13. АНАЛИЗ ПОДСЧЕТОВ

Относительно часто возникает потребность анализа альтернативных признаков с дискретной изменчивостью. Кроме того, при изучении количественных признаков иногда более удобно классифицировать животных на группы, например, по уровню продуктивности, возрасту и т.п. В таких случаях исходными являются данные, полученные на основе подсчетов животных, принадлежащих к разным группам (т.е. пропорции). Эти данные обычно используются для решения следующих задач:

- проверки гипотез при внешнем отсутствии предполагаемых закономерностей в данных учета;
- выявления принадлежности выборок к одной совокупности;
- установления сопряженности между переменными.

При проведении исследований с подобными задачами, находят соответствие экспериментальных (наблюдаемых, фактических) пропорций с теоретически ожидаемыми, согласно принятой гипотезе. Если экспериментальные пропорции соответствуют теоретическим, то это свидетельствует об истинности гипотезы. В противном случае гипотезу отвергают и выдвигают новую.

Степень соответствия наблюдаемых пропорций ожидаемым различна: от очень небольшой и чисто случайной, до достаточно значительной. Для ее измерения используют *критерий* χ^2 (хи-квадрат), который был предложен К. Пирсоном в 1900 году. χ^2 называют также критерием *соответствия* или *согласия*; относится к семейству *непараметрических* критериев.

13.1. Критерий χ^2

Формула расчета критерия χ^2 проста:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i},$$

где O_i - экспериментальная пропорция для i-го класса; E_i - соответствующая теоретическая пропорция; k - число классов (≥2).

Отклонения экспериментальных пропорций от теоретических возводят в квадрат. Поэтому значения χ^2 только положительные с асимметричным распределением (см. раздел 8.5).

Если экспериментальные и теоретические числа полностью совпадут, то тогда значение χ^2 равно нулю. По мере увеличения разницы между наблюдаемыми и ожидаемыми числами значение χ^2 возрастает. В принципе, значения χ^2 могут возрастать до бесконечности. Соответственно этому вероятности их появления убывают от 1 до 0 (или от 100% до 0).

Для критерия χ^2 число степеней свободы (df) - это то число классов (или групп), пропорции которых могут принимать любые значения, не связанные с наблюдаемыми пропорциями. В простейших случаях df=k-1.

При вычислении χ^2 необходимо, чтобы в каждом из эмпирических или теоретических классов было не менее 5 наблюдений. Если наблюдений меньше 5, то пропорции смежных классов объединяют.

В случае если проводят несколько опытов по одному и тому же вопросу, то можно вычислить частные χ^2 для каждого отдельного опыта. Затем рассчитывают обобщенный критерий χ^2 по всем опытам путем простого суммирования частных χ^2. Число степеней свободы также будет равно сумме степеней свободы складываемых χ^2. Так, например, если в каждом отдельном опыте df=1, а опытов было 5, то число степеней свободы для обобщенного χ^2 равно 5.

С другой стороны, можно обработать весь материал в целом: игнорируя отдельные опыты определить соответствующие фактические пропорции, рассчитать теоретически ожидаемые и по ним вычислить критерий χ^2.

Сравнение χ^2, полученных двумя разными способами объединения экспериментального материала, позволит судить о степени его однородности (или неоднородности), т.е. проверить гипотезу о принадлежности выборок к одной совокупности.

13.2. Проверка нулевой гипотезы

Критерий χ^2 используют для проверки гипотезы H_0 - имеет ли место согласие между экспериментальными и теоретическими пропорциями; альтернатива H_1 заключается в том, что отклонения наблюдаемых пропорций от теоретических значительны.

Общая схема применения критерия χ^2 для проверки нулевой гипотезы:

1. Определяют меру расхождения наблюдаемых и теоретических пропорций, т.е. рассчитывают χ^2.

2. Для выбранного уровня значимости α по табл. А.9 Приложения А находят критическое значение $\chi^2_{\alpha;df}$ при числе степеней свободы df.

3. Если вычисленное значение χ^2 меньше табличного, т.е. $\chi^2 < \chi^2_{\alpha;df}$, то нулевую гипотезу принимают на уровне значимости α. В противном случае ($\chi^2 \geq \chi^2_{\alpha;df}$) нулевую гипотезу отклоняют и принимают альтернативную.

В исследованиях по биологии имеет место следующее правило: если при df=1 вычисленный χ^2>3,84, то расхождения между наблюдаемыми и теоретическими пропорциями статистически значимы при α=0,05. Следовательно, гипотеза H_0, исходящая из теоретического распределения, неверна и должна быть отброшена.

Если вычисленное значение χ^2 очень мало, то это свидетельствует об очень большой вероятности согласия гипотезы с экспериментальными данными. Однако в то же время это означает, что вероятность случайного отклонения наблюдаемых данных от теоретических значений очень мала, т.е. получено *слишком хорошее* согласие с гипотезой H_0. Исследователь должен иметь это в виду и при систематическом получении малых значений χ^2 проверить методику эксперимента с целью обнаружения факторов, искусственно снижающих случайную изменчивость (например, выбраковка «нежелательных» результатов).

Формула расчета χ^2 проста и ее применение не вызывает затруднений. Проблемы возникают при определение ожидаемых пропорций. Поэтому ниже даны примеры, на которых демонстрируются приемы их расчета и нахождения критерия χ^2 [100,114].

13.3. Анализ расщепления

Наиболее широко критерий χ^2 применяют в исследованиях по генетике животных, когда нужно проверить соответствие пропорций классов, получаемых при расщеплении, свободном комбинировании или сцеплении, с пропорциями, теоретически ожидаемыми по той или иной генетической гипотезе. В этом случае для вычисления теоретических пропорций надо помножить общее число фактически изучаемых особей на соответствующую долю, теоретически возможную при данном типе расщепления.

Пример 13.1. При скрещивании серых гетерозиготных каракульских овец получено 310 ягнят серой окраски и 90 - черной. При обычном моногибридном скрещивании отношение должно быть 3:1, т.е. 300 серых и 100 черных ягнят. Для этого примера

$$\chi^2 = \frac{(310-300)^2}{300} + \frac{(90-100)^2}{100} = 1{,}33.$$

Групп животных только две, поэтому df=1. В первой строке табл. А.9 Приложения А находим критическое значение для уровня значимости α=0,05 - $\chi^2_{0,05;1}$=3,84. Так как 1,33<3,84, то адекватность фактических данных ожидаемым достаточно велика. Гипотеза о том, что имело место расщепление ягнят по окраске в отношении 3 к 1, подтверждается.

