Из кн.: Кузнецов В.М. **Основы научных исследований в животноводстве.** Киров: Зональный НИИСХ Северо-Востока, 2006.- 568 с.

11. ОЦЕНКА ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

Всю информацию о распределении количественного признака в нормальной совокупности можно свести к двум величинам: *среднему значению* и *стандартному отклонению*. Среднее служит мерой «основной тенденции», а стандартное отклонение - мерой изменчивости признака, т.е. разброса отдельных значений вокруг среднего. Кроме этих двух статистик, совокупность характеризуют мода, медиана, размах варьирования, процентили и т.п.

11.1. Выборочное среднее

Выборочной средней, $\hat{\mu}$ или $\bar{x}$, называют среднее арифметическое значение признака выборочной совокупности.

Пусть для изучения генеральной совокупности относительно количественного признака X извлечена выборка объема n. Если все значения $x_1, x_2, ..., x_i, ..., x_n$ признака различны, то

$$\hat{\mu} = \bar{x} = (x_1 + x_2 + ... + x_i + ... + x_n)/n = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{\sum x_i}{n}.$$

Среднее может быть рассчитано также путем перемножения значений x_i на их относительную частоту встречаемости, с последующим суммированием произведений. Так, если значение x_i встречается n_i раз, то относительная частота его равна $p_i = n_i / n$ (при $n_1 + n_2 + ... + n_i + ... + n_k = n$) и выборочное среднее есть

$$\hat{\mu} = \bar{x} = \sum_{i=1}^{k} p_i x_i = \sum p_i x_i.$$

Пример 11.1. Два метода расчета среднего значения:

Метод 1

x_i
9
5
4
7
5
Σx_i = 30

$\bar{x} = \frac{1}{n}\sum x_i = 6{,}0$

Метод 2

x_i		p_i		
9	×	0,2	=	1,8
5	×	0,4	=	2,0
4	×	0,2	=	0,8
7	×	0,2	=	1,4

$\bar{x} = \sum p_i x_i$ = 6,0

Математическое ожидание средней арифметической:

$$E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} n\mu = \mu .$$

Следовательно,

$$E(\bar{x}) = \mu .$$

Известно, что $Var(\bar{x}) = \sum Var(x_i)/n^2$. Так как X - случайная переменная, варианса которой σ^2, то $Var(x_i) = \sigma^2$. Тогда

$$Var(\bar{x}) = (1/n^2)\underbrace{(\sigma^2 + ... + \sigma^2 + ... + \sigma^2)}_{n} = (1/n^2)(n\sigma^2) = \sigma^2/n .$$

Из отношения следует важное свойство: *варианса средней в n раз меньше вариансы нормально распределенной случайной величины X*, т.е. средние группируются вокруг центра распределения теснее, чем отдельные значения.

Оценка $\hat{\mu}$ (или $\bar{x}$) несмещенная, эффективная (имеет минимальную вариансу) и состоятельная. Кроме того, если по нескольким выборкам достаточно большого объема из одной и той же генеральной совокупности рассчитать средние, то они будут приближенно равны между собой. В этом заключается свойство *устойчивости* выборочных средних.

Следует отметить, что если вариансы двух одинаково распределенных совокупностей равны, то близость выборочных средних к генеральным зависит от объема выборки: чем объем выборки больше, тем меньше выборочное среднее отличается от генеральной. Например, если из одной популяции отобран 1% животных, а из другой - 4%, причем объем первой выборки был больше, чем второй, то первое выборочное среднее будет меньше отличаться от соответствующей генеральной средней.

11.2. Групповая и общая средние

Допустим, что все значения количественного признака X выборки разбиты на несколько групп. Рассматривая каждую группу как самостоятельную субвыборку, можно найти ее среднее.

Групповой средней, $\hat{\mu}_j$ или $\bar{x}_j$, называют среднее арифметическое значений признака, принадлежащих группе. *Общей средней,* $\hat{\mu}$ или $\bar{x}$, называют среднее значений признака, принадлежащих всей выборке. Зная групповые средние и объемы групп, можно найти общее среднее, которое равно средней арифметической групповых средних, взвешенных на объемы групп.

Пример 11.2. Необходимо найти общее среднее выборки, состоящей из следующих двух групп:

Группа	Первая		Вторая	
Значение признака	1	6	1	5
Частота	10	15	20	30
Объем	10+15=25		20+30=50	
Групповые средние	$\overline{x}_1$=(1×10+6×15)/25=4,0		$\overline{x}_2$=(1×20+5×30)/50=3,4	
Общее среднее	$\overline{x}$=(4,0×25+3,4×50)/(25+50)=3,6			

Отклонение от общей средней и его свойство. Рассмотрим выборочную совокупность значений количественного признака X объема n:

значения признака $x_1 \quad x_2 \quad \dots \quad x_k$

частоты $n_1 \quad n_2 \quad \dots \quad n_k$.

При этом $\sum_{i=1}^{k} n_i = n$. Далее для удобства записи знак суммы $\sum_{i=1}^{k}$ заменен знаком Σ.

Общее среднее:

$$\overline{x} = (\sum n_i x_i)/n \quad \rightarrow \quad \sum n_i x_i = n\overline{x}.$$

Так как $\overline{x}$ - постоянная величина, то

$$\sum n_i \overline{x} = \overline{x} \sum n_i = n\overline{x}.$$

Отклонением называют разность, $x_i - \overline{x}$, между значением признака и общей средней. Отклонение от среднего имеет следующее свойство: сумма произведений отклонений на соответствующие частоты (или сумма всех отклонений по выборке) равна нулю:

$$\sum n_i (x_i - \overline{x}) = \sum n_i x_i - \sum n_i \overline{x} = n\overline{x} - n\overline{x} = 0.$$

Из этого следует, что среднее значение отклонений также равно нулю:

$$(\sum n_i (x_i - \overline{x}))/\sum n_i = 0/n = 0.$$

Пример 11.3. Имеется распределение количественного признака:

$$\begin{array}{llll} x_i: & 1 & 2 & 3 \\ n_i: & 10 & 4 & 6. \end{array}$$

Следует показать, что сумма произведений отклонений на соответствующие частоты равна нулю.

