Probability theory Ng and spectral PhL 3.26.05

I think this is now cleared up and Spectral Doc now has Appendix D.

1. Let's try to establish the analogy between Ng trials and pulse trains. In each "trial" a set of "features" are measured. Movie cuddly or not, action or not, etc. In the pulse train, cuddly/non-cuddly = y_m and action/no_action = y_n where y_m and y_n are the amplitudes of two cars in the train. The number of trials goes with the number of pulse trains in the ensemble and that has index i and total count I. The analogy is not quite right because y_m takes values from the same set as y_n.

2. Consider this expectation value from my Ng notes:

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx\, dy\, x\, y\; p(x,y)$$

Here each variable takes a continuum of values in (-∞.∞). The integration endpoints could be (a,b) for some other case. But I am interested in the case where X can take certain discrete values we shall denote by index k, not to be confused with i. So we have y_k and x_k and each has perhaps a separate set of discrete values. Then the above would become

$$E(XY) = \Sigma_k \Sigma_{k'}\; x_k\, y_{k'}\; p(x_k, y_{k'})$$

Here $p(x_k,y_{k'})$ is the probability in the entire ensemble that X takes the value x_k at the same time that Y takes the value $y_{k'}$. If X and Y are correlated (features or pulse train positions), we cannot factor p into a simple product of two functions. If things are uncorrelated, we factor in this way

$$p(x_k,y_{k'}) = p_X(x_k) p_Y(y_{k'})$$

and then

$$E(XY) = \Sigma_k \Sigma_{k'}\; x_k\, y_{k'}\; p(x_k,y_{k'}) = E(X)E(Y)$$

3. Now consider this claim:

$$E(XY) = (1/I)\, \Sigma_{i=1}^{I}\; x^{(i)} y^{(i)}$$

Two questions: Is this true? How is it related to item 2 above? Here $x^{(i)}$ is the value x takes in the i^{th} trial and $y^{(i)}$ is the value y takes in that same trial. How does this connect with pulse trains?

$$E(Y_m Y_n) = (1/I)\, \Sigma_{i=1}^{I}\; y_m^{(i)} y_n^{(i)} \qquad \text{perhaps m-n = k}$$

Now the subscripts m and n are not summation indices, they are just labels which distinguish what I called X and Y in the previous equation. These are static labels for the moment. If we sum over all pairs as shown over the ensemble of trials or pulse trains, that is going to be the average value of that quantity.

In this form for E(XY), there is no probability function appearing as in the previous form. Instead, we are using the actual values from experiment, and that experiment knows about the probability function! The expectation value of some quantity must be the average value of that quantity if you average over a billion trials or pulse trains in an ensemble. So I think the answer to the first question is: yes, it is true.

4. If you had the experimental data from item 3, is there some way to compute the p function of item 2? I suspect you would need all "moments" and not just this one moment. But it must be possible to go the other direction, and I have already shown how, given p, you can compute E(XY).

5. In the pulse train case, I guess I am happy then with these two equations for non-correlated cases.

$$E(y_m y_n) = E(y_m)E(y_n)$$

$$E(y_m) = [p]\ A + (1-p)\ B$$

6. What about the standard deviation stuff? We can write

$$E([X-E(X)]^2) = \Sigma_k\ [x_k - E(X)]^2\ p_k(x_k)$$

or

$$E([X-\mu_x]^2) = \Sigma_k\ [x_k - \mu_x]^2\ p_k(x_k) = \Sigma_k\ x_k^2\ p_k(x_k) + \mu_k^2\ \Sigma_k\ p_k(x_k) - \Sigma_k\ (2\mu_x x_k)\ p_k(x_k)$$

$$= E(X^2) + \mu_k^2 - 2\mu_x^2 = E(X^2) - E(X)^2 = \text{variance} = \sigma^2$$

7. OK, I think all is well. What is correlation and covariance then?

$$\text{corr}(X,Y) = E(XY) = \Sigma_k \Sigma_{k'}\ x_k\ y_{k'}\ p(x_k, y_{k'})$$ // wrong!!!

I think this might be wrong! Ouch, I will then have an Ng mistake. Luckily I only made a small Ng comment about it. My Proakis notes show no corr(X,Y) object. Where did I get that idea from? Here is a wiki note which shows me wrong.

In probability theory and statistics, the mathematical descriptions of **covariance and correlation** are very similar.[1][2] Both describe the degree of similarity between two random variables or sets of random variables.

correlation $\phi_{XY} = E[(X - E[X])(Y - E[Y])]/(\sigma_X \sigma_Y)$

covariance $\gamma_{XY} = E[(X - E[X])\,(Y - E[Y])]$

Here is another PDF source confirming the above,

Throughout this section, we will use the notation $EX = \mu_X$, $EY = \mu_Y$, $\mathrm{Var}X = \sigma_X^2$, and $\mathrm{Var}Y = \sigma_Y^2$.

Definition 4.5.1 The covariance of X and Y is the number defined by

$$\mathrm{Cov}(X,Y) = E((X - \mu_X)(Y - \mu_Y)).$$

Definition 4.5.2 The correlation of X and Y is the number defined by

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}.$$

The value ρ_{XY} is also called the correlation coefficient.

So I need to fix this in my Ng notes very soon! There is only one occurrence of corr(in that doc so that will be easy to fix. // Ng notes are now all fixed and pushed out, so we can continue here. That was a very ugly error to have in t public document.

8. I am now ready to rewrite a small part of spectral doc relating to the above subject. // Well, it ended up being a whole new Appendix D.