Title: Correction to Begg's Test for Publication Bias
Author: Haben Michael and Musie Ghebremichael
Abstract: Begg and Mazumdar proposed using a rank correlation test to for publication bias when carrying out meta-analyses. The asymptotic variance of the rank correlation test statistic was derived under assumptions unmet by this application, often resulting in a loss of power. Low power when Begg's test is used to screen for publication bias may lead to false positives in a subsequent meta-analysis. We obtain the asymptotic bias under the common conditionally normal model as a function of the distribution of primary study variances. In simulations we consider the performance of Begg's test using an approximation to the correct asymptotic variance. We then examine several meta-analyses drawn from the literature where the standard and bias-corrected versions of Begg's test lead to different conclusions.

1 Introduction

Meta-analysis is a popular technique for summarizing a body of studies. Key to the soundness of the approach is that the body of studies used be representative of the studies conducted. This requirement may fail to be met when publication bias is present, when the availability of a study is tied to its conclusion [1]. Several hypotheses tests have been proposed with the goal of identifying the presence of publication bias on the basis of the relationship between the conclusion of the study and various study characteristics.

Common to these tests is that the null is held to be no publication bias. A typical conservative analyst might be expected to treat the presence of publication bias as the null. Given the manifold sources of publication bias, devising a test under such a null does not appear practical. The result, however, is that Type II errors in the test for publication bias will often correspond to Type I errors in the subsequent meta-analysis. Assessing and improving the power of the screening test is therefore worthwhile.

One of the most common tests for publication bias, Begg's test [2], tests for correlation between the studies' reported effect sizes and their standard errors. An issue with Begg's test procedure is that it uses the asymptotic variance for a general correlation test derived under assumptions unmet by Begg's test. This nominal variance is often larger than the correct variance, as discussed below. As a result, the test does not reject as frequently as it ought, which, as mentioned above, is likely to lead to Type I error in the meta-analysis for which the publication bias test is being performed.

This issue with Begg's test has been noted previously, including by the author of the test [2, 1]. More recently, [3] showed by direct calculation that the observations are correlated, so that the usual assumptions for the rank correlation test are not met. Since the rank correlation test depends on an asymptotic approximation, this criticism isn't entirely fair unless the effect of the correlation does not vanish in the limit, as we show.

2 Asymptotic bias of Begg's test

2.1 Description of test

Begg's test is a test of correlation between the reported effect sizes and their reported variances. The premise is that a tendency to publish larger effect sizes induces a trend in effect sizes across their variances, and no such trend exists without selection. See Fig. 2 for an illustration.

The data consists of independent pairs $(y_1, \sigma_1), \ldots, (y_n, \sigma_n)$ representing the estimated effect sizes and sampling variances of n studies with a common mean effect size, say θ:

$$\begin{aligned} E(y_j \mid \sigma_j) &= \theta, j = 1, \ldots, n \\ Var(y_j \mid \sigma_j) &= \sigma_j. \end{aligned} \tag{1}$$

The null is that y_j is uncorrelated with σ_j, $j = 1, \ldots, n$. The test statistic is Kendall's rank correlation coefficient,

$$\tau = 2\binom{n}{2}^{-1} \sum_{j<k} \{(u_j - u_k)(v_j - v_k) > 0\} - 1,$$

applied to the sequence of pairs (u_j, v_j) given by the data after standardizing the effect sizes,

$$\begin{aligned} &(u_j, v_j) = \left(\frac{y_j - \hat{\theta}}{\sqrt{\sigma_j^2 - \sigma_{\hat{\theta}}^2}}, \sigma_j \right), j = 1, \ldots, n, \\ &\text{where} \\ &\hat{\theta} = (\sum_{j=1}^{n} y_j/\sigma_j^2)/(\sum_{j=1}^{n} 1/\sigma_j^2), \\ &\sigma_{\hat{\theta}}^2 = 1/\sum_{j=1}^{n}(1/\sigma_j^2). \end{aligned} \tag{2}$$

The mean estimate $\hat{\theta} = (\sum_{j=1}^{n} y_j/\sigma_j^2)/(\sum_{j=1}^{n} 1/\sigma_j^2)$ is the inverse-variance weighted estimate of the common study mean θ and $\sigma_{\hat{\theta}}^2 = 1/\sum_{j=1}^{n}(1/\sigma_j^2)$ is its variance, both conditional on the study variances. The test statistic counts the number of corresponding pairs of studentized effect sizes $u_j = (y_j - \hat{\theta})/\sqrt{\sigma_j^2 - \sigma_{\hat{\theta}}^2}$ and variances $v_j = \sigma_j$ that concord in their magnitudes, in the sense that either $u_j < u_k$ and $v_j < v_k$ or $u_j > u_k$ and $v_j > v_k$. The null of no correlation is to be interpreted as no publication bias, and is rejected at level α when $\sqrt{9n/4}|\tau| > \Phi^{-1}(1 - \alpha/2)$.

The asymptotic null variance 4/9 is derived under the assumption that the pairs form an iid sequence. This assumption does not hold for the pairs (2) due to the common terms $\hat{\theta}$ and $\sigma_{\hat{\theta}}^2$. While the latter is of order $1/n$ and typically negligible in the limit, the dependence induced by the summary statistic $\hat{\theta}$, ordinarily of order $1/\sqrt{n}$, must be accounted for in computing the asymptotic null variance of $\sqrt{n}\tau$.

