SELECTIVE SIGN-DETERMINING MULTIPLE CONFIDENCE INTERVALS WITH FCR CONTROL

Asaf Weinstein and Daniel Yekutieli

Stanford University and Tel Aviv University

Supplementary Material

S1. Determining the Direction and the Type of the Association in Genomic Association Studies

The data were taken from the WTCC1 case-control study of Type-2 diabetes (Burton et al., 2007). We analyze data for $m = 459,653$ SNPs. The data for SNP $s = 1, ..., m$, is a 2-by-3 table listing the SNPs' minor allele counts for the cases and for the controls. We denote by $n_{s,i,j}$ the number of subjects of type i ($controls = 1$, $cases = 2$) with $j = 0, 1, 2$ copies of a, the minor allele of SNP s. For example, the data for SNP $s = 1412$ is shown in Table 1.

Clarke et al. (2011) suggest using the Cochran-Armitage trend test for discovering association between SNP and disease, with weights $w =$

Table 1: Data for SNP 1412.

Genotype:	AA	Aa	aa
Controls	690	1442	804
Cases	377	989	555

(w_1, w_2, w_3) that are chosen to detect particular types of association. $w = (0, 1, 1)$ is used to test whether allele a is dominant over allele A and $w = (0, 0, 1)$ is used to test whether allele a is recessive to allele A. However, most often, $w = (0, 1, 2)$ is used to test for an additive effect of allele a. The Cochran-Armitage statistic has a chi-squared distribution with 1 d.f. under the null hypothesis of no association. It is equivalent to the score statistic for the corresponding linear logit model, and its result is very similar to the logistic regression Wald test (Agresti, 2002). Assume that $n_{s,i,j}$ are multinomial with probabilities $(\pi_{s,1,0}, ..., \pi_{s,2,3})$. Let $\gamma_s^j = \log(\pi_{s,2,j}/\pi_{s,1,j})$ denote the log-Odds for diabetes for allele a count j of SNP s. For SNP 1412, the Cochran-Armitage test with $w = (0, 1, 2)$ yielded $Z = 2.605$; while the Wald test for the the null hypothesis that $\beta_s^1 = 0$ in the corresponding linear model,

$$\gamma_s^j = \beta_s^0 + \beta_s^1 \cdot j, \tag{S1.1}$$

yielded $Z = 2.604$.

S1.1 Confidence Regions for the Dominance and Recessiveness Effects

In our analysis, we treat SNP association as a bivariate problem that corresponds to the linear model:

$$\gamma_s^j = \gamma_s^0 + \beta_s^{Dom} \cdot I(1 \leq j) + \beta_s^{Rec} \cdot I(2 \leq j). \tag{S1.2}$$

Our parameters of interest are the allele a dominance effect $\beta_s^{Dom} = \gamma_s^1 - \gamma_s^0$, and the allele a recessiveness effect $\beta_s^{Rec} = \gamma_s^2 - \gamma_s^1$. For our analysis we assume that the allele effect is monotone increasing, $\gamma_s^0 \leq \gamma_s^1 \leq \gamma_s^2$, or decreasing, $\gamma_s^0 \geq \gamma_s^1 \geq \gamma_s^2$. Thus, $(\beta_s^{Dom}, \beta_s^{Rec})$ are both either nonnegative or non-positive. In our analysis, we construct confidence regions that determine the sign of $(\beta_s^{Dom}, \beta_s^{Rec})$ and indicate whether this effect is dominant $(\beta_s^{Rec} = 0)$, additive $(\beta_s^{Rec} = \beta_s^{Dom})$ or recessive $(\beta_s^{Dom} = 0)$.

Our parameter estimators are $\hat{\beta}_s^{Dom} = \hat{\gamma}_s^1 - \hat{\gamma}_S^0$ and $\hat{\beta}_s^{Rec} = \hat{\gamma}_s^2 - \hat{\gamma}_S^1$, with $\hat{\gamma}_S^j = \log(n_{s2j}/n_{s1j})$. To construct the confidence sets for $(\beta_s^{Dom}, \beta_s^{Rec})$, we assume that $(\hat{\beta}_s^{Dom}, \hat{\beta}_s^{Rec})$ are bivariate normal with mean $(\beta_s^{Dom}, \beta_s^{Rec})$ and covariance matrix whose entries are the following estimated variances and covariance: $\hat{\text{Var}}(\hat{\beta}_s^{Dom}) = 1/n_{s20} + 1/n_{s10} + 1/n_{s21} + 1/n_{s11}$, $\hat{\text{Var}}(\hat{\beta}_s^{Rec}) = 1/n_{s21} + 1/n_{s11} + 1/n_{s22} + 1/n_{s12}$, and $\hat{\text{Cov}}(\hat{\beta}_s^{Dom}, \hat{\beta}_s^{Rec}) = -1/n_{s21} - 1/n_{s11}$. Note that these parameter estimates are the same as those produced by fitting

model (S1.2) in R. The effect estimates for SNP 1412 are $\hat{\beta}_{1412}^{Dom} = 0.227$ and $\hat{\beta}_{1412}^{Rec} = 0.006$ and the estimated covariance matrix is

$$\hat{\Sigma}_{1412} = \begin{pmatrix} 0.0058 & -0.0017 \\ -0.0017 & 0.0047 \end{pmatrix}.$$

