

Подборка экзаменов по теории
вероятностей.
Факультет экономики, НИУ-ВШЭ

Коллектив кафедры
математической экономики и эконометрики,
фольклор

3 сентября 2012 г.

Содержание

1	Описание	2
2	2005-2006	3
2.1	Контрольная работа № 1, 18.10.2005	3
2.2	Контрольная работа № 2, 21.12.2005	6
2.3	Контрольная работа № 3, 04.03.2006	8
3	2006-2007	16
3.1	Контрольная работа № 1, ??..11.2006	16
3.2	Контрольная работа № 2, 27.01.2007	19
3.3	Контрольная работа № 3, 21.02.2007	26
4	2007-2008	33
4.1	Контрольная работа № 1, 03.11.2007	33
4.2	Контрольная работа № 2, демо-версия, 21.01.2008	37
4.3	Контрольная работа № 2, 21.01.2008	40
4.4	Контрольная работа № 3, демо-версия, 01.03.2008	48
4.5	Контрольная работа № 3, 01.03.2008	50
5	2008-2009	55
5.1	Контрольная работа № 1, демо-версия, ??..11.2008	55
5.2	Контрольная работа № 1, ??..11.2008	59

5.3	Контрольная работа №2, демо-версия, 26.12.2008	63
5.4	Контрольная работа №2, 26.12.2008	67
5.5	Контрольная работа №3, 02.03.2009	74
6	2009-2010	81
6.1	Контрольная работа №2, ??..12.2009	81
6.2	Контрольная работа №4, ??..??..2010	84
7	2010-2011	86
7.1	Контрольная работа №1, ??..10.2010	86
7.2	Контрольная работа №2, ??..12.2010	88
7.3	Контрольная работа №3, ??..03.2011	90
8	2011-2012	91
8.1	Контрольная работа №1, 24.10.2011	91
8.2	Контрольная работа №2, 29.12.2011	97
8.3	Контрольная работа №3, 13.03.2012	105
8.4	Экзамен, 26.03.2012	106

1 Описание

Скачать этот документ можно с блога <http://pokrovka11.wordpress.com/>:



Прямая ссылка на скачивание, http://dl.dropbox.com/u/6806513/teaching/probability/2011/2011_prob_exam_collection.pdf:



Уникальное предложение для студентов факультета экономики ГУ-ВШЭ: Найдите ошибки в этом документе и получите дополнительные бонусы к итоговой оценке! Смысловые ошибки поощряются сильнее, чем просто опечатки. Письма с замеченными ошибками пишите на адрес boris.demeshev@gmail.com.

2 2005-2006

2.1 Контрольная работа № 1, 18.10.2005

1. Если X - случайная величина, то $Var(X) = Var(16 - X)$?
2. Функция распределения случайной величины является неубывающей ?
3. Дисперсия случайной величины не меньше, чем ее стандартное отклонение ?
4. Для любой случайной величины $E(X^2) \geq (E(X))^2$?
5. Если ковариация равна нулю, то случайные величины независимы ?
6. Значение функции плотности может превышать единицу ?
7. Если события A и B не могут произойти одновременно, то они независимы ?
8. Для любых событий A и B верно, что $P(A|B) \geq P(A \cap B)$?
9. Функция плотности не может быть периодической ?
10. Для неотрицательной случайной величины $E(X) \geq E(-X)$?
11. Я еще не видел части с задачами, но что-то мне уже домой хочется ?

Часть II Стоимость задач 10 баллов.

Задача №1

Шесть студентов, три юноши и три девушки, стоят в очереди за пирожками в случайном порядке. Какова вероятность того, что юноши и девушки чередуются?

Задача №2

Имеется три монетки. Две "правильных" и одна - с "орлами" по обеим сторонам. Петя выбирает одну монетку наугад и подкидывает ее два раза. Оба раза выпадает "орел". Какова вероятность того, что монетка "неправильная"?

Задача №3

Вася гоняет на мотоцикле по единичной окружности с центром в начале координат. В случайный момент времени он останавливается. Пусть случайные величины X и Y - это Васиной абсцисса и ордината в момент остановки. Найдите $P(X > \frac{1}{2})$, $P(X > \frac{1}{2} | Y < \frac{1}{2})$. Являются ли события $A = \{X > \frac{1}{2}\}$ и $B = \{Y < \frac{1}{2}\}$ независимыми?

Подсказка: $\cos(\frac{\pi}{3}) = \frac{1}{2}$, длина окружности $l = 2\pi R$

Задача №4

В коробке находится четыре внешне одинаковых лампочки. Две из лампочек исправны, две - нет. Лампочки извлекают из коробки по одной до тех пор, пока не будут извлечены обе исправные.

а) Какова вероятность того, что опыт закончится извлечением трех лампочек?

б) Каково ожидаемое количество извлеченных лампочек?

Задача №5

Два охотника выстрелили в одну утку. Первый попадает с вероятностью 0,4, второй - с вероятностью 0,6. В утку попала ровно одна пуля. Какова вероятность того, что утка была убита первым охотником?

Задача №6

а) Известно, что $E(Z) = -3$ и $E(Z^2) = 15$. Найдите $Var(Z)$, $Var(4 - 3Z)$ и $E(5 + 3Z - Z^2)$.

б) Известно, что $Var(X + Y) = 20$ и $Var(X - Y) = 10$. Найдите $Cov(X, Y)$ и $Cov(6 - X, 3Y)$.

Задача №7

Известно, что случайная величина X принимает три значения. Также известно, что $P(X = 1) = 0,3$; $P(X = 2) = 0,1$ и $E(X) = -0,7$. Определите чему равно третье значение случайной величины X и найдите $Var(X)$.

Задача №8

Известно, что функция плотности случайной величины X имеет вид:

$$p(x) = \begin{cases} cx^2, & x \in [-2; 2] \\ 0, & x \notin [-2; 2] \end{cases}$$

Найдите значение константы A , $P(X > 1)$, $E(X)$, $E\left(\frac{1}{X^3+10}\right)$ и постройте график функции распределения величины X .

Задача №9

Бросают два правильных игральных кубика. Пусть X - наименьшая из выпавших граней, а Y - наибольшая.

- а) Рассчитайте $P(X = 3 \cap Y = 5)$;
- б) Найдите $E(X)$, $Var(X)$, $E(3X - 2Y)$;

Задача №10

Вася решает тест путем проставления каждого ответа наугад. В тесте 5 вопросов. В каждом вопросе 4 варианта ответа. Пусть X - число правильных ответов, Y - число неправильных ответов и $Z = X - Y$.

- а) Найдите $P(X > 3)$
- б) Найдите $Var(X)$ и $Cov(X, Y)$
- в) Найдите $Corr(X, Z)$

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 11-х задач по выбору!

Задача №11-А

Петя сообщает Васе значение случайной величины, равномерно распределенной на отрезке $[0; 4]$. С вероятностью $\frac{1}{4}$ Вася возводит Петино число в квадрат, а с вероятностью $\frac{3}{4}$ прибавляет к Петиному числу 4. Обозначим результат буквой Y .

Найдите $P(Y < 4)$ и функцию плотности случайной величины Y .

Вася выбирает свое действие независимо от Петиного числа.

Требуется решить одну из двух 11-х задач по выбору!

Задача №11-В

Вы хотите приобрести некую фирму. Стоимость фирмы для ее нынешних владельцев - случайная величина, равномерно распределенная на отрезке $[0; 1]$. Вы предлагаете владельцам продать ее за называемую Вами сумму. Владельцы либо соглашаются, либо нет. Если владельцы согласны, то Вы платите обещанную сумму и получаете фирму. Когда фирма переходит в Ваши руки, ее стоимость сразу возрастает на 20%.

- а) Чему равен Ваш ожидаемый выигрыш, если Вы предлагаете цену 0,5?
- б) Какова оптимальная предлагаемая цена?

2.2 Контрольная работа № 2, 21.12.2005

Верный ответ = +1 балл, Неверный ответ = 0 баллов, Отсутствие ответа = + 0,5 балла.

1. Сумма двух нормальных независимых случайных величин нормальна ?
2. Сумма любых двух непрерывных случайных величин непрерывна ?
3. Нормальная случайная величина не может принимать отрицательные значения ?
4. Пуассоновская случайная величина является непрерывной ?
5. Сумма двух независимых равномерно распределенных величин равномерна ?
6. Дисперсия суммы зависимых величин всегда больше суммы дисперсий ?
7. Дисперсия пуассоновской с.в. равна ее математическому ожиданию ?
8. Если X - непрерывная с.в., $E(X) = 6$ и $Var(X) = 9$, то $Y = \frac{X-6}{3} \sim N(0; 1)$. ?
9. Теорема Муавра-Лапласа является частным случаем центральной предельной . ?
10. Для любой случайной величины $E(X|X > 0) \geq E(X)$?

Часть II Стоимость задач 10 баллов.

Задача №1

Вася, владелец крупного Интернет-портала, вывесил на главной странице рекламный баннер. Ежедневно его страницу посещают 1000 человек. Вероятность того, что посетитель портала кликнет по баннеру равна 0,003. С помощью пуассоновского приближения оцените вероятность того, что за один день не будет ни одного клика по баннеру.

Задача №2

Совместный закон распределения случайных величин X и Y задан таблицей:

	$Y = -1$	$Y = 0$	$Y = 2$
$X = 0$	0,2	c	0,2
$X = 1$	0,1	0,1	0,1

Найдите c , $P(Y > -X)$, $E(X \cdot Y^2)$, $E(Y|X > 0)$

Задача №3

Случайный вектор $(X_1 \ X_2)$ имеет нормальное распределение с математическим ожиданием $(2 \ -1)$ и ковариационной матрицей $\begin{pmatrix} 9 & -4,5 \\ -4,5 & 25 \end{pmatrix}$.
Найдите $P(X_1 + 3X_2 > 20)$.

Задача №4

Совместная функция плотности имеет вид

$$p_{X,Y}(x,y) = \begin{cases} x+y, & \text{if } x \in [0;1], y \in [0;1] \\ 0, & \text{otherwise} \end{cases}$$

Найдите $P(Y > X)$, $E(X)$, $E(X|Y > X)$

Задача №5

В среднем 20% покупателей супермаркета делают покупку на сумму свыше 500 рублей. Какова вероятность того, что из 200 покупателей менее 21% сделают покупку на сумму менее 500 рублей?

Задача №6

Вася и Петя метают дротики по мишени. Каждый из них сделал по 100 попыток. Вася оказался метче Пети в 59 попытках. На уровне значимости 5% проверьте гипотезу о том, что меткость Васи и Пети одинаковая, против альтернативной гипотезы о том, что Вася метче Пети.

Задача №7

Найдите $P(X \in [16; 23])$, если

- а) X нормально распределена, $E(X) = 20$, $Var(X) = 25$.
- б) X равномерно распределена на отрезке $[0; 30]$
- в) X распределена экспоненциально и $E(X) = 20$

Задача №8

Каждый день цена акции равновероятно поднимается или опускается на один рубль. Сейчас акция стоит 1000 рублей. Введем случайную величину X_i , обозначающую изменение курса акции за i -ый день. Найдите $E(X_i)$ и $Var(X_i)$. С помощью центральной предельной теоремы найдите вероятность того, что через сто дней акция будет стоить больше 1030 рублей.

Задача №9

Определите математическое ожидание и дисперсию случайной величины, если ее функция плотности имеет вид $p(t) = c \cdot \exp(-2 \cdot (t+1)^2)$.

Задача №10

Пусть случайные величины X и Y независимы и распределены по Пуассону с параметрами $\lambda_X = 5$ и $\lambda_Y = 15$ соответственно. Найдите условное распределение случайной величины X , если известно, что $X + Y = 50$.

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 11-х задач по выбору!

Задача №11-А

Допустим, что оценка X за экзамен распределена равномерно на отрезке $[0; 100]$. Итоговая оценка Y рассчитывается по формуле $Y =$

$$\begin{cases} 0, & \text{if } X < 30 \\ X, & \text{if } X \in [30; 80] \\ 100, & \text{if } X > 80 \end{cases}.$$

Найдите $E(Y)$, $E(X \cdot Y)$, $E(Y^2)$, $E(Y|Y > 0)$.

Требуется решить одну из двух 11-х задач по выбору!

Задача №11-В

Вася играет в компьютерную игру - "стрелялку-бродилку". По сюжету ему нужно убить 60 монстров. На один выстрел уходит ровно 1 минута. Вероятность убить монстра с одного выстрела равна 0,25. Количество выстрелов не ограничено. Сколько времени в среднем Вася тратит на одного монстра? Найдите дисперсию этого времени? Какова вероятность того, что Вася закончит игру меньше, чем за 3 часа?

2.3 Контрольная работа № 3, 04.03.2006

Solution!

Просто из сил выбьешься, пока вдруг как-то само не уладится; что-то надо подчеркнуть, что-то - выбросить, не договорить, а где-то - ошибиться, без ошибки такая пакость, что глядеть тошно.

В.А. Серов

Часть I . Обведите верный ответ:

1. Если $X \sim \chi_n^2$ и $Y \sim \chi_{n+1}^2$, X и Y - независимы, то X не превосходит Y . Нет.
2. В тесте Манна-Уитни предполагается нормальность хотя бы одной из сравниваемых выборок. Нет.
3. График функции плотности случайной величины, имеющей t -распределение симметричен относительно 0. Да.
4. Мощность больше у того теста, у которого вероятность ошибки 2-го рода меньше. Да.
5. Если $X \sim t_n$, то $X^2 \sim F_{1,n}$. Да.
6. При прочих равных 90% доверительный интервал шире 95%-го. Нет.
7. Несмещенная выборочная оценка дисперсии не превосходит квадрата выборочного среднего. Нет.
8. Если гипотеза отвергается при 5%-ом уровне значимости, то она будет отвергаться и при 1%-ом уровне значимости. Нет.
9. У t -распределения более толстые 'хвосты', чем у стандартного нормального. Да.
10. Р-значение показывает вероятность отвергнуть нулевую гипотезу, когда она верна. Нет.
11. Если t -статистика равна нулю, то Р-значение также равно нулю. Нет.
12. Если $X \sim N(0; 1)$, то $X^2 \sim \chi_1^2$. Да.
13. Пусть X_i - длина i -го удава в сантиметрах, а Y_i - в дециметрах. Выборочный коэффициент корреляции между этими наборами данных равен $\frac{1}{10}$. Нет.
14. Математическое ожидание выборочного среднего не зависит от объема выборки, если X_i одинаково распределены. Да.
15. Зная закон распределения X и закон распределения Y можно восстановить совместный закон распределения пары (X, Y) . Нет.

16. Если ты отвечаешь на вопросы этого теста наугад, то число правильных ответов - случайная величина, имеющая биномиальное распределение с дисперсией 4. Да.
[правильно=1/нет ответа=0/неправильно=-1]

Часть II Стоимость задач 10 баллов.

Задача 1

Пусть случайная величина X распределена равномерно на отрезке $[0; a]$, где $a > 3$. Исследователь хочет оценить параметр $\theta = P(X < 3)$. Рассмотрим следующую оценку $\hat{\theta} = \begin{cases} 1, & X < 3 \\ 0, & X \geq 3 \end{cases}$.

а) [3] Объясните, что означают термины 'несмещенность', 'состоятельность', 'эффективность'.

смотрим учебник

б) [3] Верно ли, что оценка $\hat{\theta}$ является несмещенной?

$E(\hat{\theta}) = 1 \cdot P(X < 3) + 0 \cdot P(X \geq 3) = \theta$, да является

в) [4] Найдите $E\left((\hat{\theta} - \theta)^2\right)$.

$$E\left((\hat{\theta} - \theta)^2\right) = E\left(\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2\right) =$$

Заметим, что $\hat{\theta}^2 = \hat{\theta}$

$$= \theta - 2\theta^2 + \theta^2 = \theta - \theta^2$$

Задача 2

Пусть X_1, X_2, \dots, X_n независимы и их функции плотности имеет вид:

$$f(x) = \begin{cases} (k+1)x^k, & x \in [0; 1]; \\ 0, & x \notin [0; 1]. \end{cases}$$

Найдите оценки параметра k :

а) [5] Методом максимального правдоподобия

$$L = (k+1)^n (x_1 \cdot x_2 \cdot \dots \cdot x_n)^k$$

$$l = \ln L = n \ln(k+1) + k(\sum \ln x_i)$$

$$\frac{dl}{dk} = \frac{n}{k+1} + \sum \ln x_i$$

$$\frac{n}{k+1} + \sum \ln x_i = 0$$

$$\hat{k} = -\left(1 + \frac{n}{\sum \ln x_i}\right)$$

б) [5] Методом моментов

$$E(X_i) = \int_0^1 t \cdot p(t) dt = \int_0^1 (k+1)t^{k+1} dt = \frac{k+1}{k+2}$$

$$\frac{\hat{k}+1}{\hat{k}+2} = \bar{X}$$

$$\hat{k} = \frac{2\bar{X}-1}{1-\bar{X}}$$

Задача 3

У 200 человек записали цвет глаз и волос. На уровне значимости 10% проверьте гипотезу о независимости этих признаков.

Цвет глаз/волос	Светлые	Темные	Итого
Зеленые	49	25	74
Другие	30	96	126
Итого	79	121	200

$$C = \sum \frac{(X_{i,j} - n\hat{p}_{i,j})^2}{n\hat{p}_{i,j}} \sim \chi_{(r-1)(c-1)}^2$$

$$C \sim \chi_1^2$$

$$C = 35$$

Если $\alpha = 0, 1$, то $C_{crit} = 2, 706$.

Вывод: H_0 (гипотеза о независимости признаков) отвергается.

[2 балла за формулировку H_0 и H_a]

[-1 балл за неверные степени свободы]

[-2 за неумение пользоваться таблицей]

Задача 4

На курсе два потока, на первом потоке учатся 40 человек, на втором потоке 50 человек. Средний балл за контрольную на первом потоке равен 78 при (выборочном) стандартном отклонении в 7 баллов. На втором потоке средний балл равен 74 при (выборочном) стандартном отклонении в 8 баллов.

а) [6] Постройте 90% доверительный интервал для разницы баллов между двумя потоками

Число наблюдений велико, используем нормальное распределение.

$$P\left(-1,65 < \frac{\bar{X} - \bar{Y} - \Delta}{\sqrt{\frac{\sigma_x^2}{40} + \frac{\sigma_y^2}{50}}} < 1,65\right) = 0,9$$

$$\Delta \in 4 \pm 1,65 \sqrt{\frac{49}{40} + \frac{64}{50}}$$

$$\Delta \in [1, 4; 6, 6]$$

б) [2] На 10%-ом уровне значимости проверьте гипотезу о том, что результаты контрольной между потоками не отличаются.

Используем результат предыдущего пункта: H_0 отвергается, т.к. число 0 не входит в доверительный интервал.

в) [2] Рассчитайте точное P-значение (P-value) теста в пункте 'б'

$$Z = 2,505 \text{ и } P_{value} = 0,0114$$

Задача 5

Предположим, что время жизни лампочки распределено нормально. По 10 лампочкам оценка стандартного отклонения времени жизни оказалась равной 120 часам.

а) [5] Найдите 80%-ый (двусторонний) доверительный интервал для истинного стандартного отклонения.

$$\chi_9^2 = \frac{9\hat{\sigma}^2}{\sigma^2} \in [4, 17; 14, 69]$$

$$\sigma^2 \in [8822, 3; 31080]$$

$$\sigma \in [93, 9; 176, 3]$$

б) [5] Допустим, что выборку увеличат до 20 лампочек. Какова вероятность того, что выборочная оценка дисперсии будет отличаться от истинной дисперсии меньше, чем на 40%?

$$P(|\hat{\sigma}^2 - \sigma^2| < 0,4\sigma^2) = P(0,6 < \frac{\hat{\sigma}^2}{\sigma^2} < 1,4) = P(11,4 < \chi_{19}^2 < 26,6) \approx 0,8$$

Задача 6

Из 10 опрошенных студентов часть предпочитала готовиться по синему учебнику, а часть - по зеленому. В таблице представлены их итоговые баллы.

Синий	76	45	57	65		
Зеленый	49	59	66	81	38	88

а) [8] С помощью теста Манна-Уитни (Mann-Whitney) проверьте гипотезу о том, что выбор учебника не меняет закона распределения оценки.

Разрешается использование нормальной аппроксимации

$$W_1 = 2 + 4 + 6 + 8 = 20 \text{ или } W_2 = 1 + 3 + 5 + 7 + 9 + 10 = 35$$

[3 из 8 за правильный расчет суммы рангов]

$$U_1 = 10 \text{ или } U_2 = 14$$

$$Z_1 = -0,43 = -Z_2$$

Вывод: H_0 (гипотеза об отсутствии сдвига между законами распределения) не отвергается

б) [2] Возможно ли в этой задаче использовать (Wilcoxon Signed Rank Test)?

Нет, т.к. наблюдения не являются парными.

Задача 7

Вася очень любит играть в преферанс. Предположим, что Васин выигрыш распределен нормально. За последние 5 партий средний выигрыш составил 1560 рублей, при оценке стандартного отклонения равной 670 рублям. Постройте 90%-ый доверительный интервал для математического ожидания Васиного выигрыша.

$$P(-2,13 < t_4 < 2,13) = 0,9$$

$$\mu \in 1560 \pm 2,13 \cdot \sqrt{\frac{670^2}{5}}$$

$$\mu \in [921,8; 2198,2]$$

[-3 балла за использование N вместо t]

[-1 балл за неверные степени свободы]

[-2 за неумение пользоваться таблицей]

Задача 8

Имеется две конкурирующие гипотезы:

H_0 : Величина X распределена равномерно на отрезке $[0; 100]$

H_a : Величина X распределена равномерно на отрезке $[50; 150]$

Исследователь выбрал такой критерий:

Если $X < c$, то использовать H_0 , иначе использовать H_a .

а) [3, по 1 баллу за определение] Что такое 'ошибка первого рода', 'ошибка второго рода', 'мощность теста'?

читаем книжки с картинками

б) Постройте графики зависимостей ошибок первого и второго рода от c .

$$P(1 \text{ type error}) = P(X > c | X \sim U[0; 100]) = \begin{cases} 1, & c < 0 \\ 1 - \frac{c}{100}, & c \in [0; 100] \\ 0, & c > 100 \end{cases}$$

$$P(2 \text{ type error}) = P(X < c | X \sim U[50; 150]) = \begin{cases} 0, & c < 50 \\ \frac{c-50}{100}, & c \in [50; 150] \\ 1, & c > 150 \end{cases}$$

Построение оставлено читателю в качестве самостоятельного упражнения :)

Задача 9

На плоскости выбирается точка со случайными координатами. Абсцисса и ордината независимы и распределены $N(0; 1)$. Какова вероятность того, что расстояние от точки до начала координат будет больше 2,45?

$$P(\sqrt{X^2 + Y^2} > 2,45) = P(X^2 + Y^2 > 2,45^2) = P(\chi_2^2 > 6) = 0,05$$

[-1 балл за неверные степени свободы]

[-2 за неумение пользоваться таблицей]

Задача 10

С вероятностью 0,3 Вася оставил конспект в одной из 10 посещенных им сегодня аудиторий. Вася осмотрел 7 из 10 аудиторий и конспекта в них не нашел.

а) [5] Какова вероятность того, что конспект будет найден в следующей осматриваемой им аудитории?

A = конспект забыт в 8-ой аудитории

B = конспект был забыт в другом месте (не в аудиториях)

C = конспект не был найден в первых 7-и

$$P(A|C) = \frac{P(A)}{P(C)} = \frac{0,3 \cdot 0,1}{0,3 \cdot 0,3 + 0,7} = \frac{3}{79}$$

б) [5] Какова (условная) вероятность того, что конспект оставлен где-то в другом месте?

$$P(B|C) = \frac{P(B)}{P(C)} = \frac{0,7}{0,79} = \frac{70}{79}$$

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 11-х задач по выбору!