2.2 Source of bias

The variance of $\sqrt{n}\tau$ is

$$Cov(\sqrt{n}\tau) = 4n\binom{n}{2}^{-2} \sum_{1\le i,j,k,l\le n} Cov(\{(u_i - u_j)(v_i - v_j) > 0\}, \{(u_k - u_l)(v_k - v_l) > 0\})$$

The sum has $\binom{n}{2}\binom{n-2}{2}$ terms where i, j, k, l are all distinct, $2(n-2)\binom{n}{2}$ terms where the set $\{i, j, k, l\}$ has size 3, and $\binom{n}{2}$ terms where $|\{i, j, k, l\}| = 2$, so

$$Cov(\sqrt{n}\tau) = 4\frac{(n-2)(n-3)}{n-1}Cov(\{(u_1 - u_2)(v_1 - v_2) > 0\}, \{(u_3 - u_4)(v_3 - v_4) > 0\})+$$

$$16\frac{n-2}{n-1}Cov(\{(u_1 - u_2)(v_1 - v_2) > 0\}, \{(u_1 - u_3)(v_1 - v_3) > 0\}) + O(1/n).$$

The second term on the rhs, with the $O(1)$ coefficient, converges in probability to

$$16Cov\left(\left\{\left(\frac{y_1-\theta}{\sigma_1} - \frac{y_2-\theta}{\sigma_2}\right)(\sigma_1 - \sigma_2) > 0\right\}, \left\{\left(\frac{y_1-\theta}{\sigma_1} - \frac{y_3-\theta}{\sigma_3}\right)(\sigma_1 - \sigma_3) > 0\right\}\right) = 4/9,$$

the usual asymptotic null variance of Kendall's τ. The first term on the rhs, with an $O(n)$ coefficient, is a source of bias if the covariance

$$Cov(\{(u_1 - u_2)(v_1 - v_2) > 0\}, \{(u_3 - u_4)(v_3 - v_4) > 0\}) =$$
$$Cov\left(\left\{\left(\frac{y_1-\hat{\theta}}{\sqrt{\sigma_1^2 - \sigma_{\hat{\theta}}^2}} - \frac{y_2-\hat{\theta}}{\sqrt{\sigma_2^2 - \sigma_{\hat{\theta}}^2}}\right)(\sigma_1 - \sigma_2) > 0\right\}, \left\{\left(\frac{y_1-\theta}{\sqrt{\sigma_1^2 - \sigma_{\hat{\theta}}^2}} - \frac{y_3-\theta}{\sqrt{\sigma_3^2 - \sigma_{\hat{\theta}}^2}}\right)(\sigma_4 - \sigma_4) > 0\right\}\right) \tag{3}$$

does not vanish faster than $1/n$. The false positive rate of Begg's test will exceed or fall below the nominal level when the direction of the bias is negative or positive, respectively, i.e., when the covariance (3) is positive or negative.

2.3 Direction of bias

Assume in (1) that y_j/σ_j has a fixed distribution, say F_Z, i.e., the pairs $(y_1, \sigma_1), \ldots, (y_n, \sigma_n)$ belong to a scale family. Assume further that F_Z is the distribution of a symmetric random variable. Let $S^2 = 1/\sigma^2$ denote the study precisions.

$$\begin{aligned} Z_1, \ldots, Z_n \text{ iid } &\sim F_Z \\ S_1, \ldots, S_n \text{ iid } &\sim F_S \\ Z_j &\sim -Z_j, \\ z_j \mid s_j &\sim z_j, \\ Y_j &= Z_j/S_j, j = 1, \ldots, n \end{aligned}$$

Ignoring the $O(1/n)$ terms $\sigma_{\hat{\theta}}$ in (2), the test statistic may be written in terms of z and s as

$$\tau = \sum_{j<k} \left\{ \frac{z_j - z_k}{s_j - s_k} > \hat{\theta} \right\}.$$

The covariance (3) determining the bias of the asymptotic variance relative to the 9/4 is

$$Cov\left(\left\{ \frac{z_1 - z_2}{s_1 - s_2} > \hat{\theta} \right\}, \left\{ \frac{z_3 - z_4}{s_3 - s_4} > \hat{\theta} \right\} \right).$$

The grand mean estimate $\hat{\theta} = \sum ZS / \sum S^2$ is a function of $(z_1, s_1), \ldots, (z_4, s_4)$ and induces dependence between the two terms in the covariance. By symmetry of the F_Z, $P(\{\frac{z_j - z_k}{s_j - s_k} > \hat{\theta}\} = 1/2$, so the last expression is

$$\begin{aligned} &E(\{\frac{z_1 - z_2}{s_1 - s_2} > \hat{\theta}\}\{\frac{z_3 - z_4}{s_3 - s_4} > \hat{\theta}\}) - 1/4 \\ &= P(\frac{z_1 - z_2}{s_1 - s_2} \wedge \frac{z_3 - z_4}{s_3 - s_4} > \hat{\theta}) - 1/4. \end{aligned}$$

The symmetry of F_Z further implies that $\{\frac{z_1 - z_2}{s_1 - s_2} > \hat{\theta}\}\{\frac{z_3 - z_4}{s_3 - s_4} > \hat{\theta}\}$ has the same distribution as $\{\frac{z_1 - z_2}{s_1 - s_2} < \hat{\theta}\}\{\frac{z_3 - z_4}{s_3 - s_4} < \hat{\theta}\}$, so the condition for a positive bias, $E(\{\frac{z_1 - z_2}{s_1 - s_2} > \hat{\theta}\}\{\frac{z_3 - z_4}{s_3 - s_4} > \hat{\theta}\}) < 1/4$, is

$$1/2 > P\left(\left\{ \frac{z_1 - z_2}{s_1 - s_2} > \hat{\theta} \right\} \left\{ \frac{z_3 - z_4}{s_3 - s_4} > \hat{\theta} \right\} \right) + P\left(\left\{ \frac{z_1 - z_2}{s_1 - s_2} < \hat{\theta} \right\} \left\{ \frac{z_3 - z_4}{s_3 - s_4} < \hat{\theta} \right\} \right)$$