The confidence regions for $(\beta_{1412}^{Dom}, \beta_{1412}^{Rec})$, shown in Figure 1a, are valid under the assumption that $(\hat{\beta}_{1412}^{Dom}, \hat{\beta}_{1412}^{Rec})$ is bivariate normal with mean $(\beta_{1412}^{Dom}, \beta_{1412}^{Rec})$ and covariance $\hat{\Sigma}_{1412}$. The black curves are equi-density curves that produce $1-\alpha$ confidence regions with smallest volume for $(\beta_{1412}^{Dom}, \beta_{1412}^{Rec})$. The blue arrows are drawn in the direction of the principal components of the covariance matrix and their length is proportional to the square root of their variance. For $\hat{\Sigma}_{1412}$, $PC_{1412}^{1} = (-0.805, 0.593)^T$ with variance 0.0071, $PC_{1412}^{2} = (0.593, 0.805)^T$ with variance 0.0035.

Per construction, for all s, $\hat{\beta}_s^{Dom}$ and $\hat{\beta}_s^{Rec}$ are negatively correlated. Therefore, the 1st principal component will be a weighted difference between β_s^{Dom} and β_s^{Rec} and the 2nd principal component will be a weighted sum of β_s^{Dom} and β_s^{Rec}. As the sign of PC_s^2 is the same as the signs of β_s^{Dom} and β_s^{Rec}, we use the linear combination of $(\hat{\beta}_s^{Dom}, \hat{\beta}_s^{Rec})$ in the direction of PC_s^2 which has the smallest variance of all linear combinations of $(\hat{\beta}_s^{Dom}, \hat{\beta}_s^{Rec})$ to determine the direction of association. The line passing through $(0, 0)$ that is perpendicular to PC_s^2 (for SNP 1412 it is the red

diagonal line in Figure 1) represents 0 association. We quantify the size of association with $Z(PC_s^2)$, the distance in PC_s^2 standard deviations between $(\hat{\beta}_s^{Dom}, \hat{\beta}_s^{Rec})$ and the red diagonal. For SNP 1412, $Z(PC_{1412}^2) = 2.37$.

The $1-\alpha$ confidence regions for $(\beta_s^{Dom}, \beta_s^{Rec})$ we propose for this problem, are rectangular regions formed by intersecting a marginal $1-\alpha_1$ confidence region for $(\beta_s^{Dom}, \beta_s^{Rec})$ in the PC_s^1 direction with a $1-\alpha_2$ confidence region for $(\beta_s^{Dom}, \beta_s^{Rec})$ in the PC_s^2 direction, where $1-\alpha = (1-\alpha_1)\cdot(1-\alpha_2)$. Orthogonality of the estimators of the principal components ensures $1-\alpha$ coverage probability for our rectangular intervals, thereby allowing to allocate a different degree of confidence in each direction. In the next section we use this property for constructing confidence regions that are inflated due to selection *only in the direction of* PC_s^2. Furthermore, replacing the ellipsoid confidence regions with rectangular confidence regions may also lead to sharper sign determination.

The green rectangle in Figure 1a is 0.95 confidence region for $(\beta_{1412}^{Dom}, \beta_{1412}^{Rec})$ formed by intersecting a symmetric two-sided marginal 0.96 CI in the PC_{1412}^2 direction and a symmetric two-sided marginal $1-0.0104$ CI in the PC_{1412}^1 direction. Indeed, we see that since $Z(PC_{1412}^2) > z_{.02}$ this confidence region SNP is above and to the right of the red line even though the 0.95 ellipsoid bivariate normal confidence set crosses the red diagonal, indicating

positive association of allele a with diabetes. Furthermore, the confidence region prominently covers $\beta_{1412}^{Rec} = 0$ parameter points while barely covering $\beta_{1412}^{Dom} = 0$ parameter points, suggesting that the effect of allele a is dominant.

S1.2 Using FCR-Adjusted CIs for Determining SNP Direction of Association

We begin by constructing 1-dimensional CIs for β_s^1, the logistic regression coefficient for the number of minor SNP alleles in model (S1.1), for all m SNPs and use these CIs to determine the SNPs' direction of association.

