Ideas re App G (f) on Sequences PhL 7.27.13

Review: This was a major stumbling block section. I am struggling Geoff Chew style to find some way to prove that for an Apparatus generated ensemble, $E(Y_n) = E(Y_{n+1})$. Such a simple idea, and so hard to show it is true.

Plan A. While walking near LDS hospital I got this idea that you could characterize **partial correlation** between two pulse train positions. I defined a smallness function ε(k) where k is spacing between the positions. I show that partial correlation means covariance goes to 0 as it should if correlation goes to 0 in my sense. I still think this idea is reasonable, but I never made use of it. My hope at the time was that I might assume a pulse train has some "correlation distance" and beyond that point Y_n and Y_m would be uncorrelated random variables. Maybe I was trying to use this idea to get $\langle y_n \rangle$ not depending on n, etc. I note now that such an idea might somehow ameliorate the seam of you wrapped a sequence around, but that is just a vague idea.

Plan B. My "**random seed** index s idea". The context here is letting the Apparatus go "one extra clock" so you then have two experiments each with its own ensemble of sequences, the left and the right. I think I wanted to argue that both these left and right ensembles had some kind of 1 to 1 relation, reordering of each other, and therefore $\langle y_n \rangle$ would be the same for both and thus independent of n.

Here for the first time I started writing things like $E(Y_n)$ as a sum over the **full p function**, and I needed some "symmetry rule" on the p function to prove that $E(Y_n) = E(Y_{n+1})$. But I kept being stuck with "end effects" that made things not work. I wanted to use my "correlation distance" idea maybe to say those ends effects did not matter, but I could never find a way to make that idea fly. Abrupt end.

Plan C. Here I try to make the end effects go away by treating only **infinite sequences**. Then you claim that a shift reproduces the same ensemble and you get $E(Y_n) = E(Y_{n+1})$. But arguments all shaky and I did not pursue this.

Plan D is my "**useful ensemble**" idea that includes all cyclic rotations and the argument is that seam repairs can be done and don't affect calculations much. This is the idea I "went with" in App G (f). I had trouble with notation showing that the p function was cyclic, confusion between variables and values. Had to do with the meaning of the expression used to compute the big sum over the p function. There was a paradox and then I resolved it. The final writeup of this "useful ensemble" idea is elsewhere.

Plan A

1. I was looking for a way to characterize "partial correlation". Here is one way:

$$p(y_n, y_{n+k}) = p(y_n)p(y_{n+k}) + \varepsilon(k)q(y_n, y_{n+k}) \qquad \text{where } \max(q) = 1$$

The idea is that ε(k) is some real function that decreases as k increases. This is then how you show that as k increases, things are becoming more uncorrelated, a way to do it gradually.

2. Now consider

$$\text{cov}(y_n, y_{n+k}) = \Sigma_{yn,yn+k} \; [y_n - \mu_n][\,[y_{n+k} - \mu_{n+k}]\; p(y_n, y_{n+k})$$

$$= \Sigma_{yn,yn+k} \; [y_n - \mu_n][\,[y_{n+k} - \mu_{n+k}]\; \varepsilon(k) q(y_n, y_{n+k})$$

where we know the contribution of $p(y_n)p(y_{n+k})$ will be zero. Then we have

$$\text{cov}(y_n, y_{n+k}) = \varepsilon(k)\; \Sigma_{yn,yn+k} \; [y_n - \mu_n][\,[y_{n+k} - \mu_{n+k}]\; q(y_n, y_{n+k})$$

and as k increases, this formal expression for correlation clearly decreases.

I don't know how to use this idea yet, but will ponder.

