This is a first attempt at App G's "sequence section" using the "left and right ensembles" idea. It did not fly, but I got the idea of doing something to obtain the desired cyclic property of the p function.

(f) Experiments with Sequences

Our pulse trains have a set of amplitudes $\{y_i\}$ which form a sequence. We will ignore the pulse train and just think about **sequences**. If the sequence has N elements, then it can be written $[y_1, y_2 ..... y_N]$, and each of these amplitudes can be a value in some set which we might for fun take to be {1,2,3,4,5,6}, though {1,0} is certainly more common.

Here is our first experiment. We have an Apparatus which generates sequences of length N according to some set of rules implemented within the Apparatus (perhaps these are the rules for generating AMI linecode sequences). Every sequence it generates is "legal" according to its rules. Earlier we discussed an ensemble experiment in which M dice were rolled one at a time and were left on the craps table for study and from what we saw on the table, we were able to construct a distribution function. The ensemble was M dice. In our current **ensemble experiment**, we let the Apparatus crank away and generate an ensemble of M sequences each of length N, and these M sequences are left sitting on the same craps table for us to inspect.

When many sequences are generated in this manner, one finds that the values of variable y_n in position n of the sequence can vary, and that there is some probability distribution $p(y_n)$ associated with the parameter y_n. Thus, the position n in the sequence is associated with a **random variable** we must call Y_n which takes values y_n. Notice that the notation y_n refers to the name of a variable (like x earlier). Having the subscript on y_n can lead to some confusion, so note that

$$E(X) = \Sigma_x \; x\, p(x) \;\; = \;\; \Sigma_i \; x_i \, p(x_i)$$

$$E(Y_n) = \Sigma_{y_n} \; y_n \, p(y_n) \;\; = \Sigma_i \; (y_n)_i \, p[(y_n)_i] \qquad \text{(G.36)}$$

We shall try to avoid the forms with indices like i.

If we let M be many billions, when we study the resulting ensemble and count sequences of each pattern, we will be able to construct the corresponding pmf distribution which we write as

$$p(y_1, y_2, ., , y_N) = \text{pmf}(Y_1 = y_1, Y_2 = y_2, ... Y_N = y_N) \quad . \qquad \text{(G.37)}$$

ok to here

In general there is no symmetry between arguments, so one *cannot* say $p(y_1, y_2, ., , y_N) = p(y_2, y_1, ., , y_N)$ for example. In general the dependence of $p(y_1, y_2, ., , y_N)$ on each variable in the variable list can be different. We shall return to this issue of symmetry of $p(y_1, y_2, ., , y_N)$ momentarily.

Certain probabilities are of special interest to us. Using (G.7) we may write

$$p(y_n) = \Sigma_{y_1, y_2 ..... \neq y_n} \, p(y_1, y_2 ..... y_N)$$

$$p(y_n,y_m) = \Sigma_{y_1,y_2.....\neq y_m y_n}\ p(y_1,y_2.....y_N) \qquad m \neq n \qquad \text{(G.38)}$$

and then

$$E(Y_n) = \Sigma_{y_n}\ (y_n)\ p(y_n)$$

$$E(Y_n^2) = \Sigma_{y_n}\ (y_n)^2\ p(y_n)$$

$$E(Y_n Y_m) = \Sigma_{y_n,y_m}\ y_n y_m\ p(y_n,y_m) \qquad \text{(G.39)}$$

Inserting (G.38) into (G.39) can write these same expected values out "in full" as :

$$E(Y_n) = \Sigma_{y_1,y_2....}\ (y_n)\ \ p(y_1,y_2.....y_N)$$

$$E(Y_n^2) = \Sigma_{y_1,y_2....}\ (y_n)^2\ p(y_1,y_2.....y_N)$$

$$E(Y_n Y_m) = \Sigma_{y_1,y_2....y_n y_m}\ p(y_1,y_2.....y_N) \qquad \text{(G.40)}$$

Looking again at

$$E(Y_n) = \Sigma_{y_n}\ y_n\ p(Y_n = y_n) = \Sigma_x\ x\ \ p(Y_n = x) \qquad \text{(G.41)}$$

we see that in general $E(Y_n)$ will be a function of n, and the same for $E(Y_n^2)$. Similarly, $E(Y_n Y_m)$ will be a function of both n and m.

Now we entertain a *new* ensemble experiment. We take our same sequence-generating Apparatus and we run it again to generate M sequences of length N, using the same seed list as the original run. However, this time for each sequence we throw out the first element generated and we let the Apparatus continue and create an $N+1^{th}$ element. On the first run, if we had a sequence {a,b,c,d,,,,,x} of length N, then on this second run we will get the sequence { b,c,d,,,,,x,y} of length N. Since there is a one-to-one correspondence between these two sequences of the two runs of the experiment, they must both have the same probability. That is to say.

$$p(a,b,c,d....x) = p(b,c,d.....x,y)$$

where for example Y_1 = a in the left p and Y_1 = b in the right p.

