Supplementary Material for

# Differential Expression in SAGE: Accounting for Normal Between-Library Variation

**Keith A. Baggerly***
Department of Biostatistics
UT M.D. Anderson Cancer Center
Houston, TX 77030-4009

**Li Deng**
Department of Statistics
Rice University
Houston, TX 77005

**Jeffrey S. Morris**
Department of Biostatistics
UT M.D. Anderson Cancer Center
Houston, TX 77030-4009

**C. Marcelo Aldaz**
Department of Carcinogenesis
UT M.D. Anderson Cancer Center
Houston, TX 77030-4009

January 6, 2003

Email: kabagg@mdanderson.org

## Unconditional Mean and Var of $X_i/n_i$

Here,

$$p_i \sim \text{Beta}(\alpha,\beta), \quad E(p_i) = \frac{\alpha}{\alpha+\beta}, \quad V(p_i) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

The second part of our model says that given the true proportion in a sample, the corresponding count will have a binomial distribution with the true proportion as a parameter:

$$X_i|p_i \sim \text{Binomial}(n_i, p_i).$$

to get the unconditional mean and variance of the estimated proportion $\hat{p}_i = X_i/n_i$ we make use of the tower property of conditional expectation, that $E(X) = E(E(X|Y))$. Here, this yields

$$\begin{aligned}
E(X_i) &= E\left[E(X_i|p_i, n_i)\right] \\
&= E(n_i p_i) \\
&= n_i \frac{\alpha}{\alpha+\beta} \\
E(X_i/n_i) &= \frac{\alpha}{\alpha+\beta} \\
E(X_i^2) &= E\left[E(X_i^2|p_i, n_i)\right] \\
&= E(n_i p_i(1-p_i) + (n_i p_i)^2) \\
&= n_i \frac{\alpha}{\alpha+\beta} + n_i(n_i-1)\left[\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} + \frac{\alpha^2}{(\alpha+\beta)^2}\right] \\
V(X_i) &= n_i \frac{\alpha}{\alpha+\beta} + n_i(n_i-1)\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} - n_i \frac{\alpha^2}{(\alpha+\beta)^2} \\
&= n_i^2 \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} + n_i\left[\frac{\alpha}{\alpha+\beta} - \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2}\right] \\
&= \frac{\alpha}{\alpha+\beta}\left\{n_i^2 \frac{\beta}{(\alpha+\beta)(\alpha+\beta+1)} + n_i\left[1 - \frac{\beta}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha}{\alpha+\beta}\right]\right\} \\
&= \frac{\alpha}{\alpha+\beta}\left\{n_i^2 \frac{\beta}{(\alpha+\beta)(\alpha+\beta+1)} + n_i\left[\frac{\alpha^2+2\alpha\beta+\beta^2+\alpha+\beta-\beta-\alpha^2-\alpha\beta-\alpha}{(\alpha+\beta)(\alpha+\beta+1)}\right]\right\} \\
&= \frac{\alpha}{\alpha+\beta}\left\{n_i^2 \frac{\beta}{(\alpha+\beta)(\alpha+\beta+1)} + n_i\left[\frac{\beta(\alpha+\beta)}{(\alpha+\beta)(\alpha+\beta+1)}\right]\right\} \\
&= \frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)}\left[n_i^2 \frac{1}{\alpha+\beta} + n_i\right] \\
V(X_i/n_i) &= \frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)}\left[\frac{1}{\alpha+\beta} + \frac{1}{n_i}\right].
\end{aligned}$$

## Derivation of $\hat{V}_{unb}$

We recall that the general form of the variance of a single proportion is

$$V(X_i/n_i) = \sigma_1^2 + \frac{\sigma_2^2}{n_i},$$

with the first piece (which doesn't change with $n_i$) coming from the variation between libraries and the second (which changes with $n_i$) from the sampling variability within a library. The explicit values of $\sigma_1^2$ and $\sigma_2^2$ can be found by matching terms with the value for $V(X_i/n_i)$ found in the previous section; we have simply found this notation a convenient shorthand. Given this form, the variance of a weighted combination of proportions is of the form

$$V(\sum w_i(X_i/n_i)) = \sum w_i^2 \sigma_1^2 + \sum w_i^2 \frac{\sigma_2^2}{n_i}.$$

