# Tables of Correlation Features

E. Garcia, admin@miislita.com

Date: November 15, 2010; Last Update: November 25, 2010 
Keywords: correlation features, correlation coefficients, signal, noise, Cohen's d, correlation regimes

## Introduction

Tables of correlation features are presented for the convenience of analysts and for use with statistical and practical significance tests. Analysts requiring additional theory or statistical data corresponding to confidence levels and/or degrees of freedom not covered in the tables are referred to the literature. Our tables are great for data mining studies or for those that don't have access to power analysis software packages.

Features of correlation distributions are compared up to a resolution of ± 0.005 and at several confidence levels and $n - 2$ degrees of freedom. These features are summarized in the Appendix section.

## Reproducing our tables

Our tables can be reproduced with Excel or a similar spreadsheet application by the following procedure.

In one column of a spreadsheet, paste degrees of freedom, $df$, at a given confidence level; e.g., 95% ($\alpha = 0.05$). In a second column, paste Student's $t$ values. You may want to use the *Student's t-distribution Table* available at http://en.wikipedia.org/wiki/Student's_t-distribution. With $df$ and $t$ values several features can be computed. For instance in a third column, compute correlation coefficients defined as

$$r = \frac{t}{\sqrt{df + t^2}}$$

In a fourth column, list the amount of "signal" $S$ associated to $r$, defined as

$$S = r^2$$

Also known as *Coefficient of Determination*, $S$ is the fraction of variations in the dependent variable ($y$) explained by the independent variable ($x$). The fraction of unexplained variations in the dependent variable or "noise" $N$ is therefore

$$N = 1 - S = 1 - r^2$$

and $S/N$ is a signal-to-noise ratio. Construct columns for the $N$ and $S/N$ features.

Compute also a column of $1/t$ values and a column of Cohen's $d$ values using any of the following equations

$$d \quad = \quad 2\frac{r}{\sqrt{1-r^2}} = 2\sqrt{S/N} = 2\frac{t}{\sqrt{df}}$$

where Cohen's $d$ is the difference between any two independent sample means. Note that Cohen's $d$ is approximated as a signal-to-noise function.

For the grand finale, construct columns for comparing any two features $F_1$ and $F_2$ and as follows:

$1/t$ vs $N$ $\quad$ $S$ vs $N$ $\quad$ $S$ vs $1/t$ $\quad$ $r$ vs $N$ $\quad$ $S/N$ vs $r$ $\quad$ $S/N$ vs $1/t$ $\quad$ $r$ vs $1/t$ $\quad$ $d$ vs $N$ $\quad$ $d$ vs $1/t$

You should be able to reproduce the following tables:

**Table 1. Features of a correlation distribution at a 90% confidence level (α = 0.10)**

| *df* | *t* | *r* | *S* | *N* | *S/N* | *1/t* | *d* | *1/t vs N* | *S vs N* | *S vs 1/t* | *r vs N* | *S/N vs r* | *S/N vs 1/t* | *r vs 1/t* | *d vs N* | *d vs 1/t* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 6.314 | 0.988 | 0.976 | 0.024 | 39.867 | 0.158 | | | | | | | | | | |
| 2 | 2.92 | 0.900 | 0.810 | 0.190 | 4.263 | 0.342 | | | | | | | | | | |
| 3 | 2.353 | 0.805 | 0.649 | 0.351 | 1.846 | 0.425 | | 1/t > N | | | | | | | | |
| 4 | 2.132 | 0.729 | 0.532 | 0.468 | 1.136 | 0.469 | | 1/t ≈ N | S > N | S > 1/t | | | | | | |
| 5 | 2.015 | 0.669 | 0.448 | 0.552 | 0.812 | 0.496 | 1.802 | 1/t < N | S < N | S < 1/t | | | | | | |
| 6 | 1.943 | 0.621 | 0.386 | 0.614 | 0.629 | 0.515 | 1.586 | | | | r > N | | S/N > 1/t | | | |
| 7 | 1.895 | 0.582 | 0.339 | 0.661 | 0.513 | 0.528 | 1.432 | | | | r < N | S/N > r | S/N < 1/t | | | |
| 8 | 1.86 | 0.549 | 0.302 | 0.698 | 0.432 | 0.538 | 1.315 | | | | | S/N < r | | r > 1/t | | |
| 9 | 1.833 | 0.521 | 0.272 | 0.728 | 0.373 | 0.546 | 1.222 | | | | | | | r < 1/t | | |
| 10 | 1.812 | 0.497 | 0.247 | 0.753 | 0.328 | 0.552 | 1.146 | | | | | | | | | |
| 11 | 1.796 | 0.476 | 0.227 | 0.773 | 0.293 | 0.557 | 1.083 | | | | | | | | | |
| 12 | 1.782 | 0.457 | 0.209 | 0.791 | 0.265 | 0.561 | 1.029 | | | | | | | | | |
| 13 | 1.771 | 0.441 | 0.194 | 0.806 | 0.241 | 0.565 | 0.982 | | | | | | | | | |
| 14 | 1.761 | 0.426 | 0.181 | 0.819 | 0.222 | 0.568 | 0.941 | | | | | | | | | |
| 15 | 1.753 | 0.412 | 0.170 | 0.830 | 0.205 | 0.570 | 0.905 | | | | | | | | | |
| 16 | 1.746 | 0.400 | 0.160 | 0.840 | 0.191 | 0.573 | 0.873 | | | | | | | | d > N | |
| 17 | 1.74 | 0.389 | 0.151 | 0.849 | 0.178 | 0.575 | 0.844 | | | | | | | | d ≈ N | |
| 18 | 1.734 | 0.378 | 0.143 | 0.857 | 0.167 | 0.577 | 0.817 | | | | | | | | d < N | |
| 19 | 1.729 | 0.369 | 0.136 | 0.864 | 0.157 | 0.578 | 0.793 | | | | | | | | | |
| 20 | 1.725 | 0.360 | 0.130 | 0.870 | 0.149 | 0.580 | 0.771 | | | | | | | | | |
| 21 | 1.721 | 0.352 | 0.124 | 0.876 | 0.141 | 0.581 | 0.751 | | | | | | | | | |
| 22 | 1.717 | 0.344 | 0.118 | 0.882 | 0.134 | 0.582 | 0.732 | | | | | | | | | |
| 23 | 1.714 | 0.337 | 0.113 | 0.887 | 0.128 | 0.583 | 0.715 | | | | | | | | | |
| 24 | 1.711 | 0.330 | 0.109 | 0.891 | 0.122 | 0.584 | 0.699 | | | | | | | | | |
| 25 | 1.708 | 0.323 | 0.104 | 0.896 | 0.117 | 0.585 | 0.683 | | | | | | | | | |
| 26 | 1.706 | 0.317 | 0.101 | 0.899 | 0.112 | 0.586 | 0.669 | | | | | | | | | |
| 27 | 1.703 | 0.311 | 0.097 | 0.903 | 0.107 | 0.587 | 0.655 | | | | | | | | | |
| 28 | 1.701 | 0.306 | 0.094 | 0.906 | 0.103 | 0.588 | 0.643 | | | | | | | | | |
| 29 | 1.699 | 0.301 | 0.091 | 0.909 | 0.100 | 0.589 | 0.631 | | | | | | | | | |
| 30 | 1.697 | 0.296 | 0.088 | 0.912 | 0.096 | 0.589 | 0.620 | | | | | | | | | d > 1/t |
| 40 | 1.684 | 0.257 | 0.066 | 0.934 | 0.071 | 0.594 | 0.533 | | | | | | | | | d < 1/t |
| 50 | 1.676 | 0.231 | 0.053 | 0.947 | 0.056 | 0.597 | 0.474 | | | | | | | | | |
| 60 | 1.671 | 0.211 | 0.044 | 0.956 | 0.047 | 0.598 | 0.431 | | | | | | | | | |
| 80 | 1.664 | 0.183 | 0.033 | 0.967 | 0.035 | 0.601 | 0.372 | | | | | | | | | |
| 100 | 1.66 | 0.164 | 0.027 | 0.973 | 0.028 | 0.602 | 0.332 | | | | | | | | | |
| 120 | 1.658 | 0.150 | 0.022 | 0.978 | 0.023 | 0.603 | 0.303 | | | | | | | | | |
| ∞ | 1.645 | | | | | | | | | | | | | | | |