Applying the BH procedure to $p_s = 2 \cdot (1 - \Phi(|Z(\hat{\beta}_s^1)|))$ at level $q = 0.05$ yielded 27 discoveries and at level $q = 0.10$ it yielded 43 discoveries. Here $Z(\hat{\beta}_s^1) = \hat{\beta}_s^1 / \hat{\text{sd}}(\hat{\beta}_s^1)$. Level $q = 0.05$ FCR-adjusted MQC CI with $\psi = 0.7$ yielded 35 discoveries and setting $\psi = 0.9$ yielded 36 discoveries.

We now consider the rectangular confidence regions for $(\beta_s^{Dom}, \beta_s^{Rec})$ defined in Section 5 that, using the algorithm in Definition 3 , will be inflated for selection according to the value $Z(PC_s^2)$, for determining the SNPs' direction of association and the type of association.

For the rectangular confidence regions for $(\beta_s^{Dom}, \beta_s^{Rec})$ it is necessary to specify q_1 and q_2, the non-coverage level for each principal component.

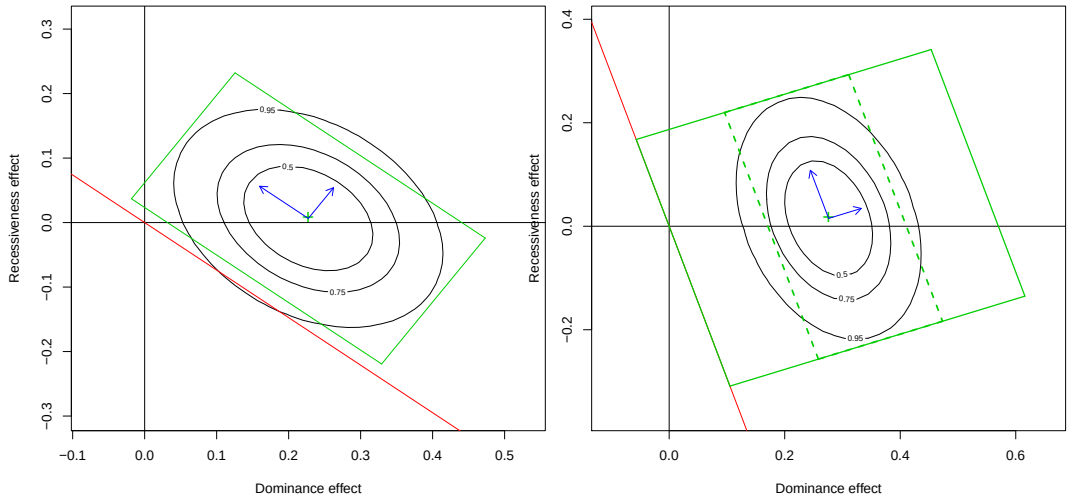
Allocating all the non-coverage probability to PC_s^2, i.e. setting $q_2 = q$ and $q_1 = 1$, reduces the confidence regions to $1 - q$ marginal CIs for PC_s^2 and level q selection rules that are based on PC_s^2. Applying the BH procedure to $p_s = 2 \cdot (1 - \Phi(|Z(PC_s^2)|))$ at level $q = 0.05$ yielded 23 discoveries and at level $q = 0.10$ it yielded 31 discoveries. Level $q = 0.05$ FCR-adjusted MQC CI with $\psi = 0.7$ yielded 24 discoveries and setting $\psi = 0.9$ yielded 30 discoveries.

We now consider confidence regions for $(\beta_s^{Dom}, \beta_s^{Rec})$ with $q_2 = 0.04$ and $q_1 = 0.0104$. As selection is only applied in the direction of PC_s^2, the CIs for PC_s^2 are level $q_2 = 0.04$ FCR-adjusted marginal CIs that are based on $Z(PC_s^2)$, and the CIs for PC_s^1 are (unadjusted) $1 - 0.0104$ marginal two-sided CIs based on $Z(PC_s^1)$. Applying the BH procedure to $P_s = 2 \cdot (1 - \Phi(|Z(PC_s^2)|))$ at level $q = 0.04$ yielded 23 discoveries as before, and at level $q = 0.08$ the BH procedure yielded 30 discoveries. Level $q = 0.04$ FCR-adjusted MQC CI with $\psi = 0.7$ yielded 24 discoveries, and setting $\psi = 0.9$ yielded 29 discoveries. Even though the distribution of $|Z(\hat{\beta}_s^1)|$ was larger than that of $|Z(PC_s^2)|$ (it yielded more discoveries) the ordering of the SNPs according to the two Z-scores was very similar—the ranking of the 19 most significant SNPs according to $Z(\hat{\beta}_s^1)$ and $Z(PC_s^2)$ was the same.