3. A Paradox? Consider

$$p(y_m) = \Sigma_{YN}\, p(y_m,y_N) = \Sigma_{YN}\, [p(y_m)\, p(y_N) \; + \varepsilon(|m-N|)\; q(y_m,y_N)]$$

$$= p(y_m)\; [\Sigma_{YN}\, p(y_N)\,] + \varepsilon(|m-N|) \;\; \Sigma_{YN}\, q(y_m,y_N)$$

$$= p(y_m) + \varepsilon(|m-N|) \;\; \Sigma_{YN}\, q(y_m,y_N)$$

and therefore

$$\varepsilon(|m-N|) \;\; \Sigma_{YN}\, q(y_m,y_N) = 0$$

and therefore

$$\Sigma_{YN}\, q(y_m,y_N) \;\; = 0$$

I guess this must be a property of q. Do I believe that? Go back to

$$p(x, y) = p(x)p(y) + \; \varepsilon(|\text{index}(x)-\text{index}(y)|)q(x, y)$$

$$p(x) = \Sigma_y\, p(x, y) \;\; = p(x) + \varepsilon(|\text{index}(x)-\text{index}(y)|)\; \Sigma_y\, q(x, y)$$

$$0 = \varepsilon(|\text{index}(x)-\text{index}(y)|)\; \Sigma_y\, q(x, y)$$

$$0 = \Sigma_y\, q(x, y)$$

It must be true. Maybe I can use this fact somehow.

$$p(x, y) = p(x)p(y) + \; \varepsilon(|\text{index}(x)-\text{index}(y)|)q(x, y)$$

$$p(y) = \Sigma_x\, p(x, y) \;\; = p(y) + \varepsilon(|\text{index}(x)-\text{index}(y)|)\; \Sigma_x\, q(x, y)$$

$0 = \varepsilon(|\text{index}(x)-\text{index}(y)|)\ \Sigma_{\mathbf{x}}\ q(x, y)$

$0 = \Sigma_{\mathbf{x}}\ q(x, y)$

Result is symmetric, fine.

4. Consider

$E(y_n) = \Sigma_{\mathbf{yn,yN}}\ (y_n)\ p(y_n,y_N)$

$= \Sigma_{\mathbf{yn,yN}}\ (y_n)\ p(y_n)p(y_N) + \Sigma_{\mathbf{yn,yN}}\ \varepsilon(|n-N|)q(y_n, y_N)$

$= \Sigma_{\mathbf{yn}}\ (y_n)\ p(y_n) + \Sigma_{\mathbf{yn,yN}}\ \varepsilon(|n-N|)q(y_n, y_N)$

But this took me nowhere.

Plan B. Let s be a "seed index". The Apparatus generates all the symbols of a given sequence based on this seed, so we really have

$E(Y_n) = \Sigma_{\mathbf{s}}\ y_n(s)\ p(Y_1= y_1(s), Y_2= y_2(s), \ldots\ Y_n= y_n(s), \ldots Y_N= y_N(s))$

The p-function here gives the probability of Exp #1 Apparatus creating sequence $\{ y_1(s), \ldots y_N(s) \}$. We know that Exp #2 creates a sequence $\{ y_2(s),\ldots y_N(s)\ y_{N+1}(s) \}$ based on the same seed s. These two sequences have their N-2 interior elements the same, by the way. We claim that this second sequence has the same probability as the first sequence (argument not stated right here), so the above becomes

$E(Y_n) = \Sigma_{\mathbf{s}}\ y_n(s)\ p(Y_1= y_2(s), Y_2 = y_3(s), \ldots ,Y_{n-1} = y_n(s), Y_n= y_{n+1}(s), \ldots Y_N= y_{N+1}(s))$

Now replace the name $y_n(s)$ by x(s) and the above says

$E(Y_n) = \Sigma_{\mathbf{s}}\ x(s)\ p(Y_1= y_2(s), Y_2 = y_3(s), \ldots ,Y_{n-1} = x(s), Y_n= y_{n+1}(s), \ldots Y_N= y_{N+1}(s))$

By definition this is $E(Y_{n-1})$!!! Thus we have shown that $E(Y_n) = E(Y_{n-1})$ even for finite N.

Let's examine this for N = 2 length sequences to see the most extreme case.

