

On Familywise Error Rate Cutoffs in a Pairwise Exchangeable Setting. A Precis.

Thomas Fung^{1,*} and Eugene Seneta²

¹ School of Mathematical and Physical Sciences, Macquarie University, North Ryde, N.S.W. 2109, Australia

² School of Mathematics and Statistics FO7, The University of Sydney, N.S.W. 2006, Australia

ISI 2023, 2023-07-18

Summary In a pairwise exchangeable dependence setting for test statistics, the elegant cutoffs of Sarkar et al. (2016) may be viewed as a first iteration improvement of Holm (1979)'s classical cutoffs, under a convexity condition on the copula. The cutoffs of Seneta and Chen (1997) which improve Holm's in the present exchangeability setting, are shown, on an analogous first iteration step, to lead to a refinement of Sarkar et al. (2016). Further, we show that the convexity condition can be circumvented in practice, computationally. Comparisons between the effects of the several cutoff sets are made by way of plots of the familywise error rate against correlation ρ in the classic setting of the multivariate Normal; and the distributional setting of the multivariate Generalized Hyperbolic for the

*Corresponding author. Email address: thomas.fung@mq.edu.au (Thomas Fung).

important Variance Gamma type subfamily, for which a convexity condition cannot be analytically verified.

1 The Stepdown Procedure

We consider the multiple test problem when there are n hypotheses H_1, H_2, \dots, H_n , and corresponding P-values R_1, R_2, \dots, R_n , assuming the test statistics X_1, X_2, \dots, X_n are from a continuous distribution. Suppose that in such a multiple test procedure the property:

$$Pr(H_s, s \in I_m, \text{ are accepted} | H_s, s \in I_m \text{ are true}) \geq 1 - \alpha \quad (1)$$

holds for prespecified size of test α (familywise error rate: FWE) for every I_m , where I_m is any non-null subset of $\{1, 2, \dots, n\}$, for every $m = 1, 2, \dots, n$. Then the FWE is said to be strongly controlled.

Let $R_{(1)}, R_{(2)}, \dots, R_{(n)}$ be the ordered P-values, and $\Delta(i), 1 \leq i \leq n$ be a strictly increasing sequence of constants, with $0 < \Delta(i) < 1$, for each i . A step-down procedure begins by testing if $R_{(1)} < \Delta(1)$. If so, reject $H_{(1)}$ and go on to test if $R_{(2)} < \Delta(2)$. If not, accept all hypotheses. In general, if $R_{(i)} < \Delta(i), 1 \leq i \leq j - 1$, then at step j the remaining hypotheses are $H_{(j)}, H_{(j+1)}, \dots, H_{(n)}$, and the inequality next to be checked is $R_{(j)} < \Delta(j)$. If it holds, reject $H_{(j)}$ and continue. Otherwise accept $H_{(j)}, H_{(j+1)}, \dots, H_{(n)}$. The process may continue until a decision is made on the basis of whether $R_{(n)} < \Delta(n)$. The step-down procedure of Holm (1979) uses the set of constants

$$\bar{\Delta}(i) = \frac{\alpha}{n - i + 1}, \quad 1 \leq i \leq n, \quad (2)$$

and Holm proved that with these constants the FWE is strongly controlled.

2 The Exchangeability Context

In the sequel we assume that $(X_i, X_j), i \neq j, i, j \in I_m$ are *exchangeable*, (so that (R_i, R_j) are also) when $I_m, m \geq 2$ is the index set of assumed true hypotheses, and that their joint bivariate distribution is the same for each such I_m . The marginal distribution of each $R_i, i \in I_m, m \geq 1$ is $U(0, 1)$. We introduce the notation $H(u) = Pr(R_1 \leq u, R_2 \leq u), u \in (0, 1)$, and $n_0 = n - i + 1$. $H(u) = Pr(R_1 \leq u, R_2 \leq u), u \in (0, 1)$ Seneta and Chen (1997), Sections 4, 5 showed that then the critical cutoff constants:

$$\tilde{\Delta}(i) = \frac{\alpha + (n_0 - 1)H(\frac{\alpha}{n_0})}{n_0}, \quad (3)$$

are monotonic increasing with i (that is: decreasing with increasing n_0), and the step down procedure based on them strongly controls the FWE. Moreover the step-down procedure with these constants provides tighter control on the FWE than Holm's since obviously

$$\tilde{\Delta}(i) > \bar{\Delta}(i) = \frac{\alpha}{n_0}, 1 \leq i \leq n - 1, \tilde{\Delta}(n) = \bar{\Delta}(n) = \alpha.$$

