Appendix D: Probability Theory: how α and β are related to μ and σ 1
(a) Random Variables X and Y 1
(b) Application to the Pulse Train Ensemble 3
(c) Main conclusions for the pulse train ensemble 4

Appendix D: Probability Theory: how α and β are related to μ and σ

We first review some of the basics of random variables X and Y: expectation values, means, covariance, variance and standard deviation. We do this in a context in which each variable takes on a set of discrete values rather than continuous values. We then translate our various facts into the language of X,Y → Y_m Y_n where the latter are random variables associated with the amplitudes at positions m and n of a pulse train. The labels m and n add a notational complexity that is a little difficult to comprehend unless things are written out explicitly, which is done below.

(a) Random Variables X and Y

If X and Y are random variables which can take discrete values x_k and y_k, one can express the mean value of each variable as

$$\mu_x = E(X) = \Sigma_k\, x_k\, p_x(x_k)$$
$$\mu_y = E(Y) = \Sigma_{k'}\, y_{k'}\, p_y(y_{k'}) \qquad (D.1)$$

where E(X) means the expectation value of X. The sum Σ_k is over all values in the set of values that X can take, and $\Sigma_{k'}$ is over all values that Y can take. Probability $p_x(x_k)$ is the probability that X takes the value x_k and similarly for $p_y(y_{k'})$. If we have a very large ensemble of I systems we can measure the above quantities in this manner,

$$\mu_x = E(X) = (1/I)\, \Sigma_{i=1}^{I}\, x^{(i)}$$
$$\mu_y = E(Y) = (1/I)\, \Sigma_{i=1}^{I}\, y^{(i)} \qquad (D.2)$$

where i labels a particular system in the ensemble, and $x^{(i)}$ and $y^{(i)}$ are the values X and Y take in ensemble system i. The probability functions like $p_x(x_k)$ do not appear in (D.2), but they are present in spirit, being embedded in the "experimental data" like $x^{(i)}$.

The expectation value E(XY) is defined by

$$E(XY) = \Sigma_k \Sigma_{k'}\, x_k\, y_{k'}\, p_{xy}(x_k, y_{k'}) \qquad (D.3)$$

where now $p_{xy}(x_k,y_{k'})$ is the joint probability that X has value x_k when Y has the value $y_{k'}$. One would measure E(XY) in this manner,

$$E(XY) = (1/I)\,\Sigma_{i=1}^{I}\, x^{(i)}\, y^{(i)} \,. \qquad (D.4)$$

If the random variables X and Y are independent (no correlation), then

$$p_{xy}(x_k,y_{k'}) = p_x(x_k)\, p_y(y_{k'}) \qquad (D.5)$$

and then it is obvious from the above definitions of E(XY), E(X) and E(Y) that

$$E(XY) = E(X)E(Y) \,. \qquad (D.6)$$

The covariance of X and Y is defined as,

$$cov(X,Y) = E(\,[X - \mu_x]\,[Y - \mu_y]\,) = \Sigma_k\Sigma_{k'}\,(x_k - \mu_x)\,(y_{k'} - \mu_y)\, p_{xy}(x_k,y_{k'}) \qquad (D.7)$$

and it could be measured for an ensemble in this way

$$cov(X,Y) = (1/I)\,\Sigma_{i=1}^{I}\,(x^{(i)} - \mu_x)\,(y^{(i)} - \mu_y) \qquad (D.8)$$

where μ_x and μ_y would be the measured values of E(X) and E(Y) as shown above. If X and Y are independent variables then we use (D.5) in (D.7) to find that

$$cov(X,Y) = E(\,[X - \mu_x]\,[Y - \mu_y]\,) = E(X - \mu_x)\,E(Y - \mu_y) \,. \qquad (D.9)$$

In the special case that X and Y are the same variable, we write cov(X,Y) = var(X), the variance of X, which is also defined as the square of the standard deviation σ_x,

$$var(X) \equiv \sigma_x^2 \equiv cov(X,X) = \Sigma_k\Sigma_{k'}\,(x_k - \mu_x)\,(x_{k'} - \mu_x)\, p_{xx}(x_k,x_{k'}) \,. \qquad (D.10)$$

But $p_{xx}(x_k,x_{k'}) = \delta_{k,k'}p_x(x_k)$ because if X has value x_k, it cannot also have a different value $x_{k'}$, so

$$p_{xx}(x_k,x_{k'}) = \delta_{k,k'}\, p_x(x_k) \qquad (D.11)$$

and then

$$var(X) \equiv \sigma_x^2 \equiv cov(X,X) = \Sigma_k\,(x_k - \mu_x)^2\, p_x(x_k) \qquad (D.12)$$

with experimental measurement

$$var(X) \equiv \sigma_x^2 \equiv cov(X,X) = (1/I)\,\Sigma_{i=1}^{I}\,(x^{(i)} - \mu_x)^2 \,. \qquad (D.13)$$

