Statistica Sinica Preprint No: SS-2017-0316	
Title	Selective Sign-Determining Multiple Confidence Intervals with FCR Control
Manuscript ID	SS-2017-0316
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202017.0316
Complete List of Authors	Asaf Weinstein and Daniel Yekutieli
Corresponding Author	Asaf Weinstein
E-mail	asafw@stanford.edu

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

Statistica Sinica

SELECTIVE SIGN-DETERMINING MULTIPLE CONFIDENCE INTERVALS WITH FCR CONTROL

Asaf Weinstein and Daniel Yekutieli

Stanford University and Tel Aviv University

Abstract: **Given m unknown parameters with corresponding independent estimators, the Benjamini–Hochberg (BH) procedure can be used to classify the signs of the parameters, such that the expected proportion of erroneous directional decisions (directional FDR) is controlled at a preset level q. More ambitiously, our goal is to construct sign-determining confidence intervals—instead of only classifying the sign—such that the expected proportion of noncovering constructed intervals (FCR) is controlled. We suggest a valid procedure that adjusts a marginal confidence interval to construct a maximum number of sign-determining confidence intervals. We propose a new marginal confidence interval, designed specifically for our procedure, that allows us to balance the trade-off between the power and the length of the constructed intervals. We apply our methods to detect the signs of correlations in a highly publicized social neuroscience study and, in a second example, to detect the direction of association for SNPs with Type-2 diabetes in GWAS data. In both examples, we compare our**

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

procedure to existing methods and obtain encouraging results.

Key words and phrases: Confidence intervals, directional decisions, False Coverage Rate, False Discovery Rate, Selective Inference, multiplicity

1. Introduction

Let f be a known density symmetric about zero, and suppose that an analyst collects independent observations $Y_i \sim f(y_i - \theta_i)$, for $i = 1, ..., m$, corresponding to unknown location parameters $\theta_i \in \mathbb{R}$. In many applications, the analyst will highlight a subset of parameters that the data suggest are interesting, and then focus the inferences on these highlighted parameters only. A prototypical case is constructing confidence intervals for parameters θ_i corresponding to only rejected null hypotheses $H_{0i} : \theta_i = \theta_{0i}$, for $i = 1, ..., m$. For example, in an RNA microarray, we may be interested in genes that are differentially expressed. More generally, we may consider a two-stage procedure, in which the analyst first attempts to answer a question of primary interest related to each of the θ_i. Then, in the second stage, and *only if* he was able to answer the primary question related to θ_i with enough certainty, the analyst will pose a secondary, follow-up question related to θ_i. The first question is often intended to detect "signals" of one or more types; the subsequent question may depend on the answer to the

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

first, and is usually intended to learn about further qualities of θ_i.

In this study, the primary question concerns the sign of the parameters. Specifically, the analyst is first interested in classifying the signs the of parameters θ_i as positive (> 0) or nonpositive (≤ 0); we refer to this as a *weak sign classification*. A third decision—declaring "inconclusive data"—is allowed when an observation size is too small to infer the sign. The sign classification problem is important in many applications. For example, when comparing several drugs to a control, it may be of interest to determine which are better (difference positive), and which are not (difference nonpositive). Bohrer (1979) and Bohrer and Schervish (1980) called the sign classification problem a multiple *three-decision* problem (in reference to Neyman et al., 1977), and studied optimality of decision rules under familywise error rate (FWER) control, that is, the probability of making at least one incorrect directional decision. In the current study, we consider rules that control the *weak directional* FDR, which we define as

$$\text{wdFDR} := \text{E}\left[\frac{V_D}{R_D \vee 1}\right], \tag{1.1}$$

where R_D is the number of parameters with signs that were classified, and V_D is the number of parameters with signs that were classified incorrectly: a nonpositive parameter declared as positive, or a positive parameter declared as nonpositive. Above and elsewhere, we use $a \vee b$ to denote the maximum

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

between two numbers a and b.

A procedure known to control wdFDR is the directional-BH procedure. Here, parameters are selected via the usual BH procedure, using two-sided p-values to test $H_{0i} : \theta_i = 0$. A directional decision is made for each selected parameter according to the sign of the estimator. In fact, the directional-BH procedure makes a *strict* sign classification for each selected parameter, that is, θ_i is declared negative (if $Y_i < 0$) or positive (if $Y_i > 0$), and still controls the expected proportion of incorrect decisions (Benjamini and Yekutieli, 2005). The latter is known as the *mixed directional* FDR (Benjamini et al., 1993), and is a stronger version of the directional FDR. Note that it may differ from the wdFDR when zero parameters are possible. The procedures we consider are required to make only weak directional decisions and control the wdFDR. With regard to terminology, when we refer to a sign classification or a directional decision, unless otherwise noted, it should be understood in the weak sense, that is, positive or nonpositive.

Whereas a directional decision may be of primary importance, in practice, it is almost always desirable to supplement such a decision with a confidence interval. For example, if the data suggest that a particular drug among a candidate set improves over a control, we would like to be able to say how big the difference is at least. If the difference is immaterial, the

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

drug may not be prescribed.

Thus, a more general objective is to construct, upon observing $Y_i, i \leq m$, confidence intervals for a *subset* of the m parameters, such that (i) each constructed confidence interval includes either only positive or only nonpositive values, and (ii) we have control at level q over the false coverage rate (Benjamini and Yekutieli, 2005)

$$\text{FCR} = \text{E}\left[\frac{V_{CI}}{R_{CI} \vee 1}\right], \tag{1.2}$$

where R_{CI} is the number of confidence intervals constructed, and V_{CI} is the number of noncovering confidence intervals constructed. A confidence interval is said to be *sign-determining* if it includes only positive or only nonpositive values. Correspondingly, we call a procedure with the property (i) a "selective sign-determining confidence intervals procedure", abbreviated as *Selective-SDCI*. We say that a Selective-SDCI procedure with property (ii) is *valid at level q*.

The directional-BH procedure corresponds to a valid Selective-SDCI procedure, that trivially constructs the interval $(0, \infty)$ for any selected parameter declared positive, and the interval $(-\infty, 0)$ (note the exclusion of zero) for any parameter declared negative. Here, we first propose a more general valid procedure that can be used to construct *nontrivial* sign-determining selective confidence intervals. In the case of a single parameter,

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

$Y \sim f(y-\theta)$, our procedure is very simple: suppose that $\mathcal{C}(y;\alpha)$ is any marginal $1-\alpha$ confidence interval; that is, $\Pr_\theta\left(\theta \notin \mathcal{C}(Y;\alpha)\right) \leq \alpha$. Then, construct $\mathcal{C}(Y;\alpha)$ if and only if it is a subset of $(0,\infty)$ or of $(-\infty,0]$. In that case,

$$\begin{aligned} \text{FCR} &= \Pr\left(\{\theta \notin \mathcal{C}(Y;\alpha)\} \cap \{\mathcal{C}(Y;\alpha) \text{ is constructed}\}\right) \\ &\leq \Pr\left(\theta \notin \mathcal{C}(Y;\alpha)\right) \leq \alpha. \end{aligned} \tag{1.3}$$

Benjamini and Yekutieli (2005) pointed out (1.3) to demonstrate that selection is handled gracefully, as long as multiplicity is not involved. A main thrust of the current work is to extend the above to the case of general m. Of course, reporting all marginal $1-q$ confidence intervals that are a subset of either $(0,\infty)$ or $(-\infty,0]$ will not lead to a valid q-level procedure. Instead, for any marginal confidence interval procedure, our method employs the FCR adjustment of Benjamini and Yekutieli (2005) to produce the largest set of *sign-determining* FCR-adjusted marginal confidence intervals.

Because the choice of the *marginal* confidence interval procedure determines our Selective-SDCI procedure entirely, it controls both the power of the procedure as a sign-determining rule and the length (and shape) of the constructed intervals. In line with the recent work of Fithian et al. (2014) and Tian et al. (2018), our procedure shows that there is a trade-off between these objectives: in general, higher power incurs a cost of lower "accuracy"

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

of the constructed intervals (as measured by their length and shape). This leads us to derive a *new* marginal confidence interval that enables our procedure to control the trade-off. The corresponding procedure determines the signs of the parameters according to a level-$(\psi \cdot 2q)$ directional-BH procedure for $1/2 \leq \psi \leq 1$, and constructs sign-determining intervals that, loosely speaking, are longer for larger ψ.

Along with the original sign problem, we offer two extensions of our procedure. The first is motivated by an example from genetics, and generalizes our procedure to the case of a two-dimensional parameter, where the primary objective is to classify the sign of the first component. The second extension goes beyond the sign problem. Here, we consider detecting parameters $\theta_i > \delta$ or $\theta_i < -\delta$ where δ is some prespecified quantity. To address this problem, we propose a procedure that constructs selective confidence intervals, such that each interval contains either only values larger than δ or only values smaller than $-\delta$.

The remainder of the paper is organized as follows. Section 2 reviews the work of Benjamini and Yekutieli (2005), which we use in Section 3 to derive a class of valid Selective-SDCI procedures. Section 4 presents a new marginal confidence interval, designed specifically to be used in a Selective-SDCI procedure. In Section 5 we generalize our method to con-

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

struct sign-determining confidence regions for parameters with more than one dimension. The results of our simulation study are reported in Section 6. In Section 7, we use our method to detect the sign of correlations in a neuroscience study. All proofs are deferred to the Appendix.

