

## ЛАБОРАТОРНАЯ РАБОТА № 2. ЭТАПЫ РЕШЕНИЯ ЗАДАЧИ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ

При изучении и анализе сложных и многообразных причинно-следственных отношений между объектами и явлениями биотехнологу приходится учитывать целый комплекс внешних и внутренних факторов, от которых в конечном итоге зависят уровень и ход наблюдаемых процессов, те или иные биологические свойства живых организмов, их динамика и разнообразие. При этом зачастую важно оценивать не только роль одного из многочисленных внешних факторов, но и их взаимодействие при совместном влиянии на технологию или организм.

*Идейная база* для изучения действия факторов содержится уже *в методе сравнения двух выборок*. Биотехнологическим содержанием операции сравнения двух выборок, в конце концов, выступает поиск факторов, ответственных за смещение средних арифметических или усиление изменчивости признаков. Развивая это направление биометрического исследования, можно не ограничиваться только двумя «дозами» фактора, но изучить серию ситуаций, в которых фактор проявлял разную силу действия на результирующий признак - от самого «слабого» до самого «сильного». При этом каждому уровню фактора будет соответствовать отдельная выборка, и общая задача получит формулировку «сравнить несколько выборок». *В терминах факториальной биометрии* вопрос о влиянии фактора на признак звучит так: сказывается ли отличие условий получения разных выборок на качестве (значениях) вариант? *В терминах статистики* вопрос звучит несколько иначе: из одной ли генеральной совокупности отобраны все выборки, оценивают ли выборочные средние арифметические одну и ту же генеральную среднюю? Вариантов ответа может быть только два: а) все выборки отобраны из одной генеральной совокупности, условия возникновения вариант одни и те же; б) выборки отобраны из разных генеральных совокупностей, условия возникновения вариант выборок различаются.

В постановке вопроса можно уловить противоречие. Выше было сказано, что по условию задачи выборки формировались в разных условиях, и тут же предполагается, что условия были одинаковые. На самом деле про

тиворечия нет, поскольку речь идет об определении чувствительности признака к действию фактора. Условия формирования выборок могут отличаться, но они могут никак и не сказаться на величине изучаемого признака, не отразиться на значениях вариантов. *Смысл статистического сравнения* в том и состоит, чтобы *оценить эффективность действия фактора на признак, доказать реальность реакции вариант выборок на разные условия их формирования.* В сферу исследования можно вовлекать как один, так и два признака, как количественные, так и качественные характеристики. В каждом случае процедура анализа несколько отличается.

### **Изучение сходства (различий) двух выборок**

В любых биологических экспериментах и наблюдениях особое значение имеют различия, на основании которых судят об эффективности действия тех или иных факторов, например, *по разности между опытной и контрольной группами делают заключение о результатах опыта.* При этом особенно важно оценить статистическую достоверность разности, т.е. определить, можно ли данное различие считать закономерным, характерным для всей генеральной совокупности и рассматривать его как результат действия особенных факторов, или же оно случайно и является следствием недостаточного количества данных и в следующих опытах может не проявиться.

### ***Параметрические критерии оценки различия двух выборок***

Обнаружение достоверных отличий статистических параметров - это первый шаг к познанию новых биологических закономерностей, причем количественно доказанных. Ответ на вопрос о достоверности или случайности отличий дают статистические критерии, среди которых самые распространенные параметрические критерии («*t*-критерий или критерий Стьюдента и «*F*»-критерий» или критерий Фишера). Вычисление их ведется по специальным формулам (различным в зависимости от сравниваемых пара-, метров и типов распределения). Полученные этим способом значения критериев (для чего в формулы подставляются экспериментальные данные); сравнивают с табличными значениями при выбранном уровне значимости/

(обычно  $p = 0,05$ ) и числе степеней свободы (объемы выборок без числа ограничений/ *Результатом* такого *сравнения* должен стать один из двух вариантов следующего *статистического вывода*.