Пример 13.2. При скрещивании короткоухих овец (гетерозигот, полученных путем скрещивания нормальных длинноухих овец с овцами, лишенными наружного уха) было получено 22 потомка, в том числе: 7 овец с нормальными ушами, 9 - короткоухих и 6 - безухих. Так как гетерозиготы по длине ушей у овец фенотипически отличаются от гомозиготных форм, то в F_2 ожидалось расщепление 1:2:1 (0,25:0,50:0,25). Сопоставление фактических результатов с ожидаемыми дано в табл. 16.

16. Вычисление критерия χ^2 для данных о расщеплении по длине ушей у овец

Пропорции		(O-E)	(O-E)^2	(O-E)^2/E
факт.(O)	теорет.(E)			
7	22×0,25= 5,5	1,5	2,25	0,410
9	22×0,50=11,0	-2,0	4,00	0,364
6	22×0,25= 5,5	0,5	0,25	0,045
Σ=22	Σ=22	-	-	Σ=0,819

При числе степеней свободы df=3-1=2 и уровне значимости α=0,05 критическое значение $\chi^2_{0,05;2}$=5,99 (табл. А.9 Приложения А). Таким образом, исходную гипотезу о расщеплении овец F_2 по длине ушей в отношении 1:2:1 можно считать правильной.

По аналогичному принципу проводят вычисления ожидаемых пропорций для более сложных типов расщепления, например 9:3:3:1; 12:3:1 и т.д.

Пример 13.3. При обычном дигибридном скрещивании наблюдалось расщепление 168 животных по фенотипу на 4 группы: АВ - 117, Аb - 26, аВ - 18 и ab -7 голов. Ожидаемое расщепление 9:3:3:1. Расчеты ожидаемых пропорций и χ^2 даны в табл. 17.

17. Вычисление критерия χ^2 для данных о расщеплении при дигибридном скрещивании

Фенотип	Пропорции		$(O-E)^2/E$
	O	E	
AB	117	(9/16)168=94,5	$(117-94,5)^2/94,5$ =5,36
Ab	26	(3/16)168=31,5	$(26-31,5)^2/31,5$ =0,96
aB	18	(3/16)168=31,5	$(18-31,5)^2/31,5$ =5,79
ab	7	(1/16)168=10,5	$(7-10,5)^2/10,5$ =1,17
Σ	168	168	13,28.

$\chi^2_{0,05;3} = 7,81$. Значение $\chi^2 = 13,28$ дало основание отвергнуть нулевую гипотезу при 5% уровне ошибки и считать, что имело место статистически значимое отклонение от ожидаемого отношения 9:3:3:1.

Пример 13.4. Имелось 5 пометов кроликов второго поколения (F_2), происходящих от скрещивания F_1 (табл. 18). Требуется вычислить χ^2 для определения того, совпадает ли наблюдаемое распределение животных по пигментации с теоретическим в соответствии с правилом Менделя о соотношении доминантных и рецессивных особей в F_2, равном 3:1.

Число df для каждого частного значения χ^2 равно 1, т.к. для каждого помета имелось 2 класса ($k=2$). Как следует из табл. 18, только для второго помета соотношение доминантных и рецессивных особей не соответствовало теоретическому (χ^2=4>3,84 при df=1).

Суммарный по пометам χ^2=6,175. При df=5 и уровне значимости α=0,05 критическое значение $\chi^2_{0,05;5}$=11,07. Таким образом, можно считать, что наблюдаемое соотношение крольчат недостоверно отличалось от теоретического (3:1). Полученные в опыте данные соответствовали гипотезе (правилу Менделя о расщеплении признаков в F_2).

18. Вычисление критерия χ^2 для данных о расщеплении кроликов F_2 по пигментации

№ помета	Крольчат, O		Σ	Крольчат, E		(O-E)²/E для		χ^2	df
	черных	белых		черных	белых	черных	белых		
1	6	4	10	7,5[1)]	2,5	0,300[2)]	0,900	1,200	1
2	6	6	12	9,0	3,0	1,000	3,000	4,000	1
3	7	1	8	6,0	2,0	0,166	0,500	0,666	1
4	5	1	6	4,5	1,5	0,055	0,160	0,215	1
5	10	4	14	10,5	3,5	0,023	0,071	0,094	1
Σ	34	16	50	-	-	-	-	6,175	5

Примечание. 1) 7,5=0,75×10; 2) 0,3=(6-7,5)²/7,5.

По объединенным данным:

$$\chi^2 = \frac{(34 - 0{,}75 \times 50)^2}{0{,}75 \times 50} + \frac{(16 - 0{,}25 \times 50)^2}{0{,}25 \times 50} = 1{,}31\,.$$

При df=1 критическое значение $\chi^2_{0,05;\,1}$=3,84. Таким образом, и по объединенным данным нет оснований отвергать нулевую гипотезу.

Между значениями χ^2, вычисленными двумя способами, наблюдалось различие, которое должно было указывать на степень однородности или неоднородности опытных данных:

	df	χ^2
По 5 пометам	5	6,125
По итоговым данным	1	1,310
Неоднородность	4	4,815

Для df=4 и α=0,05 критическое значение $\chi^2_{0,05;\,4}$=9,49. Так как 4,815<9,49, то неоднородность при уровне значимости α=0,05 не была доказана. Можно считать, что пометы однородны, и что лучшей оценкой действительного отношения крольчат по пигментации было отношение 34:16.

13.4. Анализ многопольных таблиц

Критерий χ^2 можно применять для анализа влияния различных факторов на те или иные биозоотехнические процессы и явления. Данные подобных опытов обычно группируют в *таблицы сопряженности*, состоящие из нескольких полей (2×2, 2×3, 4×4 и т.д.). Исходной нулевой гипотезой, которая должна быть или отвергнута после определения χ^2 или, наоборот, сохранена, является принятие независимости в вариации изучаемых показателей, отсутствие различий во влиянии тех или других факторов.

Пример 13.5. В табл. 19 представлены данные по иммунизации телят от туберкулеза. Исходная нулевая гипотеза заключалась в утверждении, что прививка не влияет на число заболевших телят, т.е. заболеваемость в обеих группах одинаковая (как по всему материалу).

Теоретические значения находят из пропорций итоговых цифр 60 и 40 (=100). Так, для левой верхней клетки теоретически ожидаемое значение было равно 37(60/100)=22,2, а для верхней правой - 37(40/100)=14,8. Аналогично находят ожидаемые пропорции для нижних клеток - 63 животных контрольной группы распределены пропорционально 60 и 40.

