

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

Коллектив кафедры
математической экономики и эконометрики,
фольклор

20 декабря 2015 г.

Содержание

1	Описание	3
2	2004-2005	3
2.1	Контрольная работа №2, 22.12.04	3
3	2005-2006	4
3.1	Контрольная работа №1, 18.10.2005	4
3.2	Контрольная работа №2, 21.12.2005	6
3.3	Контрольная работа №3, 04.03.2006	8
4	2006-2007	13
4.1	Контрольная работа №1, ??..11.2006	13
4.2	Контрольная работа №2, 27.01.2007	15
4.3	Контрольная работа №3, 21.02.2007	18
5	2007-2008	22
5.1	Контрольная работа №1, 03.11.2007	22
5.2	Контрольная работа №2, демо-версия, 21.01.2008	25
5.3	Контрольная работа №2, 21.01.2008	26
5.4	Контрольная работа №3, демо-версия, 01.03.2008	30
5.5	Контрольная работа №3, 01.03.2008	32
6	2008-2009	35
6.1	Контрольная работа №1, демо-версия, ??..11.2008	35
6.2	Контрольная работа №1, ??..11.2008	38
6.3	Контрольная работа №2, демо-версия, 26.12.2008	41
6.4	Контрольная работа №2, 26.12.2008	43
6.5	Контрольная работа №3, 02.03.2009	48
7	2009-2010	51
7.1	Контрольная работа №2, ??..12.2009	51
7.2	Контрольная работа №3?, ??..??..2010?	53
7.3	Контрольная работа №4, ??..??..2010	53

8	2010-2011	55
8.1	Контрольная работа №1, ??.10.2010	55
8.2	Контрольная работа №2, ??.12.2010	57
8.3	Контрольная работа №3, ??.03.2011	58
9	2011-2012	59
9.1	Контрольная работа №1, 24.10.2011	59
9.2	Контрольная работа №2, 29.12.2011	63
9.3	Контрольная работа №3, 13.03.2012	68
9.4	Экзамен, 26.03.2012	69
10	2012-2013	77
10.1	Контрольная работа №1, 14.11.2012	77
10.2	Контрольная работа №2, 26.12.2012	79
10.3	Демо-версия зачёта	82
10.4	Зачёт, 15.01.2013	83
10.5	КоКо, компьютерная контрольная №3, 13.03.13	84
10.6	Экзамен, 26.03.2013	85
11	2013-2014	90
11.1	Контрольная работа №1, 5.11.2013	90
11.2	Контрольная работа №1, i-поток, 15.11.2013	92
11.3	Контрольная работа №2, i-поток, 16-28.12.2013	94
11.4	Контрольная работа №2, 25.12.2013	98
11.5	Контрольная работа 3	100
11.6	Контрольная работа 3, i-поток, 19.03.2014.	101
11.7	Переписывание кр1, вариант 1	103
11.8	Переписывание кр1, вариант 2	107
11.9	Билеты к зачёту	110
11.10	Экзамен, 26.03.2014	111
12	2014-2015	117
12.1	Контрольная работа 1. Базовый поток. 30.10.2014	117
12.2	Решение кр 1. Базовый поток	119
12.3	Праздник номер 1, i-поток, 30.10.2014	125
12.4	Праздник номер 1 по теории вероятностей, i-поток. Решение	128
12.5	Контрольная работа 2. Базовый поток. 15.12.2014	140
12.6	Решение. Контрольная работа 2. Базовый поток. 15.12.2014	141
12.7	Праздник номер 2, i-поток, 15.12.2014	146
12.8	Решение. Праздник номер 2, i-поток	148
12.9	Пересдача за 1-ый семестр??	153
12.10	Контрольная номер 3	158
12.11	Контрольная номер 4, 05.06.2015	158
12.12	Экзамен, 15.06.2015	160
13	2015-2016	165
13.1	Контрольная номер 1, базовый поток	165
13.2	Праздник номер 1, исследователи, индивидуальный тур	165
13.3	Регата, исследователи, командный тур	165
13.4	Контрольная номер 2, поток Арктура, 12.12.2015	165
13.5	Контрольная номер 2, поток Риччи, 12.12.2015	166
13.6	Контрольная номер 2, поток Риччи, 12.12.2015, решение	168

1 Описание

Свежую версию можно скачать с блога <http://pokrovka11.wordpress.com/> или с github репозитория <https://github.com/bdemeshev/pr201/wiki>.

Уникальное предложение для студентов факультета экономики ГУ-ВШЭ:

Найдите ошибки в этом документе или пришлите отсутствующие решения в теке и получите дополнительные бонусы к итоговой оценке! Найденные смысловые ошибки поощряются сильнее, чем просто опечатки. Письма с замеченными ошибками и решениями пишите на адрес boris.demeshev@gmail.com. Перед отправкой письма, пожалуйста, сверьтесь со свежей версией этой подборки.

Неполный список благодарностей:

1. Андрей Зубанов, решения (экзамен 26.03.2012, ...)
2. Кирилл Пономарёв, решения (контрольная 1, 2014)
3. Александр Левкун, решения (контрольная 1, 2014)

2 2004-2005

2.1 Контрольная работа №2, 22.12.04

1. [1 балл] Вычислите вероятность $\mathbb{P}(|X - \mathbb{E}(X)| > 2\sqrt{\text{Var}(X)})$, если известно, что случайная величина X подчиняется нормальному закону распределения.
2. [1 балл] Определите значения математического ожидания и дисперсии случайной величины, функция плотности которой имеет вид

$$f(x) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(x+1)^2}{18}}$$

3. [1 балл] Страховая компания «Ой» заключает договор страхования от «невыезда» (невыдачи визы) с туристами, покупающими туры в Европу. Из предыдущей практики известно, что в среднем отказывают в визе одному из 130 человек. Найдите вероятность того, что из 200 застраховавшихся в «Ой» туристов, четверым потребуется страховое возмещение.
4. [1 балл] Считая вероятность рождения мальчика равной 0.52, вычислите вероятность того, что из 24 новорожденных будет 15 мальчиков.
5. [1 балл] Для случайной величины X с нулевым математическим ожиданием дисперсией 16, оцените сверху вероятность $\mathbb{P}(|X| > 15)$.
6. [2 балла] Случайные величины X и Y независимы. Известно, что $\mathbb{E}(X) = 0$, $\text{Var}(X) = 4$, $\mathbb{E}(Y) = 5$. Определите значение дисперсии случайной величины Y , если известно, что случайная величина $Z = 2X - Y$, принимает неотрицательные значения с вероятностью 0.9.
7. [2 балла] Вычислите вероятность $\mathbb{P}(|X - \mathbb{E}(X)| > 2\sqrt{\text{Var}(X)})$, если известно, что случайная величина X распределена по закону Пуассона с параметром $\lambda = 0.09$
8. [2 балла] Портфель страховой компании состоит из 1000 договоров, заключенных 1 января и действующих в течение года. При наступлении страхового случая по каждому из договоров компания обязуется выплатить 1500 рублей. Вероятность наступления страхового события по каждому из договоров предполагается равной 0.05 и не зависящей от наступления

страховых событий по другим контрактам. Каков должен быть совокупный размер резерва страховой компании для того, чтобы с вероятностью 0.95 она могла бы удовлетворить требования, возникающие по указанным договорам?

9. [2 балла] В коробке лежат три купюры, достоинством в 100, 10 и 50 рублей соответственно. Они извлекаются в случайном порядке. Пусть X_1 , X_2 и X_3 — достоинства купюр в порядке их появления из коробки.

- (a) Верно ли, что X_1 и X_3 одинаково распределены?
- (b) Верно ли, что X_1 и X_3 независимы?
- (c) Найдите дисперсию X_2

10. [2 балла] Когда Винни-Пуха не кусают пчелы, он сочиняет в среднем одну кричалку в день. Верный друг и соратник Винни-Пуха Пятачок записал, сколько кричалок сочинялось в дни укусов. Эта выборка из 36 наблюдений перед Вами:

2, 0, 0, 2, 0, 0, 0, 2, 0, 2, 0, 2, 2, 0, 2, 0, 2, 2, 0, 0, 2, 2, 0, 0, 2, 0, 2, 2, 0, 2, 2, 0, 2.

Верно ли, что укусы пчел положительно сказываются на творческом потенциале Винни-Пуха (используйте нормальную аппроксимацию биномиального распределения)?

11. [4 балла] Пусть X_t — количество бактерий, живущих в момент времени t . Известно, что $X_1 = 1$ и $X_t = A_t \cdot X_{t-1}$, где случайные величины A_t независимы и равномерно распределены на отрезке $[0; 2a]$. Величина A_t может интерпретироваться как среднее количество потомков. Можно догадаться, что данная модель приводит к экспоненциальной динамике.

- (a) Определите долгосрочный темп роста бактерий, т.е. найдите предел $\lim_{n \rightarrow \infty} \frac{\ln X_n}{n}$
- (b) При каком a темп роста будет положительным?

3 2005-2006

3.1 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 $\mathbb{E}(X^2) \geq (\mathbb{E}(X))^2$?
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 $\mathbb{P}(A|B) \geq \mathbb{P}(A \cap B)$?
9. Функция плотности не может быть периодической ?
10. Для неотрицательной случайной величины $\mathbb{E}(X) \geq \mathbb{E}(-X)$?
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?

Решение:

$$P(A) = \frac{3!3!}{6!} = 0.05$$

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»?

Ответ:

$$P(A|B) = \frac{1/3}{1/3+2/12} = 2/3$$

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 Васиные абсцисса и ордината в момент остановки. Найдите $\mathbb{P}(X > \frac{1}{2})$, $\mathbb{P}(X > \frac{1}{2} | Y < \frac{1}{2})$. 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?

Ответы:

$$1. \begin{array}{c|ccc} X & 2 & 3 & 4 \\ \hline \mathbb{P}() & 1/6 & 1/3 & 1/2 \end{array}$$

$$2. \mathbb{E}(X) = 3\frac{1}{3}$$

Задача №5

Два охотника выстрелили в одну утку. Первый попадает с вероятностью 0,4, второй - с вероятностью 0,7. В утку попала ровно одна пуля. Какова вероятность того, что утка была убита первым охотником?

$$\text{Ответ: } \mathbb{P}(A|B) = \frac{0.4 \cdot 0.3}{0.4 \cdot 0.3 + 0.6 \cdot 0.7}$$

Задача №6

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

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

Ответы:

$$1. \text{Var}(Z) = 6, \text{Var}(4 - 3Z) = 54, \mathbb{E}(5 + 3Z - Z^2) = -19$$

$$2. \text{Cov}(X, Y) = 2.5, \text{Cov}(6 - X, 3Y) = -7.5$$

Задача №7

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

$$\text{Ответы: } x = -2, \text{Var}(X) = 3.1 - 0.49 = 2.61$$

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 c , $\mathbb{P}(X > 1)$, $\mathbb{E}(X)$, $\mathbb{E}(\frac{1}{X^3+10})$ и постройте график функции распределения величины X .

$$\text{Ответы: } c = 3/16, \mathbb{P}(X > 1) = 13/16, \mathbb{E}(X) = 0, \mathbb{E}(1/(X^3+10)) = \frac{3}{8} \ln(3), F(x) = \begin{cases} 0, & x < -2 \\ \frac{x^3+8}{16}, & x \in [-2; 2] \\ 1, & x > 2 \end{cases}$$

Задача №9

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

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

Ответы: $\mathbb{P}(X = 3 \cap Y = 5) = 2/36$, $\mathbb{E}(X) = 91/36$, $Var(X) \approx 2.1$, заметим, что $X + Y = R_1 + R_2$, поэтому $\mathbb{E}(X) + \mathbb{E}(Y) = 7$, и $\mathbb{E}(3X - 2Y) = 3\mathbb{E}(X) - 2\mathbb{E}(Y) = 3\mathbb{E}(X) - 2(7 - \mathbb{E}(X)) = 5\mathbb{E}(X) - 14$

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 $\mathbb{P}(X > 3)$
б) Найдите $Var(X)$ и $Cov(X, Y)$
в) Найдите $Corr(X, Z)$

Ответы: $\mathbb{P}(X > 3) = 61/1024$, $Var(X) = 15/16$, $Cov(X, Y) = Cov(X, 5 - X) = -Cov(X, X) = -15/16$, $Corr(X, Z) = Corr(X, 2X - 5) = 1$

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 Петину числу 4. Обозначим результат буквой Y .

Найдите $\mathbb{P}(Y < 4)$ 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?

3.2 Контрольная работа №2, 21.12.2005

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

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

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 , $\mathbb{P}(Y > -X)$, $\mathbb{E}(X \cdot Y^2)$, $\mathbb{E}(Y|X > 0)$

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 $\mathbb{P}(X_1 + 3X_2 > 20)$.

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 $\mathbb{P}(Y > X)$, $\mathbb{E}(X)$, $\mathbb{E}(X|Y > X)$

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 $\mathbb{P}(X \in [16; 23])$, если

- X нормально распределена, $\mathbb{E}(X) = 20$, $Var(X) = 25$.
- X равномерно распределена на отрезке $[0; 30]$
- X распределена экспоненциально и $\mathbb{E}(X) = 20$

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 $\mathbb{E}(X_i)$ и $Var(X_i)$. С помощью центральной предельной теоремы найдите вероятность того, что через сто дней акция будет стоить больше 1030 рублей.

Решение: Если S — финальная стоимость акции, то $S = 1000 + X_1 + X_2 + \dots + X_{100}$. Тогда по ЦПТ $S \sim N(1000, 100)$ и $\mathbb{P}(S > 1030) \approx 0.001$.

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?

3.3 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 $\theta = \mathbb{P}(X < 3)$. 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?

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

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

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

$$\begin{aligned} \text{Заметим, что } \hat{\theta}^2 &= \hat{\theta} \\ &= \theta - 2\theta^2 + \theta^2 = \theta - \theta^2 \end{aligned}$$

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

$$\begin{aligned} L &= (k+1)^n (x_1 \cdot x_2 \cdot \dots \cdot x_n)^k \\ l &= \ln L = n \ln(k+1) + k(\sum \ln x_i) \end{aligned}$$

$$\begin{aligned} \frac{dl}{dk} &= \frac{n}{k+1} + \sum \ln x_i \\ \frac{n}{k+1} + \sum \ln x_i &= 0 \end{aligned}$$

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

б) [5] Методом моментов

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

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

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

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.

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

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

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

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 Р-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%?

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

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.

$$\mathbb{P}(-2,13 < t_4 < 2,13) = 0,9$$

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

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

[-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 критерей:

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 .

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

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

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?

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

[-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

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

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

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

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: $\mathbb{E}(X) = \frac{1}{2} \cdot 50 = 25$

Подсказка: Думайте!

4 2006-2007

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

1.2. $\mathbb{P}(A) = p$, $\mathbb{P}(B) = 1 - p^3$, $\mathbb{P}(A \cap B) = p(1 - p^2)$, 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. $\mathbb{P}(X = 1) = C_{10}^1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^9$

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

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

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

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. $\mathbb{P}(B) = 0,96 \cdot 0,96 + 0,04 \cdot 0,05$

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

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).

$$\lambda = np = 4$$

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

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

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
$\mathbb{P}(X = x_i)$	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
$\mathbb{P}(X = x_i)$	0,6	0,2	0,08	0,12

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

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.

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

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

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

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?

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

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

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

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

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

4.2 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., $\mathbb{E}(X) = 6$ 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:

$\mathbb{E}(X)$ - 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 , $\mathbb{P}(Y > -X)$, $\mathbb{E}(X \cdot Y^2)$, $\mathbb{E}(Y|X > 0)$

Решение:

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

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 $\mathbb{P}(X_1 + 3X_2 > 20)$.

Решение:

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

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 $\mathbb{P}(Y > 2X)$, $\mathbb{E}(X)$

Решение:

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

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

(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 $n_i = \frac{w_i \sigma_i}{\sum w_j \sigma_j}$

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

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:

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

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

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:

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

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:

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

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

Задача 8

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

а) Укажите верхнюю границу для $\mathbb{P}(X^2 > 4 \text{Var}(X))$? [5]

б) Найдите указанную вероятность, если дополнительно известно, что X нормально распределена. [5]

Решение:

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

$$\text{б) } \mathbb{P}(X^2 > 4 \text{Var}(X)) = \mathbb{P}(|Z| > 2) = 0.05$$

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:

$$\text{а) } \bar{X}$$

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 $\theta = \frac{1}{\mathbb{P}(X_i < 5)}$.

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:

$$\text{а) } \mathbb{E}(\bar{X}) = \frac{a}{2}, \theta = \frac{1}{\mathbb{P}(X_i < 5)} = \frac{1}{5/a} = \frac{1}{5}a$$

$$\hat{\theta} = \frac{2}{5}\bar{X}$$

$$\text{б) } \text{Var}(\hat{\theta}_n) = \left(\frac{2}{5}\right)^2 \cdot \frac{a^2}{12n}$$

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, равномерно).

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:

$$e_{100} = 0$$

$$e_{99} = 1$$

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

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

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

...

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

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

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:

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

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

$$\text{Разностное уравнение: } p_n = 0.7p_{n-1} + 0.2p_{n-2}$$

Подсказка: Думайте!

4.3 Контрольная работа №3, 21.02.2007

Нужные и ненужные формулы:

T - сумма чего-то там.