Задача 11-А [Hardy-Weinberg theorem]

У диплоидных организмов наследственные характеристики определяются парой генов. Вспомним знакомые нам с 9-го класса горошины чешского монаха Менделя. Ген, определяющий форму горошины, имеет две аллели: 'А' (гладкая) и 'а' (морщинистая). 'А' доминирует 'а'. В популяции бесконечное количество организмов. Родители каждого потомка определяются случайным образом, согласно имеющемуся распределению генотипов. Одна аллель потомка выбирается наугад из аллелей матери, другая - из аллелей отца. Начальное распределение генотипов имеет вид: 'АА' - 30%, 'Аа' - 60%, 'аа' - 10%.

- а) [10] Каким будет распределение генотипов в n -ом поколении?
- б) [10] Заметив закономерность, сформулируйте и докажите теорему Харди-Вайнберга для произвольного начального распределения генотипов.

О чем молчал учебник биологии 9 класса...

Если:

- а) ген имеет всего две аллели;
- б) в популяции бесконечное число организмов;
- в) одна аллель потомка выбирается наугад из аллелей матери, другая - из аллелей отца;

То:

Распределение генотипов стабилизируется уже в первом поколении (!!!).

Т.е. $AA_1 = AA_2 = \dots$ и $Aa_1 = Aa_2 = \dots$

Вероятность получить 'А' от родителя для рождающихся в поколении 1 равна: $p_1 = 0,3 \cdot 1 + 0,6 \cdot 0,5 + 0,1 \cdot 0 = 0,6$

В общем виде: $p_1 = AA_0 + 0,5 \cdot Aa_0$

$AA_1 = p_1^2 = 0,36$, $Aa_1 = 2p_1(1 - p_1) = 0,48$.

$p_2 = AA_1 + 0,5 \cdot Aa_1 = p_1^2 + p_1(1 - p_1) = p_1$

Требуется решить одну из двух 11-х задач по выбору!

Задача 11-В

В киосках продается 'открытка-подарок'. На открытке есть прямоугольник размером 2 на 7. В каждом столбце в случайном порядке находятся очередная буква слова 'подарок' и звездочка. Например, вот так:

П	*	*	А	*	О	К
*	О	Д	*	Р	*	*

Прямоугольник закрыт защитным слоем, и покупатель не видит, где буква, а где - звездочка. Следует стереть защитный слой в одном квадратике в каждом столбце. Можно попытаться угадать любое число букв. Если открыто n букв слова 'подарок' и не открыто ни одной звездочки, то открытку можно обменять на $50 \cdot 2^{n-1}$ рублей. Если открыта хотя бы одна звездочка, то открытка остается просто открыткой.

а) [15] Какой стратегии следует придерживаться покупателю, чтобы максимизировать ожидаемый выигрыш?

Безразлично.

Если я решил попробовать угадать n букв, то выигрыш вырастает, а вероятность падает в 2 раза по сравнению с попыткой угадать $(n - 1)$ -ую букву.

б) [5] Чему равен максимальный ожидаемый выигрыш?

В силу предыдущего пункта: $E(X) = \frac{1}{2} \cdot 50 = 25$

Подсказка: Думайте!

3 2006-2007

3.1 Контрольная работа № 1, ??..11.2006

Вывешенное решение может содержать неумышленные опечатки.

Заметил опечатку? Сообщите преподавателю!

1. Из семей, имеющих троих разновозрастных детей, случайным образом выбирается одна семья. Пусть событие A заключается в том, что в этой семье старший ребенок - мальчик, B - в семье есть хотя бы одна девочка.

1.1 Считая вероятности рождения мальчиков и девочек одинаковыми, выяснить, являются ли события A и B независимыми.

1.2 Изменится ли результат, если вероятности рождения мальчиков и девочек различны.

1.1. $P(A) = 0,5$, $P(B) = 1 - P(B^c) = 1 - 0,5^3 = \frac{7}{8}$, $P(A \cap B) = 0,5 \cdot (1 - 0,5^2) = \frac{3}{8}$, $P(A \cap B) \neq P(A)P(B)$, события зависимы.

1.2. $P(A) = p$, $P(B) = 1 - p^3$, $P(A \cap B) = p(1 - p^2)$, независимость событий возможна только при $p = 0$ или $p = 1$

2. Студент решает тест (множественного выбора) проставлением ответов

наугад. В тесте 10 вопросов, на каждый из которых 4 варианта ответов. Зачет ставится в том случае, если правильных ответов будет не менее 5.

2.1 Найти вероятность того, что студент правильно ответит только на один вопрос

2.2 Найти наиболее вероятное число правильных ответов

2.3 Найти математическое ожидание и дисперсию числа правильных ответов

2.4 Найти вероятность того, что студент получит зачет

Пусть X - число правильных ответов.

$$2.1. P(X = 1) = C_{10}^1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^9$$

$$2.2. k_{P(X=k) \rightarrow \max} = \lfloor p(n+1) \rfloor = \lfloor \frac{11}{4} \rfloor = 2 \text{ (можно не зная формулы просто выбрать наибольшую вероятность)}$$

$$2.3. E(X) = 10E(X_i) = \frac{10}{4}$$

$$Var(X) = 10Var(X_i) = 10 \frac{1}{4} \frac{3}{4}$$

$$2.4. \sum_{i=5}^{10} C_{10}^i \left(\frac{1}{4}\right)^i \left(\frac{3}{4}\right)^{10-i}$$

3. Вероятность изготовления изделия с браком на некотором предприятии равна 0.04. Перед выпуском изделие подвергается упрощенной проверке, которая в случае бездефектного изделия пропускает его с вероятностью 0.96, а в случае изделия с дефектом - с вероятностью 0.05. Определить:

3.1 Какая часть изготовленных изделий выходит с предприятия

3.2 Какова вероятность того, что изделие, прошедшее упрощенную проверку, бракованное

A - изделие браковано, B - изделие признано хорошим

$$3.1. P(B) = 0,96 \cdot 0,96 + 0,04 \cdot 0,05$$

$$3.2. P(A|B) = \frac{0,04 \cdot 0,05}{P(B)}$$

4. Вероятность того, что пассажир, купивший билет, не придет к отправлению поезда, равна 0.01. Найти вероятность того, что все 400 пассажиров явятся к отправлению поезда (использовать приближение Пуассона).

$$\lambda = np = 4$$

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

$$P(X = 0) = e^{-4}$$

5. Охотник, имеющий 4 патрона, стреляет по дичи до первого попадания или до израсходования всех патронов. Вероятность попадания при первом выстреле равна 0.6, при каждом последующем - уменьшается на 0.1. Найти

5.1 Закон распределения числа патронов, израсходованных охотником

5.2 Математическое ожидание и дисперсию этой случайной величины

5.1.

x_i	1	2			3	4
$P(X = x_i)$	0,6	$(1 - 0,6) \cdot 0,5$			$(1 - 0,6) \cdot (1 - 0,5) \cdot 0,4$	$1 - p_1 - p_2 - p_3$
x_i	1	2	3	4		
$P(X = x_i)$	0,6	0,2	0,08	0,12		

$$E(X) = 1,7, \text{Var}(X) \approx 1,08$$

6. Поезда метрополитена идут регулярно с интервалом 2 минуты. Пассажир приходит на платформу в случайный момент времени. Какова вероятность того, что ждать пассажиру придется не более полминуты. Найти математическое ожидание и дисперсию времени ожидания поезда.

$$P(X \leq 0,5) = \frac{0,5}{2} = 0,25, E(X) = \frac{0+2}{2} = 1 \text{ (здравый смысл)}$$

$$\text{Var}(X) = E(X^2) - (E(X))^2$$

$$E(X^2) = \int_0^2 t^2 \cdot p(t) dt = \int_0^2 t^2 \cdot 0,5 dt = \frac{4}{3}$$

7. Время работы телевизора "Best" до первой поломки является случайной величиной, распределенной по показательному закону. Определить вероятность того, что телевизор проработает более 15 лет, если среднее время безотказной работы телевизора фирмы "Best" составляет 10 лет. Какова вероятность, что телевизор, проработавший 10 лет, проработает еще не менее 15 лет?

$$E(X) = 10 = \frac{1}{\lambda}, \lambda = \frac{1}{10}, p(t) = \lambda e^{-\lambda t} \text{ при } t > 0$$

$$P(X > 15) = \int_{15}^{\infty} p(t) dt = \dots = e^{-\frac{3}{2}}$$

$$P(X > 25 | X > 10) = \frac{P(X > 25)}{P(X > 10)} = \dots = e^{-\frac{3}{2}}$$

Дополнительная задача:

Пусть случайные величины X_1 и X_2 независимы и равномерно распределены на отрезках $[-1; 1]$ и $[0; 1]$, соответственно. Найти вероятность того, что $\max\{X_1, X_2\} > 0,5$, функцию распределения случайной величины $Y = \max\{X_1, X_2\}$.

Функция распределения $F_Y(t) = P(Y \leq t) = P(\max\{X_1, X_2\} \leq t) = P(X_1 \leq t \cap X_2 \leq t) = P(X_1 \leq t)P(X_2 \leq t) = \frac{t+1}{2}t$ при $t \in [0; 1]$. При $t > 1$ получаем, что $F_Y(t) = 1$ и при $t < 0$ получаем, что $F_Y(t) = 0$.

$$P(\max\{X_1, X_2\} > 0,5) = 1 - P(\max\{X_1, X_2\} \leq 0,5) = 1 - F(0,5) = \frac{5}{8}$$

3.2 Контрольная работа № 2, 27.01.2007

Часть I. Обведите верный ответ:

1. Сумма двух нормальных независимых случайных величин нормальна. Да.
2. Нормальная случайная величина может принимать отрицательные значения. Да.
3. Пуассоновская случайная величина является непрерывной. Нет.
4. Дисперсия суммы зависимых величин всегда не меньше суммы дисперсий. Нет.
5. Теорема Муавра-Лапласа является частным случаем центральной предельной. Да.
6. Пусть X - длина наугад выловленного удава в сантиметрах, а Y - в дециметрах. Коэффициент корреляции между этими величинами равен $\frac{1}{10}$. Нет.
7. Математическое ожидание выборочного среднего не зависит от объема выборки, если X_i одинаково распределены. Да.
8. Зная закон распределения X и закон распределения Y можно восстановить совместный закон распределения пары (X, Y) . Нет.
9. Если X - непрерывная с.в., $E(X) = 6$ и $Var(X) = 9$, то $Y = \frac{X-6}{3} \sim N(0; 1)$. Нет.
10. Если ты отвечать на первые 10 вопросов этого теста наугад, то число правильных ответов - случайная величина, имеющая биномиальное распределение. Да.
11. По-моему, сегодня хорошая погода, и вместо контрольной можно было бы покататься на лыжах. Да!

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Любой ответ на 11 считается правильным.

Тест не является блокирующим.

Обозначения:

$E(X)$ - математическое ожидание

$Var(X)$ - дисперсия

Часть II Стоимость задач 10 баллов.

Задача 1

Совместный закон распределения случайных величин X и Y задан таблицей:

	$Y = -1$	$Y = 0$	$Y = 2$
$X = 0$	0,1	c	0,2
$X = 1$	0,1	0,2	0,1

Найдите c , $P(Y > -X)$, $E(X \cdot Y^2)$, $E(Y|X > 0)$

Решение:

$c = 0.3$ [1], $P(Y > -X) = 0.5$ [3], $E(XY^2) = 0.5$ [3], $E(Y|X > 0) = \frac{0.1}{0.4} = 0.25$ [3]

Задача 2

Случайный вектор $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ имеет нормальное распределение с математическим ожиданием $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ и ковариационной матрицей $\begin{pmatrix} 9 & -4,5 \\ -4,5 & 25 \end{pmatrix}$. Найдите $P(X_1 + 3X_2 > 20)$.

Решение:

$E(Y) = -1$ [2], $Var(Y) = 207$ [4], $P(Y > 20) = P(Z > \frac{21}{\sqrt{207}}) = P(Z > 1.46) = 0.07$ [4]

Задача 3

Совместная функция плотности имеет вид

$$p_{X,Y}(x,y) = \begin{cases} x+y, & \text{если } x \in [0; 1], y \in [0; 1] \\ 0, & \text{иначе} \end{cases}$$

Найдите $P(Y > 2X)$, $E(X)$

Решение:

$$P(Y > 2X) = \int_0^1 \int_0^{y/2} (x+y) dx dy = \frac{5}{24} [5]$$

$$E(X) = \int_0^1 \int_0^1 x(x+y) dx dy = \frac{7}{12} [5]$$

(если интеграл выписан верно, но не взят, то [3] вместо [5])

Задача 4

В супермаркете «Покупан» продаются различные вина:

Вина	Доля	Средняя цена за бутылку (у.е.)	Стандартное отклонение (у.е.)
Элитные	0,1	150	24
Дорогие	0,3	40	12
Дешевые	0,6	10	10

Чтобы оценить среднюю стоимость предлагаемого вина производится

случайная выборка 10 бутылок.

а) Какое количество элитных, дорогих и дешевых вин должно присутствовать в выборке, для того, чтобы выборочное среднее значение цены имело минимальную дисперсию? [5]

б) Чему равна минимальная дисперсия? [5]

Решение:

Используя метод множителей Лагранжа:

$$L = \frac{(0.1 \cdot 24)^2}{a} + \frac{(0.3 \cdot 12)^2}{b} + \frac{(0.6 \cdot 10)^2}{c} + \lambda(10 - a - b - c)$$

...

$a = 2, b = 3, c = 5$, можно было использовать готовую формулу $n_i = \frac{w_i \sigma_i}{\sum w_j \sigma_j}$

$$Var(\bar{X}^s) = 14,4$$

Задача 5

Допустим, что закон распределения X_n имеет вид:

X	-1	0	2
Prob	θ	$2\theta - 0.2$	$1.2 - 3\theta$

Имеется выборка: $X_1 = 0, X_2 = 2$.

а) Найдите оценку $\hat{\theta}$ методом максимального правдоподобия

б) Найдите оценку $\hat{\theta}$ методом моментов

Решение:

а) $(2\theta - 0.2)(1.2 - 3\theta) \rightarrow \max, \hat{\theta} = 0.25$ [5]

б) $2.4 - 7\hat{\theta} = 1, \hat{\theta} = 0.2$ [5]

Задача 6

В среднем 30% покупателей супермаркета делают покупку на сумму свыше 700 рублей. Какова вероятность того, что из 200 [случайно выбранных] покупателей более 33% сделают покупку на сумму свыше 700 рублей?

Решение:

$$P(\bar{X} > 0.33) = P\left(\frac{\bar{X} - 0.3}{\sqrt{\frac{0.3 \cdot 0.7}{200}}} > \frac{0.33 - 0.3}{\sqrt{\frac{0.3 \cdot 0.7}{200}}}\right) = P(Z > 1.03) = 0.15$$

Баллы: [3] - Var , [4] - Z , [3] - таблица

Задача 7

Пусть X_i нормально распределены и независимы. Имеется выборка из трех наблюдений: 2, 0, 1.

а) Найдите несмещенные оценки для математического ожидания и дисперсии, \bar{X} и $\hat{\sigma}^2$. [2] + [3]

б) Найдите вероятность того, что оценка дисперсии превосходит истинную дисперсию более чем в 3 раза [5]

Решение:

$$\bar{X} = 1, \hat{\sigma}^2 = 1$$

$$P(\hat{\sigma}^2 > 3\sigma^2) = P\left(2\frac{\hat{\sigma}^2}{\sigma^2} > 6\right) = P(\chi_2^2 > 6) = 0.05$$

Задача 8

Известно, что у случайной величины X есть математическое ожидание, $E(X) = 0$, и дисперсия.

- а) Укажите верхнюю границу для $P(X^2 > 4\text{Var}(X))$? [5]
 б) Найдите указанную вероятность, если дополнительно известно, что X нормально распределена. [5]

Решение:

- а) $P(X^2 > 4\text{Var}(X)) = P(|X - 0| > 2\sigma) \leq \frac{\text{Var}X}{4\text{Var}(X)} = \frac{1}{4}$
 б) $P(X^2 > 4\text{Var}(X)) = P(|Z| > 2) = 0.05$

Задача 9

Пусть X_i независимы и экспоненциально распределены, т.е. имеют функцию плотности вида $p(t) = \frac{1}{\theta}e^{-\frac{1}{\theta}t}$ при $t > 0$.

- а) Постройте оценку математического ожидания методом максимального правдоподобия [2]
 б) Является ли оценка несмещенной? [2]
 в) Найдите дисперсию оценки [2]
 г) С помощью неравенства Крамера-Рао проверьте, является ли оценка эффективной среди несмещенных оценок? [2]
 д) Является ли построенная оценка состоятельной? [2]

Решение:

- а) \bar{X}
 б) Да; в) $\text{Var}(\bar{X}) = \frac{\theta^2}{n}$; г) да: несмещенность и предел дисперсии равный нулю;

Задача 10

[Независимые] случайные величины X_i распределены равномерно на отрезке $[0; a]$, известно, что $a > 10$. Исследователь хочет оценить параметр $\theta = \frac{1}{P(X_i < 5)}$.

- а) Используя \bar{X}_n постройте несмещенную оценку $\hat{\theta}$ для θ [4]
 б) Найдите дисперсию построенной оценки [3]
 в) Является ли построенная оценка состоятельной? [3]

Решение:

- а) $E(\bar{X}) = \frac{a}{2}$, $\theta = \frac{1}{P(X_i < 5)} = \frac{1}{5/a} = \frac{1}{5}a$
 $\hat{\theta} = \frac{2}{5}\bar{X}$
 б) $\text{Var}(\hat{\theta}_n) = \left(\frac{2}{5}\right)^2 \cdot \frac{a^2}{12n}$
 в) $\lim \text{Var}(\hat{\theta}_n) = 0$, оценка несмещенная, следовательно, состоятельная.

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 11-х задач по выбору!

Задача 11А

Каждый день Кощей Бессмертный кладет в сундук случайное количество копеек (от одной до ста, равновероятно).

Сколько в среднем дней нужно Кощею, чтобы набралось не меньше рубля?

Решение:

Обозначим e_n - сколько дней осталось в среднем ждать, если уже набрано n копеек.

Тогда:

$$e_{100} = 0$$

$$e_{99} = 1$$

$$e_{98} = \frac{1}{100}e_{99} + \frac{99}{100}e_{100} + 1 = 1 + \frac{1}{100}$$

$$e_{97} = \frac{1}{100}e_{98} + \frac{1}{100}e_{99} + \frac{98}{100}e_{100} + 1 = \left(1 + \frac{1}{100}\right)^2$$

$$e_{96} = \frac{1}{100}e_{97} + \frac{1}{100}e_{98} + \frac{1}{100}e_{99} + \frac{97}{100}e_{100} + 1 = \left(1 + \frac{1}{100}\right)^3$$

...

По индукции легко доказать, что $e_n = \left(1 + \frac{1}{100}\right)^{99-n}$

Таким образом, $e_0 = \left(1 + \frac{1}{100}\right)^{99} = 2.718...$

Требуется решить одну из двух 11-х задач по выбору!

Задача 11Б

Каждый день Петя знакомится с новыми девушками. С вероятностью 0.7 ему удастся познакомиться с одной девушкой; с вероятностью 0.2 - с двумя; с вероятностью 0.1 - не удастся. Дни, когда Пете не удастся познакомиться ни с одной девушкой, Петя считает неудачными.

Какова вероятность, что до первого неудачного дня Пете удастся познакомиться [ровно] с 30-ю девушками?

Решение:

$$p_0 = 0.1, p_1 = 0.7 \cdot 0.1;$$

$$p_n = P(\text{в первый день Петя познакомился с одной девушкой})p_{n-1} + P(\text{в первый день Петя познакомился с двумя девушками})p_{n-2};$$

$$\text{Дифференциальное уравнение: } p_n = 0.7p_{n-1} + 0.1p_{n-2}$$

Подсказка: Думайте!

3.3 Контрольная работа № 3, 21.02.2007

Нужные и ненужные формулы:

T - сумма чего-то там.

Если H_0 верна, то $E(T) = \frac{n}{2}$ и $Var(T) = \frac{n}{4}$

T - сумма каких-то рангов.

Если H_0 верна, то $E(T) = \frac{n(n+1)}{4}$ и $Var(T) = \frac{n(n+1)(2n+1)}{24}$.

T - сумма каких-то рангов.

Если H_0 верна, то $E(T) = \frac{n_1(n_1+n_2+1)}{2}$, $Var(T) = \frac{n_1n_2(n_1+n_2+1)}{12}$.

$$\cos^2(x) + \sin^2(x) = 1$$

УДАЧИ!

Часть I.

Обведите нужный ответ

1. Если $X \sim N(0; 12)$, $Y \sim N(12, 24)$, $Corr(X, Y) = 0$, то $X+Y \sim N(12, 36)$.

Да. Нет.

[любой ответ считался правильным. на самом деле верный ответ - нет]

2. Если закон распределения X задан табличкой

x	0	1
Вероятность	0.5	0.5

, то X - нормально распределена. Да. Нет.

3. Непараметрические тесты неприменимы, если выборка имеет χ^2 распределение. Да. Нет.

4. Р-значение показывает вероятность отвергнуть нулевую гипотезу, когда она верна. Да. Нет.

5. Если t -статистика равна нулю, то Р-значение также равно нулю. Да. Нет.

6. Если гипотеза отвергается при 5%-ом уровне значимости, то она будет отвергаться и при 1%-ом уровне значимости. Да. Нет.

7. При прочих равных 90% доверительный интервал шире 95%-го. Да. Нет.

8. Значение функции плотности может превышать единицу. Да. Нет.

9. Для любой случайной величины $E(X^2) \geq (E(X))^2$. Да. Нет.

10. Если $Corr(X, Y) > 0$, то $E(X)E(Y) < E(XY)$. Да. Нет.

11. На экзаменационной работе не шутят! Нет, шутят.

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Ответ «да» означает истинное утверждение, ответ «нет» - ложное.

Тест не является блокирующим.

[Неправильное использование таблиц = штраф 2 балла]

[Неправильные степени свободы = штраф 2 балла]

Часть II Стоимость задач 10 баллов.

Задача 1

Из урны с 5 белыми и 7 черными шарами случайным образом вынимается 2 шара. Случайная величина принимает значение (-1), если оба шара - белые; 0, если шары разного цвета и 1, если оба шара черные.

- а) Найдите $P(X = -1)$ [2], $E(X)$ [3], $Var(X)$ [3]
- б) Постройте функцию распределения величины X [2, достаточно аккуратно выписать функцию]

Задача 2

Случайная величина X имеет функцию распределения $F_X(t) = \begin{cases} 0, & t < 0 \\ ct^2, & 0 \leq t < 1 \\ 1, & 1 \leq t \end{cases}$

- а) Найдите c [1], $P(0.5 < X < 2)$ [1], 25%-ый квантиль [1]
- б) Найдите $E(X)$ [2], $Var(X)$ [2], $Cov(X, -X)$ [1], $Corr(2X, 3X)$ [1]
- в) Выпишите функцию плотности величины X [1]

Задача 3

Доходности акций двух компаний являются случайными величинами X и Y с одинаковым математическим ожиданием и ковариационной матрицей $\begin{pmatrix} 4 & -2 \\ -2 & 9 \end{pmatrix}$.

- а) Найдите $Corr(X, Y)$ [1], $Corr = -\frac{1}{3}$
- б) [5] В какой пропорции нужно приобрести акции этих двух компаний, чтобы дисперсия доходности получившегося портфеля была наименьшей?
- в) [2] Можно ли утверждать, что величины $X + Y$ и $7X - 2Y$ независимы?
- г) [2] Изменится ли ответ на пункт «в», если дополнительно известно, что величины X и Y в совокупности нормально распределены?

Подсказка: Если R - доходность портфеля, то $R = \alpha X + (1 - \alpha)Y$

$\alpha = \frac{1}{11}$

Задача 4

Проверка 40 случайно выбранных лекций показала, что студент Халявин присутствовал только на двух из них.

- а) [4] Найдите 90%-ый доверительный интервал для вероятности увидеть Халявина на лекции.
- б) [5] Укажите минимальный размер выборки, необходимый для того, чтобы с вероятностью 0,9 выборочная доля посещаемых Халявиным лекций отличалась от соответствующей вероятности не более, чем на 0,1.
- в) [1] Какие предпосылки и теоремы использовались при ответах на

предыдущие пункты?