| Table 2. Features of a correlation distribution at a 95% confidence level (α = 0.05) | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***df*** | ***t*** | ***r*** | ***S*** | ***N*** | ***S/N*** | ***1/t*** | ***d*** | ***1/t vs N*** | ***S vs N*** | ***S vs 1/t*** | ***r vs N*** | ***S/N vs r*** | ***S/N vs 1/t*** | ***r vs 1/t*** | ***d vs N*** | ***d vs 1/t*** |
| 1 | 12.710 | 0.997 | 0.994 | 0.006 | 161.544 | 0.079 | | | | | | | | | | |
| 2 | 4.303 | 0.950 | 0.903 | 0.097 | 9.258 | 0.232 | | | | | | | | | | |
| 3 | 3.182 | 0.878 | 0.771 | 0.229 | 3.375 | 0.314 | | | | | | | | | | |
| 4 | 2.776 | 0.811 | 0.658 | 0.342 | 1.927 | 0.360 | | 1/t > N | | | | | | | | |
| 5 | 2.571 | 0.755 | 0.569 | 0.431 | 1.322 | 0.389 | | 1/t < N | S > N | | | | | | | |
| 6 | 2.447 | 0.707 | 0.499 | 0.501 | 0.998 | 0.409 | 1.998 | | S ≈ N | | | | | | | |
| 7 | 2.365 | 0.666 | 0.444 | 0.556 | 0.799 | 0.423 | 1.788 | | S < N | S > 1/t | | | | | | |
| 8 | 2.306 | 0.632 | 0.399 | 0.601 | 0.665 | 0.434 | 1.631 | | | S < 1/t | r > N | S/N > r | | | | |
| 9 | 2.262 | 0.602 | 0.362 | 0.638 | 0.569 | 0.442 | 1.508 | | | | r < N | S/N < r | | | | |
| 10 | 2.228 | 0.576 | 0.332 | 0.668 | 0.496 | 0.449 | 1.409 | | | | | | S/N > 1/t | | | |
| 11 | 2.201 | 0.553 | 0.306 | 0.694 | 0.440 | 0.454 | 1.327 | | | | | | S/N < 1/t | | | |
| 12 | 2.179 | 0.532 | 0.283 | 0.717 | 0.396 | 0.459 | 1.258 | | | | | | | | | |
| 13 | 2.160 | 0.514 | 0.264 | 0.736 | 0.359 | 0.463 | 1.198 | | | | | | | | | |
| 14 | 2.145 | 0.497 | 0.247 | 0.753 | 0.329 | 0.466 | 1.147 | | | | | | | | | |
| 15 | 2.131 | 0.482 | 0.232 | 0.768 | 0.303 | 0.469 | 1.100 | | | | | | | r > 1/t | | |
| 16 | 2.120 | 0.468 | 0.219 | 0.781 | 0.281 | 0.472 | 1.060 | | | | | | | r ≈ 1/t | | |
| 17 | 2.110 | 0.456 | 0.208 | 0.792 | 0.262 | 0.474 | 1.024 | | | | | | | r < 1/t | | |
| 18 | 2.101 | 0.444 | 0.197 | 0.803 | 0.245 | 0.476 | 0.990 | | | | | | | | | |
| 19 | 2.093 | 0.433 | 0.187 | 0.813 | 0.231 | 0.478 | 0.960 | | | | | | | | | |
| 20 | 2.086 | 0.423 | 0.179 | 0.821 | 0.218 | 0.479 | 0.933 | | | | | | | | | |
| 21 | 2.080 | 0.413 | 0.171 | 0.829 | 0.206 | 0.481 | 0.908 | | | | | | | | | |
| 22 | 2.074 | 0.404 | 0.164 | 0.836 | 0.196 | 0.482 | 0.884 | | | | | | | | | |
| 23 | 2.069 | 0.396 | 0.157 | 0.843 | 0.186 | 0.483 | 0.863 | | | | | | | | d > N | |
| 24 | 2.064 | 0.388 | 0.151 | 0.849 | 0.178 | 0.484 | 0.843 | | | | | | | | d < N | |
| 25 | 2.060 | 0.381 | 0.145 | 0.855 | 0.170 | 0.485 | 0.824 | | | | | | | | | |
| 26 | 2.056 | 0.374 | 0.140 | 0.860 | 0.163 | 0.486 | 0.806 | | | | | | | | | |
| 27 | 2.052 | 0.367 | 0.135 | 0.865 | 0.156 | 0.487 | 0.790 | | | | | | | | | |
| 28 | 2.048 | 0.361 | 0.130 | 0.870 | 0.150 | 0.488 | 0.774 | | | | | | | | | |
| 29 | 2.045 | 0.355 | 0.126 | 0.874 | 0.144 | 0.489 | 0.759 | | | | | | | | | |
| 30 | 2.042 | 0.349 | 0.122 | 0.878 | 0.139 | 0.490 | 0.746 | | | | | | | | | |
| 40 | 2.021 | 0.304 | 0.093 | 0.907 | 0.102 | 0.495 | 0.639 | | | | | | | | | |
| 50 | 2.009 | 0.273 | 0.075 | 0.925 | 0.081 | 0.498 | 0.568 | | | | | | | | | |
| 60 | 2.000 | 0.250 | 0.063 | 0.938 | 0.067 | 0.500 | 0.516 | | | | | | | | | d > 1/t |
| 80 | 1.990 | 0.217 | 0.047 | 0.953 | 0.050 | 0.503 | 0.445 | | | | | | | | | d < 1/t |
| 100 | 1.984 | 0.195 | 0.038 | 0.962 | 0.039 | 0.504 | 0.397 | | | | | | | | | |
| 120 | 1.980 | 0.178 | 0.032 | 0.968 | 0.033 | 0.505 | 0.361 | | | | | | | | | |
| ∞ | 1.960 | | | | | | | | | | | | | | | |