This is the quantity that we wish to estimate. Now, to check the bias of certain estimators, we need some expectations. Specifically, we note that

$$\begin{aligned} E(X_i/n_i) &= \mu \\ E((X_i/n_i)^2) &= \sigma_1^2 + \frac{\sigma_2^2}{n_i} + \mu^2. \end{aligned}$$

Again, the value for $\mu$ can be found by plugging in the value from the previous section. Note that the expectation of the quadratic term is different for different proportions. Starting with the estimate

$$\hat{V} = \sum w_i (X_i/n_i)^2 - \left(\sum w_i (X_i/n_i)\right)^2$$

and taking expectations yields

$$\begin{aligned} E(\hat{V}) &= \sum w_i \left(\sigma_1^2 + \frac{\sigma_2^2}{n_i} + \mu^2\right) - \left(\mu^2 + \sum w_i^2 \left(\sigma_1^2 + \frac{\sigma_2^2}{n_i}\right)\right) \\ &= \sigma_1^2 \left(\sum w_i(1-w_i)\right) + \sigma_2^2 \left(\sum w_i(1-w_i)/n_i\right). \end{aligned}$$

Now, we want the multiplier for $\sigma_1^2$ to be $\sum w_i^2$, so we could multiply $\hat{V}$ by the ratio

$$\frac{\sum w_i^2}{\sum w_i(1-w_i)}$$

to achieve this. Unfortunately, we want the multiplier for $\sigma_2^2$ to be $\sum w_i^2/n_i$, and the appropriate multiplication factor for achieving this is

$$\frac{\sum w_i^2/n_i}{\sum w_i(1-w_i)/n_i},$$

which is different than the factor suggested above. Thus, we cannot get an unbiased estimate of the quantity we want as a scalar multiple of an actual sample variance (which is guaranteed to be positive).

This problem has a few fixes. The fix we propose is associated with the idea of estimating variance components – essentially, finding unbiased estimates of $\sigma_1^2$ and $\sigma_2^2$ and combining them. Now, these estimates start from the estimate of the two combined, and involve some subtraction steps, so we can wind up with negative estimates. The result of one such excursion (based on trial and error) is

$$\hat{V}_0 = \sum w_i^2 (X_i/n_i)^2 - (\sum w_i^2) \left(\sum w_i (X_i/n_i)\right)^2 .$$

Taking expectations here gives

$$\begin{aligned} E(\hat{V}_0) &= \sum w_i^2 \left(\sigma_1^2 + \frac{\sigma_2^2}{n_i} + \mu^2\right) - (\sum w_i^2) \left[\mu^2 + \sum w_i^2 \left(\sigma_1^2 + \frac{\sigma_2^2}{n_i}\right)\right] \\ &= \left(1 - \sum w_i^2\right) * \sigma_1^2 \sum w_i^2 + \left(1 - \sum w_i^2\right) * \sigma_2^2 \sum w_i^2/n_i \end{aligned}$$

so that the quantity

$$\hat{V}_{unb} = \hat{V}_0/(1 - \sum w_i^2)$$

is actually an unbiased estimate of the variance that we want. Empirically, this estimator does give rise to negative variances in some situations, which is one reason we couple it with a functional floor in the paper.

## Some Simulation Results

In order to assess the performance of our new statistic, $t_w$, we ran some small simulations comparing it with the two-sample $t$-test (with no assumption of equal variances) and with the $\chi^2$ test (as a representative member of the class of pooled tests). For the simulations, we generated data as follows. First, the true mean proportions for group $A$, $p_A$, was generated. Proportions (as counts out of 50K) of 1, 2, 5, 10, and 100 were tried. Then, the mean proportion for group $B$, $p_B$, was specified as a multiple of $p_A$. were specified. Multiples of 1, 2, 3, 5, 10 were tried. Null proportions for each of the individual libraries in a group were then generated from a beta distribution, with the parameters of the beta chosen so as to yield a fixed multiple of overdispersion relative to the straight binomial. The number of libraries in both groups was taken to be 4, and the overdispersion multipliers were taken to be 2, 5, 10, or 50. Given the library proportions, counts of the tag of interest were then generated. This was done using Poisson sampling rather than binomial because of the faster average run time for scarce observations. This process was repeated 10000 times for each combination. In each case, we computed the $t_w$, $t$ and $\chi^2$ values, converted the values to p-values using the approximate asymptotics, and recorded the fraction of the observations found to be significant at the 10%, 5%, 1%, and 0.5% levels. The results are summarized in several tables below.

