

Об одной вероятностной модели миграций особей, сборе данных и проблеме гетероскедастичности

On stochastic model of migrations, data collection, and problem of heteroscedasticity

Л.В. Недорезов
L.V. Nedorezov

University of Nova Gorica, Vipavska 27A, Nova Gorica SI-5000 Slovenia. E-mail l.v.nedorezov@gmail.com
Университет в г. Нова Горица, Словения.

Ключевые слова: миграции особей, сбор данных, математическая модель, проблема гетероскедастичности.
Key words: migrations of individuals, data collection, mathematical model, problem of heteroscedasticity.

Резюме. В работе рассматривается вероятностная модель с дискретным временем миграции особей в ограниченной части плоскости — из узла в ближайшие узлы целочисленной решётки. Предполагается, что численность особей не меняется, а миграции из одного узла в другой зависят от численностей особей в этих узлах. Предполагается также, что существуют некие границы изменения численности, когда узел становится «привлекательным» для мигрантов; превышение некоторого порога по численности меняет статус узла — он становится «непривлекательным» для мигрантов, и особи стремятся покинуть этот узел. Через определённое число шагов по времени фиксируется численность в некоторых случайно выбранных узлах решётки (по аналогии с методом бросания рамки или рубки модельных деревьев). Изучается зависимость оценок средних значений численности и стандартных отклонений от фиксированной известной плотности и объёма выборки. В частности, даётся обоснование того, что известная проблема гетероскедастичности не может быть принципиально решена при оценке параметров экологических моделей.

Abstract. In paper stochastic model with discrete time of migrations in finite part of plane is considered. It is assumed that migrations can be from every knot of integer lattice to nearest knots only, and these migrations depend on numbers of individuals in the respective knots. Population size is assumed to be constant for every sequence of population size measurements. It is also assumed that there are two limits of local population size in knot when respective knot is attractable for migrants. Intersection by local population size of higher threshold leads to situation when knot becomes intractable for migrant, and all individuals try to leave the respective knot. After a certain number of time steps local population size in randomly selected knots is determined (it looks like «throwing of frame» or «cutting of model trees» but in considering situation it doesn't lead to changing of conditions for population). Dependence of averages and standard deviations with respect to known fixed population density and sample size are analyzed. In particular, it is shown that well-known problem of heteroscedasticity cannot be solved in principle for the situation when ecological model

parameters must be estimated using empirical or experimental time series.

Введение

Одной из наиболее актуальных в математической экологии является проблема оценки параметров модели по эмпирическим данным [Бард, 1979; Воронцов, 1978; Исаев и др., 1984, 2001; Недорезов, Садыкова, 2005; McCallum, 2000; Wood, 2001; Sadykova, Nedorezov, 2013; Nedorezov, 2012, 2013]. Поиск «хороших» оценок параметров моделей важен как для построения научно-обоснованных прогнозов различной заблаговременности, так и для поиска оптимальных методов управления численностью.

Видимо, наиболее популярным методом оценивания параметров является метод наименьших квадратов [Бард, 1979; Большев, Смирнов, 1983; Дрейпер, Смит, 1986, 1987; Лакин, 1990]. Пусть, к примеру, следующее уравнение описывает изменение численности (или плотности) популяции во времени:

$$x_{k+1} = x_k F(\alpha, x_k) \quad (1)$$

В модели (1) x_k — численность популяции в момент времени k , $k = 0, 1, 2, \dots, \alpha$ — вектор (неизвестных) параметров, F — неотрицательная нелинейная функция (коэффициент размножения). Известно большое количество моделей типа (1), которые используются для аппроксимации энтомологических данных [Kostitzin, 1937; Moran, 1950; Мейнард Смит, 1970, 1976; Hassell, 1975; Варли, Градуэлл, Хасселл, 1978; Bellows, 1981; Недорезов, 1986, 2007, 2010; Berryman, Munster-Swendsen, 1994].

Обозначим через x_0 начальное значение численности популяции, через $x = x(k, x_0, \alpha)$ — решение уравнения (1) при заданных значениях начальной

численности и параметров модели α , через x_k^* , $k = 0, 1, 2, \dots, N$ — имеющуюся эмпирическую выборку (временной ряд), где $N + 1$ — объём выборки. В рамках метода наименьших квадратов считается, что наилучшие значения оценок параметров модели получаются тогда, когда следующее выражение

$$\sum_{k=0}^N (x_k^* - x(k, x_0, \alpha))^2 \quad (2)$$

достигает своего глобального минимума. Смысл данного требования таков: во множестве всех возможных последовательностей, которые могут быть сгенерированы с помощью модели (1) при разных значениях начальной численности и параметров, выбирается такая последовательность, которая «ближе» других располагается к имеющейся эмпирической выборке.