Notation. $\mathcal{C}(y;\alpha)$ is a confidence interval that covers the true value with probability at least $1-\alpha$, and should be understood as a function of both y and α, unless the context suggests otherwise. To emphasize the dependency on α, we sometimes write $\{\mathcal{C}(\cdot;\alpha) : 0 \leq \alpha \leq 1\}$ instead, and call it a confidence interval *procedure*. Throughout, f denotes a probability density, and F is the corresponding distribution function. We denote by c_α the $1-\alpha$ quantile of a distribution, that is, the value $F^{-1}(1-\alpha)$; z_α is used for the special case of a standard normal distribution. Throughout, we write $Y_{(i)}$ for the observation with the ith *largest absolute value*; that is, $|Y_{(m)}| \leq |Y_{(m-1)}| \leq ... \leq |Y_{(1)}|$. Finally, for any set B, define $-B := \{-x : x \in B\}$. We tried to minimize the use of nonstandard acronyms, but avoiding them altogether would result in a tedious manuscript. The main acronyms used in this paper are as follows: CI=Confidence Interval; wdFDR=Weak Directional-FDR; SDCI=Sign-Determining Confidence Intervals; BY=Benjamini and Yekutieli (2005); QC=Quasi-Conventional; MQC=Modified Quasi-Conventional.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

2. Review

Benjamini and Yekutieli (2005, BY hereafter) set up a framework for selective inference when multiple parameters are considered. Let $\mathbf{Y} = (Y_1, ..., Y_m)$ be a vector of estimators, where $Y_j \sim f(y_j - \theta_j)$. Suppose that $\mathcal{S}$ is a pre-specified selection rule yielding a subset $S = \mathcal{S}(\mathbf{Y}) \subset \{1, ..., m\}$, and that a procedure, which may depend on $\mathcal{S}$ and on S, is used to construct confidence intervals for the selected parameters $\{\theta_j : j \in S\}$ only. Denote by R_{CI} the number of confidence intervals constructed, and by V_{CI} the number of noncovering confidence intervals constructed. Then, BY define the false coverage-statement rate (FCR) as the expected value of

$$Q_{CI} = \frac{V_{CI}}{R_{CI} \vee 1}.$$

Thus, the FCR depends on $\mathcal{S}$, which specifies what subset of parameters is selected given the data, and on the procedure that specifies how confidence intervals are constructed for any selected subset of parameters.

Suppose we have a marginal confidence interval procedure $\{\mathcal{C}(\cdot; \alpha) : 0 \leq \alpha \leq 1\}$, which, for any $\alpha \in [0, 1]$, specifies a $(1-\alpha)$-level marginal confidence interval for θ, based on $Y \sim f(y - \theta)$. That is, $\Pr_\theta(\theta \in \mathcal{C}(Y; \alpha)) \geq 1 - \alpha$ holds for any $\alpha \in [0, 1]$. In the sequel, we will often refer to a marginal confidence interval *procedure* simply as a confidence *interval* and write $\mathcal{C}(y; \alpha)$,

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

which should be understood as a function of both y and α. Suppose that the procedure $\mathcal{C}$ satisfies the following monotonicity requirement:

Requirement (MON 1) For any y and any $0 \leq \alpha, \alpha' \leq 1$, if $\alpha' \leq \alpha$, then $\mathcal{C}(y;\alpha) \subseteq \mathcal{C}(y;\alpha')$.

Denote $CI_i(\alpha) = \mathcal{C}(Y_i;\alpha)$. For $m > 1$, if $\mathcal{S}$ is an arbitrary selection rule, then constructing $CI_i(q)$ for each $i \in S$ does not, in general, guarantee FCR $\leq q$. This should be obvious from considering, for example, a rule that selects the parameter corresponding to the largest of $m > 1$ independent estimators (here, $R_{CI} \equiv 1$). On the other hand, constructing the marginal confidence interval at level $1 - q/m$ trivially ensures FCR $\leq q$. Indeed, denoting as NCI_i the event $\{\theta_i \notin CI_i(q/m), i \in S\}$, we have

$$\text{FCR} = \text{E}[Q_{CI}] \leq \Pr\left(\cup_{i=1}^{m} NCI_i\right) \leq \Pr\left(\cup_{i=1}^{m} \{\theta_i \notin CI_i(q/m)\}\right) \leq q.$$

However, under the independence of the estimators, BY show that the Bonferroni adjustment is conservative, and a smaller increase in the confidence level is sufficient to ensure FCR $\leq q$. Specifically, they prove that the FCR is controlled at level q under the following scheme.

Definition 1. Level-q BY FCR-Adjusted Selective-CI Procedure

1. Apply the selection criterion $\mathcal{S}$ to obtain $\mathcal{S}(\mathbf{Y})$.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

2. For each selected parameter $\theta_i,\ i \in \mathcal{S}(\mathbf{Y})$, let

$$R_{\min}(\mathbf{Y}^{(i)}) = \min_y \left\{ \left| \mathcal{S}(\mathbf{Y}^{(i)}, Y_i = y) \right| : i \in \mathcal{S}(\mathbf{Y}^{(i)}, Y_i = y) \right\}, \quad (2.4)$$

where $\mathbf{Y}^{(i)}$ is the vector obtained by omitting Y_i from $\mathbf{Y}$.

3. For each selected parameter $\theta_i,\ i \in \mathcal{S}(\mathbf{Y})$, construct the following confidence interval:

$$CI_i \left(\frac{R_{\min}(\mathbf{Y}^{(i)}) \cdot q}{m} \right).$$

For many selection criteria, such as the step-up procedure of Benjamini and Hochberg, the term $\left| \mathcal{S}(\mathbf{Y}^{(i)}, Y_i = y) \right|$ is constant for all y such that $i \in \mathcal{S}(\mathbf{Y}^{(i)}, Y_i = y)$, implying that $R_{\min}(\mathbf{Y}^{(i)}) = R_{CI}$. In that case, to adjust the confidence intervals, we simply multiply the marginal level q by the number of parameters selected, and divide by m.

3. A Class of Selective-SDCI Procedures

In this section, we propose a general scheme to produce valid Selective-SDCI procedures, that relies on a marginal confidence interval. Starting with any marginal confidence interval, we show how a valid Selective-SDCI procedure can be obtained utilizing the FCR adjustment of Benjamini and Yekutieli (2005). We then discuss how the choice of the marginal confidence interval affects the resulting selective procedure.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

Suppose that $\{\mathcal{C}(\cdot;\alpha) : 0 \leq \alpha \leq 1\}$ is any marginal confidence interval procedure satisfying Requirement (MON 1) of the previous section, as well as

Requirement (MON 2) For any $0 \leq \alpha \leq 1$, $\mathcal{C}(-y;\alpha) = -\mathcal{C}(y;\alpha)$ and the lower boundary $l(y) = \inf\{\nu : \nu \in \mathcal{C}(y;\alpha)\}$ is increasing in $y > 0$.

Define a corresponding Selective-SDCI procedure as follows.

Definition 2. Level-q BY-Adjusted Selective-SDCI Procedure

1. Let $Y_{(i)}$ be the estimate with the ith largest absolute value; that is, $|Y_{(m)}| \leq |Y_{(m-1)}| \leq ... \leq |Y_{(1)}|$.

2. Denoting $CI_i(\alpha) = \mathcal{C}(Y_i;\alpha)$, find

$$R = \max\left\{r : CI_{(r)}\left(\frac{r \cdot q}{m}\right) \text{ is contained in } (-\infty, 0] \text{ or in } (0, \infty)\right\},$$

and let $\mathcal{S}^*(\mathbf{Y}) = \left\{i : |Y_i| \geq Y_{(R)}\right\}$ be the (possibly empty) set of selected parameters.

3. For each $i \in \mathcal{S}^*(\mathbf{Y})$, construct the confidence interval

$$CI_i\left(\frac{R \cdot q}{m}\right).$$

Theorem 1. *Suppose that $Y_i \sim f(y_i - \theta_i)$, for $i = 1, ..., m$, are independent, and let $\{\mathcal{C}(\cdot;\alpha) : 0 \leq \alpha \leq 1\}$ be a marginal confidence interval procedure*

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

satisfying Requirement (MON 1) and Requirement (MON 2). Then, the procedure in Definition 2 enjoys FCR $\leq q$.

Proof. We show that the procedure of Definition 2 uses the BY FCR-adjusted confidence level for the constructed CIs; in other words, the Selective-SDCI procedure is simply the BY procedure in Definition 1 for the selection rule $\mathcal{S}^*$ in Definition 2. This will complete the proof, because the level-q BY procedure has FCR $\leq q$ for any selection rule.

It remains to show that for the procedure in Definition 2, $R_{\min}(\mathbf{Y}^{(i)}) = R$; in other words, $|\mathcal{S}^*(\mathbf{Y}^{(i)}, Y_i = y)|$ is constant over y, for all y such that $i \in \mathcal{S}^*(\mathbf{Y}^{(i)}, Y_i = y)$. Indeed, if this is true, then the constructed intervals use the BY-adjusted level, and, therefore, FCR $\leq q$. This part is proved in the appendix. □

For a given marginal confidence interval $\mathcal{C}(y; \alpha)$, the procedure described in Definition 2 constructs the largest number possible of BY FCR-adjusted CIs that determine the sign. Because the set of discoveries is determined by the adjustment of a marginal CI, our procedure is completely characterized by the choice of $\mathcal{C}(y; \alpha)$. Therefore, this choice affects both the power of the procedure as a sign-classification rule—the expected (say) number of intervals constructed—and the shape of the constructed intervals. In particular, using a marginal interval $\mathcal{C}(y; \alpha)$ that determines the sign for

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

relatively small values of $|y|$ enhances the power. On the other hand, if a marginal interval $\mathcal{C}(y;\alpha)$ with a small (relative to other marginal CIs) maximum length is used, then the constructed CIs will enjoy a small (relative to other marginal CIs adjusted at the same level) maximum length, because only the confidence level is adjusted when constructing the intervals. The following examples describe the BY-adjusted Selective-SDCI procedure corresponding to three different choices of a marginal confidence interval. We assume here that $Y_i - \theta_i$ are independently and identically distributed (i.i.d.) $N(0,1)$, and denote $z_p = \Phi^{-1}(1-p)$.

