	156	0.970544	21	0.294827	0.134615	0.403718	0.552279
79	158	0.970681	21	0.301987	0.132911	0.401899	0.545217
80	160	0.970794	21	0.30913	0.13125	0.402500	0.557531
81	162	0.970884	21	0.316252	0.12963	0.398457	0.550117
82	164	0.971022	22	0.271515	0.134146	0.396829	0.562415
83	166	0.971201	22	0.278079	0.13253	0.395241	0.574491
84	168	0.97136	22	0.284636	0.130952	0.393810	0.566885
85	170	0.9715	22	0.291182	0.129412	0.392353	0.578607
86	172	0.97162	22	0.297717	0.127907	0.390814	0.571077
87	174	0.971721	22	0.304238	0.126437	0.388793	0.582468
88	176	0.971804	22	0.310744	0.125	0.387727	0.575016
89	178	0.971931	23	0.268046	0.129213	0.385562	0.586057
90	180	0.972091	23	0.274057	0.127778	0.385000	0.59676
91	182	0.972234	23	0.280063	0.126374	0.383297	0.589333
92	184	0.972361	23	0.286062	0.125	0.381957	0.599754
93	186	0.972472	23	0.292052	0.123656	0.380484	0.592408
94	188	0.972567	23	0.298032	0.12234	0.379149	0.602564
95	190	0.972647	23	0.304	0.121053	0.377895	0.612445
96	192	0.972743	24	0.263293	0.125	0.376667	0.605186
97	194	0.97289	24	0.268818	0.123711	0.375155	0.614769
98	196	0.973022	24	0.274341	0.122449	0.373980	0.624092
99	198	0.973142	24	0.279858	0.121212	0.372323	0.616907
100	200	0.973248	24	0.28537	0.12	0.375	0.626016

Table 7 covers $101 \le m < n \le 200$. In Tables 6 and 8 the ratio n/m is always 2, in Table 6 because $n_{\max} = 2m$ from the computer search, and in Table 8 by our choice. In the range $101 \le m < n \le 200$ of Table 7, $n_{\max}/m = 2$ is not possible, but $3/2$ is and occurs whenever possible as described in Fact 2(d). Other simple ratios are possible and may or may not occur. For example, in some cases where either $4/3$ or $5/3$ would be possible, $5/3$ occurs. When $m = 175$, $n_{\max} = 176$, even though $n = 200$ would have given a simpler ratio $n/m = 8/7$; but $r_{\max}(175, 200) = 0.927656 < 0.928771 = r_{\max}(175, 176)$. Ratios occur of $n_{\max}/m = 9/7 = 198/154$, $10/7 = 190/133$, and $11/7 = 187/119$.

For given m, $mrmr$ often, but not always, occurs when $n = n_{\max}$. For example, it does when $m = 132$ and for $195 \le m \le 199$, but not for $m = 168$, for which $n_{\max} = 196$ but $mrmr$ occurs for $n = 169$.