- $E(Y_2) = \Sigma_{\mathbf{s}}\ y_2(s)\ p(Y_1= y_1(s), Y_2= y_2(s))$ seq = $\{y_1(s), y_2(s) \}$
- $p(Y_1= y_1(s), Y_2= y_2(s)) = p(Y_1= y_2(s), Y_2= y_3(s))$ seq = $\{y_2(s), y_3(s) \}$

Therefore

$E(Y_2) = \Sigma_{\mathbf{s}}\ y_2(s)\ p(Y_1= y_2(s), Y_2= y_3(s))$

$= \Sigma_{\mathbf{s}}\ x(s)\ p(Y_1= x(s), Y_2= y_3(s))$

$= E(Y_1)$

Here I never talk about anything being cyclic. STOP

Plan C. The Apparatus creates an ensemble containing an infinite number of infinite length sequences. This is Ensemble #1. We write down this ensemble, paying attention mainly to a few random variables y_i near the center of the infinite sequences. Here we schematically list the ensemble:

..... $Y_5Y_6Y_7$	
..... e f g	sequence A
....	
....	
....	Ensemble #1

Although the sequences are very long, we assume that the number of ensemble entries is extremely large, so that every legal Apparatus sequence appears in the ensemble at least once, perhaps many times. We then experimentally determine p(Y6 = x) for the ensemble and we get some result

p(Y6 = x) = f(x) .

Next, we create Ensemble #2 to shifting each sequence in Ensemble #1 one symbol to the right. We then get

..... $Y_5Y_6Y_7$	
..... d e f	sequence B
....	
....	
....	Ensemble #2

For Ensemble #2, since all sequences are just shifted one symbol, we will find that p(Y7 = x) = f(x).

Since Ensemble #1 contains all possible legal sequences, it must contain sequence B in addition to sequence A.

****

Plan D. I looked back at FT where I first talk about statistical stuff. I talk there about an ensemble with i from 1 to I. Here is an idea to make things work a little better.

The Useful Ensemble. To construct a useful ensemble of pulse trains, we first imagine generating some number M of pulse trains of length N from whatever device we have to do this, such as a line code encoder using random input data streams. We then take this ensemble of M pulse trains and increase the size of the ensemble by including in it all cyclic permutations (rotations) of the M pulse trains, so now the enlarged ensemble contains I = M*N pulse trains. Depending on the rules of the encoder generating the

pulse trains, these rotated pulse trains might be "illegal" at the boundary where the pulse train wrapped around on itself. For example with N = 5 we obtain this sequence plus its rotations,

```
A B C D E
B C D E|A
C D E|A B
D E|A B C
E|A B C D   .
```

Here the seam is marked by a bar | . If the seam were "illegal", we could probably repair the seam by replacing A with some other symbol Q to get a legal sequence of symbols,

```
A B C D E
B C D E Q
C D E Q B
D E Q B C
E Q B C D
```

Perhaps several symbols near the seam would have to be "repaired" in this manner, but we shall assume just one as shown above. Notice that each column has at most one bad symbol Q, and that is 1/N of the symbols in the column. If N = 100, then only 1% of the symbols in a column are wrong after adjustment. If a 2-symbol repair were needed, then at most 2% in a column would be wrong.

We imagine doing this for all M of our initial pulse trains to get our final set of I = M*N pulse trains. In this "useful ensemble", in each vertical column, 1/N of the symbols might be wrong.

Now we associate each column with a random variable

$$Y_1\ Y_2\ Y_3\ Y_4\ Y_5$$

Fact: The expected value $E(Y_n) \equiv \langle y_n \rangle$ is independent of n.