Note that $H(u) = Pr(\max(R_1, R_2) \leq u)$, so that $H(u), 0 < u < 1$, is a cdf. It is also the copula $C(u_1, u_2) = Pr(R_1 \leq u_1, R_2 \leq u_2), 0 < u_1, u_2 < 1$, of the bivariate joint distribution of the exchangeable pair (X_1, X_2) , evaluated on the diagonal $u_1 = u_2 = u$.

In fact, any sequence $w_1(n_0), n_0 = 1, 2, \dots, n$ satisfying $w_1(1) = \alpha$ and for $n_0 \geq 2$

$$0 < w_1(n_0) < \alpha, w_1(n_0) < w_1(n_0 - 1), G_{n_0}(w_1(n_0)) < \alpha, \quad (4)$$

where $G_{n_0}(u)$ is given by

$$G_{n_0}(u) = n_0 u - (n_0 - 1)H(u), \quad 0 < u < 1, \quad (5)$$

and $H(u) = \Pr(R_1 \leq u, R_2 \leq u), u \in (0, 1)$, provides a set of monotonic increasing cut-offs which assure strong control of the FWE in the step-down test procedure, by setting $\Delta(i) = w_1(n - i + 1), i = 1, 2, \dots, n$.

This statement is implicit in Sarkar et al's reformulation of the justification in Seneta and Chen (2005), when specialised to the exchangeable case.

The cutoffs (2), (3) satisfy the description of such a sequence $w_1(n_0), n_0 = 1, 2, \dots, n$.

Sarkar et al. (2016), motivated by Seneta and Chen (2005), make a substantial improvement on (3). They make the overarching Assumption: $H(u)$ is convex in $u \in (0, 1)$, and consequently show that the monotonic increasing cut-offs:

$$\hat{\Delta}(i) = \frac{\alpha^2/n_0}{G_{n_0}(\alpha/n_0)}, \quad (6)$$

where G_{n_0} is defined by (5) provide tighter control on the FWE than (3), since

$$\hat{\Delta}(i) > \tilde{\Delta}(i), \quad 1 \leq i \leq n - 1, \quad \tilde{\Delta}(n) = \bar{\Delta}(n) = \alpha. \quad (7)$$

Now define for $n_0 = 1, 2, \dots, n$:

$$w_2(n_0) \stackrel{\text{def}}{=} \frac{\alpha w_1(n_0)}{G_{n_0}(w_1(n_0))}. \quad (8)$$

When $w_1(n_0) = \bar{\Delta}(i) = \frac{\alpha}{n_0}$ it follows from (6) that $w_2(n_0) = \hat{\Delta}(i) = \frac{\alpha^2/n_0}{G_{n_0}(\alpha/n_0)}$. Our main result is:

Theorem: Let $\alpha, 0 < \alpha < 1$, be a fixed constant. Then

$$w_2(1) = \alpha; \text{ and for } n_0 \geq 2, \quad 0 < w_1(n_0) < w_2(n_0) < \alpha. \quad (9)$$

If we assume $H(u)/u \uparrow, u \uparrow$, then for $n_0 \geq 2$:

$$w_2(n_0) < w_2(n_0 - 1), \quad G_{n_0}(w_2(n_0)) < \alpha. \quad (10)$$

so that, from (4)(5), the cutoffs $w_2(n_0)$ strongly control the FWE.

If $w_1^*(n_0) > w_1(n_0)$, $n_0 \geq 2$ and $w_1^*(n_0)$ satisfies the same conditions (4)(5) as $w_1(n_0)$, and if $H(u)/u \uparrow, u \uparrow$, then

$$w_2^*(n_0) > w_2(n_0), \quad n_0 \geq 2. \quad (11)$$

□

Since, as we have noted from (3), $w_1^*(n_0) = \tilde{\Delta}(i)$ satisfies the conditions of the theorem, we conclude that $w_2^*(n_0) > \hat{\Delta}(i)$, so that the cutoffs $w_2^*(n_0)$ provide tighter control on FWE than (6).