(a) Symmetric confidence interval. Set $\mathcal{C}(y;\alpha) = (y - z_{\alpha/2}, y + z_{\alpha/2})$.

For any $\alpha \in (0,1)$, this CI includes values of one sign only (and possibly zero) whenever $z_{\alpha/2} \leq |y|$. Thus, the algorithm in Definition 2 selects the parameters corresponding to the $R = \max\{r : \{z_{r\cdot q/(2m)} \leq |Y_{(r)}|\}$ largest observations. Now, let $P_i = 2(1 - \Phi(|Y_i|))$ be the two-sided p-value for testing $H_{0i} : \theta_i = 0$, and let $P_{(1)} \leq P_{(2)} \leq ... \leq P_{(m)}$ be the ordered p-values (note that the subscript of the order statistic has the conventional meaning for the p-values but not for the estimators). Then, $R = \max\{r : P_{(r)} \leq r \cdot q/m\}$, and, thus, the selected parameters are exactly those corresponding to the hypotheses rejected by the BH procedure applied at level q. The constructed CI for each selected

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

parameter θ_i is $CI_i = Y_i \pm z_{R\cdot q/(2m)}$.

(b) One-sided CI . Take

$$\mathcal{C}(y;\alpha) = \begin{cases} (-\infty,\infty), & -z_\alpha < y < z_\alpha \\ (0,\infty), & z_\alpha \le y \\ (-\infty,0], & y \le -z_\alpha. \end{cases} \tag{3.5}$$

For any α, this CI includes values of one sign only already when $z_\alpha \le |y|$. Our procedure therefore selects the set of parameters corresponding to the $R = \max\left\{r : z_{r\cdot q/m} \le |Y_{(r)}|\right\} = \max\left\{r : P_{(r)} \le r \ \cdot(2q)/m\right\}$ largest observations, which is the set of parameters rejected by the BH procedure when applied at level $2q$. The constructed CI for each selected parameter θ_i, is $CI_i = (0,\infty)$ if $0 < Y_i$, and $CI_i = (-\infty, 0]$ if $Y_i < 0$.

For lack of a better term, we refer to the CI in (3.5) as "one-sided," although this name is usually reserved for a CI of the form $(y - z_\alpha, \infty)$ or $(-\infty, y + z_\alpha)$.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

(c) Pratt's CI. We can use a more sophisticated one-sided interval,

$$\mathcal{C}(y;\alpha) = \begin{cases} (y - z_\alpha, y + z_\alpha), & \text{if } |y| < z_\alpha \\ (0, y + z_\alpha), & \text{if } z_\alpha \le y \\ (y - z_\alpha, 0], & \text{if } y \le -z_\alpha \end{cases} . \tag{3.6}$$

This construction was suggested by Pratt (1961), who sought to minimize the expected length of a CI at $\theta = 0$. Pratt's interval still determines the sign at z_α but its length is finite when it determines the sign, in contrast to the usual one-sided interval. The resulting FCR-adjusted selective-SDCI procedure therefore still has $R = \max\left\{r : z_{r \cdot q/m} \le |Y_{(r)}|\right\}$, and selects according to a level-$2q$ BH procedure. However, the constructed CI for a selected parameter is now

$$CI_i = \begin{cases} (0, Y_i + z_{Rq/m}), & \text{if } z_{Rq/m} < Y_i \\ (Y_i - z_{Rq/m}, 0], & \text{if } z_{Rq/m} < Y_i, \end{cases}$$

instead of the infinitely long intervals constructed by the plain one-sided interval (3.5).

It is easy to verify that all of the aforementioned marginal CIs are valid

The original CI suggested by Pratt treats zero "symmetrically," whereas we append zero to the negative part of the line; (3.6) is therefore slightly different from the original construction, but the difference is not essential.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

(i.e., have $1-\alpha$ coverage) and satisfy the two monotonicity requirements (MON 1) and (MON 2).

For a fixed α, we would ideally want to equip the BY-adjusted Selective-SDCI procedure with a marginal interval $\mathcal{C}(y;\alpha)$ that determines the sign as early as possible, while at the same time, having the smallest possible (say, maximum) length. Unfortunately, these two requests are incompatible: early sign determination incurs longer CIs, at least for some values of y. This is demonstrated in the examples above. The two-sided marginal interval has the shortest possible maximum length, but determines the sign starting only at the $1-\alpha/2$ quantile. In contrast, the one-sided marginal interval determines the sign already at the $1-\alpha$ quantile, but has infinite maximum length. The Pratt interval improves on the length of the one-sided interval "for free," but its length is still unbounded in y, which is necessary if we require that signs are determined starting at the $1-\alpha$ quantile. Consequently, if we are to use the procedure described in Definition 2, then a trade-off between power and the maximum (potential) length of the constructed intervals is unavoidable.

Nevertheless, we are not limited to the marginal CIs in (a)–(c), in which the sign determination occurs at either of the two extremes, $F^{-1}(1-\alpha/2)$ or $F^{-1}(1-\alpha)$. Rather than insisting on the earliest possible sign determination

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

or the smallest possible maximum length, we may choose a marginal CI that balances the two factors. That is, we choose a marginal CI that determines the sign starting at a value slightly greater than the $1-\alpha$ quantile, and, in turn, has a maximum length that is only slightly larger than twice the $1-\alpha/2$ quantile. Equipped with this marginal family, the procedure of Definition 2 selects parameters according to a BH procedure at a level close to $2q$, while controlling the length of the constructed CIs.

Benjamini et al. (1998) suggested a nonequivariant marginal CI that achieves this purpose. They assume that $Y \sim f(y-\theta)$, with $f = F'$ a unimodal, symmetric density, and obtain their QC interval by inverting a family of acceptance regions. Specifically, the QC interval at y is defined as the convex hull of

$$\{\theta : y \in A_{QC}(\theta)\},$$

where

$$A_{QC}(\theta) = \begin{cases} (\theta - \bar{c}, \theta + \tilde{c}), & 0 < \theta \leq \bar{c} \\ (0, \theta + F^{-1}(1-\alpha+F(-\theta))), & \bar{c} < \theta \leq c_{\alpha/2} \\ (\theta - c_{\alpha/2}, \theta + c_{\alpha/2}), & c_{\alpha/2} < \theta \end{cases}, \tag{3.7}$$

and $A(\theta) = -A(-\theta)$ for $\theta < 0$. The acceptance region at zero is symmetric

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

in the original construction, but we take

$$A_{QC}(0) = (-\infty, c_\alpha),$$

which fits our (asymmetric) definition of sign determination. The constants $\bar{c}$ and $\tilde{c}$ are determined by a parameter $1/2 \leq \psi < 1$, and given by

$$\bar{c} = F^{-1}(1 - \psi\alpha) \qquad \tilde{c} = F^{-1}(1 - \alpha + F(-\bar{c})).$$

For any $p \in [0, 1]$, we write $c_p = F^{-1}(1-p)$ for the $(1-p)$-th quantile of F. The QC CI determines the sign for $|y| \geq \bar{c} \in (c_\alpha, c_{\alpha/2}]$, and can be shown to have a maximum length of $\tilde{c} + c_{\alpha/2} < \infty$. The parameter ψ controls the balance between the early sign determination and the maximum length of the QC interval. For $\psi = 1/2$, we have $\bar{c} = c_{\alpha/2}$, and the usual symmetric CI follows. When $\psi \to 1$, $\bar{c} \to c_\alpha$, and for any fixed y, the Pratt interval is obtained in the limit. As ψ increases from $1/2$ to 1, the sign determination occurs at a gradually earlier point at the cost of an increasing maximum length.

For any α, the QC interval with $1/2 \leq \psi < 1$ determines the sign at $\bar{c} < c_{\alpha/2}$. Thus, the BY-adjusted Selective-SDCI procedure using the QC interval will have more power than when using the symmetric interval. At the same time, the constructed intervals will be shorter than the Pratt CIs, and their length never exceeds $F^{-1}(1 - q' + F(-F^{-1}(1 - \psi q')))$ for

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

$q' = Rq/m$ (this is just $\tilde{c} + c_{\alpha/2}$ for $q = q'$). Although the QC interval already exhibits the features that enable our procedure to balance power and length, an improvement is possible. Indeed, will show that the QC interval can be slightly modified so that our procedure constructs shorter intervals at no expense.

4. A Modified Quasi-Conventional CI

In this section, we present a new marginal CI that adopts a feature from Finner (1994) to modify the QC interval of Benjamini et al. (1998). The idea is to take advantage of the fact that in a BY-adjusted Selective-SDCI procedure, only sign-determining CIs are ultimately constructed. Hence, inflating the QC CI whenever it includes values of opposite signs has no cost, but it allows us to construct shorter CIs when the sign is determined.

