Biometrika (2015), **xx**, x, *pp*. 1–6

Printed in Great Britain

Supplementary material: On improving Holm's procedure using pairwise dependencies

1. Other distributions satisfying assumption 1

The following are examples of distributions other than those given in Section 4 for the underlying null test statistics or p-values for which Assumption 1 holds.

1. Multivariate Gamma. Let $X_i = Y_0 + Y_i$, $i = 1, \ldots, n$, where Y_i, $i = 0, 1, \ldots, n$, are independent with $Y_0 \sim Gamma(\alpha_0, \beta)$, where $\alpha_0 \geq 1$, and $Y_i \sim Gamma(\alpha, \beta)$ for $i = 1, \ldots, n$.

2. Multivariate F. Let $X_i = Y_i/Y_0$, $i = 1, \ldots, n$, where Y_i, $i = 0, 1, \ldots, n$, are independent with $Y_0 \sim \chi^2_{\nu_0}/\nu_0$ and $Y_i \sim \chi^2_{\nu}/\nu$ for $i = 1, \ldots, n$.

3. Archimedean Copula. Let the distribution of the p-values generated from the test statistics be assumed to be such that the pairwise joint distribution of the null p-values can be modelled by an Archimedean copula. A bivariate copula, which is the joint cumulative distribution function of a pair of random variables on a unit square with unform marginals, is said to be Archimedean if it can be expressed by $C(u, v) = \phi^{-1}\{\phi(u) + \phi(v)\}$, $0 < u, v < 1$, for some convex decreasing function ϕ, called generator, satisfying $\phi(1) = 0$, with the convention $\phi^{-1}(u) = 0$ if $u > \phi(0)$. The following are some well-known systems of bivariate distributions belonging to this class:

(a) Clayton copula: $C_\theta(u, v) = \left\{\max(u^{-\theta} + v^{-\theta} - 1,\, 0)\right\}^{-1/\theta}$, $\theta \in [-1, \infty) \setminus \{0\}$.

(b) Gumbel copula: $C_\theta(u, v) = \exp\left[-\left\{(-\log u)^\theta + (-\log v)^\theta\right\}^{1/\theta}\right]$, $\theta \in [1, \infty)$.

(c) Frank copula: $C_\theta(u, v) = -\log\left[1 + \{\exp(-\theta u) - 1\}\{\exp(-\theta v) - 1\}/\{\exp(-\theta) - 1\}\right]/\theta$, $\theta \in (-\infty, \infty) \setminus \{0\}$.

(d) Joe copula: $C_\theta(u, v) = 1 - \left\{(1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta\right\}^{1/\theta}$, $\theta \in [1, \infty)$.

(e) Ali-Mikhail-Haq copula: $C_\theta(u, v) = uv/\{1 - \theta(1-u)(1-v)\}$, $\theta \in [-1, 1)$.

The following three lemmas prove the desired convexity property for all of the above distributions. While we prove the property for multivariate Gamma and F by establishing it for certain general families of location- and scale-mixture distributions in Lemmas 1 and 2, respectively, we do it individually for each of the Archimedean copulas in Lemma 3.

Proofs of Lemmas 1 and 2 rely on the following important result: Let X be a random variable with the density, $f(x, \theta)$, at x depending on parameter θ. Then, the expectation of an increasing function of X is increasing in θ if $f(x, \theta)$ is totally positive of order two in (x, θ); that is, $f(x, \theta)f(x', \theta') \geq f(x', \theta)f(x, \theta')$ for all $x \leq x', \theta \leq \theta'$ (Karlin, 1968).

LEMMA 1. *Let the random variables $X_1, \ldots, X_n$ be such that, given $Y = y$, they are independent and identically distributed with a common density $\phi_1(x - y)$ and $Y \sim \phi_2(y)$, for some densities ϕ_1 and ϕ_2. Then, Property 1 holds for $X_1, \ldots, X_n$ if $\phi_2(x - y)$ is totally positive of order two in (x, y).*