Finally, we note that

$$cov(X,Y) = E(\,[X - \mu_x]\,[Y - \mu_y]) = E(XY) - \mu_x E(Y) - \mu_y E(X) + \mu_x\mu_y E(1)$$

$$= E(XY) - \mu_x\mu_y - \mu_y\mu_x + \mu_x\mu_y$$

$$= E(XY) - \mu_x\mu_y \ . \qquad \text{(D.14)}$$

When X and Y are the same random variable this says

$$\text{var}(X) \equiv \sigma_x^2 \equiv \ \text{cov}(X,X) \ = E(X^2) - \mu_x^2 \qquad . \qquad \text{(D.15)}$$

(b) Application to the Pulse Train Ensemble

In the above we set

$X = Y_m$ = random variable associated with the amplitudes $y_m^{(i)}$ at pulse location m in pulse train
$Y = Y_n$ = random variable associated with the amplitudes $y_n^{(i)}$ at pulse location n in pulse train

Things are a little complicated due to the extra labels m and n, but we can nevertheless carefully translate all the X and Y equations above. Also, we shall add the following extra notation for an expectation value to be consistent with our earlier work,

$$\langle x \rangle \ \equiv \ E(X) \qquad \langle xy \rangle \ \equiv \ E(XY) \qquad \langle x^2 \rangle \ \equiv \ E(X^2) \ . \qquad \text{(D.16)}$$

The first pair of equations above translate into the following:

$$\begin{aligned} \mu_{ym} \ &= E(Y_m) = \langle y_m \rangle \ = \Sigma_k \ (y_m)_k \ p_{ym}[(y_m)_k] \ \ = \Sigma_k \ (y_m)_k \ p[(y_m)_k] \ = \mu \\ \mu_{yn} \ &= E(Y_n) = \langle y_n \rangle \ = \Sigma_k \ (y_n)_k \ p_{yn}[(y_n)_k] \ \ = \Sigma_k \ (y_m)_k \ p[(y_m)_k] \ = \mu \ . \end{aligned} \qquad \text{(D.17)}$$

Here $(y_m)_k$ is the k^{th} value in the set of all possible amplitudes that position m in the pulse train can have. For example, we might have $(y_m)_1$ = A and $(y_m)_2$ = B. For the pulse train, the set of allowed amplitude values is the same for any pulse in the train, and the probability distribution is also the same, yielding the simplifications shown above. In particular, the means μ_{ym} and μ_{yn} don't depend on position m and n, so we can just call both these means μ. Here is how the above means would be measured for an ensemble,

$$\begin{aligned} \mu = \mu_{ym} \ &= E(Y_m) = \langle y_m \rangle = \ (1/I) \ \Sigma_{i=1}^{I} \ y_m^{(i)} \\ \mu = \mu_{yn} \ &= E(Y_n) = \langle y_n \rangle = \ (1/I) \ \Sigma_{i=1}^{I} \ y_n^{(i)} \ . \end{aligned} \qquad \text{(D.18)}$$

Next, for E(XY) we translate (D.3) to get ,

$$E(Y_mY_n) = \langle y_my_n \rangle = \Sigma_k\Sigma_{k'} \ (y_m)_k \ (y_n)_{k'} \ p_{ym,yn}[(y_m)_k, (y_n)_{k'}]$$

$$E(Y_mY_n) = \langle y_my_n \rangle = (1/I) \ \Sigma_{i=1}^{I} \ y_m^{(i)} \ y_n^{(i)} \qquad . \qquad \text{(D.19)}$$

If Y_m and Y_n are uncorrelated (which requires that $m \neq n$) then in analogy with (D.5),

$$p_{ym,yn}[(y_m)_k, (y_n)_{k'}] \ = p_{ym}[(y_m)_k] \ p_{yn}[(y_n)_k] \ = p[(y_m)_k] \ p[(y_n)_k] \qquad \text{(D.20)}$$

and as in (D.6) we find that, in the case of no correlation,

$$E(Y_m Y_n) = \langle y_m y_n \rangle = E(Y_m)\, E(Y_n) = \langle y_m \rangle \langle y_n \rangle \quad . \qquad m \neq n \qquad \text{(D.21)}$$

Obviously if m = n, then Y_m and Y_n cannot be "uncorrelated" since they are the same random variable. If m = n, then we have instead this version of (D.11),

$$p_{y_m, y_m}[(y_m)_k, (y_m)_{k'}] = \delta_{k,k'}\, p_{y_m}[(y_m)_k] = \delta_{k,k'}\, p[(y_m)_k] \qquad \text{(D.22)}$$

and then

$$E(Y_m^2) = \langle y_m^2 \rangle = \Sigma_k\, [(y_m)_k]^2\, p[(y_m)_k] \qquad \text{(D.23)}$$

which in experiment (over the ensemble) would be measured as

$$E(Y_m^2) = \langle y_m^2 \rangle = (1/I)\, \Sigma_{i=1}^{I}\, (y_m^{(i)})^2 \, . \qquad \text{(D.24)}$$