3 Verification. FWER

In both Sarkar et al. (2016) and in its Supplementary Materials, much attention is focussed on the problem of verifying the assumption of convexity of $H(u)$, $0 < u < 1$. According to the Theorem above, the conditions for the copula $H(u)$, to obtain the same conclusion, are milder.

Further, the validity of (10) and (11) can be checked directly from the values of $w_2(n_0)$, and the values of $\{w_i(n_0), w_i^*(n_0), i = 1, 2\}$ respectively in a testing setting, without knowledge, such as $H(u)/u \uparrow, u \uparrow$, which removes the dependence on such a condition in statistical application. We make use of this in Section 4, in an example when the condition

$H(u)/u \uparrow, u \uparrow$ cannot be analytically verified.

For comparisons of the differential effects of various cutoffs, we use the familywise error rate (FWER) as a function of ρ as the means of comparison. This is in general defined as $Pr(\text{Reject at least one } H_i, i = 1, 2, \dots, n | H_i, i = 1, 2, \dots, n \text{ are all true})$. Thus for the stepdown procedure

$$\begin{aligned} \text{FWER} &= Pr(R_{(1)} < \Delta(1) | H_i, i = 1, 2, \dots, n \text{ are all true}) \\ &= Pr(R_{(1)} < w_j(n) | H_i, i = 1, 2, \dots, n \text{ are all true}), \\ &= Pr\left(\min_{i=1,2,\dots,n} X_i < F_X^{-1}(\Delta(1)) | H_i, i = 1, 2, \dots, n \text{ are all true}\right) \end{aligned} \quad (12)$$

for fixed $j = 1, 2$, in terms of sample values $X_i, i = 1, 2, \dots, n$, where $F_X(x)$ is the marginal cdf of X_i . The expression (12) may be estimated from repeated simulation of the sample $X_i, i = 1, 2, \dots, n$, once $F_X^{-1}(\Delta(1))$ is numerically calculated.

4 The Generalized Hyperbolic Distribution. Tests.

The multivariate skew GH distribution, $GH(0, R_n, \boldsymbol{\theta}, p, a, b)$ which we take as the joint distribution of our test-statistics $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$ to have is defined by its mean-variance mixing representation as

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\theta}W + \sqrt{W}\mathbf{Z} \quad (13)$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)^\top$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)^\top$, and $W \sim GIG(p, a, b)$ is independently distributed of $\mathbf{Z} \sim N(0, R_n)$. Here $R_n = (1 - \rho)bsI_n + \rho\mathbf{1}_n\mathbf{1}_n^\top$, with $-1 < \rho < 1$, so that $R_2 = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. We assume that the random variable W has a (univariate) Generalised

Inverse Gaussian (GIG) distribution, denoted by $GIG(p, a, b)$, that is, it has density

$$\begin{aligned} f_{GIG}(w) &= \frac{1}{2\bar{K}_p(a, b)} w^{p-1} \exp\left(-\frac{1}{2}(a^2 w^{-1} + b^2 w)\right), \quad w > 0; \\ &= 0, \quad \text{otherwise;} \end{aligned}$$

where

$$\bar{K}_p(a, b) = \begin{cases} \left(\frac{a}{b}\right)^p K_p(ab), & p \in \mathbb{R}, \text{ if } a, b > 0; \\ b^{-2p} \Gamma(p) 2^{p-1}, & p, b > 0, \text{ if } a = 0; \\ a^{2p} \Gamma(-p) 2^{-p-1}, & a > 0 \text{ and } p < 0, \text{ if } b = 0, \end{cases} \quad (14)$$

Here $K_p(\omega)$, $\omega > 0$, is the modified Bessel function of the second kind (Erdélyi et al. (1954)) with index $p \in \mathbb{R}$.