*Вывод 1.* Если полученное значение (величина) критерия больше табличного, значит, различия между параметрами при заданном уровне значимости и установленном числе степеней свободы достоверны, т.е. в разных выборках действительно проявилось действие разных факторов или разных уровней одного фактора.

*Вывод 2.* Если же полученная величина критерия меньше табличной, то при данном уровне значимости и числе степеней свободы различия между параметрами недостоверны. Последнее говорит о том, что различия случайны, никакого определенного вывода о побудительных причинах отличий сделать нельзя, нулевая гипотеза остается не опровергнутой.

При сравнении выборок *по степени выраженности признака* говорят о достоверности (недостоверности) *отличий средних арифметических и долей*, а при сравнении выборок *по уровню изменчивости показателей* - о достоверности (недостоверности) *отличий стандартных отклонений (дисперсий) и коэффициентов вариации*. Особый случай представляет сравнение двух выборок по характеру распределения (достоверность отличия частот), а также общее отличие выборок без указания определенных параметров (для признаков в полуколичественных единицах).

***Сравнение средних арифметических.*** Для выяснения вопроса о случайном или неслучайном расхождении значений некоторого параметра проводят две серии экспериментов и для каждой из них подсчитывают средние значения  $M_1$  и  $M_2$ . Если между ними есть различие, то можно допустить два равновероятных объяснения. *Первое* - предположив, что обнаруженная разность обуславливается случайным варьированием средних величин исследуемой выборки около некоторой общей, хотя и неизвестной нам в точности генеральной средней. Тогда найденные средние должны считаться принадлежащими к одной и той же качественно однородной совокупности. *Второе* - рассматриваемые две средние величины принадлежат к двум качественно отличным друг от друга статистическим совокупностям. Оба эти предположения одинаково законны и применимы ко вся- кой паре сравниваемых величин  $M$ .

Наиболее распространенный способ проведения технологического исследования основан на сравнении результатов полученных в контрольной и опытной группах с применением не парного критерия Стьюдента, который справедлив для больших ( $n > 30$ ) или малых равновеликих выборок ( $n_1 \approx n_2$ ).

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2/n_1 + m_2^2/n_2}} \quad (4.1)$$

где  $M_1, M_2$  - средние величины в контроле и опыте соответственно;  
 $m_1, m_2$  - ошибки среднеквадратического отклонения в контроле и опыте соответственно.

Следует помнить, что разность средних нужно брать по модулю, т.е. без учета знака. Для определения существенности различий необходимо обращаться к таблице Стьюдента (см. Приложение А, табл. А1). По найденному значению  $t$  и числу степеней свободы ( $k = n_1 + n_2 - 2$ ) в таблице Стьюдента находим соответствующее значение вероятности  $P$ . Результатом такого сравнения должен стать один из двух вариантов статистического вывода, о котором уже говорилось выше.

Как уже отмечалось, в биологических экспериментах принято, чтобы показатель надежности был не менее 95 % ( $P > 0,95$  или  $p < 0,05$ ). Если полученное различие меньше, к примеру, 80-90 %, то необходимо высказываться с меньшей категоричностью. Можно указать, что разность может иметь место, но в данной серии экспериментов она не доказана, и необходимо поставить дополнительные опыты.

При сравнении выборочных параметров нормального и биномиального распределений используется одна и та же формула. Например, в процессе специальных исследований было установлено, что у людей пожилого возраста (выборка из 20 человек) до лечения инсулином среднее содержание белков в крови составляло  $M_1 = 81,04 \pm m_1, m_1 = 1,7$ ; а после лечения  $M_2 = 79,33 \pm m_2, m_2 = 1,6$ . Нетрудно видеть, что полученные величины неодинаковы. Но достоверно ли это различие, закономерно ли оно? Можно ли на его основании утверждать, что лечение инсулином понижает содержание белков (в крови)? Ответ На этот вопрос может дать критерий достоверности различий средних арифметических. Согласно общей нулевой гипотезе средние не отличаются. (Проверим ее с помощью критерия; Стью-

дента. Согласно формуле (4.1)  $t_{ci} = 0,7$ . По таблице граничных значений критерия Стьюдента находим, что для уровня значимости  $p = 0,05$  и числа степеней свободы  $k = 20 + 20 - 2 = 38$  величина критерия составляет  $t_{табл}(0,05,38) = 2,03$ . Поскольку  $t_{табл} < t_{табл}$ , нулевая гипотеза сохраняется, различия между средними величинами статистически недостоверны (незначимы), случайны. Следовательно, влияние инсулина на содержание белков в крови не подтверждается и остается недоказанным, возможно, из-за недостаточного числа определений.