## Table 3. Features of a correlation distribution at a 98% confidence level ($\alpha$ = 0.02)

| *df* | *t* | *r* | *S* | *N* | *S/N* | *1/t* | *d* | *1/t vs N* | *S vs N* | *S vs 1/t* | *r vs N* | *S/N vs r* | *S/N vs 1/t* | *r vs 1/t* | *d vs N* | *d vs 1/t* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 31.820 | 1.000 | 0.999 | 0.001 | 1012.512 | 0.031 | | | | | | | | | | |
| 2 | 6.965 | 0.980 | 0.960 | 0.040 | 24.256 | 0.144 | | | | | | | | | | |
| 3 | 4.541 | 0.934 | 0.873 | 0.127 | 6.874 | 0.220 | | | | | | | | | | |
| 4 | 3.747 | 0.882 | 0.778 | 0.222 | 3.510 | 0.267 | | $1/t > N$ | | | | | | | | |
| 5 | 3.365 | 0.833 | 0.694 | 0.306 | 2.265 | 0.297 | | $1/t < N$ | | | | | | | | |
| 6 | 3.143 | 0.789 | 0.622 | 0.378 | 1.646 | 0.318 | | | | | | | | | | |
| 7 | 2.998 | 0.750 | 0.562 | 0.438 | 1.284 | 0.334 | | | | | | | | | | |
| 8 | 2.896 | 0.715 | 0.512 | 0.488 | 1.048 | 0.345 | | | $S > N$ | | | | | | | |
| 9 | 2.821 | 0.685 | 0.469 | 0.531 | 0.884 | 0.354 | 1.881 | | $S < N$ | | | | | | | |
| 10 | 2.764 | 0.658 | 0.433 | 0.567 | 0.764 | 0.362 | 1.748 | | | | | | | | | |
| 11 | 2.718 | 0.634 | 0.402 | 0.598 | 0.672 | 0.368 | 1.639 | | | $S > 1/t$ | $r > N$ | $S/N > r$ | | | | |
| 12 | 2.681 | 0.612 | 0.375 | 0.625 | 0.599 | 0.373 | 1.548 | | | $S \approx 1/t$ | $r < N$ | $S/N < r$ | | | | |
| 13 | 2.650 | 0.592 | 0.351 | 0.649 | 0.540 | 0.377 | 1.470 | | | $S < 1/t$ | | | | | | |
| 14 | 2.624 | 0.574 | 0.330 | 0.670 | 0.492 | 0.381 | 1.403 | | | | | | | | | |
| 15 | 2.602 | 0.558 | 0.311 | 0.689 | 0.451 | 0.384 | 1.344 | | | | | | | | | |
| 16 | 2.583 | 0.542 | 0.294 | 0.706 | 0.417 | 0.387 | 1.292 | | | | | | $S/N > 1/t$ | | | |
| 17 | 2.567 | 0.529 | 0.279 | 0.721 | 0.388 | 0.390 | 1.245 | | | | | | $S/N \approx 1/t$ | | | |
| 18 | 2.552 | 0.515 | 0.266 | 0.734 | 0.362 | 0.392 | 1.203 | | | | | | $S/N < 1/t$ | | | |
| 19 | 2.539 | 0.503 | 0.253 | 0.747 | 0.339 | 0.394 | 1.165 | | | | | | | | | |
| 20 | 2.528 | 0.492 | 0.242 | 0.758 | 0.320 | 0.396 | 1.131 | | | | | | | | | |
| 21 | 2.518 | 0.482 | 0.232 | 0.768 | 0.302 | 0.397 | 1.099 | | | | | | | | | |
| 22 | 2.508 | 0.472 | 0.222 | 0.778 | 0.286 | 0.399 | 1.069 | | | | | | | | | |
| 23 | 2.500 | 0.462 | 0.214 | 0.786 | 0.272 | 0.400 | 1.043 | | | | | | | | | |
| 24 | 2.492 | 0.453 | 0.206 | 0.794 | 0.259 | 0.401 | 1.017 | | | | | | | | | |
| 25 | 2.485 | 0.445 | 0.198 | 0.802 | 0.247 | 0.402 | 0.994 | | | | | | | | | |
| 26 | 2.479 | 0.437 | 0.191 | 0.809 | 0.236 | 0.403 | 0.972 | | | | | | | | | |
| 27 | 2.473 | 0.430 | 0.185 | 0.815 | 0.227 | 0.404 | 0.952 | | | | | | | | | |
| 28 | 2.467 | 0.423 | 0.179 | 0.821 | 0.217 | 0.405 | 0.932 | | | | | | | | | |
| 29 | 2.462 | 0.416 | 0.173 | 0.827 | 0.209 | 0.406 | 0.914 | | | | | | | $r > 1/t$ | | |
| 30 | 2.457 | 0.409 | 0.168 | 0.832 | 0.201 | 0.407 | 0.897 | | | | | | | $r \approx 1/t$ | $d > N$ | |
| 40 | 2.423 | 0.358 | 0.128 | 0.872 | 0.147 | 0.413 | 0.766 | | | | | | | $r > 1/t$ | $d < N$ | |
| 50 | 2.403 | 0.322 | 0.104 | 0.896 | 0.115 | 0.416 | 0.680 | | | | | | | | | |
| 60 | 2.390 | 0.295 | 0.087 | 0.913 | 0.095 | 0.418 | 0.617 | | | | | | | | | |
| 80 | 2.374 | 0.257 | 0.066 | 0.934 | 0.070 | 0.421 | 0.531 | | | | | | | | | |
| 100 | 2.364 | 0.230 | 0.053 | 0.947 | 0.056 | 0.423 | 0.473 | | | | | | | | | |
| 120 | 2.358 | 0.210 | 0.044 | 0.956 | 0.046 | 0.424 | 0.431 | | | | | | | | | $d > 1/t$ |
| $\infty$ | 2.326 | | | | | | | | | | | | | | | |