We take for, $i = 1, 2, \dots, n$, our i th null hypothesis as : $H_i : \mu_i = \theta_i = 0$, so that when all H_i are true,

$$\mathbf{X} = \sqrt{W} \mathbf{Z}, \quad (15)$$

and similarly for any subset of indices $\{1, 2, \dots, n\}$, so that exchangeability holds, and the other assumptions of Section 1.2, The Exchangeability Context, hold. Then the distribution functions of the X_i are the same: $F_i(u) = F_1(u)$. The marginal density of each of X_1 is then given by

$$f_{X_1}(x) = \frac{\bar{K}_{p-1/2}((x^2 + a^2)^{1/2}, b)}{\sqrt{2\pi} \bar{K}_p(a, b)}, \quad x \in \mathbb{R}. \quad (16)$$

The case $b = 0$ of (15) encompasses the symmetric multivariate t -distribution, which has been studied as a central example in Sarkar et al. (2016). We therefore take up an

example to illustrate the case $b > 0$. This setting for exchangeability is additional to the various examples considered by Sarkar and Chang (1997) and Sarkar et al. (2016).

The Variance Gamma (VG) distribution, which is a special case, when $a = 0$, $b = \sqrt{\frac{2}{\nu}}$, $p = \frac{1}{2}$, is important in financial mathematics modelling. The case $b > 0$ has been studied extensively in the bivariate case ($n = 2$), by Fung and Seneta (2016), and this bivariate case is central, as we have seen, to the role of the copula $H(u)$ in our exchangeable setting.

Given that we haven't been able to analytically verified that $H(u)/u \uparrow$, as $u \uparrow$, for the exchangable GH case, we have to check the conditions numerically to ensure the iterative cutoffs $w_2(n_0)$ are strongly controlling FWER. Form Table 1, we can see that $w_2(n_0)$ satisfied (10) and therefore the cutoffs $w_2(n_0)$ strongly control FWE.

In Figure 1, we plotted the error rate in the multivariate GH setting with $a = 0$, $p = 4$, $b = 0.5$, at level $\alpha = 0.05$ with $n = 16$ for various values of common correlation ρ and two cutoff constants $\hat{\Delta}(i)$ and $w_2(n_0)$ (from Sarkar et al. (2016) and by iterating the cutoffs of $\tilde{\Delta}(i)$ respectively):

$$\hat{\Delta}(i) = \frac{\alpha^2/n_0}{G_{n_0}(\alpha/n_0)},$$

$$w_2(n_0) = \frac{\alpha w_1(n_0)}{G_{n_0}(w_1(n_0))}, \quad \text{where} \quad w_1(n_0) = \tilde{\Delta}(i) = \frac{\alpha + (n_0 - 1)H\left(\frac{\alpha}{n_0}\right)}{n_0}.$$

One hundred thousand (100,000) independent replications were used in all simulations. The smoothed line is provided by a GAM fit. We can see that $w_2(n_0)$ provided tighter control of FWE for high values of ρ but performance was indistinguishable otherwise.