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

T - сумма каких-то рангов.

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

T - сумма каких-то рангов.

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

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

УДАЧИ!

Часть I.

Обведите нужный ответ

1. Если $X \sim N(0; 12)$, $Y \sim N(12, 24)$, $\text{Corr}(X, Y) = 0$, 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 отвергает 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 $\mathbb{E}(X^2) \geq (\mathbb{E}(X))^2$. Да. Нет.

10. Если $\text{Corr}(X, Y) > 0$, то $\mathbb{E}(X)\mathbb{E}(Y) < \mathbb{E}(XY)$. 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 X принимает значение (-1), если оба шара - белые; 0, если шары разного цвета и 1, если оба шара черные.

а) Найдите $\mathbb{P}(X = -1)$ [2], $\mathbb{E}(X)$ [3], $\text{Var}(X)$ [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], $\mathbb{P}(0.5 < X < 2)$ [1], 25%-ый квантиль [1]
- б) Найдите $\mathbb{E}(X)$ [2], $\text{Var}(X)$ [2], $\text{Cov}(X, -X)$ [1], $\text{Corr}(2X, 3X)$ [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 $\text{Corr}(X, Y)$ [1], $\text{Corr} = -\frac{1}{3}$
- 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 «Чертово 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, $\mathbb{E}(X_i) = 0$, $\text{Var}(X_i) = \theta$.

- [3] Постройте оценку $\hat{\theta}$ методом максимального правдоподобия.
- Проверьте свойства несмещенности, состоятельности, эффективности у построенной оценки. [каждое свойство по 2, если дано аккуратное определение, то 1]
- [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 критерий:

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

- Дайте определение ошибок первого и второго рода. [2 + 2]
- 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

- [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

- [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 $\mathbb{E}(Y)$

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

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

5 2007-2008

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

Устно: $p = 1/3$, $\mathbb{E}(X) = 1.5$

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

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

Задача 6

Допустим, что закон распределения X имеет вид:

X	1	2	3
Prob	θ	2θ	$1 - 3\theta$

а) Найдите $\mathbb{E}(X)$ б) При каких θ среднее будет наибольшим? При каких - наименьшим?
 $\mathbb{E}(X) = 3 - 4\theta$, $\theta \in [0; 1/3]$, $\theta_{max} = 0$, $\theta_{min} = 1/3$

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 $\mathbb{E}(N)$

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

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:

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

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

$$\mathbb{P}(N = 2) = 5/15$$

$$\mathbb{E}(N) = 28/15, \text{ первая.}$$

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.

$$p = 3/8$$

5.2 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 , $\mathbb{P}(Y > X)$, $\mathbb{E}(X \cdot Y)$, $\mathbb{E}(X|Y > 0)$

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 $\mathbb{P}(X_1 + X_2 > 0)$.

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 , $\mathbb{P}(3Y > X)$, $\mathbb{E}(X)$, $\mathbb{E}(X|Y > 0.5)$

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.

1. Чему равно ожидаемое значение цены через 60 дней? Дисперсия?
2. Какова вероятность того, что через 60 дней цена будет больше 900 рублей?

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

1. $\mathbb{P}(-2\sigma < X - \mu < 2\sigma), X \sim N(\mu; \sigma^2)$
2. $\mathbb{P}(8 < X < 12), X \sim U[0; 20]$
3. $\mathbb{P}(-2 < X - \mathbb{E}(X) < 2), X$ 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!

5.3 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., $\mathbb{E}(X) = 6$ 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:

$\mathbb{E}(X)$ - 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 , $\mathbb{P}(Y > -X)$, $\mathbb{E}(X \cdot Y)$, $Corr(X, Y)$, $\mathbb{E}(Y|X > 0)$

Answers:

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

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

$\mathbb{E}(Y|X > 0) = 0.25$ [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 $\mathbb{P}(X_1 + 3X_2 > 20)$. [5]

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

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

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

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

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

$Var(X_1|X_2 = 0) = 8.19$, $\mathbb{E}(X_1|X_2 = 0) = 1.82$

Есть страшные люди, которые наизусть помнят, что:

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

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

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 $\mathbb{P}(Y > 2X)$, $\mathbb{E}(X)$

Являются ли величины X и Y независимыми?

Решение:

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

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

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 $\mathbb{P}(X = 1 \cap Y = 1) = \mathbb{E}(XY) = \text{Cov}(X, Y) + \mathbb{E}(X)\mathbb{E}(Y)$

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

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

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.

$$\begin{aligned} \text{а) } \mathbb{P}(100 \cdot 1,08^N \cdot 0,95^{64-N} > 110) &= \\ \mathbb{P}(N \ln(1,08) + (64 - N) \ln(0,95) > \ln(1,1)) &= \\ \mathbb{P}\left(N > \frac{\ln(1,1) - 64 \ln(0,95)}{\ln(1,08) - \ln(0,95)}\right) \end{aligned}$$

Заметим, что N - биномиально распределена, примерно $N(64 \cdot \frac{1}{2}, 64 \cdot \frac{1}{4})$

$$Z = \frac{N-32}{4} - \text{стандартная нормальная и } \mathbb{P}(Z > -1,42) = 0,92$$

$$\text{б) } \mathbb{E}(N \ln(1,08) + (100 - N) \ln(0,95))$$

На этот раз $\mathbb{E}(N) = 50$ и $\mathbb{E}(\ln(P_{100})) = 1,28$

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.

1. Какой процент потребителей в среднем обращается за гарантийным ремонтом? [4]
2. Какова вероятность того, что из 1000 потребителей за гарантийным ремонтом обратится более 35% покупателей? [6]

Подсказка: $\exp(5/7) = 2,0427$

Solution:

$$p_{break} = 1 - \exp(-5/7) = 0,51 = \int_0^5 \frac{1}{7} e^{-\frac{t}{7}} dt$$

$$p = 0,8 \cdot 0,51 \approx 0,4$$

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

$$\mathbb{P}(S > 350) = \mathbb{P}(Z > -3,23) \approx 1$$

Задача 8

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

а) Укажите верхнюю границу для $\mathbb{P}(X^2 > 2,56 \cdot \text{Var}(X))$?

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

$$\text{а) } \mathbb{P}(X^2 > 2,56 \text{Var}(X)) = \mathbb{P}(|X - 0| > 1,6\sigma) \leq \frac{\text{Var}X}{2,56 \text{Var}(X)} = \frac{100}{256} \approx 0,4 \text{ [5]}$$

$$\text{б) } \mathbb{P}(X^2 > 2,56 \text{Var}(X)) = \mathbb{P}(|Z| > 1,6) = 0,11 \text{ [5]}$$

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

$$f(a, b) = \int_0^1 x^a (1-x)^b dx$$

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

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

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

Если кто получит этот ответ другим (более интуитивным) образом - тому большой дополнительный балл (!) - обращайтесь на boris.demeshev@gmail.com

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 $\mathbb{E}(X)$ и $\mathbb{E}(Y)$

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

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:

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

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

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

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

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

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

Поехали:

$$\text{Заметим, что } X_i = X_i + M_i F_i = M_i F$$

$$\text{Таким образом } X = M F = n F - F^2$$

$$Y_i = F - F_i - X_i$$

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

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

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

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

Подсказка: Думайте!

5.4 Контрольная работа №3, демо-версия, 01.03.2008

Демо-версия кр3!

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

Напоминалка: 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.

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

б) 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 $\hat{a}_n = \min\{X_1, X_2, \dots, X_n\}$.

а) Является ли предлагаемая оценка состоятельной?

б) Является ли предлагаемая оценка несмещенной?

Solution of 8:

Заметим, что $\hat{a}_n \geq a$.

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

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

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

б) 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 $\hat{b} = c \cdot (|X_1| + |X_2|)$.

а) При каком c оценка будет несмещенной?

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

5.5 Контрольная работа №3, 01.03.2008

Часть I. Обведите верный ответ:

1. Мощность теста можно рассчитать заранее, до проведения теста. Да.
2. Точное P -значение можно рассчитать заранее, до проведения теста. Нет.
3. Если гипотеза отвергает при 5%-ом уровне значимости, то она обязательно будет отвергаться и при 10%-ом уровне значимости. Да.
4. Мощность больше у того теста, у которого вероятность ошибки 1-го рода меньше. Нет.
5. Функция плотности F -распределения $p(t)$ не определена при $t < 0$. Нет.
6. При большом k случайную величину, имеющую χ_k^2 распределение, можно считать нормально распределенной. Да.
7. 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.

бия. 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 ($Z_{observed} = -2.12$, $Z_{critical} = -1.28$)

в) $P_{value} = 0.017$

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 таблицей:

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) $\hat{a} = \frac{5 - \bar{X}}{10}$

b) 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:

$$\chi_{observed}^2 = 1.15$$

$$\chi_{2,5\%}^2 = 5.99$$

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.

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

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

Solution:

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

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

Задача 8

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

a) Является ли предлагаемая оценка состоятельной? [8]

b) Является ли предлагаемая оценка несмещенной? [2]

Solution:

Заметим, что $\hat{a}_n \leq a$.

$$\begin{aligned}\mathbb{P}(|\hat{a}_n - a| > \varepsilon) &= \mathbb{P}(-(\hat{a}_n - a) > \varepsilon) = \mathbb{P}(\hat{a}_n < a - \varepsilon) = \mathbb{P}(\max\{X_1, X_2, \dots, X_n\} < a - \varepsilon) = \\ &= \mathbb{P}(X_1 < a - \varepsilon \cap X_2 < a - \varepsilon \cap \dots) = \mathbb{P}(X_1 < a - \varepsilon) \cdot \mathbb{P}(X_2 < a - \varepsilon) \cdot \dots = (1 - \varepsilon)^n \\ \lim_{n \rightarrow \infty} (1 - \varepsilon)^n &= 0\end{aligned}$$

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.

http://en.wikipedia.org/wiki/Simpson's_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}

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

$$\text{Несмещенность: } \mathbb{E}(\hat{x}) = a\mathbb{E}(X) + b\mathbb{E}(Y) + c\mathbb{E}(Z) = ax + by + c(x + y) = x$$

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

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

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.

6 2008-2009

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

Часть I. Обведите верный ответ:

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

Ответ: $\mathbb{P}(A) = 2/15$

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 $\mathbb{P}(X < 1)$, $\mathbb{E}(X)$

Ответ: $\mathbb{P}(X < 1) = 1/3$, $\mathbb{E}(X) = 1.5$

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?

Ответы: $c = 0.1$, $\mathbb{P}(X > 30) = e^{-3}$, $\mathbb{P}(X > 45 | X > 15) = e^{-3}$, $\mathbb{E}(X) = 10$

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

Ответы: $k^* = 2$, $\text{Var}(X) = 1.875$, $\mathbb{E}(X) = 2.5$

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 , $\mathbb{P}(Y > -X)$, $\mathbb{E}(X \cdot Y)$, $\text{Corr}(X, Y)$, $\mathbb{E}(Y | X > 0)$

Ответ: $c = 0.2$, $\mathbb{P}(Y > -X) = 0.5$, $\mathbb{E}(XY) = 0$, $\text{Corr}(X, Y) = -0.155$, $\mathbb{E}(Y | X > 0) = 1/4$

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 $\mathbb{E}(N)$

Ответ: $\mathbb{E}(N) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \mathbb{E}(X_3) = 0.9 + 0.7 + 0.5 = 2.1$

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

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

Ответы: $\mathbb{E}(X) \approx 1.7$, $\text{Var}(X) \approx 1.08$

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?

6.2 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 $\text{Var}(X) = 2008$. Неверно.
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) $\mathbb{E}(X)$ [2], $\text{Var}(X)$ [2]

б) $\mathbb{E}(X - Y)$ [3], $\text{Var}(X - Y)$ [3]

Решение:

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

Отсюда: $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 0$, $\text{Var}(X) = 16$, $\text{Var}(Y) = 9$

Задача 5

Закон распределения пары случайных величин X и Y задан табличкой:

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

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

Каждая величина по [2] очка.

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

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 $\mathbb{P}(Y < 2)$

б) Постройте функцию плотности для величины Y

в) Найдите $\mathbb{E}(Y)$, $\text{Var}(Y)$

г) Найдите $\text{Cov}(X, Y)$

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) $\mathbb{P}(Y < 2) = 1/4$ [1]

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

д) $\text{Cov}(X, Y) = 3$.(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 «уголков», возможно пересекающихся, видимых с Дурасовского переулка.

Найдите $\mathbb{E}(X)$ и $\text{Var}(X)$

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».

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

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.

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

6.3 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 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 $\frac{1}{10}$. 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 $\text{Var}(X) = 9$, 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

$\text{Var}(X)$ - 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 $\mathbb{P}(Y > 2X)$, $\mathbb{E}(X)$

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

$$\begin{pmatrix} 2 \\ -1 \end{pmatrix} \text{ и ковариационной матрицей } \begin{pmatrix} 9 & -4,5 \\ -4,5 & 25 \end{pmatrix}.$$

а) Найдите $\mathbb{P}(X_1 + 3X_2 > 20)$. [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 $\mathbb{P}(\frac{(X-30)^2}{\text{Var}(X)} < 3)$, если известно, что $\mathbb{E}(X) = 30$. 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 $\mathbb{P}(\sum (X_i - \bar{X})^2 > a\sigma^2) = 0.1$.

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 $\mathbb{E}(Y)$.

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!

6.4 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 $\mathbb{P}(X = Y) = 0.9999$, 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 $\mathbb{P}(Y > X)$, $\mathbb{E}(Y)$ [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 $\mathbb{E}(X_i) = 0$, $\text{Var}(X_i) = 1$ и $\text{Var}(X_i^2) = 2$

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

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

Sol:

а) [2 + 2]

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

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

б) [3 + 3]

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

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

Задача 3

Случайный вектор $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ имеет нормальное распределение с математическим ожиданием

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ и ковариационной матрицей } \begin{pmatrix} 9 & -5 \\ -5 & 25 \end{pmatrix}.$$

а) Найдите $\mathbb{P}(X_1 + 2X_2 > 20)$. [5]

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

Sol:

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

$$\text{Var}(X_1 + 2X_2) = \text{Var}(X_1) + 4 \text{Var}(X_2) + 4 \text{Cov}(X_1, X_2) = 89$$

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 $\beta = \mathbb{E}(X^2)$.

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

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

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

$$\mathbb{E}(Y^2) = 3a^2$$

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

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

Т.к. обе оценки несмещенные вместо сравнения дисперсий можно сравнить квадраты ожиданий $\frac{16}{81}\mathbb{E}(X^2Y^2)$ vs $\frac{1}{81}\mathbb{E}(Y^4)$

...

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

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: $3\lambda = 6$.

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

Ответ с факториалам считается полным.

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

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

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

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) $\lambda = 1/10, \mathbb{P}(X < 7) = 0.5$

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

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)

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

$Y - Z$ - это сколько правильных ответов дал лично Вася и оно не зависит от числа Z правильных списанных ответов, значит $\text{Cov}(Y - Z, Z) = 0$

Аналогично все остальные ковариации равны нулю.

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

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

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

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 $s(n) = s$.

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

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 Микоса. 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 выбрать 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. Заменим 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!

6.5 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-A Стоимость задач 10 баллов. Теория вероятностей.

Нужно решить любые **3 (три)** задачи из части II-A.

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?

Задача 2

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

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?

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 ?

Задача 5

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

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

Часть II-В Стоимость задач 10 баллов. Построение и свойства оценок.

Нужно решить любые 2 (две) задачи из части II-В.

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.

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?

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?

Часть II-С Стоимость задач 10 баллов. Проверка гипотез и доверительные интервалы.

Нужно решить любые **3 (три)** задачи из части II-С.

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.

Задача 10

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

- а) Проверить гипотезу о том, что мальчики и девочки рождаются одинаково часто против гипотезы о том, что мальчиков рождается больше, чем девочек
- б) Вычислить р-значение (минимальный уровень значимости, при котором основная гипотеза отвергается)
- в) 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?

Задача 11

Проверьте гипотезу о независимости пола респондента и предпочитаемого им сока.

	Апельсиновый	Томатный	Вишневый
Мужчины	70	40	25
Женщины	75	60	35

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).

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 из части III.

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?

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 Васины затраты при использовании оптимальной стратегии?

в) Сколько магазинов в среднем будет посещено?

Подсказка: Думайте!

7 2009-2010

7.1 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.

Занятые 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»,

<http://folklor.kulichki.net/texts/vektor.html>

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.

Решение:

Из условия: $\text{Var}(X) = 5^2 = 25$, $\text{Var}(Y) = 4^2 = 16$, $\text{Var}(X - Y) = 2^2 = 4$. Есть такое тождество, $\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y)$. Отсюда находим $\text{Cov}(X, Y) = 37/2$ и $\text{Corr}(X, Y) = 37/40$.

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 $\text{Var}(X) = 12$.
 - (a) С помощью неравенства Чебышева оцените в каких пределах лежит вероятность $\mathbb{P}(|X - 16| > 4)$
 - (b) Найдите вероятность $\mathbb{P}(|X - 16| > 4)$, если известно, что X равномерна на $[10; 22]$
 - (c) 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 $\text{Var}(X_i)$. С помощью центральной предельной теоремы найдите вероятность того, что через сто дней акция будет стоить больше 1010 рублей.