Proof: Each column has at most 1 bad symbol like Q for each initial ensemble sequence. So if there are now I = M*N pulse trains in the ensemble, then in each column there will be at most N bad symbols. When the expected value is computed by summing the elements of the column n and dividing by I,

$$E(Y_n) = (1/I) \sum_{i=1}^{I} y_n^{(i)} = (1/I) \sum_{i=1}^{I} (y_n)_i \qquad \text{// two notations for the same thing}$$

one finds that the sums for different n are almost identical with an error on the order of 1/N. In our example the first column sums to A+B+C+D+E whereas the other columns are Q+B+C+D+E . Of course the exact error depends on what symbol was used to do the repair and what the palette of symbol values is, and so on, but the general order is 1/N. As N is made large, this error approaches 0 and then all the columns have the same expected value. **QED**

Now recall that the probability distribution for our N=5 ensemble is written in two ways

$$p(a,b,c,d,e) \equiv p(Y_1=a, Y_2=b, Y_3=c, Y_4=d, Y_5=e) .$$

Fact: For our useful ensemble, the probability distribution p(a,b,c ...) has cyclic symmetry.

Example: $p(a,b,c,d,e) = p(e,a,b,c,d) = p(d,e,a,b,c) = p(e,a,b,c,d)$

Proof: This might seem obvious, but a proof is warranted. We show this for the case N = 5; the general case should then be obvious. Using the notation introduced in (G.37) we can imagine evaluating p(a,b,c,d,e) this way, where [....] indicates an ordered sequence of numbers,

$$p(a,b,c,d,e) = (1/I) \Sigma_{i=1}^{I} ([a,b,c,d,e] = [(y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i])$$

Since our ensemble contains all cyclic rotations of $[(y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i]$, we don't change the sum shown above by replacing this with $[(y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i]$, we merely re-order the sum. In our example above, we can rewrite the 5 sequences on the left in this fancier notation,

		$(y_1)_i$	$(y_2)_i$	$(y_3)_i$	$(y_4)_i$	$(y_5)_i$
α	A B C D E	$(y_1)_1$	$(y_2)_1$	$(y_3)_1$	$(y_4)_1$	$(y_5)_1$
β	B C D E A	$(y_1)_2$	$(y_2)_2$	$(y_3)_2$	$(y_4)_2$	$(y_5)_2$
δ	C D E A B	$(y_1)_3$	$(y_2)_3$	$(y_3)_3$	$(y_4)_3$	$(y_5)_3$
γ	D E A B C	$(y_1)_4$	$(y_2)_4$	$(y_3)_4$	$(y_4)_4$	$(y_5)_4$
ε	E A B C D	$(y_1)_5$	$(y_2)_5$	$(y_3)_5$	$(y_4)_5$	$(y_5)_5$

We now rotate the above columns one column to the left to get

		$(y_2)_i$	$(y_3)_i$	$(y_4)_i$	$(y_5)_i$	$(y_1)_i$
β	B C D E A	$(y_2)_1$	$(y_3)_1$	$(y_4)_1$	$(y_5)_1$	$(y_1)_1$
δ	C D E A B	$(y_2)_2$	$(y_3)_2$	$(y_4)_2$	$(y_5)_2$	$(y_1)_2$
γ	D E A B C	$(y_2)_3$	$(y_3)_3$	$(y_4)_3$	$(y_5)_3$	$(y_1)_3$
ε	E A B C D	$(y_2)_4$	$(y_3)_4$	$(y_4)_4$	$(y_5)_4$	$(y_1)_4$
α	A B C D E	$(y_2)_5$	$(y_3)_5$	$(y_4)_5$	$(y_5)_5$	$(y_1)_5$

As it most easily seen on the left, this rotation by one column has just reordered the ensemble sequences from the original order α,β,γ,δ,ε to the new order β,γ,δ,ε,α. The same set of samples is included in the sum above regardless of how they are ordered.