$$= P(\overline{Z}^T (R_1 R_2^T + R_2 R_1^T) \overline{Z}/2 > 0),$$

where $\overline{Z} = (Z_1, \ldots, Z_n)$ and (for $n \geq 5$) maybe factor out sum

$$\begin{aligned} R_1 &= (\frac{1}{S_1 - S_2} - \frac{S_1}{\sum S_j^2}, \frac{-1}{S_1 - S_2} - \frac{S_2}{\sum S_j^2}, \frac{-S_3}{\sum S_j^2}, \ldots, \frac{-S_n}{\sum S_j^2}) \\ R_2 &= (-\frac{S_1}{\sum S_j^2}, \frac{-S_2}{\sum S_j^2}, \frac{1}{S_3 - S_4} - \frac{S_3}{\sum S_j^2}, \frac{-1}{S_3 - S_4} - \frac{S_4}{\sum S_j^2}, -\frac{S_5}{\sum S_j^2}, \ldots, \frac{-S_n}{\sum S_j^2}). \end{aligned}$$

The two nonzero eigenvalues of $(R_1 R_2^T + R_2 R_1^T)/2$ are $\lambda_a \pm \lambda_b$ where

$$\lambda_a = -1/(2\sum_j S_j^2)$$

$$\lambda_b = \frac{\sqrt{((S_1 - S_2)^2/\sum_j S_j^2 - 2)((S_3 - S_4)^2/\sum_j S_j^2 - 2)}}{2(S_1 - S_2)(S_3 - S_4)}.$$

Then $\lambda_a - \lambda_b \leq 0 \leq \lambda_a + \lambda_b$, and $|\lambda_a + \lambda_b| < |\lambda_a - \lambda_b|$, the negative eigenvalue is larger in magnitude. Let v_1, v_2 denote unit eigenvectors associated respectively to the positive and negative eigenvalues. The condition for a positive bias takes the form

$$P\left(-\frac{\lambda_a}{\lambda_b} > \frac{Z^T(v_1^{\otimes 2} - v_2^{\otimes 2})Z}{Z^T(v_1^{\otimes 2} + v_2^{\otimes 2})Z}\right) > 1/2. \tag{4}$$

The ratio $-\lambda_a/\lambda_b$ is > 0 of order $1/n$. A sufficient condition for a positive bias is then $P(Z^T(v_1^{\otimes 2} - v_2^{\otimes 2})Z < 0) \geq 1/2$ or

$$P(|Z^T v_1| < |Z^T v_2|) \geq 1/2. \tag{5}$$

The projections $Z^T v_1, Z^T v_2$ are uncorrelated with mean zero. When Z is gaussian, they are iid conditionally on S, and (5) holds with equality. In general, however, whether (4) or (5) holds, and therefore whether the bias is positive or negative, depends on the relationship between the distribution of Z and of S.

2.4 Bias in the Gaussian model

The study effects are often modeled as gaussian by appealing to the CLT, e.g., in the original paper [2]. In this situation, the bias takes the form given by Theorem 1. proofs to appendix depending on journal

Theorem 1. *Given iid $(Y_j, S_j), j = 1, \ldots, n$, such that $(Y_1, \ldots, Y_n)|(S_1, \ldots, S_n) \sim N(0, \mathrm{diag}(1/S_1^2, \ldots, 1/S_n^2))$, $S > 0$ a.s., S has a continuous lebesgue density, $E(S^2) < \infty$, and $P(S^2 \leq s)$ is $O(s^\epsilon)$ for some $\epsilon > 0$, then $Var(\sqrt{n}\tau) \to 4/9 - \frac{(E|S_1 - S_2|)^2}{\pi E(S^2)}$.*

Proof. As discussed above, any bias in the asymptotic variance relative to 4/9 is due to the term (3)

$$Cov_0 = P(\frac{Z_1 - Z_2}{S_1 - S_2} \wedge \frac{Z_3 - Z_4}{S_3 - S_4} > \hat{\theta}) - 1/4. \tag{6}$$

The result follows on showing that

$$P\left(\frac{Z_1 - Z_2}{S_1 - S_2} \wedge \frac{Z_3 - Z_4}{S_3 - S_4} > \hat{\theta} \mid S_1, \ldots, S_n\right) - \frac{1}{4} = \frac{(S_1 - S_2)(S_3 - S_4)}{4\pi \sum_{j=5}^n S_j^2} + o(1/\sum_{j=5}^n S_j^2)$$

and n times the rhs is uniformly integrable, since then the LLN gives

$$
\begin{aligned}
Cov(\sqrt{n}\tau) - 4/9 &= 4nCov_0 + o(1) \\
&= 4nE\left(\frac{(S_1 - S_2)(S_3 - S_4)}{4\pi \sum_{j=5}^{n} S_j^2} + o(1/\sum_{j=1}^{n} S_j^2)\right) + o(1) \qquad (7) \\
&\to E\left(\frac{(E(S_1 - S_2))^2}{\pi E(S^2)}\right).
\end{aligned}
$$