Для учета индивидуальных свойств биообъекта (в случае, когда исследование сопряжено с парносвязанными параметрами) предпочтительно использовать парный критерий Стьюдента:

$$\frac{\bar{M}}{m_{12}} \quad (4.2)$$

где  $M_{12}$  - средняя разность величин в контроле и опыте;  $m_{12}$  - ошибка среднеквадратического отклонения разности величин в контроле и опыте;  $k$  — количество пар измерений контроль - опыт.

Таким образом, для повышения воспроизводимости результатов опытов необходимо чтобы единственным различием между контрольной и опытной группой должно быть воздействие исследуемого, и никакого другого, фактора.

*Сравнение долей.* При сравнении достоверности различия долей или процентов признаков (ЯЭ), характеризующихся альтернативным распределением, применяют критерий Фишера с  $\chi^2$ -преобразованием. Суть углового преобразования Фишера состоит в переводе процентных долей в величины центрального угла, который измеряется в радианах. Большей процентной доле будет соответствовать больший угол  $\chi^2$ , а меньшей доле - меньший угол, но соотношения здесь не линейные:

$$\chi^2 = 2 \cdot \arcsin \sqrt{VF}, \quad (4.3)$$

где  $V$ -процентная доля, выраженная в долях единицы.

Вместо процентов берут  $\chi^2$ -значения и подставляют их в формулу:

$$\frac{\chi^2_1 + \chi^2_2}{n_1 + n_2} \quad (4.4)$$

где  $\chi^2_1$  и  $\chi^2_2$  - преобразованные доли;  $n_1$  и  $n_2$  - объемы выборок.

Полученное значение сравнивают с табличным в соответствии с заданным уровнем значимости  $p = 0,05$ , и числом степеней свободы  $k_1 = 1, k_2 = 2$ .

Например, в процессе учета мелких млекопитающих в двух разных биотопах, где стояло по 200 ловушек ( $n = 200; k_1 = 1, k_2 = 398$ ), попало соответственно 5 и 15 зверьков. Отличается ли численность животных на этих площадках? Если рассматривать ловушку как вариант, способную принимать два значения - «пустая» и «сработавшая» (со зверьком), то получаем выборку вариант (ловушек) с альтернативным распределением. Число пойманных особей можно пересчитать в процент сработавших ловушек:  $Y = 2,5 \%$ ;  $W_2 = 7,5 \%$ . Согласно формуле (4.2) вычисляем значение критерия Фишера:  $F = 5,7$ . Полученная величина больше критической  $F_{0,05,1,398} \sim 3,9$  (см. Приложение Б, табл. Б1). Значит, численность мелких млекопитающих во втором биотопе достоверно выше, чем в первом.

**Сравнение показателей изменчивости.** Изучение суммарного влияния многих количественно неопределенных факторов, сочетающихся в произвольных и непонятных для исследователя соотношениях, часто приводит к неясным закономерностям и ошибочным выводам. В таких случаях следует все факторы, обуславливающие явления, разделить на основные  $a_j$ ,  $h_j$ , оказывающие наибольшее влияние на развитие явления, и дополнительные  $b_j$ , влияющие на развитие явления второстепенно. Тогда в опыте измеряют или отмечают лишь величины и свойства, характеризующие основные факторы.