19. Таблица сопряженности 2×2
для данных по иммунизации телят

Группа	Число животных		Σ
	заболело	не заболело	
Опытная (прививка)	9 (22,2) [a]	28 (14,8) [b]	[a+b]= 9+28=37
Контрольная (без прививки)	51 (37,8) [c]	12 (25,2) [d]	[c+d]=51+12=63
Σ	[a+c]=9+51=60	[b+d]=28+12=40	[a+b+c+d] = 100

Примечание. В круглых скобках теоретические значения.

Критерий χ^2 рассчитывают по общей формуле:

$$\chi^2 = \frac{(9-22{,}2)^2}{22{,}2} + \frac{(28-14{,}8)^2}{14{,}8} + \frac{(51-37{,}8)^2}{37{,}8} + \frac{(12-25{,}2)^2}{25{,}2} = 31{,}14.$$

Так как данные были представлены в виде таблицы со строками (k_1) и столбцами (k_2), то $df = (k_1 - 1)(k_2 - 1)$=(2-1)(2-1)=1.

χ^2=31,14>$\chi^2_{0,05;1}$=3,84, поэтому нулевую гипотезу отвергают, т.е. предположение об отсутствии влияния прививки на заболеваемость телят было неверным. Критерий χ^2 выявил статистически значимые различия между частотой заболеваемости в опытной и контрольной группах.

Для таблицы 2×2, пропорции в которой обозначены как a, b (для верхних полей), c и d (для нижних), а общая сумма a+b+c+d=n, можно применить упрощенную формулу расчета χ^2:

$$\chi^2 = \frac{(ad - bc)^2 n}{(a+b)(c+d)(a+c)(b+d)}.$$

Для рассматриваемого примера получим:

$$\chi^2 = \frac{(9\times12-51\times28)^2 100}{37\times63\times60\times40} = 31{,}14,$$

что соответствует результату основной формулы.

Распределение χ^2, как это видно из табл. А.9 Приложения А, является непрерывным. Распределение же групп в таблицах, подобных табл. 19, дискретно. Поэтому при оценке χ^2 имеет место некоторая неточность. Для уменьшения неточности в упрощенную формулу χ^2 вводят поправку Иэйтса на непрерывность:

$$\chi^2 = \frac{[|ad-bc|-0{,}5\times n]^2 n}{(a+b)(c+d)(a+c)(b+d)},$$

(выражение $|ad-bc|$ обозначает абсолютную величину ad-bc).

Смысл поправки в том, что наблюдаемые частоты, которые больше ожидаемых, увеличивают на 0,5, а наблюдаемые частоты, которые меньше ожидаемых, уменьшают на эту же величину.

Поправка Иэйтса всегда занижает значение χ^2. Так, для примера 13.5 получим:

$$\chi^2 = \frac{(|9\times12-51\times28|-0{,}5\times100)^2 100}{37\times63\times60\times40} = 28{,}83.$$

Поправка Иэйтса вводится и в основную формулу χ^2:

$$\chi^2 = \sum_{i=1}^{k} \frac{(|O_i - E_i|-0{,}5)^2}{E_i}.$$

Поправку Иэйтса следует использовать всегда, когда df=1.

Пример 13.6. В табл. 20 представлены данные по влиянию возраста птицы на оплодотворяемость яиц.

20. Таблица сопряженности 4×2 для данных о влиянии возраста петухов и кур на оплодотворяемость яиц

Группа	Возраст		Число яиц		Σ
	отцов	матерей	оплодот. факт.(теор.)	неоплодот. факт.(теор.)	
1	молодые	молодые	35(33,20)	5(6,80)	40
2	старые	молодые	53(49,80)	7(10,20)	60
3	молодые	старые	48(45,65)	7(9,35)	55
4	старые	старые	30(37,35)	15(7,65)	45
Σ			166	34	200

Примечание. В круглых скобках теоретически ожидаемые значения.

Опыт включал четыре класса по сочетаниям возраста (k_1=4) и два класса по состоянию яиц (k_2=2), поэтому df=(4-1)(2-1)=3.

Теоретические пропорции для каждой клетки таблицы находят исходя из нулевой гипотезы, утверждающей, что возрастной подбор родителей не влияет на оплодотворяемость яиц. Поэтому итоги по классам возрастного подбора (40, 60, 55 и 45) выражались пропорционально итогам по показателю оплодотворяемости (т.е. числам 166 и 34), соотношение между которыми (0,83 и 0,17) можно ожидать, согласно гипотезе, в каждой группе возрастного подбора (например, для первой строки 40×0,83=33,2 и 40×0,17=6,8).

Полученный по основной формуле критерий χ^2 составил 11,0, а табличное значение - $\chi^2_{0,05;\,3}$=7,81. Следовательно, наблюдаемые соотношения оплодотворенных и неоплодотворенных яиц статистически значимо отличались от теоретических. Полученные в опыте данные при уровне значимости α=0,05 не соответствовали исходной гипотезе. Возрастной подбор пар влиял на оплодотворяемость яиц.

13.5. Соответствие эмпирического распределения теоретическому

После составления вариационного ряда, вычисления средней арифметической и стандартного отклонения - часто возникает необходимость установить, насколько фактическое распределение признака соответствует тому или другому из известных теоретических распределений (нормальному, биномиальному, Пуассона). Критерий χ^2 позволяет установить степень такого соответствия. Некоторые трудности возникают при нахождении ожидаемой численностей каждого класса вариационного ряда.

Пример 13.7. Пусть имеется вариационный ряд по количеству молочного жира 300 первотелок (табл. 21) со следующими статистическими показателями: $\overline{x}$=134,3 кг, $\hat{\sigma}$=21,3 кг. Задача вычисления теоретических пропорций сводится к тому, чтобы отнести к уже имеющимся классам возможные значения пропорций, если они распределены по законам нормального распределения.

Для нахождения теоретических пропорций, необходимо трансформировать имеющиеся классы, выраженные в кг, в классы, выраженные в долях стандартного отклонения (сигмы, σ). После этого рассчитывают, сколько первотелок должно приходиться на каждый данный отрезок нормальной кривой, ограниченный определенными значениями сигмы.

Возьмем, например, класс с центральным значением «100». Границы этого класса 95,0 и 104,9. В значениях (терминах) сигмы они будут:

$$u_1 = \frac{95,0 - 134,3}{21,3} = -1,845 \quad \text{и} \quad u_2 = \frac{104,9 - 134,3}{21,3} = -1,380\,.$$

Какая же доля общего числа животных при нормальном распределении должна находиться в интервале между $-1,845\hat{\sigma}$ и $-1,380\hat{\sigma}$?