Table 7. $101 \leq m < n \leq 200$

m	n	$rmaxx$	$k_{\max}$	pvatmax	$dmaxx$	$d_0(m, 200)$	$mrmr$
101	200	0.913382	2134	0.408438	0.105644	0.369109	0.589514
102	153	0.943929	36	0.346915	0.117647	0.369314	0.57433
103	155	0.913333	1764	0.403162	0.110492	0.367913	0.56871
104	156	0.944382	36	0.358576	0.115385	0.367308	0.57608
105	175	0.93144	58	0.375393	0.110476	0.365952	0.585619
106	159	0.944769	37	0.337672	0.116352	0.364811	0.593385
107	161	0.914677	1886	0.391834	0.109479	0.363598	0.575931
108	162	0.945233	37	0.348785	0.114198	0.362593	0.586098
109	164	0.915258	1921	0.403431	0.107463	0.361468	0.592792
110	165	0.94563	37	0.35982	0.112121	0.360909	0.601528
111	185	0.932974	60	0.36867	0.108108	0.359685	0.593304
112	168	0.946023	38	0.339124	0.113095	0.358571	0.599187
113	170	0.916523	2048	0.391779	0.106611	0.357257	0.598983
114	171	0.946435	38	0.34966	0.111111	0.356667	0.607297
115	184	0.924245	96	0.395831	0.104348	0.355652	0.615512
116	174	0.946787	38	0.360125	0.109195	0.354655	0.602189
117	195	0.934419	61	0.381039	0.104274	0.353632	0.604576
118	177	0.947179	39	0.339676	0.110169	0.352627	0.612539
119	187	0.92098	134	0.40119	0.102368	0.351765	0.620418
120	180	0.947549	39	0.349682	0.108333	0.351667	0.62813
121	182	0.918795	2314	0.369177	0.105077	0.349959	0.609661
122	183	0.94787	40	0.329881	0.10929	0.34918	0.618786
123	164	0.935287	53	0.366045	0.107724	0.348049	0.624895
124	186	0.948254	40	0.339454	0.107527	0.347581	0.632313
125	200	0.926795	101	0.385868	0.101	0.348	0.639416
126	189	0.94859	40	0.348975	0.10582	0.345873	0.628722
127	191	0.92039	2493	0.367425	0.102774	0.344961	0.62876
128	192	0.948905	41	0.329447	0.106771	0.344375	0.635817
129	172	0.936549	54	0.372684	0.104651	0.34345	0.642654
130	195	0.949257	41	0.338568	0.105128	0.343077	0.649389
131	197	0.921385	2571	0.38654	0.099624	0.341794	0.634883
132	198	0.949565	41	0.347641	0.103535	0.341212	0.642219
133	190	0.923341	132	0.395393	0.099248	0.339474	0.645595
134	135	0.920683	2045	0.330121	0.113046	0.339328	0.652115
135	180	0.937714	56	0.356856	0.103704	0.338889	0.65852
136	170	0.930667	70	0.375306	0.102941	0.338235	0.65651
137	138	0.921316	2091	0.342759	0.1106	0.337372	0.648274
138	184	0.93829	56	0.370191	0.101449	0.336739	0.654592
139	140	0.921695	2121	0.351497	0.108993	0.335935	0.660788
140	175	0.931495	71	0.376012	0.101429	0.335714	0.666764
141	188	0.938842	57	0.362092	0.101064	0.33461	0.672611
142	143	0.922434	2310	0.291798	0.113759	0.333873	0.65669
143	144	0.922679	2326	0.295749	0.112956	0.333322	0.662625
144	192	0.939363	58	0.354142	0.100694	0.332778	0.668386
145	174	0.92777	87	0.381501	0.1	0.332069	0.674069