SNP 69962 has large $Z(PC^2_{69962}) = 4.461$ (ranked 27) and a relatively smaller $Z(\hat{\beta}^1_{69962}) = 4.074$ (ranked 64 – undiscoverable with level $q = 0.10$ BH procedure). The rectangle formed by the solid green lines in Figure 1b is the MQC CI rectangular confidence sets for $(\beta^{Dom}_{69962}, \beta^{Rec}_{69962})$ with $q_2 = 0.04$ and $q_1 = 0.0104$ and $\psi = 0.9$ adjusted for the 29 selected SNPs. The smaller rectangle formed by the broken green lines and the solid green lines is the unadjusted confidence sets for $(\beta^{Dom}_{69962}, \beta^{Rec}_{69962})$ with $q_2 = 0.04$ and $q_1 = 0.0104$. As SNP 69962 was selected its selection adjusted confidence set does not cross the red line, but rather is on the red line corresponding to MQC CI lower boundary that is equal 0; this indicates non-negative association with diabetes. Furthermore, the fact that the selective confidence region prominently covers $\beta^{Rec}_{69962} = 0$ parameter points while barely covering $\beta^{Dom}_{69962} = 0$ parameter points suggests that the effect of allele a is dominant.

S2. Selective-SDCI Procedures Under Dependency

Equipped with any marginal CI that satisfies the requirements (MON 1) and (MON2), a level-q Selective-SDCI procedure is guaranteed to control FCR $\leq q$ under independence of the observations. The case of dependent data is considerably more challenging. For general dependency, applying the Selective-SDCI of Definition 2 (Section 3), equipped with any marginal



(a) marginal confidence regions (b) selective confidence regions

Figure 1: (a) Confidence regions for $(\beta^{Dom}_{1412}, \beta^{Rec}_{1412})$. Green plus sign is $(\hat{\beta}^{Dom}_{1412}, \hat{\beta}^{Rec})$. Black curves are 0.50, 0.75 and 0.95 bivariate normal confidence regions. Blue arrows are the principal components of the bivariate normal distribution. Green rectangle is a 0.95 confidence set based on the two principal components. Red line is drawn at association effect equals 0. (b) Selection-adjusted confidence regions for $(\beta^{Dom}_{69962}, \beta^{Rec}_{69962})$. Green plus sign is at $(\hat{\beta}^{Dom}_{69962}, \hat{\beta}^{Rec}_{69962})$. Black curves are 0.50, 0.75 and 0.95 (unadjusted) bivariate normal confidence regions. Green rectangle captured by broken lines is the 0.95 unadjusted confidence set based on the two principal components. Solid green rectangle is the 0.95 selection-adjusted confidence set based on the two principal components. Red line is drawn at association effect equals 0.

CI) at level $q/\sum_{j=1}^{m}\frac{1}{j}$ ensures FCR $\leq q$. This follows immediately from Theorem 4 in Benjamini and Yekutieli (2005).

If the estimators are positive regression dependent on a subset (PRDS hereafter; Benjamini and Yekutieli, 2001), a consequence of Theorem 3 in Benjamini and Yekutieli (2005) is that if the level-q procedure of Definition 2 (Section 3) is equipped with a the interval $\mathcal{C}(y;\alpha) = (y - c_\alpha, \infty)$ (or $\mathcal{C}(y;\alpha) = (-\infty, y + c_\alpha)$), then the FCR is still controlled at q under PRDS; however, if sign detection is of interest, we would never want to equip the Selective-SDCI with such a CI (which would mean giving up on either detection of positive or on detection of negative parameters). Hence Theorem 3 in Benjamini and Yekutieli (2005) does not really cover selection followed by CI construction via the Selective-SDCI procedure of Definition 2 (Section 3). We would like to point out that under PRDS even the validity of the level-q directional-BH procedure, which is a special case of our Selective-SDCI procedure, has not been established before.

We examine FCR of the Selective-SDCI procedure in practice using simulations. The data simulates the brain voxel data of Section 7 under dependency; it was generated as in Rosenblatt and Benjamini (2014) and using their code, available at `https://github.com/johnros/SelectiveEstimationSimulations`. Specifically, for each configuration of nonzero effect size and proportion of nulls, and in each of 100 rounds, data representing z-transformed ($n = 16$ subjects) correlations for a "brain" of