Proof. We prove the lemma only for the pair (X_1, X_2) without any loss of generality. Let $\Phi_1(x - y) = \text{pr}(X_i \leq x \,|\, Y = y)$, and $f(x) = \int \phi_1(x - y)\phi_2(y)dy = \int \phi_1(y)\phi_2(x - y)dy$ be the common density of X_i, for $i = 1, 2$. Then, we have

$$\begin{aligned}\mathrm{pr}(X_1 \le x \mid X_2 = x) &= \{f(x)\}^{-1} \int \Phi_1(x-y)\phi_1(x-y)\phi_2(y)dy \\ &= \{f(x)\}^{-1} \int \Phi_1(y)\phi_1(y)\phi_2(x-y)dy \\ &= \int \Phi_1(y)\phi^*(y,x)dy,\end{aligned}$$

where $\phi^*(y,x) = \phi_1(y)\phi_2(x-y)/f(x)$. Since the density $\phi^*(y,x)$, with x treated as parameter, is totally positive of order two in (y,x) and $\Phi_1(y)$ is increasing in y, the above integral is increasing in x. This proves the lemma. □

Remark 1. Multivariate Gamma belongs to the family of distributions considered in Lemma 1. It is easy to see that here the density of Y_0 at y, say $\phi_2(y)$, is such that $\phi_2(x-y)$ is totally positive of order two in (x,y). Hence, these X_i's that jointly have a multivariate gamma distribution satisfy Property 1.

LEMMA 2. *Let the positive valued random variables $X_1,\ldots,X_n$ be such that, given $Y=y$, where Y is also positive valued, they are independent and identically distributed with a common density $y\psi_1(yx)$ and $Y \sim \psi_2(y)$, for some densities ψ_1 and ψ_2. Then, Property 1 holds for $X_1,\ldots,X_n$ if $\psi_2(x/y)$ is totally positive of order two in (x,y).*

Proof. As in Lemma 1, we prove the lemma only for the pair (X_1,X_2). Let $\mathrm{pr}(X_i \le x \mid Y = y) = \Psi_1(yx)$, and $f(x) = \int y\psi_1(yx)\psi_2(y)dy = x^{-2}\int y\psi_1(y)\psi_2(y/x)dy$ be the common density of X_i, for $i = 1,2$.

$$\begin{aligned}\mathrm{pr}(X_1 \le x \mid X_2 = x) &= \{f(x)\}^{-1} \int \Psi_1(yx)y\psi_1(yx)\psi_2(y)dy \\ &= \{f(x)\}^{-1}x^{-2} \int \Psi_1(y)y\psi_1(y)\psi_2(y/x)dy \\ &= \int \Psi_1(y)\psi^*(y,x)dy,\end{aligned}$$

where $\psi^*(y,x) = x^{-2}y\psi_1(y)\psi_2(y/x)/f(x)$. Since the density $\psi^*(y,x)$, with x treated as parameter, is totally positive of order two in (y,x) and $\Psi_1(y)$ is increasing in y, the above integral is increasing in x. This proves the lemma. □

Remark 2. Multivariate F belongs to the so-called scale mixture family of distributions considered in Lemma 2. It is easy to see that the density of $\chi^2_{\nu_0}$ at y, say $\psi_2(y)$, is such that $\psi_2(y/x)$ is totally positive of order two in (y,x). Hence, these X_i's that jointly have a multivariate F distribution satisfy Property 1.

LEMMA 3. *The Archimedean copulas listed above satisfy Property 1.*

Proof. The lemma will be proved by showing that $H'_\theta(u)$, the derivative of $H_\theta(u) = C_\theta(u,u)$ with respect to u, is non-decreasing in $u \in (0,1)$, and thus proving the desired convexity result, for each of these copulas.

(a). Clayton copula: $H_\theta(u) = (2u^{-\theta}-1)^{-1/\theta}$ if $2u^{-\theta} \ge 1$; otherwise $=0$, $\theta \in [-1,\infty)\setminus\{0\}$.

For this copula, $H'_\theta(u) = 2/\{(2u^{-\theta}-1)^{1/\theta+1}u^{\theta+1}\}$, which is non-decreasing in $u \in (0,1)$, since the denominator term has the following derivative, $-(1+\theta)(2u^{-\theta}-1)^{1/\theta}u^\theta$, which is

≤ 0, for $u \in (0,1)$.

(b). Gumbel copula: $H_\theta(u) = \exp[-\{2(-\ln u)^\theta\}^{1/\theta}] = u^{2^{1/\theta}},\ \theta \geq 1.$
Here, $H'_\theta(u) = 2^{1/\theta} u^{2^{1/\theta}-1}$, which is clearly non-decreasing in $u \in (0,1)$.