Решение: Если S — финальная стоимость акции, то $S = 1000 + X_1 + X_2 + \dots + X_{100}$. Тогда по ЦПТ $S \sim N(1000, 100)$ и $\mathbb{P}(S > 1010) = \mathbb{P}(Z > 1)$.

Дополнительная задача:

Вася и Петя подбрасывают несимметричную монету. Вероятность выпадения «орла» $p = 0,25$. 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.

7.2 Контрольная работа №3?, ????.2010?

1. Имеются наблюдения $-1.5, 2.6, 1.2, -2.1, 0.1, 0.9$. Найдите выборочное среднее, выборочную дисперсию. Постройте эмпирическую функцию распределения.
2. Известно, что в урне всего n_t шаров. Часть этих шаров — белые. Количество белых шаров, n_w , неизвестно. Мы извлекаем из урны n шаров без возвращения. Количество белых шаров в выборке, X , — это случайная величина и $\nu = X/n$. Найдите $\mathbb{E}(\nu)$, $\text{Var}(\nu)$. Будет ли ν состоятельной оценкой неизвестной доли $p = n_w/n_t$ белых шаров в выборке? Будет ли оценка ν несмещенной? Дайте определение несмещенной оценки.
3. Стоимость выборочного исследования генеральной совокупности, состоящей из трёх страт определяется по формуле $TC = 150n_1 + 40n_2 + 15n_3$, где n_i — количество наблюдений в выборке, относящихся к i -ой страте. Стоимость исследования фиксирована. Цель исследования — получить несмещенную оценку среднего по генеральной совокупности с наименьшей дисперсией. Сколько наблюдений нужно взять из каждой страты, если:

Страта	1	2	3
Стандартная ошибка	50	20	10
Вес	10%	30%	60%
Цена наблюдения	150	40	15

4. По выборке X_1, X_2, \dots, X_n найдите методом моментов оценку параметра θ равномерного распределения $U[0; \theta]$. Является ли она несмещенной? Является ли она состоятельной? Какая оценка эффективнее, оценка метода моментов или оценка $T = \frac{n+1}{n} \max\{X_1, \dots, X_n\}$?
5. Неправильная монетка подбрасывается n раз. Количество выпавших орлов — случайная величина X . Найдите оценку вероятности выпадения орла. Проверьте несмещенность, состоятельность и эффективность этой оценки.
6. «Насяльника» отправил Равшана и Джемшуда измерить ширину и длину земельного участка. Равшан и Джемшуд для надежности измеряют длину и ширину 100 раз. Равшан меряет длину, результат каждого измерения — случайная величина $X_i = a + e_i$, где a — истинная длина участка, а $e_i \sim N(0, 1)$ — ошибка измерения. Джемшуд меряет ширину, результат каждого измерения — случайная величина $Y_i = b + u_i$, где b — истинная ширина участка, а $u_i \sim N(0, 1)$ — ошибка измерения. Все ошибки независимы. Думая, что «насяльника» хочет измерить площадь участка, Равшан и Джемшуд каждый раз сообщают «насяльнику» только величину $X_i Y_i$. Помогите «насяльнику» оценить параметры a и b по отдельности методом моментов.

7.3 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.
 - (а) 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 c 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.
 - (a) Согласуются ли полученные данные с предположением страховой компании? (Альтернатива: вероятность страхового случая больше)
 - (b) Определить минимальный уровень значимости, при котором основная гипотеза отвергается (p-value).

8 2010-2011

8.1 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 $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. 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 $\text{Corr}(X, Y) = 100$. 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 $\mathbb{E}(X) = 0$. Да. Понятия не имею.

Задачи.

1. В жюри три человека, они должны одобрить или не одобрить конкурсанта. Два члена жюри независимо друг от друга одобряют конкурсанта с одинаковой вероятностью p . Третий член жюри для вынесения решения бросает правильную монету. Окончательное решение выносится большинством голосов. С какой вероятностью жюри одобрит конкурсанта? Что предпочтёт конкурсант: чтобы решение принимало данное жюри, или чтобы решение принимал один человек, одобряющий с вероятностью p ?

Ответы: p , всё равно

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?

Ответы: $\mathbb{P}(A) = 0.8$, $\mathbb{P}(B|A) = 0.84$

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?

Ответ: $a \geq 2 \ln 10$

4. Правильный кубик подбрасывают до тех пор, пока накопленная сумма очков не достигнет 3 очков или больше. Пусть X — число потребовавшихся подбрасываний кубика. Постройте функцию распределения величины X и найдите $\mathbb{E}(X)$ и $\text{Var}(X)$.

Ответ: $\mathbb{E}(X) = 1.36$, $\text{Var}(X) = 0.2$, $F(x) = \begin{cases} 0, & x < 1 \\ 2/3, & x \in [1; 2) \\ 35/36, & x \in [2; 3) \\ 1, & x \geq 3 \end{cases}$

5. Тест по теории вероятностей состоит из 10 вопросов, на каждый из которых предлагается 3 варианта ответа. Васе удастся списать ответы на первые 5 вопросов у отличника Лёни, который никогда не ошибается, а на оставшиеся 5 он вынужден отвечать наугад. Оценка за тест, величина X — число правильных ответов. Оценка «отлично» начинается с 8 баллов, «хорошо» — с 6, «зачёт» — с 4-х.

(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.

Ответ: $\text{Var}(X) = 1.05$, $\mathbb{E}(X) = 6.5$, $P(A) = 0.3^5$; $Y = 5 + V - (5 - V) = 2V$, $\text{Cov}(X, Y) = \text{Cov}(5 + V, 2V) = 2 \text{Var}(V) = 2.1$

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

Ответ: $\mathbb{E}(X) = 0.5$, $\mathbb{E}(Y) = 1.5$, $\text{Var}(X) = 1.65$, $\text{Cov}(X, Y) = 0.05$, $\text{Cov}(2X + 3, -3Y + 1) = -0.3$

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

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

Ответы: $\mathbb{E}(X_1) = 1$, $\text{Var}(X_1) = 1/3$, $\text{Med}(X_1) = 1$, $f(x, y) = \begin{cases} \frac{1}{4}, & x_1 \in [0; 2], x_2 \in [1, 3] \\ 0, & \end{cases}$

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

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

8.3 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 $\text{Corr}(X; Y) = 0,4$, 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: $\sum_i X_i = 540$, $\sum_i Y_i = 480$, $\sum_i \frac{X_i^2}{15} = 1706,264$, $\sum_i \frac{Y_i^2}{12} = 1958,3$, $\sum_i \frac{(X_i - 36)^2}{15} = 410,264$, $\sum_i \frac{(Y_i - 40)^2}{12} = 358,3$, $\sum_i \frac{(X_i - 40)^2}{15} = 426,264$, $\sum_i \frac{(Y_i - 36)^2}{12} = 374,3$.

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 $\mathbb{N}(\mu_1; \mu_2 - \mu_1^2)$ 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 $\mathbb{N}(\theta; 1)$ методом максимального правдоподобия оценить параметр θ . Будет ли найденная оценка эффективной? Ответ обосновать.

9 2011-2012

9.1 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 $\mathbb{P}(A) = 0.7$ и $\mathbb{P}(B) = 0.5$, 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.

- 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».

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

- 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.

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

- 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]

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

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

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

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]

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

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

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

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

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

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 $\text{Var}(X)$ 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 $\text{Var}(X) = 3/4$.

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]

$$\mathbb{E}(X) = 3/2, \mathbb{E}(Y) = 3/4, \mathbb{E}(XY) = \mathbb{E}(1) = 1, \text{ значит } \text{Cov}(X, Y) = 1 - 9/8 = -1/8 \\ \mathbb{E}(Y^2) = 3/5, \text{Var}(Y) = 3/80, \text{Corr}(X, Y) = -\sqrt{5}/3 \approx 0,75$$

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: $\mathbb{E}(X) = 1, \mathbb{E}(Y) = 4, \mathbb{E}(XY) = 8, \text{Var}(X) = \text{Var}(Y) = 9$. Для случайных величин $U = X + Y$ и $V = X - Y$ вычислите:

(a) $\mathbb{E}(U), \text{Var}(U), \mathbb{E}(V), \text{Var}(V), \text{Cov}(U, V)$

$$\mathbb{E}(U) = 5 [1 \text{ pt}], \mathbb{E}(V) = -3 [1 \text{ pt}], \text{Var}(U) = 26 [2 \text{ pts}], \text{Var}(V) = 10 [2 \text{ pts}], \text{Cov}(U, V) = 0 [2 \text{ pts}]$$

(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 $\text{Var}(X)$ [3 points]

$$\mathbb{E}(X) = 80 \cdot 0,05 = 4 \quad (15)$$

$$\text{Var}(X) = 80 \cdot 0,05 \cdot 0,95 = 4 \cdot 0,95 \quad (16)$$

(b) Найдите точную вероятность $\mathbb{P}(X = 5)$ [3 points]

$$\mathbb{P}(X = 5) = C_{80}^5 0,05^5 0,95^{75} \quad (17)$$

(с) Найдите вероятность $\mathbb{P}(X = 5)$, используя пуассоновскую аппроксимацию. [3 points]

$$\mathbb{P}(X = 5) \approx \exp(-4)4^5/5! \quad (18)$$

(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]

$$\mathbb{E}(X) = 50 \quad (23)$$

(с) 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?

9.2 Контрольная работа №2, 29.12.2011

Разрешается использование калькулятора.

При себе можно иметь шпаргалку А4.

Обозначения:

$\mathbb{P}(A)$ – вероятность A

$E(X)$ – математическое ожидание

$\text{Var}(X)$ – 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 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 $\text{Cov}(X, Y)$
- (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) $\mathbb{E}(X) = 13/20 = 0,65$, $\mathbb{E}(XY) = 2/5 = 0,4$, $\text{Cov}(X, Y) = -9/400 = -0,0225$
 - (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, $MSE = \mathbb{E}((\hat{b} - b)^2)$?

Ответы:

- (a) Заметим, что величина $|X_i|$ распределена равномерно на $[0; b]$, поэтому $\mathbb{E}(|X_i|) = b/2$ и $\text{Var}(|X_i|) = b^2/12$. Значит $\mathbb{E}(\hat{b}) = cb$ и для несмещенности $c = 1$.
- (b) Находим MSE через b и c :

$$MSE = \text{Var}(\hat{b}) + (\mathbb{E}(\hat{b}) - b)^2 = 2c^2 \cdot \frac{b^2}{12} + (c - 1)^2 \cdot b^2 = b^2 \left(\frac{7}{6}c^2 - 2c + 1 \right) \quad (26)$$

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, $\mathbb{E}(S) = 27$, $\text{Var}(S) = 27$, поэтому

$$\mathbb{P}(S \in [20; 34]) = \mathbb{P}(|S - \mathbb{E}(S)| \leq 7) \geq 1 - \frac{27}{7^2} = \frac{22}{49} \quad (28)$$

- (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

$$\mathbb{P}(S \in [430; 470]) \approx \mathbb{P}(N(0; 1) \in [-0,94; +0,94]) \approx 0,6528 \quad (30)$$

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 $\text{Var}(Y) = 1 + 9 - 2 = 8$, 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 $\text{Cov}(X_1, Y) = 1 - 1 = 0$. 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:

$$\begin{cases} \text{Var}(X_1) = 1 \\ \text{Var}(X_2) = 9 \\ \text{Cov}(X_1, X_2) = -1 \end{cases} \Leftrightarrow \begin{cases} a^2 + b^2 = 1 \\ c^2 = 9 \\ bc = -1 \end{cases} \quad (34)$$

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 $\text{Var}(\hat{p})$?
 (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) $\text{Var}(\hat{p}) = \frac{p(1-p)}{n}$. Максимально возможное значение $p(1-p)$ равно 1/4, поэтому максимально возможное значение $\text{Var}(\hat{p}) = 1/4n$.

- (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: $1/n - 1/365 \approx 0$, $n \approx 365$

9.3 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.

9.4 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... возрадуйтесь! 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 табличкой

X	-1	0	2
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 $\text{Var}(X)$ 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) $\text{Var}(X)$	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 $\mathbb{E}(X) = 3$, $\text{Var}(X) = 16$, $\mathbb{E}(Y) = 1$, $\text{Var}(Y) = 4$, $\mathbb{E}(XY) = 6$, найдите $\text{Cov}(X, Y)$

1) 0	2) 3	3) 4	4) 6	5) нет верного ответа	Ответ:
------	------	------	------	-----------------------	--------

9. Известно, что $\mathbb{E}(X) = 3$, $\text{Var}(X) = 16$, $\mathbb{E}(Y) = 1$, $\text{Var}(Y) = 4$, $\mathbb{E}(XY) = 6$, найдите $\text{Var}(2X - 7)$

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 $\text{Var}(X) = 10$, 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 — χ_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 неизвестном ожидании 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 $\text{Var}(X_i) = \sigma^2$, 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 $\text{Var}(X_i) = \sigma^2$, 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:
------	------	----------	---------------	-----------------	--------

Ответы: (поместить в строчку)

1. 2)
2. 3)
3. 2)
4. 5)
5. 4)
6. 4)
7. 3)
8. 2)
9. 5)
10. 3)
11. 2)
12. 2)
13. 2)
14. 5)
15. 3)
16. 4)
17. 4)
18. 4)
19. 5)
20. 4)
21. 4)
22. 2)
23. 3)

24. 5)

25. 3)

26. 4)

27. 5)

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.
 - (а) Какова вероятность попадания в «яблочко» при втором выстреле?
 - (б) 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 $\mathbb{E}(X_1)$, $\mathbb{E}(X_2)$, $\text{Cov}(X_1, X_2)$. 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.
 - (а) Оцените λ методом максимального правдоподобия
 - (б) Найдите наблюдаемую информацию Фишера, \hat{I} , и оцените дисперсию $\hat{\lambda}$
 - (с) Предположив, что оценка максимального правдоподобия имеет нормальное распределение постройте примерный 95%-ый доверительный интервал для λ
 - (д) 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 $I_1 = \mathbb{E} \left(\left(\frac{\partial \ln f(X_1, \theta)}{\partial \theta} \right)^2 \right)$ и $I_n = \mathbb{E} \left(\left(\frac{\partial \ln L(X_1, \dots, X_n, \theta)}{\partial \theta} \right)^2 \right)$, где $L(x_1, \dots, x_n, \theta)$ — функция правдоподобия. Как связаны между собой I_n и I_1 ?

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?

Часть решений (выверить):

1.

$$\mathbb{P}(A_2) = 0,7 \cdot 0,8 + 0,3 \cdot 0,7 = 0,56 + 0,21 = 0,77.$$

$$\mathbb{P}(A_2 \mid A_1 \cap A_3^c) = \frac{0,7 \cdot 0,8 \cdot 0,2}{0,7 \cdot 0,8 \cdot 0,2 + 0,7 \cdot 0,2 \cdot 0,3} = \frac{0,16}{0,22} = \frac{8}{11}.$$

2. Если $Y = y(X)$, то $\mathbb{E}(Y) = \int_{-\infty}^{+\infty} y(x) f_X(x) dx$, поэтому:

$$\mathbb{E}(X_1) = \int_0^{2\pi} \cos z \frac{1}{2\pi} dz = 0$$

$$\mathbb{E}(X_2) = \int_0^{2\pi} \sin z \frac{1}{2\pi} dz = -\frac{\cos z}{2\pi} \Big|_0^{2\pi} = 0$$

$$\begin{aligned}\text{Cov}(X_1; X_2) &= \mathbb{E}(X_1 X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2) = \mathbb{E}(X_1 X_2) = \\ &= \frac{1}{2\pi} \int_0^{2\pi} \sin z \cos z dz = \frac{1}{4\pi} \int_0^{2\pi} \sin 2z dz = -\frac{1}{4\pi} \cos 2z \Big|_0^{2\pi} = 0\end{aligned}$$

Случайные величины зависимы, так как $\sin^2 Z + \cos^2 Z = 1$. Если, например, $\sin Z = 1$, то не может оказаться, что $\cos Z = 1/2$.

3. Пусть X_i — случайная величина, которая равна 1, если посетитель i выбрал первый вход и 0, — если второй. $X_i \sim \text{Bi}(1; p)$. Тогда $\bar{X} = \sum_{i=1}^{1000} X_i / 1000$ — доля посетителей, вошедших через первый вход. По условию,

$$E(X_i) = \frac{1}{2}. \quad \sigma = \sqrt{\frac{\frac{1}{2} \cdot \frac{1}{2}}{1000}} = \frac{1}{20\sqrt{10}}.$$

Найдем такое k , что $\mathbb{P}(\bar{X} < k) > 0,99$:

$$\mathbb{P}\left(\frac{\bar{X} - 1/2}{\frac{1}{20\sqrt{10}}} < \frac{k - 1/2}{\frac{1}{20\sqrt{10}}}\right) > 0,99$$

$$\mathbb{P}(Z < 10\sqrt{10}(2k - 1)) > 0,99$$

$$10\sqrt{10}(2k - 1) > 2,33$$

$$k > 0,536841$$

Аналогичную долю получаем и для второго гардероба.