Qualitatively, we find that $t_w$ and the two-sample $t$-test perform quite similarly throughout, with $t_w$ exhibiting very slightly less power. The $\chi^2$ test has higher power throughout, so the sensitivity is good, but it suffers from an abysmal type I error rate, so the specificity is very bad. The specificity of the $t$ and $t_w$ tests is ok, though the asymptotic bounds we are using here are evidently conservative: We pick up a fraction smaller than the nominal type I error rate as significant when there is in fact no difference, particularly when the counts are low.

Given that the $t$ and $t_w$ tests are quite similar as far as coarse measures of power are concerned, we checked several plots of the simulation results to suggest what differences did exist. In general, the range of $t_w$ is smaller than that of $t$. This is due in large part to the presence of "explosive" $t$ values where the two groups exhibit differences but the sampling variance is less than the binomial variance. In many of the cases where this occurs, the test statistic value is reduced from one extreme value (eg, $t = 30$) to another extreme value (eg $t_w = 12$) so that the overall power of the tests is largely unaffected. Such corrections can, however, serve to reorder the list of the most significantly expressed genes. We have seen this type of correction in real data, with the most dramatic corrections occurring when the number of libraries is small and the improperly low variance estimate is associated with the higher overall proportion. This correction for explosion also happens with low-count genes, and there it often shifts the $t$ values from "significant" to "not significant". This can reduce the perceived power, but it may be more realistic, in that 4 counts of 1 in one group are not persuasively different from 4 counts of 0 in the other group.

It is possible for $t_w$ to be "more significant" than $t$; this can occur if the different variance estimates give rise to different assessements of the degrees of freedom to use. There are other cases we have seen in real data where $t_w$ actually gives a larger value than $t$. These occur when the library sizes are different, and the most outlying values within a group are associated with the smallest libraries.

| $p_B = p_A$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\chi^2$ | | | | $t$ | | | | $t_w$ | | | |
| | $p_A$ | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 |
| | 1 | 2318 | 1391 | 392 | 241 | 550 | 182 | 12 | 5 | 210 | 23 | 0 | 0 |
| | 2 | 2400 | 1557 | 597 | 423 | 771 | 343 | 56 | 33 | 504 | 106 | 0 | 0 |
| $k = 2$ | 5 | 2430 | 1680 | 674 | 483 | 854 | 395 | 67 | 30 | 632 | 213 | 3 | 1 |
| | 10 | 2467 | 1653 | 642 | 443 | 835 | 370 | 64 | 34 | 628 | 221 | 12 | 0 |
| | 100 | 2456 | 1666 | 695 | 455 | 914 | 433 | 83 | 36 | 715 | 277 | 21 | 5 |
| | 1 | 3973 | 3073 | 1534 | 1235 | 229 | 47 | 5 | 4 | 126 | 17 | 0 | 0 |
| | 2 | 4451 | 3549 | 2082 | 1709 | 447 | 159 | 20 | 7 | 363 | 83 | 0 | 0 |
| $k = 5$ | 5 | 4571 | 3770 | 2445 | 2001 | 678 | 290 | 49 | 31 | 635 | 237 | 12 | 3 |
| | 10 | 4590 | 3786 | 2391 | 2014 | 782 | 336 | 66 | 30 | 751 | 307 | 31 | 10 |
| | 100 | 4588 | 3757 | 2497 | 2118 | 917 | 425 | 84 | 32 | 897 | 402 | 71 | 28 |
| | 1 | 4531 | 3741 | 2313 | 2002 | 83 | 9 | 1 | 0 | 45 | 5 | 0 | 0 |
| | 2 | 5530 | 4747 | 3293 | 2912 | 251 | 54 | 10 | 7 | 208 | 39 | 5 | 1 |
| $k = 10$ | 5 | 5939 | 5250 | 3951 | 3536 | 503 | 196 | 23 | 12 | 492 | 182 | 10 | 3 |
| | 10 | 5989 | 5337 | 4091 | 3665 | 626 | 291 | 54 | 30 | 625 | 283 | 46 | 14 |
| | 100 | 6036 | 5308 | 4110 | 3717 | 860 | 368 | 63 | 32 | 857 | 364 | 62 | 30 |
| | 1 | 3140 | 2786 | 2158 | 2005 | 3 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| | 2 | 5017 | 4537 | 3586 | 3362 | 23 | 1 | 0 | 0 | 17 | 1 | 0 | 0 |
| $k = 50$ | 5 | 7261 | 6758 | 5759 | 5467 | 109 | 14 | 2 | 1 | 95 | 13 | 1 | 0 |
| | 10 | 7892 | 7528 | 6730 | 6485 | 269 | 64 | 9 | 4 | 265 | 61 | 8 | 4 |
| | 100 | 8224 | 7872 | 7195 | 6967 | 783 | 339 | 51 | 26 | 783 | 340 | 52 | 27 |