(c). Frank copula: $H_\theta(u) = -\theta^{-1} \log\left[\{\exp(-u\theta) - 1\}^2/\{\exp(-\theta) - 1\} + 1\right],\ \theta \in (-\infty, \infty) \setminus \{0\}$.
For this copula,

$$\begin{aligned} H'_\theta(u) &= \frac{2\exp(-2\theta u) - 2\exp(-\theta u)}{\exp(-2\theta u) - 2\exp(-\theta u) + \exp(-\theta)} \\ &= \frac{1}{1 - \left(\frac{1}{2}\right)\left\{\frac{\exp(-2\theta u) - \exp(-\theta)}{\exp(-2\theta u) - \exp(-\theta u)}\right\}}. \end{aligned}$$

It is easy to check that the term $\{\exp(-2\theta u) - \exp(-\theta)\}/\{\exp(-2\theta u) - \exp(-\theta u)\}$ is non-decreasing in $u \in (0,1)$, and so is $H'_\theta(u)$.

(d). Joe copula: $H_\theta(u) = 1 - \{2(1-u)^\theta - (1-u)^{2\theta}\}^{1/\theta},\ \theta \geq 1.$
For this copula, $H'_\theta(u) = 2\{2 - (1-u)^\theta\}^{1/\theta}\{(1-u)^\theta - 1\}/\{(1-u)^\theta - 2\}$, which is non-decreasing in $u \in (0,1)$, since both of the terms $\{2 - (1-u)^\theta\}^{1/\theta}$ and $\{(1-u)^\theta - 1\}/\{(1-u)^\theta - 2\}$ are so and are non-negative.

(e). Ali-Mikhail-Haq copula: $H_\theta(u) = u^2/\{1 - \theta(1-u)^2\},\ \theta \in [-1, 1)$.
For this copula, $H'_\theta(u) = 2u\{1 - \theta(1-u)\}/\{1 - \theta(1-u)^2\}^2$. Let $G_\theta(u) = H'_\theta(1-u)$. Then, we note that $G'_\theta(u) = 2y_\theta(u)/\{(1-\theta u^2)^3\}$, where $y_\theta(u) = -1 - \theta + 6\theta u - 3\theta u^2 - 3\theta^2 u^2 + 2\theta^2 u^3$. Since $y'_\theta(u) = 6\theta(1-u)(1-\theta u)$, we see the following:
Case 1: $\theta \geq 0$. The function $y_\theta(u)$ is non-decreasing in $u \in (0,1)$, and so takes the maximum value at $u = 1$, which is $-(1-\theta)^2 < 0$, implying that $y_\theta(u) < 0$ on $(0,1)$.
Case 2: $\theta < 0$. The function $y_\theta(u)$ is monotonically decreasing in $u \in (0,1)$, and takes the maximum value at $u = 0$, which is $-(1+\theta) \leq 0$, implying that $y_\theta(u) \leq 0$ on $(0,1)$.
Thus, $G_\theta(u)$ is non-increasing, and hence $H'_\theta(u)$ is non-decreasing, in $u \in (0,1)$. □

2. Additional simulation results

Figure 1 extends the same figure presented in the main text from $n = 16$ to $n = 100$, but focusing on the comparison between our proposed Method 1 and that of Seneta and Chen. As seen from this figure, the improvement of Method 1 over Seneta and Chen's when $n = 100$ is comparable to that when $n = 16$.

Figure 2 presents simulation results for the case of $n = 20$ for a moderately large $\rho = 0.70$. As seen here, the proposed Method 1 offers a small but noticeable power improvement over Seneta and Chen's, and its improvement over Hochberg's is significant. Although Method 1 is marginally improved by our Method 2 for this ρ, we have noticed, but not reported here, some improvement for larger ρ.

Figure 3 presents the percentage change in the average power of Method 1 over Seneta and Chen's. It is seen that the magnitude of improvement is comparable across different n's.