Наименьшее необходимое число мест в гардеробе будет равно $\lceil 1000k \rceil = \lceil 536,841 \rceil = 537$

4. (a)

$$L = \prod_{i=1}^n f_i(x) = \lambda^n e^{-\lambda \sum_{i=1}^n X_i} = \lambda^n e^{-\lambda n \bar{X}}$$

$$\ln L = n \ln \lambda - \lambda n \bar{X}$$

$$(\ln L)' = n/\lambda - n\bar{X}$$

$$(\ln L)'' = -n/\lambda^2 < 0$$

$$\hat{\lambda} = \frac{1}{\bar{X}}$$

(b) Ожидаемая информация Фишера: $I(x; \lambda) = -\mathbb{E}\left(\frac{\partial^2 \ln L}{\partial \lambda^2}\right) = \frac{n}{\lambda^2}$

Граница Крамера-Рао $\text{Var}(\hat{\lambda}) \geq \frac{1}{I} = \frac{\lambda^2}{n}$

(c) В нашем случае $\mathbb{E}(\hat{\lambda}) = \lambda$ и $\text{Var}(\hat{\lambda}) = \frac{\lambda^2}{n}$. Условие для нахождения доверительного интервала:

$$z_{0,025} < \frac{\hat{\lambda} - \lambda}{\sqrt{\text{Var}(\hat{\lambda})}} < z_{0,975}$$

Доверительный интервал:

5. Пользуясь некоррелированностью, получаем:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-0,5\left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right)} = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-0,5\frac{(x-\mu_x)^2}{\sigma_x^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_y} e^{-0,5\frac{(y-\mu_y)^2}{\sigma_y^2}} = f(x) \cdot f(y),$$

Значит величины X и Y независимы, так как $f(x, y) = f(x) \cdot f(y)$.

6.

$$I_n = \mathbb{E} \left(\left(\frac{d \ln L(X_1; X_2; \dots; X_n; \lambda)}{d\lambda} \right)^2 \right) = E \left(\left(\frac{\sum_{i=1}^n d \ln f(X_i; \lambda)}{d\lambda} \right)^2 \right)$$

Все X_i одинаково распределены, поэтому:

$$\mathbb{E} \left(\left(\frac{d \ln f(X_1; \lambda)}{d\lambda} \right)^2 \right) = \mathbb{E} \left(\left(\frac{d \ln f(X_i; \lambda)}{d\lambda} \right)^2 \right)$$

Стало быть,

$$\begin{aligned} I_n &= \mathbb{E} \left(\left(\frac{d \ln L(X_1; X_2; \dots; X_n; \lambda)}{d\lambda} \right)^2 \right) = E \left(\left(\frac{d \sum_{i=1}^n \ln f(X_i; \lambda)}{d\lambda} \right)^2 \right) = \\ &= \sum_{i=1}^n \mathbb{E} \left(\left(\frac{d \ln f(X_i; \lambda)}{d\lambda} \right)^2 \right) = n \mathbb{E} \left(\left(\frac{d \ln f(X_i; \lambda)}{d\lambda} \right)^2 \right) = n I_1 \quad (40) \end{aligned}$$

10 2012-2013

10.1 Контрольная работа №1, 14.11.2012

14 ноября 1936 года в СССР была создана Гидрометеорологическая служба.

1. Погода завтра может быть ясной с вероятностью 0.3 и пасмурной с вероятностью 0.7. Вне зависимости от того, какая будет погода, Маша даёт верный прогноз с вероятностью 0.8. Вовочка, не разбираясь в погоде, делает свой прогноз по принципу: с вероятностью 0.9 копирует Машин прогноз, и с вероятностью 0.1 меняет его на противоположный.

- (a) Какова вероятность того, что Маша спрогнозирует ясный день?
- (b) Какова вероятность того, что Машин и Вовочкин прогнозы совпадут?
- (c) Какова вероятность того, что день будет ясный, если Маша спрогнозировала ясный?
- (d) Какова вероятность того, что день будет ясный, если Вовочка спрогнозировал ясный?

Ответы:

- (a) $\mathbb{P}(A) = 0.8 \cdot 0.3 + 0.7 \cdot 0.2 = 0.38$
- (b) $\mathbb{P}(B) = 0.9$
- (c) $\mathbb{P}(C|A) = \frac{0.3 \cdot 0.8}{0.38} = 0.632$
- (d) $\mathbb{P}(C|D) = \frac{0.3 \cdot (0.9 \cdot 0.8 + 0.1 \cdot 0.2)}{0.9 \cdot 0.38 + 0.1 \cdot (1 - 0.38)} = 0.55$

2. Машин результат за контрольную, M , равномерно распределен на отрезке $[0; 1]$. Вовочка ничего не знает, поэтому списывает у Маши, да ещё может наделать ошибок при списывании. Поэтому Вовочкин результат, V , распределен равномерно от нуля до Машиного результата.

- (a) Найдите $\mathbb{P}(M > 2V)$, $\mathbb{P}(M > V + 0.1)$
- (b) Зачёт получают те, чей результат больше 0.4. Какова вероятность того, что Вовочка получит зачёт? Какова вероятность того, что Вовочка получит зачёт, если Маша получила зачёт?

Подсказка: попробуйте нарисовать нужные события в осях (V, M)

Это была задачка-неберучка!

3. Функция плотности случайной величины X имеет вид
$$f(x) = \begin{cases} \frac{3}{7}x^2, & x \in [1; 2] \\ 0, & x \notin [1, 2] \end{cases}$$

- (a) Не производя вычислений найдите $\int_{-\infty}^{+\infty} f(x) dx$
- (b) Найдите $\mathbb{E}(X)$, $\mathbb{E}(X^2)$ и дисперсию $\text{Var}(X)$
- (c) Найдите $\mathbb{P}(X > 1.5)$
- (d) Найдите функцию распределения $F(x)$ и постройте её график

Ответы:

- (a) 1
- (b) $\mathbb{E}(X) = 45/28 \approx 1.61$, $\mathbb{E}(X^2) = 93/35 \approx 2.66$, $\text{Var}(X) = 291/3920 \approx 0.07$
- (c) $37/56 \approx 0.66$
- (d)
$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{x^3-1}{7}, & x \in [1; 2] \\ 1, & x > 2 \end{cases}$$

4. Совместное распределение случайных величин X и Y задано таблицей

	$X = -2$	$X = 0$	$X = 2$
$Y = 1$	0.2	0.3	0.1
$Y = 2$	0.1	0.2	a

- (a) Определите неизвестную вероятность a .
- (b) Найдите вероятности $\mathbb{P}(X > -1)$, $\mathbb{P}(X > Y)$
- (c) Найдите математические ожидания $\mathbb{E}(X)$, $\mathbb{E}(X^2)$
- (d) Найдите корреляцию $\text{Corr}(X, Y)$

Ответы:

- (a) $a = 0.1$, [1]
 - (b) $\mathbb{P}(X > -1) = 0.7$, $\mathbb{P}(X > Y) = 0.1$ [3]
 - (c) $\mathbb{E}(X) = -0.2$, $\mathbb{E}(X^2) = 2$ [3]
 - (d) $\text{Corr}(X, Y) = 0.117$ [3]
5. Винни Пух собрался полакомиться медом, но ему необходимо принять решение, к каким пчелам отправиться за медом. Неправильные пчелы кусают каждого, кто лезет к ним на дерево с вероятностью 0,9, но их всего 10 штук. Правильные пчелы кусаются с вероятностью 0,1, но их 100 штук.
- (a) Определите математическое ожидание и дисперсию числа укусов Винни Пуха для каждого случая

- (b) Определите наиболее вероятное число укусов и его вероятность для каждого случая
- (c) К каким пчелам следует отправиться Винни Пуху, если он не может выдержать больше двух укусов?

Ответы:

- (a) Правильные: $\mathbb{E}(X) = 10$, $\text{Var}(X) = 9$, Неправильные: $\mathbb{E}(Y) = 9$, $\text{Var}(Y) = 0.9$
- (b) Наиболее вероятное число укусов равно математическому ожиданию
- (c) Лучше идти к неправильным пчелам, т.к. $\mathbb{P}(X \leq 2) < \mathbb{P}(Y \leq 2)$.

10.2 Контрольная работа №2, 26.12.2012

Тест:

1. Зная распределение компонент случайного вектора всегда можно восстановить их совместное распределение. Да. Нет.
2. Пусть X — длина наугад выловленного удава в сантиметрах, а Y — в дециметрах. Коэффициент корреляции между этими величинами равен 0.1. Да. Нет.
3. Для любой случайной величины X справедливо неравенство

$$\mathbb{P}(|X - \mathbb{E}(X)| > 2\sqrt{\text{Var}(X)}) \leq 1/4$$

Да. Нет.

4. Сумма независимых нормальных случайных величин нормальна. Да. Нет.
5. Сумма n независимых равномерно распределенных на интервале $(0, 1)$ случайных величин асимптотически нормальна. Да. Нет.
6. Квадрат стандартной нормальной случайной величины имеет хи-квадрат распределение. Да. Нет.
7. Если ковариация компонент случайного двумерного нормального вектора равна нулю, то они независимы. Да. Нет.
8. Условная дисперсия всегда больше безусловной. Да. Нет.
9. Элементы выборки без возвращения из конечной генеральной совокупности независимы. Да. Нет.
10. Математическое ожидание выборочного среднего одинаково распределенных случайных величин не зависит от объема выборки. Да. Нет.
11. Конец света по техническим причинам переносится на показ работ по теории вероятности. Да. Может быть.

Ответы: Нет, Нет, Да, Да, Да, Да, Да, Нет, Нет (зависимы), Да (не зависит), Любой верный
Задачи:

1. Купчиха Сосипатра Титовна очень любит чаевничать. Её чаепитие продолжается случайное время S , имеющее равномерное распределение от 0 до 3 часов. Встретив Сосипатру Титовну в пассаже на Петровке, её подруга Олимпиада Карповна узнала, сколько длилось вчерашнее чаепитие Сосипатры Титовны. Решив, что такая продолжительность чаепития является максимально возможной, Олимпиада Карповна устраивает чаепитие, продолжающееся случайное время T , имеющее равномерное распределение от 0 до S часов.

- (a) Найдите совместную функцию плотности величин S и T
- (b) Найдите вероятность $\mathbb{P}(S > T)$
- (c) Найдите $\mathbb{E}(T^2)$

Решение: $f(s, t) = f(s) \cdot f(t|s) = \frac{1}{3s}$ при $0 \leq t \leq s \leq 3$. Бонус тем, кто прочитал условие, $\mathbb{P}(S > T) = 1$.

$$\mathbb{E}(T^2) = \int_0^3 \int_0^s \frac{t^2}{3s} dt ds = 1$$

2. Для случайно выбранного домохозяйства случайные величины X и Y принимают значения, равные доле расходов на продукты питания и алкоголь плюс табак соответственно. Случайный вектор $(X, Y)^T$ хорошо описывается двумерным нормальным законом распределения с математическим ожиданием $(0.45, 0.16)^T$ и ковариационной матрицей

$$C = 0.144 \cdot \begin{pmatrix} 1 & -0.9 \\ -0.9 & 1 \end{pmatrix}$$

Найдите:

- (a) Вероятность того, что домохозяйство тратит более половины своих доходов на питание.
- (b) Вероятность того, что домохозяйство тратит более половины своих доходов на алкогольную и табачную продукцию и продукты питания.
- (c) Ожидаемую долю расходов на алкоголь и табак для домохозяйства, которое тратит на питание четверть своих доходов.
- (d) Вероятность того, что домохозяйство из предыдущего пункта тратит более трети своих доходов на алкогольную и табачную продукцию.
- (e) Для доли расходов на питание вычислите центральный момент 2013-го порядка.

Решение:

- (a) $\mathbb{P}(X > 0.5) = \mathbb{P}(Z > 0.1317616) \approx 0.4475864$, $\sigma_X \approx 0.3794733$
- (b) $\mathbb{P}(X + Y > 0.5) = \mathbb{P}(Z > -0.6481812) \approx 0.7415661$, $\sigma_{X+Y} = 0.1697056$, $\mathbb{E}(X + Y) = 0.61$
- (c) $X = 0.25$ при нормировке даёт $\tilde{X} = -0.5270463$. Получаем: $\mathbb{E}(\tilde{Y} | \tilde{X} = -0.5270463) = 0.4743416$, $\text{Var}(\tilde{Y} | \tilde{X} = -0.5270463) = 0.19$.
Значит $\mathbb{E}(Y | \tilde{X} = -0.5270463) = 0.34$, $\text{Var}(Y | \tilde{X} = -0.5270463) = 0.02736$.
- (d) $\mathbb{P}(Y > 1/3 | \tilde{X} = -0.5270463) = \mathbb{P}(Z > -0.0403042) = 0.5160747$
- (e) Ноль

3. Вычислите (или оцените) вероятность того, что по результатам 4000 бросаний симметричной монеты, частота выпадения герба будет отличаться от 0.5 не более, чем на 0.01. Решите задачу с помощью неравенства Чебышёва и с помощью ЦПТ.

Решение: $\mathbb{E}(\hat{p}) = 0.5$, $\text{Var}(\hat{p}) = 0.25/n = 1/16000$. По Чебышёву:

$$\mathbb{P}(|\hat{p} - 0.5| \leq 0.01) \geq 1 - \frac{\text{Var}(\hat{p})}{0.01^2} = \dots = 0.375$$

Используя нормальную аппроксимацию:

$$\mathbb{P}(|\hat{p} - 0.5| \leq 0.01) = \mathbb{P}(|Z| \leq 1.2649111) \approx 0.7940968$$

4. Компания кабельного телевидения НВТ, Новая Вершина Телевидения, анализирует возможность присоединения к своей сети пригородов N-ска. Опросы показали, что в среднем каждые 3 из 10 семей жителей пригородов хотели бы стать абонентами сети. Стоимость работ, необходимых для организации сети в любом пригороде оценивается величиной 2 080 000 у.е. При подключении каждого пригорода НВТ надеется получить 1 000 000 у.е. в год от рекламодателей. Планируемая чистая прибыль от оплаты за кабельное телевидение одной семьей в год равна 120 у.е.

Каким должно быть минимальное количество семей в пригороде для того, чтобы с вероятностью 0.99 расходы на организацию сети в этом пригороде окупились за год?

Решение:

Обозначим N — количество подключенных абонентов, тогда $N \sim \text{Bin}(n, 0.3)$. При больших n биномиальное распределение можно заменить на нормальное, $N \sim N(0.3n, 0.21n)$.

$$\mathbb{P}(120N > 1\,080\,000) = \mathbb{P}(N > 9000) = \mathbb{P}\left(Z > \frac{9000 - 0.3n}{\sqrt{0.21n}}\right) = 0.99$$

Из таблицы находим, что

$$\frac{9000 - 0.3n}{\sqrt{0.21n}} = -2.3263479$$

Решаем квадратное уравнение, находим корни, один — отрицательный, другой, $n \approx 30622$.

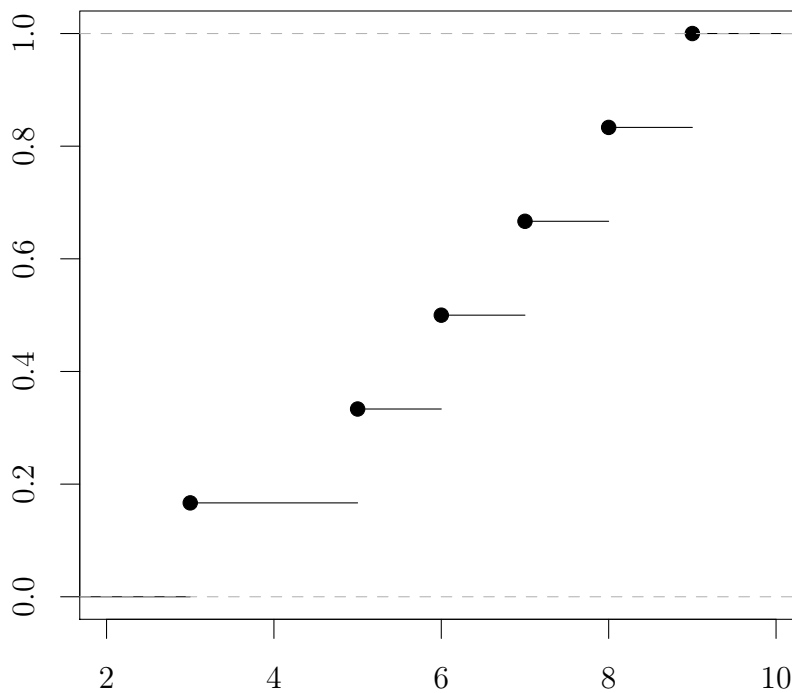
5. Оценки за контрольную работу по теории вероятностей 6 случайно выбранных студентов оказались равны:

8 5 6 7 3 9.

- (а) Выпишите вариационный ряд
- (б) Постройте график выборочной функции распределения
- (с) Вычислите значение выборочного среднего и выборочной дисперсии.

