Специальный математический практикум

с.н.с. Уфимцев Михаил Валентинович

факультет ВМК МГУ

Резервуар задач

1. В Генеральный штаб прусской армии ежегодно в течение 20 лет поступали от 10 кавалерийских корпусов данные о количестве кавалеристов, погибших в результате гибели под ними коня (всего 200 донесений). Обозначим через x число людей, погибших в одном корпусе за год. Соответствующие частоты Rx даны в табл.

Количество погибших людей, x	0	1	2	3	4	5	6	Всего
Количество донесений Rx с указанием x погибших	109	65	22	3	1	0	0	200

Как эти данные согласуются с гипотезой о том, что число погибших подчиняется распределению Пуассона ? (Указание: так как ожидаемые частоты для x > 2 малы, нужно объединить последние 4 значения x в одну ячейку группировки)

2. Измерения роста в группе из 10 человек дали: 160, 160, 167, 170, 173, 176, 178, 178, 181, 181. Проверяемая гипотеза H состоит в том, что эти данные согласуются с нормальным законом распределения с параметрами: математическим ожиданием 167 и неизвестной постоянной дисперсией D. Альтернативой K служит гипотеза, что данные согласуются с нормальным законом распределения с параметрами: математическим ожиданием > 167 и неизвестной постоянной дисперсией D. Проверить.

3. На производстве испытывали 13 образцов пряжи, чтобы выяснить, не изменяется при намокании ее способность к вытягиванию. Для этого 6 произвольно отобранных образцов были проверены на растяжение, для чего подвешивался груз заданной величины. их относительные удлинения (Y) оказались равны (в процентах) Y = 12.3, 13.7, 10.4, 11.4, 14.9, 12.6.

Остальные 7 образцов были тщательно намочены, после чего их относительные удлинения (X) при такой же проверке с тем же самым грузом дали следующие результаты: X = 15.7, 10.3, 12.6, 14.5, 12.6, 13.8, 11.9.

Требуется выяснить, не приводит ли намокание к изменению длины пряжи ? (Данные считать нормально распределенными).

4. В некотором процессе получается 2 % брака. Берется выборка из 10 независимых образцов. Какова вероятность того, что в выборке не окажется дефектов ?

5. Посмотреть графики биномиального распределения для n = 8,

a) p = 0.1 c) p = 0.5

b) p = 0.3 d) p = 0.9 .

6. В окрестностях г. Кливленда (США) среднегодовое число погибших в автокатастрофах за период 1945-1951 гг. составило 0.5. В 1952 г. там погибло 3 человека. Не указывает это на реальный рост транспортной опасности ?

7. Посмотреть графики распределения Пуассона Po(m) для

a) m = 0.5 b) m = 1.0 c) m = 6.0 .

8. При выборе случайных цифр с одинаковой вероятностью может быть взята каждая цифра 0, 1, 2, 3, ..., 9. В выборке объема 3 все три отобранные цифры превосходят 6. Можно ли считать это неожиданным событием ?

9. Вероятность попадания в цель равна 0.90.

(а) Найдите математическое ожидание числа попаданий в 10 независимых попытках;

(б) какова вероятность более чем 5 попаданий ?

(в) какова вероятность двух или менее попаданий ?

10. Как много независимых попыток нужно сделать в условиях упражнения 9, чтобы вероятность попасть хотя бы один раз была не менее .99 ?

11. По заявлению некоторой фирмы вероятность того, что данное покрытие устойчиво против коррозии (что определяется путем стандартных испытаний) , равна 0.95. Было выбрано 20 независимых образцов.

(а) Если эти покрытия столь хороши, как утверждается, то как много брака можно ожидать ?

(б) Какова вероятность того, что будет обнаружен более чем один случай брака ?

12. Процент выхода, характеризующий эффективность химического процесса, распределен симметрично около среднего 87 %. В каждом из 6 независимых лабораторных испытаний выход оказался менее 87 %.

(а) Какова вероятность такого события ?

(б) Какова вероятность того, что эффективность окажется > 87 % (для этого процесса) в 8 случаях из 10 ?

13. Максимальное сопротивление резиновой ленты должно составлять 2.0 Ом*см. Ниже приведены числа дефектов (X), полученные в 1000 выборках по 10 образцов.

X	0	1	2	3	4	5	6	7	8	9	10
f	353	382	196	55	11	2	1	0	0	0	0

(а) Сравните эти результаты с биномиальным распределением при p = 0.1 и n = 10.

(б) Сравните эти результаты с распределением Пуассона при m = 1.0.

(в) В каком случае данные аппроксимируются лучше ? Почему ?

14. Пусть вероятность получения качественного листа при нанесении на него покрытия равна 0.90.

а) Сколько бракованных листов следует ожидать в выборке объема 50 ?

б) Какова вероятность того, что в выборке объема 50 окажется более 5 плохих листов ?

в) Какова вероятность того, что окажется менее трех плохих листов ?

15. Какова вероятность двух или более дефектных образцов в выборке объема 20, если средняя доля дефектных образцов равна 0.12 ?

16. Фирма провела полевые испытания шести образцов одинакового оборудования и не обнаружила среди них неисправных. Как часто можно ожидать, что в выборке объема 6 будут присутствовать неисправные образцы, если

а) ожидаемая доля неисправных образцов равна 5 % ?

б) ожидаемая доля неисправных образцов равна 10 % ?

17. Фирма провела кампанию по уменьшению числа несчастных случаев. Среднее число таких случаев за предыдущие годы было 20. В год поведения указанной кампании их число оказалось равным 16. Можете ли вы сказать, что это произошло под влиянием неслучайных причин ? Оцените вероятность 16 и менее несчастных случаев, полагая, что условия не изменились.

18. Случайная величина X имеет распределение Пуассона со средним значением μ. Покажите, что математическое ожидание A^X равно

$$\exp\{\mu \cdot (A - 1)\}.$$

Каково среднее значение этой величины, если X имеет биномиальное распределение с параметрами n и p ?

19. Предполагается, что внешний диаметр (ВД) годных для сборки стальных стержней распределен приблизительно по нормальному закону со средним 2.30 см и средним квадратическим отклонением 0.06 см, т.е. N(2.30, 0.06). Пределы допуска 2.31$\pm$ 0.10 см. Изделие с ВД ниже нижнего предела допуска считается ломом, тогда как при превышении ВД верхнего предела возможна доработка. Нужно получить ответы на следующие вопросы:

а) Каков получающийся процент лома ?

б) Сколько процентов продукции нуждается в доработке ?

в) Каков будет ожидаемый процент лома и изделий, требующих доработки, если изменить средний ВД до 2.31 см ?

20. Наружные диаметры втулок распределены нормально. Среднее этого распределения равно 2.000 дюйма, а среднее квадратическое отклонение составляет 0.003 дюйма. Определите вероятность того, что наружный диаметр втулки

а) равен или больше, чем 2.009 дюйма ;

б) меньше, чем 1.994 дюйма ;

в) заключен между 1.997 и 2.003 дюйма ;

г) заключен между 1.994 и 2.006 дюйма.

21. Пусть установлено, что выход в граммах красителя стандартного цвета со специальным оттенком распределен нормально со средним 1550 и средним квадратическим отклонением 50. В скольким из 100 проверок вы ожидаете, что выход в среднем будет

а) ниже 1550 ;

б) выше 1650 ;

в) между 1525 и 1575 ;

г) выше 1470 ?

22. Изделия типа А должны подгоняться к изделиям типа В. Предполагается, что критический внешний размер А распределен нормально со средним 4.30 см и средним квадратическим отклонением 0.04 см. Изделия В имеют критический внутренний размер, который считается распределенным нормально со средним 4.36 см и средним квадратическим отклонением 0.04 см.

а) Какова ожидаемая доля случаев, когда изделия А и В, выбранные независимо и случайно, окажутся непригодными друг для друга ?

б) Надо ли считать необычным два случая несовместимости из 20 ?

23. Рычаг собирается из пяти секций. Исследование отдельных секций показало, что средняя длина крайних секций составляет 1.001 дюйма, а трех средних 1.999 дюйма. Средние квадратические отклонения длин всех секций равны 0.004 дюйма. Если осуществить случайную сборку и отдельные секции предполагаются распределенными нормально, то

а) Какова средняя длина сборки ?

б) Каково среднее квадратическое отклонение длины сборки ?

в) Какова вероятность того, что длина рычага превышает 8.002 дюйма ?

24. В выборку включены 224 последовательные бобины медной проволоки. Измеряется толщина пластикового покрытия проволоки. Можно ли на основе приведенных ниже данных считать, что толщина покрытия имеет нормальное распределение ? Сравните наблюдаемые и теоретические частоты.

Толщина X	146	147	148	149	150	151	152	153
Частота f	3	3	7	11	25	33	34	37
Толщина X	154	155	156	157	158	159	160	
Частота f	25	23	11	9	2	0	1	

25. Величина X распределена нормально и 1) Pr[X > 7] = 0.05; 2) Pr[X < 5] = 0.50. Найдите Pr[4 < X < 6].

26. Найдите функцию плотности вероятностей суммы Z = X + Y, если функция плотностей совместного распределения (X, Y) есть P(x, y) = 1 для 0 <= x <= 1; 0 <= y <= 1 и P(x, y) = 0 вне единичного квадрата.

27. В некотором химическом процессе стандартная жидкость из трех бутылей сливается в большой бак. Если среднее квадратическое отклонение объема жидкости в каждой бутыли равно 0.007 литра, то каково среднее квадратическое отклонение объема жидкости в большом баке ?

28. Две части A и B соединяются в одну трубу. Средние квадратические отклонения длин A и B равны 0.22 и 0.45 мм соответственно. Длина, добавляемая при соединении, имеет среднее квадратическое отклонение σ мм. Получите формулу для среднего квадратического отклонения длины C, предполагая, что A и B взаимно независимы.

29. При сборке два изделия C (см. упражнение 28) должны отличаться по длине менее чем на 1.25 мм. Предполагая нормальное распределение, найдите выражение для вероятности того, что два случайно взятых изделия C окажутся совместимыми. Как мала должна быть величина σ, чтобы эта вероятность составляла не менее 90 % ?

30. Если в упражнении 29 величина $\sigma = 0.12$ мм, а пары C отбираются случайно до тех пор, пока не найдется пара, пригодная для сборки, то каково среднее число пар, которые придется перебрать ?

31. В таблице, приведенной ниже, X - твердость по Шору сложного сплава. Постройте гистограмму и сравните ее с нормальным распределением при $\mu = 73.5$, $\sigma = 3$.

X	57-58	59-60	61-62	63-64	65-66	67-68	69-70	71-72
f	1	3	5	8	10	13	20	25
X	73-74	75-76	77-78	79-80	81-82	83-84	85-86	87-88
f	38	22	18	15	11	6	3	2

32. Следующие цифры представляют собой значения твердости 15 образцов сплава в условных единицах: 12.1; 13.7; 11.0; 11.6; 11.9; 12.9; 13.4; 12.2; 12.5; 11.9; 11.9; 11.5; 12.9; 13.0; 10.5. Проверьте гипотезу о том, что среднее значение твердости равно 12.0 (используйте двухсторонний критерий с $\alpha = 0.05$).

33. Проверьте, используя данные упражнения 32, гипотезу о том, что среднее квадратическое отклонение составляет 1.4 у. е.

34. Какой потребуется объем выборки для того, чтобы с доверительной вероятностью, равной 99 %, среднее выборочное значение твердости (Xсредн.)

отличалось от его математического ожидания не более чем на 0.3 у.е., если стандартное отклонение известно и равно 2.0 у.е. ?

35. При проверке 100 единиц продукции производственной линии было обнаружено 2 дефектных. Что можно сказать относительно процента брака в исследуемом процессе ?

36. < Нижеследующие странные числа получились при переводе фунтов и унций англо-американской системы мер в метрическую. - М.У. > Минимальный допустимый вес консервной банки равен 1 фунту = 453.6 гр. Отклонение веса более чем на 14.175 гр. ниже этого значения является основанием для отбраковки изделия. Какие ограничения надо наложить на выборочное среднее значение, чтобы вероятность браковки не превосходила 0.001 при а) n = 10; б) n = 20; в) n = 50 ?

