**Weiner Process**

The Wiener process forms the basis, for my purposes, in Stochastic Calculus. So I guess we’ll start here.

**Scalar Weiner Process**

Let’s specialize to the case where the probability distribution, rather than just the moments, of the driving random variable is known. We’ll presume it’s white noise, so that:



Well, since we have a continuous process, P should be a functional really. In general, I think we actually have:



where w(t) = dW/dt. The old formulas follow:



We can write this another way,



From this we can define the Wiener process.



Being the sum of normally distributed variables, its mean and variance would be:



and it’s probability distribution would be normal as well, given by:



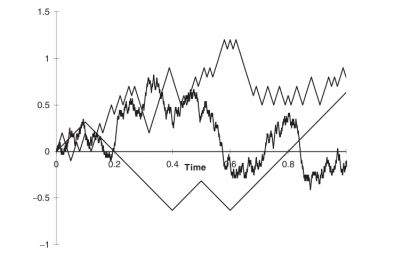
Another moment ofi interest is <W(t)W(t´)>. What’s this? I’ll take t´ > t for sake of discussion,



where we note that dW(t) and dW(t´) are independent, and so have zero correlation, unless t = t´. So we can see we have in general,



Or could say it equals D×the overlap of the two times. Because the Wiener process can be expressed as the sum of n = anything, independent and identically distributed variables, it is considered and **infinitely divisible process**. It looks something like this:



And let’s consider as well: ΔW = W(a+Δt) – W(a). Clearly <ΔW> = 0. What about the variance?



So the moments are:



and its pdf is:



Let’s split the (0, t) time interval into three parts (0, t1), (t1, t2), (t2, t). And let’s write,



What’s the correlation between the increments? Let’s note first,



And then, let i > j, we’ll notate W(t2) – W(t1) as W(ti + Δt) – W(t1), say. Then,



So the increments are independent. And if i = j,



**Example**

Is W(t) = W(t1) + W(t2) if t1 + t2 = t? We can calculate the variance of the difference and see if its nonzero. If it is, then the equality doesn’t hold.



So no. This way of judging equality between two random variables X(t), Y(t), i.e., by checking whether the variance of X(t) – Y(t) is zero, is a method we’ll use often. Consider,



Seems kind of requisite that X(t) = Y(t) for this to hold, as this is the sum of purely positive numbers otherwise.

**Algebraic Functions of Weiner Process**

Now let’s consider functions of W(t), i.e. F(W(t)). This is of course similar to the thing in Mello’s notes. This is also a random variable, and we’d like to ascertain its distribution. First we’ll consider algebraic operations on W(t). If we have F(t) = bW(t) and we want to determine average, variance, and p.d.f. of F(t). Well,



How about a linear relation:



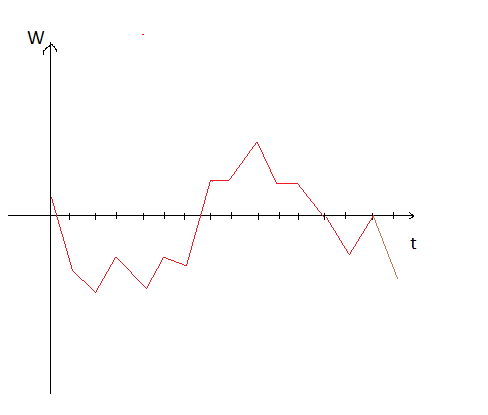
and what if we had a simple quadratic, F(t) = bW(t)2? Well in that case let’s remember that W(t) is also a Gaussian variable with average 0 and st.dev. Dt, and so we can compute statistics of this variable using the ‘GF’ techniques in the Appendix.



and this is actually a χ2 distribution, as we might recognize from discussion in Prob Stat Folder/Multi-variable pdf’s.

**Integral functions of Weiner process**

OK, and now for integration. First we’ll introduce the concept of integrating a Weiner process, and then we’ll want to work out the moments and pdf of an integrated Weiner process. So consider a Weiner process W(t) illustrated below. The process looks discretized, but we should have in mind a continuous one.



If we were to naively define an integral from 0 to t, we would chop the interval up into times tj, and write:



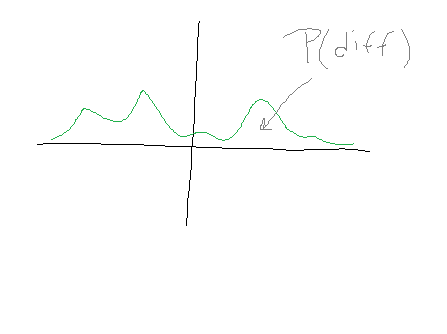
where tj\* is some time inbetween tj and tj+1. We will propose some formula for this in terms the end points, like this:



Where F is the anti-derivative of f w/r to W, and Remainder, as we’ll see, is not zero, but depends on where t\*j+1 is, inside the interval [tj, tj+1]. This is at odds with a regular Reimann sum, whose value does not depend on this detail. Equality is adjudicated in the sense that the integral and the anti-derivative + Remainder have the same probability distribution (‘cause note that each are just random variables). The easier way to enforce ‘same probability distribution’ is to require the expectation of their square difference to be zero, so that F is considered equal to ∫fdW if:



(in the limit that the partition size goes to zero). To get a sense of this criterion, consider two independent variables x and y. They are mean square equal if: <(x-y)2> = <x2> + <y2> - 2<x><y> = 0. If these had same mean and variance, then we’d have 2<x2> - 2<x>2 = 0, which would imply that their variance were 0, which would be a delta distribution – a pretty stringent condition. The condition basically means that their probability distributions can’t be different over any finite range because if they were, then that difference would contribute to the expectation and give a non-zero result. For instance, as can see, the only way to get <(diff)> = 0 is for P(difference) = 0 everywhere, except at perhaps a finite/denumerable # of *points*, strictly speaking.



Now note that the normal rules of calculus don’t apply here. For instance consider:



This cannot be true, as evinced by the non-equivalence of their expectations.



If we make tj\* = tj, then because the increment dW = W(tj+1) – W(tj) is independent of W(tj) itself, the expectation would be zero. Or another way to look at it is:



So from this perspective, we get zero not because the ΔW is independent of W, but because both terms give the same (up to a differentially small amount) the same thing. Anyway, if we were to place tj\* in between tj and tj+1, then they wouldn’t be independent, and so we probably wouldn’t get 0. Let’s consider the value of this integral in more detail. Let tj\* = tj + λ(tj+1 – tj), where λ ∈ [0,1]. Then…(basically we will want to break this expression up into all of its independent increments)…



Here we observe that the last term is the same as the integral itself. And so we have:



where in the last line we get W(t)2 – W(0)2 because that preceding sum is a telescoping sum. The last two terms are independent Gaussian deviants which we can write, in terms of normalized (μ = 0, σ = 1) Gaussian variables ε1j and ε2j. It seems to me that this would be an exact replacement too (?), since the deviants are the sum of Gaussian variables, dW, which is therefore itself Gaussian, and will have the posited variance.



Now is it ok to replace these random terms by their expectations? If so then we’d get (note <ε2> = <ε2> - <ε>2 = σ2 = 1):



Let’s check if this proposal works in the mean square sense.



Now we need only consider expectations from cross terms in that term there, as the diagonal ones will involve (ti+1 – ti)2 = (Δt/N)2, which when summed from i = 0 to N will still leave us with a factor of Δt2/N, which will go to zero in the large N limit. So then,



In the second line we note that we aren’t including i = j terms and so are missing 1 of the N terms in the sum. But this is negligible since it’s of order 1/N. So this checks out. So we do have:



And we can see that the result of our integration does depend on our choice of tj\*. There are two popular choices. One is to set λ = 0 so that tj\* = tj. And the other is to set λ = ½ so that tj\* = tj + (1/2)(tj+1 – tj).

λ = 0 is called the Ito calculus.

λ = ½ is called the Stratonvich calculus. Note that Stratonovich version obeys the ordinary rules of Calculus.

Interesting that the result is completely deterministic. Let’s consider integration of general functions more closely now. First of all,



And now consider the antiderivative F(W,t) for which ∂F/∂W = f. And let’s consider a Taylor series of F. I think we *could* series it out directly from t = 0 to t = t? Certainly in the discrete case, the resulting function would still be well defined. But in the present m.o. we’re going to chop up the time interval into pieces. And I imagine that as we shrink the time interval down, all higher order terms will vanish naturally.



Now it seems that we usually replace the ΔW2’s with the equivalent DΔt. But we’ll observe that while one could more or less do this for the first term in the brackets, since that ΔW2 is independent of the F´´ term (because the F´´ term has dW’s up to time k+λ whereas the ΔW2 term, after subtraction is taken account of, has dW’s between times k+λ and k+1), we cannot do this with the second ΔW2 term, since there will be overlap between its dW’s and the F´´’s dW’s. Still I think it’s justified to make the DΔt replacement because if we expand F´´ about tk, instead of tk+λ, only the first term in the sum ought to contribute. Well let’s check. Doing the mean square thing between the discrete integral and the proposed replacement requires the following condition:



It seems that the only term that has the possibility to survive the limit Δtk ~ Δt/N → 0 is indeed the one with just the zeroth order g term. In this case it looks like the equality will survive. If so, then we’d have:



As usual we can stop at this order in the expansion, and so then we have our result (maybe see Appendix 4 for a concise version of this): I guess you could say it’s the **Fundamental Theorem of Stochastic Calculus**.