Since $Z_1, \dots Z_n$, are independent of $S_1, \dots, S_j$, (6) is

$$
P\left(\frac{\sum_{j=5}^{n} Z_j S_j}{\sum_j S_j^2} < \left(\frac{Z_1 - Z_2}{S_1 - S_2} - \frac{\sum_{j=1}^{4} Z_j S_j}{\sum_j S_j^2}\right) \wedge \left(\frac{Z_3 - Z_4}{S_3 - S_4} - \frac{\sum_{j=1}^{4} Z_j S_j}{\sum_j S_j^2}\right) \Big| \overline{S}_n\right)
$$

$$
= E\left(\Phi\left(\frac{\sum_j S_j^2}{\sqrt{\sum_{j=5}^{n} S_j^2}}\left(\frac{Z_1 - Z_2}{S_1 - S_2} - \frac{\sum_{j=1}^{4} Z_j S_j}{\sum_j S_j^2}\right) \wedge \left(\frac{Z_3 - Z_4}{S_3 - S_4} - \frac{\sum_{j=1}^{4} Z_j S_j}{\sum_j S_j^2}\right)\right) \Big| \overline{S}_n\right)
$$

$$
= E\left(\Phi(W_0 \wedge W_1) \mid \overline{S}_n\right), \qquad (8)
$$

where W_0, W_1 are conditionally jointly normal with mean 0, variances

$$
V_0 = \frac{2}{(S_1 - S_2)^2} + \frac{2}{\sum_j S_j^2} + \frac{\sum_1^4 S^2}{(\sum_j S_j^2)^2}, \quad V_1 = \frac{2}{(S_3 - S_4)^2} + \frac{2}{\sum_j S_j^2} + \frac{\sum_1^4 S^2}{(\sum_j S_j^2)^2}
$$

and covariance

$$
\rho\sqrt{V_0 V_1} = -\frac{2}{\sum_j S_j^2} + \frac{\sum_1^4 S^2}{(\sum_j S_j^2)^2}.
$$

The density $f_{W_0 \wedge W_1}$ of the minimum of a bivariate normal pair is readily available and substitution into (8) gives

$$
\begin{aligned}
&E\left(\Phi(W_0 \wedge W_1) \mid \overline{s}_n\right) = \int_{-\infty}^{\infty} \Phi(u) f_{W_0 \wedge W_1}(u) du \\
&= \sum_{j \in \{0,1\}} \frac{1}{\sqrt{V_j}} \int_{-\infty}^{\infty} \Phi(u)\Phi\left(\frac{u}{\sqrt{1-\rho^2}}\left(\frac{\rho}{\sqrt{V_j}} - \frac{1}{\sqrt{V_{1-j}}}\right)\right)\phi\left(\frac{u}{\sqrt{V_j}}\right) du.
\end{aligned}
$$

The integral inside the sum has the form

$$
\int_{-\infty}^{\infty} \Phi(\alpha_j u)\Phi(\beta_j u)\phi(\gamma_j u) du
$$

with $\alpha_j = 1, \beta_j = \frac{u}{\sqrt{1-\rho^2}}\left(\frac{\rho}{\sqrt{V_j}} - \frac{1}{\sqrt{V_{1-j}}}\right)$, and $\gamma_j = \frac{1}{\sqrt{V_j}}$. This integral may be computed by differentiating first with respect to α and then β to put it in terms of elementary functions,

$$\int_{-\infty}^{\infty} u^2\phi(\alpha_j u)\phi(\beta_j u)\phi(\gamma_j u)du = \frac{1}{2\pi(\alpha_j^2+\beta_j^2+\gamma_j^2)^{3/2}},$$

then integrating twice to obtain $\frac{1}{2\pi\gamma_j}\arctan\left(\frac{\alpha_j\beta_j}{\gamma_j\sqrt{\alpha_j^2+\beta_j^2+\gamma_j^2}}\right)$ as the required definite integral up to a constant. The constant may be determined by taking limits as $1/(4\gamma_j)$, giving

$$E\left(\Phi(W_0 \wedge W_1) \mid \overline{s}_n\right) = \frac{1}{2} + \frac{1}{2\pi}\sum_j \arctan\left(\frac{\alpha_j\beta_j}{\gamma_j\sqrt{\alpha_j^2+\beta_j^2+\gamma_j^2}}\right).$$

Let $\Delta_0 = S_1 - S_2$ and $\Delta_1 = S_3 - S_4$. After simplification, the last expression is

$$\frac{1}{2} + \frac{1}{2\pi}\sum_j \arctan\left(-\frac{1}{\Delta_j^2}\left(\frac{1}{\Delta_0^2\Delta_1^2} + \left(\frac{1}{\Delta_0^2}+\frac{1}{\Delta_1^2}\right)\frac{1}{2\sum_j S_j^2}\right)^{-1/2}\right).$$

Let $u = 1/\sum_{j=5}^n S_j^2$, so by the LLN $u \to 0$ almost surely. When $u = 0$, the last expression is $1/2 + (2\pi)^{-1}\sum_j \arctan\left(-|\frac{\Delta_{1-j}}{\Delta_j}|\right) = 1/4$. Expanding about $u = 0$ and simplifying,

$$E\left(\Phi(W_0 \wedge W_1) \mid \overline{s}_n\right) = \frac{1}{4} + \frac{u}{2\pi}\sum_j \frac{\partial}{\partial u}\arctan\left(-\frac{1}{\Delta_j^2}\left(\frac{1}{\Delta_0^2\Delta_1^2} + \left(\frac{1}{\Delta_0^2}+\frac{1}{\Delta_1^2}\right)\frac{1}{2\sum_j S_j^2}\right)^{-1/2}\right)\Bigg|_{u=0} + o(u)$$

$$= \frac{1}{4} + \frac{u}{4\pi}\Delta_0\Delta_1 + o(u).$$

Passing the limit into the expectation in 7 follows from the eventual uniform integrability of the sequence $n \cdot u = n/\sum_{j=5}^n S_j^2, n = 1, 2, \ldots,$ which in turn follows from Lemma 1.