Использование выражения (2) на практике встречает определённый ряд трудностей, связанных не только с нахождением минимума (что само по себе является весьма сложной задачей) или обоснованием выбора вида выражения (2) (откуда следует, что нужно складывать квадраты, а, например, не абсолютные величины разностей?). Существует ещё одна важная проблема: малые значения исходной выборки практически не оказывают влияния на оценки параметров.

Действительно, каждая разность в выражении (2), $e_k = x_k^* - x(k, x_0, \alpha)$ — это разность между эмпирическим значением (значением некоторой случайной величины) и теоретическим значением, полученным с помощью модели. Поскольку предполагается, что теоретические значения являются истинными, а отклонения эмпирических значений от теоретических обоснованы, прежде всего, ошибками измерений, то, соответственно, каждая разность в (2) является случайной величиной с нулевым средним. При этом дисперсии этих случайных величин разные. Поэтому в (2) складываются квадраты значений разно-распределённых случайных величин e_k , а для того, чтобы все слагаемые оказывали одинаковое влияние на получаемые оценки параметров модели, случайные величины должны быть одинаково-распределёнными (проблема гетероскедастичности).

Эту проблему пытаются решить введением весов в выражение (2):

$$\sum_{k=0}^N w_k (x_k^* - x(k, x_0, \alpha))^2 \quad (3)$$

где все веса $w_k \geq 0$ — некоторые неотрицательные константы. Таким образом, проблема гетероскедастичности увязывается с проблемой выбора неотри-

цательных весов в выражении (3). Но к настоящему времени никаких реальных критериев выбора весов $\{w_k\}$ не существует. Точнее, существует ни на чём не основанный следующий критерий выбора: «малым значениям нужно давать большие веса, а большим значениям — малые».

Более целесообразным представляется другой подход к выбору весов $\{w_k\}$: значения весов w_k брать обратно пропорциональными дисперсиям величин e_k . В этом случае каждое слагаемое в (3) (без квадратов) будет значением случайной величины с нулевым средним и единичной дисперсией. Это, конечно, не решение проблемы гетероскедастичности, но при этом две основные характеристики — среднее и дисперсия — у слагаемых в (3) будут одинаковыми. Заметим, что при таком выборе веса $\{w_k\}$ в (3) не являются константами, и их значения зависят от соответствующих теоретических значений (что важно учитывать при поиске минимума выражения (3)).

Таким образом, проблема гетероскедастичности отчасти сводится к проблеме анализа зависимости стандартных отклонений эмпирических данных от реальных значений численности (или плотности) популяции. Если эта зависимость носит монотонный характер (в идеале — линейный), то ещё есть определённая надежда на то, что удастся как-то оценить эту зависимость; но, если эта зависимость носит немонотонный характер, то это будет означать, что в рассматриваемом случае проблема гетероскедастичности неразрешима — выбрать немонотонную последовательность весов $\{w_k\}$ для реальных условий практически невозможно.

Модель

Рассмотрим модель, в рамках которой миграционные потоки организованы достаточно просто. А именно, будем предполагать, что перемещения особей (в некоторой ограниченной области плоскости) носят случайный характер, и эти перемещения зависят от соотношений численностей в соседних участках. Суммарная численность считается неизменной (особи в популяции не погибают и не рождаются) в течение всего времени (поэтому истинная плотность популяции каждый раз точно известна), а «методы сбора данных» (в рамках модели) можно вполне отождествить, например, с методом бросания рамки, с оцениванием плотности лесных насекомых в подстилке осенью, или оценкой плотности насекомых при рубке модельных деревьев.

Пусть N — общая численность популяции, относительно которой будем предполагать, что в течение всего времени (проведения «машинных экспериментов» по оценке плотности популяции) её величина

остаётся неизменной. Пусть дана целочисленная прямоугольная решётка Z_{nm}^2 на плоскости R^2 :

$$Z_{nm}^2 = \{(i, j) : 1 \leq i \leq n, 1 \leq j \leq m\}.$$

Будем считать, что численность определена только в узлах решётки Z_{nm}^2 , и обозначим через $x_{ij}(t)$ количество особей в узле $(i, j) \in Z_{nm}^2$ в момент времени t . Будем предполагать, что особи за границы решётки Z_{nm}^2 не мигрируют, и поэтому для всех значений t выполняется соотношение:

$$\sum_{i=1}^n \sum_{j=1}^m x_{ij}(t) = N.$$

В модели предполагается, что время дискретно, $t = 0, 1, 2, \dots$.

