Упрощенная трактовка понятия стохастического дифференциала и формулы Ито
(редакция 02 июля 2012 г.)

В.Н. Тутубалин

Аннотация. Этот учебный материал предназначен для студентов специальности «механика», изучающих курс «Математическая статистика и случайные процессы», а также для студентов, изуча.щих курсы естественно-научного содержания «Финансовая статистика» и «Приложения вероятностных методов» (лектор В.Н. Тутубалин)

1.Общее введение.

В стохастическом анализе стохастические дифференциальные уравнения типа $dx(t) = a(t,x)dt + \sigma(t,x)dW(t)$, где $W(t)$ - винеровский процесс, получают строгую математическую трактовку путем переписывания их в интегральном виде и доказательства того, что решения полученных интегральных уравнений могут быть найдены методом последовательных приближений (если считать, конечно, что траектории винеровского процесса уже даны). Разумеется, при этом нужно наложить какие-то условия на коэффициент сноса $a(t,x)$ и квадратный корень из коэффициента диффузии $\sigma(t,x)$. (Ведь теорема существования и единственности для обычных уравнений, не содержащих винеровского процесса, тоже доказывается при некоторых условиях на правую часть уравнения.) Однако в приложениях трудно представить такую ситуацию, когда интересующее нас явление в любом масштабе по времени описывается траекторией диффузионного процесса. Проще всего пояснить это на примере из области финансовой математики. Известно, что одной из моделей динамики рыночных цен финансовых активов является так называемое *геометрическое броуновское движение.* Это означает, что логарифм цены совершает броуновское движение с некоторыми постоянными (т.е. не зависящими от t и x) коэффициентами сноса и диффузии. Эта модель довольно близка к истине, если рассматривать изменения цен активов за сравнительно большие промежутки времени – порядка одних или нескольких суток. Но в настоящее время нетрудно получить таблицы цен активов с шагом по времени одна секунда. Можно определить цену актива как функцию непрерывного времени, положив, что его цена в момент t равна цене ближайшей предшествующей сделки. Поскольку сделки с каждым активом совершаются, грубо говоря, раз в минуту, то при разрешении по времени в одну секунду траектории цен будут кусочно постоянными и не будут иметь ничего общего с винеровской траекторией.

(Собственно говоря, мы имеем дело с обычным приемом математической физики, когда дискретная последовательность – в данном случае, последовательность цен актива с шагом в одну секунду – представляется в виде процесса с непрерывным временем, в расчете на то, что формульные операции выглядят проще и изящней для непрерывного времени.)

Если же говорить о физических явлениях, например, о реальном броуновском движении, то, опять-таки, при большом шаге по времени картина весьма похожа на винеровский процесс (умноженный на некоторую размерную константу), но при малом шаге эта модель слишком груба. В частности, более точной моделью броуновского движения является процесс Орнштейна-Уленбека (в этой модели броуновская частица имеет скорость). При моделировании физических явлений вообще дело обычно начинается с некоторой марковской цепи, которая после замены переменных делается похожей на диффузионный процесс – но не в смысле траекторий движения чего-либо, а лишь в смысле распределений вероятностей. Приведем простейшие примеры таких моделей.

2.Простейшие примеры диффузионных моделей.

2.1. Движение по компасу.

Хорошо погуляв в лесу, мы собираемся вернуться домой. Мы знаем, что на расстоянии L=10км в направлении с севера на юг проходит шоссе с частым автобусным движением, причем остановка автобуса находится точно в направлении от нас на географический восток. Как бы, имея только школьный компас, выйти возможно точнее прямо на эту остановку и с какой точностью вообще это возможно?

Если изобразить обычную прямоугольную систему координат на плоскости, то в настоящий момент мы находимся в точке (0,0), ось абсцисс смотрит прямо на восток, а ось ординат – на север. Двигаться нужно в идеале по оси абсцисс, но реально это не получится: появятся ненулевые значения ординаты. Модель движения по компасу состоит в том, что время от времени мы останавливаемся, замечаем какой-то ориентир в нужном (восточном) направлении, идем к этому ориентиру, а оказавшись вблизи него, замечаем следующий ориентир.

(Эта модель уже является некоторым огрублением реальности, поскольку не учитывает необходимости обходить лужи или поваленные деревья. Таким образом, описывается не точная траектория движения, а лишь результат – на каком расстоянии от автобусной остановки мы выйдем на шоссе. Или как нарастает отклонение от чисто восточного напрвления в процессе движения.)