Задача 5

Изучается эффективность нового метода обучения. У группы из 40 студентов, обучавшихся по новой методике, средний бал на экзамене составил 322.12, а выборочное стандартное отклонение 54.53. Аналогичные показатели для независимой выборки из 60 студентов того же курса, обучавшихся по старой методике, приняли значения 304.61 и 62.61 соответственно.

- а) [4] Проверьте гипотезу о равенстве дисперсий оценок в двух группах.
- б) [1] Какие предпосылки использовались при ответе на «а»?
- в) [4] Постройте 90% доверительный интервал для разницы математических ожиданий оценок в двух группах
- г) [1] Можно ли считать новую методику более эффективной?

Задача 6

В парке отдыха за час 57 человек посетило аттракцион «Чертовое колесо», 48 - «Призрачные гонки» и 54 - «Американские горки».

Можно ли на 5% уровне значимости считать, что посетители одинаково любят эти три аттракциона?

Задача 7

Можно ли по имеющейся таблице утверждать о независимости пола и доминирующей руки на 5% уровне значимости?

Пол/рука	Правша	Левша
Мужчины	16	76
Женщины	25	81

Задача 8

Пусть X_i нормально распределены, независимы, $E(X_i) = 0$, $Var(X_i) = \theta$.

- а) [3] Постройте оценку $\hat{\theta}$ методом максимального правдоподобия.
- б) Проверьте свойства несмещенности, состоятельности, эффективности у построенной оценки. [каждое свойство по 2, если дано аккуратное определение, то 1]
- в) [1] Какая оценка более предпочтительна: построенная или привычная $\hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$?

Задача 9

Имеется две конкурирующие гипотезы:

H_0 : Величина X распределена равномерно на отрезке $[0; 100]$

H_a : Величина X распределена равномерно на отрезке $[50; 150]$

Исследователь выбрал такой критерей:

Если $X < c$, то использовать H_0 , иначе использовать H_a .

а) Дайте определение ошибок первого и второго рода. $[2 + 2]$

б) Постройте графики зависимостей ошибок первого и второго рода от c .
 $[3 + 3]$

Задача 10

Вася измерил длину 10 пойманных им рыб. Часть рыб была поймана на левом берегу реки, а часть - на правом. Бывалые рыбаки говорят, что на правом берегу реки рыба крупнее.

Левый берег	25	45	37	47	51
Правый берег	49	28	39	46	57

а) [10] С помощью теста Манна-Уитни (Mann-Whitney) проверьте гипотезу о том, что выбор берега реки не влияет на среднюю длину рыбы против альтернативной гипотезы, что на правом берегу рыба длиннее.

Разрешается использование нормальной аппроксимации

б) [Не оценивался] Возможно ли в этой задаче использовать (Wilcoxon Signed Rank Test)?

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 11-х задач по выбору!

Задача 11А

Имеются две монетки. Одна правильная, другая - выпадает орлом с вероятностью 0.45. Одну из них, неизвестно какую, подкинули n раз и сообщили Вам, сколько раз выпал орел. Ваша задача проверить гипотезу H_0 : «подбрасывалась правильная монетка» против H_a : «подбрасывалась неправильная монетка».

Каким должно быть наименьшее n и критерий выбора гипотезы, чтобы вероятность ошибок первого рода не превышала 10%, а вероятность ошибки второго рода не превышала 15%?

Требуется решить одну из двух 11-х задач по выбору!

Задача 11Б

Время горения лампочки – экспоненциальная случайная величина с математическим ожиданием равным θ . Вася включил одновременно 20 лампочек. Величина Y обозначает время самого первого перегорания.

а) [8] Найдите $E(Y)$

б) [6] Постройте с помощью Y несмещенную оценку для θ

в) [6] Сравните по эффективности оценку построенную в пункте «б» и обычное выборочное среднее

4 2007-2008

4.1 Контрольная работа № 1, 03.11.2007

Quote

The 50-50-90 rule: Anytime you have a 50-50 chance of getting something right, there's a 90% probability you'll get it wrong.

Andy Rooney

Часть I. Обведите верный ответ:

1. Для любой случайной величины $P(X > 0) \geq P(X + 1 > 0)$. Нет.
2. Для любой случайной величины с $E(X) < 2$, выполняется условие $P(X < 2) = 1$. Нет.
3. Если $A \subset B$, то $P(A|B) \leq P(B|A)$. Да.
4. Если X - случайная величина, то $E(X) + 1 = E(X + 1)$. Да.
5. Функция распределения случайной величины является неубывающей. Да.
6. Для любых событий A и B , выполняется $P(A|B) + P(A|B^c) = 1$. Нет.
7. Для любых событий A и B верно, что $P(A|B) \geq P(A \cap B)$, если обе вероятности существуют. Да.
8. Функция плотности может быть периодической. Нет.
9. Если случайная величина X имеет функцию плотности, то $P(X = 0) = 0$. Да.
10. Для неотрицательной случайной величины $E(X) \geq E(-X)$. Да.
11. Вероятность бывает отрицательной. Даже не знаю, что и сказать...

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Часть II Стоимость задач 10 баллов.

Задача 1

На день рождения к Васе пришли две Маши, два Саши, Петя и Коля. Все вместе с Васей сели за круглый стол. Какова вероятность, что Вася окажется между двумя тезками?

Слева должен сесть тот, у кого есть тезка. $p_1 = 4/6$

Справа должен сесть его парный. $p_2 = 1/5$

Итого: $p = p_1 \cdot p_2 = 2/15$

Задача 2

Поезда метро идут регулярно с интервалом 3 минуты. Пассажир приходит на платформу в случайный момент времени. Пусть X - время ожидания поезда в минутах.

Найдите $P(X < 1)$, $E(X)$

Устно: $p = 1/3$, $E(X) = 1.5$

Задача 3

Вы играете две партии в шахматы против незнакомца. Равновероятно незнакомец может оказаться новичком, любителем или профессионалом. Вероятности вашего выигрыша в отдельной партии, соответственно, будут равны: 0,9; 0,5; 0,3.

а) Какова вероятность выиграть первую партию?

б) Какова вероятность выиграть вторую партию, если вы выиграли первую?

Решение:

$$p_a = \frac{1}{3}(0.9 + 0.5 + 0.3) = \frac{17}{30},$$

$$p_b = \frac{1}{3}(0.9^2 + 0.5^2 + 0.3^2)/p_a = \frac{115}{170}$$

Задача 4

Время устного ответа на экзамене распределено по экспоненциальному закону, т.е. имеет функцию плотности $p(t) = c \cdot e^{-0.1t}$ при $t > 0$.

а) Найдите значение параметра c

б) Какова вероятность того, что Иванов будет отвечать более получаса?

в) Какова вероятность того, что Иванов будет отвечать еще более получаса, если он уже отвечает 15 минут?

г) Сколько времени в среднем длится ответ одного студента?

Решение:

а) либо взятие интеграла, либо готовый ответ: $c = 0.1$

б) $\int_{30}^{+\infty} p(t)dt = e^{-3} \approx 0.05$

в) такой же результат, как в «б»

г) $1/\lambda = 10$

Задача 5

Годовой договор страховой компании со спортсменом-теннисистом, предусматривает выплату страхового возмещения в случае травмы специального вида. Из предыдущей практики известно, что вероятность получения теннисистом такой травмы в любой фиксированный день равна 0,00037. Для периода действия договора вычислите

- а) Наиболее вероятное число страховых случаев
- б) Математическое ожидание числа страховых случаев
- в) Вероятность того, что не произойдет ни одного страхового случая
- г) Вероятность того, что произойдет ровно 2 страховых случая

P.S. Указанные вероятности вычислите двумя способами: используя биномиальное распределение и распределение Пуассона.

Решение:

«б» $365 \cdot 0.00037 = 0.13505$

Следовательно, «а», ближайшее целое равно 0.

Для Пуассоновского распределения: $\lambda = 0.13505$

в) $P(N = 0) = 0.99963^{365} \approx e^{-\lambda}$

г) $P(N = 2) = C_{365}^2 \cdot 0.99963^{363} \cdot 0.00037^2 \approx e^{-\lambda} \lambda^2 / 2$

Задача 6

Допустим, что закон распределения X имеет вид:

X	1	2	3
Prob	θ	2θ	$1 - 3\theta$

- а) Найдите $E(X)$
- б) При каких θ среднее будет наибольшим? При каких - наименьшим?

$E(X) = 3 - 4\theta$, $\theta \in [0; 1/3]$, $\theta_{max} = 0$, $\theta_{min} = 1/3$

Задача 7

Вася пригласил трех друзей навестить его. Каждый из них появится независимо от другого с вероятностью 0,9, 0,7 и 0,5 соответственно. Пусть N - количество пришедших гостей.

Найдите $E(N)$

$N = X_1 + X_2 + X_3$, где X_i равно 1 или 0 в зависимости от того, пришел ли друг. Значит $E(N) = E(X_1) + E(X_2) + E(X_3) = 0.9 + 0.7 + 0.5 = 2.1$

Задача 8

У спелестолога в каменоломнях сели батарейки в налобном фонаре, и он оказался в абсолютной темноте. В рюкзаке у него 6 батареек, 4 новых и 2 старых. Для работы фонаря требуется две новых батарейки. Спелестолог вытаскивает из рюкзака две батарейки наугад и вставляет их в фонарь.

Если фонарь не начинает работать, то спелестолог откладывает эти две батарейки и пробует следующие две и т.д.

- а) Найдите закон распределения числа попыток
- б) Сколько попыток в среднем потребуется?
- в) Какая попытка скорее всего будет первой удачной?

Решение:

$$P(N = 1) = \frac{C_4^2}{C_6^2} = 6/15$$

$$P(N = 3) = \frac{4 \cdot 2 \cdot 3 \cdot 1}{C_6^2 C_5^2} = 4/15$$

$$P(N = 2) = 5/15$$

$$E(N) = 28/15, \text{ первая.}$$

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух задач (9А или 9Б) по выбору!

Задача 9А

По краю идеально круглой столешницы отмечается наугад n точек. В этих точках к столешнице прикручиваются ножки. Какова вероятность того, что полученный столик с n ножками будет устойчивым?

Решение:

Имеется n способов выбрать левую точку. Оставшиеся $(n - 1)$ точка должны попасть в правую полуокружность относительно выбранной левой точки.

$$\text{Получаем } p = n \cdot (0.5)^{n-1}$$

Требуется решить одну из двух задач (9А или 9Б) по выбору!

Задача 9Б

На окружности с центром O (не внутри окружности!) сидят три муравья, их координаты независимы и равномерно распределены по окружности.

Два муравья A и B могут общаться друг с другом, если $\angle AOB < \pi/2$.

Какова вероятность того, что все три муравья смогут не перемещаясь общаться друг с другом (возможно через посредника)?

Решение:

Будем считать координату одного за точку отсчета.

На квадрате $[0; 1] \times [0; 1]$ нетрудно нарисовать нужное множество.

$$p = 3/8$$

4.2 Контрольная работа № 2, демо-версия, 21.01.2008

Демо-версия.

В контрольной на этом месте будет 10 тестовых вопросов!

Часть II Стоимость задач 10 баллов.

Задача 1

Совместный закон распределения случайных величин X и Y задан таблицей:

	$Y = -1$	$Y = 1$	$Y = 2$
$X = -1$	0,1	c	0,2
$X = 1$	0,1	0,1	0,1

Найдите c , $P(Y > X)$, $E(X \cdot Y)$, $E(X|Y > 0)$

Являются ли величины X и Y независимыми?

Задача 2

Случайный вектор $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ имеет нормальное распределение с математическим ожиданием $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ и ковариационной матрицей $\begin{pmatrix} 9 & -4 \\ -4 & 36 \end{pmatrix}$.

а) Найдите $P(X_1 + X_2 > 0)$.

б) Какое условное распределение имеет X_1 при условии, что $X_2 = -1$?

Задача 3

Совместная функция плотности имеет вид

$$p_{X,Y}(x,y) = \begin{cases} c(x-y), & \text{если } x \in [0;1], y \in [0;1], x > y \\ 0, & \text{иначе} \end{cases}$$

Найдите c , $P(3Y > X)$, $E(X)$, $E(X|Y > 0.5)$

Задача 4

Вероятность дождя в субботу 0.5, вероятность дождя в воскресенье 0.3. Корреляция между наличием дождя в субботу и наличием дождя в воскресенье равна r .

Какова вероятность того, что в выходные вообще не будет дождя?

Задача 5

Автор книги получает 50 тыс. рублей сразу после заключения контракта и 5 рублей за каждую проданную книгу. Автор предполагает, что количество книг, которые будут проданы - это случайная величина с

ожиданием в 10 тыс. книг и стандартным отклонением в 1 тыс. книг. Чему равен ожидаемый доход автора? Чему равна дисперсия дохода автора?

Задача 6

Сейчас акция стоит 1000 рублей. Каждый день цена может равновероятно либо возрасти на 3 рубля, либо упасть на 5 рублей.

- а) Какова вероятность того, что через 60 дней цена будет больше 900 рублей?
- б) Чему равно ожидаемое значение цены через 60 дней?

Задача 7

В данном регионе кандидата в парламент Обещаева И.И. поддерживает 60% населения. Сколько нужно опросить человек, чтобы с вероятностью 0,99 доля опрошенных избирателей, поддерживающих Обещаева И.И., отличалась от 0,6 (истинной доли) менее, чем на 0,01?

Задача 8

С помощью неравенства Чебышева, укажите границы, в которых находятся величины; рассчитайте также их точное значение

а) $P(-2\sigma < X - \mu < 2\sigma)$, $X \sim N(\mu; \sigma^2)$

б) $P(8 < X < 12)$, $X \sim U[0; 20]$

с) $P(-2 < X - E(X) < 2)$, X - имеет экспоненциальное распределение с $\lambda = 1$

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 9-х задач по выбору!

Задача 9А

Вы приехали в уездный город N . В городе кроме Вас живут M мирных граждан и U убийц. Каждый день на улице случайным образом встречаются два человека. Если встречаются два мирных гражданина, то они пожимают друг другу руки. Если встречаются мирный гражданин и убийца, то убийца убивает мирного гражданина. Если встречаются двое убийц, то оба погибают.

Каковы Ваши шансы выжить в этом городе? Зависят ли они от Вашей стратегии?

Требуется решить одну из двух 9-х задач по выбору!

Задача 9Б

Дед Мороз развешивает новогодние гирлянды. Аллея состоит из 2008 елок. Каждой гирляндой Дед Мороз соединяет две елки (не обязательно соседние). В результате Дед Мороз повесил 1004 гирлянды и все елки оказались украшенными. Какова вероятность того, что существует хотя бы одна гирлянда, пересекающаяся с каждой из других?

Например, гирлянда 5-123 (гирлянда соединяющая 5-ую и 123-ю елки) пересекает гирлянду 37-78 и гирлянду 110-318.

Подсказка: Думайте!

4.3 Контрольная работа № 2, 21.01.2008

Часть I. Обведите верный ответ:

1. Сумма двух нормальных независимых случайных величин нормальна. Да.
2. Нормальная случайная величина может принимать отрицательные значения. Да.
3. Пуассоновская случайная величина является непрерывной. Нет.
4. Дисперсия суммы зависимых величин всегда не меньше суммы дисперсий. Нет.
5. Теорема Муавра-Лапласа является частным случаем центральной предельной. Да.
6. Пусть X - длина наугад выловленного удава в сантиметрах, а Y - в дециметрах. Коэффициент корреляции между этими величинами равен $\frac{1}{10}$. Нет.
7. Математическое ожидание выборочного среднего не зависит от объема выборки, если X_i одинаково распределены. Да.
8. Зная закон распределения X и закон распределения Y можно восстановить совместный закон распределения пары (X, Y) . Нет.
9. Если X - непрерывная с.в., $E(X) = 6$ и $Var(X) = 9$, то $Y = \frac{X-6}{3} \sim N(0; 1)$. Нет.
10. Если ты отвечать на первые 10 вопросов этого теста наугад, то число правильных ответов - случайная величина, имеющая биномиальное распределение. Да.
11. Раз уж выпал свежий снег, то вместо контрольной можно было бы покататься на лыжах! Да.

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Любой ответ на 11 считается правильным.

Тест не является блокирующим.

Обозначения:

$E(X)$ - математическое ожидание

$Var(X)$ - дисперсия

Часть II Стоимость задач 10 баллов.

Задача 1

Совместный закон распределения случайных величин X и Y задан таблицей:

	$Y = -1$	$Y = 0$	$Y = 2$
$X = 0$	0,2	c	0,2
$X = 1$	0,1	0,2	0,1

Найдите c , $P(Y > -X)$, $E(X \cdot Y)$, $Corr(X, Y)$, $E(Y|X > 0)$

Answers:

$c = 0.2$ [1], далее $P(Y > -X) = 0.5$ [2] и $E(X \cdot Y) = 0,1$ [2]

$Corr(X, Y) = \frac{-0.02}{\sqrt{0.24 \cdot 1.41}}$ [3]

$E(Y|X > 0) = 0.25$ [2]

Задача 2

Случайный вектор $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ имеет нормальное распределение с математическим ожиданием $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ и ковариационной матрицей $\begin{pmatrix} 9 & -4,5 \\ -4,5 & 25 \end{pmatrix}$.

а) Найдите $P(X_1 + 3X_2 > 20)$. [5]

б) Какое условное распределение имеет X_1 при условии, что $X_2 = 0$? [5]

а) $E(S) = -1$, $Var(S) = 207$, $P(Z > 1.47) = 1 - 0.9292 = 0.0708$

б) $p(x_1|0) \sim \exp\left(-\frac{1}{2}\begin{pmatrix} x_1 - 2 & 0 + 1 \end{pmatrix} \begin{pmatrix} 9 & -4,5 \\ -4,5 & 25 \end{pmatrix}^{-1} \begin{pmatrix} x_1 - 2 \\ 0 + 1 \end{pmatrix}\right)$

$p(x_1|0) \sim \exp\left(-\frac{1}{2 \det} (25(x_1 - 2)^2 + 9(x_1 - 2) + 9)\right)$

$p(x_1|0) \sim \exp\left(-\frac{1}{2 \cdot 8.19} (x_1 - 1.82)^2\right)$

$Var(X_1|X_2 = 0) = 8.19$, $E(X_1|X_2 = 0) = 1.82$

Есть страшные люди, которые наизусть помнят, что:

$Var(X_1|X_2 = x_2) = (1 - \rho^2)\sigma_1^2$

$E(X_1|X_2 = x_2) = \mu_1 \rho \frac{\sigma_1}{\sigma_2} (x_2 - \mu_2)$

упоминание нормальности: [2]

Задача 3

Совместная функция плотности имеет вид

$$p_{X,Y}(x,y) = \begin{cases} x+y, & \text{если } x \in [0;1], y \in [0;1] \\ 0, & \text{иначе} \end{cases}$$

Найдите $P(Y > 2X)$, $E(X)$

Являются ли величины X и Y независимыми?

Решение:

$$P(Y > 2X) = \int_0^1 \int_0^{y/2} (x+y) dx dy = \frac{5}{24} \text{ [4]}$$

$$E(X) = \int_0^1 \int_0^1 x(x+y) dx dy = \frac{7}{12} \text{ [4]}$$

зависимы [2]

(если интеграл выписан верно, но не взят, то [3] вместо [4])

Задача 4

Вася может получить за экзамен равновероятно либо 8 баллов, либо 7 баллов. Петя может получить за экзамен либо 7 баллов - с вероятностью $1/3$; либо 6 баллов - с вероятностью $2/3$. Известно, что корреляция их результатов равна 0.7.

Какова вероятность того, что Петя и Вася покажут одинаковый результат?

Solution:

Рассмотрим $X = 8 - (\text{Васин бал})$ и $Y = (\text{Петин бал}) - 6$

$\text{Corr}(X, Y) = -0.7$ (т.к. при линейном преобразовании может поменяться только знак корреляции)

$$\text{Var}(X) = \frac{1}{2}(1 - \frac{1}{2})$$

$$\text{Var}(Y) = \frac{1}{3}(1 - \frac{1}{3})$$

Интересующая нас величина - это $P(X = 1 \cap Y = 1) = E(XY) = \text{Cov}(X, Y) + E(X)E(Y)$

$$\text{answer: } \frac{10-7\sqrt{2}}{60} \approx 0.001675$$

$$\text{key point: } \text{Cov} = -\frac{7\sqrt{2}}{60} \quad [5]$$

Расчет средних [2]

логический переход от средних и ковариации к вероятности [3]

Задача 5

В городе Туме проводят демографическое исследование семейных пар. Стандартное отклонение возраста мужа оказалось равным 5 годам, а стандартное отклонение возраста жены - 4 годам. Найдите корреляцию возраста жены и возраста мужа, если стандартное отклонение разности возрастов оказалось равным 2 годам.

$$\text{Answer: } \frac{37}{40} = 0,925$$

Задача 6

Сейчас акция стоит 100 рублей. Каждый день цена может равновероятно либо возрасти на 8%, либо упасть на 5%.

а) Какова вероятность того, что через 64 дня цена будет больше 110 рублей? [8]

б) Чему равно ожидаемое значение логарифма цены через 100 дней? [2]

Подсказка: $\ln(1,08) = 0,07696$, $\ln(0,95) = -0,05129$, $\ln(1,1) = 0,09531$

Частая ошибка в «а»- решение другой задачи, где проценты заменены на копейки.

Если неправильная задача решена полностью - ставится [4] вместо [8]

Пусть N - число подъемов акции.

$$\text{а) } P(100 \cdot 1,08^N \cdot 0,95^{64-N} > 110) =$$

$$P(N \ln(1,08) + (64 - N) \ln(0.95) > \ln(1.1)) =$$

$$P\left(N > \frac{\ln(1.1) - 64 \ln(0.95)}{\ln(1.08) - \ln(0.95)}\right)$$

Заметим, что N - биномиально распределена, примерно $N(64 \cdot \frac{1}{2}, 64 \cdot \frac{1}{4})$
 $Z = \frac{N-32}{4}$ - стандартная нормальная и $P(Z > -1,42) = 0.92$

b) $E(N \ln(1,08) + (100 - N) \ln(0.95))$

На этот раз $E(N) = 50$ и $E(\ln(P_{100})) = 1.28$

Задача 7

Допустим, что срок службы пылесоса имеет экспоненциальное распределение. В среднем один пылесос бесперебойно работает 7 лет. Завод предоставляет гарантию 5 лет на свои изделия. Предположим также, что примерно 80% потребителей аккуратно хранят все бумаги, необходимые, чтобы воспользоваться гарантией.

а) Какой процент потребителей в среднем обращается за гарантийным ремонтом? [4]

б) Какова вероятность того, что из 1000 потребителей за гарантийным ремонтом обратится более 35% покупателей? [6]

Подсказка: $\exp(5/7) = 2,0427$

Solution:

$$p_{break} = 1 - \exp(-5/7) = 0.51 = \int_0^5 \frac{1}{7} e^{-\frac{t}{7}} dt$$

$$p = 0.8 \cdot 0.51 \approx 0.4$$

$$E(S) = 1000p = 400, \text{ } Var(S) = 1000p(1-p) = 240$$

$$P(S > 350) = P(Z > -3.23) \approx 1$$

Задача 8

Известно, что у случайной величины X есть математическое ожидание, $E(X) = 0$, и дисперсия.

а) Укажите верхнюю границу для $P(X^2 > 2.56 \cdot Var(X))$?

б) Найдите указанную вероятность, если дополнительно известно, что X нормально распределена.

$$\text{а) } P(X^2 > 2.56 Var(X)) = P(|X - 0| > 1.6\sigma) \leq \frac{Var X}{2.56 Var(X)} = \frac{100}{256} \approx 0.4 \text{ [5]}$$

$$\text{б) } P(X^2 > 2.56 Var(X)) = P(|Z| > 1.6) = 0.11 \text{ [5]}$$

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 9-х задач по выбору!

Задача 9А

Св. X распределена равномерно на отрезке $[0; 1]$. Вася изготавливает неправильную монетку, которая выпадает «орлом» с вероятностью x и передает ее Пете. Петя, не зная x , и подкидывает монетку один раз. Она выпала «орлом».

