

Re: Generating a Matrix of Random values with a specified Correlation

Source: <http://sci.tech-archive.net/Archive/sci.math/2008-09/msg03091.html>

- *From:* Robert Israel <israel@xx>
 - *Date:* Thu, 25 Sep 2008 12:17:05 -0500
-

willgarrison@xxxxxxxx writes:

I'm trying to generate two columns of random values where each value is between 1 and 5 where the matrix as a whole has a specified correlation.

Ideally this is what I'm looking for:

Example algorithm:

1. Generate 20 random values between 1 and 5 in array A.
2. Generate 20 random values between 1 and 5 in array B.
3. Create a 3rd array (array C), where array B is sorted so that array A and array C have a correlation of X.

Does anyone have any ideas on how to do this?

This may be a horrible way to solve this problem. If so, does anyone have any ideas on how else I could go about it?

Start with your two arrays A and B. Let $a = \text{Var}(A)$, $b = \text{Var}(B)$, $c = \text{Cov}(A, B)$, so the (Pearson) correlation of A and B is $\rho = c / \sqrt{a b}$. Let's suppose the desired correlation $x > \rho$ (where of course $x \leq 1$). Consider $C = tA + (1-t)B$, where $0 \leq t \leq 1$. We have $\text{Cov}(A, C) = ta + (1-t)c$ and $\text{Var}(C) = t^2 a + 2t(1-t)c + (1-t)^2 b$ so the correlation of A and C is $(ta + (1-t)c) / \sqrt{a(t^2 a + 2t(1-t)c + (1-t)^2 b)}$. Set this equal to x and solve for t : there should be exactly one solution in the interval $0 \leq t \leq 1$.

On the other hand, if $-1 \leq x < \rho$, try $C = tR + (1-t)B$ where each $R_j = 6 - A_j$.

--
Robert Israel israel@xx
Department of Mathematics <http://www.math.ubc.ca/~israel>
University of British Columbia Vancouver, BC, Canada

Re: Generating a Matrix of Random values with a specified Correlation