The covariance equations translated from (D.7) and (D.8) become

$$\mathrm{cov}(Y_m, Y_n) = E(\, [Y_m - \mu]\, [Y_n - \mu]\,) = \Sigma_k \Sigma_{k'}\, ((y_m)_k - \mu)\, ((y_m)_{k'} - \mu)\, p_{y_m, y_n}[(y_m)_k, (y_n)_{k'}]$$

$$\mathrm{cov}(Y_m, Y_n) = (1/I)\, \Sigma_{i=1}^{I}\, (y_m^{(i)} - \mu)\, (y_n^{(i)} - \mu) \, . \qquad \text{(D.25)}$$

The variance (and standard deviation) translations from (D.12) and (D.13) are.

$$\mathrm{var}(Y_m) \equiv \sigma^2 \equiv \mathrm{cov}(Y_m, Y_m) = \Sigma_k\, [(y_m)_k - \mu]^2\, p[(y_m)_k]$$

$$\mathrm{var}(Y_m) \equiv \sigma^2 \equiv \mathrm{cov}(Y_m, Y_m) = (1/I)\, \Sigma_{i=1}^{I}\, (y_m^{(i)} - \mu)^2 \, . \qquad \text{(D.26)}$$

Finally, (D.14) and (D.15) become,

$$\mathrm{cov}(Y_m, Y_n) = E(Y_m Y_n) - \mu_{y_x} \mu_{y_n} = E(Y_m Y_n) - \mu^2$$

$$\mathrm{var}(Y_m) \equiv \sigma^2 \equiv \mathrm{cov}(Y_m, Y_m) = E(Y_m^2) - \mu^2 = \langle y_m^2 \rangle - \mu^2 \quad \Rightarrow \langle y_m^2 \rangle = \sigma^2 + \mu^2 \, . \qquad \text{(D.27)}$$

(c) Main conclusions for the pulse train ensemble

The following quantities are measured from the pulse train ensemble as follows,

$$\langle y_m \rangle = (1/I)\, \Sigma_{i=1}^{I}\, y_m^{(i)} = \mu \qquad \textit{(D.18)}$$

$$\langle y_m y_n \rangle = (1/I)\, \Sigma_{i=1}^{I}\, y_m^{(i)}\, y_n^{(i)} \qquad \text{any m and n} \qquad \textit{(D.19)}$$

$$\langle y_m^2 \rangle = (1/I)\, \Sigma_{i=1}^{I}\, [y_m^{(i)}]^2 = \sigma^2 + \mu^2 \qquad \textit{(D.27)} \qquad \text{(D.28)}$$

where we show how each quantity is related to symbols μ and σ. When the amplitudes at different locations in the pulse train are uncorrelated, we can further write

$$\langle y_m y_n \rangle = \langle y_m \rangle \langle y_n \rangle = \mu^2 \qquad m \neq n \ . \qquad \textit{(D.21)} \qquad \text{(D.29)}$$

We can now restate equation (35.6) in this more statistically-oriented manner,

$$\langle y_m \rangle = \mu = [p]A + [1\text{-}p]B$$

$$n \neq m \qquad \alpha \equiv \langle y_m y_n \rangle = \mu^2 = \{\, [p]A + [1\text{-}p]B \,\}^2 = [pp]\, AA + [p(1\text{-}p)]\, 2AB + [(1\text{-}p)(1\text{-}p)]BB$$

$$n = m \qquad \beta \equiv \langle y_m^2 \rangle = [p]AA + [(1\text{-}p)]\, BB = \sigma^2 + \mu^2 \qquad \text{(D.30)}$$

which is to say

$$\alpha = \mu^2$$
$$\beta = \mu^2 + \sigma^2 = \alpha + \sigma^2$$
$$\sigma^2 = (\beta\text{-}\alpha) \ . \qquad \text{(D.31)}$$

Using our expressions for α and β we can compute σ^2:

```
alpha := (p*A + (1-p)*B)^2;
```

$$\alpha := (p\,A + (1-p)\,B)^2$$

```
beta := p*A^2+(1-p)*B^2;
```

$$\beta := p\,A^2 + (1-p)\,B^2$$

```
var := beta-alpha: factor(%);
```

$$-p\,(B-A)^2\,(-1+p)$$

which says

$$\sigma^2 = \text{variance} = [\, p(1\text{-}p)]\, (A\text{-}B)^2 \ . \qquad \text{(D.32)}$$

If p = 0 or p = 1, this variance vanishes as one would expect.