Можно сделать предположение, что среднее квадратическое отклонение веса консервной банки составляет 2.835 гр. (Используйте уровень значимости 0.01).

37. Шариковые подшипники проходят проверку на овальность в специальном измерительном устройстве, которое автоматически фиксирует отклонение от заданных условий. Возникло подозрение, что время, необходимое для проведения проверки, у разных контролеров различно. Были отобраны 4 контролера, и время, необходимое каждому из них на проведение проверки, регистрировалось. Измерения повторялись каждым контролером 6 раз. Приведенные ниже данные представляют собой время измерения в сек.

а) Наблюдаются ли различия в скорости работы контролеров ?

б) Какие предположения необходимо сделать при выбранном критерии значимости ?

Контролёр	**Время измерения**					
A	13	12	11	14	12	13
B	14	14	16	15	14	16
C	11	14	11	13	14	11
D	12	15	14	16	13	16

38. Найдите 99 %-ный доверительный интервал для разности средних значений времени измерения для контролеров A и D из упражнения 37.

39. Измерялось сопротивление проволок пяти типов. Утверждается, что между сопротивлениями проволок разных типов в среднем нет различий. Результаты проверки сопротивлений проволок каждого типа в шести выборках приведены ниже.

Проволока				
A	B	C	D	E
0.126	0.121	0.121	0.129	0.128
0.131	0.121	0.119	0.132	0.135
0.126	0.124	0.126	0.136	0.134
0.127	0.122	0.128	0.139	0.129
0.124	0.120	0.126	0.130	0.135
0.130	0.124	0.124	0.132	0.132
0.128	0.125	0.122	0.137	0.134
0.124	0.120	0.127	0.136	0.126

а) Можно ли принять гипотезу H об одинаковом значении среднего сопротивления для проволок шести типов ?

б) Если принимают гипотезу Н, то что это значит ?

40. Проверьте по данным, приведенным в упражнении 39, гипотезу об однородности дисперсий. Определите 95 %-ный доверительный интервал для общего значения сопротивления, если принята нулевая гипотеза.

41. Для определения прочности на разрыв целлофановых мешков разработан специальный критерий. Исследовались 15 мешков типа А и 20 мешков типа В. Каждый из мешков наполняли и бросали до тех пор, пока он не разрывался. Обозначим число падений мешка до момента разрыва через X. Получены следующие результаты:

Тип	Среднее X	Выборочная дисперсия D	n
A	75.5	83.17	15
B	89.3	128.20	20

а) Можно ли говорить, что мешки одного типа прочнее, чем другого ?

б) Сделайте необходимые предположения и задайте значение вероятности ошибки.

42. Дисперсия предела прочности на разрыв некоторого волокна составляет 35.63 кг ^{2} . Ожидается, что внесенные в технологический процесс изменения снизят указанную дисперсию. В выборке объема 15 были зарегистрированы следующие значения предела прочности на разрыв в кг :

151	156	147	153	155
148	160	149	160	156
161	154	162	163	149

а) Привело ли изменение процесса к снижению дисперсии ?

б) Установите критическую область, где должно производиться отклонение гипотезы Н.

43. Исследовались потери веса 11-ти резиновых стержней при испытаниях на износ. От каждого стержня было отрезано по два образца для проведения испытаний. Один из них прошел вулканизацию при 80 градусах. а другой - при 150 градусах.

а) Можно ли утверждать, исходя из приведенных ниже данных, что наблюдается значимое различие между средними потерями веса образцов, прошедших различную вулканизацию, при $\alpha = 0.05$?

б) Какая гипотеза проверяется ?

Т вулканизации	СТЕРЖНИ										
	1	2	3	4	5	6	7	8	9	10	11
80	3.02	2.22	4.60	4.53	2.31	3.11	2.70	2.58	3.27	4.19	2.90
150	2.91	2.30	4.15	2.63	2.40	3.20	2.50	2.29	3.11	3.80	2.72

44. Произведено исследование предела прочности на разрыв шести различных по химической структуре твердых смол. Было взято по два образца каждой из них. При этом требовалось выяснить, вносят ли действия операторов какое-нибудь смещение в результаты наблюдений. Двум операторам А и В предложили испытать по одному образцу смол каждого типа.

а) Можно ли говорить, что наблюдаются различия между результатами измерений у разных операторов ?

б) Если да, то определите 95 %-ный доверительный интервал для математического ожидания разности.

Смола	127	135	138	139	146	152
Оператор А	**5240**	**4975**	**5050**	**5075**	**4795**	**5190**
Оператор В	**5230**	**4980**	**5020**	**5080**	**4750**	**5120**

45. В одной из общин 92 % жителей получили соматическую вакцину, а в другой общине только 87 %. Численность общин оставляет 1200 и 1520 жителей. Можно ли говорить о том, что в одной из общин вакцинация проходит более активно, чем в другой ?

46. Найдите 95 %-ный доверительный интервал для дисперсии коэффициента преломления различных образцов стекла, если по выборке, включающей 10 образцов, получены следующие результаты:

1.589	1.587	1.559	1.596	1.583
1.569	1.574	1.592	1.590	1.561

47. Ожидается, что число дефектных шин среди изготовляемых заводом А должно составлять 3 шины в неделю. В течение последних 20 недель было обнаружено 47 дефектных шин. Можно ли сказать, что наблюдается значимое снижение уровня брака ?

48. Ожидается, что после определенной термической обработки образцов стали их предел прочности на разрыв возрастет приблизительно на 1200 фунт/дюйм ^{2}. Укажите этапы процедуры статистической проверки того, что указанный показатель удалось повысить, по крайней мере, на 1200 фунт/дюйм ^{2}. Какие при этом нужно сделать предположения ? Как выбрать гипотезу, уровень значимости и мощность критерия ?

49. Измеряется сопротивление проволоки. При исследовании 8-ми образцов получены следующие результаты:

0.129; 0.132; 0.136; 0.139; 0.132; 0.137; 0.125 и 0.136.

а) Есть ли достаточные основания для того, чтобы исключить из выборки одно или несколько значений ?

б) Какие предположения были сделаны при исследовании статистического критерия ?

50. Для 10-ти выборок, каждая из которых включала по 6 образцов виниловых стержней, было проведено измерение удлинений этих стержней. Ниже приведены средние значения удлинений в процентах (%) для каждой из выборок:

781, 726, 719, 735, 742, 722, 730, 728, 742 и 736.

Можно ли классифицировать некоторые из приведенных средних значений как выбросы, если используется $\alpha = 0.05$?

51. Приведенные ниже данные получены из большой выборки, собранной на одном промышленном предприятии. Измерен диаметр шага резьбы металлического штифта. Приведены отклонения от заданного среднего, сгруппированные в интервалы длиной 0.0003 ед.

Используя критерий согласия Пирсона, исследуйте эти данные на нормальность.

X	f	X	f
от -0.0018 до -0.0015	3	от -0.0003 до -0.0000	178
от -0.0015 до -0.0012	11	от 0.0000 до 0.0003	147
от -0.0012 до -0.0009	32	от 0.0003 до 0.0006	91
от -0.0009 до -0.0006	89	от 0.0006 до 0.0009	43
от -0.0006 до -0.0003	149	от 0.0009 до 0.0012	12.

52. При работе специального счетчика в некоторых заданных условиях среднее число отсчётов в течение одной мин. $m_0 = 4.0$. За 10-минутный интервал было зарегистрировано $X_0 = 31$ отсчётов. Можно ли считать этот результат совместимым с ожидаемым ? Предположим, что нас устраивает ситуация, когда мы с 5 %-ной вероятностью совершаем ошибку, делая вывод о несовместимости выборки и совокупности, когда на самом деле они совместимы.

53. По записям строительной фирмы число несчастных случаев за период в несколько лет составляло в среднем 1.5 в месяц. В течение последнего года зарегистрировано 25 несчастных случаев. Хотелось бы выяснить, превосходит ли зарегистрированное число несчастных случаев то их количество, которое можно было бы ожидать. Размер критерия $\alpha = 0.05$.

54. Данные образуют выборку $y_1, \ldots, y_n$ из нормального $N(\mu, \sigma^2)$ распределения. Показать, что для выборочного стандартного отклонения

$$s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} ,$$

где $\bar{y}$ - выборочное среднее, в линейном приближении имеет место соотношение

$$\sigma\{s_y\} = \frac{\sigma}{\sqrt{2\cdot(n-1)}} ,$$

где $\sigma\{z\}$ - стандартное отклонение для с.в. z.

55. Данные образуют выборку $y_1, \ldots, y_n$ из нормального $N(\mu, \sigma^2)$ распределения. Показать, что для выборочной дисперсии

$$s_y{}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 ,$$

где $\bar{y}$ - выборочное среднее, для оценки точности $s_y{}^2$ в линейном приближении имеют место формулы (здесь D{z} - дисперсия, а σ {z} - стандартное отклонение для с.в. z)

$$D\{s_y{}^2\} = \frac{2\sigma^4}{n-1} , \sigma\{s_y{}^2\} = \sqrt{\frac{2}{n-1}}\cdot\sigma^2 .$$

Сравнить ее с точным вторым моментом для выборочной дисперсии в общем случае и для нормальной выборки в ([3], гл. 27).

56. Измерения эффективного сечения реакции двумя методиками дали результаты: (а) (6.4; 7.2; 7.0; 5.8) барн; (б) (7.4; 7.5; 7.0) барн. Можно ли быть уверенным, что методики неравноправны ?

57. На основании $n = 12$ измерений оценка дисперсии $s_y{}^2 = 6$. Не подтверждает ли этот результат сомнение, что $D\{y\} = D_1 = 7 > D_0 = 4$ (уровень значимости $\alpha = 0.05$, данные распределены нормально) ?

58. Резерфорд и Гейгер регистрировали число α-частиц, излучаемых диском в течение 2608 периодов по 7.5 сек. здесь *r(i)* - число случаев (периодов) с *i* α-частицами; *n*p(i)* - математические ожидания *E{r(i)}* числа случаев, если предполагать закон распределения Пуассона. Получилась следующая таблица:

i	0	1	2	3	4	5	6	7	8	9	10	Всего 2608
r(i) n*p(i)	57 54.399	203 210.523	383 407.361	525 525.495	532 508.418	408 393.515	273 253.817	139 140.325	45 67.882	27 29.189	16 17.075	2608 2608.0

(а) На основании экспериментальных данных вычислить оценку параметра θ распределения Пуассона Po(θ) и сравнить ожидаемые значения

$$n*p(i) = n*Pr\{X = i\} = n * \quad e^{-\theta}\frac{\theta^{i}}{i!}$$

с полученными Резерфордом и Гейгером.

(б) Проверить гипотезу H: X имеет распределение Пуассона (с помощью критерия Пирсона Хи-квадрат).

59. В таблице приведены данные об уровнях холестерина в крови (X) и отношениях вес/рост (Y, фунт/дюйм) для 10 человек, прошедших обследование в кардиологическом центре.

X	254	240	279	284	315	250	298	384	310	337
Y	2.71	2.96	2.62	2.19	2.68	2.64	2.37	2.61	2.12	1.94

Предположим, что пары (X, Y), помещенные в табл., являются реализациями независимых двумерных векторов с совместным нормальным распределением $N_2(\kappa; \eta, \sigma^2, \tau^2, \rho)$. Можно ли, применяя критерий

$$t = r * \text{Sqrt}((n - 2)/(1 - r^2)), \ |t| >= t(n - 2, 1 - \alpha/2),$$

где r – коэффициент выборочной корреляции, α – уровень значимости (положить $\alpha = 0.1$), заключить, что уровень холестерина в крови коррелирован с отношением рост/вес ?

60. За первый день счетчик зарегистрировал 20026 импульсов пуассоновского процесса, за второй день – 19580. Есть ли основания считать, что за второй день интенсивность поступления импульсов снизилась ? Уровень значимости $\alpha = 0.05$.

61. За первый час счетчик зарегистрировал 150 импульсов пуассоновского процесса, а за последующие 2 часа – 250 импульсов. Была ли одинаковой интенсивность поступления импульсов в единицу времени ? Уровень значимости $\alpha = 0.05$.