Таким образом, время n (номер ориентира) принимает дискретные значение 0 (начальный момент), 1, 2,… .Мы рассматриваем значения $y(n)$ ординаты нашего положения в эти моменты, причем $y(0)=0$. Напишем такое уравнение модели: $\Delta y(n) = y(n+1) - y(n) = l_n \sin\varphi_n \approx l_n\varphi_n$, в котором l_n – расстояние до ориентира, замеченного в момент n, а φ_n – ошибка в выборе направления на ориентир (т.е. угол, который делает на координатной плоскости фактическое направление перехода за время от n до n+1 с осью абсцисс: в идеале этот угол равен нулю, поскольку ось абсцисс смотрит прямо на восток). Ввиду малости ошибки ориентирования мы заменили её синус на её саму (измеренную в радианах). Теперь предстоит разобраться с вероятностными предположениями модели.

Основной смысл предположения случайности об ошибках чего-нибудь состоит в том, что случайные ошибки с нулевым математическим ожиданием взаимно компенсируются. Следовательно, их суммарное влияние будет нарастать медленнее, чем в том случае, если бы ошибки были систематическими. В оценке этого нарастания для случайных ошибок и заключается основной смысл вероятностных расчетов. Но такой компенсации нечего ожидать от величин l_n – расстояний между соседними ориентирами. Сами же статистические свойства этих величин зависят от условий видимости в лесу, так что в разных участках нашего предполагаемого пути могут быть совершенно разными. (Нарушение статистической однородности.) Кроме того, без специальных натурных исследований эти статистические свойства нельзя определить (а тем более нельзя узнать свойства вероятностной зависимости между величинами l_n и φ_n), Поэтому лучше предположить величины l_n неслучайными и равными одной и той же величине, скажем l_n=l=100метров.

Наоборот, ошибки ориентирования φ_n могут принимать положительные и отрицательные значения, а стало быть, могут компенсировать друг друга. Особое внимание надо обратить на то, чтобы добиться нулевого математического ожидания этих ошибок (в частности, нужно знать и учитывать магнитное склонение). Можно оценить по порядку величины и стандартное отклонение этих ошибок σ_φ. Действительно, цена деления школьного компаса составляет 3^0 и можно ошибиться на одно-два деления. Следовательно, принимая для стандартной ошибки σ_φ два деления, получаем 6^0, т.е. примерно 1/10 радиана. При разных n величины φ_n естественно считать независимыми.

В модели с постоянными $l_n=l=100$ метров приращение абсциссы $\Delta x(n) = x(n+1) - x(n) = l\cos\varphi_n \approx l.$ Следовательно, всего придется сделать примерно $N=L/l=100$ переходов для достижения шоссе в какой-то точке. Расстояние от автобусной остановки до этой точки есть $y(N) = l(\varphi_0 + \varphi_1 + ... + \varphi_{N-1}).$ В силу центральной предельной теоремы, это примерно нормальная случайная величина с нулевым математическим ожиданием и стандартным отклонением вида $l\sqrt{N}\sigma_\varphi = 100*10*1/10 = 100$ метров. Если же магнитное склонение (по порядку величины тоже около 1/10 радиана) не учесть, то это даст систематическую ошибку L*1/10=1 километр. У нас получился довольно оптимистический вывод относительно точности движения по компасу (стандартное отклонение 100 метров на переходе в 10 км). Спрошенные геологи не считают это чем-то невозможным.

Данная задача (оценка точности выхода в заданную точку линейного ориентира) удобнее всего решается в конечно-разностной модели. Посмотрим, однако, как эту задачу можно было бы уложить в рамки *теоремы перехода А.Н.Колмогорова* (см. [1], стр. 259 – 262). С точностью до постоянного множителя l речь идет о сумме $S_N = \varphi_0 + \varphi_1 + ... + \varphi_{N-1}$. Следует рассмотреть последовательность нормированных сумм вида $s(0) = 0, s(t_k) = \frac{1}{\sqrt{N}}(\xi_1 + \xi_2 + ... + \xi_k)$, где $t_k = k/N, \xi_i = \varphi_{i-1}/\sigma_\varphi$. Последние суммы удовлетворяют теореме Колмогорова, и их распределения вероятностей могут быть аппроксимированы распределениями винеровского процесса на отрезке [0,1]. В частности, распределение $s(1)$ аппроксимируется распределением $W(1)$, т.е. стандартным нормальным распределением. Возвращая обратно исключенные постоянные множители, мы получим, что $y(n)$ аппроксимируется распределением величины $\sigma_\varphi l\sqrt{N}W(n/N)$, которое совпадает с распределением величины $\sigma_\varphi lW(n)$. Это напоминает аппроксимацию $y(n)$ траекторией винеровского процесса, умноженного на константу. Последовательность $y(n)$ определена лишь при целых неотрицательных n. Но можно её определить при вещественных неотрицательных t, например, с помощью линейной интерполяции. Получится кусочно- линейная функция, которая в малом ничего не имеет общего с траекторией винеровского процесса (умноженного на постоянный множитель). Но при достаточно больших n их распределения вероятностей близки.