21. Соответствие эмпирического и теоретического распределения 300 первотелок по количеству жира

Центр. значения классов	Пропорции			(O-E)	$(O-E)^2$	$(O-E)^2/E$
	факт., O	теор., E				
80	1	2,17				
90	2	6,45	8,6	-5,6	31,36	3,646
100	17	15,3		1,7	2,89	1,889
110	39	29,3		9,7	94,09	3,211
120	44	44,8		-0,8	0,64	0,014
130	66	55,0		11,0	121,00	2,200
140	42	54,2		-12,2	148,84	2,746
150	34	43,0		-9,0	81,00	1,884
160	29	27,3		1,7	2,89	0,106
170	18	13,9		4,1	16,81	1,209
180	3	5,68				
190	3	1,86	8,0	0,0	0,00	0,000
200	2	0,49				
Σ	n=300	299,45		-	-	16,905

В табл. А.4 Приложения А даны накопленные частоты (последующие частоты прибавлены к предыдущим) для каждого отрезка нормальной кривой. u_1=1,845 соответствует величина 0,4673, а u_2=1,380 - величина 0,4162. Тогда *доля* особей в интервале между двумя значениями накопленных частот составит 0,4673-0,4162=0,0511. Ожидаемое *число* первотелок для данного класса при n=300 должно быть 0,0511×300=15,33. Аналогично находят теоретически ожидаемые пропорции для всех других классов (столбец 3 табл. 21).

Критерий χ^2 требует наличие в каждом из эмпирических и теоретических классов не менее 5 животных. Поэтому две верхние и три нижние строчки обоих рядов объединяют.

Число степеней свободы составило 10-3=7, так как теоретический и фактический ряды имели 3 общих элемента: n, σ и $\bar{x}$.

Полученное по общей формуле значение χ^2=16,905 при df=7 было выше критического в столбце с α=0,05, но ниже критического при α=0,01 (табл. А.9 Приложения А). Следовательно, соответствие экспериментального ряда теоретическому нормальному было недостаточным, так как границей соответствия принято считать уровень значимости 0,05.

Пример 13.8. В табл. 22 представлены результаты сравнения распределения помесных коров по жирномолочности с известным теоретическим распределением, полученном на чистопородных коровах.

22. Сравнение эмпирического ряда с известным теоретическим распределением

Ср. % жира	Пропорции		$(O-E)^2/E$
	факт., О	теор., Е.	
3,5 3,8	2 3 } 5	3 4 } 7	$(5-7)^2\,2/7 = 0,57$
4,1	5	5	$(5-5)^2\,2/5 = 0,00$
4,4	5	9	$(5-9)^2\,2/9 = 1,77$
4,7	10	21	$(10-21)^2\,2/21 = 5,76$
5,0	20	9	$(20-9)^2\,2/9 = 13,44$
5,3	10	5	$(10-5)^2\,2/5 = 5,00$
6,3 6,9	5 3 } 8	4 3 } 7	$(8-7)^2\,2/7 = 0,14$
Σ	63	63	26,68

Критерий χ^2, рассчитанный по основной формуле, составил 26,68. При df=7-2=5 и уровне значимости α=0,05 критическое значение $\chi^2_{0,05;\,5}$=11,07 (табл. А.9 Приложения А). Таким образом, распределение помесных животных по жирномолочности статистически значимо отличалось от теоретического по чистопородным животным.

Пример 13.9. В табл. 23 дано распределение числа хрячков в пометах свиноматок, в каждом из которых было по 6 поросят.

23. Сравнение наблюдаемого распределения числа хрячков в пометах с теоретически ожидаемым

Число хрячков в помете	Число пометов		О-Е	$(O-E)^2$	$(O-E)^2/E$
	факт.,О	теорет.,Е			
0 1	3 16 } 19	3,45 20,72 } 24,17	-5,17	26,73	1,11
2	53	51,80	+1,20	1,44	0,03
3	78	69,06	+8,94	79,92	1,16
4	53	51,80	+1,20	1,44	0,03
5 6	10 8 } 18	20,72 3,45 } 24,17	-6,17	38,07	1,57
Σ	221	221	-	-	χ^2 =3,90

В данном случае имела место дискретная изменчивость. Поэтому следует ожидать биномиальное распределение, в котором ряд описывается биномом $(p+q)^m$. Наблюдаемый ряд имел 7 классов, следовательно m=6. При равном соотношении полов p=q=0,5 теоретически ожидаемые доли классов будут (можно использовать треугольник Паскаля) 0,015625, 0,09375, 0,234375, 0,3125, 0,234375, 0,09375, 0,015625. Теоретически ожидаемые пропорции были получены путем перемножения этих долей на 221 - общее число пометов (3 столбце табл. 23).

Перед нахождением χ^2 необходимо объединить две верхние строчки в один класс и две нижние строчки. Дальнейшие вычисления проводят по основной формуле. Число степеней свободы в биномиальном случае равно df=k-2=5-2=3.

Фактическое значение χ^2=3,9 не превысило критическую точку $\chi^2_{0,05;\,3}$=7,81. Таким образом, при уровне значимости α=0,05 имело место относительно хорошее соответствие наблюдаемого биномиального распределения числа хрячков в помете с теоретически ожидаемым.

Пример 13.10. Пуассоновское распределение относится к явлениям, обладающим очень малой вероятностью. Поэтому оно асимметрично. Характерным признаком для него является то, что среднее арифметическое ($\bar{x}$ или λ) и варианса численно почти равны.

Теоретически ожидаемые пропорции пуассоновского распределения представляют собой следующий ряд:

$$\frac{n}{e^{\lambda}}\text{(нулевой член)},\ \frac{n\lambda}{e^{\lambda}},\ \frac{n\lambda^2}{2e^{\lambda}},\ \frac{n\lambda^3}{(2)(3)e^{\lambda}},\ \frac{n\lambda^4}{(2)(3)(4)e^{\lambda}}\ \text{и т.д.,}$$

где n - общее число наблюдений в вариационном ряду; e - основание натурального логарифма (≈2,718, а его логарифм при основании 10 равен 0,43429); λ - среднее арифметическое (=$\bar{x}$).