Определение. Два узла $(i_1, j_1), (i_2, j_2) \in Z_{nm}^2$ будем называть соседними, если выполнено равенство $|i_1 - i_2| + |j_1 - j_2| = 1$.

В рамках модели будем считать, что миграции особей из узла (i, j) возможны только в соседние узлы. Относительно характера поведения мигрирующих особей будем предполагать, что независимо от значений локальных численностей, в каждом узле всегда имеется относительно небольшая доля особей $\delta = const > 0, \delta < 1$, которые мигрируют в соседние узлы с равными вероятностями (случайные мигранты, поведение которых не зависит от локальных значений численностей). Будем предполагать, что миграции остальных особей (доля которых составляет $1 - \delta$) зависят от значений численностей особей в соседних узлах решётки Z_{nm}^2 .

Все узлы решётки Z_{nm}^2 будем делить на три качественно разные группы. Будем считать, что существуют некие критические значения D_1 и $D_2, D_1 < D_2$, которые выделяют «оптимальный интервал» по численности, и если в некотором узле (i, j) численность $x_{ij}(t)$ такова, что $D_1 \leq x_{ij}(t) \leq D_2$, то имеются только случайные мигранты (численность которых равна $[\delta x_{ij}(t)]$) из данного узла в соседние, где $[\cdot]$ — целая часть числа. Если выполняется неравенство $x_{ij}(t) < D_1$, то все особи стремятся покинуть этот узел. Если $x_{ij}(t) > D_2$, то численность находится вне оптимальной зоны, и в этом случае будем предполагать, что все особи также стремятся покинуть этот узел. В отличие от предыдущего варианта, здесь будем считать, что для каждой особи имеется определённая вероятность того, что она не покидает данный узел.

Обозначим через q_j веса, отражающие уровень привлекательности узла для мигрантов, $q_j = const > 0$. Пусть привлекательность узла, в котором численность ниже порога D_1 , равна единице, $q_1 = 1$. Привлекательность узла q_2 , в котором численность находится в оптимальном интервале, должна быть выше единицы, $q_2 = const > 1$. Будем считать, что привлекательность узла q_3 , в котором численность больше порогового значения $D_2, x_{ij}(t) > D_2$, минимальна, $q_3 < q_1$ (это предположение не всегда выполняется, и можно указать ситуации, когда миграционные потоки направлены именно в сторону с высокой плотностью популяции [Исаев и др., 1984, 2001]).

Обозначим через $q_{ij}(t)$ привлекательность узла (i, j) момент времени t (в зависимости от значения локальной численности $x_{ij}(t)$ эта величина $q_{ij}(t)$ равна q_1, q_2 или q_3). Будем считать, что для каждой особи вероятность мигрировать в какой-либо соседний узел пропорциональна привлекательности этого узла, и обратно пропорциональна сумме всех весов, в которые это перемещение возможно. Таким образом, для случая, когда $x_{ij}(t) < D_1$ и все особи стремятся покинуть этот узел, и когда узел (i, j) не принадлежит границе решётки, будем считать, что вероятность $p_{i-1,j}$ миграции каждой отдельной особи в узел $(i-1, j)$ определяется выражением:

$$p_{i-1,j} = \frac{q_{i-1,j}(t)}{q_{i-1,j}(t) + q_{i+1,j}(t) + q_{ij-1}(t) + q_{ij+1}(t)}. \quad (4)$$

Соответствующий вид имеют остальные три вероятности миграции особей, а также вероятности миграций в случае, когда узел является граничным. Для случая, когда $x_{ij}(t) > D_2$, в силу сделанных выше предположений, вероятность особи остаться на месте и не мигрировать в какой-либо соседний узел будет определяться выражением:

$$p_{ij} = \frac{q_{ij}(t)}{q_{ij}(t) + q_{i-1,j}(t) + q_{i+1,j}(t) + q_{ij-1}(t) + q_{ij+1}(t)}. \quad (5)$$

Заметим, что величина миграционного потока, определяемого формулами (4) и (5), монотонно растёт с ростом коэффициента привлекательности. Заметим также, что в (5) величина $q_{ij}(t) = q_2$.