Общее среднее

$$\overline{x} = (10 \times 1 + 4 \times 2 + 6 \times 3) / 20 = 1{,}8.$$

Взвешенная сумма произведений отклонений

$$\sum n_i (x_i - \overline{x}) = 10(1 - 1{,}8) + 4(2 - 1{,}8) + 6(3 - 1{,}8) = 8 - 8 = 0.$$

11.3. Выборочная варианса

Среднее указывает на точку, вокруг которой варьируют значения признака. Но среднее не дает никакой информацию о том, как эти значения распределены. Два разные набора данных могут иметь одно и тоже среднее, но различаться по степени варьирования признака. Например, наборы данных

3,4, 5, 6, 7 и

-10, 2, 3, 10, 20

имеют средние, равные 5, но второй набор данных более изменчивый, чем первый.

В качестве меры изменчивости значений признака используется *варианса* (*дисперсия*), или усредненный квадрат отклонений от среднего. Выборочную вариансу для признака X обозначают $\hat{\sigma}_x^2$. Она может быть вычислена двумя способами:

- непосредственно, как

$$\hat{\sigma}_x^2 = \frac{1}{n-1}\left[(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 +,...,+ (x_n - \overline{x})^2\right] =$$

$$= \frac{1}{df}\sum_{i=1}^{n}(x_i - \overline{x})^2 \quad \text{и}$$

- при известных относительных частотах каждого квадрата отклонения от среднего ($p_i = n_i / n$), как

$$\hat{\sigma}_x^2 = \frac{n}{n-1}\sum p_i (x_i - \overline{x})^2,$$

где (n-1)=df - число степеней свободы*.

Единицы измерения вариансы - это единицы анализируемого признака в *квадрате*.

* Число степеней свободы (degrees freedom, df). Если известен ряд от x_1 до x_n, состоящий из n членов или наблюдений, то для него общей характеристикой является среднее арифметическое. При изветной средней, все наблюдения ряда, кроме одного (n-1), могут принимать *любые* значения. Но значение одного наблюдения в данной совокупности будет зависеть от значений остальных наблюдений.

Пример 11.4. Оценка выборочной вариансы двумя методами:

Метод 1

x_i	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$
9	-3	9
5	-1	1
4	-2	4
7	1	1
5	-1	1

$SS_x = \sum (x_i-\bar{x})^2 = 16$

$\hat{\sigma}_x^2 = \frac{1}{n-1}\sum (x_i-\bar{x})^2 = 16/(5-1) = 4.$

Метод 2

x_i	$(x_i-\bar{x})^2$		p_i		
9	9	×	0,2	=	1,8
5	1	×	0,4	=	0,4
4	4	×	0,2	=	0,8
7	1	×	0,2	=	0,2

$\hat{\sigma}_x^2 = \frac{n}{n-1}\sum p_i (x_i-\bar{x})^2 = 4.$

Формула расчета вариансы может иметь различный вид:

$$\hat{\sigma}_x^2 = \frac{SS_x}{df} = \frac{\sum(x_i-\bar{x})^2}{n-1}$$

$$= \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$$

$$= \frac{n\sum x_i^2 - (\sum x_i)^2}{n(n-1)}$$

$$= \frac{\sum x_i^2 - n\bar{x}^2}{n-1}.$$

где SS_x - сумма квадратов по признаку X.

$\hat{\sigma}_x^2$ является несмещенной и состоятельной, но не эффективной оценкой σ_x^2.

11.4. Стандартное отклонение

Более удобно измерять изменчивость *стандартным отклонением* (SD или $\hat{\sigma}_x$) - значением корня квадратного из вариансы:

$$SD = \hat{\sigma}_x = \sqrt{\hat{\sigma}_x^2}\ .$$

Стандартное отклонение выражается в тех же единицах, что и среднее.

Для сравнения изменчивости разных признаков, или одного признака в разных выборках, используют *коэффициент изменчивости* (CV):

$$CV_x = (\hat{\sigma}_x / \bar{x})100\%.$$

Коэффициент изменчивости выражает вариацию признака в процентах. Чем более однороден изучаемый материал, тем меньше CV. Однако даже при достаточной однородности выборки, степень изменчивости разных признаков будет различной. Это определяется биологическими особенностями признаков.

Пример 11.5. Пусть имеется 10 телят со следующей живой массой ($\overline{x} = 40{,}7$ кг):

Номер теленка (i)	Вес (x_i) в кг
1	40
2	42
3	35
4	36
5	45
6	47
7	40
8	43
9	41
10	38

$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^{10} x_i^2 - n\overline{x}^2}{n-1} = \frac{16693 - 10(40{,}7)^2}{9} = $$
$$= 14{,}233\,\text{кг}^2;$$

где

$$\sum_{i=1}^{10} x_i^2 = \text{общая сумма квадратов} =$$
$$= 40^2 + 42^2 + ... + 38^2 = 16693.$$

Стандартное отклонение и коэффициент изменчивости:

$$\hat{\sigma}_x = \sqrt{14{,}233} = 3{,}77\,\text{кг};$$

$$CV = (3{,}77 / 40{,}7)100 = 9{,}3\%.$$

11.5. Надежность оценки средней

Выборочная оценка среднего, $\hat{\mu}$ (или $\overline{x}$), характеризует среднее генеральной совокупности, μ, лишь приближенно, отличаясь от него на некоторую величину. Значения средних по разным выборкам из одной генеральной совокупности варьируют. Эта вариация также имеет нормальное распределение и свое стандартное отклонение ($\hat{\sigma}_{\hat{\mu}}$), которое называется *ошибкой среднего*

($m_{\hat{\mu}}$) или *стандартной статистической ошибкой* ($SE_{\hat{\mu}}$). Ошибка среднего прямо пропорциональна изменчивости признака и обратно пропорциональна числу животных (наблюдений) в выборке:

$$\hat{\sigma}_{\hat{\mu}} = m_{\hat{\mu}} = SE_{\hat{\mu}} = \frac{\hat{\sigma}}{\sqrt{n}}.$$

Ошибка представляет собой величину расхождения между *истинным* значением параметра и его *оценкой* $(\mu - \hat{\mu})$.