Consider our starting expression

$$p(a,b,c,d,e) = (1/I) \Sigma_{i=1}^{I} ([a,b,c,d,e] = [(y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i])$$

Note that c is a specific number, as is $(y_5)_3$. Neither of these is a variable. Now suppose the ensemble has the property just noted above, that we can reorder the sum this way,

$$p(a,b,c,d,e) = (1/I)\,\Sigma_{i=1}^{I}\,(\,[\,a,b,c,d,e] \;=\; [(y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i]\,)\,.$$

The equality of the two sequences of numbers shown in the truth test [...] = [...] does not change if we alter each sequence in the same way, so rewrite again as

$$p(a,b,c,d,e) = (1/I)\,\Sigma_{i=1}^{I}\,(\,[\,e,a,b,c,d] \;=\; [(y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i,]\,)\,.$$

But the expression on the right is the definition of p(e,a,b,c,d). Thus we have shown that

$$p(a,b,c,d,e) = p(e,a,b,c,d)\,.$$

We can then show in the same way that p(e,a,b,c,d) = p(d,e,a,b,c) and so on, so in the end all cyclic permutations (rotations) of p(a,b,c,d,e) are equal for our useful ensemble. **QED**

Repeat this in order notation

****

$$p(y_1, y_2, y_3, y_4, y_5) \;\equiv\; p(Y_1{=}y_1, Y_2{=}y_2, Y_3{=}y_3, Y_4{=}y_4, Y_5{=}y_5)\,.$$

Fact: For our useful ensemble, the probability distribution $p(y_1, y_2, y_3....)$ has cyclic symmetry.

Proof: This might seem obvious, but a proof is warranted. We show this for the case N = 5; the general case should then be obvious. Using the notation introduced in (G.37) we can imagine evaluating $p(y_1, y_2, y_3, y_4, y_5)$ this way, where [....] indicates an ordered sequence of numbers,

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I)\,\Sigma_{i=1}^{I}\,(\,[y_1, y_2, y_3, y_4, y_5] \;=\; [\,(y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i\,]\,)$$

Since our ensemble contains all cyclic rotations of $[\,(y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i\,]$, we don't change the sum shown above by replacing this with $[\,(y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i\,]$, we merely re-order the sum. In our example above, we can rewrite the 5 sequences on the left in this fancier notation,

		$(y_1)_i$	$(y_2)_i$	$(y_3)_i$	$(y_4)_i$	$(y_5)_i$
α	A B C D E	$(y_1)_1$	$(y_2)_1$	$(y_3)_1$	$(y_4)_1$	$(y_5)_1$
β	B C D E A	$(y_1)_2$	$(y_2)_2$	$(y_3)_2$	$(y_4)_2$	$(y_5)_2$
δ	C D E A B	$(y_1)_3$	$(y_2)_3$	$(y_3)_3$	$(y_4)_3$	$(y_5)_3$
γ	D E A B C	$(y_1)_4$	$(y_2)_4$	$(y_3)_4$	$(y_4)_4$	$(y_5)_4$
ε	E A B C D	$(y_1)_5$	$(y_2)_5$	$(y_3)_5$	$(y_4)_5$	$(y_5)_5$

We now rotate the above columns one column to the left to get

		$(y_2)_i$	$(y_3)_i$	$(y_4)_i$	$(y_5)_i$	$(y_1)_i$
β	B C D E A	$(y_2)_1$	$(y_3)_1$	$(y_4)_1$	$(y_5)_1$	$(y_1)_1$
δ	C D E A B	$(y_2)_2$	$(y_3)_2$	$(y_4)_2$	$(y_5)_2$	$(y_1)_2$
γ	D E A B C	$(y_2)_3$	$(y_3)_3$	$(y_4)_3$	$(y_5)_3$	$(y_1)_3$
ε	E A B C D	$(y_2)_4$	$(y_3)_4$	$(y_4)_4$	$(y_5)_4$	$(y_1)_4$
α	A B C D E	$(y_2)_5$	$(y_3)_5$	$(y_4)_5$	$(y_5)_5$	$(y_1)_5$

As it most easily seen on the left, this rotation by one column has just reordered the ensemble sequences from the original order α,β,γ,δ,ε to the new order β,γ,δ,ε,α. The same set of samples is included in the sum above regardless of how they are ordered.