62. При чтении шкалы некоторого прибора, где последняя цифра оценивается на глаз, наблюдатель часто предпочитает одни цифры другим. В следующей таблице даны частоты появления каждой цифры в последнем разряде для 200 выбранных случайным образом чтений шкалы, сделанных

Цифра	0	1	2	3	4	5	6	7	8	9
Частота	35	16	15	17	17	19	11	16	30	24

некоторым наблюдателем.

Таблица показывает, что цифры 0 и 8 появляются несколько чаще, чем другие цифры. Не делает ли наблюдатель систематической ошибки ?

63. Генератор случайных чисел выдает независимые значения X, целочисленной переменной в интервале 0, ..., 99. Наблюдаемые частоты попадания в отрезки [0, 19], [20, 39], [40, 59], [60, 79], [80, 99] оказались соответственно равными 9, 6, 12, 13 и 6. Применить критерий Хи-квадрат для проверки H: Pr{ X = k } = 0.01, k = 0, ..., 99. Уровень значимости $\alpha = 0.01$.

64. При 8000 испытаний события A, B и C, составляющие полную группу, осуществились 2014, 5012 и 974 раз, соответственно. С помощью критерия Хи-квадрат уровня $\alpha = 0.05$ проверить гипотезу : Pr{ A } = 0.5 - 2*a, Pr{ B } = 0.5 + a, Pr{ C } = a, где 0 < a < 0.25.

65. При определении величины заряда электрона e0*10 $^{-10}$ Милликен получил n = 58 независимых измерений x_i величины e0. Выборочное среднее и выборочная дисперсия оказались равными соответственно

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 4.7808$$

и

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = 22981*10^{-8}.$$

Предполагая, что ошибки измерений распределены нормально, а результаты измерений равноточны и лишены систематической ошибки, построить 100*γ%-ный доверительный интервал для e0 с $\gamma = 0.9$.

66. Для определения точности регистрирующего прибора было произведено 5 независимых измерений. Получены результаты: y(1) = 2781, y(2) = 2836, y(3) = 2807, y(4) = 2763, y(5) = 2858. Определить приближенное значение дисперсии и среднего квадратического отклонения, характеризующих точность прибора, если значение измеряемой величины а) известно и равно Ave(y) = 2800; б) неизвестно. Оценить точность.

67. Найти среднюю квадратичную ошибку функции F = ln(n1 / n2). Указать границы применимости формулы.

68. Найти формулу для оценки дисперсии D логарифма интенсивности излучения y = ln(N / t).

69. Проверить значимость простой линейной регрессии для данных:

x(i)	20	22	24	26	28	30	32	34	36	38	40	42
y(i)	8.4	9.5	11.8	10.4	13.3	14.8	13.2	14.7	16.4	16.5	18.9	18.5

70. Построить доверительный интервал для линии регрессии по данным задачи 69 для $\alpha = 0.05$.

71. Проверить адекватность описания простой линейной регрессией данных, приведенных ниже:

x(i)	1.0	1.0	2.0	3.3	3.3	4.0	4.0	4.0	4.7	5.0	5.6	5.6	5.6	6.0	6.0	6.5	6.9
y(i)	2.3	1.8	2.8	1.8	3.7	2.6	2.6	2.2	3.2	2.0	3.5	2.8	2.1	3.4	3.2	3.4	5.0

72. При построении линии регрессии нужно взять полиномы степени не выше второй. Для заданных экспериментальных данных используйте наиболее подходящие полиномы. Выбор обосновать.

y(i)	x(i)
240	1500
270	1600
260	1700
275	1800
295	1900
310	2000

73. (Дрейпер, Смит; гл. 1, зад. 1).

X	-5	-4	-3	-2	-1	0	1	2	3	4	5
Y	1	5	4	7	10	8	9	13	14	13	18

В одной работе исследовалось влияние температуры на выход химического продукта. Были собраны следующие данные (в кодированной форме):

1) Какими будут НК-оценки коэффициентов β_0 и β_1 в предположении, что модель данных имеет вид: $Y = \beta_0 + \beta_1 * X + \varepsilon$? Каково уравнение для предсказания ?

2) Каковы границы доверительного интервала ($\gamma = 0.95$) для β_1 ?

3) Каковы границы доверительного интервала ($\gamma = .95$) для “истинного” среднего значения величины Y, когда X = 3 ?

74. Пусть мы наблюдаем неотрицательную непрерывную с.в. X. Её значения (округленные до 0.01 и данные в порядке возрастания) для n = 50 испытаний суть 0.01, 0.01, 0.04, 0.17, 0.18, 0.22, 0.22, 0.25, 0.25, 0.29, 0.42, 0.46, 0.47, 0.47, 0.56, 0.59, 0.67, 0.68, 0.70, 0.72, 0.76, 0.78, 0.83, 0.85, 0.87, 0.93, 1.00, 1.01, 1.01, 1.02, 1.03, 1.06, 1.32, 1.34, 1.37, 1.47, 1.50, 1.52, 1.54, 1.59, 1.71, 1.90, 2.10, 2.35, 2.46, 2.46, 2.50, 3.73, 4.07, 6.03. Рабочая гипотеза H состоит в том, что $Pr\{X <= x\} = F(x) = 1 - exp(-x), x>= 0$. Группируя данные в $k = 4$ равновероятных интервала (в предположении, что H верна), проверьте это предположение по критерию Хи-квадрат при уровне значимости $\alpha = 0.1$.

75. Число золотых песчинок X в тонком слое взвеси регистрировалось под микроскопом в течение равных интервалов времени. Используя данные из таблицы:

Число частиц i	0	1	2	3	4	5	6	7	Всего
Число событий m(i)	112	168	130	68	32	5	1	1	Σ m(i) = 517

, проверьте гипотезу H: X имеет распределение Po(θ), где θ – неизвестный параметр.

76. / Критерий Бартлетта/ При определении предела прочности на разрыв нескольких различных по структуре полимеров испытания проводились на восьми образцах каждого из шести исследуемых полимеров. Следующие числа представляют собой оценки дисперсии в кодированном виде : 3.24; 4.18; 4.06; 3.98; 4.19; 4.02 (единиц) ^{2} . Нужно проверить гипотезу $H: \sigma_1^2 = \sigma_2^2 = ... = \sigma_6^2$. Допустим, что необходимые предположения выполняются, а уровень значимости $\alpha = 0.01$.

77. Чтобы определить силу сцепления клеевых соединений двух стекол, были проведены испытания на растяжение. Исследовались образцы, у которых обработка склеиваемых поверхностей производилась двумя различными методами, называемыми 1) перекрестной шлифовкой и 2) торцевой обточкой.

Требуется решить при уровне значимости $\alpha = 0.05$, могут ли эти данные быть взяты из нормально распределенных совокупностей с одинаковыми средними значениями. Предположения: случайность и независимость выбора образцов при испытании соединений каждого типа, а также равенство дисперсий (общая дисперсия неизвестна).

Перекрестная шлифовка		Торцевая обточка		Перекрестная шлифовка		Торцевая обточка	
16	20	13	14	18	17	17	13
14	15	19	15	19	18	21	15
19	18	14	10				

Табл. 8.1. Сила сцепления образцов склеенных поверхностей.

78. Пластмассовые контейнеры должны обладать определенной прочностью. При исследовании 100 образцов контейнеров было обнаружено, что два из них не удовлетворяют критерию прочности. Что можно сказать относительно вероятности (1 - p) того, что выбранный случайным образом контейнер будет удовлетворять требованиям прочности ? Известно, что МП-оценка для $p = X/n$, где n – число дефектных изделий. Предположим, что требуется проверить гипотезу о равенстве доли дефектных изделий величине p_0 = 0.01, т.е. гипотеза Н: $p = p_0$ (= 0.01), где p – истинная доля дефектных изделий в совокупности, из которой взята выборка, а p_0 – ее предполагаемое значение. Гипотеза Н: p = 0.01 проверяется при α = 0.05 с помощью двустороннего критерия. Сравните результаты, вычисленные из функции вероятностей биномиального распределения, с полученными по аппроксимации нормальным распределением.

79. На каждом из двух имеющихся на заводе станков производят новые изделия одного типа. При пробном пуске из 350 изделий, изготовленных на станке А, при проверке по принципу “годен - не годен”, 15 оказались дефектными. Среди 320 изделий, изготовленных для сравнения на станке В, оказалось всего 8 дефектных. Можно ли, приняв уровень значимости равным α = 0.05, ожидать, что в продукции, производимой на сравниваемых станках, окажется одинаковая доля дефектных изделий ?

80. / Парные наблюдения/ Данные получены на испытаниях шин. В табл. приведены результаты исследования температуры, до которой нагревается шина при движении рейсового автобуса. Приведенные цифры - результаты испытания левых и правых передних шин 19 автобусов. Рассмотрены три варианта нагрузки на каждую шину - один при движении по поверхности с частыми и острыми ребрами и два при движении по более мелким неровностям. При этом замечено, что при движении по поверхности с частыми ребрами шины нагреваются до более высокой температуры. В приведенной таблице представлены наблюдения, полученные при движении автобусов по поверхности с частыми ребрами.

Табл. 8.3. Условные значения температур левой и правой передних шин.

Левая передняя (ЛП)	Правая передняя (ПП)	D = ПП - ЛП	ЛП	ПП	D = ПП - ЛП
36	27	-9	41	60	19
42	45	3	40	34	-6
55	84	29	100	117	17
59	84	25	58	78	20
79	70	-9	38	56	18
108	99	-9	73	85	12
79	84	5	89	65	-24
41	77	36	58	72	14
36	56	20	60	85	25
47	86	39			

На основании данной таблицы нужно проверить гипотезу H о равенстве температур правой и левой передних шин. Уровень значимости α = 0.05.

81. На сталелитейном заводе возникли подозрения относительно постоянства содержания марганца в одной из марок выпускаемой стали. В течение восьми недель производилось по 10 отливок в неделю (при разной температуре ковшей). В табл. 8.4 приведены результаты исследований этих отливок. Приведенные цифры показывают процентное содержание марганца в каждой отливке.

Номер выборки	1	2	3	4	5	6	7	8
	1.20	**1.39**	**1.18**	**1.22**	**1.21**	**1.30**	**1.33**	**1.13**
	1.17	**1.31**	**1.17**	**1.15**	**1.24**	**1.35**	**1.35**	**1.11**
	1.15	**1.29**	**1.08**	**1.17**	**1.19**	**1.25**	**1.27**	**1.20**
	1.21	**1.28**	**1.15**	**1.22**	**1.17**	**1.26**	**1.30**	**1.12**
	1.14	**1.30**	**1.18**	**1.26**	**1.15**	**1.23**	**1.24**	**1.13**
	1.17	**1.28**	**1.09**	**1.27**	**1.18**	**1.24**	**1.30**	**1.08**
	1.18	**1.34**	**1.06**	**1.19**	**1.17**	**1.26**	**1.27**	**1.15**
	1.21	**1.32**	**1.08**	**1.22**	**1.17**	**1.33**	**1.29**	**1.17**
	1.25	**1.30**	**1.15**	**1.20**	**1.25**	**1.28**	**1.34**	**1.11**
	1.14	**1.35**	**1.06**	**1.15**	**1.18**	**1.24**	**1.28**	**1.16**
Средние значения	**1.182**	**1.316**	**1.120**	**1.205**	**1.191**	**1.274**	**1.297**	**1.136**
Дисперсия $\cdot 10^4$	**12.6222**	**12.2667**	**25.3333**	**17.1666**	**10.5444**	**16.4889**	**12.011**	**12.044**

Необходимо решить, можно ли при уровне значимости $\alpha = 0.05$ считать, что отливки относятся к совокупностям с одним и тем же средним значением, т.е. $H : \mu_1 = \mu_2 = \ldots = \mu_8$. Предполагается, что выборки взяты из нормально распределенных совокупностей с неизвестной общей дисперсией. (Подсказка : см. табл. 2).

82. При изучении стабильности температуры в термостате получены данные: 21.2; 21.8; 21.3; 21.0; 21.4; 21.3. Затем к стабилизатору температуры было применено некоторое усовершенствование, после чего (на другом режиме) получены данные: 37.7; 37.6; 37.6; 37.4. Можно ли при уровне значимости α = 0.05 считать усовершенствование эффективным ?