Доверительный интервал, который «накроет» *истинное* значение среднего (μ) с доверительной вероятностью P=1-α, есть:

$$(\hat{\mu} - t_{\alpha;df}\, m_{\hat{\mu}}) < \mu < (\hat{\mu} + t_{\alpha;df}\, m_{\hat{\mu}}),$$

где $t_{\alpha;df}$ - критическая точка (значение) для уровня значимости α и df=n-1 (берут из табл. А.8 Приложения А; α для *двусторонней* критической области).

Для проверки $H_0: \mu = 0$ рассчитывают *наблюдаемый критерий* для $\hat{\mu}$ - $t_{\hat{\mu}}$ (К в разделе 10):

$$t_{\hat{\mu}} = \frac{\hat{\mu} - 0}{\hat{\sigma}_{\hat{\mu}}} = \frac{\hat{\mu}}{m_{\hat{\mu}}},$$

который сравнивают с критической точкой $t_{\alpha;df}$ (k в разделе 10; определяют по табл. А.8). Если $t_{\hat{\mu}} \geq t_{\alpha;df}$ при $df = (n-1)$, то H_0 отвергают и утверждают, что выборочное среднее статистически значимо.

В больших выборках (n>100) среднее статистически значимо отличается от 0 при $t_{\hat{\mu}} = 3{,}00$ с P=99,7%, $t_{\hat{\mu}} = 2{,}58$ с P=99% и $t_{\hat{\mu}} = 1{,}96$ с P=95% (см. табл. А.3 Приложения А; $u = t_{\hat{\mu}}$).

Выражение $t_{\alpha;df}\, m_{\hat{\mu}} = t_{\alpha;df} \frac{\hat{\sigma}}{\sqrt{n}} = \delta$ есть *точность* (погрешность) оценки истинного среднего. Его можно использовать для расчета минимального объема выборки, который обеспечит заданную точность:

$$n = t^2_{\alpha;df}\, \hat{\sigma}^2 / \delta^2.$$

11.6. Сравнение средних

Сравнение средних (например, опытной и контрольной групп животных) основывается на анализе нулевой гипотезы.

Первоначально принимают, что *обе выборки составляют одну совокупность и между средними различия нет*. Статистический анализ должен привести или к отклонению нулевой гипотезы, если будет доказана статистическая значимость полученного различия, или к ее сохранению, если таковой не будет доказана.

Отрицание нулевой гипотезы должно быть связано с принятием определенного уровня значимости. Так, если принят $\alpha = 1\%$ и если вероятность надежности оценки разности средних не удовлетворяет этому условию, т.е. она ниже 99%, например 97%, то тогда нет оснований для отбрасывания нулевой гипотезы.

Допустим, что имеются две группы животных с численностью n_1 и n_2. Нулевая гипотеза сводится к признанию того, что эти две выборки являются в сущности одной и разность между их средними,

$$\hat{d} = \hat{\mu}_1 - \hat{\mu}_2,$$

чисто случайна, т.е. лежит в пределах ошибки ($H_0 : \mu_1 = \mu_2$).

Как и для среднего значения, рассчитывают стандартную ошибку разности ($\hat{\sigma}_{\hat{d}}$ или $m_{\hat{d}}$):

$$\hat{\sigma}_{\hat{d}} = m_{\hat{d}} = \sqrt{m^2_{\hat{\mu}_1} + m^2_{\hat{\mu}_2}} \ .$$

Значимость $\hat{d}$ проверяют по t-*критерию Стьюдента*:

$$t_{\hat{d}} = \frac{|\hat{d}|}{m_{\hat{d}}},$$

который сравнивают с критической точкой - $t_{\alpha;df}$ (берут из табл. А.8 Приложения А; α для *двусторонней* критической области). Если $t_{\hat{d}} \geq t_{\alpha;df}$ при $df = n_1 + n_2 - 2$, то гипотезу H_0 на уровне значимости α отвергают (разница между средними статистически значима). В противном случае гипотезу H_0 принимают, т.е. предположение об отсутствии различий между опытной и контрольной группами не вызывает возражений.

Приближенно, при большом числе животных или наблюдений (n>100), если выборочная оценка больше своей удвоенной ошибки, то она статистически значима с вероятностью 95%, если больше утроенной ошибки - то с вероятностью 99%.

Доверительный интервал, который «накроет» *истинное* значение различия между средними двух групп животных (d) с доверительной вероятностью $P=1-\alpha$, есть:

$$(\hat{d} - t_{\alpha;df}\, m_{\hat{d}}) \; < \; d \; < (\hat{d} + t_{\alpha;df}\, m_{\hat{d}}).$$

Следует отметить, что произведение $t_{\alpha;df}\, m_{\hat{d}} = \delta_{\hat{d}}$ (точность или погрешность оценки) иногда используют как значение **наименьшей существенной разности** (НСР) при сравнении соседних средних ранжированного ряда (расположение средних в порядке возрастания их значений).

Пример 11.6. Пусть имеются данные по многоплодию свиноматок уржумской (n_x =8) и крупной белой (n_y =10) пород:

x_i	y_i	x_i^2	y_i^2
10	8	100	64
7	11	49	121
9	11	81	121
6	9	36	81
14	13	196	169
8	10	64	100
11	9	121	81
13	12	169	144
-	14	-	196
-	12	-	144
Σx=78	Σy=109	$\sum x_i^2$ =816	$\sum y_i^2$ =1221

Расчеты:

Средние: $\hat{\mu}_x = 78/8 = 9{,}75$ гол.; $\hat{\mu}_y = 109/10 = 10{,}9$ гол.;

Вариансы: $\hat{\sigma}_x^2 = \dfrac{816 - 8 \times 9{,}75^2}{8-1} \approx 7{,}93$; $\hat{\sigma}_y^2 = \dfrac{1221 - 10 \times 10{,}9^2}{10-1} \approx 3{,}66$;

Сигмы: $\hat{\sigma}_x = \sqrt{7{,}93} = 2{,}82$; $\hat{\sigma}_y = \sqrt{3{,}66} = 1{,}91$;

