Re: Randomness

Source: http://sci.tech−archive.net/Archive/sci.math/2006−09/msg00589.html

- *From*: "Matt Zellman" <matt.zellman@xxxxxxxxx>
- *Date*: 3 Sep 2006 14:04:49 −0700

mensanator@xxxxxxxxxxx wrote:

> Matt Zellman wrote:
>
> > mensanator@xxxxxxxxxxx wrote:
> >
> > > Matt Zellman wrote:
> > >
> > > > Matthias Klaey wrote:
> > > >
> > > > > "Matt Zellman"
> > > > > <matt.zellman@xxxxxxxxx>
> > > > > wrote:
> > > > >
> > > > > > mensanator@xxxxxxxxxxx
> > > > > > wrote:
> > > > > >
> > > > > > > Matt
> > > > > > > Zellman
> > > > > > > wrote:
> > > > > > >
> > > > > > > > are
> > > > > > > > the
> > > > > > > > following
> > > > > > > > string
> > > > > > > > of
> > > > > > > > digits
> > > > > > > > random?
> > > > > > >
> > > > > > > Not
> > > > > > > likely.
> > > > > >
> > > > > > I suppose

that was kind of a stupid question. Are they random by some definition of random? and if so, what definition?

0101000000000000000000000000000010000100100

0000110110111110001010100010011100110011011

1101110011101111010000011010111000001111101

0000000000000000000000000100000000010000000

1010110100111010011010100010011001100010011

1010100111100001010110111101000100001000010

here's a better question. Could the preceding six strings of digits conceivably be randomly generated?

Sure.

This question somehow reminds me of a theorem that I heard of, but never actually seen: "For

> any given set of data there is
> a test such
> that the data will pass the
> test" (or converse, "such that
> the data
> will fail the test/such that
> the null hypothesis is
> rejected")
> I believe it is attributed to
> Kolmogorov.
> Does anyone know
> something more precise
> about this?
>
> And Matt, by this theorem
> the answer to your firts
> question is "yes".
> For the second: I have no
> idea. The best reference I
> can give you is
> D.E.Knuth, The Art of
> Computer Programming,
> Volume 2, Chapter 3, Third
> Edition, Addison−Wesley
> 1998, ISBN
> 0−201−89684−2. There you
> may learn
> more about random
> sequences than you ever
> wanted to know :)
>
> Greetings, Matthias Kläy

> Thanks. The way the strings were generated
> is as follows:
>
> I started with the first however many digits
> of pi (a number whose
> digits are provably normal

They are?

my mistake. I mistook the well−evidenced conjecture for a proof. In any case, the first 6 billion digits or so are normal, so it shouldn't make too big a difference for our sample here.

>), and applied a series of tests to it:

Oh, a series of tests.

I feel like the guy with the spitoon, I thought it was all one string.

> the first sequence gives a 1 for every digit
> that is a 0 or 1

So eight out of every 10 digits will give 0. How is that random?
Any tests that doesn't produce the same number of 1's and 0's certainly isn't going to be random. Even if it does, it will still depend on how they are distributed.

Your coin is definitely biased.

How is it not random? Just because you get more of one outcome doesn't make it not random, it just makes it not fair.

A weighted die isn't nonrandom, just unfair.

> the second sequence gives a 1 for every digit
> greater than or equal to
> 5

So half will be 1's and half will be 0's. It still won't be random,
but it'll be harder to show that.

I think this one will actually be random (and fair).

> the third sequence gives a 1 for every digit
> that is odd

Again, same number of 1's as 0's.

> the fourth sequence gives a 1 for every digit
> that is the same as the
> previous digit

So you'll get too many 0's. Not random.

> the fifth sequence gives a 1 for every digit
> that is greater than the
> previous digit

Tie goes to 0, so you'll have too many 0's. Also note that under
this rule, you can't have more than 9 consecutive 1's. This will
be an obvious giveaway of non−randomness when you make the sequence big enough since, in a random distribution, all sequence lengths occur eventually.

the max run of 9 does make for an interesting restriction...

> the sixth sequence gives a 1 for every digit
> that is a 3,4,5, or 6

Again, simply too many 0's. Obviously non−random.

> The resulting sequences are not (necessarily)
> normal, but I think they
> can still be described as "random" as long as
> there is some nonzero
> chance that a digit could be either a zero or a
> one.

But you have to say up front what the probabilities are if they aren't 1/2. So, no, the sequences with differing 1 and 0 counts can't be described as random. And the fifth sequence can't be described as random even if you give the probabilities.

I don't know... I remember all my statistics textbooks explicitly designating a coin or die as fair in order to establish the normality of the distribution.

> But you didn't make such an explicit designation until after you explained what the source was. If you had said "given this random sequence, determine the weighting" or "given this weighting, determine whether this sequence is random" you have a case. But you didn't do that. You asked for two unknowns, the weighting and the randomness. An assumption has to be made and I chose the obvious one of fair weighting.

Fair enough.

> > Why would that be necessary if the concept of randomness necessarily included normality?
>
> When you ask silly questions, don't complain about the quality of the answers.

sorry, not trying to be a dork, just a little overzealous about something that seemed obvious to me, but was apparently not actually the case.

.