Continued on next page

m	n	$rmaxx$	$k_{\max}$	pvatmax	$dmaxx$	$d_0(m, 200)$	$mrmr$
146	147	0.92338	2375	0.307108	0.110661	0.331301	0.679665
147	196	0.939886	58	0.366614	0.098639	0.330714	0.665351
148	185	0.933056	73	0.376649	0.098649	0.330135	0.669828
149	150	0.924015	2423	0.318878	0.108412	0.329463	0.675366
150	200	0.940395	59	0.358464	0.098333	0.33	0.680802
151	152	0.9244	2455	0.326689	0.106962	0.328212	0.686084
152	190	0.933791	74	0.376618	0.097368	0.327895	0.691283
153	154	0.924759	2488	0.334009	0.105594	0.327059	0.676529
154	198	0.926355	132	0.384897	0.095238	0.326429	0.681813
155	186	0.929738	90	0.381638	0.096774	0.325968	0.686953
156	195	0.934499	75	0.376378	0.096154	0.325385	0.692017
157	158	0.925501	2711	0.282122	0.109288	0.324713	0.696965
158	159	0.925721	2728	0.285641	0.10859	0.324051	0.692976
159	160	0.925934	2745	0.289158	0.107901	0.323491	0.687533
160	200	0.935183	76	0.375946	0.095	0.32375	0.692448
161	162	0.926347	2780	0.29579	0.106587	0.32236	0.697268
162	189	0.927765	108	0.38127	0.095238	0.321975	0.701955
163	164	0.92674	2814	0.302799	0.105267	0.321258	0.706579
164	165	0.926928	2831	0.306296	0.104619	0.320854	0.692784
165	198	0.931538	93	0.380526	0.093939	0.320455	0.697485
166	167	0.927286	2865	0.313277	0.103348	0.319759	0.702062
167	168	0.927455	2882	0.316759	0.102723	0.319461	0.7066
168	196	0.928852	110	0.381517	0.093537	0.319048	0.711063
169	170	0.927778	2917	0.323319	0.101532	0.318254	0.715408
170	171	0.927934	2934	0.326785	0.100929	0.317941	0.702107
171	172	0.928084	2951	0.330246	0.100333	0.317368	0.706563
172	173	0.928229	2968	0.333699	0.099745	0.31686	0.710925
173	174	0.928384	3160	0.274412	0.104976	0.316329	0.715164
174	175	0.92858	3178	0.277564	0.104368	0.315805	0.719326
175	176	0.928771	3196	0.280715	0.103766	0.315714	0.723426
176	177	0.928956	3214	0.283863	0.103172	0.315	0.71059
177	178	0.929141	3233	0.286679	0.102615	0.31435	0.714702
178	179	0.929321	3251	0.289823	0.102034	0.314101	0.718797
179	180	0.929496	3269	0.292965	0.101459	0.31338	0.72281
180	181	0.929666	3287	0.296104	0.10089	0.313333	0.726721
181	182	0.929831	3305	0.299239	0.100328	0.312514	0.730593
182	183	0.929992	3323	0.302371	0.099772	0.312253	0.718232
183	184	0.930148	3341	0.3055	0.099222	0.311694	0.722166
184	185	0.930299	3359	0.308624	0.098678	0.311522	0.726004
185	186	0.930446	3378	0.311415	0.098169	0.310946	0.729824
186	187	0.930591	3396	0.314533	0.097637	0.31043	0.733571
187	188	0.930732	3414	0.317646	0.09711	0.31	0.737226
188	189	0.930867	3432	0.320755	0.096589	0.309574	0.727759
189	190	0.930999	3450	0.323859	0.096074	0.309074	0.729033
190	191	0.931125	3468	0.326959	0.095564	0.308947	0.732715
191	192	0.931267	3679	0.271066	0.100322	0.308298	0.736292

Continued on next page

m	n	$rmaxx$	$k_{\max}$	pvatmax	$dmaxx$	$d_0(m, 200)$	$mrmr$
192	193	0.931438	3699	0.27362	0.099822	0.308125	0.739829
193	194	0.931607	3718	0.276457	0.0993	0.307513	0.743299
194	195	0.931772	3737	0.279293	0.098784	0.307062	0.744568
195	196	0.931932	3756	0.282127	0.098273	0.306795	0.735198
196	197	0.932089	3775	0.284959	0.097768	0.306531	0.738695
197	198	0.932242	3794	0.287789	0.097267	0.305533	0.742108
198	199	0.932391	3813	0.290616	0.096772	0.306162	0.74544
199	200	0.932536	3832	0.293442	0.096281	0.303719	0.748761

The following Table 8 treats $100 \le m \le 300$ and $n = 2m$. So it relates to Fact 3, for $m \ge 100$ in parts (a), (b), and (d), and the first statement in part (c) for $n = 2m$.

For each such m and $n = 2m$, $r_{\max}(m, n)$ was computed. It has a numerator p-value "pvatmax" attained at $d_{\max} = k_{\max}/n$.

Throughout the table, $r_{\max}$ continues to increase, as it does in Table 6 for $m \ge 16$, and as stated in Fact 3(a).

In the last column, $rbd_{\max}$ is the maximum of $r_{ub}(m, 2m, d)$ as defined in (43) for d in the range (34). These $rbd_{\max}$ tend to increase with m, although not monotonically. All values shown are less than 0.832, which is less than $r_{\max}$ for all the values of m shown. This confirms the values of $r_{\max}$.