$CV_{\%}$: $CV_x = \dfrac{2{,}82}{9{,}75}100 = 28{,}9\%$; $CV_y = \dfrac{1{,}91}{10{,}9}100 = 15{,}5\%$;

Ошибки: $m_{\hat{\mu}_x} = \dfrac{2{,}82}{\sqrt{8}} = 0{,}997$; $m_{\hat{\mu}_y} = \dfrac{1{,}91}{\sqrt{10}} = 0{,}604$;

Критерии: $t_{\hat{\mu}_x} = \dfrac{9{,}75}{0{,}997} = 9{,}779$; $t_{\hat{\mu}_y} = \dfrac{10{,}9}{0{,}604} = 18{,}046$;

Степени свободы: $df_x = 8 - 1 = 7$; $df_y = 10 - 1 = 9$;

Критические точки из табл. А.8 Приложения А:

для $df_x = 7$ $t_{0,05} = 2{,}37$; $t_{0,01} = 3{,}5$; $t_{0,001} = 5{,}4$;

для $df_y = 9$ $t_{0,05} = 2{,}23$; $t_{0,01} = 3{,}25$; $t_{0,001} = 4{,}78$

Доверительные границы при P=95%:

для породы «X» $\hat{\mu}_{x_{min}} = 9{,}75 - 2{,}37 \times 0{,}997 = 7{,}4$ гол.

$\hat{\mu}_{x_{max}} = 9{,}75 + 2{,}37 \times 0{,}997 = 12{,}1$ гол.

для породы «Y» $\hat{\mu}_{y_{min}} = 10{,}9 - 2{,}23 \times 0{,}604 = 7{,}9$ гол.

$\hat{\mu}_{y_{max}} = 10{,}9 + 2{,}23 \times 0{,}604 = 13{,}9$ гол.

Доверительные интервалы:

для μ_x от 7,4 до 12,1 гол.

для μ_y от 7,9 до 13,9 гол.

Разность средних: $\hat{d} = 9{,}75 - 10{,}9 = -1{,}15$;

Ошибка разности: $m_{\hat{d}} = \sqrt{0{,}997^2 + 0{,}604^2} \approx 1{,}17$;

Наблюдаемый критерий: $t_{\hat{d}} = \dfrac{|-1{,}15|}{1{,}17} = 0{,}98$.

Число степеней свободы: $df = 8 + 10 - 2 = 16$.

Критические точки ($df = 16$): $t_{0,05} = 2{,}120$; $t_{0,01} = 2{,}921$; $t_{0,001} = 4{,}015$.

Границы доверия при P=95%: $\hat{d}_{min} = -1{,}15 - 2{,}12 \times 1{,}17 = -3{,}6$ гол.

$\hat{d}_{max} = -1{,}15 + 2{,}12 \times 1{,}17 = +1{,}3$ гол.

Доверительный интервал для d: от –3,6 до +1,3 гол.

Выводы: Многоплодие свиноматок уржумской породы характеризовалось значительной вариабельностью (CV=28,9%). В крупной белой породе признак был более однороден (CV=15,5%). Среднее многоплодие свиноматок уржумской породы (9,7±1,00) было на 1,1 поросенка ниже, чем крупной белой породы (10,9±0,60). Однако различие можно считать случайным, т.к. $t_{набл.} < t_{0,05}$. Истинное различие в многоплодии свиноматок двух пород при 95% доверительной вероятности находилось в интервале от –3,6 до +1,3 поросенка. Доверительный интервал включал нулевое значение, что также свидетельствовало о случайном характере межпородных различий по многоплодию свиноматок.

При сравнении **парных наблюдений** на одних и те же животных или, например, продуктивности дочерей (D) и матерей (M), используют *парный* t-*критерий*: сначала рассчитывают отклонения по каждой паре, $d_i = (x_{D_i} - x_{M_i})$, которые затем обрабатывают обычным способом, т.е. рассчитывают $\hat{\mu}_d$, $\hat{\sigma}_d$, $m_{\hat{\mu}_d}$ и $t_{\hat{\mu}_d}$. Последнюю величину сравнивают с критической точкой $t_{\alpha;df}$ при df=n-1 (табл. А.8).

11.7. Сравнение варианс

Задача сравнения варианс возникает тогда, когда требуется оценить влияние каких-либо факторов на признаки, однородность выборок или результатов экспертных оценок, сравнить точность приборов, инструментов, самих методов измерений и т.п.

Пусть имеются две нормально распределенные генеральные совокупности с вариансами σ_1^2 и σ_2^2. Необходимо проверить нулевую гипотезу о равенстве варианс, т.е.

$$H_0:\ \sigma_1^2 = \sigma_2^2 = \sigma^2,$$

относительно конкурирующей $H_1:\ \sigma_1^2 \neq \sigma_2^2$.

С целью проверки гипотезы H_0 из этих совокупностей взяты две независимые выборки объемом n_1 и n_2. Для оценки варианс σ_1^2 и σ_2^2 используют выборочные вариансы $\hat{\sigma}_1^2$ и $\hat{\sigma}_2^2$. Следовательно, задача проверки гипотезы H_0 сводится к сравнению варианс $\hat{\sigma}_1^2$ и $\hat{\sigma}_2^2$.

Нулевую гипотезу, или однородность выборочных варианс, проверяют по *критерию Фишера*. Для этого рассчитывают отношение бóльшей выборочной вариансы к меньшей, т.е. случайную величину (при $\hat{\sigma}_1^2 > \hat{\sigma}_2^2$)

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}.$$

Величина F (F-критерий) при условии справедливости гипотезы H_0 имеет распределение Фишера-Снедекора (см. гл. 8) со степенями свободы $df_1 = n_1 - 1$ и $df_2 = n_2 - 1$. Теоретически значения F-критерия могут колебаться от 0 до ∞. Если обе вариансы $\hat{\sigma}_1^2$ и $\hat{\sigma}_2^2$ равны, то тогда F=1. Если они не равны, то нужно доказать, что это неравенство неслучайно, статистически значимо.