□

Lemma 1. *If $X_1, X_2, \ldots$ are nonnegative and iid, then the sequence of reciprocals of the sample means $n/(\sum_{j=1}^n X_j), n = n_0, n_0+1, \ldots,$ is uniformly integrable for some n_0 if and only if the common cdf of the X_j is $O(x^\epsilon)$ for some $\epsilon > 0$.*

Proof. First, $n/(\sum_{j=1}^n X_j)$ has moments > 1, say $1+\epsilon$, for n large enough. As

$P(\frac{1}{n}\sum_{j=1}^n X_j < x) \le P(X_1 < nx)^n$,

$$\begin{aligned}
E\left(\left(\frac{n}{\sum_{j=1}^n X_j}\right)^{1+\epsilon}\right) &= (1+\epsilon)\int_0^\infty x^\epsilon P\left(\frac{n}{\sum_{j=1}^n X_j} > x\right) dx \\
&\le (1+\epsilon)\left(1 + \int_1^\infty x^\epsilon P\left(\sum_{j=1}^n X_j < \frac{n}{x}\right) dx\right) \\
&\lesssim (1+\epsilon)\left(1 + \int_1^\infty x^\epsilon \left(\frac{n}{x}\right)^{n\epsilon} dx\right),
\end{aligned}$$

which is finite for $n > 1/\epsilon - 1$. Next, for such n, the sample means $\frac{1}{n}\sum_{j=1}^n X_j, n = 1, 2, \ldots$, are a reverse martingale with respect to $\mathcal{F}_n = \sigma\{\sum_{j=1}^n X_j, \sum_{j=1}^{n+1} X_j, \ldots\}$. The conditional form of Jensen's inequality applied to the convex function $x \mapsto x^{-(1+\epsilon)}$ on $\mathbb{R}^+$ gives, for $k \in \mathbb{N}$,

$$\begin{aligned}
E\left(\left(\frac{n+k}{\sum_{j=1}^{n+k} X_j}\right)^{1+\epsilon}\right) &= E\left(\left(E\left(\frac{1}{n}\sum_{j=1}^n X_j \middle| \mathcal{F}_{n+k}\right)\right)^{-(1+\epsilon)}\right) \\
&\le E\left(\left(\frac{1}{n}\sum_{j=1}^n X_j\right)^{-(1+\epsilon)}\right) = E\left(\left(\frac{n}{\sum_{j=1}^n X_j}\right)^{1+\epsilon}\right).
\end{aligned}$$

The reciprocals of the sample means are therefore L^p–bounded with $p = 1+\epsilon$ for all large n, implying that they are uniformly integrable.

Conversely, if $P(X < x)$ isn't $O(x^\epsilon)$ for any ϵ, there are sequences $\epsilon_n \to 0, x_n \to 0, x_n < 1$, such that $P(X < x_n) > x_n^{\epsilon_n}$. Then as $P(\frac{1}{m}\sum_{j=1}^m X_j < x_n) > P(X < x_n)^m > x_n^{m\epsilon_n}$,

$$\begin{aligned}
E\left(\frac{m}{\sum X_j}\right) &= \int_0^\infty P\left(\frac{m}{\sum X_j} > x\right) dx \\
&\ge \int_1^\infty P\left(\frac{1}{m}\sum_j X < 1/x\right) dx \\
&\ge \sum_{j=1}^\infty x_j^{m\epsilon_j}(1/x_{j+1} - 1/x_j) \\
&\ge \sum_{j=j_0}^\infty (x_j/x_{j+1} - 1),
\end{aligned}$$

with j_0 is chosen so that $m\epsilon_j < 1$ when $j \ge j_0$. The condition $x_n \to 0$ then implies

$$\sum_{j=j_0}^{\infty}(x_j/x_{j+1}-1) \geq \sum_{j=j_0}^{\infty}\log(x_j/x_{j+1})$$
$$= \log\left(\prod_{j=j_0}^{\infty} x_j/x_{j+1}\right) = \infty,$$

so the reciprocals of the sample means aren't integrable. □

Theorem 1 states that when the study effects are gaussian the bias depends on the ratio

$$b = \frac{(E|S_1 - S_2|)^2}{\pi E(S^2)}$$

of the squared mean absolute difference to the second moment of the distribution of study precisions. The quantity is scale free, but depends on the location of S through the denominator. Jensen's inequality gives $(E|S_1 - S_2|)^2 \leq E((S_1 - S_2)^2) = 2Var(S) \leq 2E(S^2)$ so that $b \leq 2$ and the magnitude of the asymptotic bias of the estimator is bounded above by $2/\pi$. This bound is only achieved when the distribution of S degenerates to a point mass at 0. Viewing $1/E(S^2)$ as the harmonic mean of the study variances σ^2, another bound is $b \leq E(\sigma^2)(E|S_1 - S_2|)^2$.

For S uniformly distributed on $[a, b], a \geq 0$, $b = \frac{(b-a)^2}{3(a^2+b^2+ab)}$, which is maximized at 1/3 when $a = 0$. For the general beta distribution with parameters $a > 0, b > 0$, the form of b is given in Table 1. A continued fraction approximation suggests it is maximized at $(a, b) \approx (.15, .39)$ with value $\approx .66$, a bimodal density (Fig. 2.4).

For S exponentially distribution, $b = 1/2$. For a gamma distribution with shape parameter a, $b = \frac{16a}{a+1}(B(1/2, a, a+1) - 1/2)^2$, where $B(1/2, k, k+1) = \int_0^{1/2} x^k(1-x)^{k+1}dx$ is an incomplete beta integral. Numerical evaluation suggests this expression is maximized at $k \approx .54$ with value $\approx .56$.

When S has a pareto distribution with shape parameter a, i.e., $f_S(s) = a\sigma^a/s^{a+1}\{s > \sigma\}$ when the scale parameter is σ,

$$b = \frac{4a(a-2)}{((2a-1)(a-1))^2},$$

defined for $a > 2$ and maximized over a at the unique root of $-2a^3 + 6a^2 - 2a - 1$ on $(2, \infty)$,

$$a = 1 + 2\sqrt{2/3}\cos\left(1/3\arctan\left(\sqrt{101/27}\right)\right) \approx 2.53,$$

where it takes the value $\approx .14$.