Table 8 was relatively easy to compute, as there is only one n for each m, and then there are only n possible values of d. The computations for Table 6 and visibly in Table 7 were much more time-consuming because of the many relatively prime pairs (m, n) giving mn values of d. Only a very small fraction (a few percent) of such pairs requiring computation for Table 7 are shown in it because they happened to result in an $rmaxx$.

TABLE 8. $100 \le m \le 300$, $n = 2m$

m	n	$r_{\max}$	$k_{\max}$	pvatmax	$d_{\max}$	$d_0(m, 2m)$	$rbd_{\max}$
100	200	0.973248	24	0.28537	0.12	0.375	0.609268
101	202	0.973341	24	0.290874	0.118812	0.376238	0.602412
102	204	0.973421	24	0.296371	0.117647	0.372549	0.61113
103	206	0.973488	24	0.301857	0.116505	0.368932	0.619616
104	208	0.973611	25	0.262685	0.120192	0.370192	0.612852
105	210	0.973737	25	0.267779	0.119048	0.366667	0.621146
106	212	0.973852	25	0.27287	0.117925	0.367925	0.614442
107	214	0.973955	25	0.277958	0.116822	0.364486	0.622553
108	216	0.974047	25	0.283042	0.115741	0.361111	0.630454
109	218	0.974129	25	0.28812	0.114679	0.362385	0.623845
110	220	0.974199	25	0.293191	0.113636	0.359091	0.631581
111	222	0.974264	26	0.255903	0.117117	0.355856	0.639119
112	224	0.974386	26	0.260616	0.116071	0.357143	0.632607
113	226	0.974498	26	0.265329	0.115044	0.353982	0.639994
114	228	0.9746	26	0.270039	0.114035	0.350877	0.647196
115	230	0.974692	26	0.274746	0.113043	0.352174	0.640783
116	232	0.974776	26	0.279451	0.112069	0.349138	0.647847
117	234	0.97485	26	0.284151	0.111111	0.346154	0.654737
118	236	0.974915	26	0.288846	0.110169	0.347458	0.648422
119	238	0.974975	27	0.253039	0.113445	0.344538	0.655186
120	240	0.975085	27	0.25741	0.1125	0.345833	0.648926