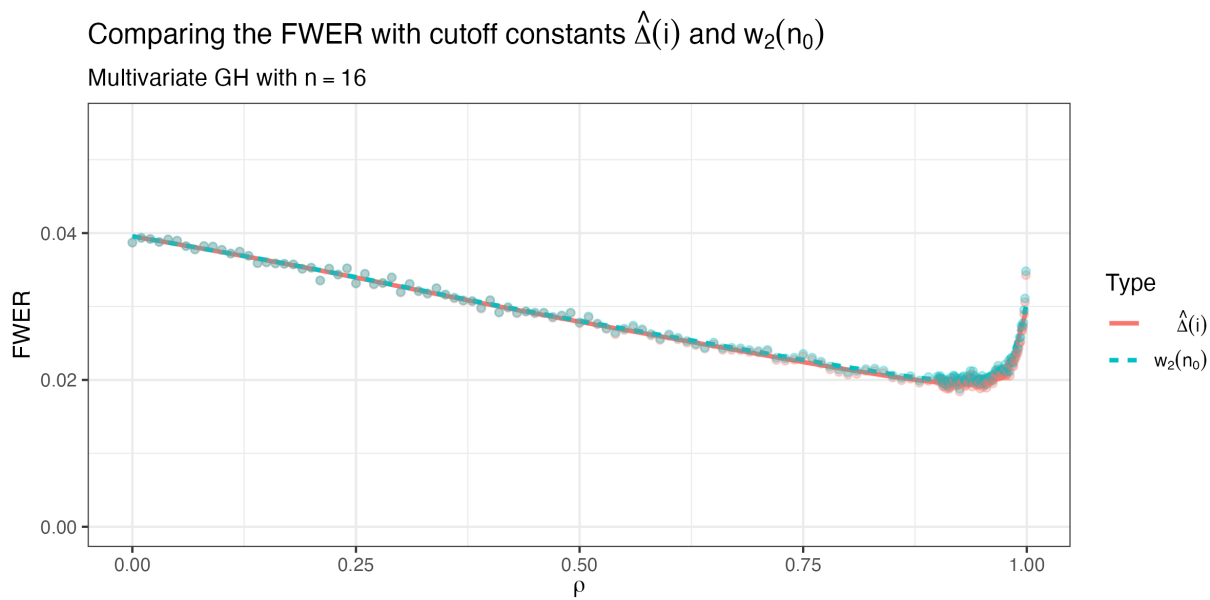


Figure 1: Comparing the familywise error rate with cutoff constants $\hat{\Delta}(i)$ and $w_2(n_0)$

References

- A. Erdélyi, W. Magnus, F. Oberhettinger, and F. Tricomi. *Bateman Manuscript Project: Tables of Integral Transform*, volume 2. McGraw-Hill, New York, 1954.
- T. Fung and E. Seneta. Tail asymptotics for the bivariate skew normal. *Journal of Multivariate Analysis*, 144:129–138, Feb. 2016. ISSN 0047259X. doi: 10.1016/j.jmva.2015.11.002.
- S. Holm. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979. ISSN 0303-6898.
- S. K. Sarkar and C.-K. Chang. The Simes Method for Multiple Hypothesis Testing With

Positively Dependent Test Statistics. *Journal of the American Statistical Association*, 92(440):1601–1608, 1997. ISSN 0162-1459. doi: 10.2307/2965431.

S. K. Sarkar, Y. Fu, and W. Guo. Improving holm’s procedure using pairwise dependencies. *Biometrika*, 103(1):237–243, 03 2016. doi: 10.1093/biomet/asv063. URL <https://academic.oup.com/biomet/article-lookup/doi/10.1093/biomet/asv063>. tex.ids= sarkarSupplementaryMaterialImproving2015.

E. Seneta and J. T. Chen. Simple Stepwise Tests of Hypotheses and Multiple Comparisons. *International Statistical Review / Revue Internationale de Statistique*, 73(1): 21–34, 2005. ISSN 0306-7734.

E. Seneta and T. Chen. A sequentially rejective test procedure. *Theory of Stochastic Processes*, 3(19):393–402, 1997. doi: [https://www.maths.usyd.edu.au/u/eseneta/SenetaChen\(1997\).pdf](https://www.maths.usyd.edu.au/u/eseneta/SenetaChen(1997).pdf). URL [https://www.maths.usyd.edu.au/u/eseneta/SenetaChen\(1997\).pdf](https://www.maths.usyd.edu.au/u/eseneta/SenetaChen(1997).pdf).

i	n_0	$w_1(n_0)$	$G(w_1(n_0))$	$w_2(n_0)$	$G(w_2(n_0))$
1	16	0.00313	0.02183	0.00716	0.04801
2	15	0.00333	0.02189	0.00761	0.04801
3	14	0.00357	0.02196	0.00813	0.04801
4	13	0.00385	0.02205	0.00872	0.04802
5	12	0.00417	0.02217	0.00940	0.04803
6	11	0.00455	0.02231	0.01019	0.04804
7	10	0.00500	0.02249	0.01112	0.04806
8	9	0.00556	0.02272	0.01222	0.04810
9	8	0.00625	0.02303	0.01357	0.04814
10	7	0.00714	0.02345	0.01523	0.04820
11	6	0.00833	0.02403	0.01734	0.04830
12	5	0.01000	0.02488	0.02009	0.04844
13	4	0.01250	0.02623	0.02382	0.04865
14	3	0.01667	0.02861	0.02913	0.04898
15	2	0.02500	0.03364	0.03715	0.04949
16	1	0.05000	0.05000	0.05000	0.05000

Table 1: Checking $w_2(n_0)$ can strongly control FWER