Чтобы устранить или, по крайней мере, уменьшить ошибку, появляющуюся вследствие деления основных факторов на  $a_j$  (поддающиеся изменениям в ходе эксперимента) и  $b_j$  (не поддающиеся изменениям в ходе эксперимента), надо при постановке опытов стремиться нейтрализовать вариацию  $b_j$  факторов. То есть необходимо создать такие условия, при которых действия  $b_j$  факторов было бы возможно более неизменно и незначительно, а величины и свойства, характеризующие эти факторы, приближались бы к постоянным. При этом исследователь должен стремиться сделать переменными величинами лишь  $a_j$  факторы. Таким образом, общим принципом исследования является постоянство всех остальных факторов

при изменении избранных  $a_j$ . Тестом на соблюдение выше перечисленных условий является *критерий равноточности измерений*.

*Измерения*, дающие дисперсии одинаковой величины, называются *равноточными*. Равноточность измерений серии опытов позволяет обобщить обработку результатов экспериментов и уменьшает суммарные ошибки исследования. Практически добиться равноточности измерений можно только в тех случаях, когда измерения будут проводить достаточно опытный человек одним и тем же прибором в одинаковых условиях.

Условием равноточности опытов служит критерий Фишера (см. Приложение Б, табл. Б1), в форме отношения дисперсий (большее значение должно стоять в числителе):

$$F = \frac{\sigma_1^2}{\sigma_2^2} \sim F_{\text{табл}}(P, k_1, k_2), \quad (4-5)$$

где  $\sigma_1 = S_1$  и  $\sigma_2 = S_2$  - среднеквадратические отклонения разных серий опытов,  $k_1 = n_1 - 1$ ;  $k_2 = n_2 - 1$ .

В случае если вычисленное значение критерия Фишера окажется меньше или равным табличному ( $F < F_{\text{табл}}$ ) результаты считаются равноточными, расхождение между ними может считаться несущественным, случайным, т.е. нулевая гипотеза ( $H_0$ ) не отвергается. В обратном случае ( $F > F_{\text{табл}}$ ) различие между дисперсиями признается достоверным, результаты считаются неравноточными (выборки не принадлежат к одной генеральной совокупности). В некоторых случаях неравноточность приводит к несравнимым результатам, к ошибочным выводам и заключениям, поэтому надо всегда добиваться равноточности измерений исследуемых величин. Для избегания получения искажения объективной информации также необходимо исключить грубые ошибки, которые возникают при проведении измерений.

Обобщение результатов эксперимента, полученных в различных опытах, проводят при помощи вычисления средневзвешенной величины и ее ошибки. Заметим, что в отличие от  $t$ -критерия  $F$ -критерий чувствителен к отклонениям исходных случайных величин от нормальности. При значительных отклонениях от нормальности, особенно при небольшом числе наблюдений, его не следует применять.

Например, при сравнении по показателю плодовитости (число эмбрионов на самку) двух популяций красной полевки с разным уровнем численности (у первой, горной, популяции плотность населения в два раза выше, чем у равнинной) оказалось, что при очень близких средних арифметических (соответственно  $M_1 = 5,8$  и  $M_2 = 5,4$ , разница статистически недостоверна) стандартные отклонения значительно различаются:  $S_1 = 1,82$  и  $S_2 = 0,52$  (при  $n_1 = 27$ ,  $n_2 = 12$ ). Согласно формуле (4.4)  $F = 12,25$ . Полученное значение критерия больше табличного  $P_{\alpha, \nu_1, \nu_2}$  (0,05; 26 и 11) = 2,6. Следовательно, нулевую гипотезу о случайности отличий можно отбросить, сделав вывод о том, что показатели изменчивости плодовитости в разных по численности популяциях достоверно отличаются, результаты неравноточные. С биологических позиций это понятно, поскольку генетические отличия между особями практически по всем признакам, включая плодовитость, в больших популяциях выше, чем в малых. Новым фактором, усиливающим изменчивость особей в выборке, становится возможность появления абберантных форм (измененных форм) в условиях более свободной панмиксии (свободное скрещивание особей в пределах популяции или какой-либо другой внутривидовой группы).