а) Какова вероятность того, что она снова выпадет «орлом»?

б) Как выглядит ответ, если Пете известно, что монетка при n подбрасываниях k раз выпала орлом?

б) Искомая вероятность равна $Prob = f(k+1, n-k)/f(k, n-k)$, где $f(a, b) = \int_0^1 x^a (1-x)^b dx$

Проинтегрировав по частям, видим, что $f(a, b) = f(a + 1, b - 1) \frac{b}{a+1}$

Отсюда $f(a, b) = \frac{a!b!}{(a+b+1)!}$

Подставляем, и получаем: $Prob = \frac{k+1}{n+2}$

Если кто получит этот ответ другим (более интуитивным) образом - тому большой дополнительный балл (!) - обращайтесь на roah@yandex.ru

Задача 9Б

В семье n детей. Предположим, что вероятности рождения мальчика и девочки равны. Дед Мороз спросил каждого мальчика «Сколько у тебя сестер?» и сложив эти ответы получил X . Затем Дед Мороз спросил каждую девочку «Сколько у тебя сестер?» и сложив эти ответы получил Y . Например, если в семье 2 мальчика и 2 девочки, то каждая девочка скажет, что у нее одна сестра, а каждый мальчик скажет, что у него 2 сестры, $X = 4$, $Y = 2$

а) Найдите $E(X)$ и $E(Y)$

б) Найдите $Var(X)$, $Var(Y)$

Solution:

Занумеруем детей в порядке появления на свет.

Обозначим M_i - индикатор того, что i -ый ребенок - мальчик

И F_i - индикатор того, что i -ый ребенок - девочка

Конечно, $F_i + M_i = 1$ и $F_i M_i = 0$

M , F - общее число мальчиков и девочек соответственно

Запасаемся простыми фактами:

$$E(F_i) = E(M_i) = E(F_i^2) = E(M_i^2) = \frac{1}{2}$$

$$E(F) = E(M) = \frac{n}{2}$$

$$Var(F_i) = Var(M_i) = \frac{1}{4}$$

$$Var(F) = Var(M) = \frac{n}{4}$$

$$E(F^2) = E(M^2) = Var(F) + E(F)^2 = \frac{n(n+1)}{4}$$

$$E(F F_i) = \frac{n+1}{4}$$

Поехали:

Заметим, что $X_i = X_i + M_i F_i = M_i F$

Таким образом $X = MF = nF - F^2$

$$Y_i = F - F_i - X_i$$

$$Y = (n-1)F - MF = (n-1)F - nF + F^2 = F^2 - F$$

Далее берем матожидание (легко) и дисперсию (громоздко):

$$E(X) = E(Y) = \frac{n(n-1)}{4}$$

... (если кто решил до сих пор, то наверняка, он смог и дальше решить)

...

Подсказка: Думайте!

4.4 Контрольная работа № 3, демо-версия, 01.03.2008

Демо-версия крЗ!

Часть I. Здесь будет тест!

Часть II Стоимость задач 10 баллов.

Задача 1

Вася и Петя метают дротики по мишени. Каждый из них сделал по 100 попыток. Вася оказался метче Пети в 59 попытках.

а) На уровне значимости 5% проверьте гипотезу о том, что меткость Васи и Пети одинаковая, против альтернативной гипотезы о том, что Вася метче Пети.

б) Чему равно точное P -значение при проверке гипотезы в п. «а»?

Задача 2

Из 10 опрошенных студентов часть предпочитала готовиться по синему учебнику, а часть - по зеленому. В таблице представлены их итоговые баллы.

Синий	76	45	57	65		
Зеленый	49	59	66	81	38	88

С помощью теста Манна-Уитни (Mann-Whitney) проверьте гипотезу о том, что выбор учебника не меняет закона распределения оценки.

Задача 3

Имеется случайная выборка X_1, X_2, \dots, X_n , где все X_i имеют распределение, задаваемое табличкой:

X	1	2	5
P	a	0.1	0.9-a

а) Постройте оценку неизвестного a методом моментов

б) Является ли построенная оценка состоятельной?

Демо-версия!

Задача 4

Имеется случайная выборка X_1, X_2, \dots, X_n , где все X_i имеют $N(27, a)$ распределение.

Найдите оценку неизвестного a методом максимального правдоподобия

Напоминка: не забудьте проверить условия второго порядка

Задача 5

На курсе два потока, на первом потоке учатся 40 человек, на втором потоке 50 человек. Средний балл за контрольную на первом потоке равен 78 при (выборочном) стандартном отклонении в 7 баллов. На втором потоке средний балл равен 74 при (выборочном) стандартном отклонении в 8 баллов.

- Постройте 90% доверительный интервал для разницы баллов между двумя потоками
- На 10%-ом уровне значимости проверьте гипотезу о том, что результаты контрольной между потоками не отличаются.

Задача 6

Проверьте независимость пола респондента и предпочитаемого им сока:

	Апельсиновый	Томатный	Вишневый
М	69	40	23
Ж	74	62	34

Задача 7

На Древе познания Добра и Зла растет 6 плодов познания Добра и 5 плодов познания Зла. Адам и Ева съели по 2 плода. Какова вероятность того, что Ева познала Зло, если Адам познал Добро?

Задача 8

Пусть X_i - независимы и имеют функцию плотности $p(t) = e^{a-t}$ при $t > a$, где a - неизвестный параметр. В качестве оценки неизвестного a используется $\hat{a}_n = \min\{X_1, X_2, \dots, X_n\}$.

- Является ли предлагаемая оценка состоятельной?
- Является ли предлагаемая оценка несмещенной?

Solution of 8:

Заметим, что $\hat{a}_n \geq a$.

$$P(|\hat{a}_n - a| > \varepsilon) = P(\hat{a}_n - a > \varepsilon) = P(\hat{a}_n > a + \varepsilon) = P(\min\{X_1, X_2, \dots, X_n\} > a + \varepsilon) =$$

$$= P(X_1 > a + \varepsilon \cap X_2 > a + \varepsilon \cap \dots) = P(X_1 > a + \varepsilon) \cdot P(X_2 > a + \varepsilon) \cdot \dots = \left(\int_{a+\varepsilon}^{\infty} e^{a-t} dt \right)^n = (e^{-\varepsilon})^n = e^{-n\varepsilon}$$

$$\lim_{n \rightarrow \infty} e^{-n\varepsilon} = 0$$

- нет, не является ни при каких n , хотя смещение с ростом n убывает

Демо-версия кр 3!

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 9-х задач по выбору!

Задача 9А

Имеются две монетки. Одна правильная, другая - выпадает орлом с вероятностью 0.45. Одну из них, неизвестно какую, подкинули n раз и сообщили Вам, сколько раз выпал орел. Ваша задача проверить гипотезу H_0 : «подбрасывалась правильная монетка» против H_a : «подбрасывалась неправильная монетка».

Каким должно быть наименьшее n и критерий выбора гипотезы, чтобы вероятность ошибок первого рода не превышала 10%, а вероятность ошибки второго рода не превышала 15%?

Задача 9Б

Пусть X_i - iid, $U[-b; b]$. Имеется выборка из 2-х наблюдений. Вася строит оценку для b по формуле $\hat{b} = c \cdot (|X_1| + |X_2|)$.

- При каком c оценка будет несмещенной?
- При каком c оценка будет минимизировать средне-квадратичную ошибку, $MSE = E((\hat{b} - b)^2)$?

4.5 Контрольная работа № 3, 01.03.2008

Часть I. Обведите верный ответ:

- Мощность теста можно рассчитать заранее, до проведения теста. Да.
- Точное P -значение можно рассчитать заранее, до проведения теста. Нет.
- Если гипотеза отвергает при 5%-ом уровне значимости, то она обязательно будет отвергаться и при 10%-ом уровне значимости. Да.
- Мощность больше у того теста, у которого вероятность ошибки 1-го рода меньше. Нет.
- Функция плотности F -распределения $p(t)$ не определена при $t < 0$. Нет.
- При большом k случайную величину, имеющую χ_k^2 распределение, можно считать нормально распределенной. Да.
- Оценки метода моментов всегда несмещенные. Нет.

8. Оценки метода максимального правдоподобия асимптотически несмещенные. Да.

9. Непараметрические тесты можно использовать, даже если закон распределения выборки неизвестен. Да.

10. Неравенство Крамера-Рао применимо только к оценкам метода максимального правдоподобия. Нет.

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Да - истинное утверждение, Нет - ложное

Тест не является блокирующим.

Обозначения:

$E(X)$ - математическое ожидание

$Var(X)$ - дисперсия

Часть II Стоимость задач 10 баллов.

Задача 1

Школьник Вася аккуратно измерял время, которое ему требовалось, чтобы добраться от школы до дома. По результатам 90 наблюдений, среднее выборочное оказалось равным 14 мин, а несмещенная оценка дисперсии - 5 мин².

а) Постройте 90% доверительный интервал для среднего времени на дорогу [4]

б) На уровне значимости 10% проверьте гипотезу о том, что среднее время равно 14,5 мин, против альтернативной гипотезы о меньшем времени. [4]

в) Чему равно точное P -значение при проверке гипотезы в п. «б»? [2]

Ответы:

а) [13.61; 14.39]

б) Отвергается ($Z_{observed} = -2.12$, $Z_{critical} = -1.28$)

с) $P_{value} = 0.017$

Задача 2

Садовник осматривал розовые кусты и записывал число цветков. Всего в саду растет 25 розовых кустов. Предположим, что количество цветков на разных кустах независимы и одинаково распределены.

Вот заметки садовника:

12,17,21,14,15;21,16,24,11,14;22,17,21,14,15;12,26,14,21,14;11,31,18,13,18.

Проверьте гипотезу о том, что медиана количества цветков равна 19

Решение:

Заменяем числа на цифры 0 и 1 (0 - меньше 19 цветков), (1 - больше)

$$\hat{p} = \frac{8}{25} = 0.32$$

$$H_0: p = 0.5$$

$$H_a: p \neq 0.5$$

$$Z = \frac{0.32 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{25}}} = -1.8$$

При уровне значимости 5%, $Z_{critical} = 1.96$

H_0 - не отвергается.

Задача 3

Имеется случайная выборка X_1, X_2, \dots, X_n , где все X_i имеют распределение, задаваемое табличкой:

X	1	2	5
P	a	2a	1-3a

а) Постройте оценку неизвестного a методом моментов [5]

б) Является ли построенная оценка несмещенной? [5]

Ответы:

а) $\hat{a} = \frac{5 - \bar{X}}{10}$

б) да, является

Задача 4

Имеется случайная выборка X_1, X_2, \dots, X_n , где все X_i имеют $N(a, 4a)$ распределение.

Найдите оценку неизвестного a методом максимального правдоподобия [8]

Напоминка: не забудьте проверить условия второго порядка [2]

Решение:

$$L = -\frac{n}{2} \ln(a) - \frac{na}{8} - \frac{\sum X_i^2}{8a} + c$$

$$L' = 0 \text{ равносильно } \hat{a}^2 + 4\hat{a} + 4 = 4 + \frac{\sum X_i^2}{n}$$

$$\hat{a} = -2 + \sqrt{4 + \frac{\sum X_i^2}{n}}$$

Задача 5

Допустим, что логарифм дохода семьи имеет нормальное распределение. В городе А была проведена случайная выборка 40 семей, показавшая выборочную дисперсию 20 (тыс.р.)². В городе Б по 30 семьям выборочная дисперсия оказалась равной 32 (тыс.р.)².

На уровне значимости 5% проверьте гипотезу о том, что дисперсия (логарифма дохода) одинакова, против альтернативной гипотезы о том, что город А более однородный.

Solution:

$$F_{29,39} = \frac{32}{20} = 1.6$$

$$F_{critical} = 1.74$$

Гипотеза о том, что дисперсия одинакова не отвергается.

Задача 6

Учебная часть утверждает, что все три факультатива («Вязание крючком для экономистов», «Экономика вышивания крестиком» и «Статистические методы в макраме») одинаково популярны. В этом году на эти факультативы соответственно записалось 35, 31 и 40 человек. Правдоподобно ли заявление учебной части?

Решение:

$$\chi^2_{observed} = 1.15$$

$$\chi^2_{2,5\%} = 5.99$$

Правдоподобно

Задача 7

Снайпер попадает в «яблочко» с вероятностью 0.8, если в предыдущий раз он попал в «яблочко» и с вероятностью 0.7, если в предыдущий раз он не попал в «яблочко» или если это был первый выстрел. Снайпер стрелял по мишени 3 раза.

а) Какова вероятность попадания в «яблочко» при втором выстреле? [5]

б) Какова вероятность попадания в «яблочко» при втором выстреле, если при первом снайпер попал, а при третьем - промазал? [5]

Solution:

$$а) p = 0.7 \cdot 0.8 + 0.3 \cdot 0.7 = 0.77$$

$$б) p = \frac{0.7 \cdot 0.8 \cdot 0.2}{0.7 \cdot 0.8 \cdot 0.2 + 0.7 \cdot 0.2 \cdot 0.3} = \frac{8}{11}$$

Задача 8

Пусть X_i - независимы и распределены равномерно на $[a - 1; a]$, где a - неизвестный параметр. В качестве оценки неизвестного a используется $\hat{a}_n = \max\{X_1, X_2, \dots, X_n\}$.

а) Является ли предлагаемая оценка состоятельной? [8]

б) Является ли предлагаемая оценка несмещенной? [2]

Solution:

Заметим, что $\hat{a}_n \leq a$.

$$P(|\hat{a}_n - a| > \varepsilon) = P(-(\hat{a}_n - a) > \varepsilon) = P(\hat{a}_n < a - \varepsilon) = P(\max\{X_1, X_2, \dots, X_n\} < a - \varepsilon) =$$

$$= P(X_1 < a - \varepsilon \cap X_2 < a - \varepsilon \cap \dots) = P(X_1 < a - \varepsilon) \cdot P(X_2 < a - \varepsilon) \cdot \dots = (1 - \varepsilon)^n$$

$$\lim_{n \rightarrow \infty} (1 - \varepsilon)^n = 0$$

б) нет, не является ни при каких n , хотя смещение с ростом n убывает

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 9-х задач по выбору!

Задача 9А

Два лекарства испытывали на мужчинах и женщинах. Каждый человек принимал только одно лекарство. Общий процент людей, почувствовавших улучшение, больше среди принимавших лекарство А. Процент мужчин, почувствовавших улучшение, больше среди принимавших лекарство В. Процент женщин, почувствовавших улучшение, больше среди принимавших лекарство В. Возможно ли это?

Да.

en.wikipedia.org/wiki/Simpson%27s_paradox

Требуется решить одну из двух 9-х задач по выбору!

Задача 9Б

Есть два золотых слитка, разных по весу. Сначала взвесили первый слиток и получили результат X . Затем взвесили второй слиток и получили результат Y . Затем взвесили оба слитка и получили результат Z . Допустим, что ошибка каждого взвешивания - это случайная величина с нулевым средним и дисперсией σ^2 .

- а) Придумайте наилучшую оценку веса первого слитка.
- б) Сравните придуманную Вами оценку с оценкой, получаемой путем усреднения двух взвешиваний первого слитка.

Solution:

- а) Пусть истинные веса слитков равны x , y и z .

Назовем оценку буквой \hat{x}

$$\hat{x} = aX + bY + cZ$$

$$\text{Несмещенность: } E(\hat{x}) = aE(X) + bE(Y) + cE(Z) = ax + by + c(x + y) = x$$

$$a + c = 1, b + c = 0$$

$$\hat{x} = (1 - c)X + (-c)Y + cZ$$

Эффективность:

$$\text{Var}(\hat{x}) = ((1 - c)^2 + c^2 + c^2) \cdot \sigma^2 = (3c^2 - 2c + 1)\sigma^2$$

Чтобы минимизировать дисперсию нужно выбрать $c = 1/3$

$$\text{Т.е. } \hat{x} = \frac{2}{3}X - \frac{1}{3}Y + \frac{1}{3}Z$$

$$\text{б) } \text{Var}(\hat{x}) = \frac{2}{3}\sigma^2$$

$$\text{Var}\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{2}\sigma^2$$

Усреднение двух взвешиваний первого слитка лучше.

5 2008-2009

5.1 Контрольная работа № 1, демо-версия, ??..11.2008

Часть I. Обведите верный ответ:

1. Для любой случайной величины $P(X > 0) \geq P(X + 1 > 0)$. Да. Нет.
2. Для любой случайной величины с $E(X) < 2$, выполняется условие $P(X < 2) = 1$. Да. Нет.
3. Если $A \subset B$, то $P(A|B) \leq P(B|A)$. Да. Нет.
4. Если X - случайная величина, то $E(X) + 1 = E(X + 1)$. Да. Нет.
5. Функция распределения случайной величины является неубывающей. Да. Нет.
6. Для любых событий A и B , выполняется $P(A|B) + P(A|B^c) = 1$. Да. Нет.
7. Для любых событий A и B верно, что $P(A|B) \geq P(A \cap B)$, если обе вероятности существуют. Да. Нет.
8. Функция плотности может быть периодической. Да. Нет.
9. Если случайная величина X имеет функцию плотности, то $P(X = 0) = 0$. Да. Нет.
10. Для неотрицательной случайной величины $E(X) \geq E(-X)$. Да. Нет.
11. Вероятность бывает отрицательной. Да. Нет.

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Часть II Стоимость задач 10 баллов.

Задача 1

На день рождения к Васе пришли две Маши, два Саши, Петя и Коля. Все вместе с Васей сели за круглый стол. Какова вероятность, что Вася окажется между двумя тезками?

Задача 2

Поезда метро идут регулярно с интервалом 3 минуты. Пассажир приходит на платформу в случайный момент времени. Пусть X - время ожидания поезда в минутах.

Найдите $P(X < 1)$, $E(X)$

Задача 3

Жители уездного города N независимо друг от друга говорят правду с вероятностью $\frac{1}{3}$. Вчера мэр города заявил, что в 2014 году в городе будет проведен межпланетный шахматный турнир. Затем заместитель мэра подтвердил эту информацию.

Какова вероятность того, что шахматный турнир действительно будет проведен?

Задача 4

Время устного ответа на экзамене распределено по экспоненциальному закону, т.е. имеет функцию плотности $p(t) = c \cdot e^{-0.1t}$ при $t > 0$.

- а) Найдите значение параметра c
- б) Какова вероятность того, что Иванов будет отвечать более получаса?
- в) Какова вероятность того, что Иванов будет отвечать еще более получаса, если он уже отвечает 15 минут?
- г) Сколько времени в среднем длится ответ одного студента?

Задача 5

Студент решает тест (множественного выбора) проставлением ответов наугад. В тесте 10 вопросов, на каждый из которых 4 варианта ответов. Зачет ставится в том случае, если правильных ответов будет не менее 5.

- а) Найдите вероятность того, что студент правильно ответит только на один вопрос
- б) Найдите наиболее вероятное число правильных ответов
- в) Найдите математическое ожидание и дисперсию числа правильных ответов
- г) Найдите вероятность того, что студент получит зачет

Задача 6

Совместный закон распределения случайных величин X и Y задан таблицей:

	$Y = -1$	$Y = 0$	$Y = 2$
$X = 0$	0,2	c	0,2
$X = 1$	0,1	0,2	0,1

Найдите c , $P(Y > -X)$, $E(X \cdot Y)$, $Corr(X, Y)$, $E(Y|X > 0)$

Задача 7

Вася пригласил трех друзей навестить его. Каждый из них появится независимо от другого с вероятностью 0,9, 0,7 и 0,5 соответственно. Пусть N - количество пришедших гостей.

Найдите $E(N)$

Задача 8

Охотник, имеющий 4 патрона, стреляет по дичи до первого попадания или до израсходования всех патронов. Вероятность попадания при первом выстреле равна 0.6, при каждом последующем - уменьшается на 0.1. Найдите

- Закон распределения числа патронов, израсходованных охотником
- Математическое ожидание и дисперсию этой случайной величины

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух задач (9А или 9Б) по выбору!

Задача 9А

У Мистера Х есть n зонтиков. Зонтики мистер Х хранит дома и на работе. Каждый день утром мистер Х едет на работу, а каждый день вечером - возвращается домой. При этом каждый раз дождь идет с вероятностью 0.8 независимо от прошлого, (т.е. утром дождь идет с вероятностью 0.8 и вечером дождь идет с вероятностью 0.8 вне зависимости от того, что было утром). Если идет дождь и есть доступный зонтик, то мистер Х обязательно возьмет его в дорогу. Если дождя нет, то мистер Х поедет без зонтика.

Какой процент поездок окажется для мистера Х неудачными (т.е. будет идти дождь, а зонта не будет) в долгосрочном периоде?

Требуется решить одну из двух задач (9А или 9Б) по выбору!

Задача 9Б

Начинающая певица дает концерты каждый день. Каждый ее концерт приносит продюсеру 0.75 тысяч евро. После каждого концерта певица может впасть в депрессию с вероятностью 0.5. Самостоятельно выйти из депрессии певица не может. В депрессии она не в состоянии проводить концерты. Помочь ей могут только цветы от продюсера. Если подарить цветы на сумму $0 \leq x \leq 1$ тысяч евро, то она выйдет из депрессии с вероятностью \sqrt{x} .

Какова оптимальная стратегия продюсера?

5.2 Контрольная работа № 1, ??..11.2008

Часть I. Верны ли следующие утверждения? Обведите ваш выбор.

1. Пуассоновская случайная величина является непрерывной. Нет.
2. Не существует случайной величины с $E(X) = 2008$ и $Var(X) = 2008$. Да.
3. $P(A|B) = P(A \cap B|B)$ для любых событий A и B . Да.
4. $E(X/Y) = E(X)/E(Y)$ для любых случайных величин X и Y . Нет.
5. При увеличении t величина $P(X \leq t)$ не убывает. Да.
6. Для любых событий A и B , выполняется $P(A|B) + P(A|B^c) = 1$. Нет.
7. События A и B независимы, если они не могут наступить одновременно. Нет.
8. Функция плотности может принимать значения больше 2008. Да.
9. Если $P(A) = 0.7$ и $P(B) = 0.5$, то события A и B могут быть несовместными. Нет.
10. Если X - неотрицательная случайная величина, то $P(X \leq 0) = 0$. Нет.

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Ответ «Да» означает, что утверждение верно.

Ответ «Нет» означает, что утверждение неверно.

Любой ответ на 11 вопрос считается верным.

Часть II Стоимость задач 10 баллов.

Задача 1

Вася купил два арбуза у торговли тети Маши и один арбуз у торговли тети Оли. Арбузы у тети Маши спелые с вероятностью 90% (независимо друг от друга), арбузы у тети Оли спелые с вероятностью 80%.

- а) Какова вероятность того, что все три Васиных арбуза будут спелыми?
- б) Какова вероятность того, что хотя бы два арбуза из Васиных будут

спелыми?

в) Каково ожидаемое количество спелых арбузов у Васи?

Решение:

а) [3] $0.9 \cdot 0.9 \cdot 0.8$

б) [4] $2 \cdot 0.1 \cdot 0.9 \cdot 0.8 + 0.9 \cdot 0.9 \cdot 1 = 0.9(0.16 + 0.9) = 0.9 \cdot 1.06 = 0.954$

в) [3] $0.9 + 0.9 + 0.8 = 2.6$

Задача 2

Случайная величина X может принимать только значения 5 и 9, с неизвестными вероятностями

а) Каково наибольшее возможное математическое ожидание величины X ?

б) Какова наибольшая возможная дисперсия величины X ?

Ответ:

а) [5] 9 (если взять 9 с вероятностью один)

б) [5] 4 (если взять 5 и 9 равновероятно)

Задача 3

Предположим, что социологическим опросам доверяют 70% жителей. Те, кто доверяют, опросам всегда отвечают искренне; те, кто не доверяют, отвечают наугад. Социолог Петя в анкету очередного опроса включил вопрос «Доверяете ли Вы социологическим опросам?»

а) Какова вероятность, что случайно выбранный респондент ответит «Да»?

б) Какова вероятность того, что он действительно доверяет, если известно, что он ответил «Да»?