In practice, one may attempt to estimate b using the data. This approach may introduce bias of its own, particularly when using common ratio estimators in conjunction with smaller meta-analyses.

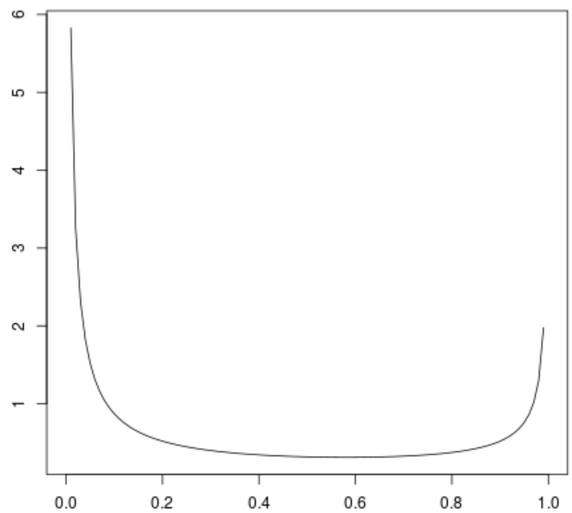

Figure 1: The beta distribution maximizing ρ (obtained numerically).

3 Simulations

We consider the Types I and II error rates for the standard and debiased versions of Begg's test. To carry out the debiased version of Begg's test, we used both the true asymptotic bias, which depends on the distribution of S, and an approximation based on the data. For the latter, given a sample of study precisions $s_1, \ldots, s_n$, $b = (E|S_1 - S_2|)^2/E(S^2)$ was estimated by

$$\left(\binom{n}{2}^{-1} \sum_{j<k} |s_j - s_k| \right)^2 \Big/ \overline{s^2}.$$

The scripts used to run the simulations and produce the figures in this section, as well as a supporting `R` package, are available at `https://github.com/haben-michael/begg-public`.

1. FPR control

   To examine the Type 1 error rate, data is generated under the null model (2.3). The distributions of S considered were uniform, beta, exponential, gamma, and pareto. Three meta-analysis sample sizes were considered, 25, 75, and 150. The test was conducted at a nominal level of 5%. There were 1000 monte carlo repetitions.

The results are presented in Table 2. The unadjusted Begg test is conservative and consistent with the discussion in Section 2.2 this bias does not go away with increased sample size. The magnitude of the asymptotic bias follows the order one expects from Table 1: i.e., beta, gamma, uniform, pareto from most to least severe. For the small meta-analyses the corrected test exceeds the nominal level by 1–3%, matching the nominal level for the larger meta-analysis. There does not appear to be much loss in approximating the correct variance using the data rather than using the true variance, even for the smaller meta-analysis.

2. Type II error rate

   We examine the Type II error rate under an alternative considered by [2]. Under this alternative, a study with p-value p is selected for publication with probability $\propto \exp(-bp)$, with $b \geq 0$. The parameter b controls the strength of selection, with $b = 0$ corresponding to no selection. The choice of selection function was informed by studies of selection bias contemporaneous with [2]. The distrbutions of S considered for this simulation were uniform, beta, exponential and gamma, and the size of the meta-analyses considered were 25, 75, and 150.

   Power curves are presented in Fig. 3. For interpretability, the alternatives are parameterized by the proportion of studies selected, rather than b. The improvement in power across distributions and sizes of has a median value of 17%. The estimator based on the approximation to the true asymptotic variance performs similarly to the oracle estimator. maybe show this by simulation.

4 Data analysis

As an application, we describe three meta-analyses chosen to illustrate different conclusions drawn by the standard and bias-corrected Begg's test. In the typical situation that Begg's test reports a p-value far from the analyst's chosen level, the bias-corrected test will agree with the biased test. In these examples, the standard Begg's test reports p-values in the range $5 - -10\%$ and would be insignificant at the 5% level. The three examples were chosen to contrast the stength of evidence for publication bias conveyed by a funnel plot, the conventional informal test for publication bias. In the first, the authors see little evidence for publication bias based on a funnel plot. In the second, the authors are unable to determine the risk of publication bias. In the third, the authors caution that studies have likely been omitted.

In a 2005 analysis, [6] assess the therapeutic effects of alpha-glucosidase inhibitors in treating type 2 diabetes mellitus. As part of this analysis, they examine the change in body weight under treatment. In a meta-analysis based on 13 randomized trials of at least 12 weeks' duration together involving 864 subjects, the authors find little or no effect of the treatment on weight, contrary to expectations. The authors assess the likelihood of publication bias as low based on a funnel plot (Fig. 4). On this data, Begg's test gives a p-value of 7.4%, whereas the bias-corrected test rejects at the 5% level with a p-value of 2.7%. In this case the possibility of publication bias, suggested by the bias-corrected test, is arguably of less concern as the subsequent meta-analysis can't be a false positive, failing to be significant at the chosen level.

In a 2016 analysis, [5] assess the therapeutic effects of intravenously administered paracetamol in treating postoperative pain. As part of their analysis, they examine the reduction in opioids administered under treatment by paracetamol. In a meta-analysis of 13 randomized trials together involving 777 subjects, the authors find a highly significant reduction of 1.92 mg, with a 95% CI $(-2.41, -1.42)$. The authors assess the quality of the data as moderate, noting that the risk of selective reporting is unclear. In this case, a standard Begg's test gives a p-value 5.2%, just insignificant at the 5% level, whereas the bias-corrected test returns a p-value 2.6%, suggesting publication bias is present and casting doubt on the validity of the subsequent meta-analysis.