Значения $F_{\alpha;\, df_1;\, df_2}$, критические для признания статистической значимости разницы между вариансами, приводятся в специальных таблицах, где учитываются разные объемы сравниваемых групп (вернее для разных чисел степеней свободы этих групп) и принимаются различные уровни значимости (см., например, [14]). В сокращенном виде (только для уровня значимости α=0,05 и α=0,01) критические точки, $F_{\alpha;df_1;df_2}$, представлены в табл. А.10 и А.11 Приложении А. Таблицы позволяют осуществлять проверку гипотезы H_0 на 5%-ном и 1%-ном уровнях значимости при использовании односторонней критической области, и на 10%-ном и 2%-ном уровнях значимости - при двусторонней критической области.

Для того чтобы при заданном уровне значимости α проверить гипотезу H_0 о равенстве генеральных варианс нормально распределенных совокупностей при конкурирующей гипотезе $H_1: \sigma_1^2 \neq \sigma_2^2$ (двусторонняя критическая область), надо:

- вычислить отношение большей выборочной вариансы к меньшей, т.е. F-критерий, и
- по таблице F-распределения Фишера-Снедекора (А.10 или А.11 Приложении А) по уровню значимости α/2 (вдвое меньшем заданного) и числам степеней свободы df_1 и df_2 ($df_1 > df_2$) найти критическую значение $F_{\alpha/2;\ df_1;df_2}$.

Если наблюдаемое значение будет равно или больше табличного, $F \geq F_{\alpha/2;\ df_1;df_2}$, то гипотезу H_0 отвергают, различие между вариансами признают статистически значимым (при принятом уровне α). Оно не может быть объяснено случайными причинами, а является следствием того, что сами генеральные вариансы различны. Например, если различие выборочных варианс результатов измерения содержания жира в молоке двумя приборами (или методами) оказалось статистически значимым, то это свидетельствует о различной точности приборов (методов).

Если $F < F_{\alpha/2;\ df_1;df_2}$, то оснований отвергнуть гипотезу H_0 нет. Она справедлива, т.е. генеральные вариансы одинаковы (однородны, гомогенны), различие выборочных варианс объясняется случайными причинами, в частности случайным отбором элементов выборки. Например, если различие выборочных варианс результатов измерения содержания жира в молоке двумя приборами (методами) оказалось незначимым, то можно утверждать, что приборы (методы) имеют одинаковую точность.

Пример 11.7. По двум независимым выборкам, объемы которых равны $n_1 = 21$ и $n_2 = 16$, извлеченным из нормальных генеральных совокупностей X и Y, найдены оценки выборочных варианс: $\hat{\sigma}_1^2 = 1{,}23$ и

$\hat{\sigma}_2^2 = 0{,}41$. При α=0,1 требуется проверить гипотезу H_0 о равенстве генеральных варианс при конкурирующей гипотезе $H_1: \sigma_1^2 \neq \sigma_2^2$.

Решение. Найдем отношение бóльшей выборочной variансы к меньшей:

$$F = \frac{1{,}23}{0{,}41} = 3.$$

В табл. А.10, по уровню значимости, вдвое меньшем заданного, т.е. при α/2=0,1/2=0,05, и числам степеней свободы $df_1 = 21 - 1 = 20$, $df_2 = 16 - 1 = 15$ находим критическую точку $F_{0,1/2;\ 20;15} = 2{,}33$.

Так как $F > F_{0,1/2;\ 20;15}$, то нулевую гипотезу о равенстве генеральных варианс отвергают. Другими словами, выборочные вариансы различаются значимо. Например, если бы рассматриваемые вариансы характеризовали точность двух методов измерения, то следует предпочесть тот метод, который имеет меньшую вариансу (0,41).

Пример 11.8. Нужно сравнить изменчивость по высоте в холке групп черно-пестрого и красно-пестрого скота. Для первого $n_1 = 100$ и $\hat{\sigma}_1^2 = 16{,}32$, для второго $n_2 = 42$ и $\hat{\sigma}_2^2 = 14{,}44$. Проверить гипотезу H_0 о равенстве варианс при уровне значимости α=0,02. Имеем

$$F = 16{,}32/14{,}44 = 1{,}13.$$

В табл. А.11 Приложения А в вертикальных столбцах нет цифры 100. Тогда надо взять $df_1 = \infty$. По горизонтали же можно взять $df_2 = 40$. Критическое значение составило $F_{0,02/2;\ 100;40} = 1{,}8$. Фактическое значение F ниже этой величины. Отсюда можно сделать вывод, что хотя черно-пестрый и красно-пестрый скот отличаются по масти, но различие варианс по высоте в холке статистически незначимо. Вероятность различия между вариансами, как случайного, более 0,02. Нулевая гипотеза о равенстве варианс сохраняет свое значение и остается неопровергнутой. Можно считать, что группы черно-пестрого и красно-пестрого скота по высоте в холке составляют одну популяцию.

Практическое значение F-критерия велико, особенно, в дисперсионном (или вариансном) анализе. Если различия между вариансами групп в опытах, где анализируется влияние разных факторов (способ содержания, рацион, схема лечения, кровность, линейность и т.п.) на животных, могут быть признаны статистически значимыми, то это позволяет установить влияние того или иного фактора и их взаимодействия на изучаемые признаки или биологические свойства.

Из кн.: Кузнецов В.М. **Основы научных исследований в животноводстве.** Киров: Зональный НИИСХ Северо-Востока, 2006.- 568 с.

11.8. Другие статистики

Кроме выборочных средней и вариансы нормальное распределение характеризуют такие описательные статистики, как мода, медиана, ранг, процентили, децили, квинтили и квартили.

Модой (M_0) называют значение, которое имеет наибольшую частоту. Например, для ряда

значение	1	4	7	9
частота	5	1	20	6

мода равна 7.

Медианой (m_e) называют значение переменной (признака), которое делит вариационный ряд на две части, равные по числу объектов. Если число объектов нечетно (n=2k+1), то

$$m_e = x_{k+1};$$

при четном ($n = 2k$)

$$m_e = (x_k + x_{k+1})/2.$$

Например, для ряда 2, 3, 5, 6 и 7 медиана равна 5; для ряда 2, 3, 5, 6, 7 и 9 медиана равна (5+6)/2=5,5.