Начальное состояние популяции моделировалось следующим образом: каждая особь с равными вероятностями могла оказаться в одном из узлов решётки Z_{nm}^2 . После этого «запускался» процесс перемещения особей (в соответствии с формулами (4) и (5)), и в течении T (при расчётах принималось $T = 1000$) шагов по времени модель работала «вхолостую»; работа «вхолостую» нужна для того, чтобы

установилось некое стационарное состояние процесса, или, точнее, состояние, определяемое именно законами перемещения особей по области. После этого происходил «сбор данных»: в ста случайно выбранных различных точках фиксировалась численность популяции (иными словами, сто раз в поле случайно «бросалась рамка»). После этого модель снова работала вхолостую T шагов перед следующим «сбором данных». Эта процедура повторялась 100 раз для каждого выбранного значения численности N .

Для полученных таким образом данных рассматривались различные задачи и, в частности, исследовалась зависимость стандартного отклонения s_μ (квадратного корня из выборочной дисперсии) значений относительно истинного значения плотности $\mu = N/mn$ при различных значениях объёмов выборок (различного числа «бросаний рамки»):

$$s_\mu = \left(\frac{1}{K-1} \sum_{k=1}^K (x_k^{**} - \mu)^2 \right)^{0.5},$$

где x_k^{**} — элементы выборки (средние значения для всех независимых выборок длины M), K — общее количество выборок соответствующей длины. При $m = n = 100$ и $M = 3$ $K = 3300$, $M = 4$ — $K = 2500$, $M = 5$ — $K = 2000$, ..., $M = 10$ — $K = 1000$. Таким образом, объёмы выборок достаточно велики для получения «хороших» оценок стандартных отклонений s_μ для различных значений M .

Для полученных данных также рассматривался вопрос о зависимости частоты ошибок от плотности популяции при использовании стандартных статистических процедур. Под ошибкой понималась такая ситуация, когда истинное значение плотности μ лежит вне границ 95 % доверительного интервала. Кроме этого, для $M = 6$ и $M = 10$ проверялись «на нормальность» 100 независимых выборок для малых значений плотности популяции.

Результаты расчётов

При следующих значениях параметров модели: $m = n = 100$, $D_1 = 10$, $D_2 = 30$, $\delta = 0,1$, $q_1 = 1$, $q_2 = 10$, $q_3 = 0,2$ — наблюдалась «типичная» картина изменения стандартного отклонения в зависимости от плотности популяции («типичная» в том смысле, что подобный характер изменения наблюдался в модели и при различных других значениях параметров). При достаточно низких значениях плотности популяции (от 0,1 до 1,5) наблюдается практически линейный рост стандартного отклонения (результаты расчётов по модели для различных объёмов выборки $M = 3, 4, \dots, 10$ представлены на рис. 1; резуль-

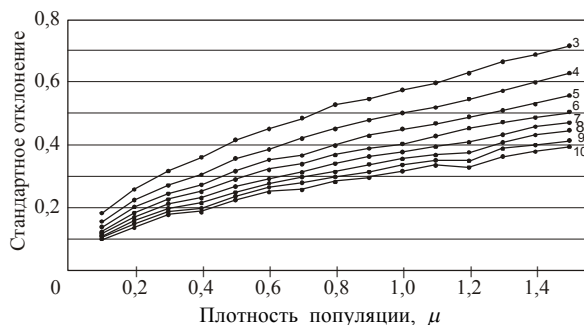


Рис. 1. Изменения значений стандартного отклонения в зависимости от плотности популяции μ ($0,1 \leq \mu \leq 1,5$) и объёма выборки $M = 3, 4, \dots, 10$.

Fig. 1. Changing of values of standard deviation on population density μ ($0,1 \leq \mu \leq 1,5$) and sample size $M = 3, 4, \dots, 10$.

таты аппроксимации данных линейными функциями представлены в таблице 1; y — значение стандартного отклонения).

На интервале $[1,5, 1,6]$ происходит переход через некоторое пороговое значение, после которого наблюдается не только резкое увеличение значений стандартного отклонения, но меняется и характер зависимости s_μ от плотности популяции (рис. 2). Достаточно хорошая аппроксимация данных, представленных на рис. 2, получается при использовании парабол (табл. 2). Во всех случаях коэффициент R^2 близок к единице. Заметим, что максимум стандартного отклонения наблюдается примерно на середине интервала «благоприятных» значений плотности популяции.