83. При восьмикратном исследовании стального бруска на разрыв получено среднее значение предела прочности Ave(x) = 62 кг/(мм)^2, причем среднеквадратическое отклонение s = 1.3. Минимальное значение x_6 = 58.6 вызывает сомнение. Проверить гипотезу о том, что x_6 – выброс.

84. Измеряемая с.в. x подчиняется закону распределения $N(10, 5)$. Найти симметричный относительно математического ожидания μ_x интервал, в который с вероятностью p попадет измеренное значение. Рассмотреть следующие числовые значения: а) p =0.9974; б) p = 0.9544; в) p = 0.50.

85. Случайная величина $x \sim N(\mu, \sigma^2)$. Пользуясь таблицей функции нормального распределения, вычислить вероятность p_k того, что отклонение с.в. x от ее математического ожидания не превзойдет величины $k{\cdot}\sigma$ (ответ получить для трех значений k = 1, 2, 3).

86. В нормально распределенной совокупности 15% значений x меньше 12 и 40% значений x больше 16.2. найти математическое ожидание и стандартное отклонение данного распределения.

87. Деталь, изготовленная автоматом, считается годной, если отклонение x от номинала не превышает 10 мм. Точность изготовления деталей характеризуется стандартным отклонением σ . Считая, что для данной технологии σ = 5 и x распределено нормально, выяснить, сколько процентов годных деталей изготовляет автомат.

88. В условиях предыдущей задачи выяснить, какой должна быть точность изготовления, чтобы процент годных деталей повысился до 98 ?

89. Деталь изготовляется на станке. Ее размер x представляет собой случайную величину, распределенную по $N(20$ см $,0.2$ см $^2)$. Какую относительную точность изделия можно гарантировать с вероятностью 0.95 ?

90. Случайная величина $x \sim N(1\ ,\sigma^2)$. Известно, что P{ $x < 2$ } = 0.99. Вычислить $E\{x^2\}$ и P{ $x^2 > 2$ }.

Комментарий к зад. 91 – 94: Из популяции с распределением $N(\mu\ ,\sigma^2)$ извлечены две независимые выборки объемов n_1 и n_2. Выборочные оценки математического ожидания и дисперсии по этим выборкам суть $\bar{x}_1, \bar{x}_2, s_1^2, s_2^2$. Объединенные оценки м.о. и дисперсии по выборке объема $n_1 + n_2$ суть

$$\bar{x} = \frac{n_1\,\bar{x}_1 + n_2\,\bar{x}_2}{n_1 + n_2}, \quad S^2 = \frac{(n_1 - 1)\,s_1^2 + (n_2 - 1)\,s_2^2}{n_1 + n_2 - 2} .$$

Если σ^2 известна, то доверительный интервал для μ равен

$$\bar{x} - \frac{\sigma}{\sqrt{n_1 + n_2}} \cdot u_{1-\alpha/2} < \mu < \bar{x} + \frac{\sigma}{\sqrt{n_1 + n_2}} \cdot u_{1-\alpha/2} . \qquad (*)$$

Если же σ^2 неизвестна, то $\hat{\sigma}^2 = S^2$, и доверительный интервал для μ

$$\bar{x} - \frac{S}{\sqrt{n_1 + n_2}} \cdot t_{1-\alpha/2}\,(n_1 + n_2 - 2) < \mu < \bar{x} + \frac{S}{\sqrt{n_1 + n_2}} \cdot t_{1-\alpha/2}\,(n_1 + n_2 - 2) . \qquad (**)$$

В задачах 91-94 найти из (*) и (**) 90% - и 99%-ные доверительные интервалы для μ.

91. Емкость конденсатора, если в ходе первого выборочного обследования было получено $n_1 = 16$, $\bar{x}_1$ = 20 мкФ, σ = 4 мкФ известно, а n_2 = 9 и $\bar{x}_2$ = 18 мкФ.

92. Время безотказной работы электронной лампы, если для первой выборки $n_1 = 100$, $\bar{x}_1$ = 500 час., σ известно и равно 10 час. Для второй выборки n_2 = 64 и $\bar{x}_2$ = 480 час.

93. Диаметр вала, если $n_1 = 9$, $\bar{x}_1$ = 30 мм, s_1^2 = 9 мм^2; $n_2 = 16$, $\bar{x}_2$ = 29 мм, s_2^2 = 4.5 мм^2.

94. Содержание углерода в единице продукта, если $n_1 = 25$, $\bar{x}_1$ = 18 г, s_1^2 = 16 г^2; $n_2 = 9$, $\bar{x}_2$ = 18.8 г, s_2^2 = 20 г^2.

95. При измерении производительности двух агрегатов получены следующие результаты (в кг вещества за час работы) :

№ замера	1	2	3	4	5
Агрегат А	14.1	10.1	14.7	13.7	14.0
Агрегат В	14.0	14.5	13.7	12.7	14.1

Можно ли считать, что производительности агрегатов А и В одинаковы, в предположении, что обе выборки получены из нормально распределенных популяций ? Принять $\alpha = 0.10$.

96. Проверить гипотезу о равенстве дисперсий трех популяций, используя следующие результаты наблюдений :

№ выборки	№ наблюдения					$\sum x_i$	$\sum x_i^2$
	1	2	3	4	5		
1	6	5	12	9	10	42	386
2	14	11	5	6	-	36	378
3	12	4	7	-	-	23	209

Принять $\alpha = 0.05$.

97. Радиоаппаратура состоит из 1000 элементов. Вероятность отказа одного элемента в течение одного года работы равна 0.001 и не зависит от состояния других элементов. Какова вероятность отказа двух и не менее двух электронных элементов в год ?

98. Математическое ожидание числа отказов радиоаппаратуры за 10000 часов работы равно 10. Определить вероятность отказа радиоаппаратуры за 100 часов работы.

99. Аппаратура содержит 2000 одинаково надежных элементов, вероятность отказа для каждого из которых равна $p = 0.0005$. Какова вероятность отказа аппаратуры, если он наступает при отказе хотя бы одного из элементов ?

100. В аппаратурный отсек космической ракеты за время ее полета с вероятностью

$$p(r,\lambda) = e^{-\lambda}\frac{\lambda^r}{r!}$$

попадает r элементарных частиц. Условная вероятность для каждой из них попасть при этом в уязвимый блок равна p . Найти вероятность попадания в блок : а) ровно k частиц; б) хотя бы одной частицы.

101. Математическое ожидание числа бракованных аппаратов при проверке их на безотказность действия определяется формулой

$$T = N\cdot [1 - (1 - p / (\Omega\cdot N))^m] ,$$

где p - вероятность того, что испытание одного из аппаратов будет признано зачетным; Ω - среднее число зачетных испытаний до получения отказа в действии аппарата; N - число аппаратов, участвующих в проверке; m - число испытаний (зачетных и незачетных), приходящихся на один аппарат.

Пользуясь методом линеаризации, определить зависимость математического ожидания и дисперсии с.в. T от m , если N, p и Ω - независимые с.в., математические ожидания и дисперсии которых соответственно равны: $E[N] = 5$, $E[p] = 0.8$, $E[\Omega] = 4$, $D[N] = 1$, $D[p] = 0.1$, $D[\Omega] = 0.2$.

102. Максимальная высота полета спутника определяется формулой

$$Y = y_0 + (R + y_0)\cdot [\,(1 + l)/(1 - \lambda)/2 - 1\,],$$

где $\lambda = v^2/(2gR)\,[1 + y_0/R]$, $l = \sqrt{1-4\lambda(1-\lambda)\cos^2\theta}$, y_0 - высота активного участка траектории, g - ускорение силы тяжести на поверхности Земли, R - радиус Земли.

Функция Y в области практически возможных значений случайных аргументов линеаризуется. Начальная скорость v и угол бросания θ - нормальные с.в. с функцией плотности вероятностей (ф.п.в.)

$$f(v,\theta) = \frac{1}{2\pi\sigma_v\sigma_\theta\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left[\left(\frac{v-\bar{v}}{\sigma_v}\right)^2+\left(\frac{\theta-\bar{\theta}}{\sigma_\theta}\right)^2-2r\cdot\left(\frac{v-\bar{v}}{\sigma_v}\right)\cdot\left(\frac{\theta-\bar{\theta}}{\sigma_\theta}\right)\right]\right\}$$

,где через $\bar{v},\bar{\theta}$ обозначены математические ожидания с.в. v и θ, соответственно. Найти приближенное значение дисперсии для максимальной высоты полета спутника.

103. Найти срединное отклонение с.в., имеющей ф.п.в.

$$f(x) = 0.5\cdot\exp\{-|x|\}.$$

104. Ф.п.в. случайной величины x имеет вид

$$f(x) = \begin{cases} \frac{1}{2l}, |x-a| \le l \\ 0, |x-a| > l. \end{cases}$$

Определить a) $E\{x\}$; b) $D\{x\}$; c) найти связь между средним квадратическим и срединным отклонениями с.в. x.

105. Определить срединную ошибку прибора, если систематических ошибок он не имеет, а случайные распределены по нормальному закону и с вероятностью 0.8 не выходят за пределы $\pm$ 20 м.

106. Количество тепла Q в калориях, выделяемое в проводнике с сопротивлением R при прохождении тока I в течение времени T, определяется формулой

$$Q = 0.24\, I^2 RT.$$

Ошибки измерения величин I, R, T являются независимыми нормальными с.в. с математическими ожиданиями $E\{I\} = 10a$, $E\{R\} = 30$ ом, $E\{T\} = 10$ мин и срединными отклонениями $E_I = 0.1\ a$, $E_R = 0.2$ ом, $E_T = 0.5$ сек. Найти приближенное значение срединного отклонения с.в. Q.

107. Частота основного тона струны определяется формулой

$$\Omega = 1/2\sqrt{\frac{P}{ML}},$$

где P - сила натяжения, M - масса струны, L - длина струны.

Известны м.о. μ_P, μ_M, μ_L и средние квадратические отклонения σ_P, σ_M и σ_L. Определить рассеивание частоты основного тона струны из-за разброса силы натяжения, массы и длины струны, если соответствующие коэффициенты корреляции равны r_{PL}, r_{PM}, r_{ML} .

108. Сопротивление участка электрической цепи определяется формулой

$$Z = \sqrt{R^2 + \left(\Omega L - \frac{1}{\Omega C}\right)^2} \ ,$$

где R - омическое сопротивление, L - индуктивность проводника тока, C - его емкость, Ω - частота тока.

Определить срединную ошибку в величине сопротивления из-за ошибок при независимых измерениях R, L, C и Ω, если заданы μ_R, μ_L, μ_C, μ_Ω и срединные отклонения E_R, E_L, E_C, E_Ω.

109. При параллельном соединении элементов сила тока в цепи определяется формулой

$$I = E / (R + W/n) \ ,$$

где E - ЭДС элемента, W - его внутреннее сопротивление, n - число элементов, R - сопротивление внешней части цепи.

Пользуясь методом линеаризации, определить м.о. и дисперсию силы тока, если с.в. E, R и W независимы, а μ_E, μ_R, μ_W и σ_E, σ_R, σ_W заданы.

110. Найти приближенное значение средних квадратических отклонений прямоугольных координат случайной точки

$$X = H \, ctg \, \varepsilon \, \cos \beta \, ,$$
$$Y = H \, ctg \, \varepsilon \, \sin \beta \, ,$$
$$Z = H \, ,$$

если с.в. H, ε и β независимы, а м.о. и средние квадратические отклонения их соответственно равны: $\mu_H = 6200$ м, μ_ε = 45°, μ_β = 30°, σ_H = 25 м, $\sigma_\beta = \sigma_\varepsilon$ = 0.001 рад.

111. Переход от сферических координат к декартовым производится по формулам

$$X = R \sin \theta \, \cos \varphi \, ,$$
$$Y = R \sin \theta \, \sin \varphi \, ,$$
$$Z = R \cos \theta \ .$$

Ошибки в определении R, θ и φ независимы со средними квадратическими отклонениями σ_R = 10 м, $\sigma_\theta = \sigma_\varphi$ = 0.001 рад. Определить приближенное значение средних квадратических ошибок прямоугольных координат, если $\mu_\theta = \mu_\varphi$ = 45°, μ_R = 10000 м.