Решение:

а) [4] $0.7 + 0.3 \cdot 0.5 = 0.85$

б) [6] $\frac{0.7}{0.85} = \frac{14}{17} \approx 0.82$

Задача 4

Случайные величины X и Y независимы и имеют функции плотности $f(x) = \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{32}(x-1)^2}$ и $g(y) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{18}y^2}$ соответственно.

Найдите:

а) $E(X)$ [2], $Var(X)$ [2]

б) $E(X - Y)$ [3], $Var(X - Y)$ [3]

Решение:

Нормальная случайная величина имеют функцию плотности $p(t) = c \cdot \exp(-\frac{1}{2\sigma^2}(t - \mu)^2)$

Отсюда: $E(X) = 1$, $E(Y) = 0$, $Var(X) = 16$, $Var(Y) = 9$

Задача 5

Закон распределения пары случайных величин X и Y задан таблицей:

	$X = -1$	$X = 0$	$X = 2$
$Y = 1$	0.2	0.1	0.2
$Y = 2$	0.1	0.2	0.2

Найдите: $E(X)$, $E(Y)$, $Var(X)$, $Cov(X, Y)$, $Cov(2X + 3, -3Y + 1)$

Каждая величина по [2] очка.

$E(X) = 0.5$, $E(Y) = 1.5$, $Var(X) = 1.65$, $Cov(X, Y) = -0.05$, $Cov(2X + 3, -3Y + 1) = 0.3$

Задача 6

Время устного ответа на экзамене распределено по экспоненциальному закону, т.е. имеет функцию плотности $p(t) = c \cdot e^{-0.2t}$ при $t > 0$.

- а) [2] Найдите значение параметра c
- б) [2] Какова вероятность того, что Иванов будет отвечать более двадцати минут?
- в) [3] Какова вероятность того, что Иванов будет отвечать еще более двадцати минут, если он уже отвечает 10 минут?
- г) [3] Сколько времени в среднем длится ответ одного студента?

Ответы:

0.2 , $\frac{1}{e^4}$, $\frac{1}{e^4}$, 5

Задача 7

Полугодовой договор страховой компании со спортсменом-теннисистом, предусматривает выплату страхового возмещения в случае травмы специального вида. Из предыдущей практики известно, что вероятность получения теннисистом такой травмы в любой фиксированный день равна $0,00037$. Для периода действия договора вычислите

- а) [3] Математическое ожидание числа страховых случаев
- б) [3] Вероятность того, что не произойдет ни одного страхового случая
- в) [4] Вероятность того, что произойдет ровно 2 страховых случая

Любое разумное понимание «полугодовой» принимается. Т.е. подходят 182, 183, и если посчитаны только рабочие дни, и если взят пример марсианского теннисиста с указанием кол-ва дней в марсианском году и пр.

И биномиальные и пуассоновские ответы принимаются

Для 182:

- а) $182 \cdot 0.00037 = 0.06734$
- б) $(1 - 0.00037)^{182} \approx \exp(-0.06734)$
- с) $C_{182}^2 p^2 (1 - p)^{180} \approx 0.5 \exp(-0.06734) 0.06734^2$

Задача 8

Большой Адронный Коллайдер запускают ровно в полночь. Оставшееся время до Конца Света - случайная величина X распределенная равномерно от 0 до 16 часов. Когда произойдет Конец Света, механические часы остановятся и будут показывать время Y .

а) Найдите $P(Y < 2)$

б) Постройте функцию плотности для величины Y

в) Найдите $E(Y)$, $Var(Y)$

г) Найдите $Cov(X, Y)$

Комментарий: по остановившимся механическим часам, к примеру, невозможно отличить, прошло ли от пуска Коллайдера 2,7 часа или 14,7 часа, т.к. Y принимает значения только на отрезке от 0 до 12 часов.

а) $P(Y < 2) = 1/4$ [1]

б) два отрезка: на высоте $2/16$ (от 0 до 4) и $1/16$ (от 4 до 12) [3] в) $E(Y) = 5$ [1], $Var(Y) = 12.(3)$ [2]

д) $Cov(X, Y) = 3.(3)$ [3]

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух задач (9А или 9Б) по выбору!

Задача 9А

На даче у мистера А две входных двери. Сейчас у каждой двери стоит две пары ботинок. Перед каждой прогулкой он выбирает наугад одну из дверей для выхода из дома и надевает пару ботинок, стоящую у выбранной двери. Возвращаясь с прогулки мистер А случайным образом выбирает дверь, через которую он попадет в дом и снимает ботинки рядом с этой дверью. Сколько прогулок мистер А в среднем совершит, прежде чем обнаружит, что у выбранной им для выхода из дома двери не осталось ботинок?

Источник: American Mathematical Monthly, problem E3043, (1984, p.310; 1987, p.79)

Решение:

Составляется граф по которому «блуждает» мистер А. Пишутся рекуррентные соотношения. Получается 12 или 13 в зависимости от того, считать ли прогулку «босиком» или нет. Оба ответа считать правильными.

Тем, кто подумал, что у двери стоят по одной паре ставится [10] за всю задачу (т.к. это существенно упрощает граф), ответ у них выходит 5.

За правильный граф (с верными вероятностями) [10], за решение системы уравнений - еще [10].

Требуется решить одну из двух задач (9А или 9Б) по выбору!

Задача 9Б

Если смотреть на корпус Ж здания Вышки с Дурасовского переулка, то видно 70 окон расположенных прямоугольником 7×10 (7 этажей, т.к. первый не видно, и 10 окон на каждом этаже). Допустим, что каждое из них освещено вечером независимо от других с вероятностью одна вторая. Назовем «уголком» комбинацию из 4-х окон, расположенных квадратом, в которой освещено ровно три окна (не важно, какие). Пусть X - число «углков», возможно пересекающихся, видимых с Дурасовского переулка. Найдите $E(X)$ и $Var(X)$

Примечание - для наглядности:

X	X
	X

,

X	
X	X

,

	X
X	X

,

X	X
X	

- это «уголки».

X	X	X
	X	
X	X	

- в этой конфигурации три «уголка»;

X		X
	X	
X		X

- а здесь - ни одного «уголка».

Решение:

X раскладывается в сумму индикаторов.

Имеется $6 \cdot 9$ позиций для потенциального «уголка».

$$E(X) = 6 \cdot 9 \cdot 1/4 = 13.5 \text{ [8]}$$

Имеется $6 \cdot 5 + 5 \cdot 9$ «боковых» пересечений потенциальных позиций.

Имеется $5 \cdot 8$ «угловых» пересечений потенциальных позиций.

Только они и могут дать ковариацию.

$$Var(X) = 54 \cdot 1/4 \cdot 3/4 + 2 \cdot (6 \cdot 8 + 5 \cdot 9) \cdot 3/32 + 2 \cdot 5 \cdot 8 \cdot 5/64 = 541/16 \text{ [12]}$$

5.3 Контрольная работа № 2, демо-версия, 26.12.2008

Часть I. Обведите верный ответ:

1. Сумма двух нормальных независимых случайных величин нормальна. Да. Нет.
2. Нормальная случайная величина может принимать отрицательные значения. Да. Нет.
3. Пуассоновская случайная величина является непрерывной. Да. Нет.
4. Дисперсия суммы зависимых величин всегда не меньше суммы диспер-

сий. Да. Нет.

5. Теорема Муавра-Лапласа является частным случаем центральной предельной. Да. Нет.

6. Пусть X - длина наугад выловленного удава в сантиметрах, а Y - в дециметрах. Коэффициент корреляции между этими величинами равен $\frac{1}{10}$. Да. Нет.

7. Математическое ожидание выборочного среднего не зависит от объема выборки, если X_i одинаково распределены. Да. Нет.

8. Зная закон распределения X и закон распределения Y можно восстановить совместный закон распределения пары (X, Y) . Да. Нет.

9. Если X - непрерывная с.в., $E(X) = 6$ и $Var(X) = 9$, то $Y = \frac{X-6}{3} \sim N(0; 1)$. Да. Нет.

10. Если ты отвечать на первые 10 вопросов этого теста наугад, то число правильных ответов - случайная величина, имеющая биномиальное распределение. Да. Нет.

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Обозначения:

$E(X)$ - математическое ожидание

$Var(X)$ - дисперсия

Часть II Стоимость задач 10 баллов.

Задача 1

Совместная функция плотности имеет вид

$$p_{X,Y}(x,y) = \begin{cases} x+y, & \text{если } x \in [0; 1], y \in [0; 1] \\ 0, & \text{иначе} \end{cases}$$

Найдите $P(Y > 2X)$, $E(X)$

Являются ли величины X и Y независимыми?

Задача 2

Случайный вектор $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ имеет нормальное распределение с математическим ожиданием $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ и ковариационной матрицей $\begin{pmatrix} 9 & -4,5 \\ -4,5 & 25 \end{pmatrix}$.

а) Найдите $P(X_1 + 3X_2 > 20)$. [5]

б) Какое условное распределение имеет X_1 при условии, что $X_2 = 0$? [5]

Задача 3

Компания заключила 1000 однотипных договоров. Выплаты по каждому договору возникают независимо друг от друга с вероятностью 0,1. В случае наступления выплат их размер распределен экспоненциально со средним значением 1000 рублей.

а) Найдите дисперсию и среднее значение размера выплат по одному контракту.

б) Какова вероятность того, что компании потребуется более 110 тысяч рублей на выплаты по всем контрактам?

Задача 4

Определите, в каких границах может лежать $P(\frac{(X-30)^2}{Var(X)} < 3)$, если известно, что $E(X) = 30$. Можно ли уточнить ответ, если дополнительно известно, что X - экспоненциально распределена.

Задача 5

Предположим, что величины X_1, X_2, \dots, X_{13} - независимы и распределены нормально $N(\mu, \sigma^2)$. Найдите число a , если известно, что $P(\sum (X_i - \bar{X})^2 > a\sigma^2) = 0.1$.

Задача 6

Предположим, что оценки студентов на экзамене распределены равномерно на отрезке $[0; a]$. Вася хочет оценить вероятность того, что отдельно взятый студент наберет больше 30 баллов. Васе известно, что экзамен

сдавали 100 человек и 15 из них набрали более 60 баллов. Помогите Васе построить несмещенную оценку!

Коля напрямую узнал у наугад выбранных 50 студентов, получили ли они больше 30 баллов. Какая оценка вероятности имеет меньшую дисперсию, Васина или Колина?

Задача 7

К продавцу мороженого подходят покупатели: мамы, папы и дети. Предположим, что это независимые Пуассоновские потоки с интенсивностями 12, 10 и 16 чел/час.

- а) Какова вероятность того, что за час будет всего 30 покупателей?
- б) Какова вероятность того, что подошло одинаковое количество мам, пап детей, если за некий промежуток времени подошло ровно 30 покупателей?

Задача 8

Известно, что $X \sim N(\mu, \sigma^2)$ и $Y = \exp(X)$. В таком случае говорят, что Y имеет лог-нормальное распределение. Найдите $E(Y)$.

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 9-х задач по выбору!

Задача 9А

There are two unfair coins. One coin has 0.7 probability head-up; the other has 0.3 probability head-up. To begin with, you have no information on which is which. Now, you will toss the coin 10 times. Each time, if the coin is head-up, you will receive \$1; otherwise you will receive \$0. You can select one of the two coins before each toss. What is your best strategy to earn more money?

Требуется решить одну из двух 9-х задач по выбору!

Задача 9Б

Дед Мороз развешивает новогодние гирлянды на аллее. Вдоль аллеи высажено 2008 елок. Каждой гирляндой Дед Мороз соединяет две елки (не обязательно соседние). В результате Дед Мороз повесил 1004 гирлянды и все елки оказались украшенными. Какова вероятность того, что существует хотя бы одна гирлянда, пересекающаяся с каждой из других? Например, гирлянда 5-123 (гирлянда соединяющая 5-ую и 123-ю елки) пересекает гирлянду 37-78 и гирлянду 110-318.

Подсказка: Думайте!

5.4 Контрольная работа № 2, 26.12.2008

Часть I. Обведите верный ответ:

1. Если пара величин (X, Y) имеет совместное нормальное распределение, то каждая случайная величина по отдельности также имеет нормальное распределение. Верно.
2. Неравенство Чебышева неприменимо к дискретным случайным величинам. Нет.
3. Нормальная случайная величины является дискретной. Нет.
4. Дисперсия любой несмещенной оценки не превосходит дисперсию любой смещенной. Нет.
5. При большом количестве степеней свободы хи-квадрат распределение похоже на нормальное. Верно.
6. Сумма ста независимых равномерных на $[0; 1]$ величин является равномерной случайной величиной на $[0; 100]$. Нет.
7. Ковариация всегда больше корреляции по модулю. Нет.
8. Если величины X и Y одинаково распределены и $P(X = Y) = 0.9999$, то корреляция X и Y близка к единице. Нет.
Комментарий: корреляция показывает насколько согласованно величины изменяются. Например, взяв X с законом распределения:

X	1/2	1	2
$Prob$	0.00005	0.9999	0.00005

и $Y = 1/X$ получим отрицательную корреляцию между X и Y .
9. Нормально распределенная величина X и биномиально распределенная величина Y могут быть зависимы. Запросто.
10. Дисперсия суммы положительных величин всегда больше суммы дисперсий. Нет.
11. Раз уж выпал свежий снег, то вместо контрольной можно было бы покататься на лыжах! Неплохо бы.

[правильно=+1 балл; нет ответа=неправильно=0 баллов]

Любой ответ на 11 считается правильным.

Часть II Стоимость задач 10 баллов.

Задача 1

Совместная функция плотности имеет вид

$$p_{X,Y}(x,y) = \begin{cases} \frac{3}{2}x + \frac{1}{2}y, & \text{если } x \in [0; 1], y \in [0; 1] \\ 0, & \text{иначе} \end{cases}$$

а) Найдите $P(Y > X)$, $E(Y)$ [4 + 4]

б) Являются ли величины X и Y независимыми? [2]

а) $\int_0^1 \int_x^1 p(x,y) dy dx = 5/12$

$\int_0^1 \int_0^1 y \cdot p(x,y) dy dx = 13/24$

б) нет, т.к. совместная функция плотности не разлагается в произведение индивидуальных

За правильно выписанный интеграл 3 балла из 4-х.

Задача 2

Пусть X_i - независимы и одинаково распределены, причем $E(X_i) = 0$, $Var(X_i) = 1$ и $Var(X_i^2) = 2$

а) С помощью неравенства Чебышева оцените $P(|X_1 + X_2 + \dots + X_7| > 14)$ и $P(X_1^2 + X_2^2 + \dots + X_7^2 > 14)$

б) Найдите эти вероятности, если дополнительно известно, что X_i - нормально распределены.

Sol:

а) [2 + 2]

$$P(|X_1 + X_2 + \dots + X_7| > 14) \leq \frac{7}{14^2} = \frac{1}{28}$$

$$P(X_1^2 + \dots + X_7^2 > 14) = P(X_1^2 + \dots + X_7^2 - 7 > 7) = P(|X_1^2 + \dots + X_7^2 - 7| > 7) \leq \frac{2 \cdot 7}{7^2} = \frac{2}{7}$$

б) [3 + 3]

$$P(|X_1 + \dots + X_7| > 14) = P(|N(0; 1)| > 14/\sqrt{7}) = P(|N(0; 1)| > 5.29) \approx 0$$

$$P(X_1^2 + X_2^2 + \dots + X_7^2 > 14) \approx 0.05$$

Задача 3

Случайный вектор $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ имеет нормальное распределение с математическим ожиданием $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ и ковариационной матрицей $\begin{pmatrix} 9 & -5 \\ -5 & 25 \end{pmatrix}$.

а) Найдите $P(X_1 + 2X_2 > 20)$. [5]

б) Какое условное распределение имеет X_1 при условии, что $X_2 = 0$? [5]

Sol:

а) $X_1 + 2X_2 \sim N(5; 89)$, $P(Z > 1.59) = 0.056$

б) нормальное, причем $N(1.4; 8)$
 не указавшим нормальность (или функцию плотности явно), а только
 нашедшим среднее и дисперсию - штраф 2 балла.
 корреляция равна $-1/3$

Задача 4

Случайные величины X и Y независимы и равномерно распределены: X - на отрезке $[0; a]$, а Y - на отрезке $[0; 3a]$.

Вася знает значение XY и хочет оценить неизвестный параметр $\beta = E(X^2)$.

Петя знает значение Y^2 и хочет оценить тот же параметр β

а) Какую несмещенную оценку может построить Вася? [3]

б) Какую несмещенную оценку может построить Петя? [3]

в) У какой оценки дисперсия меньше? [4]

Sol

$$\beta = \frac{1}{3}a^2$$

$$E(XY) = \frac{3}{4}a^2$$

$$E(Y^2) = 3a^2$$

$$\hat{\beta}_1 = \frac{4}{9}XY$$

$$\hat{\beta}_2 = \frac{1}{9}Y^2$$

Т.к. обе оценки несмещенные вместо сравнения дисперсий можно сравнить квадраты ожиданий

$$\frac{16}{81}E(X^2Y^2) \text{ vs } \frac{1}{81}E(Y^4)$$

...

$$16a^4 \text{ vs } \frac{81}{5}a^4$$

Дисперсия васиной оценки меньше.

Задача 5

Вася играет в компьютерную игру, где нужно убить 80 однотипных монстров, чтобы пройти уровень. Количество патронов, которое Вася тратит на одного монстра имеет Пуассоновское распределение со средним значением 2 патрона.

а) Какова вероятность того, что на трех первых монстров придется потратить 6 патронов? [5]

б) Какова вероятность того, на всех монстров уровня придется потратить более 200 патронов? [5]

Решение:

Заметим, что Пуассоновская величина с положительной вероятностью принимает значение ноль, значит бывает, что монстрыдохнут от одного устрашающего взгляда Васи :)

а) Сумма трех независимых пуассоновских величин - пуассоновская с параметром: $3\lambda = 6$.

$$P(X = 6) = \exp^{-6} \frac{6^6}{6!} \approx 0.16$$

Ответ с факториалам считается полным.

б) Сумма 80 величин имеет пуассоновское распределение, но при большом количестве слагаемых пуассоновское мало отличается от нормального.

$$E(S) = 160, \text{Var}(S) = 160$$

$$P(S > 200) = P\left(\frac{S-160}{\sqrt{160}} > 3.16\right) \approx 0$$

Задача 6

Допустим, что срок службы пылесоса имеет экспоненциальное распределение. В среднем один пылесос бесперебойно работает 10 лет. Завод предоставляет гарантию 7 лет на свои изделия. Предположим для простоты, что все потребители соблюдают условия гарантии.

а) Какой процент потребителей в среднем обращается за гарантийным ремонтом? [5]

б) Какова вероятность того, что из 1000 потребителей за гарантийным ремонтом обратится более 55% покупателей? [5]

Подсказка: $\ln(2) \approx 0.7$

а) $\lambda = 1/10, P(X < 7) = 0.5$

б) $P(\bar{X} > 0.55) = P(N(0; 1) > \frac{0.05\sqrt{1000}}{0.5}) = P(N(0; 1) > 3.16) \approx 0$

Задача 7

Вася и Петя решают тест из 10 вопросов по теории вероятностей (на каждый вопрос есть два варианта ответа). Петя кое-что знает по первым пяти вопросам, поэтому вероятность правильного ответа на каждый равняется 0.9 независимо от других. Остальные пять вопросов Пете непонятны и он отвечает на них наугад. Вася списывает у Пети вопросы с 3-го по 7-ой, а остальные отвечает наугад.

Пусть X - число правильных ответов Пети, а Y - число правильных ответов Васи.

Найдите $\text{Var}(X)$, $\text{Var}(Y)$, $\text{Var}(X - Y)$.

$$\text{Var}(X) = 5 \cdot 0.1 \cdot 0.9 + 5 \cdot 0.5 \cdot 0.5 = 1.7$$

$$\text{Var}(Y) = 3 \cdot 0.1 \cdot 0.9 + 7 \cdot 0.5 \cdot 0.5 = 2.02$$

Пусть Z - число правильных ответов на вопросы с 3-го по 7-ой (у Пети и у Васи)

$$\begin{aligned} \text{Cov}(X, Y) &= \text{Cov}(Z + (X - Z), Z + (Y - Z)) = \text{Var}(Z) + \text{Cov}(X - Z, Z) + \\ &+ \text{Cov}(Z, Y - Z) + \text{Cov}(X - Z, Y - Z) = \text{Var}(Z) \end{aligned}$$

$Y - Z$ - это сколько правильных ответов дал лично Вася и оно не зависит

от числа Z правильных списанных ответов, значит $Cov(Y - Z, Z) = 0$
 Аналогично все остальные ковариации равны нулю.

$$Var(Z) = 3 \cdot 0.1 \cdot 0.9 + 2 \cdot 0.5 \cdot 0.5 = 0.77$$

Задача 8

Стоимость выборочного исследования генеральной совокупности, состоящей из 3 страт, определяется по формуле: $TC = n_1 \cdot c_1 + n_2 \cdot c_2 + n_3 \cdot c_3$,
 Где c_i - стоимость наблюдения из i -ой страты, n_i - число наблюдений в выборке, относящихся к страте i .

Предполагая, что стоимость исследования TC фиксирована и равна 7000, определите значения n_i , при которых дисперсия соответствующего выборочного стратифицированного среднего достигает наименьшего значения, если:

Страта	1	2	3
Среднее значение	40	80	150
Стандартная ошибка	10	20	60
Вес	20%	20%	60%
Цена наблюдения	4	16	25

Примечание: Округлите полученные значения до ближайших целых.

Любые совпадения с курсом экономической и социальной статистики случайны и непреднамеренны.

Чтобы оценка среднего по всем трем стратам была несмещена, она должна строиться по формуле:

$$\bar{X} = w_1 \bar{X}_1 + w_2 \bar{X}_2 + w_3 \bar{X}_3 \quad (\text{здесь } \bar{X}_i - \text{среднее арифметическое по } i\text{-ой страте})$$

Поэтому $Var(\bar{X})$ (минимизируемая функция) равняется:

$$Var(\bar{X}) = \sum \frac{w_i^2 \sigma_i^2}{n_i}$$

Принцип кота Матроскина¹ (ака бюджетное ограничение): $4n_1 + 16n_2 + 25n_3 = 7000$

Решаем Лагранжем и получаем ответ: 35, 35, 252.

Некоторые маньяки наизусть знают:

$$n_i = \frac{C}{\sum w_i \cdot \sigma_i \cdot \sqrt{c_i}} \frac{w_i \cdot \sigma_i}{\sqrt{c_i}}$$

Часть III Стоимость задачи 20 баллов.

Требуется решить одну из двух 9-х задач по выбору!

¹ «Чтобы продать что-нибудь ненужное, нужно сначала купить что-нибудь ненужное. А у нас денег нет!»

Задача 9А [20]

Усама бен Ладен хочет сделать запас в 1000 тротильных шашек в пещере А. Тротильные шашки производят на секретном заводе бесплатно. При транспортировке от завода до пещеры каждая шашка взрывается с небольшой вероятностью p . Если взрывается одна шашка, то взрываются и все остальные, перевозимые вместе с ней. Сам Усама при взрыве всегда чудом остается жив. Какими партиями нужно переносить шашки, чтобы минимизировать среднее число переносок?

[В стартовой пещере бесконечный запас шашек].

Solution:

0. Замечание: неудачные переноски считаются, т.к. иначе решение тривиально - пробовать нести по 1000 шашек.

1. т.к. p небольшая будем считать, что $\ln(1-p) \approx -p$. Уже страшно, да?

2. Допустим, что $s(n)$ оптимальная стратегия, указывающая, сколько нужно брать сейчас шашек, если осталось перенести n шашек. Возможно, что s зависит от n .

Обозначим $e(n)$ ожидаемое количество переносов при использовании оптимальной стратегии.

3. Начинаем:

$$s(1) = 1, e(1) = 1/(1-p)$$

$$s(n) = \operatorname{argmin}_a (1/(1-p)^a + e(n-a)), e(n) = \min_a (1/(1-p)^a + e(n-a))$$

Замечаем, что поначалу (где-то до $1/p$ шашек) все идет хорошо, а затем плохо...