Continued on next page

m	n	$r_{\max}$	$k_{\max}$	pvatmax	$d_{\max}$	$d_0(m, 2m)$	$rbd_{\max}$
121	242	0.975187	27	0.261782	0.11157	0.342975	0.655569
122	244	0.975281	27	0.266152	0.110656	0.340164	0.662054
123	246	0.975366	27	0.270521	0.109756	0.341463	0.65589
124	248	0.975444	27	0.274888	0.108871	0.33871	0.662264
125	250	0.975514	27	0.279251	0.108	0.336	0.668489
126	252	0.975576	27	0.283611	0.107143	0.337302	0.662421
127	254	0.97563	27	0.287967	0.106299	0.334646	0.668544
128	256	0.975721	28	0.253321	0.109375	0.332031	0.674525
129	258	0.975816	28	0.257387	0.108527	0.333333	0.668552
130	260	0.975904	28	0.261453	0.107692	0.330769	0.674441
131	262	0.975985	28	0.265518	0.10687	0.328244	0.680195
132	264	0.976059	28	0.269582	0.106061	0.329545	0.674315
133	266	0.976126	28	0.273644	0.105263	0.327068	0.679984
134	268	0.976187	28	0.277703	0.104478	0.324627	0.685527
135	270	0.976241	28	0.28176	0.103704	0.325926	0.679737
136	272	0.976302	29	0.24855	0.106618	0.323529	0.685202
137	274	0.976392	29	0.252341	0.105839	0.321168	0.690545
138	276	0.976476	29	0.256133	0.105072	0.322464	0.684845
139	278	0.976553	29	0.259924	0.104317	0.320144	0.690117
140	280	0.976625	29	0.263715	0.103571	0.317857	0.695273
141	282	0.976691	29	0.267505	0.102837	0.319149	0.689661
142	284	0.976752	29	0.271294	0.102113	0.316901	0.694752
143	286	0.976806	29	0.27508	0.101399	0.314685	0.699733
144	288	0.976855	29	0.278865	0.100694	0.315972	0.694206
145	290	0.976921	30	0.246802	0.103448	0.313793	0.699127
146	292	0.977002	30	0.250345	0.10274	0.311644	0.703944
147	294	0.977077	30	0.253889	0.102041	0.309524	0.70866
148	296	0.977148	30	0.257433	0.101351	0.310811	0.70326
149	298	0.977213	30	0.260976	0.100671	0.308725	0.707922
150	300	0.977274	30	0.264519	0.1	0.306667	0.712488
151	302	0.97733	30	0.268061	0.099338	0.307947	0.707169
152	304	0.97738	30	0.271602	0.098684	0.305921	0.711685
153	306	0.977426	30	0.275142	0.098039	0.303922	0.716109
154	308	0.977485	31	0.244214	0.100649	0.305195	0.710868
155	310	0.97756	31	0.247532	0.1	0.303226	0.715246
156	312	0.97763	31	0.250851	0.099359	0.301282	0.719536
157	314	0.977695	31	0.254171	0.098726	0.302548	0.714371
158	316	0.977756	31	0.25749	0.098101	0.300633	0.71862
159	318	0.977813	31	0.26081	0.097484	0.298742	0.722784
160	320	0.977865	31	0.264129	0.096875	0.296875	0.726865
161	322	0.977914	31	0.267447	0.096273	0.298137	0.721817
162	324	0.977958	31	0.270764	0.095679	0.296296	0.725862
163	326	0.978004	32	0.240951	0.09816	0.294479	0.729828
164	328	0.978074	32	0.244064	0.097561	0.295732	0.72485
165	330	0.978139	32	0.247179	0.09697	0.293939	0.728782
166	332	0.978201	32	0.250294	0.096386	0.292169	0.732638

Continued on next page

m	n	$r_{\max}$	$k_{\max}$	pvatmax	$d_{\max}$	$d_0(m,2m)$	$rbd_{\max}$
167	334	0.978259	32	0.25341	0.095808	0.293413	0.72773
168	336	0.978313	32	0.256526	0.095238	0.291667	0.731554
169	338	0.978364	32	0.259642	0.094675	0.289941	0.735306
170	340	0.97841	32	0.262758	0.094118	0.288235	0.738988
171	342	0.978453	32	0.265873	0.093567	0.289474	0.734187
172	344	0.978492	32	0.268987	0.093023	0.287791	0.73784
173	346	0.978549	33	0.240075	0.095376	0.286127	0.741426
174	348	0.978611	33	0.243003	0.094828	0.287356	0.736689
175	350	0.97867	33	0.245932	0.094286	0.285714	0.740249
176	352	0.978726	33	0.248862	0.09375	0.284091	0.743743
177	354	0.978778	33	0.251792	0.09322	0.282486	0.747174
178	356	0.978827	33	0.254723	0.092697	0.283708	0.742539
179	358	0.978873	33	0.257654	0.092179	0.282123	0.745947
180	360	0.978915	33	0.260584	0.091667	0.280556	0.749294
181	362	0.978955	33	0.263514	0.09116	0.281768	0.744718
182	364	0.978991	33	0.266444	0.090659	0.28022	0.748044
183	366	0.979048	34	0.238431	0.092896	0.278689	0.75131
184	368	0.979105	34	0.24119	0.092391	0.277174	0.754519
185	370	0.979159	34	0.243949	0.091892	0.278378	0.750039
186	372	0.979211	34	0.246709	0.091398	0.276882	0.753229
187	374	0.979259	34	0.24947	0.090909	0.275401	0.756364
188	376	0.979304	34	0.252231	0.090426	0.276596	0.751939
189	378	0.979347	34	0.254992	0.089947	0.275132	0.755056
190	380	0.979386	34	0.257753	0.089474	0.273684	0.75812
191	382	0.979423	34	0.260515	0.089005	0.272251	0.761131
192	384	0.979457	34	0.263276	0.088542	0.273438	0.756797
193	386	0.97951	35	0.236154	0.090674	0.272021	0.759793
194	388	0.979563	35	0.238756	0.090206	0.270619	0.762738
195	390	0.979613	35	0.24136	0.089744	0.271795	0.758455
196	392	0.979661	35	0.243964	0.089286	0.270408	0.761387
197	394	0.979706	35	0.246569	0.088832	0.269036	0.764269
198	396	0.979749	35	0.249175	0.088384	0.267677	0.767103
199	398	0.979789	35	0.251781	0.08794	0.268844	0.762905
200	400	0.979827	35	0.254387	0.0875	0.2675	0.765728
201	402	0.979862	35	0.256993	0.087065	0.266169	0.768503
202	404	0.979894	35	0.259599	0.086634	0.264851	0.771234
203	406	0.979938	36	0.233354	0.08867	0.26601	0.767118
204	408	0.979988	36	0.235813	0.088235	0.264706	0.769838
205	410	0.980036	36	0.238273	0.087805	0.263415	0.772514
206	412	0.980081	36	0.240735	0.087379	0.262136	0.775146
207	414	0.980124	36	0.243196	0.086957	0.263285	0.771109
208	416	0.980165	36	0.245659	0.086538	0.262019	0.773733
209	418	0.980203	36	0.248122	0.086124	0.260766	0.776315
210	420	0.980239	36	0.250586	0.085714	0.261905	0.772321
211	422	0.980273	36	0.25305	0.085308	0.260664	0.774895
212	424	0.980305	36	0.255513	0.084906	0.259434	0.777428