112. Пусть $Z = \sin X Y$, где X и Y - независимые с.в. Найти приближенное значение σ_Z, если $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 0.001$.

113. Используя метод линеаризации, определить приближенное значение дисперсии с.в. $Z = \sqrt{kX^2 + Y^2}$, где $X = \sin V$, $Y = \cos V$, с.в. V распределена равномерно на отрезке (0., $\pi/2$), а k - известная постоянная.

114. Для измерения объема конуса измерены: а) диаметр основания и высота; б) диаметр основания и длина образующей . В каком из этих двух случаев дисперсия ошибки определения объема конуса меньше, если м.о. высоты конуса $\mu_h = 8$ дм, диаметра основания $\mu_d = 12$ дм, длины образующей $\mu_l = 10$ дм, а $\sigma_h = \sigma_d = \sigma_l = 0.1$ дм ?

115. Случайная величина X подчиняется нормальному закону с ф.п.в.

$$f(x) = \frac{1}{10\sqrt{2\pi}} \exp\left\{-\frac{(x+5)^2}{200}\right\} .$$

Определить приближенное значение м.о. и дисперсии с.в. $Y = 1 / X$, учитывая первые два и три члена разложения в ряд Тейлора.

116. Учитывая три первых члена разложения функции $Y = \varphi(X)$ в ряд Тейлора, определить м.о. и дисперсию с.в. Y , если X нормальна.

117. Высота горной вершины H определяется по наклонной дальности D и углу места ε:

$$H = D \sin \varepsilon .$$

Найти приближенное значение срединной ошибки определения высоты, если $E_D = 80$ м, $E_\varepsilon = 0.001$, а наивероятнейшие значения соответственно равны $\mu_d = 12300$ м и $\mu_\varepsilon = 31.2°$ (с.в. D и ε независимы и нормальны).

118. Построить графики $f(x)$ – функции плотности нормального $N(\mu, \sigma^2)$ распределения со значениями параметров: (а) $\mu = 0.$, $\sigma = 1.$; (б) $\mu = -5.$, $\sigma = 1.$; (в) $\mu = +5.$, $\sigma = 1.$; (г) $\mu = 2.5$, $\sigma = 0.3$; (д) $\mu = -2.5$, $\sigma = 3$. Интерпретируйте смысл параметров нормального распределения, объясните особенности поведения $f(x)$.

119. Вычислить математическое ожидание и дисперсию случайной величины ξ , имеющей равномерное распределение $R(\mu, \omega)$ с функцией плотности вероятностей

$$f(x) = \begin{cases} \frac{1}{\omega} & , x \in [\mu - \frac{\omega}{2}, \mu + \frac{\omega}{2}]; \\ 0 & , x \notin [\mu - \frac{\omega}{2}, \mu + \frac{\omega}{2}] \end{cases} .$$

120. Пусть $\xi \sim R(\mu, \omega)$. Показать, что с.в.

$$\eta = (\xi - \mu + \omega/2) / \omega \sim R(1/2,\ 1).$$

121. Пусть $\xi \sim R(1/2,\ 1)$. Доказать, что с.в. $\eta = -2\ln\xi$ имеет распределение $\chi_2^{\ 2}$ (Хи-квадрат с двумя степенями свободы) с функцией плотности вероятностей

$$f(y) = \frac{1}{2}\exp\{-y/2\} .$$

122. Показать, что если ξ – с.в., имеющая нормальное распределение $N(\mu, \sigma^2)$, то

$$E\{ | \xi - \mu | \} = \sqrt{\frac{2}{\pi}} \cdot \sigma .$$

123. Два независимых равноточных измерения угла дали результаты (в °) 20.76 и 20.98. Другие шесть независимых и равноточных измерений были выполнены с помощью другого прибора и дали результаты 21.64, 21.54, 22.32, 20.56, 21.43 и 21.07. Случайные ошибки распределены нормально, и известно, что первый прибор менее точен, чем второй, в том смысле, что ему соответствует дисперсия, превышающая в 4 раза дисперсию второго прибора. Найти доверительные интервалы для разности систематических ошибок этих приборов. Уровень доверия $\gamma = 0.95$.

124. Предполагается, что применение новой технологии в производстве микросхем приведёт к увеличению выхода годной продукции. Результаты контроля двух партий продукции, изготовленных по старой и новой технологии, приведены ниже.
Подтверждают ли эти результаты предположение об увеличении выхода готовой продукции ? Принять $\alpha = 0.01$.

изделия	Технология	
	старая	новая
годные	140	185
негодные	10	15
Всего	150	200

125. Две лаборатории измеряли концентрацию (в %) серы в стандартных образцах дизельного топлива. Шесть независимых измерений в первой лаборатории дали 0.869, 0.874, 0.867, 0.875, 0.870, 0.869. В пяти подобных измерениях во второй лаборатории получено 0.865, 0.870, 0.866, 0.871, 0.868. В предположении нормального распределения ошибок постройте 90%–ный доверительный интервал для отношения измерений дисперсии в двух лабораториях. Если есть основания считать, что дисперсии одинаковы, постройте доверительный интервал для разности систематических ошибок в обоих лабораториях.

126. Пусть вероятность появления “орла” при бросании монеты есть $p \in (0,\ 1)$. В эксперименте де Бюффона “орёл” выпал $k = 2048$ раз в $n = 4040$ бросаниях монеты. Нужно проверить, согласуются ли результаты этого опыта с гипотезой H_0 , что монета симметричная. Использовать уровень значимости (а) 0.05; (б) 0.1.

127. В $n = 4000$ независимых испытаниях события A_1 , A_2 , A_3 , которые составляют полную группу, произошли 1905, 1015, 1080 раз, соответственно. Выбрав уровень значимости 0.05, проверьте, согласуются ли эти данные с гипотезой H_0 : $p_1 = 1/2, p_2 = p_3 = 1/4,$ где $p_i = P\{A_i\}$.

128. Проверяют 500 случайно выбранных в магазине часов. Пусть i – номер интервала между i–ым и $(i + 1)$–ым часом, $i = 0, ..., 11,$ и пусть h_i – число часов, показывающих i–ый интервал. Результаты наблюдений сгруппированы в следующей таблице:

i	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Всего**
h_i	41	34	54	39	49	45	41	33	37	41	47	39	n= 500

Проверяется гипотеза H_0 , что время, показываемое часами, равномерно распределено на интервале (0, 12). Уровень значимости $\alpha = 0.1$. Верна ли H_0 ? Для какого α гипотеза H_0 принимается ?

129. Случайно выбираются числа 0,1,...,9. В n = 10000 испытаний числа, не превосходящие 4, встретились h = 5089 раз. Проверить гипотезу H_0 , что числа появляются равновероятно. Для какого уровня значимости следует отвергнуть H_0 ?

130. Число π, записанное в десятичном виде, содержит в первых 10002 позициях после запятой цифры 0, 1, ..., 9 соответственно 968, 1026, 1021, 974, 1014, 1046, 1021, 970, 948, 1014 раз. Можем ли мы считать для уровня значимости $\alpha = 0.05$ верной гипотезу H_0 о том, что эти цифры – случайные равновероятные ? Для какого уровня значимости следует отвергнуть H_0 ?

131. Эксперимент состоит в бросаниях двенадцати игральных костей. Наблюдаемая с.в. ξ равна числу костей, на которых выпало 4, 5 или 6 очков. Пусть h_i – число испытаний, в которых наблюдались значения $\xi = i$, $i = 0, 1, ..., 12$. Для $n = 4096$ испытаний получилась следующая таблица:

i	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
$\mathbf{h_i}$	0	7	60	198	430	731	948	847	536	257	71	11	0

Всего $n = 4096$.

Согласуются ли эти данные с гипотезой о том, что кость симметричная ?

132. Пусть в условиях предыдущей задачи с.в. η равна числу костей с шестью очками. Полученные результаты табулированы:

i	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**≥ 7**	**Всего**
$\mathbf{h_i}$	447	1145	1181	796	380	115	24	8	n=4096

Теперь H_0 : $\eta \sim Bi(12, 1/6)$. Верна ли H_0 ?

133. В своей классической работе Мендель рассматривал частоты различных горошин, полученных при гибридизации желтых и сморщенных зелёных горошин. Эти данные и соответствующие вероятности, которые предсказывала теория наследственности Менделя, даны в следующей таблице:

Горошины	**Частота**	**Вероятность**
Круглая и жёлтая	315	9/16
Сморщенная и жёлтая	101	3/16
Круглая и зелёная	108	3/16
Сморщенная и зелёная	32	1/16
Всего	n = 556	1

Проверьте гипотезу H_0 , что частоты согласуются с теоретическими вероятностями (при уровне значимости $\alpha \le 0.9$).

134. Приведённая ниже таблица даёт число m_i участков площадью 0.25 км ^{2} в южной части Лондона, каждый из которых подвергался бомбардировке i раз в течение II–й мировой войны. Проверьте гипотезу H_0 , что эти данные согласуются с законом распределения Пуассона на уровне значимости $\alpha = 0.05$.

i	*0*	*1*	*2*	*3*	*4*	*≥5*	*Всего*
m_i	229	211	93	35	7	1	576

135. Из 2020 семей с двумя детьми 527 семей имеют двух мальчиков, 476 семей – двух девочек, а оставшиеся 1017 семей – по мальчику и девочке. Дан уровень значимости $\alpha = 0.05$. Верна ли гипотеза, что число мальчиков в семьях с двумя детьми есть биномиальная с.в. ?

136. Среди 2000 пациентов в течение промежутка наблюдений 181 один раз болел гриппом, 9 – дважды, а остальные 1810 были здоровы. Согласуются ли эти данные с гипотезой H_0 о том, что число заболеваний у пациентов является биномиальной с.в. ? (Уровень значимости $\alpha = 0.05$).

137. Поступающие в университет разбиты на две группы по 300 человек в каждой группе. В первой группе 33, 43, 80, 144 человека получили оценки “2”, “3”, “4”, “5”, соответственно. Данные для второй группы – 39, 35, 72, 154 человека, соответственно. Можем ли мы рассматривать обе группы однородными на уровне значимости $\alpha = 0.05$?

138. Таблица

n_j	*1072*	*1133*	*2455*	*1195*
ν_j	22	23	49	33

даёт данные по числу смертей матерей, имеющих своего первого ребёнка, в четыре периода времени (n_j – число матерей, ν_j – число смертей). Проверьте гипотезу о том, что вероятность смерти в эти периоды одна и та же.

139. Проверьте гипотезу независимости для следующей двумерной таблицы сопряжённости признаков:

		ξ_2			
		b_1	b_2	b_3	Σ
ξ_1	a_1	3009	2832	3008	8849
	a_2	3047	3051	2997	9095
	a_3	2974	3038	3018	9030
	Σ	9030	8921	9023	26974

Уровень значимости $\alpha = 0.05$.

140. Из 300 поступающих в университет, которые выдержали вступительный экзамен, 97 получили “5” в школе, 48 получили “5” на вступительном экзамене, но только 18 получили “5” как в школе, так и на экзамене. Проверьте гипотезу H_0 о том, что школьные отметки и результаты экзамена в университете независимы (уровень значимости $\alpha = 0.1$).

141. Задан уровень значимости $\alpha = 0.001$. Рассматривается последовательность чисел 1.05, 1.12, 1.37, 1.50, 1.51, 1.73, 1.85, 1.98. Может ли она быть реализацией случайного вектора, все восемь компонент которого – независимые одинаково распределённые с.в. ?

142. Воспользуйтесь программой ERR1.EXE для генерации серии $n = 100$ равномерно распределённых случайных чисел. С помощью χ^2 –критерия (программа CHI_SQR.EXE) с $k = 5$ проверьте качество генератора случайных чисел (Бины разбиения одинаковы).

143. Примените критерий Колмогорова вместо χ^2 –критерия в предыдущей задаче (программа KOLMOGOR.EXE).