*Коэффициенты вариации* также можно использовать для сравнения изменчивости разных показателей. Достоверность отличий коэффициентов вариации оценивается с помощью критерия Стьюдента по формуле:

$$\frac{|C_1 - C_2|}{\sqrt{m_1^2 + m_2^2}} \quad \text{Дабл } (0,05, n_1 + n_2 - 2) \quad (4.6)$$

где  $C_1$ ,  $C_2$  и  $m_1, m_2$  - значения и ошибки коэффициентов вариации.

Описанные выше статистические критерии (t-критерий, F-критерий и др.) относятся к параметрическим, т. е. используют стандартные параметры распределений ( $\mu$ ,  $S$ ,  $n$ ). Они связаны с законом нормального распределения и применяются для оценки расхождения между генеральными параметрами по выборочным показателям сравниваемых совокупностей.

Существенным достоинством параметрических критериев служит их большая статистическая мощность, т.е. широкие разрешающие возможности, а недостатком является трудоемкость расчетов, неприме

*нимость к распределениям, сильно отклоняющимся от нормального, а также при исследовании качественных признаков.*

Правильное применение параметрических критериев для проверки статистических гипотез основано на предположении о нормальном распределении совокупностей, из которых взяты сравниваемые выборки. Однако это не всегда имеет место, так как не все биологические признаки распределяются нормально. Немаловажным является и то обстоятельство, что исследователю приходится иметь дело не только с количественными, но и с качественными признаками, многие из которых выражаются порядковыми номерами, индексами и другими условными знаками. В таких случаях необходимо использовать непараметрические критерии.

#### ***4.1.1. Сравнение выборок с помощью непараметрических критериев***

*Непараметрические методы проверки гипотез* - это методы проверки нулевой гипотезы, предполагающие, что изучаемые переменные измерены с помощью номинальной или порядковой шкал.

Непараметрические критерии обладают широкой областью применения, устойчивостью выводов, простотой математических средств. Они значительно менее трудоемки, а при распределениях, далеких от нормального, более эффективны и точны, чем параметрические. *В основе непараметрических статистических критериев лежит оперирование частотами или рангами эмпирических данных*, при этом, в отличие от параметрических критериев, тип распределения данных не обязательно должен соответствовать нормальному. Для расчета непараметрических критериев результаты измерений должны быть представлены в шкале наименований рангов или в шкале интервалов, если распределение интервальных данных значимо отличается от нормального, что бывает довольно часто при малом размере выборки. Тип распределения данных при использовании непараметрических критериев может быть любым, но при этом необходимо учитывать ограничения, специфичные для некоторых критериев.

Процедура *расчета непараметрических критериев* состоит из *трех этапов*:

- упорядочивание и ранжирование вариант;

подсчет сумм рангов в соответствии с правилами данного критерия;

- сравнение полученной величины с табличным значением критерия.

При этом с параметрическими критериями их роднит общая идеологическая подоплека. Нулевая гипотеза, как правило, состоит в том, что сравниваемые выборки взяты из одной и той же генеральной совокупности, значит, характер распределения вариантов в этих выборках должен быть сходным. Поскольку вместо самих значений вариант используются ранги, *все непараметрические методы исследуют один вопрос, насколько равномерно варианты разных выборок «перемешаны» между собой*. Если варианты разных выборок более или менее регулярно чередуются в общем упорядоченном ряду, значит, они распределены сходным образом и отличий между совокупностями нет. Если же выборки пересекаются не полно (смешиваются только краями распределений, либо одна поглощает другую), то становится ясно, что эти выборки взяты из разных генеральных совокупностей (со смещенными центрами или разными дисперсиями).

***U-критерий (Уилкоксона — Манна - Уитни)***. Этот метод сравнения двух выборок признается наиболее чувствительным и мощным среди прочих непараметрических критериев. Согласно нулевой гипотезе, сравниваемые совокупности имеют одинаковые распределения. Техника метода состоит в том, что все варианты сравниваемых совокупностей ранжируют в одном общем ряду - каждому значению присваивают ранг, порядковый номер. При этом одинаковым (повторяющимся) значениям вариант должен соответствовать один и тот же средний ранг (они как бы «делят места»).