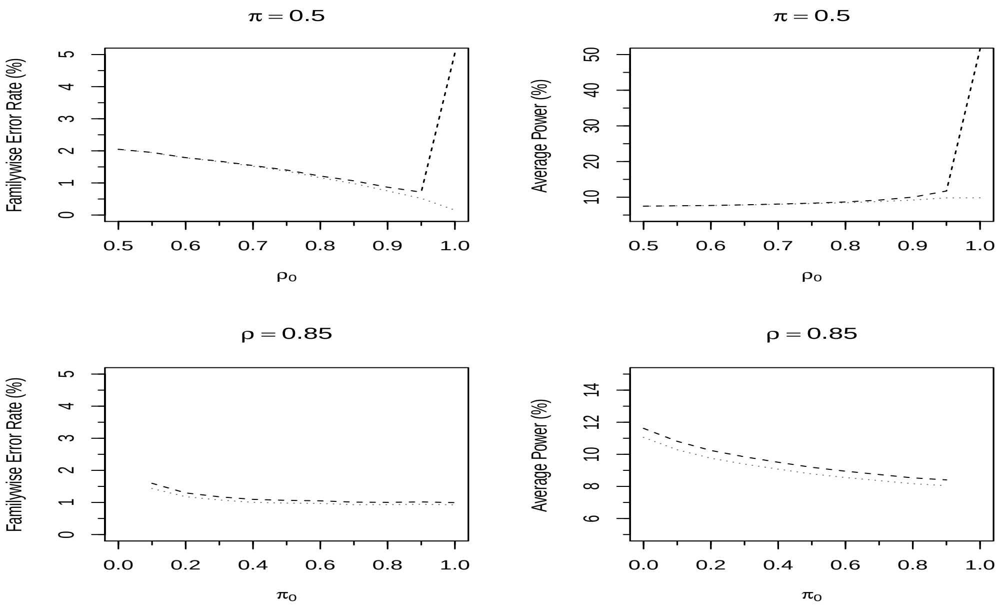

Fig. 1. Comparison Seneta and Chen (dotted) and the proposed Method 1 (dashed) when $n = 100$.

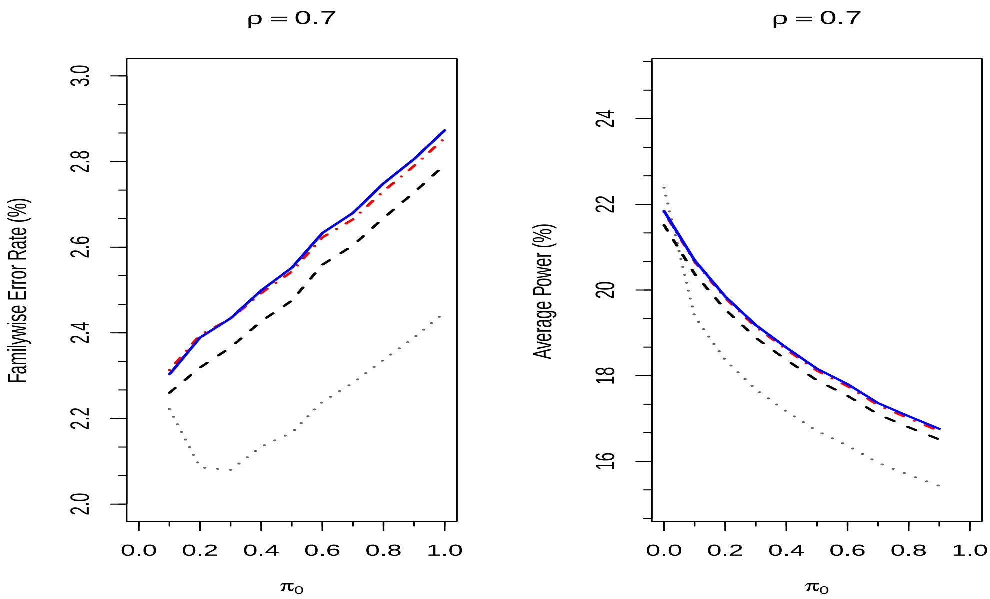

Fig. 2. Comparison of four methods, Hochberg (dotted), Seneta and Chen (dashed), proposed Method 1 (dot-dash), and proposed Method 2 (solid), when $n = 20$.

3. Estimating H and checking its convexity from Data

To illustrate how to implement the proposed procedures in practice without making any distributional assumptions allowing one to have a known form for G or H with the concavity or convexity condition, we considered analyzing a commonly used gene expression data from the

