

Руководство по работе с программой "Калькулятор Харди – Вайнберга" (Гены популяции)

Назначение программы.

Эта программа рассчитывает вероятности аллелей одного локуса на основании данных о поголовье (количестве особей, объектов) фенотипов.

Расчет производится на основании закона Харди - Вайнберга.

Расчет производится сразу по трем критериям - по критерию максимума правдоподобия, минимуму суммы квадратов отклонений (разности фактических и расчетных частот фенотипов) и минимуму суммы квадратов нормированных (к дисперсии оценки вероятности фенотипов) отклонений.

Это позволяет оценить устойчивость решения.

Работа пользователя программы, в основном, сводится к формированию списка аллелей, выбору принципа формирования фенотипов и записи данных о поголовье (количестве особей, объектов, далее - поголовье) фенотипов. При правильном занесении этих данных производится расчет выходных данных.

Ограничения.

Количество аллелей ограничено - максимум 50 аллелей.

Количество фенотипов ограничено - максимум 1500 фенотипов.

Количество тестов (антигенов, внешних признаков, наименований анализов и т.п.), на основании которых выделяются (различаются) фенотипы – максимум 50 тестов.

Правила формирования списка аллелей.

Наименование аллеля определяется перечнем тестов (антигенов, внешних признаков, наименований анализов и т.п.), обнаруживаемых (дающих положительную реакцию) при наличии в генотипе соответствующего аллеля.

В наименовании тестов заглавные и строчные буквы считаются разными буквами - DA и Da - это разные тесты.

Тесты в наименовании аллеля разделяются одним или несколькими пробелами.

Внутри названий тестов пробелов не должно быть.

Порядок перечисления тестов не имеет значения. Повторы тестов в наименовании аллеля (занесенные ошибочно) будут устранены автоматически.

Каждый тест может реагировать на несколько аллелей. Не должно быть двух и более аллелей с одинаковым списком тестов.

Количество наименований тестов ограничено - максимум 50 тестов.

Список фенотипов формируется программой автоматически на основании наименований аллелей.

Предусмотрены два способа формирования списка фенотипов - кодоминирование аллелей по отношению ко всем тестам (тесты реагируют на оба гена) и доминирование (тесты реагируют только на доминирующий аллель). В последнем случае перечень аллелей следует записывать в порядке убывания доминирования - каждый аллель списка является доминирующим по отношению к последующим аллелям и рецессивным по отношению к предшествующим.

Запись поголовья фенотипов.

Поголовье фенотипов записывается пользователем в столбце "Факт" в таблице фенотипов. В этот столбец можно заносить целые числа от 0 до 20 000 000. Если в фенотипе нет поголовья, то в этом столбце можно оставить пустое поле (0 можно не заносить).

На этой же форме можно указать критерий, результаты расчета по которому Вы хотите видеть в качестве основных, то есть на этом же окне. Это указание не скажется на полных отчетах о результатах расчета.

Результаты расчета.

Программа рассчитывает частоты (вероятности) аллелей в популяции. На основании этих результатов рассчитываются также расчетные частоты фенотипов, расчетное поголовье фенотипов, расчетные частоты геномов, частоты гетерозигот и гомозигот, фактические и расчетные частоты положительных реакций тестов.

Для оценки соответствия экспериментальных данных (задачи) условиям применимости закона Харди-Вайнберга программа рассчитывает по критерию ХИ квадрат вероятность того, что фактическое распределение поголовья по фенотипам соответствует закону Харди - Вайнберга.

Эта оценка будет статистически значимой, если сумма поголовья (объектов исследования) составляет несколько сотен и более. При меньшем поголовье значение этой вероятности может считаться весьма ориентировочным.

Кроме того, при расчете объединяются фенотипы с числом объектов, меньшим 7. Это делается для того, чтобы исключить чрезмерное влияние маловероятных фенотипов.

Если указанная вероятность меньше 0,05 - соответствие закону Харди - Вайнберга очень плохое (его почти наверняка нет).

Если указанная вероятность меньше 0,1 - соответствие закону Харди - Вайнберга очень маловероятно (его скорее всего нет).

Если указанная вероятность меньше 0,3 - соответствие закону Харди - Вайнберга слабое.

Если указанная вероятность лежит в диапазоне 0,3 -0,95 - соответствие закону Харди - Вайнберга вполне возможно. Экспериментальные данные неплохо (или хорошо) согласуются с законом Харди - Вайнберга.

Если указанная вероятность превышает 0,95, есть основания сомневаться в честности исходных данных - слишком уж все гладко. Это может указывать также на малый объем исходных данных.

Все эти пояснения, конечно, весьма субъективны. Объективно число, а не его трактовка.

Результаты расчетов можно увидеть в соответствующих экранных формах. Результаты расчета можно импортировать в книгу Excel (если он установлен на Вашем компьютере) – есть соответствующая кнопка. В книге Excel Вы имеете возможность редактирования (например, убрать лишние данные) и печати соответствующих таблиц на принтере. Отметим, что для некоторых таблиц предусмотрена книжная, а для других – альбомная ориентация при печати. Проверить можно в предварительном просмотре.

Пояснения к таблицам результатов расчета.