Continued on next page

m	n	$r_{\max}$	$k_{\max}$	pvatmax	$d_{\max}$	$d_0(m, 2m)$	$rbd_{\max}$
213	426	0.980337	37	0.230127	0.086854	0.258216	0.779921
214	428	0.980384	37	0.232454	0.086449	0.259346	0.776002
215	430	0.98043	37	0.234782	0.086047	0.25814	0.778488
216	432	0.980473	37	0.237111	0.085648	0.256944	0.780936
217	434	0.980514	37	0.239441	0.085253	0.25576	0.783346
218	436	0.980553	37	0.241771	0.084862	0.256881	0.779499
219	438	0.980591	37	0.244103	0.084475	0.255708	0.781903
220	440	0.980626	37	0.246434	0.084091	0.254545	0.78427
221	442	0.980659	37	0.248767	0.08371	0.253394	0.786602
222	444	0.980691	37	0.251099	0.083333	0.254505	0.782823
223	446	0.98072	37	0.253432	0.08296	0.253363	0.785149
224	448	0.980754	38	0.228757	0.084821	0.252232	0.787441
225	450	0.980798	38	0.230962	0.084444	0.251111	0.789698
226	452	0.98084	38	0.233169	0.084071	0.252212	0.785987
227	454	0.98088	38	0.235376	0.0837	0.251101	0.78824
228	456	0.980918	38	0.237585	0.083333	0.25	0.790459
229	458	0.980954	38	0.239794	0.082969	0.248908	0.792647
230	460	0.980989	38	0.242004	0.082609	0.25	0.788999
231	462	0.981022	38	0.244215	0.082251	0.248918	0.791183
232	464	0.981053	38	0.246426	0.081897	0.247845	0.793335
233	466	0.981083	38	0.248637	0.081545	0.246781	0.795455
234	468	0.981111	38	0.250849	0.081197	0.247863	0.79187
235	470	0.981142	39	0.226879	0.082979	0.246809	0.793988
236	472	0.981183	39	0.228972	0.082627	0.245763	0.796076
237	474	0.981222	39	0.231066	0.082278	0.244726	0.798134
238	476	0.98126	39	0.233162	0.081933	0.245798	0.794609
239	478	0.981296	39	0.235258	0.08159	0.24477	0.796664
240	480	0.98133	39	0.237355	0.08125	0.24375	0.798691
241	482	0.981363	39	0.239452	0.080913	0.242739	0.800689
242	484	0.981394	39	0.241551	0.080579	0.243802	0.797222
243	486	0.981424	39	0.24365	0.080247	0.242798	0.799219
244	488	0.981452	39	0.245749	0.079918	0.241803	0.801188
245	490	0.981478	39	0.247849	0.079592	0.240816	0.803129
246	492	0.981505	40	0.224576	0.081301	0.24187	0.799718
247	494	0.981543	40	0.226564	0.080972	0.240891	0.801659
248	496	0.98158	40	0.228554	0.080645	0.239919	0.803573
249	498	0.981616	40	0.230545	0.080321	0.238956	0.80546
250	500	0.98165	40	0.232537	0.08	0.24	0.802104
251	502	0.981683	40	0.234529	0.079681	0.239044	0.803991
252	504	0.981714	40	0.236523	0.079365	0.238095	0.805853
253	506	0.981744	40	0.238517	0.079051	0.237154	0.807689
254	508	0.981772	40	0.240512	0.07874	0.238189	0.804385
255	510	0.9818	40	0.242507	0.078431	0.237255	0.806221
256	512	0.981825	40	0.244503	0.078125	0.236328	0.808034
257	514	0.98185	40	0.246499	0.077821	0.235409	0.809822
258	516	0.981881	41	0.223807	0.079457	0.234496	0.811586