144. Сделайте то же самое (т.е. исследование задач 142 и 143) для 100 нормально $N(0\ ,\ 1)$ распределённых случайных чисел (программы ERR1.EXE и KOLMOGOR.EXE). <*Указание:* для χ^2 –критерия возьмите равновероятные интервалы разбиения. Для $k = 5$ эти интервалы суть: ($-\infty, -0.8416), (-0.8416, -0.25335), (-0.25335, +0.25335), (0.25335, 0.8416), (0.8416, +\infty)$.>

145. С помощью программы ERR1.EXE получите выборку объёма $n = 48$. Разбейте её на две совокупности $\{X_{2i}\}$, $i = 1, 2, ..., 24$ и $\{X_{2i+1}\}$, $i = 0, 1, ..., 23$ чётных и нечётных номеров. С помощью критерия Смирнова

$$\sup |F_{1,n}(z) - F_{2,m}(z)|,$$

где $F_{1,n}(z)$ и $F_{2,m}(z)$ – эмпирические функции распределения для двух выборок $x_1, \ldots, x_n$ и $y_1, \ldots, y_m$ (в нашем случае $n = m$) проверьте гипотезу, что эти две выборки взяты из одного и того же распределения.

146. В результате проведения стандартной процедуры проверки коэффициента упругости образцов резины установлено, что среднее квадратическое отклонение при измерении этого коэффициента составляет 18.0 единиц. взята выборка объёма 20 и получено $s = 23.2$ единицы. Обосновано ли предположение о нестабильности стандартной процедуры проверки коэффициента упругости ?

147. Чтобы выяснить, варьирует ли от одного дня к другому величина изменчивости температуры высокоскоростного аппарата, в первый день было проведено 12 измерений, а во второй 10. Среднее квадратическое отклонение в первый день (A) составило 23°, а во второй 30°. Можно ли сказать, что различий в дисперсиях, измеренных в разные дни, не обнаружено ?

148. Табачная фирма хотела бы знать, можно ли отправлять заказчикам сигареты и трубочный табак в одной упаковке. Если при этом качество сигарет или трубочного табака не ухудшится, то можно существенно сократить затраты по перевозке. Признаком ухудшения качества сигарет является изменение их аромата (возможно, из-за сильного запаха трубочного табака). Для проведения исследований изготовили 400 картонных коробок и в 250 из них положили табачные изделия обоих типов. В оставшиеся 150 коробок были положены только сигареты. Через месяц коробки открыли, и все 400 упаковок сигарет расположили в случайном порядке. Несколько экспертов анализировали аромат сигарет и пытались обнаружить его отличие от (предполагаемого) исходного. Результаты экспертизы приведены в таблице. Можно ли сказать, что связь между ароматом сигарет и видом упаковки отсутствует ?

Мнение об аромате	**Вид упаковки** **Совместная**	 **Раздельная**	**Всего**
Не изменился	72	119	191
Изменился	178	31	209
Всего	250	150	400

149. Среднее время выхода кислородного конвертера на температурный режим для двух последовательных дней составило 22.0 и 20.2 мин.

а) Определите 99%–ный доверительный интервал для математического ожидания изменения этого времени от одного дня к другому, если известно, что дисперсия каждого из этих выборочных средних составляет 2.00 мин ^{2}.

б) Подтверждает ли этот доверительный интервал предположение о равенстве средних значений в разные дни ?

150. Найдите 99% –ный доверительный интервал для разности средних значений времени измерения для контролёров A и D из задачи 37.

151. Проверьте по данным, приведённым в задаче 39, гипотезу об однородности дисперсий. Определите 95% –ный доверительный интервал для общего значения сопротивления, если принята нулевая гипотеза.

152. Полный вес упаковки с натуральным красителем должен составлять 22.680 ± 0.0567 кг. Известно, что дисперсия равна $7.3139 \cdot 10^{-4}$ кг. Приведённые ниже результаты получены при исследовании выборки 20 упаковок (вес указан в кг).

а) Соответствует ли средний вес упаковки, вычисленный по приведённым ниже данным, предъявляемым требованиям ?

б) Какие нужно сделать предположения ?

(Извините за обилие знаков после запятой – это получилось в результате перевода результатов измерений в фунтах в метрическую систему).

22.72486	22.57064	22.66589	22.63414
22.71125	22.71579	22.62507	22.72486
22.64775	22.76115	22.78382	22.73847
22.59332	22.58425	22.81558	22.6296
22.54342	23.50503	22.78836	22.68404

153. На двух станках производят одинаковую продукцию. Критическим размером изделий является внешний диаметр. Установлено, что за один и тот же период времени дисперсия этой величины для первого станка А составила 1.07 мм ^{2}, а для станка В – 0.84 мм ^{2}. Со станка А была взята выборка, включающая 15 изделий, а со станка В – выборка из 10 изделий. Можно ли утверждать, что математические ожидания исследуемых величин равны между собой, если $\bar{x}_A = 45.3$ мм, а $\bar{x}_B = 46.1$ мм ?

154. Анализ ранних произведений Жозефа Зайлича показал, что предложения, включающие разное число слов, распределены в них следующим образом:

Число слов	**2-3**	**4-5**	**6-8**	**9-12**	**13-16**	**17-20**	**21-24**	**25-28**	**Свыше 28**
Доля предложений	0.010	0.034	0.067	0.091	0.210	0.174	0.181	0.143	0.09

Джо Зилч, наследник автора, заявил о находке рукописи ранее не опубликованного произведения своего предка. В выборке, состоящей из 2000 предложений, было обнаружено следующее распределение фраз по длине :

Число слов	**2-3**	**4-5**	**6-8**	**9-12**	**13-16**	**17-20**	**21-24**	**25-28**	**Свыше 28**
Доля предложений	15	51	118	227	476	401	352	239	121

Проверьте, могут ли эти данные подтвердить заявление Джо Зилча.

155. Определите 95% –ный доверительный интервал для разности двух средних значений $\mu_A - \mu_B$, если $\bar{x}_A = 13.8$, $\bar{x}_B = 16.2$, $n_A = 15$, $n_B = 20$, а также известно, что дисперсии сравниваемых совокупностей равны $\sigma^2 = 4.41$. Вычислите 95% –ные доверительные интервалы раздельно для μ_A и μ_B Сравните первый из доверительных интервалов с двумя последними.

156. Выборка А обладает дисперсией 55.4 фунт ^{2}, а выборка В – дисперсией 87.3 фунт ^{2}.

а) Можно ли сказать, что выборки взяты из совокупностей с одним и тем же значением дисперсии, если их объёмы равны соответственно 15 и 12?

б) Какие при этом необходимо сделать предположения ?

157. Исследован предел прочности на разрыв у шести плавок титанового сплава. Из слитков каждой плавки было изготовлено 5 образцов, и на них проверена однородность дисперсий предела прочности. Шесть значений дисперсий, измеренных в 10000 фунт/дюйм ^{2}, оказались равными: 16.1; 30.3; 36.4; 7.3; 40.6 и 39.5.

а) Можно ли сказать, что все шесть дисперсий одинаковы ?

б) Какие при этом необходимо сделать предположения ?

в) Считая дисперсии однородными, найдите 95% –ный доверительный интервал для предполагаемой величины общего значения дисперсии.

158. Температура в автоклаве регистрируется через равные промежутки времени. Для проведения некоторого эксперимента потребовалось поддерживать заданную температуру. Температура регистрировалась в течение двух последовательных дней в случайные моменты времени. В первый день было зафиксировано 16 значений температуры со средним квадратическим отклонением 15.6, а во второй день – 21 значение со средним квадратическим отклонением 9.8. Можно ли утверждать, что наблюдения относятся к одной совокупности ?

159. Пяти лабораториям было поручено участвовать в проведении химического анализа образцов каменного угля с целью определения содержания в них золы. Один образец был расколот на сорок частей, и в каждую из лабораторий отправили по восемь кусков. Дисперсии результатов измерений в разных лабораториях получились следующими: 3.86; 4.27; 1.35; 3.90 и 1.64. Можно ли отклонить гипотезу об однородности дисперсий ?

160. Одним из методов количественного анализа величины износа шины является измерение глубины проникновения щупа в канавку на рисунке протектора в определённом месте шины. Есть подозрение, что значительная часть дисперсии измерений связана с действиями контролёров. Чтобы выделить из общей дисперсии указанную часть, трёх контролёров попросили провести по 12 независимых измерений в одной и той же точке. Результаты измерений приведены ниже.

Контролёр X			Контролёр Y			Контролёр Зилч		
121	121	126	120	129	128	141	111	118
130	127	131	136	117	112	110	135	134
127	124	125	138	124	119	134	113	123
119	126	123	136	135	134	113	139	129

а) Однородны ли дисперсии измерений, проведённых разными контролёрами ?

б) Отличается ли дисперсия измерений, проведённых контролёром Зилчем, от дисперсий, которые имеют место при измерениях, проводимых контролёрами X и Y, если предположить, что измерения, проводимые двумя последними, характеризуются одинаковой дисперсией ?

в) Объясните различие между а) и б).

161. Проверьте, есть ли выбросы в результатах измерений контролёра Зилча, приведённых в задаче 160 (используйте $\alpha = 0.05$).

162. На одном крупном газовом месторождении фирма за последние 10 лет пробурила 75 скважин. Ниже приведены данные о распределении запасов газа на скважину в единицах 10^9 фут^3.

x	f	x	f
0 – 1.999	30	10 – 11.999	1
2 – 3.999	15	12 – 13.999	2
4 – 5.999	11	14 – 15.999	1
6 – 7.999	9	16 – 17.999	2
8 – 9.999	3	....	
		26 – 27.999	1 .

Подберите по этим данным экспоненциальное распределение

$$p_X(x) = \theta e^{-\theta x} \ (0 \le x < \infty),$$

оценивая θ как $(\overline{X})^{-1}$, вычислите теоретические частоты. Используя критерий согласия, исследуйте эти данные.

163. Исследовалась сила сцепления при комнатной температуре двух типов клейких веществ: изобутила 2-цианакрилата и MBR-4197. Приведённые ниже данные представляют собой значения силы сцепления. измеренные в фунтах/дюйм^2 для каждого из образцов:

Изобут.	MBR	Изобут.	MBR	Изобут.	MBR	Изобут.	MBR
365	518	146	419	222	461	227	477
169	403	163	437	121	437	250	498
210	473	213	396	205	402	218	330
297	329	300	424	283	346	202	352
228	457	218	363	134	571	234	495

а) Проверьте гипотезу о том, что выборки имеют одинаковые дисперсии.

б) Взяты ли эти выборки из совокупностей с одинаковыми средними значениями ?

в) Какие предположения были сделаны в пп. а) и б) ? Подтверждаются ли они ?

164. В рамках исследований, описанных в задаче 168, для дополнительной проверки было испытано ещё 20 образцов клейких веществ каждого типа. Клейкое вещество на каждый образец было нанесено за 24 часа до испытаний. Результаты приводятся ниже.

Изобут.	MBR	Изобут.	MBR	Изобут.	MBR	Изобут.	MBR
311	517	298	419	328	465	358	477
330	403	297	437	362	442	329	498
249	472	353	396	260	402	332	330
351	329	319	424	543	346	324	352
335	457	293	363	308	571	328	495

а) Можно ли утверждать, что выдержка существенно влияет на силу сцепления ?

б) Определите, что понимать под словами “существенно влияет”.

в) Определите по двум выборкам, есть ли различия в средней силе сцепления у сравниваемых веществ после выдержки.

165. Два оператора провели 14 независимых опытов по исследованию температуры воспламенения эмали одного состава. Каждый оператор проверил семь образцов. Ниже приведены результаты опытов.

Оператор А	1425	1420	1410	1370	1360	1270	1450
Оператор В	1420	1380	1320	1320	1290	1280	1430

а) Наблюдается ли значимое различие между средними значениями результатов, полученных разными операторами ?

б) Какая проверялась гипотеза, какие были сделаны предположения и принят уровень значимости ?