Размахом варьирования (ранг, R) называют разность между наибольшим и наименьшим значениями переменной.

Например, для ряда 1, 3, 4, 5, 6 и 10 размах равен 10-1=9. Размах является простейшей характеристикой рассеяния вариационного ряда.

Квантиль (термин впервые использовал Кендалл в 1940 г.) распределения значений - это такое число x_p, что значения p-й части выборки меньше или равны x_p.

Наиболее часто используют следующие квантили:

$x_{0,5}$	- медиана;
$x_{0,25}, x_{0,5}, x_{0,75}$	- квартили;
$x_{0,1}, x_{0,2}, ..., x_{0,9}$	- децили;
$x_{0,01}, x_{0,22}, ..., x_{0,99}$	- процентили.

Например, квантиль 0,25 (также называемая 25-й процентилью или нижней квартилью) для некоторой переменной - это такое число (x_p), которое не превосходят 25% (p) значений переменной. Аналогично, квантиль 0,75 (также называемая 75-й процентилью или верхней квартилью) - это такое число, ниже которого попадают 75% значений переменной.

Процентиль (перцентиль, персентиль; термин был впервые использован Галтоном в 1885 г.) представляет собой процент наблюдений, значения которых равны данному наблюдению или меньше него. Процентили делят распределение на сто равных частей. В частности, медиана является 50-м процентилем.

Для ранжированного ряда наблюдений (например, удоя 101 коровы) 65-му процентилю соответствует 65,65-е значение уровня продуктивности (101×0,65=65,65). Но такого наблюдения в распределении нет (довольно частая проблема при расчете процентилей). Поэтому условились, что в подобных случаях величина процентиля должна определяться пропорционально к расстоянию между соответствующими наблюдениями. В данном случае величина 65-го процентиля равна 0,65 от расстояния между 65-м и 66-м наблюдениями, которым соответствуют, допустим, значения удоя 6700 и 6800 кг молока. Следовательно, 65-й процентиль равен 6765 кг.

Обратная задача - поиск процентильного ранга данного наблюдения. Пусть необходимо определить процентильный ранг удоя в 2200 кг. Число коров, удой которых не превосходит это значение, - 17 из 101. Таким образом, процентильный ранг уровня удоя в 2200 кг равен 0,168 или 16,8≈17 процентиля.

Децили, **квинтили** и **квартили**. Децили делят распределение на 10 частей, квинтили - на пять, квартили - на четыре.

Нижняя и верхняя квартили (термин был впервые использован Галтоном, 1882) равны соответственно 25-й и 75-й проценталем распределения. 25-я процентиль переменной - это такое значение, ниже которого попадают 25% значений переменной. Аналогично, 75-я процентиль - это такое значение, ниже которого попадают 75% значений переменной.

Квартильный размах переменных равен разности значений 75-й процентали и 25-й процентали. Таким образом, это тот диапазон вокруг медианы, который содержит 50% наблюдений.

11.9. Альтернативные признаки

Для признаков с альтернативной изменчивостью распределение может быть представлено в виде двух классов: животные с отсутствием признака (n^-) и животные с присутствием его (n^+), причем $n = (n^- + n^+)$. Среднее для таких признаков есть *относительная частота* животных определенного класса в общей совокупности:

$$\hat{\mu}^- = \hat{p} = \frac{n^-}{n} \quad \text{и} \quad \hat{\mu}^+ = \hat{q} = \frac{n^+}{n} = \frac{n - n^-}{n} = 1 - \hat{p},$$

а варианса - произведение относительных частот:

$$\hat{\sigma}^2_{\text{альт.}} = \hat{p}(1 - \hat{p}) = \hat{p}\,\hat{q}.$$

Соответственно, стандартное отклонение есть

$$\hat{\sigma}_{\text{альт.}} = \sqrt{\hat{p}\,\hat{q}}.$$

Ошибки оценок относительных частот равны:

$$m_{\hat{p}} = m_{\hat{q}} = \frac{\hat{\sigma}_{\text{альт.}}}{\sqrt{n}} = \sqrt{\frac{\hat{\sigma}^2_{\text{альт.}}}{n}} = \sqrt{\frac{\hat{p}\,\hat{q}}{n}}.$$

Расчеты надежности оценок частот и различий частот аналогичны таковым для количественных признаков. Однако если сравнивают частоты из одной совокупности ($\hat{d} = \hat{p} - \hat{q}$), то

$$m_{\hat{d}} = \sqrt{m^2_{\hat{p}} + m^2_{\hat{q}}}.$$

Но если имеются две группы данных (n_1 и n_2) и в них получены оценки частот $\hat{p}_1$ и $\hat{p}_2$ ($\hat{q}_1$ и $\hat{q}_2$), то для расчета надежности разности частот, ($\hat{p}_1 - \hat{p}_2$) или ($\hat{q}_1 - \hat{q}_2$), используют оценку вариансы в единой совокупности, из которой были взяты обе выборки ($n = n_1 + n_2$). Для такой совокупности вычисляют собственные частоты:

$$\hat{p} = \frac{n_1\,\hat{p}_1 + n_2\,\hat{p}_2}{n} \quad \text{и} \quad \hat{q} = 1 - \hat{p}.$$

Так как $\hat{\sigma}^2_{\text{альт}} = \hat{p}\,\hat{q}$, то ошибки $\hat{p}_1$ и $\hat{p}_2$ рассчитывают из

$$\dot{m}_{\hat{p}_1} = \sqrt{\frac{\hat{p}\,\hat{q}}{n_1}} \quad \text{и} \quad \dot{m}_{\hat{p}_2} = \sqrt{\frac{\hat{p}\,\hat{q}}{n_2}}.$$

Тогда ошибка разности частот, $\hat{d} = \hat{p}_1 - \hat{p}_2$, будет:

$$\dot{m}_{\hat{d}} = \sqrt{\dot{m}^2_{\hat{p}_1} + \dot{m}^2_{\hat{p}_2}}$$

и значение наблюдаемого t-критерия Стьюдента:

$$t_{\hat{d}} = |\hat{d}| / \dot{m}_{\hat{d}}.$$

Пример 11.9. Допустим, что в стаде истобенского скота из 200 коров 80 коров имело белую полосу на хребте, а у 120 голов ее не было.