In a 2020 analysis, [4] assess the effect of low-fat intake on body weight in populations not seeking to lose weight. The authors conduct a meta-analysis on 26 randomized trials of at least six months' duration involving $50,907$ subjects. They find that a low-fat diet is associated with a -1.56 kg difference in weight, with a 95% CI $(-1.88, -1.23)$, and a p-value reported as $< .00001$. Though the trials in this analysis were chosen in part for their low summary risk of bias, the authors warn of the possibility of selection bias based on a funnel plot (Fig. 4) and other analyses. Begg's test, however, gives a p-value 7.8%, whereas the corrected test gives a p-value 3.3%, consistent with the authors' suspcicions.

5 Conclusion

We have examined the causes of a known bias in Begg's test and quantified it in the common model of normally distributed studies. We have also suggested a debiased estimator that matches the nominal variance in the limit as the number of studies grows, unlike the standard estimator, which remains conservative. In simulations the corrected estimator is somewhat anticonservative on smaller meta-analyses, although from the perspective of an analyst concerned primarily about the integrity of the meta-analysis, exceeding the nominal level may be preferable to falling under it, as discussed earlier. maybe mention the poor performance of beggs test

References

[1] Colin Begg. Publication bias. In Harries Cooper and Larry Hedges, editors, *The Handbook of Research Synthesis*, pages 399–409. Russel Sage Foundation, New York, 1994.

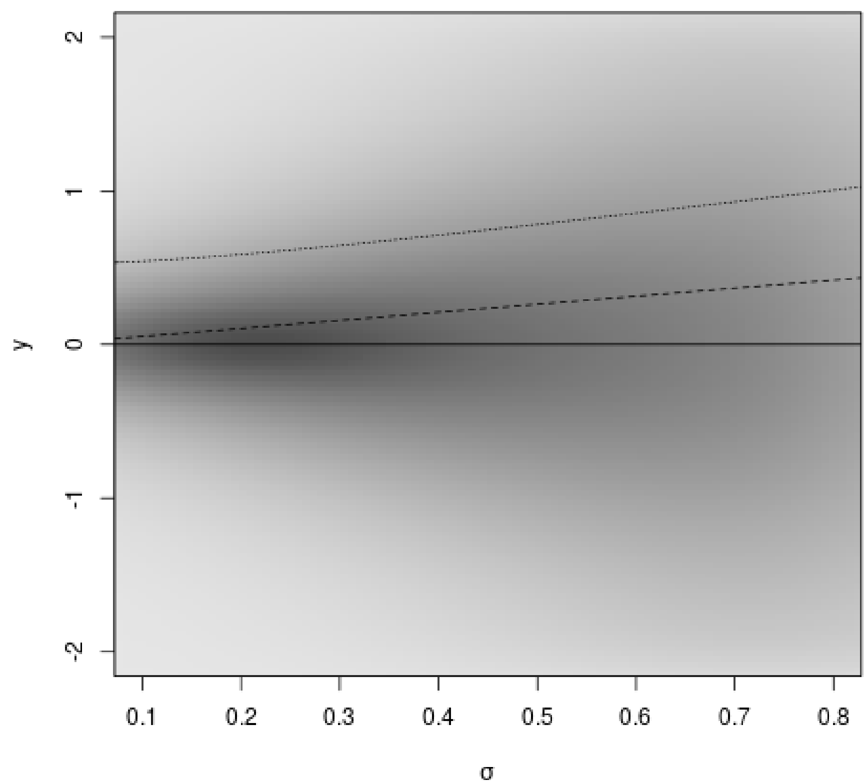

Figure 2: The effect of a simple hard thresholding selection model on the study means. Overlaid on the joint density of Y and σ is the mean of $Y|\sigma$ before selection, $\theta = 0$, exhibiting no trend (solid line), a threshold obtained by rejecting studies with a one-sided p-value $> .3$ (dashed line), and the mean of the studies after selection, exhibiting a trend that Begg's test can pick up (dotted line).

distribution of S	b	max asymptotic bias
uniform	$\frac{(b-a)^2}{3(a^2+b^2+ab)}$	.11
beta	$\left(\frac{4B(a+b,a+b)}{B(a,a)B(b+b)}\right)^2 \frac{a+b+1}{a(a^2+ab+a+b)}$	.21
gamma	$\frac{16a}{a+1}(B(1/2,a,a+1)-1/2)^2$	.18
pareto	$\frac{4a(a-2)}{((2a-1)(a-1))^2}$	.04

Table 1: Shape families of some nonnegative RVs, $\rho = \text{bias} \times \pi = (E|S_1 - S_2|)^2/E(S^2)$ in terms of the shape parameters, and the maximum bias of the asymptotic variance over the shape family.

	meta-analysis size		
precision distribution	25	75	150
uniform	0.03, 0.07, 0.07	0.03, 0.06, 0.06	0.03, 0.05, 0.05
exponential	0.02, 0.08, 0.08	0.01, 0.05, 0.05	0.01, 0.05, 0.05
gamma	0.02, 0.07, 0.07	0.02, 0.06, 0.06	0.01, 0.05, 0.05
beta	0.02, 0.08, 0.08	0.01, 0.05, 0.05	0.01, 0.05, 0.05
pareto	0.04, 0.06, 0.06	0.04, 0.05, 0.05	0.04, 0.06, 0.05

Table 2: False positive rates for standard Begg's test, after debiasing using true asymptotic bias, and after debiasing using estimated asymptotic bias [fix n=76 to 75].

[2] Colin Begg and Madhuchhanda Mazumdar. Operating characteristics of a rank correlation test for publication bias. *Biometrics*, pages 1088–1101, 1994.