Таким образом, можно условно выделить два критических значения на графике изменения стандартного отклонения: это $\mu \approx 1,6$, после которого наблюдается более быстрый рост значений (что вызвано влиянием областей с «благоприятными» значениями плотности), и $\mu \approx 2,0$, после которого наблюдается снижение значений стандартных отклонений. Вопрос можно поставить следующим

Таблица 1. Результаты аппроксимации данных линейными функциями

Table 1. Results of datasets approximations by linear functions

M	Уравнение регрессии	R^2
3	$y = 0,3578\mu + 0,2054$	0,9739
4	$y = 0,3151\mu + 0,1750$	0,9761
5	$y = 0,2788\mu + 0,1576$	0,9740
6	$y = 0,2558\mu + 0,1434$	0,9739
7	$y = 0,2350\mu + 0,1339$	0,9760
8	$y = 0,2227\mu + 0,1228$	0,9771
9	$y = 0,2069\mu + 0,1179$	0,9727
10	$y = 0,1955\mu + 0,1122$	0,9715

Таблица 2. Результаты аппроксимации данных парабололами
 Table 2. Results of datasets approximations by parabolic functions

M	Уравнение регрессии	R^2
3	$y = -0,0290\mu^2 + 1,0878\mu + 1,1604$	0,9638
4	$y = -0,0251\mu^2 + 0,9417\mu + 1,0013$	0,9626
5	$y = -0,0225\mu^2 + 0,8449\mu + 0,8777$	0,9629
6	$y = -0,0205\mu^2 + 0,7671\mu + 0,8435$	0,9599
8	$y = -0,0177\mu^2 + 0,6640\mu + 0,7214$	0,9580
10	$y = -0,0159\mu^2 + 0,5958\mu + 0,6292$	0,9595

образом: можно ли в реальных условиях оценить эти значения? Если да, то тогда возникает принципиальная возможность решения проблемы гетероскедастичности, возможность подбора оптимальных весов для выражения (3). Если же нет, то проблема гетероскедастичности неразрешима.

Учитывая то, что в реальных условиях значительные трудности могут быть связаны с оценкой плотности популяции, а не только с оценкой границ оптимальной зоны (по плотности), то можно утверждать, что в настоящее время проблема гетероскедастичности не может быть принципиально решена. Соответственно, результаты, полученные при введении в (2) и (3) неких (монотонно изменяющихся) весов, не улучшают получаемые таким образом результаты.

По полученным в рамках модели данным оценивались вероятности ошибок, полученных для различных значений объемов выборок M и плотности популяции μ . Под ошибкой понималась такая ситуация, когда истинное (известное) значение плотности популяции μ находилось вне 95 % доверительного интервала, определяемого по стандартной схеме: $\bar{\mu} \pm s_\mu \cdot t(M - 1, 0,05)$, — где $\bar{\mu}$ — оценка среднего, s_μ — стандартное отклонение, $t(M - 1, 0,05)$ — соответствующее значение t-критерия Стьюдента для $M - 1$ степени свободы и уровня значимости 0,05 [Лакин, 1990].

Анализ данных показывает, что при достаточно малых значениях плотности популяции μ оценка вероятности ошибки может превышать 74 % (рис. 3). Это означает, что результаты анализа данных, проведенного с помощью стандартных методов, с большой вероятностью не соответствуют действительности. С ростом плотности популяции наблюдается снижение вероятности ошибки, но при переходе через критическое значение (по плотности популяции) вероятность ошибки начинает возрастать (рис. 4). Увеличение объема выборки M приводит к тому, что максимум вероятности ошибки смещается «влево» — наблюдается при меньших значениях плотно-

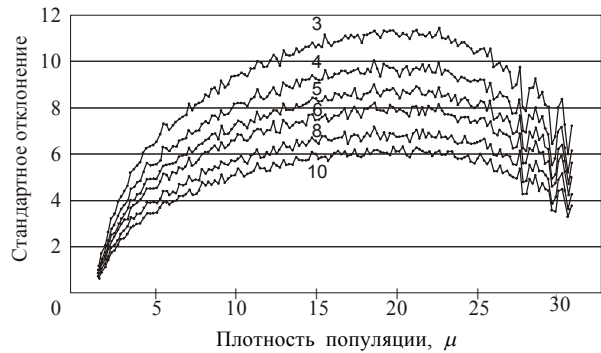


Рис. 2. Изменения значений стандартного отклонения в зависимости от плотности популяции μ ($1,7 \leq \mu \leq 30,9$) и объема выборки $M = 3, 4, 5, 6, 8, 10$.