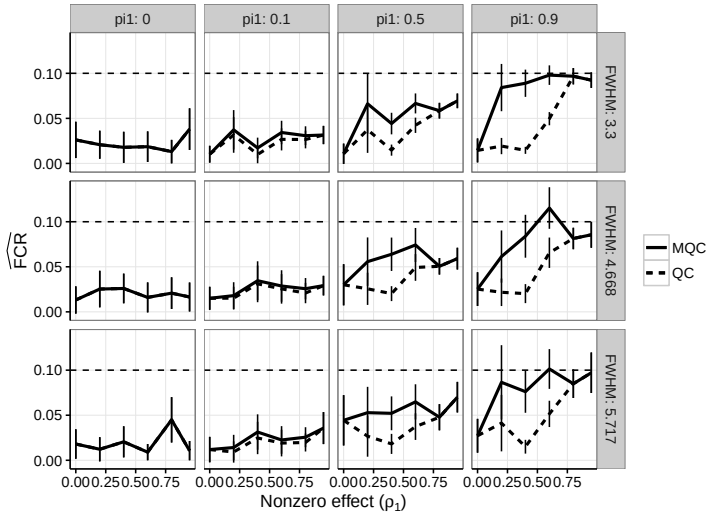


Figure 2: Estimated FCR of the Selective-SDCI procedure under dependency. Each point in the figure is estimated FCR for a specific configuration, when applying the procedure of Definition 2 (Section 3) at level 0.1 to Fisher-transformed voxel correlations (sample size is $n = 16$). Vertical bars are drawn at $\pm$ two standard errors. In each simulation round the underlying signal for each voxel was independently set to ρ_1 w.p. π_1 or zero w.p. $1 - \pi_1$. The two line types correspond to the procedure using the QC marginal interval (broken) and the MQC marginal interval (solid). For the no-signal case ($\pi_1 = 0$) the two lines coincide.

$10 \times 10 \times 10$ voxels was generated as the sum of a signal field and a smooth Gaussian noise field. The smoothness of the noise field—controlling the spatial covariance—is represented by the parameter FWHM, which was varied at 3 different levels $\{3.3, 4.7, 5.7\}$ (4.7 being the estimated quantity from the actual data analyzed in Section 7 . The smooth Gaussian random field

used in generating the data is PRDS (Nichols and Hayasaka, 2003), hence is an appropriate case to examine the actual FCR of our procedure for PRDS data. More details describing how data was generated are available in Rosenblatt and Benjamini (2014, Appendix C.2).

In each round we applied our procedure at level $q = 0.1$, first using the QC interval as the marginal CI, and second using the MQC as the marginal CI. For each of the two methods we recorded the proportion of non-covering constructed CIs (FCP) as well as the number of constructed sign-determining intervals. The results are presented in Figure 2, which shows that, overall, the situation is qualitatively similar to the independent case: for almost each simulation configuration the estimated FCR is under 0.1, while it is much closer to the nominal level for the MQC-equipped procedure. Specifically, for larger proportion of non-zero effects ($\pi_1 = 0.5, 0.9$), the estimated FCR of the MQC-equipped procedure is larger than $q/2 = 0.05$ for all configurations of smoothness (FWHM) and levels of non-zero signal strength (≥ 0.2).

S3. Connections to Existing Work on Post-Selection Inference

In the procedure of Definition 2 (Section 3) a CI is constructed for θ_i only if $i \in \mathcal{S}(\mathbf{Y})$ where $\mathcal{S}$ is a prespecified rule. While selection needs to be taken

into account when constructing the CIs, adjusting the level of marginal CIs is not the only way to achieve FCR control. To restore validity of post-selection CIs, a common approach, which is not limited to the multiplicity setup or to selection rules we consider in the current paper, is to construct intervals that have the nominal coverage level *conditional* on selection. Thus, in the setup of the current paper, whenever $i \in \mathcal{S}(\mathbf{Y})$ the conditional approach would construct an interval $CI_i(q)$ with the property

$$\Pr\left(\theta_i \in CI_i(q) \middle| i \in \mathcal{S}(\mathbf{Y})\right) \geq 1 - q. \tag{S3.3}$$

Conditional CIs based on a truncated univariate normal distribution were suggested in Zhong and Prentice (2008) and Weinstein et al. (2013) . A Recent line of work, including Lee et al. (2016); Taylor et al. (2014); Sun and Taylor (2014) (among others), greatly increased the applicability of the conditional approach by developing exact methods for constructing CIs when selection corresponds to truncating a multivariate normal distribution to a polyhedron; these results were in turn extended to generalized linear models in Fithian et al. (2014), who also provided theoretical support.