| $p_B = 2 * p_A$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\chi^2$ | | | | $t$ | | | | $t_w$ | | | |
| | $p_A$ | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 |
| $k = 2$ | 1 | 3975 | 2858 | 1391 | 1056 | 1411 | 633 | 107 | 66 | 1003 | 204 | 0 | 0 |
| | 2 | 5081 | 4176 | 2344 | 1879 | 2296 | 1194 | 232 | 139 | 1977 | 783 | 7 | 0 |
| | 5 | 7618 | 6856 | 5125 | 4435 | 4545 | 2827 | 732 | 408 | 4323 | 2502 | 302 | 50 |
| | 10 | 9301 | 8948 | 7875 | 7361 | 7109 | 5254 | 1863 | 1137 | 7001 | 5100 | 1524 | 651 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9814 | 9342 | 10000 | 9999 | 9841 | 9447 |
| $k = 5$ | 1 | 5132 | 4205 | 2514 | 2128 | 630 | 186 | 37 | 25 | 479 | 85 | 0 | 0 |
| | 2 | 5754 | 4993 | 3534 | 3090 | 1108 | 471 | 75 | 37 | 1039 | 380 | 9 | 0 |
| | 5 | 7079 | 6503 | 5288 | 4843 | 2427 | 1266 | 259 | 135 | 2389 | 1238 | 199 | 50 |
| | 10 | 8350 | 7946 | 6972 | 6583 | 3900 | 2321 | 554 | 302 | 3875 | 2311 | 527 | 253 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 9969 | 9794 | 7476 | 5888 | 9969 | 9795 | 7523 | 5969 |
| $k = 10$ | 1 | 5685 | 4884 | 3304 | 2895 | 237 | 40 | 5 | 2 | 186 | 28 | 0 | 0 |
| | 2 | 6465 | 5776 | 4466 | 4091 | 629 | 193 | 35 | 14 | 584 | 161 | 10 | 1 |
| | 5 | 7150 | 6632 | 5596 | 5232 | 1391 | 573 | 98 | 63 | 1387 | 566 | 78 | 27 |
| | 10 | 7957 | 7550 | 6773 | 6428 | 2383 | 1222 | 266 | 146 | 2379 | 1221 | 263 | 134 |
| | 100 | 9990 | 9987 | 9971 | 9968 | 9316 | 8202 | 4142 | 2720 | 9316 | 8204 | 4164 | 2744 |
| $k = 50$ | 1 | 4307 | 3864 | 3045 | 2845 | 4 | 0 | 0 | 0 | 3 | 0 | 0 | 0 |
| | 2 | 6241 | 5716 | 4755 | 4501 | 75 | 11 | 1 | 0 | 57 | 8 | 1 | 0 |
| | 5 | 7964 | 7561 | 6782 | 6529 | 266 | 51 | 5 | 0 | 250 | 47 | 4 | 0 |
| | 10 | 8380 | 8070 | 7442 | 7210 | 600 | 161 | 32 | 16 | 598 | 160 | 30 | 15 |
| | 100 | 9501 | 9411 | 9259 | 9187 | 4010 | 2428 | 601 | 326 | 4010 | 2428 | 601 | 327 |