Тогда: $\hat{p} = 80/200 = 0,4$ и $\hat{q} = 1 - 0,4 = 0,6$;

$$\hat{\sigma} = \sqrt{0,4 \times 0,6} = 0,49 \quad \text{и} \quad m = 0,49/\sqrt{200} = 0,03.$$

Следовательно, $p = 0,4 \pm 0,03$ и $q = 0,6 \pm 0,03$, что является статистически значимым (ошибки меньше частот более, чем в три раза). Истинная частота белохребтовости с вероятностью 95% находилась в границах от 0,33 до 0,47.

В абсолютных единицах:

$$m = \sqrt{80 \times 120/200} = 6,9 \ \text{голов}.$$

Следовательно, белохребтовых 80±6,9, остальных - 120±6,9 голов.

Пример 11.10. Имелось две группы коров n_1=284 и n_2=50. Прореагировало на туберкулез в 1-ой группе M_1=83, во 2-ой группе M_2=6 животных. Требовалось определить статистическую значимость различия в реакции на туберкулез.

Различие в частоте животных, прореагировавших на туберкулез:

$$\hat{p}_1 = 83/284 = 0,29 \quad \text{и} \quad \hat{p}_2 = 6/50 = 0,12.$$

$$\hat{d} = \hat{p}_1 - \hat{p}_2 = 0,29 - 0,12 = 0,17.$$

Общая частота животных, прореагировавших на туберкулез:

$$\hat{p} = \frac{83+6}{284+50} = 0,27 \quad \text{и} \quad \hat{q} = 1 - \hat{p} = 0,73.$$

Ошибки для $\hat{p}_1$ и $\hat{p}_2$:

$$\dot{m}_{\hat{p}_1} = \sqrt{\frac{0,27 \times 0,73}{284}} = 0,026 \quad \text{и} \quad \dot{m}_{\hat{p}_2} = \sqrt{\frac{0,27 \times 0,73}{50}} = 0,063.$$

Тогда

$$m_{\hat{d}} = \sqrt{0,026^2 + 0,063^2} = 0,068 \ \text{и}$$

$$t_{\hat{d}} = 0,17/0,068 = 2,5.$$

По табл. А8 Приложения А критические значения для двусторонней области при $df = n_1 + n_2 - 2 = 334 - 2 = 332$ было $t_{0,05;332} = 1,982$. Так как $t_{\hat{d}} > t_{0,05;332}$, то при уровне значимости α=0,05 нулевая гипотеза может быть отвергнута.

Вывод: Различия в реакции коров двух групп на туберкулез были существенны и статистически значимы. Риск ошибиться в выводах составлял менее 5%.

В заключение данной главы два замечания относительно точности вычисления и правиле округления:

1. Среднее вычисляют с точностью, на порядок большей, чем точность измерения; стандартное отклонение и ошибку - с точностью, на порядок большей, чем точность вычисления среднего.
2. Если округляемая цифра меньше 5, то ее отбрасывают; если больше или равна 5, то в предыдущий разряд добавляют единицу.

Дополнения:

Принцип практической уверенности. Если вероятность события А в данном испытании очень мала, $P(A)=\alpha<<1$, то при *однократном выполнении испытания* можно быть уверенным в том, что событие А не произойдет, и в практической деятельности вести себя так, как будто событие А вообще невозможно.

Этот принцип нельзя доказать математически; он подтверждается практическим опытом человеческой деятельности. Например, выходя из подъезда дома мы не задумываемся о возможности гибели от упавшего с крыши кирпича, хотя некоторая (очень малая) вероятность такого события все же имеется.

Принцип практической уверенности сформулирован относительно «*однократного выполнения испытания*». При n испытаниях вероятность реализации события А повышается в n раз, становится равной $n\alpha$. В это случае считать событие А маловероятным практически невозможно.

Вопрос о том, насколько мала должна быть вероятность α события А, чтобы его можно было считать практически невозможным, выходит за рамки математической теории и решается с учетом важности последствий, вытекающих из наступления события А. В одних случаях можно пренебречь событиями, имеющих вероятность менее 0,05 или 5%. Если же речь идет, например, о здоровье животных, то нельзя пренебрегать событиями, которые могут происходить с вероятностью, равной 0,001 или 0,1%.

В компьютерных статистических программах при выборочной оценке *любого* параметра рассчитывается **точное** значение уровня значимость (обозначается, как «*p-valio*», или «*p-value*», или «*p-level*»). Поэтому необходимость в расчете t- и F-статистик и сравнении их с критическими значениями отпадает. В этом случае проверку H_0 осуществляют так:

- при $p > 0{,}10$ - гипотезу H_0 принимают;
- при $0{,}10 \geq p > 0{,}05$ - результат считают неопределенным или говорят о «имеющей место тенденции»;
- при $p \leq 0{,}05$ - гипотезу H_0 отклоняют, причем:
  - если $0{,}05 \geq p \geq 0{,}01$, то оценку считают ***значимой***;
  - если $0{,}01 > p > 0{,}001$, то оценку считают ***умеренно значимой***;
  - если $p \leq 0{,}001$, то оценку считают ***высокозначимой***.

Для двусторонних гипотез значение *p-valio* удваивают!

Двусторонняя критическая область необходима тогда, когда при сравнении выборочной оценки и параметра генеральной совокупности требуется оценить, *в **общем**,* значимость или незначимость различий между ними. В этом случае учитываются отклонения и в том и в другом направлении. Использование двустороннего критерия несколько повышает надежность выводов, когда предположения, лежащие в основе применения, например, t-критерия, выполняются не совсем точно.

Односторонний критерий считается более адекватным в случаях, когда необходимо проверить гипотезу о возможном направлении полученных отклонений. Так, если надо убедиться в том, что одна величина *в среднем* строго *больше* или *меньше* другой, то используют одностороннюю критическую область (право- или левостороннюю).