As in Benjamini et al. (1998), we make the further assumption that $f = F'$ is a unimodal density. We obtain the MQC interval by modifying the acceptance regions in (3.7). Hence, consider

$$A_{MQC}(\theta) = \begin{cases} (-\bar{c}, g(\theta)), & 0 < \theta \leq \bar{c} + c_{\alpha/2} \\ (\theta - c_{\alpha/2}, \theta + c_{\alpha/2}), & \bar{c} + c_{\alpha/2} < \theta \end{cases}, \tag{4.8}$$

with $A_{MQC}(\theta) = -A_{MQC}(-\theta)$, for $\theta < 0$, and $A_{MQC}(0) = (-\infty, c_\alpha)$. For

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

$1/2 \leq \psi < 1$,

$$\bar{c} = F^{-1}(1 - \psi\alpha) \qquad \tilde{c} = F^{-1}(1 - \alpha + F(-\bar{c})),$$

and

$$g(\theta) = \theta + F^{-1}\{1 - \alpha + F(-\bar{c} - \theta)\}.$$

As before, ψ is a parameter that controls how early the CI determines the sign of θ, and is chosen in advance.

The MQC interval is obtained as the convex hull of $\{\theta : y \in A_{MQC}(\theta)\}$. Inverting the family of acceptance regions in (4.8) is more complicated than it is for the QC acceptance regions, because we need to distinguish between three cases: (i) $0 < \psi \leq \psi_1$ (ii) $\psi_1 < \psi \leq \psi_2$, and (iii) $\psi_2 < \psi$. Here,

$$\psi_1 = \psi_1(\alpha) \text{ is the value of } \psi, \text{ such that } \tilde{c} = 2\bar{c} + c_{\alpha/2}$$

$$\psi_2 = \psi_2(\alpha) \text{ is the value of } \psi, \text{ such that } \tilde{c} = \bar{c} + 2c_{\alpha/2}.$$

From a practical point of view, however, at least when f is the standard normal density, ψ_1 tends to be very close to one. For example, when f is the standard normal density and $\alpha = 0.1$, $\psi_1 > 0.999$, and it is even closer to one for smaller α. This means that for typical, small values of α, unless ψ is chosen extremely close to one, we fall within case (i) above. For clarity, we present the confidence bounds for the first case only; a full specification of the CI, that includes the other two cases, is provided in the Supplementary

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

Material. Hence, for $0 < \psi \leq \psi_1$, the convex hull of $\{\theta : y \in A_{MQC}(\theta)\}$ is given by

$$\mathcal{C}_{MQC}(y;\alpha) = \begin{cases} (-\bar{c} - c_{\alpha/2}, \bar{c} + c_{\alpha/2}), & 0 \leq y < \bar{c} \\ (0, y + c_{\alpha/2}), & \bar{c} \leq y < \tilde{c} \\ (g^{-1}(y), y + c_{\alpha/2}), & \tilde{c} \leq y \leq g(\bar{c} + c_{\alpha/2}) \\ (\bar{c} + c_{\alpha/2}, y + c_{\alpha/2}), & g(\bar{c} + c_{\alpha/2}) < y < \bar{c} + 2c_{\alpha/2} \\ (y - c_{\alpha/2}, y + c_{\alpha/2}), & \bar{c} + 2c_{\alpha/2} \leq y, \end{cases} \tag{4.9}$$

with $\mathcal{C}(-y;\alpha) = -\mathcal{C}(y;\alpha)$. In (4.9), $g^{-1}(t)$ is well defined because g is strictly increasing to ∞ on $-\bar{c} + c_{\alpha/2} < t$, and in particular, on $\tilde{c} < t$. The assumption that f is unimodal (and symmetric) ensures that (4.9) is indeed the convex hull of $\{\theta : y \in A_{QC}(\theta)\}$.

Remark 1. The MQC interval is scale-invariant in the following sense: if $Y \sim (\theta, \sigma^2)$ and $Y' = Y/\sigma$, and $\mathcal{C}(y;\alpha)$ and $\mathcal{C}'(y';\alpha)$ are the MQC CIs (for any fixed ψ) based on Y and Y', respectively, then $\mathcal{C}(y;\alpha) = \sigma \cdot \mathcal{C}'(y/\sigma;\alpha)$.

The QC and MQC intervals determine the sign of θ starting at the same value of $|y|$, but the latter constructs shorter intervals on a subset of $\{y : |y| > \bar{c}\}$ at the expense of wider intervals for all $|y| < \bar{c}$. On this subset, for each of the three cases, the lower endpoint is farther away from zero; in

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

the last two cases—that is, when $\psi_1 < \psi$—there is a discontinuity point for the lower bound just as the CI separates from zero ($y = \tilde{c}$).

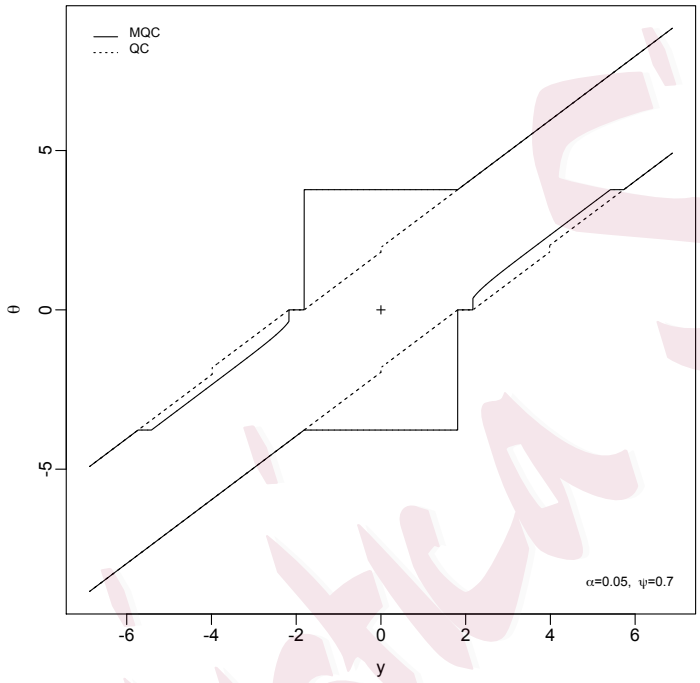


Figure 1: MQC interval vs. the QC interval of Benjamini et al.. The plot is for $\alpha = .05$ and $\psi = 0.7$. Both CIs (weakly) determine the sign of θ whenever $|y| \geq \bar{c} = 1.81$. When the sign is determined, the MQC bounds are farther away from zero for a range of y values that begins when the CI separates from zero. The interval around zero where the MQC confidence limits are constant in y, is the region where the QC interval includes both negative and positive values.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

The BY-adjusted Selective-SDCI procedure, equipped with any marginal CI that satisfies requirements (MON 1) and (MON 2), has FCR $\leq q$. The actual FCR level depends on the marginal CI employed. For the two-sided CI—that is, for the BH-selected BY-adjusted procedure—Benjamini and Yekutieli (2005) show that the FCR is also lower bounded by $q/2$. We show a similar result for the MQC interval when the estimators are normally distributed.

Theorem 2. *For independent, normally distributed estimators with a known variance, the Selective-SDCI procedure of Definition 2 using the MQC interval, with* $0 < \psi < 0.9$*, enjoys FCR* $\geq q/2$ *if* $0 < q < 0.25$.

Although Theorem 2 asserts that, under the stated conditions, using the MQC interval ensures $FCR \geq q/2$, it is typically close to q. Indeed, for standard normal observations and $0 < \alpha < 0.25$ and $0 < \psi < 0.9$, the probability in (A.11) of the Appendix is approximately q for all θ, except for a small region where it may decrease to as low as $\alpha/2$. For example, when $\alpha = .01$ and $\psi = .85$, as long as $|\theta| \notin (0, .48)$ and $|\theta| \notin (6.43, 7.4)$, the probability in (A.11) is at least 0.99α. Hence, the inequality in (A.17) can often be made much tighter, and FCR $\approx q$. Note that if the original QC interval is used instead of the MQC interval in the Selective-SDCI procedure, the FCR may fall significantly below $q/2$, as demonstrated in

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

the simulation in Section 6.

5. Sign Determination by Confidence Regions

In this section, we extend our methodology from the one-dimensional case to the case where $\boldsymbol{\theta}_i \in \mathbb{R}^k$, with $k > 1$. In principle, it is possible to classify $\boldsymbol{\theta}_i$ as having one of 2^k possible signs. Here, we consider a straightforward extension of the one-dimensional case in which $\boldsymbol{\theta}_i = (\boldsymbol{\theta}_{i1}, \theta_{i2})$, with $\boldsymbol{\theta}_{i1} \in \mathbb{R}^{k-1}$ and $\theta_{i2} \in \mathbb{R}^1$, where we try to classify $\boldsymbol{\theta}_i$ according to the sign of θ_{i2}. In Section S1 of the Supplementary Material, we apply this methodology to classify the signs of the genetic associations of almost half a million SNPs, and then to assess whether the effect of the SNP is recessive or dominant.