Решение: Вариационный ряд: 3, 5, 6, 7, 8, 9. $\bar{X} \approx 6.3333333$, $\frac{\sum(X_i - \bar{X})^2}{n-1} \approx 4.6666667$, $\frac{\sum(X_i - \bar{X})^2}{n} \approx 3.8888889$

Эмпирическая функция распределения



10.3 Демо-версия зачёта

1. Двое подельников, Маша и Саша, украли десять миллионов евро. Через некоторое время Саша был найден убитым, а Маша была арестована. Из свидетельских показаний ясно следует, что Маша и Саша ругались по поводу делёжки. Защита и обвинение выясняют, убила ли Маша Сашу. Из статистических данных известно, что:

- А 20% подельников-мужчин ругаются по поводу делёжки
- В 20% оставшихся в живых подельников-мужчин ругаются по поводу делёжки
- С 5% мужчин убивают
- Д 3% мужчин убивают их подельники
- Е 90% убитых мужчин-подельников ругались по поводу делёжки

Располагая этой информацией,

- (а) Найдите вероятность того, что Маша убила Сашу, если известно, что Маша и Саша ругались по поводу делёжки.
 - (б) Найдите вероятность того, что Маша убила Сашу, если известно, что Маша и Саша ругались по поводу делёжки, и Саша был найден убитым.
2. Маша подкидывает 300 игральных кубиков. Те, что выпали не на шестёрку, она перекидывает один раз. Обозначим буквой N количество шестёрок на всех кубиках после возможных перекидываний.
 - (а) Найдите $\mathbb{E}(N)$, $\text{Var}(N)$

- (b) Какова примерно вероятность того, величина N лежит от 50 до 70?
- (c) Укажите любой интервал, в который величина N попадает с вероятностью 0.9
3. На лукоморье набегают волны. Кот Учёный заметил, что размер каждой волны, X_i , — случайная величина, имеющая равномерное распределение от 0 до 1, а размеры волн независимы. Кот Учёный считает волну *большой*, если она больше предыдущей и следующей. Случайная величина R_i равна 1, если i -ая волна была *большой*, и 0 иначе.
- (a) Найдите $\mathbb{P}(R_i = 1)$, $\mathbb{E}(X_i)$
- (b) Найдите $\mathbb{E}(X_i | R_i = 1)$
- (c) Найдите $\text{Cov}(R_1, R_2)$, $\text{Cov}(R_1, R_3)$
4. Ермолай Лопухин решил приступить к вырубке вишневого сада. Однако выяснилось, что растут в нём не только вишни, но и яблони. Причём, по словам Любви Андреевны Раневской, среднее количество деревьев (а они периодически погибают от холода или жары, либо из семян вырастают новые) в саду распределено в соответствии с нормальным законом (X — число яблонь, Y — число вишен) со следующими параметрами:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 25 \\ 125 \end{pmatrix}; \begin{pmatrix} 5 & 4 \\ 4 & 10 \end{pmatrix} \right) \quad (41)$$

- (a) Найдите вероятность того, что Ермолаю Лопухину придется вырубить более 150 деревьев.
- (b) Каково ожидаемое число подлежащих вырубке вишен, если известно, что предприимчивый и последовательный Лопухин, не затронув ни одного вишнёвого дерева, начал очистку сада с яблонь и все 35 яблонь уже вырубил? Какова при этом вероятность того, что Лопухину придется вырубить более 100 вишен?
5. Вопрос из интервью в Морган-Стэнли. Есть две независимых равномерных на отрезке $[0; 1]$ случайных величины, X и Y . Как их нужно преобразовать, чтобы корреляция между ними оказалась равна ρ ?

10.4 Зачёт, 15.01.2013

1. Самолёт упал либо в горах, либо на равнине. Вероятность того, что самолёт упал в горах, равна 0,75. Для поиска пропавшего самолёта выделено 3 вертолёт. Каждый вертолёт можно использовать только в одном месте. Как распределить имеющиеся вертолёты, если вероятность обнаружения пропавшего самолёта отдельно взятым вертолётom равна 0,6?
2. Совместная функция плотности величин X и Y имеет вид

$$f(x, y) = \frac{1}{x} e^{-x}, \text{ при } 0 < y < x$$

- (a) Найдите $\mathbb{P}\left(\frac{Y}{X} < 0.7\right)$
- (b) Найдите $\mathbb{E}(X)$
- (c) Являются ли X и Y независимыми?
- (d) Как распределена величина $Z = Y/X$?

3. Величины X_1, \dots, X_n независимы и имеют биномиальное распределение, $X_i \sim \text{Bin}(10, p)$. Используя неравенство Чебышёва найдите наименьшее число t , чтобы выполнялось условие

$$\mathbb{P}(|\bar{X} - \mathbb{E}(\bar{X})| \geq t) \leq 0.01$$

4. 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.
- Какой процент потребителей в среднем обращается за гарантийным ремонтом?
 - Какова вероятность того, что из 1000 потребителей за гарантийным ремонтом обратится более 55% покупателей?

Подсказка: $\ln 2 \approx 0.7$

5. Вася попадает мячом в корзину с вероятностью 0.2, Петя — с вероятностью 0.25. Каждый из них сделал по 100 бросков мяча.
- Какова вероятность того, что Петя попал на 10 раз больше Васи?
 - Какое минимальное количество бросков мяча нужно сделать каждому, чтобы вероятность, того, что Петя попал на 10 раз больше Васи достигла 0.99?

10.5 КоКо, компьютерная контрольная №3, 13.03.13

Продолжительность 1 час 10 минут, разрешено пользоваться конспектами, книжками, заготовками программ, нельзя общаться, использовать Интернет. Текст работы в группах с R:

- Величины X и Y независимы. Величина X распределена нормально, $X \sim N(4.4, 6.1)$, величина Y распределена экспоненциально, $Y \sim \exp(\lambda = 2.9)$. Используя симуляционный подход примерно посчитайте
 - $\mathbb{P}(X + Y > 5.3)$
 - $\mathbb{E}(X/(X + 3Y))$
 - $\text{Var}(XY)$
 - $\text{Cov}(XY, X/Y)$
- Загрузите данные по стоимости квартир в Москве, goo.gl/zL5JQ, в табличку с именем `h`. Обозначим буквой `a` ответ на первый вопрос первой задачи. Отберите индивидуальную выборку лично для себя, выполнив команды:

```
set.seed(round(100*a)) # здесь "a" --- это ответ на первый пункт первой задачи
h <- h[sample(1:nrow(h), 1000),]
```

Постройте 90%-ый доверительный интервал для:

- Доли кирпичных домов, `brick==1`
 - Доли кирпичных домов, `brick==1`, среди домов находящихся близко от метро, `walk==1`
 - Разницы доли кирпичных домов среди домов расположенных близко и далеко от метро
3. Сгенерируйте искусственные данные, выполнив команды:

```
set.seed(round(100*a)+42) # здесь "a" --- это ответ на первый пункт первой задачи
x <- rexp(200,rate=2)
```

Величины X_i независимы и имеют функцию плотности $f(x) = e^{-xe^b}$ при $x > 0$.

- Оцените неизвестный параметр b
- Оцените дисперсию полученной оценки
- Постройте 90%-ый доверительный интервал для b
- Используя результат предыдущего пункта, на 10%-ом уровне значимости проверьте гипотезу $H_0: b = 0.7$ против альтернативной гипотезы $H_a: b \neq 0.7$.

10.6 Экзамен, 26.03.2013

- Вероятность выигрыша по лотерейному билету равна 0.05. Вероятность того, что из трёх купленных билетов ровно два окажутся выигрышными примерно равна

1) 0.002	2) 0.0025	3) 0.007	4) 0.1	5) 0.3	Ответ:
----------	-----------	----------	--------	--------	--------

- Закон распределения случайной величины задан табличкой

X	-1	0	1
$\mathbb{P}()$	0.4	0.2	?

Дисперсия величины X , $\text{Var}(X)$, равняется

1) 0	2) 0.02	3) 0.3	4) 0.8	5) 2	Ответ:
------	---------	--------	--------	------	--------

- Если $f(x)$ — функция плотности, то $\int_{-\infty}^x f(u) du$ равен

1) 0	2) 1	3) $\mathbb{E}(X)$	4) $\text{Var}(X)$	5) $F(x)$	Ответ:
------	------	--------------------	--------------------	-----------	--------

- События A и B называются независимыми, если

1) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ 2) $\mathbb{P}(A) \cdot \mathbb{P}(B) = \mathbb{P}(A \cap B)$ 3) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ 4) $\mathbb{P}(A \cap B) = 0$ 5) нет верного	Ответ:
---	--------

- Правильную монетку подбрасывают два раза. Рассмотрим два события: A — при втором броске выпала «решка», 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:
------	--------	--------	--------	------	--------

6. Есть пять случайных величин: $X \sim \chi_{10}^2$, $Y \sim F_{5,10}$, $T \sim t_{10}$, $Z \sim N(0, 1)$, $W \sim N(10, 1)$. Какие из величин распределены симметрично относительно 0?

1) X, Y, Z	2) Z, W	3) Z, T	4) Z	5) X, Y	Ответ:
------------	---------	---------	------	---------	--------

7. Известно, что $\mathbb{E}(X) = 3$, $\text{Var}(X) = 1$, $\mathbb{E}(Y) = 4$, $\text{Var}(Y) = 9$, $\mathbb{E}(XY) = 13$, найдите $\text{Cov}(X, Y)$

1) 0	2) -3	3) 18	4) 3	5) 1	Ответ:
------	-------	-------	------	------	--------

8. Известно, что $\mathbb{E}(X) = 3$, $\text{Var}(X) = 1$, $\mathbb{E}(Y) = 4$, $\text{Var}(Y) = 9$, $\mathbb{E}(XY) = 6$, найдите $\text{Var}(2X + Y)$

1) 13	2) 7	3) 1	4) 17	5) нет верного ответа	Ответ:
-------	------	------	-------	-----------------------	--------

9. 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:
------	--------	------	--------------------	--------------	--------

10. Если $X \sim N(-4; 1)$, то $\mathbb{P}(3X + 571 > 0)$ примерно равна

1) 0	2) 0.5	3) 1	4) $+\infty$	5) нет верного ответа	Ответ:
------	--------	------	--------------	-----------------------	--------

11. Про закон распределения величины X ничего не известно. Укажите самую точную оценку сверху для вероятности $\mathbb{P}(|X - \mathbb{E}(X)| > 3\sqrt{\text{Var}(X)})$

1) 0.(3)	2) 0.6(3)	3) 0.(1)	4) 1	5) нет верного ответа	Ответ:
----------	-----------	----------	------	-----------------------	--------

12. Функция распределения, $F(x) = \mathbb{P}(X \leq x)$ может не являться

1) непрерывной	2) непрерывной справа	3) монотонно неубывающей	4) ограниченной	5) неотрицательной	Ответ:
----------------	-----------------------	--------------------------	-----------------	--------------------	--------

13. Ковариационной может быть матрица:

1) $\begin{pmatrix} -1 & 1 \\ 1 & 2 \end{pmatrix}$	2) $\begin{pmatrix} 1 & 0.5 \\ 1 & 2 \end{pmatrix}$	3) $\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$	4) $\begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix}$	5) $\begin{pmatrix} 1 & 1 \\ -0.7 & 2 \end{pmatrix}$	Ответ:
--	---	---	---	--	--------

14. Если X и Y независимые случайные величины, то неверным может быть утверждение

1) $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ 2) $\mathbb{E}(X/Y) = \mathbb{E}(X)/\mathbb{E}(Y)$ 3) $\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$ 4) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ 5) $\text{Cov}(X, Y) = 0$	Ответ:
--	--------

15. Известно, что $\text{Cov}(X, Y) = 0$, $\text{Var}(X) = 10$, $\text{Var}(Y) = 10$.
Неверным может быть утверждение

1) $\text{Corr}(X, Y) = 0$ 2) $\text{Corr}(X + a, Y + b) = 0$ 3) $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$ 4) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ 5) X и Y независимы	Ответ:
---	--------

16. $Z_1, Z_2, \dots, Z_n \sim N(0, 1)$. Тогда величина $\frac{Z_1 \sqrt{n-2}}{\sqrt{\sum_{i=3}^n Z_i^2}}$ имеет распределение

1) $N(0, 1)$	2) t_n	3) $F_{1,n-2}$	4) χ_n^2	5) t_{n-2}	Ответ:
--------------	----------	----------------	---------------	--------------	--------

17. Количество страниц в книгах авторов X и Y распределено нормально с дисперсиями σ_X^2 и σ_Y^2 соответственно. Для тестирования гипотезы о равенстве дисперсий было выбрано n книг автора X и m книг — автора Y . Какое распределение имеет статистика, используемая в данном случае?

1) $\chi_{\min(m,n)}^2$	2) $\chi_{\max(m,n)}^2$	3) $F_{m,n}$	4) $F_{m-1,n-1}$	5) $F_{m+1,n+1}$	Ответ:
-------------------------	-------------------------	--------------	------------------	------------------	--------

18. Если величина X имеет χ_k^2 -распределение, величина Y — χ_n^2 -распределение, и они независимы, то дробь $nX/(kY)$ имеет распределение

1) $F_{k,n}$	2) $F_{n,k}$	3) $F_{k-1,n-1}$	4) χ_{n-k}^2	5) $F_{n-1,k-1}$	Ответ:
--------------	--------------	------------------	-------------------	------------------	--------

19. Смещенной оценкой математического ожидания по выборке независимых, одинаково распределенных случайных величин X_1, X_2, X_3 является оценка

1) $(X_1 + X_2)/2$ 2) $(X_1 + X_2 + X_3)/3$ 3) $0.7X_1 + 0.2X_2 + 0.1X_3$ 4) $0.3X_1 + 0.3X_2 + 0.3X_3$ 5) $X_1 + X_2 - X_3$	Ответ:
--	--------

20. Если величины X и Y независимы и равномерно распределены на $[0; 1]$, а $F(x, y)$ — их совместная функция распределения, то $F(0.5, 3)$ равно

1) 0	2) 0.5	3) 1	4) 1.5	5) не существует	Ответ:
------	--------	------	--------	------------------	--------

21. Если X_i независимы и имеют нормальное распределение $N(-1; 2013)$, то $\sqrt{n}(1 + \bar{X})/\hat{\sigma}$ имеет распределение

1) $N(0; 1)$	2) t_{n-1}	3) χ_{n-1}^2	4) $N(0; 2013/n)$	5) t_n	Ответ:
--------------	--------------	-------------------	-------------------	----------	--------

22. Последовательность оценок $\hat{\theta}_1, \hat{\theta}_2, \dots$ называется состоятельной, если

1) $\mathbb{E}(\hat{\theta}_n) = \theta$ 2) $\text{Var}(\hat{\theta}_n) \rightarrow 0$ 3) $\mathbb{P}(\hat{\theta}_n - \theta > t) \rightarrow 0$ для всех $t > 0$ 4) $\mathbb{E}(\hat{\theta}_n) \rightarrow \theta$ 5) $\text{Var}(\hat{\theta}_n) \geq \text{Var}(\hat{\theta}_{n+1})$	Ответ:
---	--------

23. Величины X_1, \dots, X_5 равномерны на отрезке $[0; a]$. Известно, что $\sum_{i=1}^5 x_i = 25$. При использовании первого момента оценка методом моментов неизвестного a равна

1) 1	2) 5	3) 10	4) 20	5) нет верного ответа	Ответ:
------	------	-------	-------	-----------------------	--------

24. При построении доверительного интервала для отношения дисперсий по двум независимым нормальным выборкам из n наблюдений каждая, используется статистика, имеющая распределение

1) $F_{n-1, n-1}$	2) t_{n-1}	3) χ_{n-1}^2	4) χ_n^2	5) t_n	Ответ:
-------------------	--------------	-------------------	---------------	----------	--------

25. Ботаники² строят доверительный интервал для математического ожидания числа иголок у ёжа. Количество иголок на одном еже предполагается нормально распределенным. Среднее число иголок у пойманных 100 ёжиков равно 1500, выборочная дисперсия — 400. На 5% уровне значимости какой примерно доверительный интервал должны построить ботаники?

1) [1499; 1501] 2) [1498; 1502] 3) [1497; 1503] 4) [1496; 1504] 5) нет верного ответа	Ответ:
---	--------

26. Функция правдоподобия, построенная по случайной выборке X_1, \dots, X_n из распределения с функцией плотности $f(x) = (\theta + 1)x^\theta$ при $x \in [0; 1]$ имеет вид

1) $(\theta + 1)x^{n\theta}$	2) $\sum(\theta + 1)x_i^\theta$	3) $(\theta + 1)^{\sum x_i}$	4) $(\sum x_i)^\theta$	5) $(\theta + 1)^n \prod x_i^\theta$	Ответ:
------------------------------	---------------------------------	------------------------------	------------------------	--------------------------------------	--------

²В отличие от ботаников, зоологи точно знают, сколько иголок у ёжиков!

27. Если X_i независимы, $\mathbb{E}(X_i) = \mu$ и $\text{Var}(X_i) = \sigma^2$, то математическое ожидание величины $Y = \sum_{i=1}^n (X_i - \bar{X})^2$ равно

1) $\hat{\sigma}^2$	2) $(n-1)\sigma^2$	3) μ	4) σ^2	5) σ^2/n	Ответ:
---------------------	--------------------	----------	---------------	-----------------	--------

28. Если P -значение меньше уровня значимости α , то гипотеза $H_0: \sigma = \sigma_0$

1) отвергается 2) не отвергается 3) отвергается только если $H_a: \sigma \neq \sigma_0$ 4) отвергается только если $H_a: \sigma < \sigma_0$ 5) недостаточно информации	Ответ:
--	--------

29. Если H_0 верна, то P -значение имеет распределение

1) $U[0; 1]$	2) $N(0, 1)$	3) t_n	4) t_{n-1}	5) χ_{n-1}^2	Ответ:
--------------	--------------	----------	--------------	-------------------	--------

Экзамен по теории вероятностей! Суперигра!