| $p_B = 3 * p_A$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\chi^2$ | | | | $t$ | | | | $t_w$ | | | |
| | $p_A$ | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 |
| $k = 2$ | 1 | 6380 | 5349 | 3423 | 2851 | 2996 | 1585 | 341 | 208 | 2671 | 959 | 2 | 0 |
| | 2 | 8266 | 7645 | 5917 | 5224 | 5179 | 3227 | 901 | 552 | 4971 | 2832 | 158 | 2 |
| | 5 | 9840 | 9705 | 9248 | 8985 | 8553 | 6876 | 2891 | 1871 | 8513 | 6828 | 2536 | 1104 |
| | 10 | 9998 | 9996 | 9974 | 9957 | 9841 | 9269 | 5821 | 4187 | 9843 | 9276 | 5921 | 4213 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9990 | 10000 | 10000 | 9999 | 9990 |
| $k = 5$ | 1 | 6393 | 5620 | 4026 | 3568 | 1284 | 463 | 74 | 47 | 1149 | 306 | 3 | 0 |
| | 2 | 7564 | 7043 | 5816 | 5365 | 2610 | 1247 | 244 | 127 | 2565 | 1152 | 81 | 11 |
| | 5 | 9212 | 8953 | 8347 | 8047 | 5325 | 3322 | 958 | 531 | 5314 | 3310 | 909 | 430 |
| | 10 | 9870 | 9816 | 9653 | 9556 | 7885 | 5933 | 2231 | 1380 | 7882 | 5935 | 2262 | 1383 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9866 | 9437 | 10000 | 10000 | 9869 | 9448 |
| $k = 10$ | 1 | 6644 | 5975 | 4474 | 4070 | 585 | 147 | 18 | 9 | 508 | 113 | 2 | 0 |
| | 2 | 7547 | 7055 | 5948 | 5567 | 1312 | 476 | 80 | 43 | 1275 | 436 | 32 | 7 |
| | 5 | 8595 | 8269 | 7595 | 7324 | 3112 | 1555 | 356 | 206 | 3110 | 1553 | 339 | 165 |
| | 10 | 9502 | 9396 | 9108 | 8964 | 5266 | 3250 | 888 | 505 | 5267 | 3250 | 901 | 509 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 9998 | 9975 | 8763 | 7383 | 9998 | 9975 | 8767 | 7393 |
| $k = 50$ | 1 | 5216 | 4774 | 3843 | 3628 | 37 | 2 | 0 | 0 | 30 | 0 | 0 | 0 |
| | 2 | 7221 | 6770 | 5740 | 5465 | 141 | 17 | 1 | 0 | 118 | 13 | 0 | 0 |
| | 5 | 8533 | 8250 | 7655 | 7424 | 562 | 150 | 20 | 8 | 554 | 146 | 15 | 6 |
| | 10 | 8890 | 8689 | 8228 | 8066 | 1358 | 494 | 69 | 33 | 1357 | 493 | 65 | 30 |
| | 100 | 9978 | 9972 | 9962 | 9954 | 7993 | 6050 | 2229 | 1363 | 7993 | 6050 | 2229 | 1365 |

| $p_B = 5 * p_A$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\chi^2$ | | | | $t$ | | | | $t_w$ | | | |
| | $p_A$ | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 |
| $k = 2$ | 1 | 9111 | 8717 | 7289 | 6724 | 6224 | 3954 | 1070 | 648 | 6054 | 3531 | 58 | 0 |
| | 2 | 9916 | 9837 | 9560 | 9356 | 8816 | 6982 | 2829 | 1780 | 8796 | 6905 | 2133 | 424 |
| | 5 | 10000 | 10000 | 9999 | 9997 | 9971 | 9723 | 6756 | 5101 | 9971 | 9729 | 6829 | 5204 |
| | 10 | 10000 | 10000 | 10000 | 10000 | 10000 | 9996 | 9142 | 7875 | 10000 | 9996 | 9205 | 7987 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 |
| $k = 5$ | 1 | 8343 | 7918 | 6781 | 6385 | 3063 | 1382 | 250 | 143 | 2990 | 1235 | 53 | 1 |
| | 2 | 9379 | 9168 | 8607 | 8359 | 5519 | 3231 | 824 | 468 | 5512 | 3204 | 629 | 185 |
| | 5 | 9972 | 9963 | 9913 | 9891 | 8830 | 7094 | 2806 | 1788 | 8831 | 7099 | 2823 | 1786 |
| | 10 | 10000 | 10000 | 10000 | 9999 | 9909 | 9414 | 5769 | 4097 | 9909 | 9416 | 5794 | 4137 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9982 | 10000 | 10000 | 10000 | 9982 |
| $k = 10$ | 1 | 8023 | 7551 | 6497 | 6103 | 1465 | 492 | 64 | 31 | 1387 | 424 | 21 | 5 |
| | 2 | 8933 | 8678 | 8087 | 7854 | 3185 | 1393 | 270 | 155 | 3173 | 1374 | 205 | 59 |
| | 5 | 9789 | 9739 | 9569 | 9491 | 6420 | 4010 | 1091 | 648 | 6418 | 4011 | 1095 | 636 |
| | 10 | 9985 | 9981 | 9963 | 9953 | 8852 | 7119 | 2796 | 1760 | 8951 | 7119 | 2807 | 1775 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9925 | 9585 | 10000 | 10000 | 9925 | 9586 |
| $k = 50$ | 1 | 6598 | 6120 | 5132 | 4877 | 121 | 8 | 0 | 0 | 95 | 5 | 0 | 0 |
| | 2 | 8359 | 8021 | 7260 | 7028 | 458 | 85 | 8 | 3 | 428 | 78 | 5 | 1 |
| | 5 | 9195 | 9037 | 8719 | 8570 | 1503 | 503 | 75 | 41 | 1499 | 497 | 68 | 34 |
| | 10 | 9545 | 9464 | 9266 | 9193 | 3266 | 1511 | 316 | 171 | 3266 | 1510 | 317 | 170 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 9930 | 9457 | 5785 | 4120 | 9930 | 9457 | 5785 | 4122 |