[3] Miriam Gjerdevik and Ivar Heuch. Improving the error rates of the begg and mazumdar test for publication bias in fixed effects meta-analysis. *BMC Medical Research Methodology*, 14(1):1–16, 2014.

[4] Lee Hooper, Asmaa S Abdelhamid, Oluseyi F Jimoh, Diane Bunn, and C Murray Skeaff. Effects of total fat intake on body fatness in adults. *Cochrane Database of Systematic Reviews*, (6), 2020.

[5] Ewan D McNicol, McKenzie C Ferguson, Simon Haroutounian, Daniel B Carr, and Roman Schumann. Single dose intravenous paracetamol or intravenous propacetamol for postoperative pain. *Cochrane database of systematic reviews*, (5), 2016.

[6] Floris A Van de Laar, Peter LBJ Lucassen, Reinier P Akkermans, Eloy H Van de Lisdonk, Guy EHM Rutten, and Chris Van Weel. Alpha-glucosidase inhibitors for type 2 diabetes mellitus. *Cochrane database of systematic reviews*, (2), 2005.

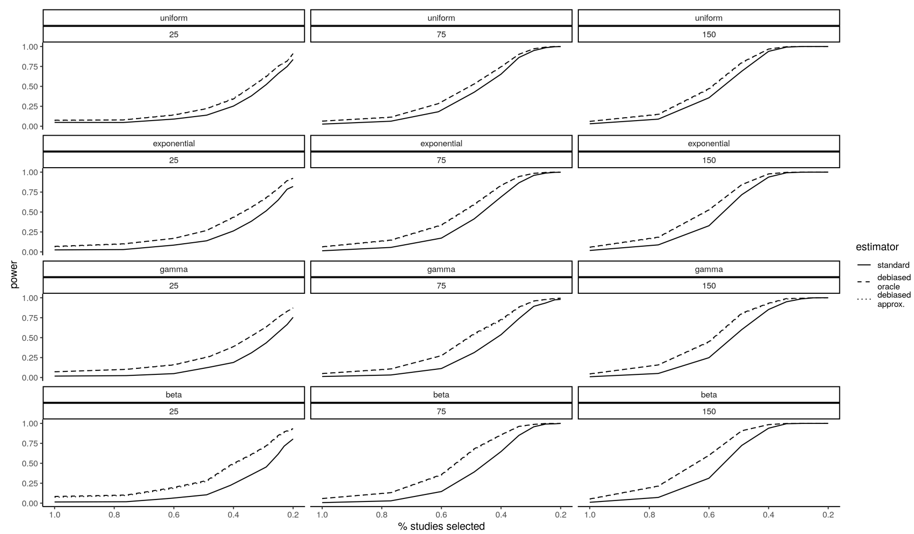

Figure 3: Power curves of standard Begg's test, after debiasing using true asymptotic bias, and after debiasing using estimated asymptotic bias. The alternatives are parameterized by the proportion of studies selected. The estimator debiased using an estimate of the true bias and the true bias itself overlap. increase size of plot

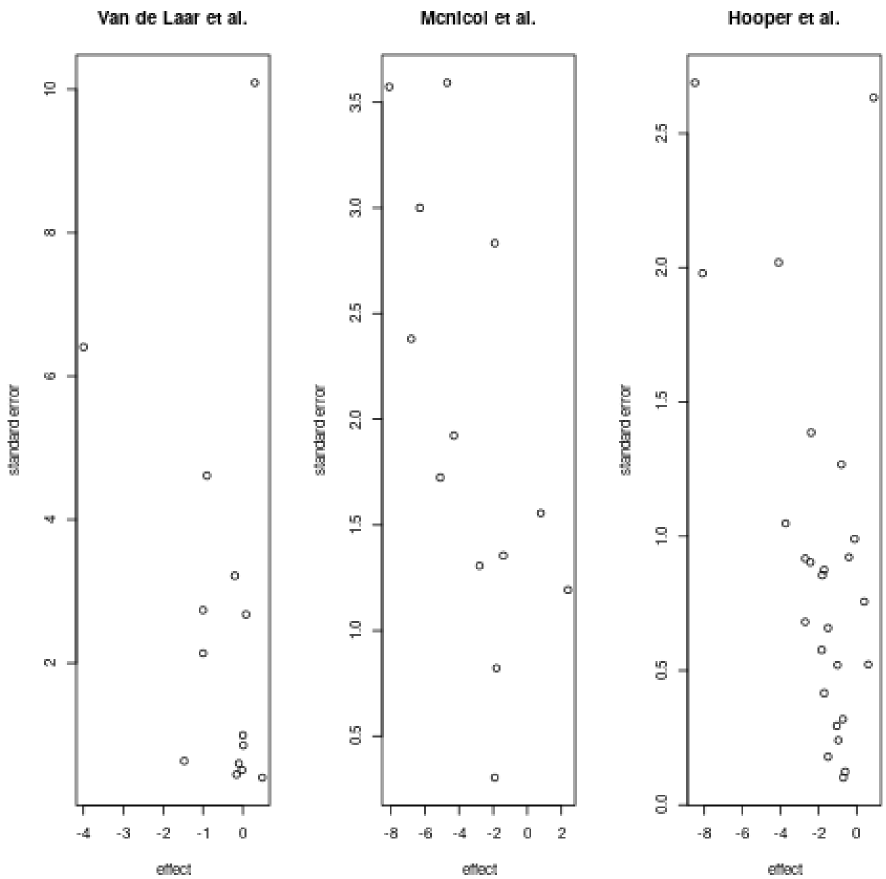

Figure 4: Funnel plots of the three meta-analyses described in Section 4. These were interpreted as suggesting lower, moderate, and higher possibilities of publication bias, going from left to right.