Under independence of the Y_i, we obtain a valid two-stage procedure by selecting parameters through a level-$2q$ BH procedure for testing $H_0^i : \theta_i = 0$, classifying θ_i as positive or non-positive according as $Y_i > 0$ or $Y_i < 0$;

then for each classified parameter constructing a $1-q$ conditional CI. From Weinstein et al. (2013, Section 7) it follows that if for each $i \in \mathcal{S}(\mathbf{Y})$ a CI $CI_i = \mathcal{C}(Y_i; \sigma^2, \hat{c}, q)$ is constructed where $\mathcal{C}(Y_i; \sigma^2, c, \alpha)$ has the property that

$$\Pr_{Y \sim N(\theta, \sigma^2)} \left(\theta \in \mathcal{C}(Y; \sigma^2, c, \alpha) \big| |Y| > c \right) \geq 1 - \alpha \tag{S3.4}$$

and where

$$\hat{c} = \Phi^{-1}(1 - i^* q/m), \qquad i^* = \max\{i : P_{(i)} \leq i(2q)/m\}$$

for $P_i = 2(1-\Phi(|Y_i|))$ and $P_{(1)} \leq P_{(2)} \leq ...P_{(m)}$, then (S3.3) is satisfied. This is because the conditional distribution of $Y_i \big| \left(\mathbf{Y}^{(i)}, i \in \mathcal{S}(\mathbf{Y}) \right)$ is that of a normal variable truncated to $\{y : |y| > \hat{c}\}$ where $\hat{c}$ is a constant determined by $\mathbf{Y}^{(i)}$.

The procedure described above controls both wdFDR $\leq q$ and FCR $\leq q$ as required; however, a disadvantage is that—unlike the procedure of Definition 2 (Section 3)—it cannot ensure that a constructed CI_i does not cross zero; this would contradict with the fact that the sign of θ_i was detected. In the example of Section 7 , if we use the conditional CI of Weinstein et al. (2013, CQC with $\lambda = 0.4$) after BH selection at level 0.2, then 28,082 out of the 43,804 constructed intervals include both positive and negative values. Hence 64% of the parameters whose sign was classified at the first stage, are supplemented with intervals that do not determine the sign. In fact, the

construction of Weinstein et al. (2013) is designed specifically to promote sign detection, and still it is not able to guarantee it.

We now show that the situation would be similar if any other conditional CI is used instead. Indeed, for a constant c let $\bar{\mathrm{P}}_\theta(\cdot)$ indicate probability under the conditional distribution of $Y \sim N(\theta, 1)$ given $|Y| > c$, and suppose that there exists a procedure $\mathcal{C}(y;\alpha) = \mathcal{C}(y;\sigma^2 = 1, c, \alpha)$ with the property (S3.4) such that $\mathcal{C}(y;\alpha) \subseteq (-\infty, 0]$ or $\mathcal{C}(y;\alpha) \subseteq (0,\infty)$ for all $|y| > c$. Let $\alpha < 1/2$. Then

$$\begin{aligned} 1 &= \bar{\mathrm{P}}_0(\mathcal{C}(Y;\alpha) \subseteq (-\infty, 0]) + \bar{\mathrm{P}}_0(\mathcal{C}(Y;\alpha) \subseteq (0,\infty)) \\ &= \lim_{\theta \to 0^+} \bar{\mathrm{P}}_\theta(\mathcal{C}(Y;\alpha) \subseteq (-\infty, 0]) + \bar{\mathrm{P}}_0(\mathcal{C}(Y;\alpha) \subseteq (0,\infty)) \leq \alpha + \alpha = 2\alpha \end{aligned}$$

which is a contradiction. The equality in the second line is by continuity of $\bar{\mathrm{P}}_\theta(\mathcal{C}(y;\alpha) \subseteq (-\infty, 0])$ at $\theta = 0$. The inequality in the third line is because for any $\theta > 0$, coverage property of the interval implies necessarily that $\bar{\mathrm{P}}_\theta(\mathcal{C}(y;\alpha) \subseteq (-\infty, 0]) \leq \alpha$, and similarly $\bar{\mathrm{P}}_0(\mathcal{C}(y;\alpha) \subseteq (0,\infty)) \leq \alpha$.

S4. A Full Specification of the MQC Confidence Interval

In Section 4 (equation 9) we provided a specification of the MQC CI when $0 < \psi \leq \psi_1$. A complete specification of the MQC CI follows. The constants $\bar{c}, \tilde{c}, \psi_1, \psi_2$ and the function $g(\theta)$ are all as defined in Section 4 .

If $0 < \psi \leq \psi_1$,

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

If $\psi_1 < \psi \leq \psi_2$,

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

If $\psi_2 < \psi$,

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

with $\mathcal{C}(-y;\alpha) = -\mathcal{C}(y;\alpha)$.