166. Для того чтобы определить, сокращается ли время сварки на отливках, если при литье вместо сырой формовочной смеси использовать сухую смесь или формовочную смесь с CO_2, было проведено специальное исследование. Совершенно ясно, что стоимость литья в случае сухой формовочной смеси или смеси с CO_2 выше, но есть мнение, что это может быть оправдано, если стоимость сварки значимо уменьшится. Ниже приведены значения времени сварки в минутах при использовании формовочных смесей разных типов.

а) Можно ли сказать, что имеет место значимое уменьшение времени сварки, если предположить, что все факторы, кроме рассматриваемого, поддерживались на одном уровне ?

б) Какие нужно сделать добавочные предположения ?

Сырая смесь	19	28	14	29	15
Сухая смесь	21	15	11	12	21
Смесь с CO_2	14	20	15	19	17

167. Проверить независимость между (1) умственными способностями школьников и (2) качеством их одежды. Классифицированы 1725 школьников. Для первого признака были использованы следующие градации: A - умственно отсталый, B - медлительный и тупой, C - тупой, D - медлительный, но умный, E - достаточно умный, F - явно способный, G - очень способный.

Способности /Как одевается	**A и B**	**C**	**D**	**E**	**F**	**G**	**Сумма**
Очень хорошо	33	48	113	209	194	39	636
Хорошо	41	100	202	255	138	15	751
Сносно	39	58	70	61	33	4	265
Очень плохо	17	13	22	10	10	1	73
Сумма	130	219	407	535	375	59	1725

168. В следующей таблице приведены результаты экспериментов по иммунизации крупного рогатого скота против туберкулёза.

Проверить, подтверждается ли гипотеза о независимости между вакцинацией и восприимчивостью к туберкулёзу (уровень значимости $\alpha = 0.1$).

	Число погибших или серьёзно пострадавших от туберкулёза	**Число легко пострадавших или не пострадавших**	**Сумма**
Привитые вакциной	6	13	19
Не привитые или привитые контрольными средствами	8	3	11
Сумма	14	16	30

169. Рассмотрим следующий эксперимент по определению водопроницаемости листов строительного материала, изготавливаемого двумя различными машинами. Из накопленного опыта известно, что логарифм водопроницаемости распределён приближённо нормально и что вариабельность от машины к машине также распределена приближённо нормально. Логарифмы водопроницаемости имели следующие значения:

x (машина I)	1.845	1.790	2.042
y (машина II)	1.583	1.627	1.282

Проверяем гипотезу H, согласно которой математическое ожидание логарифма водопроницаемости для обеих машин совпадает. Уровень значимости $\alpha = 0.05$.

Постройте также 95%-ный доверительный интервал для разности математических ожиданий логарифма водопроницаемости.

170. Следующие данные получены в эксперименте по изучению зависимости между производством фуражных кормов весной и количеством органических удобрений, вносимых в почву предыдущей осенью. Контрольные измерения (x) соответствуют 0 фунтов удобрений на акр, измерения на обработанных участках (y) соответствуют 500 фунтам удобрений на акр. Производство фуражных кормов также измерялось в фунтах на акр:

x	794	1800	576	411	897
y	2012	2477	3498	2092	1808

Предположим, что справедлива нормальная модель с равными дисперсиями и двумя выборками.

а) Найти 95%-ный доверительный интервал для $\mu_2 - \mu_1$.

б) Можно ли утверждать, что внесение 500 фунтов удобрений на акр значительно улучшает производство кормов ? Воспользуйтесь $\alpha = 0.05$.

в) Найти 90%-ный доверительный интервал для σ, пользуясь центральной статистикой $\frac{s^2}{\sigma^2}$.

171. При скрещивании двух типов кукурузы Линдштром обнаружил во втором поколении четыре различных типа растений. Простая менделевская модель предсказывает появление четырёх типов с вероятностями 9/16, 3/16, 3/16 и 1/16. Линдштром произвёл наблюдения над 1301 растением и получил $N_1 = 773$, $N_2 = 231$, $N_3 = 238$, $N_4 = 59$. Пользуясь критерием Хи-квадрат, проверьте применимость модели Менделя к данным Линдштрома.

172. Для проверки подлинности рукописи (писем Квинта Куртиуса Снодграсса), приписываемой Марку Твену, Ч.С. Бринегар произвёл подсчёт числа слов длиной $k = 1, \ldots, 12$ в произведениях Твена и в письмах Снодграсса. Выяснилось, что слова из 2, 3 и 4 букв встречаются в произведениях Твена с частотой (вероятностью):

k	2	3	4	Остальные
p_k	0.177	0.232	0.191	0.400

В 13175 словах писем Снодграсса слова той же длины встречаются с частотами, приведёнными в следующей таблице.

Проверить по критерию Хи-квадрат гипотезу, согласно которой письма

k	2	3	4	Остальные
N_k	2685	2752	2302	5436

Снодграсса представляют собой случайную выборку из произведений Твена.

173. По мнению врачей, приём некоторого витамина как-то сказывается на профилактике простудных заболеваний. Проведён следующий эксперимент: 200 человек случайным образом разделены на две равные группы, одной группе дали витамин, а другой – "пустышку", но всем 200 подопытным сказано, что им дан витамин. Результаты обследования приведены в таблице:

	Меньше простудных заболеваний	Больше простудных заболеваний	Без изменений
Контрольная группа	39	21	40
Группа, принимавшая витамин	51	20	29

Проверить на 5%-ном уровне значимости гипотезу независимости простудных заболеваний от приёма витамина.

174. (Дрейпер, Смит: гл. 1, зад. 3) Тринадцать образцов медно-никелевых сплавов, каждый с определённым содержанием железа, были испытаны на коррозию в установке с колесом. Колесо вращалось в солёной морской воде со скоростью 30 футов/с в течение 60 дней. Коррозия определялась по потере в весе в миллиграммах/дм ^{2} в день (МДД). Были собраны следующие данные: (см. табл.).

Определить, можно ли описать в рамках линейной модели влияние содержание содержания железа на сопротивление коррозии сплавов Cu-Ni в морской воде. Принять уровень значимости $\alpha = 0.05$.

x (Fe)	y (потери в МДД)	x (Fe)	y (потери в МДД)
0.01	127.6	1.44	93.3
0.48	124.0	0.71	113.1
0.71	110.8	1.96	83.7
0.95	103.9	0.01	128.0
1.19	101.5	1.44	91.4
0.01	130.1	1.96	86.2
0.48	122.0		

175. (Дрейпер, Смит: гл. 1, зад. 9) Считается, что число консервных банок, повреждённых при перевозках в товарных вагонах – это функция скорости вагонов при толчках. Методом случайного отбора были выбраны 13 вагонов для проверки того, насколько это соответствует действительности. Собраны следующие данные:

Скорость вагонов (x)	Число повреждённых банок (y)	Скорость вагонов (x)	Число повреждённых банок (y)
4	27	4	75
3	54	3	53
5	86	5	33
8	136	7	168
4	65	3	47
3	109	8	52
3	28		

Каковы ваши выводы ? (возьмите $\alpha = 0.05$).

176. (Дрейпер, Смит: гл. 1, зад. 10) Экспериментально было установлено влияние температуры процесса дезодорации на цвет конечного продукта.

Температура	Цвет	Температура	Цвет	Температура	Цвет
460	0.3	410	0.5	410	0.7
450	0.3	450	0.5	400	0.6
440	0.4	440	0.6	420	0.6
430	0.4	430	0.6	410	0.6
420	0.6	420	0.6	400	0.6

Получены следующие данные:

1) Постройте модель $y = \theta_0 + \theta_1 x + \varepsilon$.

2) Имеет ли эта модель смысл ? (используйте $\alpha = 0.05$).

177. (Дрейпер, Смит: гл. 4, зад. 5) Величина зазора затворной пластинки и её температура влияют на процент упаковок, которые благополучно проходят контрольную проверку на машинах по упаковке мыла. Были собраны некоторые данные об этих факторах; они приведены в таблице:

Затвор затворной пластинки x_1	Температура затворной пластинки x_2	Процент заготовок, проходящих зазор, y	x_1	x_2	y
130	190	35.0	139	240	56.7
174	176	81.7	188	230	84.4
134	205	42.5	175	200	94.3
191	210	98.3	156	218	44.3
165	230	52.7	190	220	83.3
194	192	82.0	178	210	91.4
143	220	34.5	132	208	43.5
186	235	94.5	148	225	51.7

1) Используется линейная модель $y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \varepsilon$; определите оценки наименьших квадратов для θ_0, θ_1 и θ_2.

2) Значима ли регрессия в целом (используйте $\alpha = 0.05$).

3) Можно ли сказать, что одна из двух переменных более полезна, чем другая, для предсказания отклика ?

4) Какие рекомендации вы можете дать относительно работы упаковочной машины ?

178. (Дрейпер, Смит: гл. 1, зад. 8) При организации производства некоторого продукта было предложено заменить анализ потерь в бутылках анализом потерь в чашках, который дороже, чем применяемая лабораторная методика. Зато анализ потерь в чашках дал бы возможность лучше управлять процессом, так как он экономит время. Если бы удалось доказать, что потери в чашках – функция потерь в бутылках, то стоило бы принять решение о переходе к анализу потерь в чашках. Имеются следующие данные:

Потери в бутылках, % X	Потери в чашках, % Y	Потери в бутылках, % X	Потери в чашках, % Y	Потери в бутылках, % X	Потери в чашках, % Y
3.0	3.1	2.7	3.6	3.2	4.1
3.1	3.9	3.1	3.1	2.1	2.6
3.0	3.4	2.7	3.6	3.0	3.1
3.6	4.0	2.7	2.9	2.6	2.8
3.8	3.6	3.3	3.6		

Каковы ваши выводы ? (Возьмите $\alpha = 0.05$).

179. (Дрейпер, Смит: гл. 1, зад. 11) Приведённые ниже данные содержат 34 пары значений переменных x и y, где

x – среднее арифметическое значение уровня афлатоксина <токсин, вырабатываемый отдельными видами плесени; есть основания подозревать, что он является канцерогеном. - *Прим. перев.*> в единичной минимальной пробе земляного ореха, равной 120 фунтам (единиц на миллиард),

y – процент заражённых орехов в партии орехов.

Y	*x*	*y*	*x*	*y*	*x*
0.971	3.0	0.942	18.8	0.863	46.8
0.979	4.7	0.932	18.9	0.811	46.8
0.982	8.3	0.908	21.7	0.877	58.1
0.971	9.3	0.970	21.9	0.798	62.3
0.957	9.9	0.985	22.8	0.855	70.6
0.961	11.0	0.933	24.2	0.788	71.1
0.956	12.3	0.858	25.8	0.821	71.3
0.972	12.5	0.987	30.6	0.830	83.2
0.889	12.6	0.958	36.2	0.718	83.6
0.961	15.9	0.909	39.8	0.642	99.5
0.982	16.7	0.859	14.3	0.658	111.2
0.975	18.8				

1) Нанесите данные на график и без всяких вычислений "на глазок" проведите прямую, которая вам кажется "самой лучшей прямой, какую только можно провести через данные точки". Сохраните этот рисунок, он вам понадобится позже для сравнения. Считаете ли вы, что ваша прямая – это "хорошее приближение" ?
2) Теперь вычислите $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum y_i^2$, $\sum x_i y_i$, где все суммы берутся от $i = 1$ до 34.
3) Постройте методом наименьших квадратов модель $y = \beta_0 + \beta_1 * x + \varepsilon$. Нанесите эту прямую на ваш график и проверьте, хороша ли была "подгонка на глазок" ?
4) Найдите остатки, сохраняя три знака после запятой. Убедитесь, что $\sum e_i = 0$ с точностью до ошибки округления.
5) Постройте таблицу дисперсионного анализа.
6) Найдите стандартные ошибки коэффициентов b_0 и b_1 .
7) Выведите формулу для стандартной ошибки значения $\hat{y}$ и постройте 95 %-ные доверительные границы для "истинного" среднего значения *y*.
8) Проверьте всю регрессию по *F*-критерию и установите, какую долю вариации относительно среднего $\bar{y}$ объясняет подобранная прямая.