2.2.Движение с помощью прибора инерциальной навигации.

Рассмотрим ту же задачу выхода в заданную точку на линейном ориентире, но для движущейся машины, сделанной, в основном, из стали. Снабжать такую машину магнитным компасом бесполезно из-за неизбежной девиации компаса, но можно представить себе инерциальный компас в виде прибора, способного сохранять (с той или иной точностью) раз установленное направление (скажем, на географический север) независимо от движения машины. Однако постепенно набегает ошибка этого прибора – как сумма (допустим, независимых) случайных величин, связанных с перегрузками при движении машины по различным участкам пути. Такая ошибка аппроксимируется, как было показано выше, в смысле распределений вероятностей (но не в смысле траекторий) процессом вида $\varphi(t) = \sigma W(t)$.

(Это при отсутствии систематических ошибок; параметр σ нужно узнавать из испытаний: теперь цена деления шкалы школьного компаса не поможет.)

Нарастание отклонения положения машины от заданной траектории (т.е. от движения прямо на восток) будет описываться моделью $dy(t) = v(t)\sin\varphi(t)dt \approx v(t)\varphi(t)dt = \sigma v(t)W(t)dt,$ где $v(t)$ –

скорость машины в момент t. Следовательно, отклонение $y(T)$ от расчетной траектории в момент $t=T$ выражается интегралом

$y(T)=\sigma\int\limits_0^T v(t)W(t)dt.$ В случае постоянной скорости $v(t)=v$ дисперсия последнего выражения имеет вид

$$\mathbf{D}\mathbf{y}(T)=\sigma^2 v^2\int_0^T\int_0^T \mathbf{E}(W(s)W(t))dsdt=\sigma^2 v^2\int_0^T\int_0^T \min(s,t)dsdt=\frac{2}{3}\sigma^2 v^2 T^3.$$

Таким образом, с ростом времени движения T дисперсия отклонения растет со скоростью T^3, а следовательно, само отклонение – со скоростью $T^{3/2}$. В случае магнитного компаса порядок роста отклонения (от расчетной траектории) есть $T^{1/2}$. Таковы оценки в предположении отсутствия систематических ошибок. При наличии постоянных во времени систематических ошибок для магнитного компаса отклонение растет, как T, а для инерциального (при его систематическом уходе) как T^2.

Сейчас непрерывный вариант модели более изящен, чем дискретный, однако это изящество достигнуто ценой введения неизвестного параметра σ. Однако качественный вывод о быстром нарастании отклонения от расчетной траектории при использовании инерциальной навигации заслуживает внимания.

2.3.Математический маятник под воздействием случайных толчков.

За счет перехода к безразмерным переменным уравнение движения математического маятника можно привести к следующему виду: $d^2x/dt^2+x=0$. Выражение $r^2=x^2+\dot{x}^2$ в силу этого уравнения имеет нулевую производную по времени, так что на фазовой плоскости $(x,\dot{x})$ движение маятника представляется бесконечным равномерным вращением по окружности (по часовой стрелке). Теперь представим себе, что в лабораторию, где качается маятник, проникли мальчишки и принялись стрелять в этот маятник горошинами. (Двое мальчишек расположились в плоскости качания маятника по разные стороны от него.) Большей частью они промахиваются, но иногда попадают. Удар горошины будем моделировать таким образом, что положение маятника не изменяется, а скорость меняется скачком. Тогда движение маятника представится на фазовой плоскости как вращение по окружностям с редкими перескоками с одной окружности на другую. Какой элемент этой картины может быть аппроксимирован диффузионным процессом? (Сами траектории движения на фазовой плоскости на диффузионный процесс никак не похожи.)