Continued on next page

m	n	$r_{\max}$	$k_{\max}$	pvatmax	$d_{\max}$	$d_0(m,2m)$	$rbd_{\max}$
259	518	0.981916	41	0.225699	0.079151	0.235521	0.808356
260	520	0.98195	41	0.227593	0.078846	0.234615	0.810121
261	522	0.981983	41	0.229488	0.078544	0.233716	0.811863
262	524	0.982014	41	0.231383	0.078244	0.232824	0.813581
263	526	0.982045	41	0.23328	0.077947	0.23384	0.8104
264	528	0.982074	41	0.235177	0.077652	0.232955	0.81212
265	530	0.982101	41	0.237075	0.077358	0.232075	0.813817
266	532	0.982128	41	0.238973	0.077068	0.231203	0.815493
267	534	0.982153	41	0.240872	0.076779	0.23221	0.812359
268	536	0.982177	41	0.242772	0.076493	0.231343	0.814035
269	538	0.982199	41	0.244672	0.076208	0.230483	0.81569
270	540	0.982232	42	0.22256	0.077778	0.22963	0.817325
271	542	0.982265	42	0.224363	0.077491	0.230627	0.814236
272	544	0.982296	42	0.226167	0.077206	0.229779	0.815871
273	546	0.982327	42	0.227973	0.076923	0.228938	0.817486
274	548	0.982356	42	0.229779	0.076642	0.228102	0.819081
275	550	0.982385	42	0.231585	0.076364	0.227273	0.820656
276	552	0.982412	42	0.233393	0.076087	0.228261	0.817633
277	554	0.982438	42	0.235201	0.075812	0.227437	0.819209
278	556	0.982462	42	0.23701	0.07554	0.226619	0.820766
279	558	0.982486	42	0.238819	0.075269	0.225806	0.822304
280	560	0.982509	42	0.240629	0.075	0.226786	0.819323
281	562	0.98253	42	0.242439	0.074733	0.225979	0.820863
282	564	0.982561	43	0.220904	0.076241	0.225177	0.822384
283	566	0.982592	43	0.222624	0.075972	0.224382	0.823886
284	568	0.982621	43	0.224345	0.075704	0.223592	0.825371
285	570	0.98265	43	0.226066	0.075439	0.224561	0.82245
286	572	0.982678	43	0.227788	0.075175	0.223776	0.823937
287	574	0.982705	43	0.229511	0.074913	0.222997	0.825405
288	576	0.98273	43	0.231235	0.074653	0.222222	0.826857
289	578	0.982755	43	0.23296	0.074394	0.223183	0.823975
290	580	0.982778	43	0.234685	0.074138	0.222414	0.825429
291	582	0.982801	43	0.236411	0.073883	0.221649	0.826865
292	584	0.982823	43	0.238137	0.07363	0.22089	0.828284
293	586	0.982843	43	0.239864	0.073379	0.220137	0.829687
294	588	0.98287	44	0.218899	0.07483	0.221088	0.826863
295	590	0.982899	44	0.220541	0.074576	0.220339	0.828268
296	592	0.982927	44	0.222183	0.074324	0.219595	0.829657
297	594	0.982955	44	0.223826	0.074074	0.218855	0.831029
298	596	0.982981	44	0.22547	0.073826	0.219799	0.828242
299	598	0.983007	44	0.227115	0.073579	0.219064	0.829617
300	600	0.983031	44	0.228761	0.073333	0.218333	0.830976