Let $\mathbf{Y}_i = (Y_{i1}, Y_{i2})$, with independent $Y_{i1} \sim f(y_{i1} - \theta_{i1})$ and $Y_{i2} \sim f(y_{i2} - \theta_{i2})$. $CI_{i1}(\alpha) = CI_{i1}(\alpha; Y_{i1})$ is a marginal $1 - \alpha$ CI for θ_{i1} and $CI_{i2}(\alpha) = CI_{i2}(\alpha; Y_{i2})$ is a marginal $1 - \alpha$ CI for θ_{i2} (assume that the coverage probability is exactly $1 - \alpha$, not more). We use $CI_{i1}(\alpha_1)$ and $CI_{i2}(\alpha_2)$ to form a $1 - \alpha_1 \cdot \alpha_2$ confidence set for $\boldsymbol{\theta}_i$,

$$\widetilde{CI}_i(\alpha_1, \alpha_2) = \{\boldsymbol{\theta}_i : \ \theta_{i1} \in CI_{i1}(\alpha_1), \theta_{i2} \in CI_{i2}(\alpha_2) \ \}.$$

For independent $\mathbf{Y}_1, ..., \mathbf{Y}_m$ and a selection rule that has $R_{\min}(\mathbf{Y}^{(i)}) \equiv R_{CI}$, BY show that the FCR is equal to $\sum_{r=1}^{m} \sum_{i=1}^{m} \Pr(|\mathcal{S}(\mathbf{Y})| = r, \tilde{NCI}_i)/r$, where $\tilde{NCI}_i$ is the event that $\boldsymbol{\theta}_i$ is selected and $\boldsymbol{\theta}_i \notin \widetilde{CI}_i$. We denote by

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

$N\tilde{C}I_{i1}$ the event that $\boldsymbol{\theta}_i$ is selected and $\theta_{i1} \notin CI_{i1}$, and by $N\tilde{C}I_{i2}$ the event that $\boldsymbol{\theta}_i$ is selected and $\theta_{i2} \notin CI_{i2}$. Thus, $N\tilde{C}I_i = N\tilde{C}I_{i1} \cup N\tilde{C}I_{i2}$, and to evaluate FCR we express $N\tilde{C}I_i$ as the disjoint union

$$N\tilde{C}I_i = N\tilde{C}I_{i1} \cup (\{\theta_{i1} \in CI_{i1}\} \cap N\tilde{C}I_{i2}).$$

We consider selection rules $\mathcal{S}_2(\mathbf{Y}_{\bullet 2})$ that are determined by $\mathbf{Y}_{\bullet 2} = (Y_{12}, ..., Y_{m2})$ only.

Definition 3. Level-(q_1, q_2) FCR adjustment for selection rules determined by $\mathbf{Y}_{\bullet 2}$

1. Apply the selection criterion $\mathcal{S}_2$ to obtain $\mathcal{S}_2(\mathbf{Y}_{\bullet 2})$.

2. For each selected parameter $\boldsymbol{\theta}_i,\ i \in \mathcal{S}_2(\mathbf{Y}_{\bullet 2})$, let

$$R_{\min}(\mathbf{Y}_{\bullet 2}^{(i)}) = \min_t \left\{ \left| \mathcal{S}(\mathbf{Y}_{\bullet 2}^{(i)}, Y_{i2} = t) \right| : i \in \mathcal{S}(\mathbf{Y}_{\bullet 2}^{(i)}, Y_{i2} = t) \right\}, \quad (5.10)$$

   where $Y_{\bullet 2}^{(i)}$ is the vector obtained by omitting Y_{i2} from $\mathbf{Y}_{\bullet 2}$.

3. For each selected parameter $\boldsymbol{\theta}_i,\ i \in \mathcal{S}_2(\mathbf{Y}_{\bullet 2})$, construct the following CI:

$$\widetilde{CI}_i \left(q_1, \frac{R_{\min}(\mathbf{Y}_{\bullet 2}^{(i)}) \cdot q_2}{m} \right).$$

Theorem 3. *Let $\mathbf{Y}_1, ..., \mathbf{Y}_m$ be independent, where $\mathbf{Y}_i = (Y_{i1}, Y_{i1})$ for independent Y_{i1} and Y_{i2}. Then, the FCR of the level-(q_1, q_2) adjusted confidence*

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

sets of Definition 3 for $\mathcal{S}_2(\mathbf{Y}_{\bullet 2})$ *is*

$$FCR(\widetilde{CI}_{\bullet}; \mathcal{S}; q_1, q_2) = q_1 + (1 - q_1) \cdot FCR(CI_{\bullet 2};\ \mathcal{S}_2;\ q_2),$$

where $FCR(CI_{\bullet 2};\ \mathcal{S}_2;\ q_2)$ *is the FCR of the level* q_2 *BY FCR-adjusted CI for* $\mathcal{S}_2$.

6. Simulation Study

We conducted two simulations that demonstrate the performance of the BY-adjusted Selective-SDCI procedure using the MQC interval. Additionally, in Section S2 of the Supplementary Material, we report the results of a simulation in which we examined the Selective-SDCI procedures under dependency. The first simulation illustrates the asymmetric shape of the MQC intervals and its increased power over the BH directional procedure to classify the sign of parameters. We took $m = 200$ parameters, where $\theta_1, ..., \theta_{160}$ were sampled from an exponential distribution with mean 0.5, and $\theta_{161}, ..., \theta_{200}$ were sampled from a $N(3, 1)$ distribution. Each θ_i was then randomly assigned a positive or a negative sign. The independent observations are $Y_1, ..., Y_{200}$, with $Y_i \sim N(\theta_i, 1)$. Figure 2 shows the constructed intervals for positive θ_i when the procedure of Definition 2 is equipped with the MQC interval ($\psi = 0.85$) and applied at level $q = 0.2$. A total of 74 sign-determining CIs were constructed, 32 of them for positive

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

observations. The number of parameters selected is almost as large as the number selected with a BH directional procedure at level $2q = 0.4$ (77) and much larger than a BH directional procedure at level $q = 0.2$ (55). At the same time, the MQC constructed CIs (vertical segments in the figure) are relatively short—mostly even shorter than the symmetric FCR-adjusted CIs for level-q BH-selected parameters (partly because additional parameters are selected). Of the 74 constructed CIs, 14 did not cover the respective parameter (six of which for positive observations), yielding a proportion of 0.19. The procedure using the QC interval ($\psi = 0.85$) instead of the MQC, constructed the same number of intervals with a false coverage proportion of 0.15.

The second simulation compares the (actual) FCR of the MQC-equipped procedure with that of the QC-equipped procedure. We first sampled $\theta_1, ..., \theta_{300}$ from a $N(0, 4)$ distribution. For $N = 10^4$ data sets $\mathbf{Y} = (Y_1, ..., Y_{300})$, with $Y_i \sim N(\theta_i, 1)$, we computed the FCP (denoted by Q_{CI} in section 2) for a level $q = 0.05$ BY-adjusted Selective-SDCI procedure using the MQC interval ($\psi = 0.85$), and for the same procedure using a QC interval ($\psi = 0.85$). We used $\psi = 0.85$ for both the QC and the MQC intervals so that the sign determination occurs at the same value for both intervals. The average FCP for QC was 0.018 ($\widehat{SD} = 1.3 \cdot 10^{-4}$), and for MQC it was

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

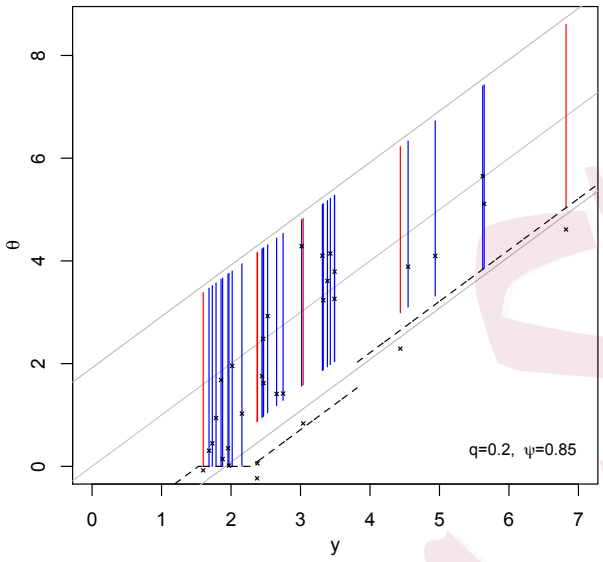


Figure 2: MQC sign-determining CIs. Level $q = 0.2$ BY-adjusted sign-determining MQC CIs were constructed for 74 out of $m = 300$ parameters, 14 of the CIs do not cover the respective parameter. Vertical lines display MQC-adjusted CIs for the 32 positive observations: 26 of them cover the respective parameter (blue bars), and six do not (red bars). The diagonal line running through the origin is the identity line. Markers denote pairs (Y_i, θ_i). Solid gray lines mark the position of level 0.2 FCR-adjusted two-sided CIs for level 0.2 BH-selected parameters. The broken line represents the lower boundary of the constructed QC intervals.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

0.048 ($\widehat{SD} = 2.2 \cdot 10^{-4}$). These results confirm that, as discussed in the previous section, the FCR when using the MQC interval is often very close to q, whereas it may fall below $q/2$ when using the QC interval.

7. Detecting the Signs of Correlations in a Social Neuroscience Study

Tom et al. (2007) carried out an experiment in an attempt to associate neural activity in the brain with behavioral "loss aversion." Their study received high publicity, and the collected data were re-analyzed by Poldrack and Mumford (2009) and Rosenblatt and Benjamini (2014). The original data were made available through the *OpenfMRI* initiative at `https://openfmri.org/dataset/ds000005`, and described in detail in the paper by Tom et al.. For each of 16 subjects, a behavioral loss-aversion index was measured, along with a neural index at each brain voxel. The voxel-specific correlations between the behavioral index and the neural index were then used to detect brain regions associated with loss aversion. Rosenblatt and Benjamini (2014) revisited this data set, exploring different methods of constructing CIs that account for the selection bias in reported voxels. Their approach is, in general, to employ a two-stage procedure. The first stage is, in principle, designed to detect nonzero correlations. In the second

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

stage, they construct a CI for each parameter selected (rejected) at the first stage, while attempting to control the FCR below some prespecified level.