Предполагая удар горошины малым вмешательством, получим следующую модель. Если невозмущенное движение маятника описывалось до удара горошиной формулами $x(t)=r\cos t, \dot{x}(t)=-r\sin t,$ то в момент удара скорость меняется на $\{-r\sin t+\delta\}$. В таком случае величина $r^2=x^2+\dot{x}^2$ получает приращение $\Delta r^2=-2r\delta\sin t+\delta^2$. Если, имея в виду диффузионное приближение, считать случайную величину δ малой, то малым оказывается и последнее приращение. Речь может идти, следовательно, о каких-то предположениях, которые позволили бы описать изменение величин r^2 диффузионным процессом. Следует понимать, что при построении моделей математической физики уровень строгости обычно не может быть математическим. Математик назовет такие рассуждения скорее «диалектикой», чем физикой или математикой. (Напомним, что словарь Даля определяет диалектику как «искусство убедительного пустословия».) Но в физике иначе нельзя: физические модели обосновываются некоей комбинацией математики, диалектики и эксперимента.

Первое, что надо сделать, - это представить эволюцию величины r^2 во времени в виде марковской цепи. Уговоримся наблюдать эту величину, с одной стороны, достаточно редко (через несколько периодов качания маятника) и притом в случайные моменты времени, чтобы маятник ничего не помнил, кроме величины r^2. С другой стороны, будем наблюдать все же достаточно часто, чтобы изменения r^2 в промежутки времени между соседними наблюдениями оказались малыми. Обеспечив (на уровне диалектики) марковскую зависимость, перейдем к вычислению коэффициентов сноса и диффузии. Предположим, что момент удара горошины настолько случаен, что величина $\sin t$ имеет такое же распределение вероятностей, как если бы t имело равномерное распределение на отрезке $[0, 2\pi]$. Далее, распределение вероятностей величины δ (эффект удара горошиной) вполне могло бы зависеть от положения и скорости маятника. Но мы предположим, что скорость горошин велика в сравнении со скоростью маятника (а масса мала) и вероятность попадания не зависит от положения и скорости маятника. Тогда окажется, что случайная величина δ не зависит от случайной величины t (момент удара). В этих предположениях вычисляем коэффициент сноса $\mathbf{E}\Delta r^2 = \mathbf{E}\{-2r\delta\sin t + \delta^2\} = -2r\mathbf{E}\delta\mathbf{E}\sin t + \mathbf{E}\delta^2 = \sigma^2$, поскольку при наших предположениях $\mathbf{E}\sin t = 0$ (использовано обозначение $\mathbf{E}\delta^2 = \sigma^2$).

(Интересно, что ответ от $\mathbf{E}\delta$ не зависит, в частности, неравенство нулю этого математического ожидания не приводит к систематическим эффектам, что объясняется случайностью $\sin t$.)

Далее, вычисляем коэффициент диффузии, удерживая в получаемых формулах лишь члены порядка δ^2 и пренебрегая членами высшего порядка (как это принято в технике диффузионных процессов). Имеем $\mathbf{E}(\Delta r^2)^2 \approx \mathbf{E}\{4r^2\delta^2\sin^2 t\} = 2r^2\sigma^2$, поскольку при наших предположениях $\mathbf{E}\sin^2 t = 1/2$. Это мы вычислили коэффициенты сноса и диффузии за один переход в марковской цепи, которой пытаемся описать динамику r^2. Если ввести для этой цепи «медленное время» τ так, чтобы $\Delta\tau = \sigma^2$, и обозначить $z = r^2$, то для $z(\tau)$ получится диффузионный процесс с коэффициентом сноса 1 и коэффициентом диффузии $2z$.

Умозрительная проверка сделанных предположений невозможна. Если бы мы занимались настоящей физикой, то надо было бы найти более интересное физическое явление, чем стрельба горошинами в маятник. Дальше должен был бы последовать эксперимент с наблюдением траекторий. Рассчитывать параметр σ, исходя из какого-то моделирования физического явления, вряд ли получится. Содержательный вопрос состоял бы в том, можно ли описать какие-то вероятностные свойства наблюдаемых траекторий моделью диффузионного процесса при каком-нибудь выборе этого единственного параметра. Для этого нужно уметь рассчитывать распределения вероятностей диффузионного процесса, к чему мы и переходим.