Для расчета критерия необходимо:

1. Расположить числовые значения сравниваемых выборок в возрастающем порядке в один общий ряд и пронумеровать члены общего ряда от 1 до  $N(N = n_1 + n_2)$ . Эти номера и будут «рангами» членов ряда
2. Отдельно для каждой выборки найти суммы рангов  $R_i$  определить величины, которые отображают связь между суммами рангов первой и второй выборки.

3. В качестве  $F$ -критерия использовать меньшую величину  $S^2$ , которую сравнить с табличным значением  $U_{st}$ . Условием для сохранения принятой Яо-гипотезы служат неравенство  $[7\phi > U_{st}$ .

После этого ранги вариант суммируют отдельно по каждой выборке и вычисляют величину критерия.

$$R_i = \sum_{j=1}^k r_{ij}, \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, n_2.$$

$$U = \frac{0,5 - \frac{1}{n_2} \sum_{i=1}^k R_i}{1 - \frac{1}{n_2} \sum_{i=1}^k R_i} \quad (47)$$

где  $U = \max(U_1, U_2)$  - максимальное значение из двух величин.

$$U_j = n \sum_{i=1}^k n_{ij} + 0,5 - \frac{1}{n} \sum_{i=1}^k R_i \quad (48)$$

Если выборка достаточно велика ( $n > 20$ ), величина статистики  $t$  сравнивается с табличным значением критерия Стьюдента для  $k = \infty$  и  $\alpha = 0,1$  (т.е. только для верхней 95 % области нормального распределения). Считается, что метод хорошо работает для выборок объемом больше 10. В случае с меньшими выборками нужно пользоваться таблицами Уилкоксона — Манна — Уитни.

**Q-критерий Розенбаума.** Этот критерий, как и предыдущий, оценивает достоверность различий двух эмпирических распределений, но в отличие от него почти не требует вычислений. Сравним два ряда цифр, характеризующих привесы барашков (в граммах) одного возраста при добавлении в корм специальной подкормки (234, 277, 214, 201, 174, 167, 184, 157, 196, 173, 190, 191, 141, 150, 191) и без нее (183, 154, 175, 159, 157, 189, 198, 165, 176, 124, 173, 182, 204, 151, 147). Устанавливаем максимальные значения (277 и 204) и минимальные значения (141 и 124), и определяем порядковый номер сравниваемых совокупностей. В качестве первой следует принять выборку с наибольшей вариантой (277).

Далее находим количество значений первой выборки, превышающих максимальное значение второй выборки (204):  $Q_1 = 3$  (это варианты 234, 277, 214). Затем определяем число вариантов второй выборки, уступающих по величине минимальному значению первой выборки (141):  $Q_2 = 1$  (варианта 124). Далее определяем критерий Розенбаума как сумму полученных чисел:  $Q = Q_1 + Q_2 = 3 + 1 = 4$ . По таблице ( $F$ -критерия находим критическое значение  $F_{табл}(0,05; 15, 15) = 6$ .

Поскольку  $\chi^2_{\text{эмпир}} < \chi^2_{\text{табл}} (4 < 6)$ , приходим к выводу об отсутствии достоверного отличия выборок друг от друга, а значит, и влияния подкормки на привесы барашков. Следует все же иметь в виду, что возможности этого метода ограничены, он дает лишь прикидочный результат и оказывается эффективным только в случае сравнительно больших различий между выборками.

**Сравнение двух частотных распределений. Критерий** В практике биологических исследований часто бывает необходимо проверить ту или иную гипотезу, т.е. выяснить, насколько полученный экспериментатором фактический материал подтверждает теоретическое предположение, насколько анализируемые данные совпадают с теоретически ожидаемыми. Возникает задача статистической оценки разницы между фактическими данными и теоретическим ожиданием, установления того, в каких случаях и с какой степенью вероятности можно считать эту разницу достоверной и, наоборот, когда ее следует считать несущественной, незначимой, находящейся в пределах случайности. В последнем случае сохраняется гипотеза, на основе которой рассчитаны теоретически ожидаемые данные или показатели. Таким вариационно-статистическим приемом проверки гипотезы служит метод *хи-квадрат* ( $\chi^2$ ). Этот метод часто называют «критерием соответствия» или «критерием согласия» Пирсона. С его помощью можно с той или иной вероятностью судить о степени соответствия эмпирически полученных данных теоретически ожидаемым.