Для удобства расчетов ряд теоретических пропорций можно представить в виде:

$$\frac{n}{e^{\lambda}},\ \left(\frac{n}{e^{\lambda}}\right)\lambda,\ \left(\frac{n\lambda}{e^{\lambda}}\right)\left(\frac{\lambda}{2}\right),\ \left(\frac{n\lambda^2}{2e^{\lambda}}\right)\left(\frac{\lambda}{3}\right),\ \left(\frac{n\lambda^3}{2\times 3e^{\lambda}}\right)\left(\frac{\lambda}{4}\right)\ \text{и т.д.}$$

Тогда достаточно вычислить первый элемент, а все последующие рассчитывают из предыдущих путем умножения на

$$\lambda,\ \tfrac{1}{2}\lambda,\ \tfrac{1}{3}\lambda,\ \tfrac{1}{4}\lambda,\ \text{и т.д.}$$

Допустим, что выборка численностью n=98 имела λ=3,0204 и σ^2=3,0204. Первый элемент ряда равен

$$\frac{98}{e^{3,0204}} = 4,78,$$

второй - 4,78×3,0204=14,44, третий - 14,44×(½)×3,0204=21,81 и т. д.

После вычисления теоретически ожидаемых пропорций для всех классов составляют таблицу, аналогично табл. 23, и рассчитывают критерий χ^2. Если число наблюдений в каком-либо из крайних классов будет меньше 5, то их объединяют с 1-2 соседними. Объединение проводят как по фактическому, так и по теоретически ожидаемому рядам. Число степеней свободы такое же, как и при биномиальном распределении - df=k-2.

13.6. Соответствие двух эмпирических распределений

Иногда возникает потребность сравнения двух наблюдаемых распределений друг с другом. Такое сравнение позволяет ответить на вопрос: принадлежат ли сравниваемые ряды к одной генеральной совокупности или к разным?

Пример 13.11. В табл. 24 даны два эмпирических распределения с одинаковой численностью кур породы леггорн и кур новой породной группы. Допускалось, что теоретическое распределение неизвестно.

24. Сравнение двух эмпирических рядов с одинаковой численностью (n=100)

Живая масса кур, кг	O		E= =½(G_1+G_2)	$(O-E)^2/E$		χ^2_{1+2}
	G_1	G_2		χ^2_1	χ^2_2	
1,2	9	5	7	0,57	0,57	1,14
1,5	10	6	8	0,50	0,50	1,00
1,8	16	8	12	1,33	1,33	2,66
2,1	20	10	15	1,66	1,66	3,32
2,4	16	20	18	0,17	0,17	0,34
2,7	14	14	14	0,00	0,00	0,00
3,0	5	17	11	3,27	3,27	6,54
3,3	5	15	10	2,50	2,50	5,00
3,6	5	5	5	0,00	0,00	0,00
Σ	100	100	100	-		20,00

В качестве пропорций неизвестной генеральной совокупности берут полу суммы пропорций по каждому классу сравниваемых рядов. Критерий χ^2 составил 20. При df=9-1=8 и α=0,05 критическое значение $\chi^2_{0,05;8}$=15,51 (табл. А.9 Приложения А). Нулевая гипотеза не подтвердилась. Живая масса кур новой породной группы при уровне значимости α=0,05 существенно отличалась от таковой кур породы леггорн. Можно считать созданную породную группу качественно иной, принадлежащей к другой генеральной совокупности.

Пример 13.12. В табл. 25 дан анализ соответствия наблюдаемых распределений по плодовитости овец двух выборок разной численности.

Теоретически ожидаемые пропорции для обоих рядов рассчитывают путем умножения наблюдаемых пропорций по объединенной выборке (20, 45, 45, 30, 20) на долю животных данной выборки во всем материале (100/160 или 60/160). Значения $(O\text{-}E)^2/E$, вычисленные для каждого класса обоих рядов, суммированы по рядам в пределах класса. Полученные числа суммированы затем по классам, что дало значение χ^2=4,15.

25. Сравнение 2-х эмпирических рядов с разной численностью

Плодовитость	O для ряда		Σ	E для ряда		$(O\text{-}E)^2/E$ для ряда		Σ
	1	2		1	2	1	2	
1	10	10	20	12,5[1)]	7,50[2)]	0,50[3)]	0,83	1,33
2	25	20	45	28,1	16,87	0,34	0,58[4)]	0,92
3	30	15	45	28,1	16,87	0,13	0,21	0,34
4	20	10	30	18,7	11,25	0,09	0,14	0,23
5	15	5	20	12,5	7,50	0,50	0,83	1,33
Σ	100	60	160	~100	~60	-		4,15

Примечание. 1) 20×(100/160)=12,5; 2) 20×(60/160)=7,5;
3) $(10\text{-}12{,}5)^2$ /12,5=0,50; 4) $(20\text{-}16{,}87)^2$ /16,87=0,58.

Для рассматриваемого примера число степеней свободы есть $df=(k_1-1)(k_2-1)=(5-1)(2-1)=4$ (где k_1 - число классов, т.е. строк; k_2 - число наблюдаемых рядов). По табл. А.9 Приложения А для α=0,05 и df=4 критическая точка была $\chi^2_{0,05;4}$=9,49, т.е. $\chi^2<\chi^2_{0,05;4}$. Таким образом, нулевая гипотеза при уровне значимости α=0,05 подтвердилась и, следовательно, обе выборки можно отнести к одной совокупности.

При проведении исследований может возникнуть ситуация, когда наблюдаемые ряды будут иметь не только разную численность, но и разное число классов. Формула расчета критерия χ^2 в этом случае имеет вид

$$\chi^2=\frac{1}{n_1 n_2}\sum\frac{(f_1 n_2+f_2 n_1)^2}{f_1+f_2},$$

где f_1 и f_2 - пропорции классов первого и второго рядов, а n_1 и n_2 - общее число животных в каждом ряду.

Пример 13.13. Сопоставление двух наблюдаемых рядов с разным числом классов (табл. 26).

26. Вариационные ряды двух выборок телят по живой массе

Вес, кг	Пропорции		f_1n_2	f_2n_1	$f_1n_2+f_2n_1$	$(f_1n_2+f_2n_1)^2$	$\frac{(f_1n_2+f_2n_1)^2}{f_1+f_2}$
	f_1	f_2					
140	1	7	54	532	586	343396	42924,5
141	1	5	54	380	434	188356	31392,7
142	8	14	432	1064	1496	2288016	101728,0
143	3	8	162	608	770	592900	53900,0
144	9	9	486	684	1170	1368900	76050,0
145	13	6	702	456	1158	1340964	70577,1
146	20	3	1080	228	1308	1710864	74385,4
147	6	2	324	152	476	226576	28322,0
148	11	-	594	-	594	352836	32076,0
149	2	-	108	-	108	11664	5832,0
150	2	-	108	-	108	11664	5832,0
	$n_1=76$	$n_2=54$	-	-	-	-	523019,7

Критерий соответствия:

$$\chi^2 = \frac{1}{76\times 54}\times 523019{,}7 = 127{,}44\,.$$

Так как единственным связующим элементом двух рядов было то, что их классы сопоставлялись попарно, то df=11-1=10. Проверка по табл. А.9 Приложения А показала статистически значимое *несоответствие* рядов. Гипотеза, что выборки взяты из одной популяции, не подтвердилась.