DON'T PANIC

- В группе учатся 30 студентов, 20 девушек и 10 юношей. Они входят в аудиторию в случайном порядке. Рассчитайте вероятности событий:
 - Маша Петрова³ войдёт девятой по счёту
 - Девятый вошедший окажется девушкой
 - Перед Машей Петровой войдут ровно 5 юношей
 - Перед Машей Петровой войдут ровно 5 юношей, если известно, что Маша Петрова вошла девятой
 - Маша Петрова войдёт девятой по счёту, если известно, что перед ней вошло ровно 5 юношей
- В поселке 2500 жителей. Каждый из них в среднем 6 раз в месяц ездит в город, выбирая день поездки независимо от других людей. Поезд ходит в город один раз в сутки.
 - Какой наименьшей вместимостью должен обладать поезд, чтобы он переполнялся в среднем не чаще 1 раза в 100 дней?
 - Сколько в среднем человек будет ехать в таком поезде, если предположить, что при переполнении часть людей полностью откажется от поездки?

Источник: экзамен РЭШ

- Случайные величины X_1, X_2, \dots, X_n независимы и имеют пуассоновское распределение с неизвестным параметром λ .

³Маша Петрова — единственная и неподражаемая!

- (a) С помощью метода максимального правдоподобия постройте оценки для λ и для $\exp(\lambda)$
- (b) Предположим, что исследователь не знает, чему равны X_i . Ему известно лишь, равно ли каждое из X_i нулю или нет. С помощью метода максимального правдоподобия постройте оценки для λ и для $\exp(\lambda)$. Всегда ли существуют предложенные оценки?
4. В таблице представлены данные по количеству пассажиров «Титаника», поделенные на группы по классу каюты:

	1 класс	2 класс	3 класс
Погиб	122	167	528
Выжил	203	118	178

Проверьте гипотезу о независимости шансов выжить от класса каюты.

5. Перед Вами две внешне неотличимых монетки. Одна из них выпадает «орлом» вверх с вероятностью 0.7, другая — с вероятностью 0.3. Вы имеете право на 4 подбрасывания. Перед каждым подбрасыванием Вы можете выбирать подбрасываемую монетку. За каждого выпавшего «орла» вы получаете 1 рубль.
- (a) Какова оптимальная стратегия?
- (b) Каков ожидаемый выигрыш при использовании оптимальной стратегии?

11 2013-2014

11.1 Контрольная работа №1, 5.11.2013

1. Вероятность застать Васю на лекции зависит от того, пришли ли на лекцию Маша и Алена. Данная вероятность равна 0.18, если девушек нет; 0.9 — если обе девушки пришли на лекцию; 0.54 — если пришла только Маша и 0.36 — если пришла только Алена. Маша и Алена посещают лекции независимо друг от друга с вероятностями 0.4 и 0.6 соответственно.
- (a) Определите вероятность того, что на лекции присутствует Алена, если в аудитории есть Вася.
- (b) Кого чаще можно застать на тех лекциях, на которых присутствует Вася: Машу или Алену?
2. Страховая компания страхует туристов, выезжающих за границу, от невыезда и наступления страхового медицинского случая за границей. Застраховано 100 туристов. Вероятность «невыезда» за границу случайно выбранного туриста — 0.002, а страховые выплаты в этом случае — 2000 у.е.; вероятность обращения за медицинской помощью за границей — 0.01, а страховые выплаты — 3000 у.е. Для каждого туриста рассмотрим две случайные величины: X_i , равную 1 при невыезде за границу и 0 иначе, и Y_i , равную 1 при обращении за медицинской помощью и нулю иначе. Обозначим $X = \sum_{i=1}^{100} X_i$ и $Y = \sum_{i=1}^{100} Y_i$.
- (a) Определите $\mathbb{P}(X = 5)$, $\mathbb{E}(X)$, $\text{Var}(X)$
- (b) Наиболее вероятное число не выехавших туристов.

- (с) Вычислите математическое ожидание и дисперсию величины совокупных страховых выплат

Подсказка: Число обращений в страховую компанию для каждого туриста может быть записано в виде $X_i + X_i Y_i$, так как медицинский страховой случай может наступить только, если турист выехал за границу. Случайные величины X_i и Y_i независимы.

3. Функция плотности случайной величины X имеет вид:

$$f(x) = \begin{cases} ce^{-x}, & x \geq 0 \\ ce^x, & x < 0 \end{cases} \quad (42)$$

- (a) Найдите c , $\mathbb{P}(X \in [\ln 0.5, \ln 4])$, $\mathbb{E}(X)$, $\text{Var}(X)$
 (b) Моменты всех порядков случайной величины x

Подсказка: $\int_0^\infty x^n e^{-x} dx = n!$

4. Известно, что $\mathbb{E}(X) = -1$, $\mathbb{E}(Y) = 1$, $\text{Var}(X) = 9$, $\text{Var}(Y) = 4$, $\text{Corr}(X, Y) = 1$. Найдите
 (a) $\mathbb{E}(Y - 2X - 3)$, $\text{Var}(Y - 2X - 3)$
 (b) $\text{Corr}(Y - 2X - 3, X)$
 (c) Можно ли выразить Y через X ? Если да, то запишите уравнение связи.
5. Совместное распределение доходов акций двух компаний Y и X задано в виде таблицы

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

- (a) Найдите частные распределения случайных величин X и Y
 (b) Найдите $\text{Cov}(X, Y)$
 (c) Можно ли утверждать, что случайные величины X и Y зависимы?
 (d) Найдите условное распределение случайной величины X при условии $Y = -1$
 (e) Найдите условное математическое ожидание $\mathbb{E}(X | Y = -1)$

Некоторые ответы:

- xx
- $\mathbb{P}(X = 5) = C_{100}^5 0.002^5 0.998^{95}$, $\mathbb{E}(X) = 0.2$, $\text{Var}(X) = 0.2 \cdot 0.998$, наиболее вероятно событие $X = 0$
- $c = 1/2$, $P = 5/8$, $\mathbb{E}(X) = 0$, $\text{Var}(X) = 2$, $\mathbb{E}(X^{2k+1}) = 0$, $\mathbb{E}(X^{2k}) = (2k)!$
- xx
- xx

11.2 Контрольная работа №1, i-поток, 15.11.2013

Часть 1

- В жюри три человека, они должны одобрить или не одобрить конкурсанта. Два члена жюри независимо друг от друга одобряют конкурсанта с одинаковой вероятностью p . Третий член жюри для вынесения решения бросает правильную монету. Окончательное решение выносится большинством голосов.
 - С какой вероятностью жюри одобрит конкурсанта?
 - Что выгоднее для конкурсанта: чтобы решение принимало данное жюри, или чтобы решение принимал один человек, одобряющий с вероятностью p ?
- Вероятность застать Васю на лекции зависит от того, пришли ли на лекцию Маша и Алена. Данная вероятность равна p , если девушек нет; $5p$ — если обе девушки пришли на лекцию; $3p$ — если пришла только Маша и $2p$ — если пришла только Алена. Маша и Алена посещают лекции независимо друг от друга с вероятностями 0.6 и 0.3 соответственно.
 - Определите вероятность того, что на лекции присутствует Алена, если в аудитории есть Вася.
 - Кого чаще можно застать на тех лекциях, на которых присутствует Вася: Машу или Алену?
 - При каком значении p Вася посещает половину всех лекций?
- Страховая компания страхует туристов, выезжающих за границу, от невыезда и наступления страхового медицинского случая за границей. Застраховано 100 туристов. Вероятность «невыезда» за границу случайно выбранного туриста — 0.002, а страховые выплаты в этом случае — 2000 у.е.; вероятность обращения за медицинской помощью за границей — 0.01, а страховые выплаты — 3000 у.е.
 - Определите вероятность того, что ровно пятеро туристов не смогут выехать за границу.
 - Найдите математическое ожидание, дисперсию и наиболее вероятное число не выехавших туристов.
 - Вычислите математическое ожидание и дисперсию величины совокупных страховых выплат
 - Вычислите ковариацию между выплатами по двум видам страхования.
- Известно, что $\mathbb{E}(X) = -1$, $\mathbb{E}(Y) = 1$, $\text{Var}(X) = 9$, $\text{Var}(Y) = 4$, $\text{Corr}(X, Y) = 1$. Найдите
 - $\mathbb{E}(Y - 2X - 3)$, $\text{Var}(Y - 2X - 3)$
 - $\text{Corr}(Y - 2X - 3, X)$
 - Можно ли выразить Y через X ? Если да, то запишите уравнение связи.

Решение: корреляция равна 1, значит есть линейная взаимосвязь между переменными. Пусть $Y + \beta X = \text{const}$, тогда $\text{Var}(Y + \beta X) = 0$. Решая уравнение находим, что $\beta = -2/3$.

- Совместное распределение доходов акций двух компаний Y и X задано в виде таблицы

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

Найдите:

- (a) Частные распределения случайных величин X и Y
- (b) $\text{Cov}(X, Y)$
- (c) Можно ли утверждать, что случайные величины X и Y зависимы?
- (d) У инвестора портфель, в котором доля акций X составляет α , а доля акций Y — $(1 - \alpha)$. Каковы должны быть доли, чтобы риск портфеля (дисперсия дохода) был бы минимальным?
- (e) Условное распределение случайной величины X при условии $Y = -1$
- (f) Условное математическое ожидание $\mathbb{E}(X \mid Y = -1)$

6. Докажите, что из сходимости в среднем порядка $s > 0$ следует сходимость по вероятности.

Часть 2

1. Муравей находится внутри спичечного коробка, в вершине A . В противоположной вершине B есть маленькая дырочка, через которую муравей сможет выбраться на поверхность. В вершине C , соседней с вершиной A , лежит крупинка сахара. Муравей ползает только по рёбрам коробка, выбирая каждый раз равновероятно одно из доступных в вершине рёбер наугад. Например, он может поползти обратно.
 - (a) Какова вероятность того, что муравей найдет крупинку сахара до того, как выберется?
 - (b) Сколько в среднем перемещений понадобится муравью, чтобы выбраться?
 - (c) Какова дисперсия количества перемещений, которые понадобятся муравью, чтобы выбраться?
2. В очереди стояло 20 человек, когда касса внезапно закрылась. Поэтому 10 случайных людей из очереди решили покинуть очередь. В результате этого очередь оказалась разбита на случайное число кусков X . Найдите $\mathbb{E}(X)$, $\text{Var}(X)$.
3. Предположим, что три возможных генотипа aa , Aa и AA изначально встречаются с частотами p_1 , p_2 и p_3 , где $p_1 + p_2 + p_3 = 1$. Ген не сцеплен с полом, поэтому частоты p_1 , p_2 и p_3 одинаковы для мужчин и для женщин.
 - (a) У семейных пар из этой популяции рождаются дети. Назовём этих детей первым поколением. Каковы частоты для трёх возможных генотипов в первом поколении?
 - (b) У семейных пар первого поколения тоже рождаются дети. Назовём этих детей вторым поколением. Каковы частоты для трёх возможных генотипов во втором поколении?
 - (c) Каковы частоты для трёх возможных генотипов в n -ном поколении?
 - (d) Заметив явную особенность предыдущего ответа сформулируйте теорему о равновесии Харди-Вайнберга. Прокомментируйте утверждение: «Любой рецессивный ген со временем исчезнет».
4. Световая волна может быть разложена на две поляризованные составляющие, вертикальную и горизонтальную. Поэтому состояние отдельного поляризованного фотона может быть описано⁴ углом α . Поляризационный фильтр описывается углом поворота θ . Фотон в состоянии α задерживается поляризационным фильтром с параметром θ с вероятностью $p = \sin^2(\alpha - \theta)$ или проходит сквозь фильтр с вероятностью $1 - p$, переходя при этом в состояние θ .

⁴На самом деле внутренний мир фотона гораздо разнообразнее.

- (a) Какова вероятность того, что поляризованный фотон в состоянии α пройдет сквозь фильтр с параметром $\theta = 0$?
- (b) Имеется два фильтра и поляризованный фотон в состоянии α . Первый фильтр — с $\theta = 0$, второй — с $\theta = \pi/2$. Какова вероятность того, что фотон пройдет через оба фильтра?
- (c) Имеется три фильтра и поляризованный фотон в состоянии α . Первый фильтр — с $\theta = 0$, второй — с $\theta = \beta$, третий — с $\theta = \pi/2$. Какова вероятность того, что фотон пройдет через все три фильтра? При каких α и β она будет максимальной и чему при этом она будет равна?
- (d) Объясните следующий фокус. Фокусник берет два специальных стекла и видно, что свет сквозь них не проходит. Фокусник ставит между двумя стеклами третье, и свет начинает проходить через три стекла.

Некоторые ответы:

- 1. $\mathbb{P}(A) = 2/3$
- 2. $\mathbb{E}(X) = 5.5$

11.3 Контрольная работа №2, i-поток, 16-28.12.2013

Заочная R-часть

- 1. Случайная величина X имеет t -распределение с 5-тью степенями свободы.
 - (a) На одном графике постройте функцию плотности случайной величины X и функцию плотности стандартного нормального распределения.
 - (b) На одном графике постройте функцию распределения случайной величины X и функцию распределения стандартной нормальной случайной величины.
 - (c) Постройте график зависимости вероятности $\mathbb{P}(a < X < a + 10)$ от a . Если возможно, найдите такое число a , при котором эта вероятность равна 0.8.
 - (d) Постройте график зависимости вероятности $\mathbb{P}(b < X < 2b)$ от b при $b > 0$. Если возможно, найдите такое число b , при котором эта вероятность равна 0.2.
 - (e) С помощью 10^6 симуляций оцените $\mathbb{P}(X^3 + X > 3)$, $\mathbb{E}(1/(X^2 + 3))$, $\text{Var}(1/(X^2 + 3))$.
 - (f) На одном графике постройте гистограмму получившейся случайной выборки из 10^6 значений и функцию плотности X . Для сравнимости гистограммы и функции плотности масштаб гистограммы нужно выбрать так, чтобы площадь под ней равнялась единице.
 - (g) С помощью этой же случайной выборки найдите самый короткий интервал, куда $1/(X^2 + 1)$ попадает с вероятностью 0.9.
- 2. Слагаемые X_i независимы и экспоненциально распределены с параметром $\lambda = 2$. Обозначим сумму буквой S_n , т.е. $S_n = X_1 + \dots + X_n$.
 Для $n = 5$, $n = 10$, $n = 50$, $n = 100$:
 - (a) Сгенерируйте случайную выборку из 10^4 значений S_n .
 - (b) Постройте выборочную функцию распределения S_n .
 - (c) Найдите выборочное среднее и выборочную дисперсию S_n . Сравните их с настоящим математическим ожиданием $\mathbb{E}(S_n)$ и настоящей дисперсией $\text{Var}(S_n)$.

- (d) На одном графике в общем масштабе постройте гистограмму для S_n и функцию плотности нормально распределенной случайной величины с математическим ожиданием, равным $\mathbb{E}(S_n)$, и дисперсией, равной $\text{Var}(S_n)$.
- (e) Оцените по построенной случайной выборке вероятность.

$$\mathbb{P}(S_n \in [0.5\sqrt{\text{Var}(S_n)}; 2\sqrt{\text{Var}(S_n)}])$$

- (f) Оцените ту же вероятность, используя нормальную аппроксимацию.
- (g) Сравните две полученные оценки вероятности между собой.

3. Вектор (X, Y) имеет совместное нормальное распределение, $\mathbb{E}(X) = \mathbb{E}(Y) = 0$, $\text{Var}(X) = 1$, $\text{Var}(Y) = 9$, $\text{Corr}(X, Y) = \rho$.

Для $\rho = -0.9$, $\rho = 0$, $\rho = 0.5$:

- (a) Найдите ковариационную матрицу вектора (X, Y) , найдите её собственные числа и собственные векторы
- (b) Постройте график совместной функции плотности
- (c) Найдите $\mathbb{P}(X \in [0, 1], Y \in [-2, 1])$
- (d) Сгенерируйте случайную выборку из 10^3 пар значений (X_i, Y_i)
- (e) Найдите выборочную ковариацию и выборочную корреляцию между X_i и Y_i , сравните их с истинными ковариацией и корреляцией
- (f) На диаграмме рассеяния дополнительно постройте линии уровня совместной функции плотности $f(x, y)$
- (g) На диаграмме рассеяния дополнительно постройте собственные векторы с длиной равной корню из соответствующего собственного значения. Каков геометрический смысл собственных векторов ковариационной матрицы?
- (h) Оцените $\mathbb{P}(Y > X + 1)$ двумя способами: с помощью имеющейся случайной выборки и численно взяв интеграл от совместной функции плотности

* *Необязательная задача.* Вдоль края стоянки идёт неразмеченная парковка длиной 100 метров. Машины приезжают по очереди и паркуются перпендикулярно тротуару на случайное место, выбираемое равномерно из возможных для парковки. Водитель считает место возможным для парковки, если расстояние до машин слева и справа не менее 50 сантиметров. Ширина автомобиля 1.7 метра.