В таблице аллелей приведены частоты (вероятности) аллелей (их наименование в первом столбце) в популяции. Во втором столбце приведены данные расчета частоты по критерию максимума правдоподобия (МП), в третьем – по критерию минимума суммы квадратов (метод наименьших квадратов - МНК) отклонений фактических и расчетных данных о количестве особей в фенотипах, в последнем - по критерию минимума суммы нормированных квадратов (метод нормированных наименьших квадратов - МННК). Результаты расчетов по разным критериям должны быть достаточно близки. Близость результатов характеризует точность и устойчивость результатов расчета.

В качестве финального результата мы рекомендуем брать расчет по МП. Этот критерий более обоснован в часто встречающихся условиях, чем остальные.

Таблица тестов объединяет сведения о фактическом и расчетном количестве (и частотах) “срабатываний” тестов. Результаты также даны для трех критериев расчета. Расчетные и фактические данные должны быть близки, если популяция соответствует требованиям по применению закона Харди-Вайнберга.

Вероятность “срабатываний” теста близкая к 0 или к 1 свидетельствует о его низкой информативности по крайней мере для исследуемой популяции. Возможно, использование таких тестов экономически не целесообразно...

Таблицы генотипов и фенотипов построены по тому же принципу, что и две предыдущие. Они обобщают данные по генотипам и фенотипам. Расчетные и фактические данные по фенотипам должны быть близки, если популяция соответствует требованиям по применению закона Харди-Вайнберга.

В том же окне, где представлены расчеты для фенотипов, даны сведения по расчету статистической вероятности соответствия исходных данных закону Харди-Вайнберга, полученные по критерию Хи квадрат. Указаны значения критерия, количество степеней свободы и собственно отмеченная вероятность.

Повторим, что эти расчеты статистически значимы, только когда число обследованных особей достаточно велико – несколько сотен. Повторим, что **при низкой вероятности полученные**

результаты расчетов нельзя использовать, поскольку исследуемая популяция не соответствует требованиям по применению закона Харди-Вайнберга.

Ошибки ввода исходных данных.

Если при вводе списка аллелей допущены грубые ошибки или Вы забыли указать поголовье фенотипов, то программа выдаст соответствующие сообщения.

Однако могут быть допущены ошибки, не обнаруживаемые программно.

Пользователь должен обнаружить их сам, анализируя списки аллелей, тестов и фенотипов на экране ввода поголовья фенотипов. Если пользователь видит в этих списках нечто, чего не должно быть, или не видит то, что быть должно, то это означает, что допущена ошибка в подготовке списка аллелей или что его задача выходит за рамки возможностей программы.

Для исправления ошибок можно нажать кнопку **КОРРЕКТИРОВАТЬ АЛЛЕЛИ** и произвести исправления.

Основные причины ошибок при записи аллелей:

- 1) В одном и том же тесте ошибочно использованы в одном месте латинские буквы, а в другом - кириллические. А кириллическое и А латинское - разные буквы. Обратите внимание также на а, р, Р, О, о, С, с, Х, х, т, Т, Е, е и не только.
- 2) В одном и том же тесте ошибочно использованы в одном месте заглавные, а в другом - строчные буквы. Это разные буквы.
- 3) Внутри наименования теста ошибочно вставлен пробел.
- 4) Наименования тестов в аллелях ошибочно не разделены пробелом.

При изменении аллелей неизбежно изменится список фенотипов. Соответственно, количество поголовья фенотипов (Факт), указанное Вами ранее, потеряет смысл. Поэтому после изменения аллелей (во избежание ошибок) введенное ранее количество фенотипов **БУДЕТ ОБНУЛЕНО**.

Архив.

Задачи, по которым получены результаты расчетов, могут быть сохранены в архиве. В дальнейшем их можно использовать для повторных расчетов или как шаблон для решения аналогичной задачи, но для другой популяции (придется только ввести соответствующие данные о поголовье). То есть сохраненные задачи можно использовать как шаблон для формирования новых задач.

Исходно в архиве в качестве примера в архив записаны задачи “Лошади1”, “Лошади2”, “Лошади3”. Эти образцовые задачи взяты из статьи “Генетическая структура популяций лошадей рысистых пород по аллелям D-системы групп крови”. Авторы С.П. Князев, Р.М. Дубровская, Н.С. Фадеева, Н.В. Гугорова, Н.В. Данильченко, С.В. Никитин. Статья в журнале “Сельскохозяйственная биология” в № 4 за 2003 год.

Результаты расчета программы достаточно хорошо согласуются с данными статьи. А там, где есть незначительные расхождения, то там результаты, выдаваемые нашей программой более качественные, что не трудно проверить расчетами значения критерия максимума правдоподобия по данным статьи и по данным программы.....

Совет.

Очень советуем для лучшего понимания работы с программой ознакомиться с примером использования программы. Его можно посмотреть, щелкнув по значку “Пример” на рабочем столе.

Об ограничении ответственности.

Автор программы гарантирует, что в программе нет каких-либо структур, предназначенных для злонамеренного искажения результатов расчета или для доставления каких-либо неприятностей пользователю. Программа не выходит самостоятельно в Интернет и никому (кроме пользователя) не передает никакие данные.

Автор программы и поставщик не несут никакой ответственности за ошибочные результаты расчетов (если они обнаружатся), которые могут возникнуть, например, из-за

не выявленных ошибок или недостатков программы или из-за сбоев компьютера или его программного обеспечения или по любой другой причине.

Автор программы и поставщик не несут никакой ответственности за трактовку результатов расчетов и их использование в любом виде.

УДАЧИ И УСПЕХОВ В НАУКЕ И ПО ЖИЗНИ!!!