С формальных позиций сравниваются два вариационных ряда, две совокупности: одна - эмпирическое распределение, другая представляет собой выборку с теми же параметрами ( $n$ ,  $M$ ,  $S^2$  и др.), что и эмпирическая, но ее частотное распределение построено в точном соответствии с выбранным теоретическим законом (нормальным, Пуассона, биномиальным и др.), которому предположительно подчиняется поведение изучаемой случайной величины.

В общем виде формула критерия соответствия может быть записана следующим образом:

$$\chi^2 = \sum \frac{(a - A)^2}{A} \quad (4.8)$$

где  $a$  - фактическая частота наблюдений;  $A$  - теоретически ожидаемая частота для данного класса.

Нулевая гипотеза предполагает, что достоверных различий между сравниваемыми распределениями нет. Для оценки существенности этих различий следует обратиться к специальной таблице критических значений хи-квадрат (см. Приложение В, табл. В1) и, сравнив вычисленную величину  $\chi^2$  с табличной, решить, достоверно или не достоверно отклоняется эмпирическое распределение от теоретического. Тем самым гипотеза об отсутствии этих различий будет либо опровергнута, либо оставлена в силе. Если вычисленная величина равна или превышает табличную / решают, что эмпирическое распределение от теоретического отличается достоверно. Тем самым гипотеза об отсутствии этих различий будет опровергнута. Если же  $\chi < \chi_{\text{табл}}$ , нулевая гипотеза остается в силе. Обычно

принято считать допустимым уровень значимости  $p = 0,05$ , т. е. в этом случае остается только 5 % шансов, что нулевая гипотеза правильна и, следовательно, есть достаточно оснований (95 %), чтобы от нее отказаться.

Определенную проблему составляет правильное определение числа степеней свободы ( $k$ ), для которых из таблицы берут значения критерия. Для определения числа степеней свободы из общего числа классов  $N$  нужно вычесть число ограничений (т.е. число параметров, использованных для расчета теоретических частот).

В зависимости от типа распределения изучаемого признака формула для расчета числа степеней свободы будет меняться. Для *альтернативного распределения* ( $N=2$ ) в расчетах участвует только один параметр (объем выборки), следовательно, число степеней свободы составляет  $k = N - 1 = 1$ . Для *полиномиального распределения* формула аналогична:  $k = N - 1$ . Для проверки соответствия вариационного ряда *распределению Пуассона* используются уже два параметра - объем выборки и среднее значение (численно совпадающее с дисперсией); число степеней свободы  $k = N - 2$ . При проверке соответствия эмпирического распределения варианту *нормальному* или *биномиальному распределению* число степеней ) свободы берется как число фактических классов минус три условия построения рядов — объем выборки, средняя и дисперсия,  $k = N - 3$ .

**Сразу стоит отметить, что критерий  $\chi^2$  работает только для выборок объемом не менее 25 вариантов, а частоты отдельных классов должны быть не ниже 4.**

Проиллюстрируем применение критерия  $\chi^2$  на примере анализа *альтернативной изменчивости*. В одном из опытов по изучению наследственности у томатов было обнаружено 3629 красных и 1176 желтых плодов. Теоретическое соотношение частот при расщеплении признаков во втором гибридном поколении должно быть 3 : 1 (75 % к 25 %). Выполняется ли оно? Иными словами, взята ли данная выборка из той генеральной совокупности, в которой соотношение частот 3 : 1 или 0,75 : 0,25?