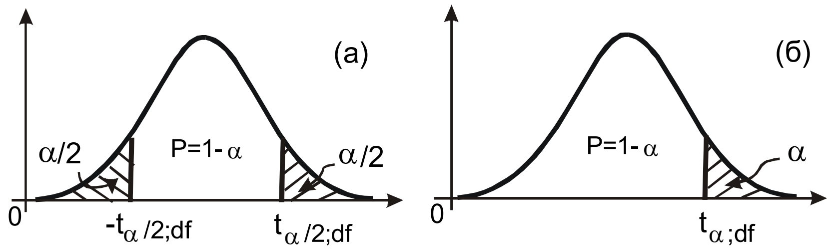
Проверка статистической гипотезы не дает логического доказательства ее верности или неверности.

Принятие (отклонение) гипотезы H_0 следует расценивать не как раз и навсегда установленный, абсолютно верный содержащийся в ней факт, а лишь как достаточно правдоподобное, не противоречащее *данному* эксперименту утверждение.

Из кн.: Кузнецов В.М. **Основы научных исследований в животноводстве.** Киров: Зональный НИИСХ Северо-Востока, 2006.- 568 с.

А.8. Критические значения t-распределения Стьюдента

(здесь и далее Р - доверительная вероятность)



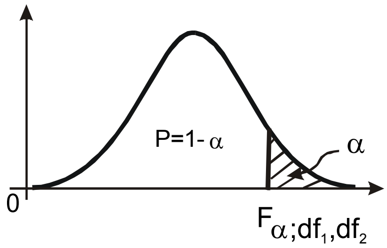
df	Уровень значимости (ошибка, α)				
	Двусторонняя критическая область (а)				
	0,100	**0,050**	**0,020**	**0,010**	**0,001**
	Односторонняя критическая область (б)				
	0,050	**0,025**	**0,010**	**0,005**	**0,0005**
1	6,314	12,706	31,821	63,657	637
2	2,920	4,303	6,965	9,925	31,598
3	2,353	3,182	4,541	5,841	12,941
4	2,132	2,776	3,747	4,604	8,610
5	2,015	2,571	3,365	4,032	6,859
6	1,943	2,447	3,143	3,707	5,959
7	1,895	2,365	2,998	3,499	5,405
8	1,860	2,306	2,896	3,355	5,041
9	1,833	2,262	2,821	3,250	4,781
10	1,812	2,228	2,764	3,169	4,587
11	1,796	2,201	2,718	3,106	4,437
12	1,782	2,179	2,681	3,055	4,318
13	1,771	2,160	2,650	3,012	4,221
14	1,761	2,145	2,624	2,977	4,140
15	1,753	2,131	2,602	2,947	4,073
20	1,725	2,086	2,528	2,845	3,850
25	1,708	2,060	2,485	2,787	3,725
30	1,697	2,042	2,457	2,750	3,646
35	1,690	2,030	2,432	2,724	3,591
40	1,684	2,021	2,408	2,704	3,551
50	1,676	2,008	2,384	2,678	3,496
100	1,661	1,982	2,360	2,625	3,390
∞	1,645	1,960	2,326	2,576	3,291

Примечание. В последней строке даны значения нормированной случайной величины $t = u \sim N(0;1)$.

Из кн.: Кузнецов В.М. **Основы научных исследований в животноводстве.** Киров: Зональный НИИСХ Северо-Востока, 2006.- 568 с.

А.10. Критические значения F-распределения Фишера-Снедекора (α=0,05)

df_1 - число степеней свободы большего среднего квадрата (вариансы);
df_2 - число степеней свободы меньшего среднего квадрата.



df_2	Уровень значимости α=0,05												
	df_1												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**20**	**30**	**∞**
1	161	200	216	225	230	234	237	239	241	242	248	250	254
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,45	19,46	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,66	8,62	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,94	5,80	5,75	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,56	4,50	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,87	3,81	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,44	3,38	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,15	3,08	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	2,94	2,86	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,77	2,70	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,65	2,57	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,54	2,47	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,46	2,38	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,39	2,31	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,33	2,25	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,28	2,19	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,23	2,15	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,19	2,11	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,16	2,07	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,12	2,04	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,10	2,01	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,07	1,98	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,05	1,96	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,03	1,94	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,01	1,92	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	1,99	1,90	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	1,97	1,88	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	1,96	1,87	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	1,94	1,85	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	1,93	1,84	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	1,84	1,74	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,75	1,65	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,66	1,55	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,57	1,46	1,00

Из кн.: Кузнецов В.М. **Основы научных исследований в животноводстве.** Киров: Зональный НИИСХ Северо-Востока, 2006.- 568 с.

А.11. Критические значения F-распределения Фишера-Снедекора (α=0,01)

df_1-число степеней свободы большего срднего квадрата (варианcы);
df_2-число степеней свободы меньшего срднего квадрата (варианcы).

df_2	Уровень значимости α=0,01												
	df_1												
	1	2	3	4	5	6	7	8	9	10	20	30	∞
1	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6209	6261	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,45	99,47	99,50
3	34,12	30,82	29,46	28,71	28,42	27,91	27,67	27,49	27,35	27,23	26,69	26,50	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,02	13,84	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,55	9,38	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,40	7,23	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,16	5,99	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,36	5,20	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	4,81	4,65	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,41	4,25	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,10	3,94	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	3,86	3,70	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,66	3,51	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,51	3,35	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,37	3,21	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,26	3,10	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,16	3,00	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,08	2,92	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,61	3,52	3,43	3,00	2,84	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	2,94	2,78	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	2,88	2,72	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	2,83	2,67	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	2,78	2,62	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	2,74	2,58	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,70	2,54	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,66	2,50	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,63	2,47	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,60	2,44	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,57	2,41	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,55	2,39	2,01
40	7,31	5,18	4,31	3,85	3,51	3,29	3,12	2,99	2,89	2,80	2,37	2,20	1,80
60	7,08	4,98	4,13	3,63	3,34	3,12	2,95	2,82	2,72	2,63	2,20	2,03	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,03	1,86	1,38
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	1,88	1,70	1,00