S5. Detecting Large Correlations

The focus in Section 7 was on detecting the sign of correlations. In practice it may be of interest to detect instead only large correlations, namely, correlations $\rho_i > \rho_0$ or $\rho_i < -\rho_0$ for some prespecified constant $\rho_0 \in (0, 1)$, while still controlling the proportion of incorrect decisions. Hodges and Lehmann

(1954) referred to testing of the interval hypothesis $H_0 : \rho \in [-\rho_0, \rho_0]$ as testing for "material significance". Also, as pointed out by Finner (1994), our pursuit reflects in some sense the opposite goal of the bioequivalence problem, where the aim is to detect parameters $|\rho| < \rho_0$. We offer here an extension of our Selective-SDCI procedure and use it to detect large correlations in the study of Section 7.

As before, let $Y_i \sim f(y_i - \theta_i)$ $i = 1, ..., m$, with f a unimodal and symmetric density. Fix $\delta \in (0, \infty)$. We are interested in a procedure which, for a subset $S = \mathcal{S}(\mathbf{Y}) \in \{1, ..., m\}$, constructs a CI CI_i for each $\theta_i, i \in S$, such that $CI_i \subseteq (\delta, \infty)$ or $CI_i \subseteq (-\infty, -\delta)$, and at the same time FCR $\leq q$. Adapting the Selective-SDCI procedure from the original sign problem to the current situation is starightforward: simply replace R in Definition 3 (Section 3) with

$$R = \max \left\{ r : CI_{(r)} \left(\frac{r \cdot q}{m} \right) \text{ includes only values } > \delta \text{ or only values } < -\delta \right\}$$

and leave the procedure of Definition 2 (Section 3) otherwise unchanged. As was the case for the sign problem, any marginal CI satisfying Requirements (MON 1) and (MON 2) can be used with the Selective-SDCI procedure. However, in order to obtain a powerful procedure, we would like to use a marginal CI which is suited to the current task rather than to the sign problem. Specifically, we adapt the MQC marginal interval of Section 4 to obtain an interval that avoids intersecting $[-\delta, \delta]$ starting at an observation

value smaller (in absolute value) than $\delta + c_{\alpha/2}$.

Like the MQC interval of Section 4 , the new interval is obtained by inverting a family of acceptance regions. The $1 - \alpha$ acceptance regions describing the CI are

$$A_{\mathrm{MQC}_\delta}(\theta) = \begin{cases} (-\delta - \bar{c}, \delta + \bar{c}), & 0 \le \theta \le \delta \\ (-\delta - \bar{c}, g_\delta(\theta)), & \delta \le \theta < \delta + \bar{c} + c_{\alpha/2} \\ (\theta - c_{\alpha/2}, \theta + c_{\alpha/2}), & \delta + \bar{c} + c_{\alpha/2} < \theta \end{cases}$$

with $A_{\mathrm{MQC}_\delta}(\theta) = -A_{\mathrm{MQC}_\delta}(\theta)$ for $\theta < 0$. Above, the constant $\bar{c}$ is determined by δ through

$$F(\delta + \bar{c} - \delta) - F(-\delta - \bar{c} + \theta) = 1 - \alpha \tag{S5.5}$$

and the function g_δ is given by

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

The convex hull of $\{\theta : y \in A_{\mathrm{MQC}_\delta}(\theta)\}$ is then

$$\mathcal{C}(y;\alpha) = \begin{cases} (-\delta - \bar{c} - c_{\alpha/2}, \delta + \bar{c} + c_{\alpha/2}), & 0 \le x < \delta + \bar{c} \\ (g^{-1}(x), x + c_{\alpha/2}), & \delta + \bar{c} \le x < g(\delta + \bar{c} + c_{\alpha/2}) \\ (\delta + \bar{c} + c_{\alpha/2}, x + c_{\alpha/2}), & g(\delta + \bar{c} + c_{\alpha/2}) \le x < \delta + \bar{c} + 2c_{\alpha/2} \\ (x - c_{\alpha/2}, x + c_{\alpha/2}), & \delta + \bar{c} + 2c_{\alpha/2} \le x \end{cases}$$

Figure 3 shows the resulting interval for a normal distribution and $\delta = 0.5$. Note that while the MQC interval was parametrized by ψ (or, equivalently, by $\bar{c}$) which determined the tradeoff between early sign determination and maximum length of the CI, MQC_δ is not indexed by such a parameter. Indeed, for MQC_δ there is no flexibility in choosing how early the interval stops crossing δ (or $-\delta$): to any δ corresponds a constant $\bar{c}$ given by (S5.5). Note also that while MQC guarantees only weak determination of the sign (≤ 0 or > 0) whenever it does not include values of both signs, MQC_δ is always an open interval, and it separates from δ ($-\delta$) immediately at $\delta + \bar{c}$, whereas MQC separates from zero at $\tilde{c} > \bar{c}$. These are consequences of the difference between the sign problem and the problem of detecting large effects.