Fig. 2. Changing of values of standard deviation on population density μ ($1,7 \leq \mu \leq 30,9$) and sample size $M = 3, 4, 5, 6, 8, 10$.

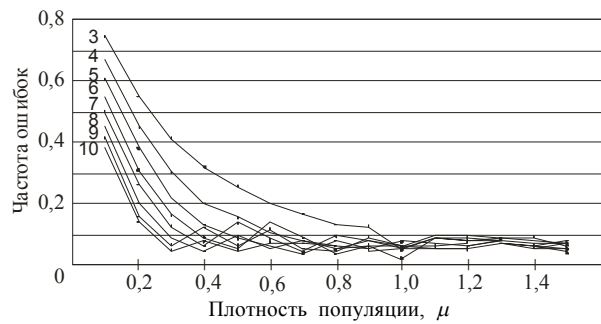


Рис. 3. Изменения значений частоты ошибок в зависимости от плотности популяции μ ($0,1 \leq \mu \leq 1,5$) и объема выборки $M = 3, 4, \dots, 10$.

Fig. 3. Changing of values of frequency of errors on population density μ ($0,1 \leq \mu \leq 1,5$) and sample size $M = 3, 4, \dots, 10$.

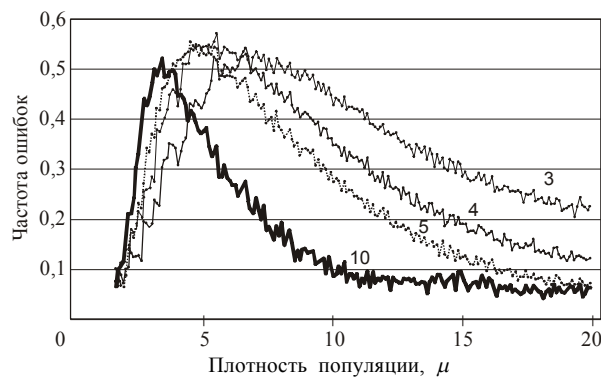


Рис. 4. Изменения значений частоты ошибок в зависимости от плотности популяции μ ($1,7 \leq \mu \leq 20$) и объема выборки $M = 3, 4, 5, 10$.

Fig. 4. Changing of values of frequency of errors on population density μ ($1,7 \leq \mu \leq 20$) and sample size $M = 3, 4, 5, 10$.

сти популяции. Величина вероятности ошибки даже при достаточно большом значении M превышает 50 % (рис. 4). Заметим, что наблюдаемая достаточно парадоксальная ситуация, когда с вероятностью 0,95 истинное значение плотности должно лежать в доверительном интервале, а с вероятностью 0,74 оно лежит вне этого интервала, не связана с какими-либо

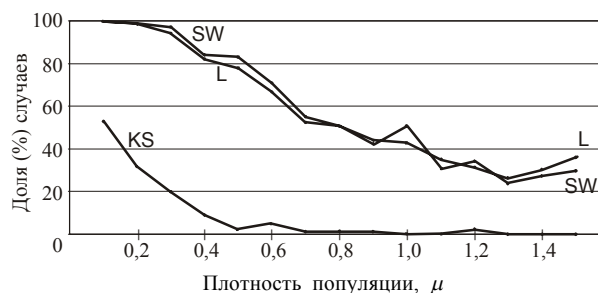


Рис. 5. Доля (в процентах) случаев, когда гипотеза о нормальности выборки с $M = 6$ отклоняется с 5 % уровнем значимости. KS — данные для теста Колмогорова-Смирнова, L — данные для теста Лиллиефорса, SW — данные для теста Шапиро-Вилка.

Fig. 5. Quota (in percentages) of situations when hypothesis about normality of a sample with $M = 6$ must be rejected with 5 % significance level. KS — values obtained with Kolmogorov-Smirnov test, L — values obtained with Lilliefors test, SW — values obtained with Shapiro-Wilk test.