Сформируем таблицу, заполнив значениями эмпирических частот и результатами расчета теоретических частот по формуле:  $A = n \cdot p$ , где  $p$  - теоретические частоты (доли вариант данного типа),  $n$  - объем выборки.

Значение (цвет плода), $X_j$	Фактическая частота, $a$	Теоретическая частота, $p$	Теоретическая частота, $A$	$\frac{(a-A)^2}{A}$
Красный	3629	0,75	3604	0,1734
Желтый	1176	0,25	1201	0,5204
Сумма	$n = \sum a = 4805$	1	$n = \sum A = 4805$	$\chi^2 = 0,69$

Далее вычисляем число степеней свободы (при двух классах и одном ограничении (объеме выборки))  $k = 2 - 1 - 1$ . Находим критическое значение  $\chi^2(0,05, 1) = 3,8$  (см. Приложение В, табл. В1). Поскольку полученная величина (0,69) меньше табличной (3,8), различия сравниваемых распределений статистически недостоверны. Иначе говоря, фактические частоты хорошо согласуются с теоретически ожидаемыми. Полученные данные не отвергают принятую гипотезу о том, что в нашем случае имеется соотношение 3:1.

Здесь следует еще раз обратить внимание читателей на то обстоятельство, что сохранение нулевой гипотезы нельзя считать доказательством справедливости нулевой гипотезы. Результатами представленных вычислений теория о расщеплении по фенотипам в соотношении 3 : 1 не доказана, хотя и не опровергнута. Статистика доказывает только факт отличий, а не их отсутствие. Чтобы доказать теорию, нужно: предположить

антитеорию (например, соотношение 1 : 1) и опровергнуть ее с помощью статистических приемов.

*Выбор статистического метода для анализа результатов.* Независимо от конкретных биотехнологических особенностей, для ответа на вопросы, поставленные перед исследованием, с помощью различных критериев приходится проверять нулевую гипотезу об отсутствии эффекта или различиях в эффекте, сравнивая при этом выборочные средние, доли, кривые выживаемости и т.п. Табл. 12 может помочь ориентироваться во множестве статистических критериев (параметрических и непараметрических), обычно применяемых для анализа результатов клинических исследований. Чтобы правильно выбрать статистический метод, необходимо учитывать, прежде всего, характер интересующего нас признака (количественный, порядковый или качественный) и тип распределения (нормальное или какое-либо другое). Для сравнения качественных переменных с числом градаций больше двух применяются статистические методы сравнения распределений, например критерий или таблицы сопряженности соответствующей размерности. Кроме того, необходимо различать ситуации, связанные со сравнением двух или более групп (при этом применяются различные критерии), а также сравниваются ли независимые группы или одна группа получает сравниваемые воздействия. В результате применения любого из этих критериев наличие статистически значимых различий будут или не будут установлены.

Таблица 12 - Статистические критерии, обычно применяемые для решения типичных задач в области клинических исследований

Признак	Две независимые группы	Более двух независимых групп	Одна группа, связанные измерения	Одна группа, несколько связанных измерений
<i>Параметрические методы</i>				
Количественный, нормальное распределение	Критерий Стьюдента, дисперсионный анализ, критерий Тьюкки, критерий Шеффе	Дисперсионный анализ, критерий Стьюдента для множественных сравнений, критерий Тьюкки, критерий Даннета, критерий Шеффе, критерий Ньюмена-Кейлса	Критерий Стьюдента для связанных пар, дисперсионный анализ повторных измерений	Дисперсионный анализ повторных измерений. критерий Шеффе для зависимых выборок
<i>Непараметрические методы</i>				
Количественный, распределение отличается от нормального, порядковый	Критерий Уилкоксона - Манна - Уитни, медианный критерий	Критерий Краскела - Уоллиса, медианный критерий	Г-критерий Уилкоксона, критерий знаков, критерий знаковых рангов Уилкоксона	Критерий Фридмана
<i>Методы сравнения долей</i>				
Качественный, альтернативное распределение	Критерий $u_j$ , точный критерий Фишера	Критерии $\chi^2$	Критерий Мак - Нимара	Критерий Кокрена