| Table 4. Features of a correlation distribution at a 99% confidence level (α = 0.01) | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **df** | **t** | **r** | **S** | **N** | **S/N** | **1/t** | **d** | **1/t vs N** | **S vs N** | **S vs 1/t** | **r vs N** | **S/N vs r** | **S/N vs 1/t** | **r vs 1/t** | **d vs N** | **d vs 1/t** |
| 1 | 63.66 | 1.000 | 1.000 | 0.000 | 4052.596 | 0.016 | | | | | | | | | | |
| 2 | 9.925 | 0.990 | 0.980 | 0.020 | 49.253 | 0.101 | | | | | | | | | | |
| 3 | 5.841 | 0.959 | 0.919 | 0.081 | 11.372 | 0.171 | | | | | | | | | | |
| 4 | 4.604 | 0.917 | 0.841 | 0.159 | 5.299 | 0.217 | | | | | | | | | | |
| 5 | 4.032 | 0.875 | 0.765 | 0.235 | 3.251 | 0.248 | | 1/t > N | | | | | | | | |
| 6 | 3.707 | 0.834 | 0.696 | 0.304 | 2.290 | 0.270 | | 1/t < N | | | | | | | | |
| 7 | 3.499 | 0.798 | 0.636 | 0.364 | 1.749 | 0.286 | | | | | | | | | | |
| 8 | 3.355 | 0.765 | 0.585 | 0.415 | 1.407 | 0.298 | | | | | | | | | | |
| 9 | 3.25 | 0.735 | 0.540 | 0.460 | 1.174 | 0.308 | | | S > N | | | | | | | |
| 10 | 3.169 | 0.708 | 0.501 | 0.499 | 1.004 | 0.316 | 2.004 | | S ≈ N | | | | | | | |
| 11 | 3.106 | 0.684 | 0.467 | 0.533 | 0.877 | 0.322 | 1.873 | | S < N | | | | | | | |
| 12 | 3.055 | 0.661 | 0.437 | 0.563 | 0.778 | 0.327 | 1.764 | | | | | | | | | |
| 13 | 3.012 | 0.641 | 0.411 | 0.589 | 0.698 | 0.332 | 1.671 | | | | | | | | | |
| 14 | 2.977 | 0.623 | 0.388 | 0.612 | 0.633 | 0.336 | 1.591 | | | | r > N | S/N > r | | | | |
| 15 | 2.947 | 0.606 | 0.367 | 0.633 | 0.579 | 0.339 | 1.522 | | | | r < N | S/N < r | | | | |
| 16 | 2.921 | 0.590 | 0.348 | 0.652 | 0.533 | 0.342 | 1.461 | | | S > 1/t | | | | | | |
| 17 | 2.898 | 0.575 | 0.331 | 0.669 | 0.494 | 0.345 | 1.406 | | | S < 1/t | | | | | | |
| 18 | 2.878 | 0.561 | 0.315 | 0.685 | 0.460 | 0.347 | 1.357 | | | | | | | | | |
| 19 | 2.861 | 0.549 | 0.301 | 0.699 | 0.431 | 0.350 | 1.313 | | | | | | | | | |
| 20 | 2.845 | 0.537 | 0.288 | 0.712 | 0.405 | 0.351 | 1.272 | | | | | | | | | |
| 21 | 2.831 | 0.526 | 0.276 | 0.724 | 0.382 | 0.353 | 1.236 | | | | | | | | | |
| 22 | 2.819 | 0.515 | 0.265 | 0.735 | 0.361 | 0.355 | 1.202 | | | | | | S/N > 1/t | | | |
| 23 | 2.807 | 0.505 | 0.255 | 0.745 | 0.343 | 0.356 | 1.171 | | | | | | S/N < 1/t | | | |
| 24 | 2.797 | 0.496 | 0.246 | 0.754 | 0.326 | 0.358 | 1.142 | | | | | | | | | |
| 25 | 2.787 | 0.487 | 0.237 | 0.763 | 0.311 | 0.359 | 1.115 | | | | | | | | | |
| 26 | 2.779 | 0.479 | 0.229 | 0.771 | 0.297 | 0.360 | 1.090 | | | | | | | | | |
| 27 | 2.771 | 0.471 | 0.221 | 0.779 | 0.284 | 0.361 | 1.067 | | | | | | | | | |
| 28 | 2.763 | 0.463 | 0.214 | 0.786 | 0.273 | 0.362 | 1.044 | | | | | | | | | |
| 29 | 2.756 | 0.456 | 0.208 | 0.792 | 0.262 | 0.363 | 1.024 | | | | | | | | | |
| 30 | 2.75 | 0.449 | 0.201 | 0.799 | 0.252 | 0.364 | 1.004 | | | | | | | | | |
| 40 | 2.704 | 0.393 | 0.155 | 0.845 | 0.183 | 0.370 | 0.855 | | | | | | | r > 1/t | d > N | |
| 50 | 2.678 | 0.354 | 0.125 | 0.875 | 0.143 | 0.373 | 0.757 | | | | | | | r < 1/t | d < N | |
| 60 | 2.66 | 0.325 | 0.105 | 0.895 | 0.118 | 0.376 | 0.687 | | | | | | | | | |
| 80 | 2.639 | 0.283 | 0.080 | 0.920 | 0.087 | 0.379 | 0.590 | | | | | | | | | |
| 100 | 2.626 | 0.254 | 0.065 | 0.935 | 0.069 | 0.381 | 0.525 | | | | | | | | | |
| 120 | 2.617 | 0.232 | 0.054 | 0.946 | 0.057 | 0.382 | 0.478 | | | | | | | | | d > 1/t |
| ∞ | 2.576 | | | | | | | | | | | | | | | |