Consider our starting expression

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I)\,\Sigma_{i=1}^{I}\,([y_1, y_2, y_3, y_4, y_5] = [\,(y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i\,]) .$$

Since our useful ensemble has the property just noted above, we can reorder the sum this way,

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I)\,\Sigma_{i=1}^{I}\,([y_1, y_2, y_3, y_4, y_5] = [(y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i]) .$$

The equality of the two sequences of numbers shown in the truth test [...] = [...] does not change if we alter each sequence in the same way, so rewrite again as

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I)\,\Sigma_{i=1}^{I}\,([y_5, y_1, y_2, y_3, y_4] = [\,(y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i\,]) .$$

But the expression on the right is the definition of $p(y_5, y_1, y_2, y_3, y_4)$. Thus we have shown that

$$p(y_1, y_2, y_3, y_4, y_5) = p(y_5, y_1, y_2, y_3, y_4) .$$

We can then show in the same way that $p(y_5, y_1, y_2, y_3, y_4) = p(y_4, y_5, y_1, y_2, y_3)$ and so on, so in the end all cyclic permutations (rotations) of $p(y_1, y_2, y_3, y_4, y_5)$ are equal for our useful ensemble when the sequence length N is large (we use N = 5 for demonstration only). **QED**

****

Why did this derivation fail in my other notation?

Let's just repeat each of the steps in the other notation,

In the shorthand on the left, one must imagine the Y_i values going Y_1 to Y_5 left to right, regardless of what particular values are used as arguments of p. Because our ensemble includes all cyclic permutations

of every sequence in the original ensemble, the p function is also cyclic in its argument values. This is because p(...) measures the counts of symbols of each type

In terms of probabilities, we would have for example

$$p(Y_3 = y_3) = \Sigma_{y1,y2,y4,y5}\, p(Y_1=y_1,Y_2= y_2,Y_3= y_3,Y_4= y_4,Y_5= y_5)$$

or in our condensed notation

$$p(y_3) = \Sigma_{y1,y2,y4,y5}\, p(y_1, y_2, y_3, y_4, y_5)$$

where $p(y_1, y_2, y_3, y_4, y_5)$ is the distribution associated with the ensemble of pulse trains. Due to our method of construction, this distribution function p is cyclic in its arguments. Thus

$p(Y_3 = x)$

$= \Sigma_{y1,y2,y4,y5}\, p(Y_1=y_1,Y_2= y_2,Y_3 = x,Y_4= y_4,Y_5= y_5)$ // definition of $p(Y_3 = x)$

$= \Sigma_{y2,y3,y5,y1}\, p(Y_1=y_2,Y_2= y_3,Y_3= x,Y_4= y_5,Y_5= y_1)$ // rename dummy summation variables

$= \Sigma_{y1,y2,y3,y5}\, p(Y_1=y_2,Y_2= y_3,Y_3= x,Y_4= y_5,Y_5= y_1)$ // trivially reorder sums on Σ

$= \Sigma_{y1,y2,y3,y5}\, p(Y_1=y_1,Y_2= y_2,Y_3= y_3,Y_4= x,Y_5= y_5)$ // use cyclic property of p function

$= p(Y_4 = x)$ // definition of $p(Y_4 = x)$

Thus, for any random variable in the sequence we have the same single-variable probability function we could then call p(x). We have just shown that

Fact: The single-random-variable probability function p(x) is the same for all positions in the pulse train provided N is reasonably large.

Of course if this is true, we recover the previous Fact since

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

Fact: For our ensemble of interest, p(a,b,c.....) has cyclic symmetry.