Specifically, one of the schemes they used is selection via the BH procedure. Because we are interested in sign classification rather than two-sided testing, we view the BH procedure here as a directional procedure, that is, as a procedure that classifies the sign of each reported parameter as strictly positive or strictly negative. If we are willing to settle for weak (rather than strict) sign determination, our method suggests an alternative that tends to discover additional parameters. Thus, we apply our method to the z-scores computed for each voxel of the Fisher-transformed correlations. The data were processed by Rosenblatt and Benjamini (2014) and kindly made available to us. The concern about validity of our procedure under dependency, which is likely to be present in the current example, is mitigated by the simulation results from Section S2 of the Supplementary Material.

A level 0.1 directional-BH procedure applied to the two-sided p-values found 18,844 voxels for which a strict sign decision could be made. However, a level 0.1 BY-adjusted Selective-SDCI procedure using the MQC interval with $\psi = 0.85$ was able to weakly classify the signs of a total of 36,131 correlations, of which 27,117 were strict sign classifications. In comparison, the BY-adjusted Selective-SDCI procedure using a one-sided (or Pratt's)

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

CI, which selects according to a BH procedure at level 0.2, reports 43,804 parameters, where all signs are weakly classified. Hence, the BH at half the level makes 57% fewer discoveries, all with a strict sign classification, whereas the MQC-equipped BY-adjusted Selective-SDCI at half the level makes only 18% fewer discoveries, the majority of them with a strict sign classification. Figure 3a displays the MQC CIs constructed for the 33,856 correlations classified as positive, along with the QC intervals. The symmetric intervals corresponding to the selection according to a level 0.1 BH procedure are also shown for reference. The figure shows that the majority of the discoveries, the lower endpoint of the MQC interval is farther away from zero than that of the QC interval, even though the latter yields the same set of discoveries. Note that the gap between the lower endpoint of the MQC CI (black points in the figure) and the lower endpoint of QC CI (gray line in figure), is largest immediately as the two intervals separate from the horizontal axis. This is precisely where we would like the gap to be largest, because it is more important to be able to quote an endpoint farther from zero for a small detected correlation than it is for a very large detected correlation.

Note that the intervals constructed by the Selective-SDCI procedure using any of the configurations (i.e., any of the marginal CIs) above are sign-

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

determining. Hence, this is a partial response to the request of Rosenblatt and Benjamini (2014), who comment that "it might be of interest to develop CIs that are dual to the selection methods used in neuroimaging."

Rather than detecting positive or negative correlations, it is reasonable to assume that a researcher would be interested in detecting the large correlations, be they positive or negative. In Section S5 of the Supplementary Material, we extend the BY-adjusted Selective-SDCI procedure to detect correlations $\rho_i > \rho_0$ or $\rho_i < -\rho_0$, for some prespecified constant $\rho_0 \in (0, 1)$, and supplement the decisions with compatible CIs. Figure 3b displays the 9 constructed BY-adjusted Selective CIs for $\rho_0 = 0.2$.

8. Discussion

"Selective inference" refers to the general situation where the target of inference is chosen adaptively (i.e., only after seeing the data). We concentrated on a setup where selective inference arises in connection with multiplicity. The analyst collects noisy observations on a (typically large) number m of unknown parameters, which he will first use to try to answer a primary question about each parameter. Then, he constructs CIs for only those parameters for which there was enough evidence to answer the primary question. Specifically, we considered the problem of detecting the signs

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

of parameters, and supplementing each directional decision made with a CI. Because the same data are used for detection and for constructing the follow-up CIs, selection needs to be accounted for.

Requiring weak directional-FDR control at the first stage and FCR control at the second stage, a natural approach to the problem is to treat the two stages separately. Thus, we can first apply the directional-BH procedure to select a subset of parameters, the signs of which will be classified. Then, for each selected parameter, construct a CI that is valid *conditionally* on selection (e.g., using the methods of Weinstein et al., 2013). Constructing conditional CIs is appealing from various aspects, one of which is that a conditional CI (usually) has the property of converging to the unadjusted CI for a large value of the observation. However, there is a drawback to constructing conditional CIs: we cannot guarantee that the CI is compatible with the directional decision of the first stage. In other words, there is no conditional CI that, for all values of the parameter and with probability one, includes either only positive or only nonpositive values. See Section S3 of the Supplementary Material, where we discuss connections to existing work on post-selection inference. Hence, to ensure compatibility of the follow-up CI with the directional decision, we combine the two inferential goals by requiring simply the construction of (a selective set of)

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

sign-determining CIs.

Although our focus was predominantly on the sign problem, the approach we suggest is quite general. For example, in the Supplementary Material (Section S5), we show how to modify the procedure of Definition 2 so that, instead of sign classification, the primary goal is to detect parameters larger than δ or smaller than $-\delta$. In general, suppose that the data are $\mathbf{Y}_i \overset{ind}{\sim} f(\mathbf{y}; \boldsymbol{\theta}_i)$, for $i = 1, ..., m$, where $\mathbf{Y}_i \in \mathcal{Y}$ and $\boldsymbol{\theta}_i \in \Theta$. Let $\Theta_j \subseteq \Theta$, for $j = 1, ..., k$, be disjoint subsets in the parameter space. The primary task is to detect the membership of $\boldsymbol{\theta}_i$ in any of the Θ_j; the secondary task is to construct a confidence set C_i for each classified parameter, such that $C_i \subseteq \Theta_j$ if $\boldsymbol{\theta}_i$ was classified to Θ_j. For hypothesis testing, $k = 1$ and Θ_1 is the set of alternatives; for the (weak) sign problem, $k = 2$, $\Theta_1 = (0, \infty)$, and $\Theta_2 = (-\infty, 0]$; the example in Section S5 of the Supplementary Material corresponds to $k = 2$, $\Theta_1 = (\delta, \infty)$, and $\Theta_2 = (-\infty, -\delta)$; in Section S1 of the Supplementary Material, $\mathcal{Y} = \Theta = \mathbb{R}^2$, $k = 2$, $\Theta_1 = (-\infty, \infty) \times (0, \infty)$, and $\Theta_2 = (-\infty, \infty) \times (-\infty, 0]$. In principle, the extension of the procedure in Definition 2 to the general case would be to construct the maximum number of FCR-adjusted confidence sets, such that each confidence set is contained in one of the subsets Θ_j, for $j = 1, ..., k$.

There are certainly remaining challenges. When $Y_i \sim N(\theta_i, \sigma^2)$ and σ is

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

unknown, Finner (1994, Section 3) pointed out that a disadvantage of CIs based on the t-statistic, is that they are unbiased for the (natural) parameter θ_i/σ, not θ_i, and suggested an alternative CI that improves uniformly over the t procedure. It might be of interest to modify Finner's CI to produce an interval with similar properties to those of the MQC interval; for the corresponding procedure of Definition 2 to be valid, the monotonicity requirements would need to be checked, which might not be trivial. Another direction worth exploring is constructing sign-determining CIs for the coefficients β_j in a linear regression model; Barber and Candes (2016) address sign classification under directional-FDR control in a Gaussian linear model. Supplementing such directional decisions with compatible confidence bounds is of clear practical importance.

Lastly, we believe it is important that we establish a benchmark against which our procedure can be evaluated: while our procedure balances the power and length of constructed CIs, it is indexed by a single scalar parameter (ψ). Thus, it is natural to ask whether more can be gained—for example, in the form of shorter CIs—by allowing greater flexibility in constructing selective CIs that determine the sign. To be able to compare procedures, a reasonable option is to set up a formal criterion that takes into account both the power and the shape of the constructed intervals.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

Supplementary Material

The online Supplementary Material includes an application of the methods of Section 5 to a genomic example, further simulation studies under dependency of the observations, a discussion of related existing work on selective inference, where we contrast the conditional approach with ours, a full specification of the MQC interval of Section 4, and an extension of the Selective-SDCI procedure for detecting only large correlations, with an application to the example of Section 7.

Acknowledgments

The work of DY was supported by ISF grant 1049/16. Section S1 of the Supplementary Material uses data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available from `www.wtccc.org.uk`. Funding for the project was provided by the Wellcome Trust under award 076113.

References

Barber, R. F. and E. J. Candes (2016). A knockoff filter for high-dimensional selective inference. *arXiv preprint arXiv:1602.03574*.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

Benjamini, Y., Y. Hochberg, and Y. Kling (1993). False discovery rate control in pairwise comparisons. *Research Paper, Department of Statistics and Operations Research, Tel Aviv University*.

Benjamini, Y., Y. Hochberg, and P. Stark (1998). Confidence intervals with more power to determine the sign: Two ends constrain the means. *Journal of the American Statistical Association 93*(441), 309–317.

Benjamini, Y. and D. Yekutieli (2005). False discovery rate–adjusted multiple confidence intervals for selected parameters. *Journal of the American Statistical Association 100*(469), 71–81.

Bohrer, R. (1979). Multiple three-decision rules for parametric signs. *Journal of the American Statistical Association 74*(366a), 432–437.

Bohrer, R. and M. J. Schervish (1980). An optimal multiple decision rule for signs of parameters. *Proceedings of the National Academy of Sciences 77*(1), 52–56.

Finner, H. (1994). Two-sided tests and one-sided confidence bounds. *The Annals of Statistics*, 1502–1516.

Fithian, W., D. Sun, and J. Taylor (2014). Optimal inference after model selection. *arXiv preprint arXiv:1410.2597*.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

Neyman, J., S. Gupta, and D. Moore (1977). Synergistic effects and the corresponding optimal version of the multiple comparison problem. *Statistical Decision Theory and Related Topics 2*, 297–311.