## Feature regimes

The tables have been color coded to demonstrate the presence of several feature regimes. The following convention was used:

- If $F_1 > F_2$ a cell is colored in blue.
- If $F_1 < F_2$ a cell is colored in pink.
- If $F_1$ - $F_2$ is less than $\pm 0.005$, a cell is colored in yellow.

So, borderline values were resolved up to that resolution. This procedure can be repeated for other confidence levels. We can keep expanding the tables by computing additional features from power analysis theory. See Figure 1.

## Results

Several "row" and "column" correlation regimes have been identified, color-coded, and labeled at their interfaces. For instance from Table 2 (i.e., at a 95% confidence level, $\alpha = 0.05$) two opposite and mathematically feasible "row" regimes have been determined:

- when $r \geq 0.811$, $1/t > N$, $S > N$, $S > 1/t$, $r > N$, $S/N > r$, $S/N > 1/t$, $r > 1/t$, $d > N$, $d > 1/t$
- when $r \leq 0.217$, $1/t < N$, $S < N$, $S < 1/t$, $r < N$, $S/N < r$, $S/N < 1/t$, $r < 1/t$, $d < N$, $d < 1/t$

Note that we refer to these as mathematically feasible regimes since $d$ values greater than 2 are computable from $d = 2\frac{r}{\sqrt{1-r^2}} = 2\sqrt{S/N} = 2\frac{t}{\sqrt{df}}$, but have no statistical meaning. If we ignore such values then the effective regimes are

- when $r \geq 0.811$, $1/t > N$, $S > N$, $S > 1/t$, $r > N$, $S/N > r$, $S/N > 1/t$, $r > 1/t$
- when $r \leq 0.456$, $1/t < N$, $S < N$, $S < 1/t$, $r < N$, $S/N < r$, $S/N < 1/t$, $r < 1/t$

"Column" regimes have also been identified from our tables.

For instance from the ***S vs N*** column of Table 2:

- when $r = 0.755$, $S > N$
- when $r = 0.666$, $S < N$

In general the exact threshold of $r$ values occurs when $r$ values are above or below $0.707$. See next section, Exercise 1.

Mixed regimes can be identified from other cells. The tables can also be used to discriminate between $r$ values.

For instance, any two coefficients $r_1$ and $r_2$, are statistically significant if $r_{observed} > r_{table}$. We can arrive at the same conclusion by considering signal-to-noise data since for statistical significance $S/N_{observed} > (t_{table}^2)/(df)$.

For similar samples and number of observations, an $r$ value where $S > N$ is generally considered more useful than one where $N > S$. When $N > S$ for any two values, $r_1$ and $r_2$, these can be discriminated by visually inspecting the tables or by computing individual features. Finally, the tables can be used to discriminate between identical feature values at different confidence levels.

## Using the tables

As mentioned earlier in this article, the tables herein described are meant to be used as complementary shortcuts for statistical and practical significance work.

To illustrate, let use Table 2 for the following examples:

1. Classify $r$ values in terms of the relative amount of signal and noise present in a set of paired variables.

   From the $S$ *vs* $N$ column we can see that when $S \approx N$, $r \approx 1/\sqrt{2} = 0.707$.

   Therefore, values above and below $0.707$ define specific ranges.

2. From 20 pairs of measurements you find a correlation coefficient of $0.495$. Can you conclude that a significant correlation exists in the population, at the $0.05$ level of significance?

   For statistical significance, $r_{observed} > r_{table}$ .

   Table 2 shows for $df = 18$ that $r_{table} = 0.444$. Since $r_{observed} = 0.495 > 0.444$, we conclude that a significant correlation does exist. Note that this is a lot easier than computing and then comparing a $t_{observed}$ against a $t_{table}$ .

3. At which experimental conditions $d = 1/t$? What does this mean?

   From $d = 2\frac{t}{\sqrt{df}}$ when $d = 1/t$ it is clear that $df = \frac{4}{d^4}$.