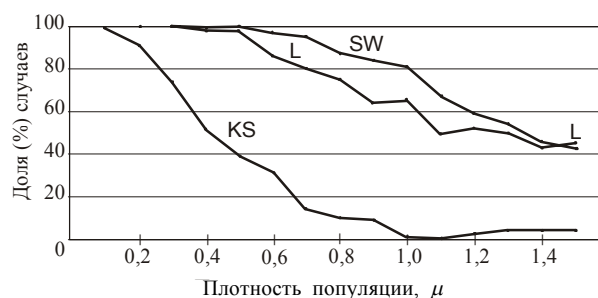


Рис. 6. Доля (в процентах) случаев, когда гипотеза о нормальности выборки с $M = 10$ отклоняется с 5 % уровнем значимости. KS — данные для теста Колмогорова-Смирнова, L — данные для теста Лиллиефорса, SW — данные для теста Шапиро-Вилка.

Fig. 6. Quota (in percentages) of situations when hypothesis about normality of a sample with $M = 10$ must be rejected with 5 % significance level. KS — values obtained with Kolmogorov-Smirnov test, L — values obtained with Lilliefors test, SW — values obtained with Shapiro-Wilk test.

проблемами проведённых машинных экспериментов (например, с возникновением зависимостей в полученных выборках). Для каждого конкретного значения плотности значения случайных величин (в соответствии с которыми моделировались миграционные процессы, а также определялись места фиксации локальных численностей) выбирались независимо.

Возникновение обсуждаемой ситуации может быть связано с частотой фиксации локальных численностей (1000 шагов процесса между моментами «сбора данных») может оказаться недостаточно большим), но, скорее всего, оно связано с определённой неоднородностью состояний узлов решётки по численности — наличия «притягивающих» и «отталкивающих» областей. Поскольку в реальных условиях и те, и другие области существуют, то отсюда можно сделать только один вывод: к анализу рассмотренной ситуации (как и аналогичных ситуаций)

некоторые стандартные статистические методы неприменимы.

Для применения параметрических методов статистики к анализу имеющихся выборок требуется, чтобы исходные выборки были нормально распределены. Применение трёх методов проверки — теста Колмогорова-Смирнова, теста Лиллиефорса, и теста Шапиро-Вилка [Lilliefors, 1967; Shapiro et al., 1968; Большев, Смирнов, 1983] — к 100 независимым выборкам для $M = 6$ и $M = 10$ показывает (рис. 5 и 6), что с ростом плотности популяции наблюдается тенденция к снижению числа случаев, когда гипотеза о нормальности выборки отклоняется с 5 % уровнем значимости. Для обоих рассматриваемых случаев при достаточно малом значении плотности тесты Лиллиефорса, и Шапиро-Вилка позволяют отклонить гипотезу о нормальности в 100 % случаев. Однако и при достаточно большом значении плотности ($\mu \approx 1,5$) эти тесты позволяют отклонить гипотезу более чем в 20 % случаев. Важно также отметить, что все три теста показывают, что увеличение объёма выборки при одних и тех же значениях плотности популяции приводит к увеличению числа случаев, когда гипотеза отклоняется.

Заметим, что, когда гипотеза о нормальности не отклоняется, то это не означает, что мы *должны* принимать эту гипотезу. Это лишь означает, что используемые тесты не позволяют отклонить гипотезу. Таким образом, следует признать, что использование параметрических статистических методов при оценке плотности популяции, когда плотность достаточно низка, не имеет под собой оснований.

Обсуждение

Анализ изменения квадратов отклонений оценок плотности популяции (для различных объёмов выборок) от истинных значений плотности показывает, что изменение этих величин (стандартных отклонений) не только не линейно, но и немонотонно (рис. 1, 2). На практике ситуация является гораздо более сложной, и нет никакой возможности оценить где именно, при каких значениях плотности нарушается монотонность (тем более, что на практике у нас нет истинного значения плотности, а есть только его оценки). В свою очередь, это означает, что нет никакой возможности выбора системы весов (3) при оценке параметров моделей: нет никакой возможности выбрать веса так, чтобы все слагаемые минимизируемого функционала «оказывали» одинаковое влияние на оценки параметров.

Если же исходить только из принципа «меньшим значениям давать большие веса», то с неизбежностью возникнет ситуация (вследствие немонотонности изменения стандартных отклонений), когда боль-

шему значению нужно будет давать больший, а не меньший вес (как это следует из принципа). Соответственно, следование принципу приведёт к непредсказуемому искажению результатов. Таким образом, метод введения весов (выражения (2) и (3)) следует признать несостоятельным, и, кроме этого, следует признать неразрешимость проблемы гетероскедастичности при оценке параметров нелинейных экологических моделей.