3. Упрощенный подход к понятию стохастического дифференциала.

Уговоримся понимать запись $dx(t) = a(t,x)dt + \sigma(t,x)dW(t)$ в следующем смысле. Распределения вероятностей случайного процесса $x(t)$ являются предельными при $\Delta t \to 0$ для последовательности марковских цепей, задаваемых конечно-разностными уравнениями вида $\Delta x(t) = a(t,x)\Delta t + \sigma(t,x)\Delta W(t)$, где под $\Delta W(t)$ понимается нормальная случайная величина с нулевым средним и дисперсией Δt, которую можно считать приращением винеровского процесса. Впрочем, в силу теоремы Колмогорова о предельном переходе, в качестве $\Delta W(t)$ можно употреблять не обязательно нормальные, а и любые другие случайные величины с указанными средними и дисперсиями. Математически строгим этот подход будет в том случае, если мы сумеем доказать, что дифференциальное уравнение Колмогорова на левом конце с заданными коэффициентами сноса и диффузии имеет гладкое решение (при гладком начальном условии). Это в данной концепции предлагается делать средствами дифференциальных уравнений. Например, можно предъявить явное решение уравнения Колмогорова для винеровского процесса со сносом (постоянные снос и диффузия), для геометрического броуновского движения и для процесса Орнштейна-Уленбека ($a(t,x) = -ax$, диффузия постоянна).

Заметим, что подход стохастического анализа (переход к интегральному уравнению) не требует привлечения средств дифференциальных уравнений и даже может использоваться для доказательства чисто вероятностными средствами желаемых свойств уравнений с частными производными. Не используя траекторий, мы многое проигрываем в математическом смысле, но вряд ли в смысле приложений.

Зато в вычислительном смысле конечно-разностный подход не требует какой-либо затраты мозгов (вместо которых употребляется мощь компьютера). Начиная с какого-либо начального условия $x(t_0) = x_0$ моделируются методом Монте-Карло приращения цепи, путем постепенного прибавления которых создается траектория цепи. И так столько раз, сколько необходимо для приближенного вычисления каких-либо вероятностей, связанных с траекториями. Затем шаг по времени уменьшается и т. д., пока не будет достигнута определенная степень нечувствительности результатов к измельчению шага. В случае многомерного $x(t)$ процедура мало чем отличается. (Теорема Колмогорова применима и в многомерном случае.) Только надо извлекать квадратный корень из матрицы коэффициентов диффузии (если считается заданной именно матрица коэффициентов диффузии).

4. Упрощенный вариант формулы Ито.

Сначала посмотрим, о чем идет речь, на примере обыкновенных дифференциальных уравнений без каких либо случайностей. Пусть имеется дифференциальное уравнение вида $dx(t)/dt = f(t,x),$ которое мы перепишем в виде $dx(t) = f(t,x)dt.$ Пусть также имеется гладкая функция $F(t,x),$ в которую мы вместо *x* подставим решение дифференциального уравнения $x(t)$ и получим $y(t) = F(t,x(t)).$ Никто не дерзнет усомниться в том, что $dy(t) = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial x}dx(t) = \left(\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x}f(t,x)\right)dt.$ Деля последнее уравнение на *dt,* получим дифференциальное уравнение для $y(t).$ Это уравнение вместе с уравнением для $x(t)$ образует систему уравнений, решая которую (если она хорошо решается в смысле теоремы существования и единственности) можно получить $y(t)$ как решение дифференциального уравнения (сначала, конечно, решив уравнение для $x(t)$, поскольку эта функция подставляется в частные производные функции *F*).

Оказывается, что нечто похожее можно сделать и в случае стохастических дифференциальных уравнений. Сценарий следующий. Пусть процесс $x(t)$ удовлетворяет стохастическому уравнению $dx(t) = a(t,x)dt + \sigma(t,x)dW(t)$, а $y(t) = F(t,x(t)).$ Тогда для последнего процесса можно получить такое стохастическое дифференциальное уравнение, что решая его вместе с уравнением для $x(t)$, мы будем получать траектории $y(t)$. Понять, в чем дело, несравненно проще в терминах разностных уравнений, т.е. для распределений вероятностей, а не для траекторий.