We could do another experiment again with the same random seed list and this time discard the first two elements and add 2 new ones at the end. Then we get

$$p(a,b,c,d....x) = p(b,c,d.....x,y) = p(c,d.....x,y,z)$$

There is not much we can do with this fact without making a new assumption about our Apparatus. We have said that the sequences of length N produced by the Apparatus have the distribution $p(y_1, y_2, .,, y_N)$. The assumption we now make is this:

Assumption: Imagine that N is large, perhaps N= 1000, and that we are interested in two random variables Y_n and Y_m which are both somewhat near the center of the sequence. The assumption is that as we move n and m farther apart from each other, there is less correlation between Y_n and Y_m. The decrease in correlation as the spacing increases need not be exactly monotonic, but it is ongoing and approaches 0 as the distance increases more and more. This assumption is that there is some effective correlation length parameter for the sequence. By the way, we know how to compute the correlation, and it is proportional to the covariance,

$$\text{cov}(Y_n Y_{n+k}) = E((Y_n - \mu_n)(Y_{n+k} - \mu_{n+k}))$$

$$= \Sigma_{y_1,y_2 \ldots y_N} (y_n-\mu_n)(y_{n+k}-\mu_{n+k}) \; p(y_1, y_2, \ldots y_n \ldots y_{n+k} \ldots y_N)$$

$$p(y_n,y_m) = \Sigma'_{y_1,y_2 \ldots y_N} p(y_1, y_2, \ldots y_n \ldots y_m \ldots y_N)$$

If N is very large, we might put the far right end of the sequence "out of mind" and write the above as

$$p(a,b,c,d....) = p(b,c,d.....) = p(c,d,e.....) = \text{etc}$$

Ignoring what happens at the far right end, we make this claim:

Fact: The probability distribution for our ensemble experiment has this property:

$$p(y_1,y_2,y_3.....) = p(y_2,y_3,y_4 .....) = \text{invariant under a shift of the parameters}$$

If we focus our interest out somewhere in the middle of the sequence, and ignore both the left and right ends, then we can shift in either direction and have invariance.

We could formalized this idea by using (G.7) and focus our attention just on a finite interior set of random variables X,Y,Z so that

$$p(x,y,z...) = \Sigma'_{a,b,c \ldots} \; p(a,b,c.....)$$

where we sum over all the left and right end parameters and don't sum on x,y,z.... We shall however dispense with this formality and say that Fact *** is valid for our "region of interest" in the interior of the sequence.

What then are the implications of Fact ** ? Consider again

$$E(Y_3) = \Sigma_{y_1,y_2 \ldots} (y_3) \; p(y_1,y_2,y_3.....)$$

$$E(Y_2) = \Sigma_{y_1,y_2 \ldots} (y_2) \; p(y_1,y_2,y_3.....)$$

In the second line, all the y_k are dummy summation indices, we can permute these names one position forward and not change the expression for $E(Y_2)$, so

$$E(Y_2) = \Sigma_{y_2,y_3....} (y_3) \quad p(y_2,y_3,y_4.....)$$

But then we apply our Fact *** to say

$$E(Y_2) = \Sigma_{y_2,y_3....} (y_3) \quad p(y_1,y_2,y_3.....)$$

We cannot strictly claim that the function is $p(y_1,y_2.....y_N)$ is cyclic in its arguments -- it is not. However, if we now think of N as a very large number, we can make this claim

$$p(y_1,y_2.....y_N) \approx p(y_N, y_1,y_2.....y_{N-1}) = \text{all other cyclic rotations}$$

In other words, we can treat $p(y_1,y_2.....y_N)$ as being cyclic if we are willing to ignore "end effects" which cause this fact to not really be true. We shall soon be interested in $N \to \infty$ and in that limit we can surely ignore the end effects. For the moment we keep N "very large".

Fact: For N very large, $p(y_1,y_2.....y_N)$ is cyclic in its arguments with some small "end effect" error.

However, this time on each run we have it generate an $N+1^{th}$ seqence element, an

and we use it as before to generate an ensemble of sequences, but this time we throw out the first generated symbol and we have the apparatus generate an extra symbol at the end so the new sequences still have N elements. For example

{a,b,c,,,,,x} old sequence, a is an actual symbol value like a = 1.

{b,c,,,,,x,y} new sequence

If in the original experiment we assume that a huge number of sequences is generated, and this number is large enough so that every possible sequence is generated (many times) which is "legal" according to the rules of the apparatus. After all, we need to have a huge number of sequences in order to get details of our distribution function $p(y_1, y_2, .,, y_N)$. Since the "new sequence" shown above is a "legal sequence", it

must be included somewhere in the original ensemble. What this means is that all sequences generated in the second experiment are the same as those generated in the first experiment.