Poldrack, R. A. and J. A. Mumford (2009). Independence in roi analysis: where is the voodoo? *Social Cognitive and Affective Neuroscience 4*(2), 208–213.

Pratt, J. W. (1961). Length of confidence intervals. *Journal of the American Statistical Association 56*(295), 549–567.

Rosenblatt, J. D. and Y. Benjamini (2014). Selective correlations; not voodoo. *Neuroimage 103*, 401–410.

Tian, X., J. Taylor, et al. (2018). Selective inference with a randomized response. *The Annals of Statistics 46*(2), 679–710.

Tom, S. M., C. R. Fox, C. Trepel, and R. A. Poldrack (2007). The neural basis of loss aversion in decision-making under risk. *Science 315*(5811), 515–518.

Weinstein, A., W. Fithian, and Y. Benjamini (2013). Selection adjusted confidence intervals with more power to determine the sign. *Journal of the American Statistical Association 108*(501), 165–176.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

Appendix

A. Proofs

A.1 A proof that $R_{\min}(\mathbf{Y}^{(i)}) = R_{CI}(\mathbf{Y})$ in Theorem 1

Without loss of generality, we show that $|\mathcal{S}^*(\mathbf{Y}^{(1)}, Y_1 = y)|$ is constant over y for all y is such that $i \in \mathcal{S}^*(\mathbf{Y}^{(1)}, Y_1 = y)$. Let

$$g(\alpha) = \inf\{y \geq 0 : \mathcal{C}(y; \alpha) \text{ includes values of one sign only}\},$$

and let $\tau(i) = g\left(\frac{i}{m}q\right),\ i = 1, ..., m$. Recall that $Y_{(i)}$ is the estimate with the i-th largest absolute value, hence $|Y_{(1)}| \geq |Y_{(2)}| \geq \cdots \geq |Y_m|$.

Because $\mathcal{C}(y; \alpha)$ satisfies the monotonicity requirements (MON 1) and (MON 2), $i^* = \max\left\{i : \tau(i) \leq Y_{(i)}\right\}$ and $\tau(i)$ is a decreasing sequence. Define now a vector $\tilde{\mathbf{Y}} = (\tilde{\mathbf{Y}}^{(1)}, \tilde{Y}_1)$, which depends on $\mathbf{Y}^{(1)}$ only, by $\tilde{\mathbf{Y}}^{(1)} = \mathbf{Y}^{(1)}$, $\tilde{Y}_1 = \infty$. Let $\tilde{Y}_{(i)}$ the element among $\tilde{Y}_1, ..., \tilde{Y}_m$ with the i-th largest absolute value. Furthermore, let

$$\tilde{i}^* = \max\{1 \leq i \leq m : \tau(i) \leq |\tilde{Y}_{(i)}|\}.$$

We will show that if $1 \in \mathcal{S}^*(\mathbf{Y})$ then $i^* = \tilde{i}^*$, hence if $1 \in \mathcal{S}^*(\mathbf{Y})$ then $|\mathcal{S}^*(\mathbf{Y})| = i^*$, which does not depend on y.

First, note that $1 \in \mathcal{S}^*(\mathbf{Y}) \iff \tau(\tilde{i}^*) \leq |y|$. Indeed, suppose that $y < \tau(\tilde{i}^*)$. For all $i \geq \tilde{i}^*$, $|Y_{(i)}| \leq \min(|\tilde{Y}_{(i)}|, \tau(\tilde{i}^*))$. Therefore, for all $i \geq \tilde{i}^*$,

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

$|Y_{(i)}| < \tau(i)$, which together with the fact that $\tau(i)$ is decreasing implies that $i \notin \mathcal{S}^*(\mathbf{Y})$ if $|Y_i| < \tau(\tilde{i}^*)$. In particular, $1 \notin \mathcal{S}^*(\mathbf{Y})$. On the other hand, if $\tau(\tilde{i}^*) \leq |y|$, then $|Y_{(\tilde{i}^*)}| = \min(|\tilde{Y}_{(\tilde{i}^*)}|, |y|) \geq \tau(\tilde{i}^*)$, which together with the fact that $\tau(i)$ is decreasing implies that $i \in \mathcal{S}^*(\mathbf{Y})$ if $\tau(\tilde{i}^*) \leq |Y_i|$. In particular, $1 \in \mathcal{S}^*(\mathbf{Y})$.

To complete the proof, observe that when $\tau(\tilde{i}^*) \leq |y|$, (i) $|Y_{(i)}| < \tau(i)$ for $i > \tilde{i}^*$, which implies $i^* \leq \tilde{i}^*$, and (ii) $|Y_{(\tilde{i}^*)}| = \min(|\tilde{Y}_{(\tilde{i}^*)}|, |y|) \geq \tau(\tilde{i}^*)$, which implies that $i^* \leq \tilde{i}^*$. We conclude that $i^* = \tilde{i}^*$, as required.

A.2 Proof of Theorem 2

By the remark in Section 4, it is enough to prove the theorem for the case $\mathrm{Var}(Y_i) = 1$. Indeed, for $\sigma^2 = \mathrm{Var}(Y_i)$, letting $Y_i' = Y_i/\sigma$ and $\theta_i' = \theta_i/\sigma$ we have that $\theta_i \notin \mathcal{C}(Y_i; \alpha) \iff \theta_i' \notin \mathcal{C}'(Y_i'; \alpha)$ where $\mathcal{C}(y; \alpha)$ and $\mathcal{C}'(y'; \alpha)$ are the MQC CIs corresponding to the distributions of Y and Y', respectively. Therefore the FCR of the procedure defined for the Y_i (w.r.t. the θ_i) is the same as the procedure defined for the Y_i' (w.r.t. the θ_i').

First we claim that for $\psi < 0.9$, the MQC interval is given by (4.9) for all $0 < \alpha < 0.25$. We need to check that $\psi_1 > 0.9$ for all $0 < \alpha < 0.25$. It can be verified that ψ_1 is a decreasing function of α on $0 < \alpha < 0.25$, and we have $\psi_1 = 0.978 > 0.9$, which together imply that $0.9 < \inf\{\psi_1 : 0 < \alpha < 0.25\}$

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

as required.

Let $0 < \alpha < 0.25$ and $0 < \psi < 0.9$. We now consider a single parameter, θ, and a corresponding estimator $Y \sim N(\theta, 1)$, and show that the probability that a sign-determining noncovering confidence interval is constructed for θ, is no less than $\alpha/2$ for all θ. Formally, let NCI be the event that $CI := \mathcal{C}_{MQC}(Y; \alpha)$ (i) determines the sign, i.e., does not include values of opposite signs and (ii) does not include the true value θ. Then we show that $\Pr_\theta(NCI) \geq \alpha/2$ for all θ. Since for the MQC interval, sign determination occurs if and only if $|Y| \geq \bar{c}$, we have

$$\Pr_\theta(NCI) = \Pr_\theta(|Y| \geq \bar{c},\ \theta \notin CI). \tag{A.11}$$

If the confidence interval were obtained simply by inverting the $1-\alpha$ acceptance regions $A(\theta)$ in (4.8) the event $\theta \notin CI$ could be replaced by $Y \notin A(\theta)$; however, the confidence interval is obtained by taking the convex hull of the inverse set, in which case it is possible that $Y \notin A(\theta)$ and yet $\theta \in CI$. We can overcome this difficulty by considering the "effective" acceptance regions, $\bar{A}(\theta)$, which take into account the fact that the convex hull of $\{\theta : Y \in A(\theta)\}$ is taken, in that $CI = \{\theta : Y \in \bar{A}(\theta)\}$ (here without the convex hull). Denoting by $l(\theta)$ and $u(\theta)$ the lower and upper endpoints of $A(\theta)$, respectively, and denoting by $\bar{l}(\theta)$ and $\bar{u}(\theta)$ the lower and upper ends of $\bar{A}(\theta)$, respectively, it holds that $\bar{l}(\theta) = \max\{u(\tilde{\theta}) : \tilde{\theta} \leq \theta\}$ and

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

$\bar{u}(\theta) = \min\{l(\tilde{\theta}) : \tilde{\theta} \geq \theta\}$. Explicitly,

$$\bar{A}(\theta) = \begin{cases} (-c_{\alpha/2}, c_{\alpha/2}), & \theta = 0 \\ (-\bar{c}, \tilde{c}), & 0 < \theta \leq \tilde{c} - \bar{c} \\ (-\bar{c}, g(\theta)), & \tilde{c} - \bar{c} < \theta \leq \bar{c} + c_{\alpha/2} \\ (\theta - c_{\alpha/2}, \theta + c_{\alpha/2}), & \bar{c} + c_{\alpha/2} < \theta \end{cases} \tag{A.12}$$

with $\bar{A}(\theta) = -\bar{A}(-\theta)$ for $\theta < 0$ and where $g(\theta) = \theta + F^{-1}\{2 - \alpha - F(\bar{c} + \theta)\}$.

Now we can write

$$\Pr_{\theta}(NCI) = \Pr_{\theta}(|Y| \geq \bar{c},\ Y \notin \bar{A}(\theta)), \tag{A.13}$$

and we note that for $0 < \theta < \bar{c} + c_{\alpha/2}$, $(-c, c) \subset \bar{A}(\theta)$, hence $\Pr_{\theta}(NCI) = \Pr_{\theta}(Y \notin \bar{A}(\theta))$. For $\theta = 0$, this is exactly α.