   From the several tables we have developed, when $d$ decreases $1/t$ increases and eventually reaches a plateau. From these tables, clearly $d = 1/t$ can only occur at the 95% confidence level ($\alpha = 0.05$) and when $d = 0.5$.

   A medium effect of $d = 0.5$ is visible to the naked eye of a careful observer. So, if for a medium effect $d = 1/t = 0.5$, $df = 64$. This is a unique situation for a medium effect size.

   Note. A small effect of $d = 0.2$ is noticeably smaller than medium but not so small as to be trivial. A large effect of $d = 0.8$ is the same distance above the medium as small is below it.

   From Figure 1 $d$ can be thought of as the average percentile standing of the average treated (experimental) relative to the average untreated (control) participant. Cohen's $d$ can also be interpreted in terms of the percent of non overlap of the treated group's scores with those of the untreated group.

| Remarks | d | Percentile Standing | Percent of Non overlap |
|---|---|---|---|
| | 2.0 | 97.7 | 81.1% |
| | 1.9 | 97.1 | 79.4% |
| | 1.8 | 96.4 | 77.4% |
| | 1.7 | 95.5 | 75.4% |
| | 1.6 | 94.5 | 73.1% |
| | 1.5 | 93.3 | 70.7% |
| | 1.4 | 91.9 | 68.1% |
| | 1.3 | 90 | 65.3% |
| | 1.2 | 88 | 62.2% |
| | 1.1 | 86 | 58.9% |
| | 1.0 | 84 | 55.4% |
| | 0.9 | 82 | 51.6% |
| Large | 0.8 | 79 | 47.4% |
| | 0.7 | 76 | 43.0% |
| | 0.6 | 73 | 38.2% |
| Medium | 0.5 | 69 | 33.0% |
| | 0.4 | 66 | 27.4% |
| | 0.3 | 62 | 21.3% |
| Small | 0.2 | 58 | 14.7% |
| | 0.1 | 54 | 7.7% |
| | 0.0 | 50 | 0% |

Figure 1. See Becker, Lee. *Effect* Sizes. http://www.uccs.edu/~faculty/lbecker/es.htm

4. What is the correlation associated to a $0.5$ difference between two independent sample means of same number of observations at $\alpha = 0.05$?

   From Table 2, $d$ is close to $0.517$, so $r$ must be less than $0.250$. A refined result is obtained by expressing $r$ in terms of $d$; i.e.

   $r^2 = \frac{d^2}{d^2 + 4}$. Therefore, $r = \frac{d}{\sqrt{d^2+4}} = \frac{0.5}{\sqrt{0.5^2+4}} \approx 0.243$

5. In the previous problem, how many observations in each group are required for a power of $p = 0.80$?

   Let $\beta$ be the probability of a Type II Error. Power is defined as the $1 - \beta$ probability area of a population normal distribution. There is a cut-off $t$ value associated to this area and its critical $t$ value that depends on $df$. These cut-off values can be computed with the tool available at http://www.statdistributions.com/t/?p=0.1&df=12&tail=2 .

| $df$ | $60$ | $80$ | $100$ | $120$ | $\infty$ |
|---|---|---|---|---|---|
| $t_{critical}$ | $2.000$ | $1.990$ | $1.984$ | $1.980$ | $1.960$ |
| $t_{cut\text{-}off}$ | $-0.848$ | $-0.846$ | $-0.845$ | $-0.845$ | $-0.842$ |

Figure 2. Values associated to a power of 0.80.

   In Figure 2 we have computed values by using the right-tail option of the tool and setting power $p$ to $0.80$. We used these values since from Table 2 we can see that for a single sample $df > 60$ and $t < 2$.

As $df$ increases $t_{cut\text{-}off}$ is about $-0.84$, so we can safely use this cut-off value. In general for a two-tailed *t*-tests, the difference between two independent means, with $\alpha = 0.05$ is given by

$$n = \frac{2(2-k)^2}{d^2} \;,$$

where $k$ is a $t$ value that cuts off the upper portion of the *t* distribution corresponding to the desired level of power.

So, $n = 64.52$ with $df = 62.52$ and $t_{critical} = \frac{d\sqrt{df}}{2} = 1.977$. Note for both samples that a total size of $2n \approx 129$ and $df = 2n - 2 \approx 127$ are needed. These results are close to those obtained with software like *G*Power*. See Figure 3.