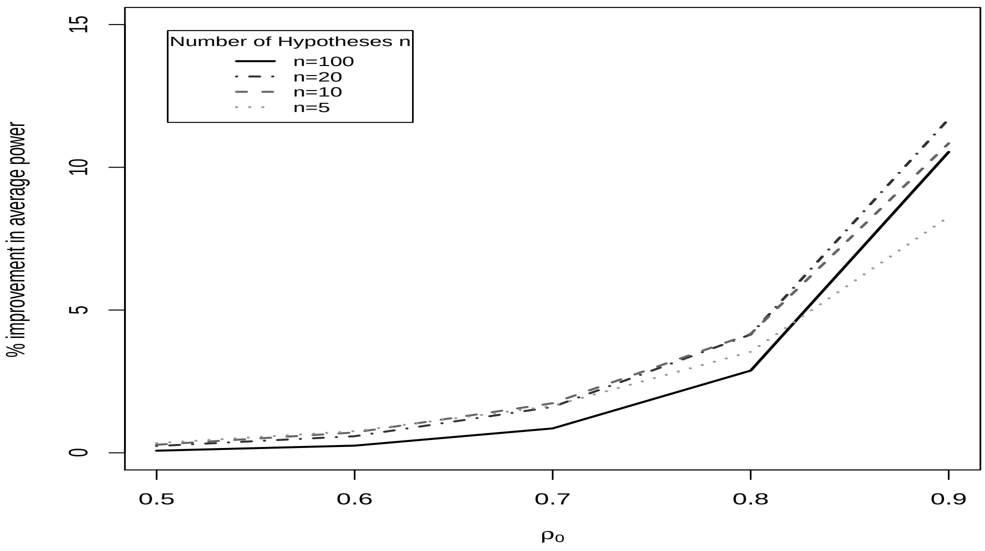

Fig. 3. Improvement of the proposed Method 1 over Seneta and Chen's under a mixture configuration of true and false null hypotheses. Specifically, 20% of the nulls are true, with mean of 0, and 80% of the false nulls are equally spread out into the four different values for the alternative mean – $0.5, 1.0, 1.5$, and 2.0.

leukemia microarray study of Golub et al. (1999). The data consist of 3051 gene expression levels across 38 tumor mRNA samples, of which 27 are of acute lymphoblastic leukemia and 11 are of acute myeloid leukemia. The data were log-transformed and normalized.

The goal of the study is to determine which genes are differentially expressed by testing H_{0i} : $\mu_{1i} = \mu_{2i}$ against $H'_{0i} : \mu_{1i} \neq \mu_{2i}$ simultaneously for $i = 1, \ldots, 3051$, where μ_{1i} and μ_{2i} are the gene specific mean expressions respectively for the acute lymphoblastic leukemia type and the acute myeloid leukemia type. For ease of illustration, we considered using a subset of data containing the first $n = 20$ gene expression levels.

We shall use two-sample t-test statistics for testing the 20 hypotheses and generating the corresponding p-values, $P_i, i = 1, \ldots, 20$. While applying our proposed procedures to the above p-values, we assume that the true null p-values are exchangeable; that is, we use Procedures 1 and 2 whose critical values are given in (2) and (5) respectively. To determine the critical values of Procedure 1, we need only to estimate the values of $H\{\alpha/(n - i + 1)\}$; whereas for Procedure 2, we need to estimate as well the values of $h\{\alpha/(n - i + 1)\}$, for $i = 1, \ldots, n = 20$.

In the two-group experimental setting, we used the permutation approach described in Dudoit and van der Laan $(2008, \S 2)$ to generate the distribution of the true null p-values. We considered generating $B = 100,000$ permutations between the two groups that correspond to the two leukemia types. For each permutated data, we used the two-sample t-test to calculate the corresponding p-value $P_i^{(b)}$ for each gene, where $i = 1, \ldots, n = 20$ and $b = 1, \ldots, B$. Then, for each pair of p-values, $(P_i^{(b)}, P_j^{(b)})$, where $1 \leq i < j \leq n = 20$, we calculated their maximum value, $\widetilde{P}_k^{(b)} = \max(P_i^{(b)}, P_j^{(b)})$, where $k = 1, \ldots, N$ and $N = n(n-1)/2$. Based on these calculated values, $\widetilde{P}_k^{(b)}, k = 1, \ldots, N$ and $b = 1, \ldots, B$, we computed its empirical distribution $\widehat{H}$, an estimate of H. Also, we derived an estimate $\widehat{h}$ of h using the R function density. Thus, we obtained

Table 1. *Estimated values of H and h and the calculated critical values in the order of Holm's procedure, Procedures 1 and 2 for $n = 20$ and $\alpha = 0.05$.*