13.7. Оценка сопряженности

По данным, полученных путем подсчетов, можно определить также степень связи между двумя переменными. Наиболее удобной нулевой гипотезой является предположение об их независимости. Эта гипотеза базируется на теории вероятностей:

«*Если два события независимы друг от друга, то вероятность одновременного наступления двух событий будет равна произведению вероятностей их отдельного наступления*».

При наличии совместной изменчивости двух альтернативных признаков с двумя классами говорят об их *ассоциации*, взаимосвязности. Если качественные признаки принимают более двух значений, то связь между ними называют *контингенцией*. В обоих случаях выясняют вопрос: встречается ли совпадение присутствия обоих качественных признаков или присутствия одного и, наоборот, отсутствия другого чаще, чем это должно быть по случайным причинам?

Коэффициент ассоциации (Ф, «фи»):

$$\Phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}.$$

По своей конструкции коэффициент Ф соответствует коэффициенту корреляции. Коэффициент Ф принимает значения в интервале от –1 до +1. При a=d=0, Ф= -1; если c=b=0, то Ф= +1.

Пример 13.14. По данным, представленным в табл. 27, необходимо определить наличие связи между заболеваемостью туберкулезом и способом содержания животных (привязным или беспривязным).

27. Число заболевших животных при различных системах содержания

Состояние животного	Содержание животного		Σ
	привязное	беспривязное	
больное	137 [a]	72 [b]	209
здоровое	152 [c]	149 [d]	301
Σ	289	221	510

Коэффициент ассоциации:

$$\Phi = \frac{137 \times 149 - 72 \times 152}{\sqrt{209 \times 301 \times 289 \times 221}} = +0{,}149.$$

По значению Ф можно *предположить* о наличии слабой связи между системой содержания животных и их заболеваемостью туберкулезом.

Было показано, что

$$\Phi = \sqrt{\chi^2 / n}.$$

С поправкой на непрерывность критерий χ^2 составил 11,32. Для уровня значимости $\alpha = 0{,}05$ и df=1 критическое значение $\chi^2_{0,05;\,1} = 3{,}84$ (табл. А.9 Приложения А). Так как $\chi^2 > \chi^2_{0,05;\,1}$, то следует вывод о *статистически значимой* связи между системой содержания и заболеваемостью животных.

Коэффициент контингенции. В том случае, когда вариацию качественного признака можно разбить не на две группы, а на несколько групп, то соответствующий числовой материал располагают в виде таблицы с несколькими строками и столбцами. Такая таблица называется таблицей *контингенции* или m×n-таблицей.

Пример 13.15. Требуется исследовать влияние консерванта на сохраняемость корма. Для этого было взято 1000 проб корма, из которых 300 проб были без добавки, 500 - с незначительной добавкой, и 200 - с большой добавкой консерванта. По степени сохраняемости корм был разбит на группы: плохая, средняя, хорошая и очень хорошая сохраняемость (табл. 28).

28. Сопряженность количества консерванта и сохраняемости корма

Обработка консервантом		Сохраняемость				Σ	
		плохая	средняя	хорошая	отличная		
без консерванта	факт.	51	175	74	-	300	q_1
	теор.	*29,4*	*129,6*	*115,8*	*25,2*		
незначительная добавка	факт	45	210	191	54	500	q_2
	теор.	*49*	*216*	*193*	*42*		
большая добавка	факт	2	47	121	30	200	q_3
	теор.	*19,6*	*86,4*	*77,2*	*16,8*		
Σ		98 Q_1	432 Q_2	386 Q_3	84 Q_4	1000	

Расчет теоретически ожидаемых значений:

- относительная частота (вероятность) для i-ой строки

$$p_i = \frac{q_i}{n}, \quad \text{где } i = 1,2,...,v;$$

- относительная частота для j-го столбца

$$P_j = \frac{Q_j}{n}, \quad \text{где } j = 1,2,...,w;$$

- относительная частота (вероятность) появления значения в i-ой строке и j-м столбце (в i, j-ой клетке таблицы)

$$p_{ij} = p_i \times P_j = \frac{q_i \times Q_j}{n^2}.$$

С учетом распределения сумм пропорций по строкам и столбцам (q_i и Q_j), теоретически ожидаемая пропорция, q'_{ij}, будут:

$$q'_{ij} = n p_{ij} = \frac{q_i Q_j}{n}.$$

Значения $q'_{11}, q'_{12}, ..., q'_{21}, q'_{22}, ...$ даны в клетках табл. 28 (правое выравнивание, курсив) под фактическими частотами (левое выравнивание).

Расчет χ^2, как меры сопряженности признаков:

$$\chi^2 = \sum_{j=1}^{w} \sum_{i=1}^{v} \frac{(q_{ij} - q'_{ij})^2}{q'_{ij}}$$

с df=(v-1)(w-1) степенями свободы (v - число строк, w - число столбцов). Фактическое значение χ^2 сравнивают с критическим значением $\chi^2_{\alpha,df}$, определенным по табл. А.9 Приложения А при заданном уровне значимости α и соответствующем df. Если $\chi^2 \geq \chi^2_{\alpha,df}$, то при уровне значимости α нулевую гипотезу отвергают и утверждают о наличии связи между признаками.

Критерий χ^2 - мера *значимости* связи вообще, но не мера *степени* связи. Мерой степени связи между переменными являются коэффициенты контингенции:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} \quad \text{и} \quad K = \sqrt{\frac{\chi^2}{n(d-1)}} \quad \text{или} \quad K' = \sqrt{\frac{\chi^2}{n\sqrt{df}}},$$

где d - наименьшее из двух чисел: v или w, т.е. числа строк или столбцов.

Значения C и K (K') лежат в границах между 0 и 1: если значение близко к 1, то связь сильная; при 0,30…0,35 - связь слабая; значение близкое к 0 указывает на отсутствие связи.