Случайная величина N — количество машин, которые смогут припарковаться на данной парковке. С помощью 10^6 симуляций ответьте на вопросы:

- (a) Сколько машин в среднем умещается на парковке?
- (b) Сколько места при случайной парковке пропадает в среднем «впустую» по сравнению с максимально аккуратной «размеченной» парковкой?
- (c) Чему равна дисперсия величины N ?
- (d) Найдите самый короткий интервал, в который N попадает с вероятностью 0.8.
- (e) Похоже ли распределение N на биномиальное? Для ответа на этот вопрос постройте на одном графике выборочную гистограмму для N и гистограмму истинных вероятностей для биномиального распределения со средним и дисперсией равным оценкам среднего и дисперсии для N .

Требования к оформлению домашнего задания:

1. Сдается в распечатанном виде в срок. Отмазки в духе «инопланетяне украли принтер утром, когда всё уже было готово» принимаются только вместе с видео-записью гуманоидов, похищающих принтер.
2. Обязательно использование языка R и пакета **knitr** с автоматическим созданием pdf-файла из Rnw-файла. Работы со шрифтом Times New Roman будут торжественно сожжены на кафедре до проверки! Код всех команд должен быть открыт для проверки.
3. Работа должна быть написана на русском языке. Do You speak English? Sprechen Sie Deutsch? Parlez-Vous Français?
4. На графиках должны быть подписаны оси. Convincing, <http://xkcd.com/833/>.
5. Обязательно в работе должны быть указаны: фамилия, имя, номер группы, e-mail. Необязательно — номер кредитной карточки с cvv кодом и сроком действия.

Очная часть, 25.12.2013

Самая важная формула:

$$\frac{1}{(\sqrt{2\pi})^n \det(C)} \exp\left(-\frac{1}{2}(x - \mu)'C^{-1}(x - \mu)\right)$$

Неравенство Берри-Эсеена:

$$|F_n(x) - \Phi(x)| \leq \frac{C_0 \mathbb{E}|X_1 - \mu|^3}{\sigma^3 \sqrt{n}}, \quad 0.4 < C_0 < 0.48$$

1. Складываются $n = 120$ чисел, каждое из которых округлено с точностью до 0.1. Предположим, что ошибки округления независимы и равномерно распределены в интервале $(-0.05, 0.05)$.
 - (a) Найдите пределы, в которых с вероятностью не меньшей 0.98 лежит суммарная ошибка.
 - (b) Вычислите максимальную погрешность, с которой истинная вероятность попадания в найденный интервал суммарной ошибки округления отличается от 0.98.

Подсказка: Следует искать симметричный относительно нуля интервал.

2. 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 поодиночке и выбирают входы равновероятно.
 - (a) 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?
 - (b) Предположим, что в каждом гардеробе ровно 500 мест. Найдите математическое ожидание числа зрителей, которым придется перейти в другой гардероб.

3. Рост в сантиметрах, X , и вес в килограммах, Y , взрослого мужчины является двумерным нормальным вектором $Z = (X, Y)$ с математическим ожиданием $\mathbb{E}(Z) = (175, 74)$ и ковариационной матрицей

$$\text{Var}(Z) = \begin{pmatrix} 49 & 28 \\ 28 & 36 \end{pmatrix}$$

Лишний вес характеризуется случайной величиной $U = X - Y$. Считается, что человек страдает избыточным весом, если $U < 90$.

- (a) Определите процент мужчин, чей рост отклоняется от среднего более, чем на 10 см.
- (b) Определите процент мужчин, чей вес отклоняется от среднего более, чем на 10 кг.
- (c) Каково распределение величины U ? Выпишите функцию плотности
- (d) Определите вероятность того, что человек страдает избыточным весом
- (e) Каково условное распределение веса при фиксированном росте? Выпишите функцию плотности
- (f) Какова вероятность того, что при росте 180 см человек будет обладать весом, меньшим 60 кг?

4. Аня, Боря и Вова сдают устный экзамен. Экзамен принимают два преподавателя. Время ответа каждого студента — экспоненциальная случайная величина со средним в 20 минут. Аня и Боря начали отвечать одновременно первыми. Вова начнет отвечать, как только кто-то из них освободиться. Длительности ответов независимы.
- Сколько времени пройдет в среднем от начала экзамена до первого ответившего?
 - Какова вероятность того, что Аня закончит отвечать позже всех?
 - Сколько в среднем времени пройдет от начала экзамена до окончания ответа Вовы?
5. Вася и Петя решают тест из 10 вопросов, на каждый вопрос есть ровно 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 равномерно. Вася списывает у Пети вопросы с 3-го по 7-ой, а остальные отвечает наугад равномерно.
- Пусть X — число правильных ответов Пети, а Y — число правильных ответов Васи.
- Найдите $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{E}(X - Y)$
 - Найдите $\text{Var}(X)$, $\text{Var}(Y)$, $\text{Cov}(X, Y)$, $\text{Var}(X - Y)$.
6. На плоскости закрашен круг с центром в нуле и единичным радиусом. Внутри этого круга равномерно случайно выбирается одна точка. Пусть X и Y — абсцисса и ордината этой точки.
- Выпишите совместную функцию плотности X и Y
 - Найдите частную функцию плотности X
 - Верно ли, что X и Y независимы?
 - Какова вероятность того, что $X + Y > 1$?
 - Найдите ожидаемое расстояние от точки до начала координат

11.4 Контрольная работа №2, 25.12.2013

Самая важная формула:

$$\frac{1}{(\sqrt{2\pi})^n \sqrt{\det(C)}} \cdot e^{-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)}$$

Неравенство Берри-Эссеена:

$$|\hat{F}_n(x) - \Phi(x)| \leq \frac{C_0 \mathbb{E}|X_n - \mu|^3}{\sigma^3 \sqrt{n}}, \quad 0.4 < C_0 < 0.48$$

Тест

- Зная распределение компонент случайного вектора всегда можно восстановить их совместное распределение. Да. Нет.
- Пусть X — длина наугад выловленного удава в сантиметрах, а Y — в дециметрах. Коэффициент корреляции между этими величинами равен 0.1. Да. Нет.
- Для любой случайной величины X (с конечной дисперсией) справедливо неравенство: $P(|X - \mathbb{E}(X)| > 2\sqrt{\text{Var}(X)}) \leq \frac{1}{4}$. Да. Нет.
- Сумма независимых нормальных случайных величин нормальна. Да. Нет.

5. Сумма n независимых равномерно распределенных на интервале $(0, 1)$ случайных величин асимптотически нормальна. Да. Нет.
6. Квадрат стандартной нормальной случайной величины имеет хи-квадрат распределение. Да. Нет.
7. Если ковариация компонент случайного двумерного нормального вектора равна нулю, то они независимы. Да. Нет.
8. Дисперсия суммы случайных величин всегда больше суммы их дисперсий. Да. Нет.
9. Центральная предельная теорема — частный случай теоремы Муавра-Лапласа. Да. Нет.
10. Математическое ожидание выборочной доли не зависит от объема выборки. Да. Нет.
11. «Математику уже затем учить надо, что она ум в порядок приводит» (М. В. Ломоносов) Да. Нет.

Задачи

1. (10) Совместная функция плотности с. в. (X, Y) имеет вид:

$$f(x, y) = \begin{cases} x + y & \text{при } x \in (0, 1), y \in (0, 1) \\ 0 & \text{иначе;} \end{cases}$$

Найдите:

- (a) $P(Y < X^2)$
 - (b) функцию плотности и математическое ожидание с. в. X
 - (c) условную функцию плотности и условное математическое ожидание с. в. X при условии, что $Y = 2$
2. (10) Случайный вектор $(X, Y)^T$ имеет двумерное нормальное распределение с математическим ожиданием $(0, 0)^T$ и ковариационной матрицей

$$C = \begin{pmatrix} 9 & -1 \\ -1 & 4 \end{pmatrix};$$

Найдите:

- (a) $P(X > 1)$
 - (b) $P(2X + Y > 3)$
 - (c) $P(2X + Y > 3 | X = 1)$
 - (d) $P\left(\frac{X^2}{9} + \frac{Y^2}{4} > 12\right)$
 - (e) Запишите совместную функцию плотности $(X, Y)^T$
3. Вычислите:

$$(a) P\left(\frac{X_1}{\sqrt{X_3^2 + X_4^2 + X_5^2}} > \frac{5}{4\sqrt{3}}\right)$$

$$(b) P\left(\frac{X_1 + 2X_2}{\sqrt{X_3^2 + X_4^2 + X_5^2}} < 4.5\right)$$

$$(c) P\left(\frac{X_1^2}{X_2^2 + X_3^2} > 17\right)$$

4. (15) Оценка за зачет по теории вероятности i -го студента — неотрицательная с. в. X_i с $\mathbb{E}(X_i) = \frac{1}{2}$ и $\text{Var}(X_i) = \frac{1}{12}$. Для случайной выборки из 36 студентов оцените или вычислите следующие вероятности $\left(\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i\right)$:

$$(a) P(|X_i - 0.5| \geq 0.3)$$

$$(b) P(X_i \geq 0.8)$$

$$(c) P(\bar{X} \geq 0.8)$$

Пусть дополнительно известно, что $X_i \sim U(0, 1)$:

$$(d) \text{Вычислите вероятность } \mathbb{P}(|X_i - 0.5| \geq 0.3)$$

$$(e) \text{Оцените погрешность вычисленной вероятности } \mathbb{P}(\bar{X} \geq 0.8)$$

$$(f) \text{Покажите, что средняя оценка за экзамен сходится по вероятности к } 0.5$$

5. При проведении социологических опросов в среднем 20 % респондентов отказываются отвечать на вопрос о личном доходе. Сколько нужно опросить человек, чтобы с вероятностью 0.99 выборочная доля отказавшихся отвечать на вопрос о доходе не превышала 0.25? Насколько изменится ответ на предыдущий вопрос, если средний процент отказывающихся отвечать неизвестен?
6. Оценки за контрольную работу по теории вероятностей 6 случайно выбранных студентов оказались равны: 8, 4, 5, 7, 3, 9.
- (a) Выпишите вариационный ряд;
- (b) Постройте выборочную функцию распределения;
- (c) Вычислите значение выборочного среднего и выборочной дисперсии.

11.5 Контрольная работа 3

Вычислите константы $B_1 = \{\text{Цифра, соответствующая первой букве Вашей фамилии}\}$ и $B_2 = \{\text{Цифра, соответствующая первой букве Вашего имени}\}$. Уровень значимости для всех проверяемых гипотез 0.0α , уровень доверия для всех доверительных интервалов $(1 - 0.0\alpha)$, где $\alpha = 1 + \{\text{остаток от деления } B_1 \text{ на } 5\}$.

А	Б	В	Г	Д	Е	Ж	З	И	К	Л	М	Н	О
1	2	3	4	5	6	7	8	9	10	11	12	13	14
П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Э	Ю	Я
15	16	17	18	19	20	21	22	23	24	25	26	27	28

1. Вес упаковки с лекарством является нормальной случайной величиной с неизвестными математическим ожиданием μ и дисперсией σ^2 . Контрольное взвешивание $(10 + B_1)$ упаковок показало, что выборочное среднее $\bar{X} = (50 + B_2)$, а несмещенная оценка дисперсии равна $B_1 \cdot B_2$. Постройте доверительные интервалы для математического ожидания и дисперсии веса упаковки (для дисперсии односторонний с нижней границей).
2. Экзамен принимают два преподавателя, случайным образом выбирая студентов. По выборкам из 85 и 100 наблюдений, выборочные доли не сдавших экзамен студентов составили соответственно $\frac{1}{B_1+1}$ и $\frac{1}{B_2+1}$. Можно ли утверждать, что преподаватели предъявляют к студентам одинаковый уровень требований? Вычислите минимальный уровень значимости, при котором основная гипотеза (уровень требований одинаков) отвергается (p-value).
3. Даны независимые выборки доходов выпускников двух ведущих экономических вузов А и В, по $(10 + B_1)$ и $(10 + B_2)$ выпускников соответственно: $\bar{X}_A = 45$, $\hat{\sigma}_A = 5$, $\bar{X}_B = 50$, $\hat{\sigma}_B = 6$. Предполагая, что распределение доходов подчиняется нормальному закону, проверьте гипотезу об отсутствии преимуществ выпускников вуза В.
4. По выборке независимых одинаково распределенных случайных величин X_1, \dots, X_n с функцией плотности $f(x) = \frac{1}{\theta} x^{-1+\frac{1}{\theta}}$, $x \in (0, 1)$, найдите оценки максимального правдоподобия параметра θ . Сформулируйте определения свойств несмещенности, состоятельности и эффективности и проверьте, выполняются ли эти свойства для найденной оценки.

Примечание. В помощь несчастным, забывшим формулу интегрирования по частям и таблицу неопределенных интегралов, или просто ленивым студентам:

$$\int_0^1 t^\alpha \ln(t) dt = -\frac{1}{(\alpha + 1)^2}$$

11.6 Контрольная работа 3, i-поток, 19.03.2014.

1. Дед Мазай подбирает зайцев. Предположим, что длина левого уха зайца имеет экспоненциальное распределение с плотностью $f(x) = a \exp(-ax)$ при $x \geq 0$. По 100 зайцам оказалось, что $\sum x_i = 2000$.
 - (a) Найдите оценку \hat{a} методом моментов
 - (b) Оцените стандартную ошибку $se(\hat{a})$
 - (c) Постройте 90%-ый доверительный интервал для неизвестного a
 - (d) На уровне значимости $\alpha = 0.05$ проверьте гипотезу $H_0: a = 15$ против $a > 15$. Найдите точное Р-значение.
2. По совету Лисы Волк опустил в прорубь хвост и поймал 100 чудо-рыб. Веса рыб независимы и имеют распределение Вейбулла, $f(x) = 2 \exp(-x^2/a^2) \cdot x/a^2$ при $x \geq 0$. Известно, что $\sum x_i^2 = 120$.
 - (a) Найдите оценку \hat{a} методом максимального правдоподобия
 - (b) Оцените стандартную ошибку $se(\hat{a})$
 - (c) Постройте 90%-ый доверительный интервал для неизвестного a
 - (d) На уровне значимости $\alpha = 0.05$ проверьте гипотезу $H_0: a = 1$ против $a > 1$. Найдите точное Р-значение.

3. [R] Как известно, Фрекен-Бок пьет коньяк по утрам и иногда видит привидения. За 110 дней имеются следующие статистические данные

Рюмок	1	2	3
Дней с привидениями	10	25	20
Дней без привидений	20	25	10

Вероятность увидеть привидение зависит от того, сколько рюмок коньяка было выпито утром, а именно, $p = \exp(a + bx)/(1 + \exp(a + bx))$, где x — количество рюмок, а a и b — неизвестные параметры.

- Оцените неизвестные параметры с помощью максимального правдоподобия.
 - На уровне значимости $\alpha = 0.05$ помощью теста отношения правдоподобия проверьте гипотезу о том, что одновременно $a = 0$ и $b = 0$. В чем содержательный смысл этой гипотезы? Найдите точное Р-значение.
4. Кот Васька поймал 5 воробьев, взвесил и отпустил. Предположим, что веса воробьев независимы и имеют нормальное распределение $N(\mu, \sigma^2)$. Известно, что $\sum x_i = 10$ и $\sum x_i^2 = 25$.
- Постройте 90% доверительный интервал для σ^2 , симметричный по вероятности
 - [R] Постройте самый короткий 90% доверительный интервал для σ^2
5. Задача о немецких танках. Всего выпущено неизвестное количество n танков. Для упрощения предположим, что на каждом танке написан его порядковый номер⁵. В бою было подбиты 4 танка с номерами 2, 5, 7 и 12.
- Найдите оценку общего выпуска танков n с помощью метода максимального правдоподобия
 - Является ли оценка максимального правдоподобия несмещенной?
 - Является ли максимум из номеров подбитых танков достаточной статистикой?
 - Является ли максимум из номеров подбитых танков полной статистикой?
 - Постройте с помощью оценки максимального правдоподобия несмещенную эффективную оценку неизвестного n .
6. Гражданин Фёдор решает проверить, не жульничает ли напёрсточник Афанасий, для чего предлагает Афанасию сыграть 5 партий в напёрстки. Фёдор решает, что в каждой партии будет выбирать один из трёх напёрстков наугад, не смотря на движения рук ведущего. Основная гипотеза: Афанасий честен, и вероятность правильно угадать напёрсток, под которым спрятан шарик, равна $1/3$. Альтернативная гипотеза: Афанасий каким-то образом жульничает (например, незаметно прячет шарик), так что вероятность угадать нужный напёрсток равна $1/5$. Статистический критерий: основная гипотеза отвергается, если Фёдор ни разу не угадает, где шарик.
- Найдите уровень значимости критерия
 - Найдите вероятность ошибки второго рода

⁵В реальности во время Второй мировой войны при оценке количества танков «Пантера» выпущенных в феврале 1944 использовались номера колес. Двух подбитых танков оказалось достаточно, чтобы оценить выпуск в 270 танков. По немецким архивам фактический объем выпуска оказался равен 276 танков.

7. [R] В службе единого окна 5 клиентских окошек. В каждое окошко стоит очередь. Я встал в очередь к окошку номер 5 ровно в 15:00, передо мной 5 человек. Предположим, что время обслуживания каждого клиента — независимые экспоненциальные величины с параметром λ . Первый человек с момента моего прихода был обслужен в окошке 1 в 15:05. Второй человек с момента моего прихода был обслужен в окошке 2 в 15:10.