4. Ищем упрощенное решение вида $s(n) = s$.

Ожидаемое число переносок равно $\frac{1000}{s} \frac{1}{(1-p)^s}$

Минимизируем по s . Получаем: $s = -1/\ln(1-p) \approx 1/p$.

5. Для тех кому интересно, точный график (10000 шашек, $p=0.01$):

[Здесь оставлено место для картины Усама-Бен-Ладен будь он не ладен таскает шашки.]

ps. В оригинале мы сканировали ксерокопию учебника Микоша. Сканер был очень умный: в него нужно положить стопку листов, а на выходе он выдавал готовый pdf файл. Проблема была в том, что он иногда жевал бумагу. В этом случае, он обрывал сканирование и нужно было начинать все заново. Возник вопрос, какого размера должна быть партия, чтобы минимизировать число подходов к ксероксу.

Требуется решить одну из двух 9-х задач по выбору!

Задача 9Б [20]

У Пети нет денег, но он может сыграть 100 игр следующего типа.

В каждой игре Петя может по своему желанию:

- либо без риска получить 1 рубль,
- либо назвать натуральное число $n > 1$ и выиграть n рублей с вероятностью $\frac{2}{n+1}$ или проиграть 1 рубль с вероятностью $\frac{n-1}{n+1}$.

Чтобы выбирать вторую альтернативу Петя должен иметь как минимум рубль. Пете позарез нужно 200 рублей. Как выглядит Петина оптимальная стратегия?

Solution:

1. Если сейчас 0 долларов, то брать 1 доллар.

Назовем ситуацию, «шоколадной» если можно выиграть без риска. Т.е. если игр осталось больше, чем недостающее количество денег.

2. Если игрок не в шоколаде, то оптимальным будет рисковать на первом ходе.

Почему? Получение одного доллара можно перенести на попозже.

3. В любой оптимальной стратегии достаточно одного успеха для выигрыша.

Почему? Допустим стратегии необходимо два успеха в двух рискованных играх. Заменяем их на одну рискованную игру. Получим большую вероятность.

Оптимальная стратегия:

Если сейчас 0 долларов, то брать доллар.

Пусть d - дефицит в долларах, а k - число оставшихся попыток.

Если $d \leq k$, то брать по доллару.

Если $d > k$, то с риском попробовать захватить $1 + d - k$ долларов.

Подсказка: Думайте!

5.5 Контрольная работа № 3, 02.03.2009

Часть I. Обведите верный ответ:

1. Если $X \sim N(0; 1)$, то $X^2 \sim \chi_1^2$. Верно. Нет.
2. Если $X \sim t_n$ и $Y \sim t_m$, то $\frac{X/n}{Y/m} \sim F_{n,m}$. Верно. Нет.
3. Если основная гипотеза отвергается при 1% уровне значимости, то она будет отвергаться и при 5% уровне значимости. Верно. Нет.
4. Неравенство Рао-Крамера справедливо только для оценок максимального правдоподобия. Верно. Нет.

5. Оценки метода максимального правдоподобия всегда несмещенные. Верно. Нет.

6. Ошибка второго рода происходит при отвержении основной гипотезы, когда она верна. Верно. Нет.

7. Из несмещенности оценки следует её состоятельность. Верно. Нет.

8. Длина доверительного интервала увеличивается при увеличении уровня доверия (доверительной вероятности) Верно. Нет.

9. Выборочное среднее независимых одинаково распределенных случайных величин с конечной дисперсией имеет асимптотически нормальное распределение. Верно. Нет.

10. Теорема Муавра-Лапласа является частным случаем ЦПТ. Верно. Нет.

11. Оценка, получаемая за эту контрольную, является несмещенной. Верно. Нет

[правильно=+1 балл; нет ответа=неправильно=0 баллов]
Любой ответ на 11 считается правильным.

Часть II Стоимость задач 10 баллов.

Теория вероятностей. Стоимость задач 10 баллов.

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 1

При контроле правдивости показаний подозреваемого на «детекторе лжи» вероятность признать ложью ответ, не соответствующий действительности, равна 0,99, вероятность ошибочно признать ложью правдивый ответ равна 0,01. Известно, что ответы, не соответствующие действительности, составляют 1% всех ответов подозреваемого.

Какова вероятность того, что ответ, признанный ложью, и в самом деле не соответствует действительности?

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 2

Предположим, что вероятность того, что среднегодовой доход наугад выбранного жителя некоторого города не превосходит уровень t , равна $P(I \leq t) = a + be^{-t/300}$ при $t \geq 0$.

Найдите:

- а) Числа a и b
- б) Средний доход жителей этого города (математическое ожидание, моду и медиану распределения). Какую из данных характеристик следует использовать для рапорта о высоком уровне жизни?

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 3

Доходности акций двух компаний являются случайными величинами X и Y с одинаковым математическим ожиданием и ковариационной матрицей

$$\begin{pmatrix} 4 & -2 \\ -2 & 9 \end{pmatrix}$$

- а) Найдите $\text{Corr}(X, Y)$
- б) В какой пропорции нужно приобрести акции этих двух компаний, чтобы дисперсия доходности получившегося портфеля была наименьшей?
Подсказка: Если R - доходность портфеля, то $R = \alpha X + (1 - \alpha)Y$
- в) Можно ли утверждать, что величины $X + Y$ и $7X - 2Y$ независимы?

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 4 Волшебный Сундук

Если присесть на Волшебный Сундук, то сумма денег, лежащих в нем, увеличится в два раза. Изначально в Сундуке был один рубль. Предположим, что «посадки» на Сундук - Пуассоновский процесс с интенсивностью λ . Каково ожидаемое количество денег в Сундуке к моменту времени t ?

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 5

На окружности единичной длины случайным образом равномерно и независимо друг от друга выбирают два отрезка: длины 0,3 и длины 0,4.

- а) Найдите функцию распределения длины пересечения этих отрезков.
- б) Найдите среднюю длину пересечения.

Построение и свойства оценок. Стоимость задач 10 баллов.
Нужно решить любые 2 (две) задачи на этой странице.

Задача 6

Асимметричная монета подбрасывается n раз. При этом X раз выпал «орел».

- а) Методом максимального правдоподобия найдите оценку вероятности «орла»
- б) Проверьте является ли полученная оценка состоятельной, несмещенной и эффективной.
- в) Считая, что n велико, укажите, в каких пределах с вероятностью 0,95 должно находиться значение оценки, если монета симметрична.

Нужно решить любые 2 (две) задачи на этой странице.

Задача 7

Вася попадает по мишени с вероятностью p при каждом выстреле независимо от других. Он стрелял до 3-х промахов (не обязательно подряд). При этом у него получилось X попаданий.

- а) Постройте оценку p с помощью метода максимального правдоподобия.
- б) Является ли полученная оценка несмещенной?

Нужно решить любые 2 (две) задачи на этой странице.

Задача 8

Известно, что величины X_1, \dots, X_n независимы и равномерны на $[0; b]$. Пусть Y - это минимум этих n величин. Вася знает n и Y .

- а) Найдите оценку b методом моментов.
- б) Является ли полученная оценка несмещенной?

Проверка гипотез и доверительные интервалы. Стоимость задач 10 баллов.

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 9

Вес выпускаемого заводом кирпича распределен по нормальному закону. По выборке из 16 штук средний вес кирпича составил 2.9 кг, выборочное стандартное отклонение 0,3. Постройте 80% доверительные интервалы для истинного значения веса кирпича и стандартного отклонения.

Примечание: можно строить односторонний интервал для стандартного отклонения, если таблицы не хватает, чтобы построить двусторонний.

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 10

В городе N за год родились 520 мальчиков и 500 девочек. Считая вероятность рождения мальчика неизменной:

- Проверить гипотезу о том, что мальчики и девочки рождаются одинаково часто против гипотезы о том, что мальчиков рождается больше, чем девочек
- Вычислить р-значение (минимальный уровень значимости, при котором основная гипотеза отвергается)
- Каким должен быть размер выборки, чтобы с вероятностью 0,95 можно было утверждать, что выборочная доля отличается от теоретической не более, чем на 0,02?

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 11

Проверьте гипотезу о независимости пола респондента и предпочитаемого им сока.

	Апельсиновый	Томатный	Вишневый
Мужчины	70	40	25
Женщины	75	60	35

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 12

Даны независимые выборки доходов выпускников двух ведущих экономических вузов А и В, по 50 выпускников каждого вуза: $\bar{X}_A = 650$, $\bar{X}_B = 690$, $\hat{\sigma}_A = 50$, $\hat{\sigma}_B = 70$.

Предполагая, что распределение доходов подчиняется нормальному закону, проверьте гипотезу об отсутствии преимуществ выпускников вуза В (уровень значимости 0,05).

Нужно решить любые **3 (три)** задачи на этой странице.

Задача 13

Величины X_1, X_2, \dots, X_{100} независимы и распределены $N(10, 16)$. Вася знает дисперсию, но не знает среднего. Поэтому он строит 60% доверительный интервал для истинного среднего значения. Какова вероятность того, что:

- а) Доверительный интервал покрывает настоящее среднее?
- б) Доверительный интервал покрывает число 9?

Часть III Стоимость задачи 20 баллов.

Нужно решить любую 1 (одну) задачу на этой странице.

Задача 14А [20]

Набранную книгу независимо друг от друга вычитывают два корректора. Первый корректор обнаружил m_1 опечаток, второй заметил m_2 опечаток. При этом m опечаток оказались обнаруженными и первым, и вторым корректорами.

- а) Постройте любым методом состоятельную оценку для общего числа опечаток (замеченных и незамеченных).
- б) Является ли построенная оценка несмещенной?

Нужно решить любую 1 (одну) задачу на этой странице.

Задача 14Б [20]

Вася хочет купить чудо-швабру! Магазинов, где продается чудо-швабра, бесконечно много. Любое посещение магазина связано с издержками равными $c > 0$. Цена чудо-швабры в каждом магазине имеет равномерное распределение на отрезке $[0; M]$. Цены в магазинах не меняются, т.е. при желании Вася может вернуться в уже посещенный им магазин для совершения покупки.

- а) Как выглядит оптимальная стратегия Васи? (Вася нейтрален к риску).
- б) Каковы ожидаемые Васиные затраты при использовании оптимальной стратегии?
- в) Сколько магазинов в среднем будет посещено?

Подсказка: Думайте!

6 2009-2010

6.1 Контрольная работа № 2, ??..12.2009

Часть I. Обведите верный ответ:

- 1. Сумма двух нормальных независимых случайных величин нормальна. Да. Нет.
- 2. Нормальная случайная величина может принимать отрицательные значения. Да. Нет.

3. Пуассоновская случайная величина является непрерывной. Да. Нет.
4. Дисперсия суммы зависимых величин всегда не меньше суммы дисперсий. Да. Нет.
5. Теорема Муавра-Лапласа является частным случаем центральной предельной. Да. Нет.
6. Пусть X - длина наугад выловленного удава в сантиметрах, а Y - в дециметрах. Коэффициент корреляции между этими величинами равен 0, 1. Да. Нет.
7. Для цепи Маркова невозвратное состояние это то, в которое невозможно вернуться. Да. Нет.
8. Последовательность независимых случайных величин является цепью Маркова. Да. Нет.
9. Зная закон распределения X и закон распределения Y можно восстановить совместный закон распределения пары (X, Y) . Да. Нет.
10. Если отвечать на первые 10 вопросов этого теста наугад, то число правильных ответов - случайная величина, имеющая биномиальное распределение. Да. Нет.
11. Если четыре причины возможного незачета устранены, то всегда найдется пятая. Да. Нет.
12. На дне глубокого сосуда
Лежат спокойно эн шаров.
Попеременно их оттуда
Таскают двое дураков.
Заняты это им приятно,
Они таскают тэ минут,
И каждый шар они обратно,
Его исследовав, кладут.
Ввиду занятия такого
Как вероятность велика,
Что был один глупей другого
И что шаров там было ка?

вероятно Виктор Скитович, автор «Раскинулось поле по модулю пять»

Часть II. Стоимость задач 10 баллов.

1. В городе Туме проводят демографическое исследование семейных пар. Стандартное отклонение возраста мужа оказалось равным 5 годам, а стандартное отклонение возраста жены - 4 годам. Найдите корреляцию возраста жены и возраста мужа, если стандартное отклонение разности возрастов оказалось равным 2 годам.
2. С.в. X и Y независимы и стандартно нормально распределены. Вычислите $\mathbb{P}(X < \sqrt{3})$ и $\mathbb{P}(X^2 + Y^2 < 6)$
3. Про случайную величину X известно, что $\mathbb{E}(X) = 16$ и $Var(X) = 12$.
 - (а) С помощью неравенства Чебышева оцените в каких пределах лежит вероятность $\mathbb{P}(|X - 16| > 4)$
 - (б) Найдите вероятность $\mathbb{P}(|X - 16| > 4)$, если известно, что X равномерна на $[10; 22]$
 - (с) Найдите вероятность $\mathbb{P}(|X - 16| > 4)$, если известно, что X нормально распределена
4. Случайный вектор $(X; Y)$ имеет нормальное распределение с математическим ожиданием $(-1; 4)$ и ковариационной матрицей $\begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}$.
Найдите $\mathbb{P}(2X + Y > 1)$ и $\mathbb{P}(2X + Y > 1 \mid Y = 2)$
5. Каждый день цена акции равновероятно поднимается или опускается на один рубль. Сейчас акция стоит 1000 рублей. Введем случайную величину X_i , обозначающую изменение курса акции за i -ый день. Найдите $\mathbb{E}(X_i)$ и $Var(X_i)$. С помощью центральной предельной теоремы найдите вероятность того, что через сто дней акция будет стоить больше 1010 рублей.

Дополнительная задача:

Вася и Петя подбрасывают несимметричную монету. Вероятность выпадения «орла» $= 0,25$. Если выпадает «орел», Вася отдает Пете 1 рубль, если «решка» – Петя отдает Васе 1 рубль. В начале игры у Васи – один рубль, у Пети – три рубля. Игра прекращается, как только у одного из игроков заканчиваются деньги.

1. Описать множество возможных состояний (указать тип состояния) и найти матрицу переходов из состояния в состояние.
2. Определить среднее время продолжительности игры
3. Определить вероятность того, что игра закончится победой Васи.

6.2 Контрольная работа № 4, ????.2010

1. Сколько нужно бросить игральных костей, чтобы вероятность выпадения хотя бы одной шестерки была не меньше 0,9?
2. Снайпер попадает в «яблочко» с вероятностью 0.8, если он в предыдущий выстрел попал в «яблочко» и с вероятностью 0.7, если в предыдущий раз не попал в «яблочко». Вероятность попасть в «яблочко» при первом выстреле также 0.7. Снайпер стреляет 2 раза.
 - (a) Определить вероятность попасть в «яблочко» при втором выстреле
 - (b) Какова вероятность того, что снайпер попал в «яблочко» при первом выстреле, если известно, что он попал при втором.
3. Случайная величина X моделирует время, проходящее между двумя телефонными звонками в справочную службу. Известно, что X распределена экспоненциально со стандартным отклонением равным 11 минутам. Со времени последнего звонка прошло 5 минут. Найдите функцию распределения и математическое ожидание времени, оставшегося до следующего звонка.
4. Известно, что для двух случайных величин X и Y : $E(X) = 1$, $E(Y) = 2$, $E(X^2) = 2$, $E(Y^2) = 8$, $E(XY) = 1$.
 - (a) Найдите ковариацию и коэффициент корреляции величин X и Y
 - (b) Определить, зависимы ли величины X и Y
 - (c) Вычислите дисперсию их суммы
5. Предположим, что время «жизни» X энергосберегающей лампы распределено по нормальному закону. По 10 наблюдениям среднее время «жизни» составило 1200 часов, а выборочное стандартное отклонение 120 часов.
 - (a) Постройте двусторонний доверительный интервал для математического ожидания величины X с уровнем доверия 0.90.
 - (b) Постройте двусторонний доверительный интервал для стандартного отклонения величины X с уровнем доверия 0.80.
 - (c) Какова вероятность, что несмещенная оценка для дисперсии, рассчитанная по 20 наблюдениям, отклонится от истинной дисперсии меньше, чем на 40%?

6. Учебная часть утверждает, что все три факультатива «Вязание крючком для экономистов», «Экономика вышивания крестиком» и «Статистические методы в макраме» одинаково популярны. В этом году на эти факультативы соответственно записалось 35, 31 и 40 человек. Правдоподобно ли заявление учебной части?

7. Имеются две конкурирующие гипотезы:

H_0 : Случайная величина X распределена равномерно на $(0,100)$

H_a : Случайная величина X распределена равномерно на $(50,150)$

Исследователь выбрал следующий критерий: если $X < c$, принимать гипотезу H_0 , иначе H_a .

(a) Дайте определение «ошибки первого рода», «ошибки второго рода», «мощности критерия».

(b) Постройте графики зависимости вероятностей ошибок первого и второго рода от c .

(c) Вычислите и вероятность ошибки второго рода, если уровень значимости критерия равен 0,05.

8. Из 10 опрошенных студентов часть предпочитала готовиться по синему учебнику, а часть по зеленому. В таблице представлены их итоговые баллы.

Учебник	Выборка					
Синий	76	45	57	65		
Зеленый	49	59	66	81	38	88

С помощью теста Манна-Уитни (Вилкоксона) проверьте гипотезу о том, что выбор учебника не меняет закона распределения оценки.

9. Случайная величина X , характеризующая срок службы элементов электронной аппаратуры, имеет распределение Релея: $F(x) = 1 - e^{-x^2/\theta}$ при $x \geq 0$. По случайной выборке X_1, X_2, \dots, X_n найдите оценку максимального правдоподобия параметра θ .

10. По случайной выборке X_1, X_2, \dots, X_n из равномерного на интервале $[\theta; \theta + 10]$ распределения методом моментов найдите оценку параметра θ . Дайте определение несмещенности и состоятельности оценки и определите, будет ли обладать этими свойствами найденная оценка.

11. При расчете страхового тарифа страховая компания предполагает, что вероятность наступления страхового случая 0.005. По итогам

прошедшего года из 10000 случайно выбранных договоров страховых случаев наблюдалось 67.

- (а) Согласуются ли полученные данные с предположением страховой компании? (Альтернатива: вероятность страхового случая больше)
- (б) Определить минимальный уровень значимости, при котором основная гипотеза отвергается (p-value).

7 2010-2011

7.1 Контрольная работа № 1, ?? .10.2010

Тест.

1. Если случайные события не могут произойти одновременно, то они независимы. Да. Нет.
2. Для любых случайных событий A, B, C верно что $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C)$. Да. Нет.
3. Функция плотности может быть периодической. Да. Нет.
4. Пусть $F(x)$ – функция распределения величины X . Тогда $\lim_{x \rightarrow \infty} F(x) = 0$. Да. Нет.
5. Для любых величин выполняется условие $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. Да. Нет.
6. Для любых величин выполняется условие $Var(X + Y) = Var(X) + Var(Y)$. Да. Нет.
7. Из совместной функции распределения величин X и Y можно получить функцию распределения величины $X + Y$. Да. Нет.
8. Пусть случайная величина X – длина удава в сантиметрах, а величина Y – его же длина в метрах. Тогда $Corr(X, Y) = 100$. Да. Нет.
9. Если две случайные величины независимы, то их ковариация равна 0. Да. Нет.
10. Если ковариация случайных величин равна 0, то они независимы. Да. Нет.

11. Пусть функция плотности величины X имеет вид $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. Тогда $E(X) = 0$. Да. Понятия не имею.

Задачи.

1. В жюри из трех человек два человека независимо друг от друга принимают правильное решение с вероятностью p , а третий для вынесения решения бросает монету. Окончательное решение выносится большинством голосов. Другое жюри – из одного человека выносит справедливое решение с вероятностью p . Какое из этих жюри выносит справедливое решение с большей вероятностью?
2. Васю можно застать на лекции с вероятностью 0,9, если на эту лекцию пришла Маша, и с вероятностью 0,5, если Маши на лекции нет. Маша бывает в среднем на трех лекциях из четырех. Найдите вероятность застать Васю на случайно выбранной лекции. Какова вероятность, что на лекции присутствует Маша, если на лекции есть Вася?
3. Число изюминок в булочке распределено по Пуассону. Сколько в среднем должны содержать изюма булочки, чтобы вероятность того, что в булочке найдется хотя бы одна изюминка, была не меньше 0.99?
4. Правильный кубик подбрасывают до тех пор, пока накопленная сумма очков не превысит 2. Пусть X – число подбрасываний кубика. Постройте функцию распределения величины X и найдите $E(X)$ и $Var(X)$.
5. Тест по теории вероятностей состоит из 10 вопросов, на каждый из которых предлагается 3 варианта ответа. Васе удается списать ответы на первые 5 вопросов у отличника Лёни, который никогда не ошибается, а на оставшиеся 5 он вынужден отвечать наугад. Оценка за зачет, величина X – число правильных ответов.
 - (a) Найдите математическое ожидание и дисперсию величины X , вероятность того, что Вася получит «отлично»
 - (b) Новый преподаватель предлагает усовершенствовать систему оценивания и вычитать балл за каждый неправильный ответ. Найти вероятность того, что Вася получит зачет по новой системе и ковариацию Васиных оценок в двух системах.

6. Закон распределения пары случайных величин X и Y и задан таблицей

	$X = -1$	$X = 0$	$X = 2$
$Y = 1$	0,2	0,1	0,2
$Y = 2$	0,1	0,2	0,2

Найдите $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $Var(X)$, $Cov(X, Y)$, $Cov(2X + 3, 1 - 3Y)$

7. Пусть величины X_1 и X_2 независимы и равномерно распределены на интервалах $[0; 2]$ и $[1; 3]$ соответственно. Найдите

- $\mathbb{E}(X_1)$, $Var(X_1)$, медиану X_1
- Совместную функцию распределения X_1 и X_2
- Функцию распределения и функцию плотности величины $W = \max\{X_1, X_2\}$

7.2 Контрольная работа № 2, ??.12.2010

1. Совместная плотность распределения случайных величин X и Y задана формулой:

$$f(x, y) = \frac{1}{2\pi} \frac{1}{\sqrt{1 - \rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)}$$

Найти $\mathbb{E}(X)$, $Var(Y)$, $Cov(X, Y)$, $\mathbb{P}(\{X > Y - 1\})$.

2. Случайные величины X , Y , Z независимы и стандартно нормально распределены. Вычислите $\mathbb{P}(\{X < \sqrt{2}\})$, $\mathbb{P}\left(\left\{\frac{|X|}{\sqrt{XY^2 + Z^2}} > 1\right\}\right)$, $\mathbb{P}(\{X^2 + Y^2 > 4\})$.

3. Доходности акций двух компаний являются случайными величинами X и Y , имеющими совместное нормальное распределение с математическим ожиданием $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ и ковариационной матрицей $\begin{pmatrix} 4 & -2 \\ -2 & 9 \end{pmatrix}$.

Найти $\mathbb{P}(\{X > 0\} \mid \{Y = 0\})$. В каком соотношении нужно приобрести акции этих компаний, чтобы риск (дисперсия) получившегося портфеля был минимальным? *Подсказка:* если R — доходность портфеля, то $R = \alpha X + (1 - \alpha)Y$. Можно ли утверждать, что случайные величины $X + Y$ и $7X - 2Y$ независимы?

4. Пусть $X_1; \dots; X_n$ — независимые одинаково распределённые случайные величины с плотностью распределения $f(x) = \frac{3}{x^4}, x \geq 1$.

Применим ли к данной последовательности закон больших чисел? С помощью неравенства Чебышева определить, сколько должно быть наблюдений в выборке, чтобы $\mathbb{P}(|\bar{X} - \mathbb{E}(X)| > 0,1) \leq 0,02$.