References

[1] D. André (1887). "Solution directe du problème resolu par M. Bertrand." *Comptes Randus Acad. Sci.* Paris **105**, 436–437.

[2] L. Bachelier (1901). "Théorie mathématique du jeu." *Ann. Sci. Ecole Nat. Sup.*, 3^e sér., **18**, 143-209.

[3] L. Bachelier (1912)*. *Calcul des probabilités* , **1**. Gauthier–Villars, Paris.

[4] L. Bachelier (1939). *Les nouvelles méthodes du calcul des probabilités.* Gauthier–Villars, Paris.

[5] J. Blackman (1956). An extension of the Kolmogorov distribution. *Ann. Math. Statist.* **27**, 513–520. [cf. [6]]

[6] J. Blackman (1958). "Correction to 'An extension of the Kolmogorov distribution.'" *Ann. Math. Statist.* **29**, 318–322.

[7] A. Dvoretzky, J. Kiefer, J. Wolfowitz (1956). "Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator." *Ann. Math. Statist.* **27**, 642–669.

[8] M. Dwass (1967). "Simple random walk and rank order statistics." *Ann. Math. Statist.* **38**, 1042–1053.

[9] B. V. Gnedenko, V. S. Korolyuk (1951). "On the maximum discrepancy between two empirical distributions." *Dokl. Akad. Nauk SSSR* **80**, 525–528 [Russian]; *Sel. Transl. Math. Statist. Probab.* **1** (1961), 13–16.

[10] J. L. Hodges (1957). "The significance probability of the Smirnov two sample test." *Arkiv för Matematik* **3**, 469-486.

[11] P. J. Kim and R. I. Jennrich (1970), "Tables of the exact sampling distribution of the two-sample Kolmogorov–Smirnov criterion, D_{mn}, $m \leq n$, in *Selected Tables in Mathematical Statistics*, Institute of Mathematical Statistics, ed. H. L. Harter and D. B. Owen; Repub. Markham, Chicago, 1973.

[12] P. Massart (1990). "The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality." *Ann. Probability* **18**, 1269–1283.

[13] T. S. Nanjundiah (1959). "Note on Stirling's Formula." *Amer. Math. Monthly.* **66**, 701–703.

[14] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery (1992). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2d. ed., Cambridge University Press.

[15] M. A. Stephens (1970). Use of the Kolmogorov–Smirnov, Cramér–Von Mises and related statistics without extensive tables. *J. Royal Statistical Soc. Ser. B (Methodological)* **32**, 115-122.

[16] F. Wei and R. M. Dudley (2011). "Two-sample Dvoretzky–Kiefer–Wolfowitz Inequalities." Preprint.

* An asterisk indicates items of which we learned from secondary sources but which we have not seen in the original.

(Fan Wei) M.I.T.
E-mail address, Fan Wei: `fan_wei@mit.edu`

(R. M. Dudley) M.I.T. Mathematics Department
E-mail address, R. M. Dudley: `rmd@math.mit.edu`