180. (Дрейпер, Смит: гл. 1, зад. 15) На одном из слушаний Федеральной комиссии по энергетике (США) были представлены следующие данные:
Переменные таковы:

x	*y*	*x*	*y*
13.3	3.5	26.3	6.0
16.9	5.1	30.1	9.5
19.9	4.8	42.6	8.1
23.2	6.7		

предиктор *x* – процент жидкости в добываемом из скважин газе;
отклик *y* – единичная стоимость в центах процесса добычи газа.

Подберите по этим данным прямую и, пользуясь таблицей дисперсионного анализа, найдите остаточную вариацию.

Свидетели, приведшие эти данные, экстраполировали прямую до точек $x = 0$ и $x = 100$ для подсчёта единичных стоимостей в таких ситуациях, когда добыча вообще не содержит жидкостей, и, наоборот, когда она сплошь жидкая. Найдите значения $\hat{y}$ *(0)* и $\hat{y}$ *(100)* и определите 95 %-ные доверительные пределы для "истинного" среднего значения y при $x = 0$ и $x = 100$. Каковы ваши заключения ?

181. (Дрейпер, Смит: гл. 1, зад. 23) В газете The Chicago Maroon за пятницу 10 ноября 1972 г. сообщалось, что на оптовом рынке ожидаются следующие цены на марочные портвейны в расчёте на бутылку:

Год	Цена, дол.	Год	Цена, дол.	Год	Цена, дол.	Год	Цена, дол.
1890	50.00	1934	15.00	1941	10.00	1952	4.99
1900	35.00	1935	13.00	1944	5.99	1955	5.98
1920	25.00	1940	6.98	1948	8.98	1960	4.98
1931	11.98			1950	6.98		

1) Нанесите данные на график и исследуйте их. Имеет ли смысл строить регрессию для отклика "цена" от предиктора "год" ?
2) Какое преобразование переменной "год" имело бы смысл ? (Указание. Предположите, например, что сейчас 1972 г. Как вы обычно говорите о своём "годе рождения" ?) Постройте график зависимости цены от вашего нового предиктора и исследуйте его. Какого рода преобразование цены кажется здесь разумным, чтобы сделать данные "лучше ложащимися на прямую" ?
3) Постройте график для данных $y = \ln(\textit{цена})$ при $z = \textit{возраст бутылки}$. Методом наименьших квадратов подберите по этим данным прямую, вычислите остатки и постройте таблицу дисперсионного анализа (**ANOVA).**
4) Что вы можете сказать про цены на марочный портвейн на основе такого представления данных и вашего анализа ? Какого (с точностью до ближайшего цента) прироста цены за год вы ожидаете, если подобная структура цен сохранится и на будущее ?
5) В следующем объявлении в той же газете три года спустя, 25 ноября 1975 г., говорилось, что предлагается марочный портвейн 1937 г. по цене 20.00 долл. За бутылку. Если бы можно было считать, что прямолинейная зависимость сохранилась и что она приложима также и к этой новой точке, то каким представляется прирост цены в год на бутылку за истекшие три года ? Сохраняет ли силу ваш ответ на вопрос из пункта 4) или похоже, что ежегодное нарастание цены ускорилось ?

182. (Дрейпер, Смит: гл. 3, зад. 2) Для производителей мыла важное значение имеет такой показатель качества, как высота слоя пены ("мылкость"). Проведён эксперимент, в котором варьировалось количество мыла и измерялась высота пены в стандартном лотке при определённой степени перемешивания. Получены следующие данные:

Колич-во мыла в граммах, x	Высота пены, y	Колич-во мыла в граммах, x	Высота пены, y
4.0	38	6.0	53
4.5	42	6.5	61
5.0	45	7.0	62
5.5	51		

Допустим, что модель имеет вид: $y = \beta_0 + \beta_1 * x + \varepsilon$.

1) Получите наилучшее уравнение для предсказания.
2) Проверьте его статистическую значимость.
3) Найдите остатки и оцените, имеются ли какие-нибудь предпосылки, указывающие на то, что здесь лучше подходит более сложная модель.

183. (Дрейпер, Смит: гл. 3, зад. 3). При проведении эксперимента, аналогичного описанному в зад. 182, исследователь заметил, что "модель бессмысленна, если не принять, что $\beta_0 = 0$, поскольку каждый знает, что если не положить сколько-нибудь мыла в лоток, то не будет и пены". Таким образом, он настаивал не модели $y = \beta_1 * x + \varepsilon$.

Им получены следующие данные (в обозначениях зад. 182):

Колич-во мыла в граммах, x	Высота пены, y	Колич-во мыла в граммах, x	Высота пены, y
3.5	24.4	6.0	51.4
4.0	32.1	6.5	61.9
4.5	37.1	7.0	66.1
5.0	40.4	7.5	77.2
5.5	43.3	8.0	79.2

1) Исходя из модели, предложенной экспериментатором, найдите наилучшую оценку коэффициента β_1.
2) Используя это уравнение, найдите $\hat{y}$ для каждого значения x.
3) Исследуйте остатки.
4) Сделайте выводы и сформулируйте рекомендации для экспериментатора.

184. (Дрейпер, Смит: гл. 3, зад. 4). Приведённые ниже данные отражают соотношение между количеством кокаина в водном растворе и прозрачностью раствора по показаниям колориметра Постройте регрессионную зависимость на основе модели $y = \beta_0 + \beta_1 * x + \varepsilon$, получите остатки, исследуйте их и рассмотрите вопрос об адекватности модели.

Концентрация (мг/миллилитр), x	Показания колориметра, y	Концентрация (мг/миллилитр), x	Показания колориметра, y
40	69	90	415
50	175	40	72
60	272	60	265
70	335	80	492
80	490	50	180

185. ("Сквозная" задача в Дрейпере и Смите). В выпарном аппарате на промышленном предприятии y – количество используемого пара в фунтах ежемесячно. Изучается расход пара в зависимости от 9 предикторных переменных. В данной задаче рассматриваются 2 предикторные переменные: x_6 – число рабочих дней в месяце; x_8 – средняя температура воздуха. Получены следующие данные:

Номер наблюдения	x_8	x_6	y
1	35.3	20	10.98
2	29.7	20	11.13
3	30.8	23	12.51
4	58.8	20	8.40
5	61.4	21	9.27
6	71.3	22	8.73
7	74.4	11	6.36
8	76.7	23	8.50
9	70.7	21	7.82
10	57.5	20	9.14
11	46.4	20	8.24
12	28.9	21	12.19
13	28.1	21	11.88
14	39.1	19	9.57
15	46.8	23	10.94

16	48.5	20	9.58
17	59.3	22	10.09
18	70.0	22	8.11
19	70.0	11	6.83
20	74.5	23	8.88
21	72.1	20	7.68
22	58.1	21	8.47
23	44.6	20	8.86
24	33.4	20	10.36
25	28.6	22	11.08

Рассматривается линейная модель первого порядка вида:

$$y = \beta_0 * x_0 + \beta_8 * x_8 + \beta_6 * x_6 + \varepsilon,$$

где x_0 – фиктивная переменная, которая всегда равна единице. Нужно применить МНК для оценивания параметров β_0, β_6, β_8 модели, провести статистический анализ результатов, построить графики остатков.

186. (Дрейпер, Смит: гл. 5, зад. 9). Точка (или температура) помутнения жидкости служит мерой кристаллизации исходного материала и её можно измерить с помощью коэффициента преломления (показателя рефракции). Предполагается, что процент фракции в исходном сырье прекрасно предсказывает точку помутнения с помощью модели второго порядка:

$$y = \beta_0 + \beta_1 * x + \beta_{11} * x^2 + \varepsilon .$$

Были собраны следующие данные:

% фракции, x	Точка помутнения, y	x	Y	x	y	X	y
0	22.1	2	26.1	5	28.9	0	22.8
1	24.5	4	28.5	6	30.0	3	27.3
2	26.0	6	30.3	7	30.4	6	29.8
3	26.8	8	31.5	8	31.4	9	31.8
4	28.2	10	33.1	0	21.9		

1) Определите наилучшую модель второго порядка.
2) Приняв $\alpha = 0.05$, проверьте полную регрессию.
3) Проверьте адекватность.
4) Будет ли достаточной модель первого порядка $y = \beta_0 + \beta_1 * x + \varepsilon$? Используйте остатки от этой простой модели как основу для ваших заключений.
5) Объясните применение подобранной модели второго порядка как предсказывающего уравнения.

187. Для двух случайных выборок из нормальных совокупностей А и В, объёма 9 и 7, соответственно, получены следующие выборочные средние и выборочные дисперсии: ($\bar{x}$ = 12.6, s^2 = 10.02) для А, ($\bar{x}$ = 8.9, s^2 = 16.08) для В. Построить 95%-ный доверительный интервал для разности математических ожиданий $\mu_A - \mu_B$.

188. Определить, какого объёма нужно взять выборку из нормальной совокупности, если известно, что σ = 10, и требуется, чтобы $\bar{x}$ находилось не далее чем на расстоянии 1.2 единиц от μ. Предполагается, что уровень доверия равен 0.95 (следовательно, уровень значимости α =0.05).

189. При шести измерениях концентрации в процентах получены следующие результаты: 1.20; 1.27; 1.33; 1.19; 1.09 и 1.24. Требуется оценить ожидаемое значение концентрации в процентах, предполагая, что результаты измерения являются случайными независимыми величинами и имеют нормальное распределение с дисперсией, равной 0.003. Постройте 95%-ный доверительный интервал для величины $E\{\xi\} = \mu$.

190. В условиях предыдущей задачи, считая дисперсию неизвестной величиной, постройте 99%-ный доверительный интервал для μ .

191. Найдите 90%-ный доверительный интервал для отношения дисперсий предела прочности на разрыв по двум приведённым ниже выборкам (предположительно однородным), полученным при испытаниях на прочность на двух разных испытательных стендах.

Стенд А		Стенд В	
1.324	1.309	1.358	1.398
1.322	1.374	1.318	1.375
1.327	1.319	1.318	1.341
1.342	1.392	1.321	1.366
1.301	1.396	1.414	1.408

192. При производстве синтетического волокна для уменьшения последующей усадки продукция, движущаяся непрерывным потоком, подвергается термической обработке. Ниже приведена величина усадки в процентах для волокон после обработки при температуре 120 градусов: 3.45; 3.64; 3.60; 3.49; 3.64; 3.56; 3.52; 3.53; 3.57; 3.44; 3.56; 3.43. Определить 98%-ный доверительный интервал для μ .

193. Чтобы выяснить, варьирует ли от одного дня к другому величина изменчивости температуры высокоскоростного аппарата, в первый день было произведено 12 измерений, а во второй 10. Среднее квадратическое отклонение в первый день составило 23° F, а во второй 30° F. Построить 95%-ный доверительный интервал для отношения дисперсий. Можно ли отсюда заключить, что дисперсии одинаковы ?

194. При следующих температурах происходит сгорание десяти образцов керамических покрытий: 1430, 1520, 1460, 1470, 1480, 1340, 1460, 1520, 1450 и 1500. Определить, содержится среди этих данных выброс (уровень значимости $\alpha = 0.05$).

195. Проверка скорости полимеризации проводится на нескольких образцах полимеров. Предполагаемая скорость полимеризации составляет 24% в час. В восьми опытах получены следующие результаты: 23.6; 22.8; 25.7; 24.8; 26.4; 24.3; 23.9 и 25.0 % в час. Построить доверительный интервал для μ и проверить справедливость утверждения о предполагаемой скорости полимеризации.

196. Выборка А обладает дисперсией 55.4 фунт ^{2} , а выборка В – дисперсией 87.3 фунт ^{2}

а) Можно ли сказать, что выборки взяты из совокупностей с одним и тем же значением дисперсии, если их объёмы равны соответственно 15 и 12 ?

б) какие при этом необходимо сделать предположения ?

197. Температура в автоклаве регистрируется через равные промежутки времени. Для проведения некоторого эксперимента потребовалось поддерживать заданную температуру. Температура регистрировалась в течение двух последовательных дней в случайные моменты времени. В первый день было зафиксировано 16 значений температуры со средним квадратическим отклонением 15.6, а во второй день – 21 значение со средним квадратическим отклонением 9.8. Можно ли утверждать, что наблюдения относятся к одной совокупности ?

198. Проверить, есть ли выбросы в данных задачи 152 темы 2.