We set $q = 0.1$ and applied our procedure, equipped with the MQC_δ interval, to the data from Tom et al. (2007) for detecting correlations $\rho_i > 0.2$ or $\rho_i < -0.2$; the value $\rho_0 = 0.2$ was chosen to represent a "sufficiently large" correlation size. The results are shown in Figure 3(b) of Section 7 . Out of the 382,362 voxels originally considered in our analysis, our procedure finds that only 9 can be declared to have a correlation larger than 0.2 or smaller than -0.2. All 9 reported intervals are for positive estimated correlations, and cover only values larger than 0.2. Hence, while

detection of the sign was possible for as many as 36,131 parameters (in the previous subsection), there is a dramatic decrease in the number of finding already when we aim at discovering correlations of size at least 0.2.

Note that the 9 constructed intervals which lie above 0.2 are FCR-adjusted at a more stringent level than the sign-determining intervals constructed before, as less parameters are selected. This inflation is needed to ensure that FCR is controlled for the new set of findings: reporting those among the 36,131 sign-determining intervals which further lie above 0.2 or below -0.2 would, of course, suffer from the same type of selection bias problem as would reporting in the first place all unadjusted 90% intervals which lie above 0.2 or below -0.2.

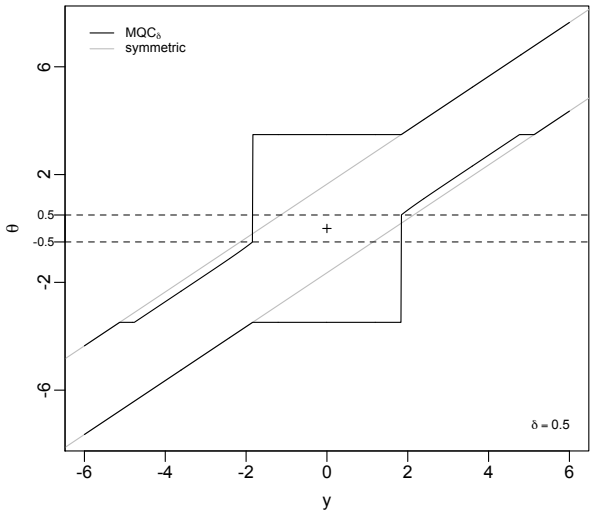


Figure 3: Modified Quasi-Conventional CI MQC_δ for early detection of large effects. Here $Y \sim N(\theta, 1)$ and the CI for θ is designed to exclude values in $[-\delta, \delta]$ as early as possible, i.e., for small values of $|y|$. The plot is for $\alpha = 0.1, \delta = 0.5$. For this configuration the interval lies completely above 0.5 or completely below -0.5 starting at $|y| = 1.84$; compare to $\delta + z_{1-0.1/2} = 2.14$ for the symmetric interval, shown in gray. Broken lines are drawn at $\pm\delta$ and the origin is marked with a plus sign.

Bibliography

Agresti, A. (2002). Categorical data analysis. *A John Wiley and Sons, Inc. Publication, Hoboken, New Jersey, USA*.

Benjamini, Y. and D. Yekutieli (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of statistics*, 1165–

1188.

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.

Burton, P. R., D. G. Clayton, L. R. Cardon, N. Craddock, P. Deloukas, A. Duncanson, D. P. Kwiatkowski, M. I. McCarthy, W. H. Ouwehand, N. J. Samani, et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature 447*(7145), 661–678.

Clarke, G. M., C. A. Anderson, F. H. Pettersson, L. R. Cardon, A. P. Morris, and K. T. Zondervan (2011). Basic statistical analysis in genetic case-control studies. *Nature protocols 6*(2), 121–133.

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*.

Hodges, J. and E. Lehmann (1954). Testing the approximate validity of statistical hypotheses. *Journal of the Royal Statistical Society. Series B (Methodological)*, 261–268.

Lee, J. D., D. L. Sun, Y. Sun, J. E. Taylor, et al. (2016). Exact post-selection inference, with application to the lasso. *The Annals of Statistics 44*(3), 907–927.

Nichols, T. and S. Hayasaka (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical methods in medical research 12*(5), 419–446.

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

Sun, Y. and J. E. Taylor (2014). Valid post-correction inference for censored regression problems. *arXiv preprint arXiv:1403.3457*.

Taylor, J., R. Lockhart, R. Tibshirani, and R. Tibshirani (2014). Exact post-selection inference for sequential regression procedures. *Journal of the American Statistical Association 111*(514), 600–620.

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.

Zhong, H. and R. L. Prentice (2008). Bias-reduced estimators and confidence intervals for odds ratios in genome-wide association studies. *Biostatistics 9*(4), 621–634.