Применительно к примеру 13.15, получим:

$$q'_{11} = \frac{300 \times 98}{1000} = 29{,}4; \quad q'_{12} = \frac{300 \times 432}{1000} = 129{,}6, \quad \ldots, \quad q'_{34} = \frac{200 \times 84}{1000} = 16{,}8.$$

Отсюда

$$\chi^2 = \frac{(51 - 29{,}4)^2}{29{,}4} + \frac{(175 - 129{,}6)^2}{129{,}6} + \ldots + \frac{(30 - 16{,}8)^2}{16{,}8} = 144{,}997.$$

Число степеней свободы составило df=(3-1)(4-1)=6. Для уровня значимости α=0,05 критическое значение $\chi_{0,05;6} = 12{,}59$ (табл. А.9 Приложения А). При отсутствии связи между добавкой консерванта и сохраняемостью корма χ^2 *несущественно* отличался бы от 0. Так как $\chi^2 > \chi_{0,05;6}$, то можно сделать вывод, что консервирующее средство оказывает статистически значимое влияние на сохраняемость корма.

Коэффициенты контингенции составили:

$$C = \sqrt{\frac{144{,}997}{144{,}997 + 1000}} = 0{,}356,$$

$$K = \sqrt{\frac{144{,}997}{1000(3-1)}} = 0{,}269,$$

$$K' = \sqrt{\frac{144{,}997}{1000\sqrt{6}}} = 0{,}243.$$

Все коэффициенты указывают на существование слабой связи между добавкой консерванта и сохраняемостью корма.

Величина С зависит от числа столбцов и строк таблицы. Поэтому вычисляют исправленный коэффициент контингенции с поправкой - C_{max}. Значение C_{max} с увеличением числа строк и столбцов таблицы приближается к +1. Для квадратной таблицы контингенции

$$C_{max} = \sqrt{(w-1)/w},$$

где w - число столбцов или строк.

Для m×n-таблицы приблизительное значение C_{max} вычисляют как среднее из максимальных значений С соответствующих квадратных таблиц контингенции. Так, для таблицы 3×4 имеем $C_3 = 0{,}826$, $C_4 = 0{,}866$, из чего следует, что

$$C_{max} = (0{,}816 + 0{,}866)/2 = 0{,}841.$$

Скорректированное значение коэффициента:

$$C_{скорр} = C/C_{max} = 0{,}356/0{,}841 = 0{,}423.$$

Поправка важна при малом числе строк и столбцов таблицы.

Точечно-бисериальная корреляция. Когда значения одного нормально распределенного признака, X, метрически шкалированы по нескольким ступеням, а другой признак, Y, альтернативный, типа «да-нет», то связь между подобными переменными называют *двухстрочечной*, или *бисериальной* корреляцией.

Если признак Y в действительности принимает различные значения, которые лишь условно разбиваются на две альтернативные группы, то связь между X и Y называют *непрерывной двухстрочечной* корреляцией. Если же признак Y дихотомический, т.е. может принимать только два взаимоисключающих значения, то говорят о *дискретной двухстрочечной* корреляции или *точечно-бисериальной* корреляцией (r_b).

Для двух альтернативных групп, численностью n_1 и n_2 ($n_1 + n_2 = n$), коэффициент r_b находят путем определения групповых средних $\bar{x}_1$ и $\bar{x}_2$ и общего стандартного отклонения $\hat{\sigma}$:

$$r_b = \frac{\bar{x}_1 - \bar{x}_2}{\hat{\sigma}_x} \sqrt{\frac{n_1 n_2}{n(n-1)}}.$$

Пример 13.16. Требуется определить наличие корреляционной связи между удоем 50 коров и типами гемоглобина: Hb AA и Hb AB (табл. 29).

29. Распределение коров по продуктивности и типу гемоглобина

Удой, тыс. кг (X)	Тип гемоглобина (Y)		
	Hb AA	Hb AB	всего
3	2	-	2
4	5	2	7
5	8	6	14
6	6	10	16
7	1	7	8
8	-	3	3
Σ	22	28	50

Для этих данных $\bar{x}_1 = 4{,}95$, $\bar{x}_2 = 6{,}11$, $\hat{\sigma}_x = 1{,}2$ (по всему материалу; см. стр. 165÷169). Тогда

$$r_b = \frac{4{,}95 - 6{,}11}{1{,}2} \sqrt{\frac{22 \times 28}{50(50-1)}} = -0{,}484.$$

Для двухстрочечных показателей связи *знак при коэффициенте* ***не имеет значения***. Полученный точечно-бисериальный коэффициент корреляции указывает на относительно сильную связь между удоем коров и типом гемоглобина в крови.

13.8. Критерий Фишера-Ирвина

Критерий χ^2 не рекомендуют применять для выборок, имеющих менее 20-30 наблюдений. В таких случаях используют «точный» критерий Фишера-Ирвина (только для таблиц 2×2):

$$P = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{n!\,a!\,b!\,c!\,d!},$$

где символ x! читается, как x-факториал.

При $P \geq 0{,}05$ - нулевую гипотезу принимают,
если $P < 0{,}01$ - нулевую гипотезу отвергают.

Пример 13.17. Из 10 больных животных, которых лечили методом А, у трех состояние улучшилось, у семи - осталось без изменения. Из 7 больных животных, которых лечили методом В, соответственно, 5 и 2 животных (табл. 30). Требуется доказать преимущество метода В.

30. Данные о влиянии метода лечения на состояние животных

Наличие улучшения	Способ лечения		Сумма
	А	В	
+	3[a]	5[b]	8[a+b]
-	7[c]	2[d]	9[c+d]
Сумма	10[a+c]	7[b+d]	17[n]

После подстановки данных в формулу Фишера, имеем:

$$P = \frac{8!9!10!7!}{17!3!5!7!2!}.$$

Сократив 7! (7-факториал) и выписав все множители, получим:

$$P = \frac{(1\times2\times3\times4\times5\times6\times7\times8)(1\times2\times3\times4\times5\times6\times7\times8\times9)(1\times2\times3\times4\times5\times6\times7\times8\times9\times10)}{(1\times2\times3\times4\times5\times6\times7\times8\times9\times10\times11\times12\times13\times14\times15\times16\times17)(1\times2\times3)(1\times2\times3\times4\times5)(1\times2)}.$$

Большинство чисел сокращается и решение будет:

$$P = \frac{4\times7\times9}{11\times13\times17} = 0{,}104.$$

Так как 0,104>0,05, то нулевая гипотеза не отвергается - преимущество метода В не доказано.