5. В большом-большом городе N 80 % аудиокиосков торгуют контрафактной продукцией. Какова вероятность того, что в наугад выбранных 90 киосках более 60 будут торговать контрафактной продукцией? Каким должен быть объём выборки, чтобы выборочная доля отличалась от истинной менее чем на 0,02 с вероятностью 0,95?
6. У входа в музей в корзине лежат 20 пар тапочек 36–45 размера (по 2 пары каждого размера). Случайным образом из корзины вытаскивается 2 тапочка. Пусть X_1 — размер первого тапочка, X_2 — размер второго. Являются ли случайные величины X_1 и X_2 зависимыми? Какова их ковариация? Найти математическое ожидание и дисперсию среднего размера $\frac{X_1 + X_2}{2}$.
7. В страховой компании «Ай» застрахованные автомобили можно условно поделить на 3 группы: недорогие (40 %), среднего класса (50 %) и дорогие (10 %). Из предыдущей практики известно, что средняя стоимость ремонта автомобиля зависит от его класса следующим образом:

	Недорогие	Среднего класса	Дорогие
Математическое ожидание	1	2,5	5
Стандартная ошибка	0,3	0,5	1

В каком соотношении в выборке объёма n должны быть представлены классы автомобилей, чтобы оценка средней стоимости ремонта (стратифицированное среднее) была наиболее точной?

8. Реализацией выборки $X = X_1; \dots; X_6$ являются следующие данные: $-0,8; 2,9; 4,4; -5,6; 1,1; -3,2$. Найти выборочное среднее и выборочную дисперсию, вариационный ряд и построить эмпирическую функцию распределения.
9. По выборке $X_1; \dots; X_n$ из равномерного распределения $\mathcal{U} \sim [0; \theta]$ с неизвестным параметром $\theta > 0$ требуется оценить θ . Будут ли оценки $T_1 = 2\bar{X}$, $T_2 = (n+1)X_{(1)}$ несмещёнными? Какая из них является более точной (эффективной)? Являются ли эти оценки состоятельными?

Дополнительная задача (не является обязательной). Случайные величины X и Y независимы, причём $\mathbb{P}(\{X = k\}) = \mathbb{P}(\{Y = k\}) = pq^{k-1}$, $0 < p < 1$, $q = 1 - p$, $k = 1; 2; \dots$. Найти $\mathbb{P}(\{X = k\} \mid \{X + Y = n\})$, $\mathbb{P}(\{Y = k\} \mid \{X = Y\})$.

7.3 Контрольная работа № 3, ??.03.2011

Решение задач с обозначением «(MIN)» необходимо и достаточно для получения удовлетворительной оценки за данную контрольную работу.

Задача 1. Во время эпидемии гриппа среди привитых людей заболевают в среднем 15 %, среди непривитых — 20 %. Ежегодно прививаются 10 % всего населения (прививка действует один год).

1. (MIN) Какой процент населения заболевает во время эпидемии гриппа?
2. Каков процент привитых среди заболевших людей?

Задача 2. Известно, что случайная величина $X \sim N(3; 25)$.

1. (MIN) Найти вероятности $\mathbb{P}(\{X > 4\})$ и $\mathbb{P}(\{4 < X \leq 5\})$.
2. Если известно также, что случайная величина Y имеет распределение $N(1; 16)$, что X и Y имеют совместное нормальное распределение и что $Corr(X; Y) = 0,4$, то найти $\mathbb{P}(\{X - 2Y < 4\})$.
3. Случайная величина $Z \sim N(6; 49)$ обладает тем свойством, что $D\left(X - 2Y + \frac{1}{\sqrt{7}}Z\right) = 88$. Найти условную вероятность $\mathbb{P}(\{X - 2Y < 4\} \mid \{Z > 8\})$.

Задача 3. Опрос домохозяйств, проживающих в Южном и Юго-Западном административных округах города Москвы, выявил следующие результаты:

Южный АО. Доходы, тыс. руб. Первая выборка, X .

8,4	15,6	21,2	15,2	38,2	28,3	19,1	44,1	68,2	56,0	34,5	33,8	84,2	45,0	28,2
-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Юго-Западный АО. Доходы, тыс. руб. Вторая выборка, Y .

54,8	26,6	14,4	22,0	23,9	43,3	65,1	18,0	69,2	32,0	46,7	64,0
------	------	------	------	------	------	------	------	------	------	------	------

Вычислены следующие суммы: $\sum_i X_i = 540$, $\sum_i Y_i = 480$, $\sum_i \frac{X_i^2}{15} = 1\,706,264$,

$$\sum_i \frac{Y_i^2}{12} = 1\,958,3, \quad \sum_i \frac{(X_i - 36)^2}{15} = 410,264, \quad \sum_i \frac{(Y_i - 40)^2}{12} = 358,3, \quad \sum_i \frac{(X_i - 40)^2}{15} = 426,264, \quad \sum_i \frac{(Y_i - 36)^2}{12} = 374,3.$$

1. (MIN) Постройте 90 % доверительный интервал для математического ожидания дохода в Юго-Западном АО.
2. На 5 % уровне значимости проверьте гипотезу о том, что средний доход в Юго-Западном АО не превышает среднего дохода в Южном АО, предполагая, что распределения доходов нормальны.
3. Проверьте гипотезу о равенстве распределений доходов в двух округах, используя статистику Вилкоксона—Манна—Уитни, на 5 %

уровне значимости. (Разрешается использование нормальной аппроксимации.)

Задача 4. Вася решил проверить известное утверждение о том, что бутерброд падает маслом вниз. Для этого он провёл серию из 200 испытаний. Ниже приведена таблица с результатами:

Бутерброд с маслом	Хлебом вверх	Хлебом вниз
Число наблюдений	105	95

(MIN) Можно ли утверждать, что бутерброд падает маслом вниз так же часто, как и маслом вверх? (Уровень значимости 0,05.)

Задача 5.

1. (MIN) По случайной выборке $X_1; \dots; X_n$ из нормального распределения $N(\mu_1; \mu_2 - \mu_1^2)$ методом моментов оценить параметры μ_1, μ_2 . Дать определения несмещённости и состоятельности и проверить выполнение этих свойств для оценки μ_1 .
2. По случайной выборке $X_1; \dots; X_n$ из нормального распределения $N(\theta; 1)$ методом максимального правдоподобия оценить параметр θ . Будет ли найденная оценка эффективной? Ответ обосновать.

8 2011-2012

8.1 Контрольная работа №1, 24.10.2011

Quote

...all models are approximations. Essentially, all models are wrong, but some are useful. However, the approximate nature of the model must always be borne in mind...

George Edward Pelham Box

УДАЧИ!

Часть I. Верны ли следующие утверждения? Отметьте плюсом верные утверждения, а минусом – неверные.

Утверждение	Верно?
1. События A и B зависимы, если $\mathbb{P}(A B) > \mathbb{P}(A)$.	+
2. При умножении случайной величины на 2, ее функция плотности умножается на 2.	–
3. Ковариация всегда лежит на отрезке $[-1; 1]$.	–
4. Если $\mathbb{P}(A B) = \mathbb{P}(B A)$, то $\mathbb{P}(A) = \mathbb{P}(B)$.	+
5. Если $\mathbb{P}(A B) > \mathbb{P}(A)$, то $\mathbb{P}(B A) < \mathbb{P}(B)$.	–
6. У экспоненциальной случайной величины может не быть функции плотности.	–
7. При умножении случайной величины на 2, дисперсия домножается на 2.	–
8. У нормальной случайной величины среднее и дисперсия равны.	–
9. Функция распределения не может принимать значений больших 2011.	+
10. Если $P(A) = 0.7$ и $P(B) = 0.5$, то события A и B могут быть независимы.	+
11. Вероятность встретить на улице динозавра равна 0,5.	–

правильно=+1 балл; неправильно=0 баллов, нет ответа=0.3 балла

Любой ответ на 11 вопрос считается верным.

Часть II Стоимость задач 10 баллов.

1. Из карточек составлено слово «СТАТИСТИКА». Из этих карточек случайно без возвращения выбирают 5 карточек. Найдите вероятность того, что из отобранных карточек можно составить слово «ТАКСИ».

$$P(A) = \frac{3 \cdot 2^3}{C_{10}^5} = \frac{2}{21} \approx 0,095 \quad (1)$$

2. При рентгеновском обследовании вероятность обнаружить туберкулез у больного туберкулезом равна 0,9. Вероятность принять здорового за больного равна 0,01. Доля больных туберкулезом по отношению ко всему населению равна 0,001. Найдите вероятность того, что человек здоров, если он был признан больным при обследовании.

$$P(A|B) = \frac{0,999 \cdot 0,01}{0,999 \cdot 0,01 + 0,001 \cdot 0,9} \approx 0,917 \quad (2)$$

3. При переливании крови надо учитывать группы крови донора и больного. Человеку, имеющему четвертую (AB) группу крови, можно перелить кровь любой группы. Человеку со второй (A) или третьей (B) группой можно перелить кровь той же группы или первой. Человеку с первой (0) группой крови только кровь первой группы. Среди населения 33,7% имеют первую, 37,5% – вторую, 20,9% – третью и 7,9% – четвертую группы крови.

- (a) Найдите вероятность того, что случайно взятому больному можно перелить кровь случайно взятого донора. [5 points]

$$P(A_1) = 0,079 + 0,209(0,209 + 0,337) + 0,375(0,375 + 0,337) + 0,337 \cdot 0,337 \approx 0,574 \quad (3)$$

- (b) Найдите вероятность того, что переливание можно осуществить, если есть два донора. [5 points]

$$P(A_2) \approx 0,778 \quad (4)$$

4. Вася сидит на контрольной работе между Дашей и Машей и отвечает на 10 тестовых вопросов. На каждый вопрос есть два варианта ответа, «да» или «нет». Первые три ответа Васе удалось списать у Маши, следующие три – у Даши, а оставшиеся четыре пришлось проставить наугад. Маша ошибается с вероятностью 0,1, а Даша – с вероятностью 0,7.

- (a) Найдите вероятность того, что Вася ответил на все 10 вопросов правильно. [3 points]

$$P(X_v = 10) = 0,9^3 \cdot 0,3^3 \cdot 0,5^4$$

- (b) Вычислите корреляцию между числом правильных ответов Васи и Даши, Васи и Маши. [7 points]

$$Var(X_m) = 0,9, Var(X_d) = 2,1, Var(X_v) = 0,27 + 0,63 + 1 = 1,9$$

$$Corr(X_v, X_d) = \frac{0,27}{\sqrt{1,9 \cdot 2,1}} \quad (5)$$

$$Corr(X_v, X_m) = \frac{0,63}{\sqrt{1,9 \cdot 0,9}} \quad (6)$$

Подсказка: иногда задача упрощается, если представить случайную величину в виде суммы.

5. Случайная величина X имеет функцию плотности

$$f(x) = \begin{cases} cx^{-4}, & x \geq 1 \\ 0, & x < 1 \end{cases} \quad (7)$$

Найдите

- (a) Значение c [1 point]

$$c = 3$$

- (b) Функцию распределения $F(x)$ [3 points]

$$F(x) = \begin{cases} 0, & x < 1 \\ 1 - x^{-3}, & x \geq 1 \end{cases} \quad (8)$$

- (c) Вероятность $\mathbb{P}(0,5 < X < 1,5)$ [3 points]

$$\mathbb{P}(0,5 < X < 1,5) = 1 - 1,5^{-3} = \frac{19}{27} \approx 0,70 \quad (9)$$

- (d) Математическое ожидание $\mathbb{E}(X)$ и дисперсию $Var(X)$ случайной величины X [3 points]

Заметим, что $\mathbb{E}(X^a) = 3/(3-a)$. Поэтому $\mathbb{E}(X) = 3/2$ и $\mathbb{E}(X^2) = 3$. Значит $Var(X) = 3/4$.

6. Случайная величина X имеет функцию плотности

$$f(x) = \begin{cases} cx^{-4}, x \geq 1 \\ 0, x < 1 \end{cases} \quad (10)$$

Найдите

(a) Функцию плотности случайной величины $Y = 1/X$ [4 points]

$$F(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(1/X \leq y) = \mathbb{P}(X \geq 1/y) = \begin{cases} 0, y < 0 \\ y^3, y \in [0; 1] \\ 1, y \geq 1 \end{cases} \quad (11)$$

$$p(y) = \begin{cases} 3y^2, y \in [0; 1] \\ 0, y \notin [0; 1] \end{cases} \quad (12)$$

(b) Корреляцию случайных величин Y и X . [6 points]

$$\mathbb{E}(X) = 3/2, \mathbb{E}(Y) = 3/4, \mathbb{E}(XY) = E(1) = 1, \text{ значит } Cov(X, Y) = 1 - 9/8 = -1/8$$

$$\mathbb{E}(Y^2) = 3/5, Var(Y) = 3/80, Corr(X, Y) = -\sqrt{5}/3 \approx 0,75$$

7. Для случайной величины X , имеющей функцию плотности

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (13)$$

вычислите центральный момент порядка 2011.

Функция плотности симметрична около нуля, поэтому:

$$\mathbb{E}((X - \mathbb{E}(X))^{2011}) = \mathbb{E}(X^{2011}) = 0 \quad (14)$$

8. Для случайных величин X и Y заданы следующие значения: $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 4$, $\mathbb{E}(XY) = 8$, $Var(X) = Var(Y) = 9$. Для случайных величин $U = X + Y$ и $V = X - Y$ вычислите:

(a) $\mathbb{E}(U)$, $Var(U)$, $\mathbb{E}(V)$, $Var(V)$, $Cov(U, V)$

$$\mathbb{E}(U) = 5 \text{ [1 pt]}, \mathbb{E}(V) = -3 \text{ [1 pt]}, Var(U) = 26 \text{ [2 pts]}, Var(V) = 10 \text{ [2 pts]}, Cov(U, V) = 0 \text{ [2 pts]}$$

- (b) Можно ли утверждать, что случайные величины U и V независимы? [2 points]

Нет, даже нулевой ковариации недостаточно для того, чтобы говорить о независимости случайных величин.

9. Белка нашла 80 орехов. Каждый орех оказывается пустым независимо от других с вероятностью 0,05. Случайная величина X – это количество пустых орехов у белки.

- (a) Найдите $\mathbb{E}(X)$ и $Var(X)$ [3 points]

$$\mathbb{E}(X) = 80 \cdot 0,05 = 4 \quad (15)$$

$$Var(X) = 80 \cdot 0,05 \cdot 0,95 = 4 \cdot 0,95 \quad (16)$$

- (b) Найдите точную вероятность $\mathbb{P}(X = 5)$ [3 points]

$$P(X = 5) = C_{80}^5 0,05^5 0,95^{75} \quad (17)$$

- (c) Найдите вероятность $\mathbb{P}(X = 5)$, используя пуассоновскую аппроксимацию. [3 points]

$$P(X = 5) \approx \exp(-4) 4^5 / 5! \quad (18)$$

- (d) Оцените максимальную ошибку при расчете вероятности с использованием пуассоновской аппроксимации. [1 point]

$$\Delta \leq \min\{p, np^2\} = \min\{0,05, 4 \cdot 0,05\} = 0,05 \quad (19)$$

10. Охраняемая Сверхсекретная Зона – это прямоугольник 50 на 100 метров с вершинами в точках $(0;0)$, $(100;50)$, $(100;0)$ и $(0;50)$. Охранник обходит Зону по периметру по часовой стрелке. Пусть X и Y – координаты охранника в случайный момент времени.

- (a) Найдите $\mathbb{P}(X > 20)$, $\mathbb{P}(X > 20 | X > Y)$, $\mathbb{P}(X > Y | X > 20)$ [1+2+2 points]

$$\mathbb{P}(X > 20) = \frac{80 + 80 + 50}{300} = 0,7 \quad (20)$$

$$\mathbb{P}(X > 20 | X > Y) = \frac{80 + 50 + 50}{100 + 50 + 50} = 0,9 \quad (21)$$

$$\mathbb{P}(X > Y | X > 20) = \frac{80 + 50 + 50}{80 + 80 + 50} = \frac{6}{7} \quad (22)$$

(b) Найдите $\mathbb{E}(X)$ [1 point]

$$\mathbb{E}(X) = 50 \quad (23)$$

(c) Постройте функцию распределения случайной величины X . [2 points]

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{6} + \frac{4}{600}x, & x \in [0; 100) \\ 1, & x \geq 100 \end{cases} \quad (24)$$

У функции два скачка высотой по $1/6$, в точках $x = 0$ и $x = 100$. На остальных участках функция линейна.

(d) Верно ли, что случайные величины X и Y независимы? [2 points]

Нет, например, если $Y = 50$ мы можем быть уверены в том, что $X \notin [10; 90]$.

Часть III Стоимость задачи 20 баллов.

Задача. Мы подбрасываем правильную монетку до тех пор пока не выпадет три орла подряд или три решки подряд. Если игра оканчивается тремя орлами, то мы не получаем ничего. Если игра оканчивается тремя решками, то мы получаем по рублю за каждую решку непосредственно перед которой выпадал орел. Каков средний выигрыш в эту игру?

8.2 Контрольная работа №2, 29.12.2011

Разрешается использование калькулятора.

При себе можно иметь шпаргалку А4.

Обозначения:

$\mathbb{P}(A)$ – вероятность A

$\mathbb{E}(X)$ – математическое ожидание

$Var(X)$ – дисперсия

\bar{A} – отрицание события A

Quote

“Can you do addition?” the White Queen asked. “What’s one and one and one and one and one and one and one and one and one?” “I don’t

know,” said Alice. “I lost count.” “She can’t do addition,” said the Red Queen.

Lewis Carroll

УДАЧИ!

Часть I. Верны ли следующие утверждения? Отметьте плюсом верные утверждения, а минусом – неверные.

Утверждение	Верно?
1. Нормальное распределение является частным случаем Пуассоновского.	Ложно
2. Оценка не может быть одновременно несмещенной и эффективной.	Ложно
3. Среднее выборочное является состоятельной оценкой для математического ожидания.	Верно
4. Из некоррелированности случайных величин, имеющих совместное нормальное распределение, следует их независимость.	Верно
5. Зная закон распределения вектора (X, Y) всегда можно найти закон распределения X .	Верно
6. Неравенство Чебышева неприменимо к нормальным случайным величинам.	Ложно
7. Сумма двух независимых стандартных нормальных величин является стандартной нормальной.	Ложно
8. Корреляция между любыми равномерными случайными величинами равна нулю.	Ложно
9. Корреляция между температурой завтра в Москве по Цельсию и по Фаренгейту равна единице.	Верно
10. Состоятельная оценка может быть смещенной.	Верно
11. Я хорошо себя вел в этом году и Дед Мороз подарит мне хорошую оценку по теории вероятностей.	

правильно=+1 балл; неправильно=0 баллов, нет ответа=0.3 балла

Любой ответ на 11 вопрос считается верным.

Часть II Стоимость задач 10 баллов.

1. Совместная функция плотности величин X и Y имеет вид

$$f(x, y) = \begin{cases} 2(x^3 + y^3), & \text{если } x \in [0; 1], y \in [0; 1] \\ 0, & \text{иначе} \end{cases} \quad (25)$$

- (a) [2] Найдите $\mathbb{P}(X + Y > 1)$
- (b) [6] Найдите $Cov(X, Y)$
- (c) [1] Являются ли величины X и Y независимыми?
- (d) [1] Являются ли величины X и Y одинаково распределенными?

Ответы:

- (a) $\mathbb{P}(X + Y > 1) = 4/5$. Здесь нужно брать интеграл...
- (b) $\mathbb{E}(X) = 13/20 = 0,065$, $\mathbb{E}(XY) = 2/5 = 0,6$, $Cov(X, Y) = -9/400 = -0,0225$
- (c) Нет, так как функция плотности не раскладывается в произведение $h(x) \cdot g(y)$.
- (d) Да, так как функция плотности симметрична по x и y

2. Величины X_1 и X_2 независимы и равномерны на отрезке $[-b; b]$.
Вася строит оценку для b по формуле $\hat{b} = c \cdot (|X_1| + |X_2|)$.

- (a) [5] При каком c оценка будет несмещенной?
- (b) [5] При каком c оценка будет минимизировать средне-квадратичную ошибку, $MSE = E((\hat{b} - b)^2)$?

Ответы:

- (a) Заметим, что величина $|X_i|$ распределена равномерно на $[0; b]$, поэтому $\mathbb{E}(|X_i|) = b/2$ и $Var(|X_i|) = b^2/12$. Значит $\mathbb{E}(\hat{b}) = cb$ и для несмещенности $c = 1$.
- (b) Находим MSE через b и c :

$$MSE = Var(\hat{b}) + (\mathbb{E}(\hat{b}) - b)^2 = 2c^2 \cdot \frac{b^2}{12} + (c-1)^2 \cdot b^2 = b^2 \left(\frac{7}{6}c^2 - 2c + 1 \right) \quad (26)$$

Отсюда $c = \frac{6}{7}$.

3. Вася пишет 3 контрольные работы по микроэкономике, обозначим их результаты величинами X_1 , X_2 и X_3 . Кроме того, Вася пишет 3 контрольные работы по макроэкономике, обозначим их результаты величинами Y_1 , Y_2 и Y_3 . Предположим, что результаты всех контрольных независимы друг от друга. В среднем Вася пишет на один и тот же балл, $\mathbb{E}(X_i) = \mathbb{E}(Y_i) = \mu$. Дисперсия результатов по микро — маленькая, $\text{Var}(X_i) = \sigma^2$, дисперсия результатов по макро — большая, $\text{Var}(Y_i) = 2\sigma^2$.

- (a) [3] Является ли оценка $\hat{\mu}_1 = (X_1 + X_2 + X_3 + Y_1 + Y_2 + Y_3)/6$ несмещенной для μ ?
- (b) [7] Найдите самую эффективную несмещенную оценку вида $\hat{\mu}_2 = \alpha \bar{X} + \beta \bar{Y}$

Ответы:

- (a) $\mathbb{E}(\hat{\mu}_1) = 6\mu/6 = \mu$, несмещенная
- (b) $\mathbb{E}(\hat{\mu}_2) = \alpha\mu + \beta\mu$ и $\text{Var}(\hat{\mu}_2) = \alpha^2 \frac{\sigma^2}{3} + \beta^2 \frac{2\sigma^2}{3}$ Для несмещенности необходимо условие $\alpha + \beta = 1$. Для минимизации дисперсии получаем уравнение

$$\alpha - 2(1 - \alpha) = 0 \quad (27)$$

Отсюда оценка имеет вид $\frac{2}{3}\bar{X} + \frac{1}{3}\bar{Y}$

4. Каждую весну дед Мазай плавая на лодке спасает в среднем 9 зайцев, дисперсия количества спасенных зайцев за одну весну равна 9. Количество спасенных зайцев за разные года — независимые случайные величины. Точный закон распределения числа зайцев неизвестен.

- (a) [3] Оцените в каких пределах лежит вероятность того, что за три года дед Мазай спасет от 20 до 34 зайцев.
- (b) [3] Оцените в каких пределах лежит вероятность того, что за одну весну дед Мазай спасет более 11 зайцев.
- (c) [4] Используя нормальную аппроксимацию, посчитайте вероятность того, что за 50 лет дед Мазай спасет от 430 до 470 зайцев.

Ответы:

- (a) $S = X_1 + X_2 + X_3$, слагаемых мало, использовать нормальное распределение некорректно. Можно использовать неравенство Чебышева, $\mathbb{E}(S) = 27$, $Var(S) = 27$, поэтому

$$\mathbb{P}(S \in [20; 34]) = \mathbb{P}(|S - \mathbb{E}(X)| \leq 7) \geq 1 - \frac{27}{7^2} = \frac{22}{49} \quad (28)$$

- (b) Используем неравенство Маркова:

$$\mathbb{P}(X_1 \geq 12) \leq \mathbb{E}(X_1)/12 = 9/12 = 0,75 \quad (29)$$

- (c) Если $S = X_1 + \dots + X_{50}$, то можно считать, что $S \sim N(450; 450)$, поэтому

$$\mathbb{P}(S \in [430; 470]) \approx \mathbb{P}(N(0; 1) \in [-0,94; +0,94]) \approx 0,6542 \quad (30)$$

5. Вектор $\vec{X} = (X_1; X_2)$ имеет совместное нормальное распределение

$$\vec{X} \sim N \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}; \begin{pmatrix} 1 & -1 \\ -1 & 9 \end{pmatrix} \right) \quad (31)$$

- (a) [2] Найдите $\mathbb{P}(X_1 + X_2 > 1)$
 (b) [4] Какое совместное распределение имеет вектор $(X_1; Y)$, где $Y = X_1 + X_2$?
 (c) [4] Какой вид имеет условное распределение случайной величины X_1 , если известно что $X_2 = 2$?