Example:

$$p(Y_1= a, Y_2= b, Y_3 = c) = p(Y_1= b, Y_2= c, Y_3 = a) = p(Y_1= c, Y_2= a, Y_3 = b)$$

or

$$p(a,b,c) = p(b,c,a) = p(c,a,b)$$

Paradox: Start with

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I) \Sigma_{i=1}^{I} (\{ (y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i) \} = \{ y_1, y_2, y_3, y_4, y_5) \})$$

Without changing a thing, I restate the equality this way

$$\{ (y_2)_i, (y_1)_i, (y_3)_i, (y_4)_i, (y_5)_i) \} = \{y_2, y_1, y_3, y_4, y_5) \}$$

Then I have

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I) \Sigma_{i=1}^{I} (\{ (y_2)_i, (y_1)_i, (y_3)_i, (y_4)_i, (y_5)_i) \} = \{y_2, y_1, y_3, y_4, y_5) \})$$

But this is the definition of $p(y_2, y_1, y_3, y_4, y_5)$. [wrong, this is now a different ensemble] Thus I have proven a symmetry which I don't think is true. What went wrong here?

**********8

$$p(Y_1=y_1,Y_2=y_2,Y_3=y_3) = (1/I)\Sigma_i (\{Y_1Y_2Y_3 =$$

The variable names $y_1...y_5$ are just dummy names on both sides of this equation, we could rename them any way we like, such as $y_1...y_5 \rightarrow$ a,b,c,d,e. But let's rename them this way

$$y_1, y_2, y_3, y_4, y_5 \rightarrow y_2, y_1, y_3, y_4, y_5$$

$$p(y_2, y_1, y_3, y_4, y_5) = (1/I) \Sigma_{i=1}^{I} (\{ (y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i) \} = \{ y_1, y_2, y_3, y_4, y_5) \})$$

to get

$$p(y_2, y_3, y_4, y_5, y_1) = (1/I) \Sigma_{i=1}^{I} (\{ (y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i) \} = \{ y_2, y_3, y_4, y_5, y_1) \})$$

$$p(y_2, y_3, y_4, y_5, y_1) = (1/I) \Sigma_{i=1}^{I} (\{ (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i) \} = \{ y_2, y_3, y_4, y_5, y_1) \})$$

In the

and do the Σ_i reordering as just described to get

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I) \Sigma_{i=1}^{I} (\{ (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i) \} = \{ y_1, y_2, y_3, y_4, y_5) \}) .$$

The variable names $y_1 ... y_5$ are just dummy names on both sides of this equation, we could rename them any way we like, such as $y_1 ... y_5 \rightarrow$ a,b,c,d,e. But let's rename them this way

$$y_1, y_2, y_3, y_4, y_5 \rightarrow y_2, y_3, y_4, y_5, y_1$$

Then the last equation above becomes

$$p(y_2, y_3, y_4, y_5, y_1) = (1/I) \Sigma_{i=1}^{I} (\{ (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i) \} = \{ y_2, y_3, y_4, y_5, y_1) \}) .$$

Next, we reorder elements within each set of the truth comparison, and in doing so we make no alteration to the truth test

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I) \Sigma_{i=1}^{I} (\{ (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i) \} = \{ y_2, y_3, y_4, y_5, y_1) \})$$

But this object is exactly the definition of $p(y_2, y_3, y_4, y_5, y_1)$.

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I) \Sigma_{i=1}^{I} (\{ (y_1)_i, (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i) \} = \{ y_1, y_2, y_3, y_4, y_5) \})$$
// initial form

$$p(y_1, y_2, y_3, y_4, y_5) = (1/I) \Sigma_{i=1}^{I} (\{ (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i) \} = \{ y_1, y_2, y_3, y_4, y_5) \})$$
// reordering sum as above

$$p(y_2, y_3, y_4, y_5, y_1) = (1/I) \Sigma_{i=1}^{I} (\{ (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i, (y_1)_i) \} = \{ y_2, y_3, y_4, y_5, y_1) \})$$
// reordering sum as above

$$= (1/I) \Sigma_{i=1}^{I} (\{ (y_1)_i , (y_2)_i, (y_3)_i, (y_4)_i, (y_5)_i,) \} = \{ y_2, y_3, y_4, y_5, y_1) \})$$
// reorder both sets with no change made in truth of equality