Анализ простой модели миграций особей на плоскости также показывает, что применение параметрических методов статистики (для анализа выборок по оценке плотности популяции) не имеет под собой оснований, если истинное значение плотности достаточно мало. В этой ситуации получаемые выборки (с помощью метода «бросания рамки» или «вырезания квадратного метра лесной подстилки осенью») не являются нормальными (рис. 5 и 6).

Литература

- Бард Й. 1979. Нелинейное оценивание параметров. М.: Статистика. 349 с.
- Большев Л.Н., Смирнов Н.В. 1983. Таблицы математической статистики. М.: Наука. 416 с.
- Варли Д.К., Градуэлл Д.Р., Хасселл М.П. 1978. Экология популяции насекомых. М.: Колос. 222 с.
- Воронцов А.И. 1978. Патология леса. М.: Лесная промышленность. 272 с.
- Дрейпер Н., Смит Г. 1986. Прикладной регрессионный анализ. Т.1. М.: Финансы и статистика. 366 с.
- Дрейпер Н., Смит Г. 1987. Прикладной регрессионный анализ. Т.2. М.: Финансы и статистика. 351 с.
- Исаев А.С., Хлебопрос Р.Г., Недорезов Л.В., Киселёв В.В., Кондаков Ю.П. 1984. Динамика численности лесных насекомых. Новосибирск: Наука. 224 с.
- Исаев А.С., Хлебопрос Р.Г., Недорезов Л.В., Киселёв В.В., Кондаков Ю.П., Суховольский В.Г. 2001. Популяционная динамика лесных насекомых. М.: Наука. 347 с.
- Лакин Г.Ф. 1990. Биометрия. М.: Высшая школа. 352 с.
- Мейнард Смит Дж. 1970. Математические идеи в биологии. М.: Мир. 179 с.
- Мейнард Смит Дж. 1976. Модели в экологии. М.: Мир. 184 с.
- Недорезов Л.В. 1986. Моделирование массовых размножений лесных насекомых. Новосибирск: Наука. 125 с.
- Недорезов Л.В. 2007. Влияние качества корма и паразитизма на циклические колебания серой листовенной листовертки // Евразийский энтомологический журнал. Т.6. Вып.2. С.229–244.
- Недорезов Л.В. 2010. Анализ динамики численности сосновой пяденицы с помощью дискретных математических моделей // Математическая биология и биоинформатика. Т.5. No.2. С.114–123.
- Недорезов Л.В., Садыкова Д.Л. 2005. К проблеме выбора математической модели динамики популяции (на примере зелёной дубовой листовёртки) // Евразийский энтомологический журнал. Т.4. Вып.4. С.263–272.
- Bellows T.S. Jr. 1981. The descriptive properties of some models for density dependence // Journal of Animal Ecology. Vol.50. P.139–156.
- Berryman A.A., Munster-Swendsen M. 1994. Simple theoretical models and population predictions // Grasman J., van Straten G. (Eds.): Predictability and nonlinear modeling in natural sciences and economics. N.Y.: Kluwer Academic. P.67–71.
- McCallum H. 2000. Population parameters estimation for ecological models. Brisbane: Blackwell Sciences. 348 p.
- Hassell M.P. 1975. Density-dependence in single-species populations // Journal of Animal Ecology. Vol.44. P.283–295.
- Kostitzin V.A. 1937. La Biologie Mathematique. Paris: A.Colin. 236 p.
- Lilliefors H.W. 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown // Journal of the American Statistical Association. Vol.64. P.399–402.
- Moran P.A.P. 1950. Some remarks on animal population dynamics // Biometrika. Vol.6. No.3. P.250–258.
- Nedorezov L.V. 2012. Gauze' experiments vs. mathematical models // Population Dynamics: Analysis, Modelling, Forecast. Vol.1. No.1. P.47–58.
- Nedorezov L.V. 2013. Dynamics of Pine Beauty (*Panolis flammea*): search for stationary dynamical regimes // Population Dynamics: Analysis, Modelling, Forecast. Vol.2. No.3. P.105–116.
- Sadykova D.L., Nedorezov L.V. 2013. Larch bud moth dynamics: can we explain periodicity of population fluctuations by the time lag dependence in birth rate? // Population Dynamics: Analysis, Modelling, Forecast. Vol.2. No.4. P.154–181.
- Shapiro S.S., Wilk M.B., Chen H.J. 1968. A comparative study of various tests of normality // Journal of the American Statistical Association. Vol.63. P.1343–1372.
- Wood S.N. 2001. Partially specified ecological models // Ecological Monographs. Vol.71. P.1–25.

Поступила в редакцию 13.1.2013