Subject: Re: BEGINNER ASKS FOR HELP!!!
Posted by steinhh on Mon, 02 Feb 1998 08:00:00 GMT
View Forum Message <> Reply to Message

In article <6at4ri$21e@lace.colorado.edu> "Edoardo \"Dado\" Marcora"
<marcora@colorado.edu> writes:

> For example, the input by the user would be:
>
> NUMBER OF ELEMENTS PER ROW e.g., 6
>
> RATIO OF 1 / (-1) e.g., 0.5 (at the level of the matrix not of the row)
>
> NUMBER OF 1 AND (-1) (total) IN EACH ROW e.g., 3 (for example 0 0 0 1 1 -1)
>
> I do not need the ratio and number of |1| per row to be exact, also an
> approximation it's good.

One way of solving this is to make a row with the specified
number of ones/zeroes/minus ones, e.g., [0,0,1,1,-1], and
then *ordering* those elements in a random way, which is
easily done by sorting a parallell, random array.

I.e.,

```
ELEMENTS = [0,0,0,1,1,-1]
N = N_ELEMENTS(ELEMENTS)
FOR ROW_NO=0,N_ROWS DO BEGIN
  DATA(*, ROW_NO) = ELEMENTS[SORT(RANDOMU(SEED,N))]
END
```

This will result in a fixed number of elements of each kind,
but randomly ordered.

If you'd like the number of ones/zeroes/minus ones to vary,
with probabilities according to the specifications,
another approach is better. First, we need to calculate
the probabilities of each "outcome".

Given N (number of elements per row), R (ratio of the number
of ones to minus ones), and NONZ (number of nonzero
elements), the probability of a single element being nonzero
will be:

```
P_NONZ = NONZ/FLOAT(N)  ; Avoid integer arithmetic
```

The probability of an element being zero will (of course) be:

P_ZERO = 1-P_NONZ

The probabilities of an element being either plus or minus
one of course add up to P_NONZ: P_PLUS + P_MINUS = P_NONZ.
We also have P_PLUS/P_MINUS = R, which gives:

P_PLUS = P_NONZ * R/(1.0+R)
P_MINUS = P_NONZ * 1.0/(1.0+R)

Now, if I generate a number that is uniformly distributed
between 0 (inclusive) and 1, the probability of that number
being between zero and P_ZERO is exactly P_ZERO. The
probability of the nuber being between P_ZERO and
P_ZERO+P_PLUS is exactly P_PLUS, and the probability of the
number being between P_ZERO+P_PLUS and 1 is P_MINUS (since
P_ZERO, P_PLUS and P_MINUS add up to one).

Thus,

```
ARRAY = RANDOMU(SEED,N,N_ROWS)
ZMASK = ROW LT P_ZERO
PMASK = ROW GE P_ZERO AND ROW LT P_ZERO+P_PLUS
MMASK = ROW GE P_ZERO+P_PLUS
ROW = 0*ZMASK + 1*PMASK + (-1)*MMASK
```

..should do it, I think. Of course that "0*ZMASK" part is not
necessary, it's only there for clarity.

Disclaimers about typos etc. apply, of course...

Stein Vidar