For $0 < \theta < \tilde{c} - \bar{c}$, $\Pr_{\theta}(Y \notin \bar{A}(\theta)) = \Pr_{\theta}(Y \notin (-\bar{c}, \tilde{c}))$, which is minimized at $\theta = (\tilde{c} - \bar{c})/2$. In order that $\Pr_{(\tilde{c}-\bar{c})/2}(Y \notin \bar{A}(\theta))$ be less than $\alpha/2$, in which case $\Pr_{(\tilde{c}-\bar{c})/2}(NCI) < \alpha/2$, it must hold that $\tilde{c} + \bar{c} > 2c_{\alpha/4}$. We claim that this cannot be the case. Hence, for any α, let ψ^* be the value of ψ for which $\tilde{c} + \bar{c} = 2c_{\alpha/4}$. Then for a fixed α, $\psi < \psi^*$ implies that $\tilde{c} + \bar{c} < 2c_{\alpha/4}$. Now, it can be verified that $\lim_{\alpha \to 0} \psi* > 0.9$ (but $\lim_{\alpha \to 0} \psi* < 0.94$) and that ψ^* is an increasing function of α on $0 < \alpha < 0.25$, which imply that $\psi^* > 0.9$ for all $0 < \alpha < 0.25$. It follows that $\tilde{c} + \bar{c} < 2c_{\alpha/4}$ for all $0 < \alpha < 0.25$, and we conclude that $\Pr(NCI) \geq \alpha/2$ also for $0 < \theta < \tilde{c} - \bar{c}$.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

For $\tilde{c} - \bar{c} < \theta \leq \bar{c} + c_{\alpha/2}$, $A(\theta) = \bar{A}(\theta)$, and since $\Pr_\theta(Y \in A(\theta)) = 1 - \alpha$, we have that $\Pr_\theta(NCI) = \alpha$.

Finally, for $\theta > \bar{c} + c_{\alpha/2}$ we have $\Pr_\theta(NCI) = \Pr_\theta(|Y| > \bar{c},\ |Y| > \theta + c_{\alpha/2}) \geq \alpha/2$.

In any case, $\Pr_\theta(NCI)$ does not drop below $\alpha/2$.

To evaluate the FCR, we follow a computation similar to that in BY. Let $0 < q < 0.25$ and $0 < \psi < 0.9$. Denote by $CI_i(\alpha) = \mathcal{C}_{MQC}(Y_i; \alpha)$ a level $1 - \alpha$ MQC interval using parameter ψ, and by $\bar{c}(\alpha) = \Phi^{-1}(1 - \psi \cdot \alpha)$ the value of the quantity $\bar{c}$ associated with it. Furthermore, let $C_k^{(i)} = \{Y^{(i)} : R_{\min}(Y^{(i)}) = k\}$. For the selective-SDCI procedure of Definition 2 $R_{\min} = R_{CI}$, in which case BY show that

$$\mathrm{FCR} = \sum_{i=1}^{m} \sum_{k=1}^{m} \frac{1}{k} \Pr \left\{ C_k^{(i)},\ i \in \mathcal{S}(\mathbf{Y}),\ \theta_i \notin CI_i \left(\frac{k \cdot q}{m} \right) \right\}. \qquad \text{(A.14)}$$

Using the fact that $i \in \mathcal{S}(\mathbf{Y})$ if and only if $|Y_i| \geq \bar{c}\left(\frac{R_{CI} \cdot q}{m}\right)$, we can replace

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

the right hand side of the last equality by

$$= \sum_{i=1}^{m}\sum_{k=1}^{m}\frac{1}{k}\Pr\left\{C_k^{(i)},\ |Y_i| \geq \bar{c}\left(\frac{k\cdot q}{m}\right),\ \theta_i \notin CI_i\left(\frac{k\cdot q}{m}\right)\right\} \tag{A.15}$$

$$= \sum_{i=1}^{m}\sum_{k=1}^{m}\frac{1}{k}\Pr\left\{C_k^{(i)}\right\} \times \Pr\left\{|Y_i| \geq \bar{c}\left(\frac{k\cdot q}{m}\right),\ \theta_i \notin CI_i\left(\frac{k\cdot q}{m}\right)\right\} \tag{A.16}$$

$$\geq \sum_{i=1}^{m}\sum_{k=1}^{m}\frac{1}{k}\Pr\left\{C_k^{(i)}\right\} \times \frac{kq}{2m} \tag{A.17}$$

$$= \frac{q}{2} \tag{A.18}$$

where inequality (A.17) follows from the preceding part of the proof as $\frac{k\cdot q}{m} \leq q < 0.25$.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316

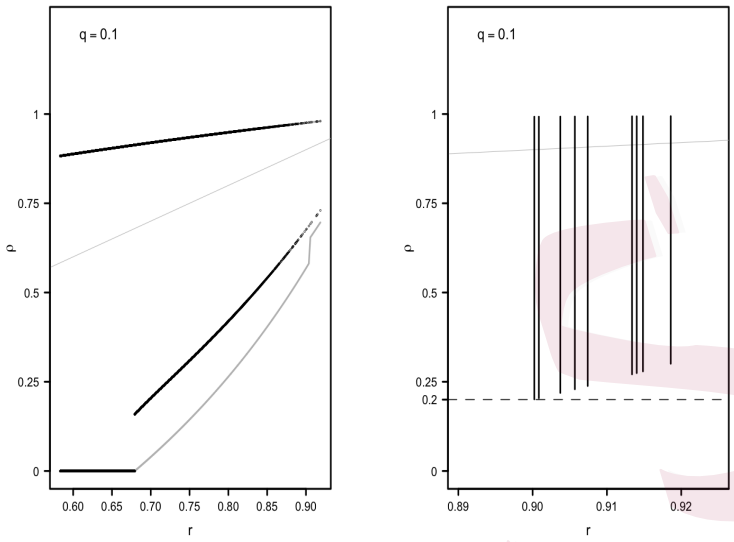
A.3 Proof of Theorem 3

Beginning with an expression for FCR as appears in Benjamini and Yekutieli (2005),

$$\begin{aligned}
\mathrm{FCR}(\widetilde{CI}_\bullet; \mathcal{S}; q_1, q_2) &= \sum_{r=1}^{m}\sum_{i=1}^{m} \frac{1}{r} \cdot \Pr(|\mathcal{S}_2| = r, \tilde{NCI}_i) \\
&= \sum_{r=1}^{m}\sum_{i=1}^{m} \frac{1}{r} \cdot \{ \Pr(\ |\mathcal{S}_2| = r,\ \tilde{NCI}_{i1}) + \Pr(\ |\mathcal{S}_2| = r,\ \theta_{i1} \in CI_{i1},\ \tilde{NCI}_{i2})\ \} \\
&= \sum_{r=1}^{m}\sum_{i=1}^{m} \frac{1}{r} \cdot \Pr(\ |\mathcal{S}_2| = r,\ i \in \mathcal{S}_2,\ \theta_{i1} \notin CI_{i1}) \\
&\quad + \sum_{r=1}^{m}\sum_{i=1}^{m} \frac{1}{r} \cdot \Pr(\ |\mathcal{S}_2| = r,\ \theta_{i1} \in CI_{i1},\ \tilde{NCI}_{i2}) \\
&= \sum_{r=1}^{m}\sum_{i=1}^{m} \frac{1}{r} \cdot \Pr(\theta_{i1} \notin CI_{i1}) \cdot \Pr(\ |\mathcal{S}_2| = r,\ i \in \mathcal{S}_2) \\
&\quad + \Pr(\theta_{i1} \in CI_{i1}) \cdot \sum_{r=1}^{m}\sum_{i=1}^{m} \frac{1}{r} \cdot \Pr(\ |\mathcal{S}_2| = r,\ \tilde{NCI}_{i2}) \\
&= q_1 \cdot \sum_{r=1}^{m} \frac{1}{r} \cdot \sum_{i=1}^{m} \Pr(\ |\mathcal{S}_2| = r,\ i \in \mathcal{S}_2) + \ (1 - q_1) \cdot \sum_{r=1}^{m}\sum_{i=1}^{m} \frac{1}{r} \cdot \Pr(\ |\mathcal{S}_2| = r,\ \tilde{NCI}_{i2})
\end{aligned}$$

To complete the proof, note that for any $\mathcal{S}_2$, $\sum_{i=1}^{m} \Pr(\ |\mathcal{S}_2| = r,\ i \in \mathcal{S}_2) = r \cdot \Pr(\ |\mathcal{S}_2| = r)$, and that $\mathrm{FCR}(CI_{\bullet 2};\ \mathcal{S}_2;\ q_2) = \sum_{r=1}^{m}\sum_{i=1}^{m} \Pr(|\mathcal{S}_2| = r, \tilde{NCI}_{\bullet 2})/r$.

Statistica Sinica: Preprint
doi:10.5705/ss.202017.0316



(a) (b)

Figure 3: Selective CIs that determine the sign (left panel), and selective CIs that detect large correlations (right panel); data from Tom et al. (2007). CIs are shown as vertical bars on right panel; on left panel only lower and upper endpoints of CI are shown. In both panels, observed correlations are on the horizontal axis, and the vertical axis represents true correlation values; light gray solid line is the identity line. (a) Black points correspond to MQC CIs, and gray lines to QC CIs, for the 33,856 correlations classified as positive. The upper endpoints for the two methods coincide, but the lower endpoint of MQC is farther away from zero. (b) Requiring that the selective CIs include only correlation values $\rho > 0.2$ or only values $\rho < -0.2$, the MQC_δ-equipped procedure constructs such intervals for nine out of the original 382,362 voxels. No correlations < -0.2 were detected.