Ответы:

- (a) Если $Y = X_1 + X_2$, то $\mathbb{E}(Y) = 3$ и $Var(Y) = 1 + 9 - 2 = 8$, значит $\mathbb{P}(Y > 1) = \mathbb{P}(N(0; 1) > -2/\sqrt{8}) \approx \mathbb{P}(N(0; 1) > -0,71) \approx 0,7602$
 (b) Находим $Cov(X_1, Y) = 1 - 1 = 0$. Итого: вектор имеет совместное нормальное распределение с

$$(X_1, Y) \sim N \left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}; \begin{pmatrix} 1 & 0 \\ 0 & 8 \end{pmatrix} \right) \quad (32)$$

- (с) Стандартизируем величины. Т.е. мы хотим представить их в виде:

$$\begin{cases} X_1 = 1 + aZ_1 + bZ_2 \\ X_2 = 2 + cZ_2 \end{cases} \quad (33)$$

Единица и двойка — это математические ожидания X_1 и X_2 . Мы хотим, чтобы величины Z_1 и Z_2 были $N(0; 1)$ и независимы. Получаем систему:

$$\begin{cases} Var(X_1) = 1 \\ Var(X_2) = 9 \\ Cov(X_1, X_2) = -1 \end{cases} \Leftrightarrow \begin{cases} a^2 + b^2 = 1 \\ c^2 = 9 \\ bc = -1 \end{cases} \quad (34)$$

Одно из решений этой системы : $c = 3$, $b = -1/3$, $a = 2\sqrt{2}/3$
Используя это разложение получаем:

$$\begin{aligned} (X_1 | X_2 = 2) &\sim \left(1 + \frac{2\sqrt{2}}{3}Z_1 - \frac{1}{3}Z_2 \mid 2 + 3Z_2 = 2 \right) \sim \\ &\sim \left(1 + \frac{2\sqrt{2}}{3}Z_1 - \frac{1}{3}Z_2 \mid Z_2 = 0 \right) \sim \left(1 + \frac{2\sqrt{2}}{3}Z_1 \right) \sim N(1; 8/9) \end{aligned} \quad (35)$$

Еще возможные решения: выделить полный квадрат в совместной функции плотности, готовая формула, etc

6. В большом-большом городе наугад выбирается n человек. Каждый из них отвечает, любит ли он мороженое эскимо на палочке. Обозначим \hat{p} долю людей в нашей выборке, любящих эскимо на палочке.

- (a) [3] Чему равно максимально возможное значение $Var(\hat{p})$?
(b) [7] Какое минимальное количество человек нужно опросить, чтобы вероятность того, что выборочная доля \hat{p} отличалась от истинной доли более чем на 0.02, была менее 10%?

Ответы:

- (a) $Var(\hat{p}) = \frac{p(1-p)}{n}$. Максимально возможное значение $p(1-p)$ равно 1/4, поэтому максимально возможное значение $Var(\hat{p}) = 1/4n$.

(b) У нас задано неравенство:

$$\mathbb{P}(|\hat{p} - p| > 0,02) < 0,1 \quad (36)$$

Делим внутри вероятности на $\sqrt{\text{Var}(\hat{p})}$:

$$\mathbb{P}\left(|N(0; 1)| > 0,02\sqrt{4n}\right) < 0,1 \quad (37)$$

По таблицам получаем $0,02\sqrt{4n} \approx 1,65$ и $n \approx 1691$

Если вместо ЦПТ использовать неравенство Чебышева, то можно получить менее точный результат $n = 6250$.

7. Злобный препод приготовил для группы из 40 человек аж 10 вариантов, по 4 экземпляра каждого варианта. Случайная величина X_1 — номер варианта, доставшийся отличнице Машеньке, величина X_2 — номер варианта, доставшийся двоечнику Вовочке. Величина $\bar{X} = (X_1 + X_2)/2$ — среднее арифметическое этих номеров.

- (a) [4] Найдите $\mathbb{E}(X_1)$, $\text{Var}(X_1)$, $\text{Cov}(X_1, X_2)$
- (b) [3] Найдите $\mathbb{E}(\bar{X})$, $\text{Var}(\bar{X})$
- (c) [2] Являются ли X_1 и X_2 одинаково распределенными?
- (d) [1] Являются ли X_1 и X_2 независимыми?

Подсказка:

$$\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6} \quad (38)$$

Решение:

- (a) $\mathbb{E}(X_i) = (1 + 10)/2 = 5,5$, $\mathbb{E}(X_1^2) = \frac{1}{10} \frac{10 \cdot 11 \cdot 21}{6} = 77/2$, $\text{Var}(X_i) = 33/4 = \sigma^2$. Можно найти $\text{Cov}(X_1, X_2)$ по готовой формуле, но мы пойдем другим путем. Заметим, что сумма номеров всех вариантов — это константа, поэтому $\text{Cov}(X_1, X_1 + \dots + X_{40}) = 0$. Значит $\text{Var}(X_1) + 39\text{Cov}(X_1, X_2) = 0$. В итоге получаем $\text{Cov}(X_1, X_2) = -\frac{1}{39}\sigma^2$
- (b) $\mathbb{E}(\bar{X}) = 11/2$, $\text{Var}(\bar{X}) = 4\frac{1}{52}$
- (c) Да, являются, т.к. и X_1 и X_2 — это номер случайно выбираемого варианта
- (d) Нет, если известно чему равно X_1 , то шансы получить такой же X_2 падают

Часть III Стоимость задачи 20 баллов.

На заводе никто не работает, если хотя бы у одного работника день рождения. Сколько нужно нанять работников, чтобы максимизировать ожидаемое количество рабочих человеко-дней в году?

Решение:

Если мы наняли n работников, то ожидаемое количество рабочих человеко-дней равно:

$$\mathbb{E}(X) = 365 \cdot n \cdot \left(\frac{364}{365}\right)^n \quad (39)$$

Для удобства берем логарифм $\ln(\mathbb{E}(X)) = c + \ln(n) + n \ln(364/365)$ и получаем условие первого порядка $1/n + \ln(364/365) = 0$. Пользуясь разложением в ряд Тейлора $\ln(1+t) \approx t$ получаем: $1/n - 1/365 \approx 0$, $n \approx 365$

8.3 Контрольная работа №3, 13.03.2012

Условия: 80 минут, без официальной шпаргалки.

1. Наблюдения X_1, X_2, \dots, X_n независимы и одинаково распределены с функцией плотности $f(x) = \lambda \exp(-\lambda x)$ при $x \geq 0$.
 - (a) Методом максимального правдоподобия найдите оценку параметра λ
 - (b) Найдите оценку максимального правдоподобия \hat{a} для параметра $a = 1/\lambda$
 - (c) Сформулируйте определение несмещенности оценки и проверьте выполнение данного свойства для оценки \hat{a}
 - (d) Сформулируйте определение состоятельности оценки и проверьте выполнение данного свойства для оценки \hat{a}
 - (e) Сформулируйте определение эффективности оценки и проверьте выполнение данного свойства для оценки \hat{a}
 - (f) Оцените параметр λ методом моментов.

Подсказка: $\mathbb{E}(X_i^2) = 2/\lambda^2$

2. В ходе анкетирования 100 сотрудников банка «Альфа» ответили на вопрос о том, сколько времени они проводят на работе ежедневно. Среднее выборочное оказалось равно 9.5 часам при выборочном стандартном отклонении 0.5 часа.

- (a) Постройте 95% доверительный интервал для математического ожидания времени проводимого сотрудниками на работе
 - (b) Проверьте гипотезу о том, что в среднем люди проводят на работе 10 часов, против альтернативной гипотезы о том, что в среднем люди проводят на работе меньше 10 часов, укажите точное Р-значение.
 - (c) Сформулируйте предпосылки, которые были использованы для проведения теста
3. В ходе анкетирования 20 сотрудников банка «Альфа» ответили на вопрос о том, сколько времени они проводят на работе ежедневно. Среднее выборочное оказалось равно 9,5 часам при стандартном отклонении 0,5 часа. Аналогичные показатели для 25 сотрудников банка «Бета» составили 9,8 и 0,6 часа соответственно.
- (a) Проверьте гипотезу о равенстве дисперсий времени, проводимого на работе, сотрудниками банков «Альфа» и «Бета». Укажите необходимые предпосылки относительно распределения наблюдаемых значений.
 - (b) Проверьте гипотезу о том, что сотрудники банка «Альфа» проводят на работе столько же времени, что и сотрудники банка «Бета». Укажите необходимые предпосылки относительно распределения наблюдаемых значений.

8.4 Экзамен, 26.03.2012

Часть 1.

1. На каждый вопрос предлагается 5 вариантов ответа
2. Ровно один из ответов — верный
3. В графу «Ответ» требуется вписать номер правильного ответа
4. Неправильные ответы не штрафуются.
5. Если Вы считаете, что на вопрос нет правильного ответа или есть несколько правильных ответов, то... возразьтесь! Ибо такой вопрос будет засчитан всем как верный.
6. Было дано 45 минут. Возможно это было много.
7. Удачи!

1. Закон распределение случайной величины задан таблицей

X	-1	0	2
\mathbb{P}	0.4	0.3	0.3

$\mathbb{E}(X^2)$ равняется

1) 0.02	2) 1.6	3) 0.52	4) 0.04	5) 0.4	Ответ:
---------	--------	---------	---------	--------	--------

2. Дисперсия $Var(X)$ считается по формуле

1) $\mathbb{E}(X^2)$ 2) $\mathbb{E}(X^2) + \mathbb{E}^2(X)$ 3) $\mathbb{E}(X^2) - \mathbb{E}^2(X)$ 4) $\mathbb{E}^2(X) - \mathbb{E}(X^2)$ 5) $\mathbb{E}^2(X)$	Ответ:
--	--------

3. Если $f(x)$ — это функция плотности, то $\int_{-\infty}^{+\infty} f(x) dx$ равен

1) 0	2) 1	3) $\mathbb{E}(X)$	4) $Var(X)$	5) $F(x)$	Ответ:
------	------	--------------------	-------------	-----------	--------

4. Если случайная величина X равномерна на отрезке $[1; 5]$ и $F(x)$ — это ее функция распределения, то $F(4)$ равняется

1) 0	2) 0.1	3) 0.2	4) 0.25	5) 0.75	Ответ:
------	--------	--------	---------	---------	--------

5. Условная вероятность $\mathbb{P}(A | B)$ считается по формуле

1) $\frac{\mathbb{P}(A)}{\mathbb{P}(B)}$	2) $\mathbb{P}(A) \cdot \mathbb{P}(B)$	3) $\frac{\mathbb{P}(A \cup B)}{\mathbb{P}(B)}$	4) $\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$	5) $\mathbb{P}(A) - \mathbb{P}(B)$	Ответ:
--	--	---	---	------------------------------------	--------

6. Правильную монетку подбрасывают два раза. Рассмотрим два события: A — при первом броске выпал «орёл», B — «орёл» выпал хотя бы один раз. Найдите $\mathbb{P}(A|B)$

1) 0	2) 1/3	3) 1/2	4) 2/3	5) 1	Ответ:
------	--------	--------	--------	------	--------

7. Известно, что величина X распределена нормально, Y — биномиально, Z — по Пуассону, W — экспоненциально, R — имеет χ^2 распределение. Непрерывными величинами являются

1) все	2) X, Y, Z	3) X, W, R	4) Y, W, R	5) X, R	Ответ:
--------	--------------	--------------	--------------	-----------	--------

8. Известно, что $\mathbb{E}(X) = 3$, $Var(X) = 16$, $\mathbb{E}(Y) = 1$, $Var(Y) = 4$, $\mathbb{E}(XY) = 6$, найдите $Cov(X, Y)$

1) 0	2) 3	3) 4	4) 6	5) нет верного ответа	Ответ:
------	------	------	------	-----------------------	--------

9. Известно, что $\mathbb{E}(X) = 3$, $Var(X) = 16$, $\mathbb{E}(Y) = 1$, $Var(Y) = 4$, $\mathbb{E}(XY) = 6$, найдите $Var(2X - 7)$

1) 16	2) 8	3) 1	4) 9	5) нет верного ответа	Ответ:
-------	------	------	------	-----------------------	--------

10. Если $F(x)$ — это функция распределения, то $\lim_{x \rightarrow +\infty} F(x)$ равен

1) 0	2) 0.5	3) 1	4) $\mathbb{E}(X)$	5) $+\infty$	Ответ:
------	--------	------	--------------------	--------------	--------

11. Если $X \sim N(-3; 25)$, то $\mathbb{P}(2X + 6 > 0)$ равна

1) 0	2) 0.5	3) 1	4) $+\infty$	5) нет верного ответа	Ответ:
------	--------	------	--------------	-----------------------	--------

12. Если $\mathbb{E}(X) = 5$ и $Var(X) = 10$, то, согласно неравенству Чебышева, $\mathbb{P}(|X - 5| \geq 5)$ лежит в интервале

1) $[0; 1]$	2) $[0; 0.4]$	3) $[0.4; 1]$	4) $[0; 0.6]$	5) $[0.6; 1]$	Ответ:
-------------	---------------	---------------	---------------	---------------	--------

13. Если P -значение больше уровня значимости α , то гипотеза H_0 :
 $\mu = \mu_0$

1) отвергается 2) не отвергается 3) отвергается только если $H_a: \mu > \mu_0$ 4) отвергается только если $H_a: \mu < \mu_0$ 5) недостаточно информации	Ответ:
---	--------

14. Функция плотности обязательно является

1) непрерывной	2) непрерывной справа	3) монотонно неубывающей	4) кусочно-постоянной	5) неотрицательной	Ответ:
----------------	-----------------------	--------------------------	-----------------------	--------------------	--------

15. Совместная функция распределения $F(x, y)$ двух случайных величин X и Y это

1) $\mathbb{P}(X \leq x) \cdot \mathbb{P}(Y \leq y)$ 2) $\mathbb{P}(X \leq x \mid Y \leq y)$ 3) $\mathbb{P}(X \leq x, Y \leq y)$ 4) $\mathbb{P}(X \leq x) + \mathbb{P}(Y \leq y)$ 5) $\mathbb{P}(X \leq x) / \mathbb{P}(Y \leq y)$	Ответ:
--	--------

16. Если случайная величина X , имеющая функцию распределения $Q(x)$, и случайная величина Y , имеющая функцию распределения $G(y)$, независимы, то для их совместной функции распределения $F(x, y)$ справедливо

1) $F(x, y) = Q(x) + G(y)$ 2) $F(x, y) = Q(x) / G(y)$ 3) $F(x, y) = Q(x)G(y) / (Q(x) + G(y))$ 4) $F(x, y) = Q(x) \cdot G(y)$ 5) $F(x, y) = \mathbb{E}(Q(X)G(Y))$	Ответ:
--	--------

17. Если X и Y независимые случайные величины, то *неверным* является утверждение:

1) $\mathbb{E}(aX) = a\mathbb{E}(X)$ 2) $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ 3) $\mathbb{E}(c) = c$ 4) $\mathbb{E}(X/Y) = \mathbb{E}(X)/\mathbb{E}(Y)$ 5) $\mathbb{E}(X - Y) = \mathbb{E}(X) - \mathbb{E}(Y)$	Ответ:
--	--------

18. Коэффициент корреляции $\text{corr}(X, Y)$ не обладает свойством

1) $\text{corr}(X, Y) = 0$ для независимых случайных величин X и Y 2) $\text{corr}(X + a, Y + b) = \text{corr}(X, Y)$ 3) $\text{corr}(X, X) = 1$ 4) $\text{corr}(X, 2Y) = 2\text{corr}(X, Y)$ 5) $\text{corr}(X, Y) = \text{corr}(Y, X)$	Ответ:
--	--------

19. Если случайная величина X стандартно нормально распределена, то случайная величина $Z = X^2$ имеет распределение

1) $N(1; 0)$	2) $N(0; 1)$	3) $F_{1,1}$	4) t_2	5) χ_1^2	Ответ:
--------------	--------------	--------------	----------	---------------	--------

20. Если величина X имеет χ_k^2 распределение, величина $Y \sim \chi_n^2$ распределение и они независимы, то их сумма, $X + Y$ имеет распределение

1) $\chi_{\min(k,n)}^2$	2) $\chi_{\max(k,n)}^2$	3) χ_{kn}^2	4) χ_{k+n}^2	5) χ_{k+n-1}^2	Ответ:
-------------------------	-------------------------	------------------	-------------------	---------------------	--------

21. Смещенной оценкой математического ожидания по выборке независимых, одинаково распределенных случайных величин X_1, \dots, X_4 является оценка

1) X_1 2) $0.25 \sum_{i=1}^n X_i$ 3) $0.1X_1 + 0.2X_2 + 0.3X_3 + 0.4X_4$ 4) $0.5X_1 + 0.5X_2 + 0.5X_3 + 0.5X_4$ 5) $X_3 - X_2 + X_1$	Ответ:
--	--------

22. Если X_i независимы и имеют нормальное распределение $N(\mu; \sigma^2)$, то $\sqrt{n}(\bar{X} - \mu)/\hat{\sigma}$ имеет распределение

1) $N(0; 1)$	2) t_{n-1}	3) χ_{n-1}^2	4) $N(\mu; \sigma^2)$	5) нет верного ответа	Ответ:
--------------	--------------	-------------------	-----------------------	-----------------------	--------

23. При построении доверительного интервала для дисперсии по выборке из n наблюдений при неизвестной дисперсии используется статистика, имеющая распределение

1) $N(0; 1)$	2) t_{n-1}	3) χ_{n-1}^2	4) χ_n^2	5) t_n	Ответ:
--------------	--------------	-------------------	---------------	----------	--------

24. Известно, что X_i нормальны $N(\mu; \sigma^2)$ и независимы, $\sum_{i=1}^8 X_i = 32$, $\sum_{i=1}^8 (X_i - \bar{X})^2 = 14$ и $t_{0.01;7} = 3$. Левая граница 98%-го доверительного интервала для μ примерно равна

1) -0.25	2) 0	3) 1	4) 2	5) 2.5	Ответ:
----------	------	------	------	--------	--------

25. Логарифм функции правдоподобия может принимать следующие значения

1) $[0; 1]$	2) $(-\infty; 0]$	3) $(-\infty; +\infty)$	4) $[0; +\infty)$	5) $[-1; 1]$	Ответ:
-------------	-------------------	-------------------------	-------------------	--------------	--------

26. Если X_i независимы, $\mathbb{E}(X_i) = \mu$ и $Var(X_i) = \sigma^2$, то математическое ожидание величины $Y = \sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)$ равно

1) 0	2) 1	3) μ	4) σ^2	5) σ^2/n	Ответ:
------	------	----------	---------------	-----------------	--------

27. Если X_i независимы, $\mathbb{E}(X_i) = \mu$ и $Var(X_i) = \sigma^2$, то дисперсия величины $Y = \sum_{i=1}^n X_i / n$ равна

1) 0	2) 1	3) μ	4) σ^2	5) σ^2/n	Ответ:
------	------	----------	---------------	-----------------	--------

Часть 2.

1. Продолжительность — 2 часа.
2. Можно пользоваться шпаргалкой А4.
3. Имели право участвовать те, кто набрал на тесте удовлетворительно.

1. Снайпер попадает в «яблочко» с вероятностью 0.8, если в предыдущий раз он попал в «яблочко»; и с вероятностью 0.7, если в предыдущий раз он не попал в «яблочко» или если это был первый выстрел. Снайпер стрелял по мишени 3 раза.
 - (a) Какова вероятность попадания в «яблочко» при втором выстреле?
 - (b) Какова вероятность попадания в «яблочко» при втором выстреле, если при первом снайпер попал, а при третьем - промазал?
2. Случайная величина Z равномерно распределена на отрезке $[0; 2\pi]$, $X_1 = \cos(Z)$ и $X_2 = \sin(Z)$. Найдите $E(X_1)$, $E(X_2)$, $Cov(X_1, X_2)$. Являются ли величины X_1 и X_2 независимыми?
3. Театр имеет два различных входа. Около каждого из входов имеется свой гардероб. Эти гардеробы ничем не отличаются. На спектакль приходит 1000 зрителей. Предположим, что зрители приходят по одиночке и выбирают входы равновероятно. Сколько мест должно быть в каждом из гардеробов для того, чтобы в среднем в 99 случаях из 100 все зрители могли раздеться в гардеробе того входа, через который они вошли?
4. Кот Мурзик ловит мышей. Время от одной мышки до другой распределено экспоненциально с функцией плотности $f(x) = \lambda e^{-\lambda x}$ при $x > 0$. На поимку 20 мышей у Мурзика ушло 2 часа.
 - (a) Оцените λ методом максимального правдоподобия
 - (b) Найдите наблюдаемую информацию Фишера, \hat{I} , и оцените дисперсию $\hat{\lambda}$
 - (c) Предположив, что оценка максимального правдоподобия имеет нормальное распределение постройте примерный 95%-ый доверительный интервал для λ
 - (d) С помощью статистики отношения правдоподобия проверьте гипотезу о том, что на одну мышку в среднем уходит 9 минут на 5% уровне значимости

Hint: $\ln(6) \approx 1.79$, $\ln(9) \approx 2.20$

5. Докажите, что из некоррелированности компонент двумерного нормально распределенного случайного вектора следует их независимость.

6. Пусть X_i одинаково распределены и независимы с функцией плотности $f(x, \theta)$. Введем обозначения $i = \mathbb{E} \left(\left(\frac{d \ln f(X_1, \theta)}{d\theta} \right)^2 \right)$ и $I = \mathbb{E} \left(\left(\frac{d \ln L(X_1, \dots, X_n, \theta)}{d\theta} \right)^2 \right)$, где $L(X_1, \dots, X_n, \theta)$ — функция правдоподобия. Докажите, что $I = ni$.
7. Вашему вниманию представлены результаты прыжков в длину Васи Сидорова на двух тренировках. На первой среди болельщиц присутствовала Аня Иванова: 1,83; 1,64; 2,27; 1,78; 1,89; 2,33. На второй Аня среди болельщиц не присутствовала: 1,26; 1,41; 2,05; 1,07; 1,59; 1,96; 1,29.
- С помощью теста Манна-Уитни на уровне значимости 5% проверьте гипотезу о том, что присутствие Ани Ивановой положительно влияет на результаты Васи Сидорова. Можно считать статистику Манна-Уитни нормально распределенной.
8. Вася Сидоров утверждает, что ходит в кино в два раза чаще, чем в спортзал, а в спортзал в два раза чаще, чем в театр. За последние полгода он 10 раз был в театре, 17 раз - в спортзале и 39 раз - в кино. Проверьте гипотезу о том, что имеющиеся данные не противоречат Васиному утверждению на уровне значимости 5%.

9. Известно, что X_i независимы и нормальны, $N(\mu; 900)$. Исследователь проверяет гипотезу $H_0: \mu = 10$ против $H_A: \mu = 30$ по выборке из 20 наблюдений. Критерий выглядит следующим образом: если $\bar{X} > c$, то выбрать H_A , иначе выбрать H_0 .
- (a) Рассчитайте вероятности ошибок первого и второго рода, мощность критерия для $c = 25$.
 - (b) Что произойдет с указанными вероятностями при росте количества наблюдений ($c \in (10; 30)$)?
 - (c) Каким должно быть c , чтобы вероятность ошибки второго рода равнялась 0,15?

И Последняя задача...

10. Пирсон придумал хи-квадрат тест на независимость признаков около 1900 года. При этом он не был уверен в правильном количестве степеней свободы. Он разошелся во мнениях с Фишером. Фишер считал, что для таблицы два на два хи-квадрат статистика будет иметь три степени свободы, а Пирсон — что одну. Чтобы выяснить истину, Фишер взял большое количество таблиц два на два с заведомо независимыми признаками и посчитал среднее значение хи-квадрат статистики.
- (a) Чему оно оказалось равно?
 - (b) Как это помогло определить истину?