

Re: Test of Independence

Source: <http://sci.tech-archive.net/Archive/sci.stat.math/2005-09/msg00122.html>

- *From:* Richard Ulrich <Rich.Ulrich@xxxxxxxxxxxx>
 - *Date:* Thu, 08 Sep 2005 17:45:56 -0400
-

On 8 Sep 2005 07:20:15 -0700, montey_man11@xxxxxxxxxxxx wrote:

> Hi,
> I am dealing with a population with two variables. These two variables
> should represent two random digit sets (uniform distribution). What I
> need to do is to ensure that these two random digit sets are random and
> independent. I am looking for randomness and that they are independence
> in order to conduct two simultaneous tests on this same population
> without one test biasing the other. Is this possible in SAS?

You want to "conduct two ... tests ... without one test biasing the other" — on two series of putatively random numbers, which each consist of integers from 0–99 (information posted later).

It certainly does matter, "random with regards to WHAT", but "two tests" does not give much to go on.

Is "pairwise" the only dependency that matters, or
– Does it matter if one is a lagged version of the other?
– Does it matter if one is a re-arranged version of the other?

>
> Since this is an inherited dataset, I am not sure how the two digits
> were generated. I want to 1) make sure that they follow a uniform
> distribution, 2) that it is random, and 3) that the two variables are
> Independent.
>
> for #1, all I needed to do is to run proc freq on the variables and
> validate that it is distributed following a uniform distribution.

Does it matter if either one is TOO perfectly uniform?

>
> for #2 I am stuck! I am not sure how I can determine the distribution
> of values is random (rather than, say, ordered by some value)
>
> for #3 I am also stuck. I am not sure how to test the independence of
> the two groups. What I am thinking of doing is to get the value of the
> correlation coefficient. However, from my LIMITED knowledge of

Re: Test of Independence

> Statistics, I know that a correlation of zero does not mean
> independence. Also, the correlation only tests linear relationships.

Typically, to test Randomizers, you crosstab lags -- by value;
by grouped values. With two randomizers, I suppose you
lag one against the other, too. Does this matter, for your
version of 'dependency'?

Do you have hundreds of values, thousands, or millions?

--

Rich Ulrich, wpilib@xxxxxxxxx
<http://www.pitt.edu/~wpilib/index.html>

• **References:**

◆ **Test of Independence**

◇ From: montey_man11

• Prev by Date: **Re: KS-test**

• Next by Date: **Re: ANOVA question**

• Previous by thread: **Re: Test of Independence**

• Next by thread: **Result for linear transformation of covariance matrix**

• Index(es):

◆ **Date**

◆ **Thread**