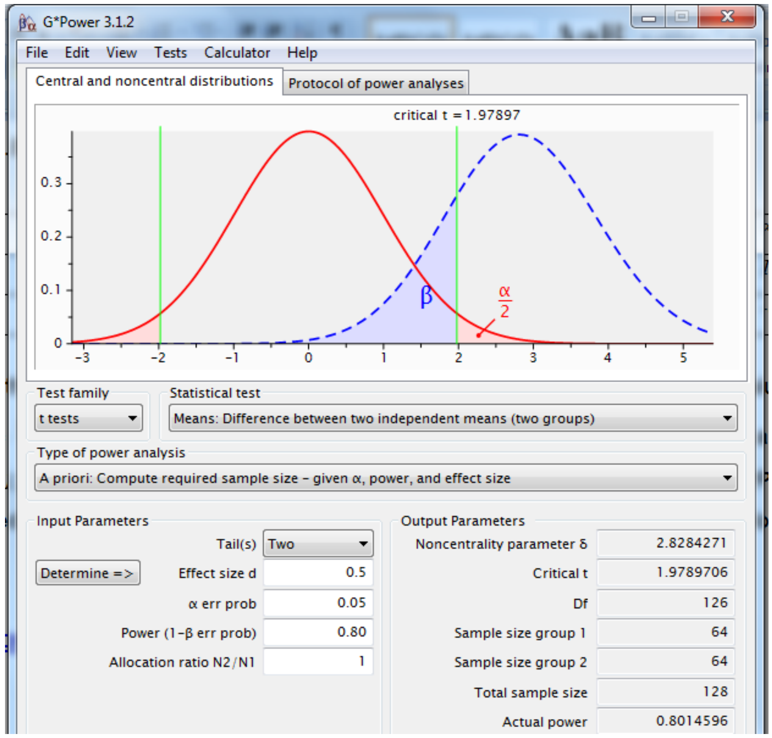


Figure 3. G*Power results.

*G*Power* gives $n = 64$ and $t_{critical} = 1.979$. For both samples a total of $128$ observations and $df = 126$ are needed. So for those with no access to *G*Power*, the above procedure can be used with comparable results.

## Appendix: List of Symbols and Features

| | | |
|---|---|---|
| $n$ | = | sample size defined as number of measurements |
| $t$ | = | statistical confidence value |
| $df$ | = | degrees of freedom defined as $n - 2$ |
| $r$ | = | correlation coefficient, computed as $\frac{t}{\sqrt{df + t^2}}$ |
| $S$ | = | $r^2$, Signal or fraction of explained variations in the dependent variable and due to variations in the independent variable |
| $N$ | = | $1 - r^2$, Noise or fraction of unexplained variations in the dependent variable |
| $S/N$ | = | Signal-to-Noise ratio |
| $1/t$ | = | Inverse confidence value |
| $d$ | = | Cohen's $d$ value, approximated as $2\frac{r}{\sqrt{1-r^2}} = 2\sqrt{S/N} = 2\frac{t}{\sqrt{df}}$ |

**Note.** Cohen's $d$ is listed up to $d \le 2$ (Becker, 2000; Cohen 1988). Procedures for transforming $d$ to $r$ and vice versa have been discussed by Cohen (1983, 1988, 1992), Friedman (1968), Glass, et. al. (1981), Rosenthal (1984), and Wolf (1986), using the following formula

$$r^2 = \frac{d^2}{d^2 + 4}$$

From the work of Hedges and Olkin (1985) and Aaron, Kromrey, and Ferron (1998), it is clear that

$$r^2 = \frac{d^2}{d^2 + (n_1 + n_2 - 2)(n_1 + n_2)/(n_1 n_2)}$$

For the special case wherein $n_1 = n_2 = n$

$$r^2 = \frac{d^2}{d^2 + 4\left(\frac{n-1}{n}\right)}$$

And the first expression is obtained for $n >> 1$. Finally solving for $d$,

$$d = 2\frac{r}{\sqrt{1-r^2}} = 2\sqrt{S/N} = 2\frac{t}{\sqrt{df}}$$

which approximates Cohen's $d$ as a signal-to-noise function.

## References


Aaron, B., Kromrey, J. D., and Ferron, J. (1998). *Equating r-based and d-based effect size indices: Problems with a commonly recommended formula*. Annual Meeting of the Florida Educational Research Association, Orlando, Florida, November 3, 1998.
http://www.eric.ed.gov/PDFS/ED433353.pdf

Becker, L. A (2000). *Effect size*. Revised 2000, 1999, and 1998.
http://www.uccs.edu/~faculty/lbecker/es.htm

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, J. (1992). A *Power Primer*. Psychological Bulletin, 1992, Vol. 112. No. 1,155-159. American Psychological Association, Inc.
http://www.uc.cl/letras/laboratoriodefonetica/html/materiales_talleres_internos/Cohen1992.pdf

Cohen, J. and Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, Nj: LEA.

Friedman, H. (1968). *Magnitude of experimental effect and a table for its rapid estimation*. Psychological Bulletin. 70:245 - 251.

Glass, G. V., McGraw, B., and Smith, M. L. (1981). *Meta-analysis in social research*. Sage, Beverly Hills.

Hedges, L. V. and Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press: Orlando.

Rosenthal, R. (1984). *Parametric measures of effect size*, In H. Cooper & L. V. Hedges, (eds.), The Handbook of Research Synthesis (pp 232 - 244).

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Sage University Paper series on quantitative applications in the behavioral sciences, 07-059. Sage, Beverly Hills.