| $p_B = 10 * p_A$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\chi^2$ | | | | $t$ | | | | $t_w$ | | | |
| | $p_A$ | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 | 0.1 | 0.05 | 0.01 | 0.005 |
| $k = 2$ | 1 | 9990 | 9982 | 9937 | 9899 | 9533 | 8121 | 3453 | 2254 | 9531 | 8116 | 2854 | 497 |
| | 2 | 10000 | 10000 | 10000 | 10000 | 9980 | 9748 | 6552 | 4744 | 9980 | 9754 | 6613 | 4803 |
| | 5 | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 9461 | 8315 | 10000 | 9999 | 9478 | 8365 |
| | 10 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9981 | 9804 | 10000 | 10000 | 9982 | 9813 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 |
| $k = 5$ | 1 | 9809 | 9730 | 9518 | 9411 | 6802 | 4122 | 1023 | 580 | 6793 | 4116 | 747 | 180 |
| | 2 | 9992 | 9989 | 9975 | 9968 | 9210 | 7399 | 2793 | 1675 | 9210 | 7403 | 2800 | 1640 |
| | 5 | 10000 | 10000 | 10000 | 10000 | 9987 | 9795 | 6575 | 4752 | 9987 | 9797 | 6580 | 4780 |
| | 10 | 10000 | 10000 | 10000 | 10000 | 10000 | 9996 | 9070 | 7610 | 10000 | 9996 | 9073 | 7624 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 |
| $k = 10$ | 1 | 9520 | 9389 | 9075 | 8926 | 4051 | 1846 | 335 | 168 | 4037 | 1819 | 233 | 59 |
| | 2 | 9914 | 9879 | 9792 | 9745 | 6867 | 4226 | 1049 | 587 | 6866 | 4229 | 1044 | 536 |
| | 5 | 10000 | 9998 | 9996 | 9996 | 9583 | 8263 | 3540 | 2181 | 9583 | 8265 | 3545 | 2183 |
| | 10 | 10000 | 10000 | 10000 | 10000 | 9990 | 9787 | 6584 | 4814 | 9990 | 9787 | 6587 | 4822 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 10000 | 9999 | 10000 | 10000 | 10000 | 9999 |
| $k = 50$ | 1 | 8639 | 8308 | 7559 | 7342 | 529 | 96 | 8 | 4 | 479 | 84 | 4 | 2 |
| | 2 | 9455 | 9348 | 9059 | 8964 | 1523 | 462 | 54 | 26 | 1506 | 456 | 49 | 17 |
| | 5 | 9818 | 9783 | 9698 | 9664 | 4251 | 1903 | 355 | 179 | 4251 | 1903 | 356 | 180 |
| | 10 | 9969 | 9961 | 9940 | 9933 | 7059 | 4354 | 1051 | 594 | 7059 | 4354 | 1051 | 594 |
| | 100 | 10000 | 10000 | 10000 | 10000 | 10000 | 9997 | 9112 | 7601 | 10000 | 9997 | 9112 | 7601 |