Расчет можно упростить, используя тот факт, что логарифм Р равен сумме логарифмов чисел, стоящих в числителе, минус сумма логарифмов чисел, стоящих в знаменателе:

$$\log P = (\mathrm{lod}\,8! + \log 9! + \log 10! + \log 7!) - (\mathrm{lod}\,17! + \log 3! + \log 5! + \log 7! + \log 2!).$$

31. Логарифмы факториалов чисел

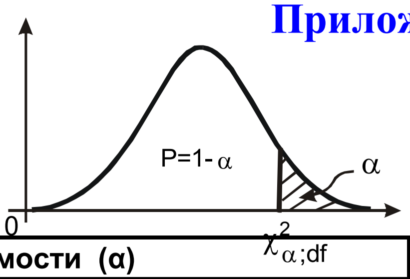
n	Log n!	n	log n!	n	log n!	n	log n!
	-	11	7,601	21	19,708	31	33,915
2	0,301	12	8,680	22	21,051	32	35,420
3	0,778	13	9,794	23	22,412	33	36,939
4	1,380	14	10,940	24	23,793	34	38,470
5	2,079	15	12,112	25	25,191	35	40,014
6	2,857	16	13,321	26	26,606	36	41,571
7	3,702	17	14,551	27	28,037	37	43,139
8	4,606	18	15,806	28	29,484	38	44,719
9	5,560	19	17,085	29	30,947	39	46,310
10	6,560	20	18,386	30	32,424	40	47,912

Если logP<-2, то нулевую гипотезу отвергают; при logP≥-1,3, нулевую гипотезу принимают (т.к. log0,01=-2,0 и log0,05=-1,3).

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

Приложение

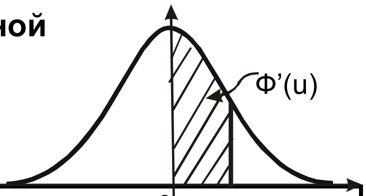
А.9. Критические значения χ^2-распределения Пирсона



df	Уровень значимости (α)									
	0,99	**0,95**	**0,90**	**0,75**	**0,50**	**0,25**	**0,10**	**0,05**	**0,025**	**0,010**
1	…	…	0,02	0,10	0,45	1,32	2,71	3,84	5,02	6,63
2	0,02	0,10	0,21	0,58	1,39	2,77	4,61	5,99	7,38	9,21
3	0,11	0,35	0,58	1,21	2,37	4,11	6,25	7,81	9,35	11,34
4	0,30	0,71	1,06	1,92	3,36	5,39	7,78	9,49	11,14	13,28
5	0,55	1,15	1,61	2,67	4,35	6,63	9,24	11,07	12,83	15,09
6	0,87	1,64	2,20	3,45	5,35	7,84	10,64	12,59	14,45	16,81
7	1,24	2,17	2,83	4,25	6,35	9,04	12,02	14,07	16,01	18,48
8	1,65	2,73	3,49	5,07	7,34	10,22	13,36	15,51	17,53	20,09
9	2,09	3,33	4,17	5,90	8,34	11,39	14,68	16,92	19,02	21,67
10	2,56	3,94	4,87	6,74	9,34	12,55	15,99	***18,31***	20,48	23,21
11	3,05	4,57	5,58	7,58	10,34	13,70	17,28	19,68	21,92	24,72
12	3,57	5,23	6,30	8,44	11,34	14,85	18,55	21,03	23,34	26,22
13	4,11	5,89	7,04	9,30	12,34	15,98	19,81	22,36	24,74	27,69
14	4,66	6,57	7,79	10,17	13,34	17,12	21,06	23,68	26,12	29,14
15	5,23	7,26	8,55	11,04	14,34	18,25	22,31	25,00	27,49	30,58
16	5,81	7,96	9,31	11,91	15,34	19,37	23,54	26,30	28,85	32,00
17	6,41	8,67	10,09	12,79	16,34	20,49	24,77	27,59	30,19	33,41
18	7,01	9,39	10,86	13,68	17,34	21,60	25,99	28,87	31,53	34,81
19	7,63	10,12	11,65	14,56	18,34	22,72	27,20	30,14	32,85	36,19
20	8,26	10,85	12,44	15,45	19,34	23,83	28,41	31,41	34,17	37,57
21	8,90	11,59	13,24	16,34	20,34	24,93	29,62	32,67	35,48	38,93
22	9,54	12,34	14,04	17,24	21,34	26,04	30,81	33,92	36,78	40,29
23	10,20	13,09	14,85	18,14	22,34	27,14	32,01	35,17	38,08	41,64
24	10,86	13,85	15,66	19,04	23,34	28,24	33,20	36,42	39,36	42,98
25	11,52	14,61	16,47	19,94	24,34	29,34	34,38	37,65	40,65	44,31
26	12,20	15,38	17,29	20,84	25,34	30,43	35,56	38,89	41,92	45,64
27	12,88	16,15	18,11	21,75	26,34	31,53	36,74	40,11	43,19	46,96
28	13,56	16,93	18,94	22,66	27,34	32,62	37,92	41,34	44,46	48,28
29	14,26	17,71	19,77	23,57	28,34	33,71	39,09	42,56	45,72	49,59
30	14,95	18,49	20,60	24,48	29,34	34,80	40,26	43,77	46,98	50,89
40	22,16	26,51	29,05	33,66	39,34	45,62	51,80	55,76	59,34	63,69
50	29,71	34,76	37,69	42,94	49,33	56,33	63,17	67,50	71,42	76,15
60	37,48	43,19	46,46	52,29	59,33	66,98	74,40	79,08	83,30	88,38
70	45,44	51,74	55,33	61,70	69,33	77,58	85,53	90,53	95,02	100,42
80	53,54	60,39	64,28	71,14	79,33	88,13	96,58	01,88	06,63	112,33
90	61,75	69,13	73,29	80,62	89,33	98,64	07,56	13,14	18,14	124,12
100	70,06	77,93	82,36	90,13	99,33	109,14	18,50	24,34	29,56	135,81

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

А.4. Площади (вероятности, Ф'(u)) под нормальной кривой распределения в интервале между μ и (μ+uσ) или (μ-uσ) (накопленные частоты нормального распределения)



u	Сотые доли u									
	0,00	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1551	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,498C	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	-	-	-	-	-	-	-	-	-

Пример. Имеем стадо 100 коров со средним удоем μ=4000 кг молока и σ=500 кг. Сколько коров будет в группе с удоем выше среднего, но менее 4500 кг?

Решение: $$u = \frac{4500-4000}{500} = 1 \quad \rightarrow \quad \Phi'(u) = 0{,}3413 .$$

Таким образом, 34 коровы в стаде будут иметь удой выше среднего, но менее 4500 кг молока (столько же в группе с удоем от 3500 до 4000 кг).