Действительно, выпишем приращение

$$\Delta y(t) = y(t+\Delta t) - y(t) = F(t+\Delta t, x(t)+\Delta x(t)) - F(t,x(t)) = \frac{\partial F}{\partial t}\Delta t + \frac{\partial F}{\partial x}\Delta x(t) + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(\Delta x(t))^2 + R,$$

где *R* обозначает остаточный член, выписанный для того, чтобы формула Тейлора была точной. В этот остаточный член входит произведение $\Delta t \Delta x,$ и произведения степеней этих величин, поэтому при вычислении $\mathbf{E}R$ получатся члены порядка $o(\Delta t).$ Следовательно, при вычислении коэффициентов сноса и диффузии для $\Delta y(t)$ остаток даст нулевой вклад в порядке $\Delta t.$ Остается разобрать, какой вклад дает $\mathbf{E}(\Delta x(t))^2.$ Поскольку в конечно-разностной схеме имеем $\Delta x(t) = a(t,x)\Delta t + \sigma(t,x)\Delta W(t),$ вклад порядка Δt в $\mathbf{E}(\Delta x(t))^2$ получится лишь от члена $(\sigma(t,x)\Delta W(t))^2.$ Этот вклад равен $\sigma^2(t,x)\Delta t,$ поскольку $\mathbf{E}(\Delta W(t))^2 = \Delta t.$

(*Замечание.* Иногда встречается утверждение, что $(\Delta W(t))^2 = \Delta t + o(\Delta t)$. Это грубо неверно, поскольку разница между случайной величиной $(\Delta W(t))^2$ и ее математическим ожиданием Δt имеет порядок величины Δt.)

Таким образом, мы получаем, что $\Delta y(t)$ может быть приближенно (но с точностью, достаточной для вычисления коэффициентов сноса и диффузии) выражено в виде $\Delta y(t) \approx \left(\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F}{\partial x^2} \right)\Delta t + \frac{\partial F}{\partial x}\Delta x(t)$. (Частные производные функции F берутся в точке $(t, x(t))$). Это можно записать также в форме $dy(t) = \left(\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F}{\partial x^2} \right)dt + \frac{\partial F}{\partial x}dx(t)$.

Не возбраняется, конечно, заменить в последней формуле $dx(t)$ на $a(t,x)dt + \sigma(t,x)dW(t)$ и в полученном выражении «привести подобные члены» по правилам школьной алгебры. Тогда получится, что $dy(t)$ есть снова некоторая линейная комбинация dt и $dW(t)$. Выражение для $dy(t)$ в любой форме записи и называется *формулой Ито.*

Равенство для $dy(t)$ является не приближенным, а точным в следующем смысле. Запись приращений случайных процессов в виде стохастических дифференциалов позволяет строить распределения этих процессов как предельные распределения для последовательности марковских цепей. В силу теоремы Колмогорова, при этом требуется, чтобы для цепей правильно вычислялись коэффициенты сноса и диффузии в порядке величины Δt. Запись стохастического дифференциала в вышеуказанной форме (считая эту запись точной) позволяет построить последовательность марковских цепей с правильно вычисленными коэффициентами сноса и диффузии. Следует еще заметить, что процесс $y(t) = F(t, x(t))$, вообще говоря, не является марковским. Марковской является пара $\{x(t), y(t)\}$. Отсюда вытекает необходимость рассмотрения многомерных диффузионных процессов. В частности, таким же образом, как выше, рассматривается преобразование коэффициентов сноса и диффузии для многомерного диффузионного процесса при гладкой замене координат. Оказывается, что набор коэффициентов сноса, часто называемый *вектором* сноса, преобразуется не как вектор (входят вторые производные от функций, задающих преобразование).

Понятно, что использованное выше приближение для $\Delta y(t)$ позволяет правильно рассчитать (в порядке величины Δt) и $\mathbf{E}[\Delta x(t)\Delta y(t)]$. Это позволяет применить теорему Колмогорова к двумерной цепи $\{x(t), y(t)\}$.

Конечно, следует понимать, что обоснование перехода от последовательности марковских цепей к диффузионному процессу в виде предельной теоремы представляет собой некую математическую условность, поскольку остается неизвестным, как скоро в этом предельном переходе (и как точно) достигается близость распределений вероятностей. На самом деле получаемые из диффузионного приближения формулы должны (в идеале) сопоставляться с каким-либо экспериментом.

Литература

1.Тутубалин В.Н. Теория вероятностей и случайных процессов. М., Изд-во МГУ, 1992. (Книга имеется в интернете).