Note that we do not have an independent way of evaluating any of the integrals on the RHS if they have any dependence on W. The only thing we ‘know’ is the LHS, which is simply the antiderivative of f (w/r to W). Solving for the integral term we have, in other notation (f = ∂F/∂W):



The last two integrals are ‘normal’ Riemann integrals w/r to time. So we see that the ‘Remainder’ term referenced above is simply those two integrals on the RHS, or really, the last integral. I wonder if there is reason to do line, surface, or volume integrals in this context. Let’s take a look at the special cases of the Ito (λ = 0) and Stratonovich (λ = ½) integrals. These would be (they have their own special notation):



And note these are related via:



We can connect them directly, via the following heuristic manipulations (see Appendix):



Let’s do some quick examples.

**Example 1**

Let’s do the one we started out with:



**Example 2**

Or how about just a function of t? Consider,



Let’s call it F(t). Is it just a weighted sum of Gaussians and is therefore a Gaussian itself? We can find the mean and variance,



What are higher moments?



Let’s look at:



Using,



and path integral integration – see folder, we can see that this is just:



So,



which is exactly how a normally distributed variable would behave, I.e., 3σ2. So isn’t it the case that:



where N(0,1) is a unit normal variable? Yes I think so. So,



I’m pretty sure our boxed formula is correct. Note the use of ~, rather than =, which means that the two have the same probability distribution. But not that they are identical per se´. This is in the same sense that two random variables X and Y can have the same probability distribution but be independent variables, or identical (completely dependent) variables, or perhaps just somewhat dependent too.

Does this mean integration is not a linear operation?

**Differential functions of Wiener process**

Now let’s go backwards, sort of. Consider a function F(W(t)). And let’s consider what its derivative might mean. We’ll define the derivative as:



Tnen we’ll expand the numerator about some point within the interval:



Making those replacements of WW´ → <WW´> is really only justified within the context of integration. But…and then dividing both sides by tk+1 – tk, we have:



So then our formula is:



So the λ would govern where, in between the (t,t+dt) interval, the partial derivatives are evaluated. The Ito version evaluates at the left end point, for which λ = 0, and the Stratonovich version evaluates at the midpoint, for which λ = ½, and the last term vanishes. Product Rule works like this:



We can do similarly for the chain rule. So we have:



**Integration by Parts**

Consider product rule,



Multiplying by dt, we have:



I think the last term is called the ‘quadratic variation’. If we integrate both sides, then we have:



And solving for one of the integrals, I think we arrive at our IBP formula, basically.



**Example**

So for example, let’s apply it to:



If we let G = W(t), and dF = f´(t)dt, then we can write:



So,



Using our result above for such integrals as we have on the right, we can conclude:



where N(0,1) is a unit normal distribution and the ~ indicates equivalence of probability distributions of the two expressions. But this is only somewhat useful. It doesn’t tell us how to combine the two terms. Note we cannot say in the first term that W(t) = √(Dt)N(0,1) and then add the two N(0,1) according to the familiar aN(0,1) + bN(0,1) = √(a2 + b2)N(0,1) because this presumes the two N(0,1) are independent. But our N(0,1) are not independent as they both depend on W(t) in some sense. Alas, nor does it follow that we can say N(0,1) = W(t)/√(Dt) in the second term and then add the two W(t) together according to aW(t) + bW(t) = (a + b)W(t), because N(0,1) is not identically equal to W(t)/√(Dt) per se´: they are two different unit normal variables (I guess not completely independent, but not completely identical either). So to put this in a form where we can combine them, we can take advantage of the linearity of the integral to write:



So we have:



**Example**

Let’s do,



So using our boxed formula:



we have:



**Example**

Let’s do:



Well, using our boxed formula,



we have:



**Example 3**

Let’s try,



We can use,



to simplify this to:



Then use,



with f´(t) = 1, to say



Not sure if this can be simplified any. We can combine the terms in the bracket into 1 term,



and evaluate it as before, but that won’t allow us to combine it with the W3(t) term. So not sure what to do at this point.

**Appendix 1**

Next question is…is this a legitimate replacement?



We have to check,



Now what about the i ≠ j terms?



This requires:



So it works, and with the expected value of c. So, looking ahead, it seems we can write:



\* only the overlap between the two random variables matters, and this is why we can change Wi+1 – Wi to Wi+λ – Wi.

**Appendix 2**

Now let’s consider something else. I’ll say 0 < λ < λ´ < 1. Is the following true?



Well,



Continuing,



So we must check,



If the off diagonal terms go to zero, then the whole thing will as well. So we just need,



Interesting. So,



Apparently, only the overlapping ΔW’s matter, and so the part from λ to λ´ is irrelevant. So I’d reckon that we can say:



This probably should be enshrined.

**Appendix 3**

Now let’s consider something else again, where 0 < λ < λ´ < 1. What can we do with this?



Well,



And so we see how our highlighted result makes this manipulation into an average-able form, quite easy.

**Appendix 4**

For short, could say,



So yeah,