i	$\widehat{H}\{\alpha/(n-i+1)\}$	$\widehat{h}\{\alpha/(n-i+1)\}$	$\alpha/(n-i+1)$	$\tilde{\alpha}_i$	α_i^*
1	0.0000069	0.0090777	0.0025000	0.0025066	0.0025066
2	0.0000078	0.0092484	0.0026316	0.0026390	0.0026390
3	0.0000085	0.0094381	0.0027778	0.0027859	0.0027859
4	0.0000096	0.0096500	0.0029412	0.0029503	0.0029503
5	0.0000112	0.0098885	0.0031250	0.0031355	0.0031356
6	0.0000125	0.0101588	0.0033333	0.0033450	0.0033451
7	0.0000144	0.0104677	0.0035714	0.0035849	0.0035850
8	0.0000169	0.0108241	0.0038462	0.0038618	0.0038619
9	0.0000192	0.0112399	0.0041667	0.0041843	0.0041844
10	0.0000225	0.0117863	0.0045455	0.0045660	0.0045661
11	0.0000275	0.0124501	0.0050000	0.0050249	0.0050251
12	0.0000342	0.0132615	0.0055556	0.0055861	0.0055863
13	0.0000429	0.0142756	0.0062500	0.0062878	0.0062880
14	0.0000555	0.0156962	0.0071429	0.0071908	0.0071911
15	0.0000741	0.0175943	0.0083333	0.0083955	0.0083960
16	0.0001061	0.0204084	0.0100000	0.0100856	0.0100863
17	0.0001611	0.0247778	0.0125000	0.0126220	0.0126231
18	0.0002785	0.0322859	0.0166667	0.0168544	0.0168564
19	0.0006144	0.0474240	0.0250000	0.0253110	0.0253146
20	0.0023727	0.0942955	0.0500000	0.0500000	0.0500000

the estimated values of both $H\{\alpha/(n-i+1)\}$ and $h\{\alpha/(n-i+1)\}$ for $i = 1, \ldots, n$, and computed the critical values of Procedures 1 and 2, which are presented in Table 1.

We then generated the plots of $\widehat{H}$ and $\widehat{h}$, and graphically checked whether H is convex and h is increasing in u on an interval $(0, \alpha_0)$, with $\alpha_0 > \alpha$. As seen from Figure 4, $\widehat{H}$ is indeed convex on $(0, 1)$ and $\widehat{h}$ is increasing on $(0, 0.9)$. Thus, the desired conditions for H and h are satisfied in this real data example.

The above example illustrates how the distribution function H and the density function h of the pairwise maxima of null p-values can be estimated from data by an appropriate resampling method, and doing so, the assumptions pertaining to H and h can be checked graphically.

References

DUDOIT, S. & VAN DER LAAN, M. (2008). *Multiple Testing Procedures with Applications to Genomics*. New York: Springer.

GOLUB, T., SLONIM, D., TAMAYO, P., HUARD, C., GAASENBEEK, M., MESIROV, J., COLLER, H., LOH, M., DOWNING, J., CALIGIURI, M., BLOOMFIELD, C., & LANDER, E. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531-537.

KARLIN, S. (1968). *Total Positivity*. Stanford: Stanford Univeristy Press.

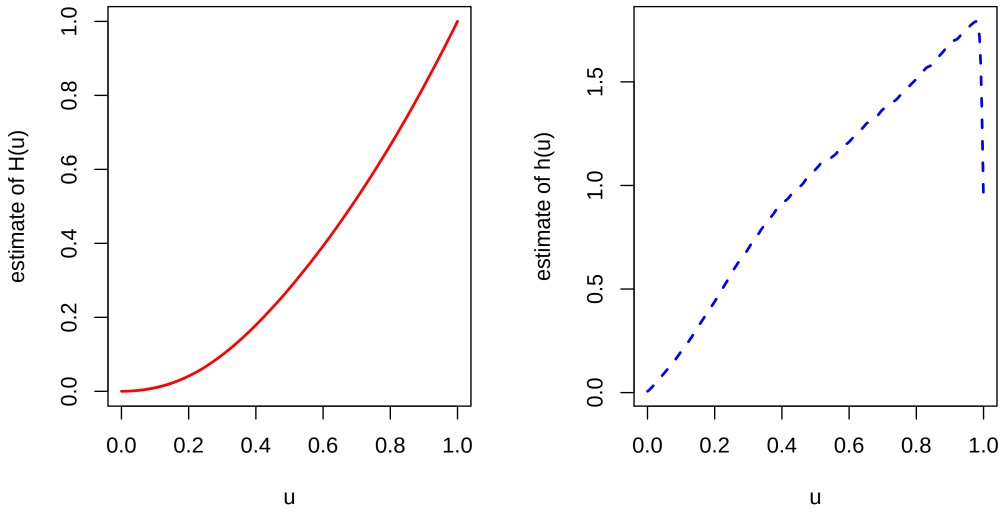

Fig. 4. Estimates of H and h