- (a) Оцените с помощью максимального правдоподобия параметр λ
 (b) Оцените, сколько мне еще стоять в очереди.

11.7 Переписывание кр1, вариант 1

Задача 1. Вероятности попадания в мишень для трех стрелков равны $4/5$, $3/4$ и $2/3$ соответственно. В результате одновременного выстрела трех стрелков в мишени образовалось две пробоины. Какова вероятность того, что 3-ий стрелок попал в мишень?

Решение. Положим $A_i = \{\text{“попал } i\text{-й стрелок”}\}$, $i = 1, 2, 3$, и $B = (A_1^c \cap A_2 \cap A_3) \cup (A_1 \cap A_2^c \cap A_3) \cup (A_1 \cap A_2 \cap A_3^c)$. Имеем

$$\begin{aligned} \mathbb{P}(A_3|B) &= \frac{\mathbb{P}(A_3 \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_1^c \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2^c \cap A_3)}{\mathbb{P}(A_1^c \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2^c \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3^c)} = \\ &= \frac{\mathbb{P}(A_1^c)\mathbb{P}(A_2)\mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3)}{\mathbb{P}(A_1^c)\mathbb{P}(A_2)\mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3^c)} = \frac{7}{13} \end{aligned}$$

□

Задача 2. В лифт 11-этажного дома на первом этаже вошли 5 человек.

- (a) Найдите вероятность того, что хотя бы один из них выйдет на 6-ом этаже.
 (b) Вычислите среднее значение тех из них, кто не выйдет на 6-ом этаже.

Решение. (a) Рассмотрим случайные величины

$$X_i = \begin{cases} 1, & \text{если } i\text{-ый пассажир вышел на 6-ом этаже,} \\ 0, & \text{в противном случае,} \end{cases}$$

$i = 1, \dots, 5$. Поскольку в условии задачи не сказано ничего иного, считаем, что пассажиры ведут себя независимо друг от друга, и каждый из них может выйти из лифта на любом этаже со второго по одиннадцатый. Поэтому случайные величины X_1, \dots, X_5 независимы и $X_i \sim \text{Be}(1/10)$, $i = 1, \dots, 5$. Случайная величина $X := X_1 + \dots + X_5$ означает число пассажиров, которые вышли на 6-ом этаже. Тогда используя то, что $X \sim \text{Bi}(5, 1/10)$, получаем искомую вероятность в пункте (a)

$$\mathbb{P}(\{X > 0\}) = 1 - \mathbb{P}(\{X = 0\}) = 1 - C_5^0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^5 = 1 - \left(\frac{9}{10}\right)^5.$$

- (b) Заметим, что случайная величина $Y = 5 - X$ означает число пассажиров, которые не вышли на 6-ом этаже. Поэтому $\mathbb{E}[Y] = 5 - \mathbb{E}[X] = 5 - 5 \cdot \left(\frac{1}{10}\right) = \frac{9}{2}$. □

Задача 3. Пусть случайная величина X имеет функцию распределения

$$F_X(x) = \begin{cases} 0 & \text{при } x < -10, \\ 1/4 & \text{при } -10 \leq x < 0, \\ 3/4 & \text{при } 0 \leq x < 10, \\ 1 & \text{при } x \geq 10. \end{cases}$$

Найдите

- (a) $\mathbb{P}(\{X = -10\})$, $\mathbb{P}(\{X = 0\})$, $\mathbb{P}(\{X = 10\})$,
 (b) $\mathbb{E}[X]$,
 (c) $D(X)$.

Решение. (a) Известно, что для любого $a \in \mathbb{R}$ имеет место

$$\mathbb{P}(\{X = a\}) = F_X(a) - \lim_{n \rightarrow \infty} F_X(a - \frac{1}{n}).$$

Поэтому $\mathbb{P}(\{X = -10\}) = \frac{1}{4} - 0 = \frac{1}{4}$, $\mathbb{P}(\{X = 0\}) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$ и $\mathbb{P}(\{X = 10\}) = 1 - \frac{3}{4} = \frac{1}{4}$.

(b) Из пункта (a) следует, что распределение случайной величины X задается таблицей

X	-10	0	10
\mathbb{P}_X	$1/4$	$1/2$	$1/4$

Поэтому $\mathbb{E}[X] = -10 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 10 \cdot \frac{1}{4} = 0$.

(c) Наконец, $\mathbb{E}[X^2] = (-10)^2 \cdot \frac{1}{4} + 0^2 \cdot \frac{1}{2} + 10^2 \cdot \frac{1}{4} = 50$. Следовательно, $D(X) = \mathbb{E}[X^2] - [\mathbb{E}X]^2 = 50$.
 \square

Задача 4. Плотность распределения случайной величины X имеет вид

$$f_X(x) = \begin{cases} 0 & \text{при } x < 0, \\ x + 1/2 & \text{при } 0 \leq x \leq 1, \\ 0 & \text{при } x > 1. \end{cases}$$

Найдите

- (a) $\mathbb{P}(\{X = 1/2\})$, $\mathbb{P}(\{X \in [1/2; 2]\})$,
 (b) $F_X(x)$,
 (c) $\mathbb{E}[X]$,
 (d) $D(X)$.

Решение. (a) Известно, что если случайная величина X является абсолютно непрерывной, то для любого множества $B \subseteq \mathbb{R}$, для которого определена вероятность $\mathbb{P}(\{X \in B\})$, имеет место формула

$$\mathbb{P}(\{X \in B\}) = \int_B f_X(x) dx.$$

Поэтому $\mathbb{P}(\{X = 1/2\}) = \mathbb{P}(\{X \in [1/2; 1/2]\}) = \int_{1/2}^{1/2} f_X(x) dx = 0$ и

$$\mathbb{P}(\{X \in [1/2; 2]\}) = \int_{1/2}^2 f_X(x) dx = \int_{1/2}^1 (x + 1/2) dx = 5/8.$$

(b) Если $x < 0$, то

$$F_X(x) = \int_{-\infty}^x f_X(t) dt = \int_{-\infty}^x 0 dt = 0.$$

Если $0 \leq x \leq 1$, то

$$F_X(x) = \int_{-\infty}^0 f_X(t) dt + \int_0^x f_X(t) dt = \int_{-\infty}^0 0 dt + \int_0^x (t + 1/2) dt = \frac{t^2}{2} \Big|_{t=0}^{t=x} + \frac{x}{2} = \frac{x(x+1)}{2}.$$

Если $x > 1$, то

$$F_X(x) = \int_{-\infty}^0 f_X(t)dt + \int_0^1 f_X(t)dt + \int_1^x f_X(t)dt = \int_{-\infty}^0 0dt + \int_0^1 (t + 1/2)dt + \int_1^x 0dt = 1.$$

В итоге

$$F_X(x) = \begin{cases} 0 & \text{при } x < 0, \\ \frac{x(x+1)}{2} & \text{при } 0 \leq x \leq 1, \\ 1 & \text{при } x > 1. \end{cases}$$

$$(c) \mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x \left(x + \frac{1}{2}\right) dx = \int_0^1 \left(x^2 + \frac{x}{2}\right) dx = \frac{x^3}{3} \Big|_{x=0}^{x=1} + \frac{x^2}{4} \Big|_{x=0}^{x=1} = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}.$$

$$(d) \mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 x^2 \left(x + \frac{1}{2}\right) dx = \int_0^1 \left(x^3 + \frac{x^2}{2}\right) dx = \frac{x^4}{4} \Big|_{x=0}^{x=1} + \frac{x^3}{6} \Big|_{x=0}^{x=1} = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}.$$

$$\text{Следовательно, } D(X) = \frac{5}{12} - \frac{49}{144} = \frac{60-49}{144} = \frac{11}{144}. \quad \square$$

Задача 5. Совместное распределение случайных величин X и Y задано при помощи таблицы

	$Y = 1$	$Y = 2$	$Y = 3$
$X = 0$	0.2	0.1	0.2
$X = 1$	0.1	0.3	0.1

- (a) Являются ли случайные величины X и Y независимыми? Ответ обоснуйте.
- (b) Постройте графики функций распределения $F_X(x)$ и $F_Y(y)$.
- (c) Постройте таблицу распределения случайной величины XY .
- (d) Найдите $\mathbb{E}[X]$, $\mathbb{E}[Y]$, $\mathbb{E}[XY]$ и $\text{cov}(X, Y)$.
- (e) Являются ли случайные величины X и Y некоррелированными? Ответ обоснуйте.
- (f) Постройте таблицу условного распределения случайной величины Y при условии $\{X = 1\}$.
- (g) Найдите $\mathbb{E}[Y|\{X = 1\}]$.

Решение. (a) Случайные величины X и Y не являются независимыми, т.к., например, $0.2 = \mathbb{P}(\{X = 0\} \cap \{Y = 1\}) \neq \mathbb{P}(\{X = 0\}) \cdot \mathbb{P}(\{Y = 1\}) = 0.5 \cdot 0.3$.

(b) Таблицы распределения случайных величин X и Y имеют вид

X	0	1
\mathbb{P}_X	0.5	0.5

Y	1	2	3
\mathbb{P}_Y	0.3	0.4	0.3

Поэтому функции распределения равны

$$F_X(x) = \begin{cases} 0 & \text{при } x < 0, \\ 0.5 & \text{при } 0 \leq x < 1, \\ 1 & \text{при } x > 1, \end{cases}$$

$$F_Y(x) = \begin{cases} 0 & \text{при } x < 1, \\ 0.3 & \text{при } 1 \leq x < 2, \\ 0.7 & \text{при } 2 \leq x < 3, \\ 1 & \text{при } x > 3. \end{cases}$$

(с) Таблица распределения случайной величины XY имеет вид

XY	0	1	2	3
\mathbb{P}_{XY}	0.5	0.1	0.3	0.1

(d) $\mathbb{E}[X] = 0 \cdot 0.5 + 1 \cdot 0.5 = 0.5$, $\mathbb{E}[Y] = 1 \cdot 0.3 + 2 \cdot 0.4 + 3 \cdot 0.3 = 2$,
 $\mathbb{E}[XY] = 0 \cdot 0.5 + 1 \cdot 0.1 + 2 \cdot 0.3 + 3 \cdot 0.1 = 1$, $\text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 1 - 0.5 \cdot 2 = 0$.

Имеется также альтернативный способ подсчета $\mathbb{E}[XY]$, который использует совместное распределение случайных величин X и Y :

$$\mathbb{E}[XY] = 0 \cdot 1 \cdot 0.2 + 0 \cdot 2 \cdot 0.1 + 0 \cdot 3 \cdot 0.2 + 1 \cdot 1 \cdot 0.1 + 1 \cdot 2 \cdot 0.3 + 1 \cdot 3 \cdot 0.1 = 1.$$

(е) Поскольку $\text{cov}(X, Y) = 0$, то случайные величины X и Y являются некоррелированными.

(f) Находим условные вероятности

$$\mathbb{P}(\{Y = 1\}|\{X = 1\}) = \frac{\mathbb{P}(\{Y = 1\} \cap \{X = 1\})}{\mathbb{P}(\{X = 1\})} = \frac{0.1}{0.5} = 0.2,$$

$$\mathbb{P}(\{Y = 2\}|\{X = 1\}) = \frac{\mathbb{P}(\{Y = 2\} \cap \{X = 1\})}{\mathbb{P}(\{X = 1\})} = \frac{0.3}{0.5} = 0.6,$$

$$\mathbb{P}(\{Y = 3\}|\{X = 1\}) = \frac{\mathbb{P}(\{Y = 3\} \cap \{X = 1\})}{\mathbb{P}(\{X = 1\})} = \frac{0.1}{0.5} = 0.2.$$

Поэтому таблица условного распределения случайной величины Y при условии $\{X = 1\}$ имеет вид

Y	1	2	3
$\mathbb{P}_{Y \{X=1\}}$	0.2	0.6	0.2

(g)

$$\begin{aligned} \mathbb{E}[Y|X = 1] &= 1 \cdot \mathbb{P}(\{Y = 1\}|\{X = 1\}) + 2 \cdot \mathbb{P}(\{Y = 2\}|\{X = 1\}) + 3 \cdot \mathbb{P}(\{Y = 3\}|\{X = 1\}) = \\ &= 1 \cdot 0.2 + 2 \cdot 0.6 + 3 \cdot 0.2 = 2 \quad (43) \end{aligned}$$

□

11.8 Переписывание кр1, вариант 2

1. Вероятности попадания в цель для трех стрелков равны $4/5$, $3/4$ и $2/3$ соответственно. Для поражения цели в нее нужно попасть не менее двух раз. В результате одновременного выстрела трех стрелков цель была поражена. Какова вероятность того, что 3-й стрелок попал в цель?

Решение. Положим $A_i = \{\text{"}i\text{-й стрелок попал в цель"}\}$, $i = 1, 2, 3$, и $B = (A_1^c \cap A_2 \cap A_3) \cup (A_1 \cap A_2^c \cap A_3) \cup (A_1 \cap A_2 \cap A_3^c) \cup (A_1 \cap A_2 \cap A_3)$. Имеем

$$\begin{aligned} \mathbb{P}(A_3|B) &= \frac{\mathbb{P}(A_3 \cap B)}{\mathbb{P}(B)} = \\ &= \frac{\mathbb{P}(A_1^c \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2^c \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3)}{\mathbb{P}(A_1^c \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2^c \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3^c) + \mathbb{P}(A_1 \cap A_2 \cap A_3)} = \\ &= \frac{\mathbb{P}(A_1^c)\mathbb{P}(A_2)\mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)}{\mathbb{P}(A_1^c)\mathbb{P}(A_2)\mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3) + \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3^c) + \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)} = \frac{19}{25} \end{aligned} \quad (44)$$

□

2. Пусть X — число единиц, а Y — число шестерок, выпадающих при подбрасывании шести игральных костей. Найдите математическое ожидание и дисперсию суммы $X + Y$.

Решение. Положим

$$X_i = \begin{cases} 1, & \text{если при } i\text{-м подбрасывании выпала единица,} \\ 0, & \text{в противном случае,} \end{cases}$$

$$Y_i = \begin{cases} 1, & \text{если при } i\text{-м подбрасывании выпала шестерка,} \\ 0, & \text{в противном случае,} \end{cases}$$

$i = 1, \dots, 6$. Пусть $Z_i := X_i + Y_i$ и $Z := Z_1 + \dots + Z_6$. Имеем

$$\begin{aligned} \mathbb{E}[Z] &= \mathbb{E}[Z_1 + \dots + Z_6] = 6\mathbb{E}[Z_1] = 6\mathbb{E}[X_1 + Y_1] = 6\mathbb{E}[X_1] + 6\mathbb{E}[Y_1] = 6 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 2, \\ D[Z] &= D[Z_1 + \dots + Z_6] = 6D[Z_1] = 6D[X_1 + Y_1] = 6(D[X_1] + 2\text{cov}(X_1, Y_1) + D[Y_1]) = \\ &= 6(D[X_1] + 2\mathbb{E}[X_1 Y_1] - 2\mathbb{E}[X_1]\mathbb{E}[Y_1] + D[Y_1]) = 6\left(\frac{1}{6} \cdot \frac{5}{6} + 2 \cdot 0 - 2 \cdot \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{5}{6}\right) = \frac{4}{3}. \quad \square \end{aligned}$$

3. Пусть $\mathbb{E}[X] = -1$, $\mathbb{E}[Y] = 2$, $D[X] = 1$, $D[Y] = 2$, $\text{cov}(X, Y) = -1$. Найдите

- (a) $\mathbb{E}[2X + Y - 4]$,
- (b) $D[X + Y - 1]$,
- (c) $D[X - Y + 1]$,
- (d) $\text{cov}(X + Y, X - Y)$,
- (e) $\text{corr}(X + Y, X - Y)$.

Решение. (a) $\mathbb{E}[2X + Y - 4] = 2\mathbb{E}[X] + \mathbb{E}[Y] - 4 = -2 + 2 - 4 = -4$.

(b) $D[X + Y - 1] = D[X + Y] = D[X] + 2\text{cov}(X, Y) + D[Y] = 1 - 2 + 2 = 1$.

(c) $D[X - Y + 1] = D[X - Y] = D[X] - 2\text{cov}(X, Y) + D[Y] = 1 + 2 + 2 = 5$.

(d)

$$\begin{aligned} \text{cov}(X + Y, X - Y) &= \text{cov}(X, X) + \text{cov}(Y, X) - \text{cov}(X, Y) - \text{cov}(Y, Y) = D[X] - D[Y] = \\ &= 1 - 2 = -1. \end{aligned} \quad (45)$$

(e) $\text{corr}(X + Y, X - Y) = \frac{\text{cov}(X+Y, X-Y)}{\sqrt{D[X+Y]}\sqrt{D[X-Y]}} = \frac{-1}{\sqrt{1}\sqrt{5}} = -\frac{1}{\sqrt{5}}. \quad \square$

4. Плотность распределения случайной величины X имеет вид

$$f_X(x) = \begin{cases} 0 & \text{при } x < -3, \\ -x^2/36 + 1/4 & \text{при } -3 \leq x \leq 3, \\ 0 & \text{при } x > 3. \end{cases}$$

Найдите

- (a) $\mathbb{P}(\{X = 2\})$, $\mathbb{P}(\{X \in [0; 2]\})$,
- (b) $F_X(x)$,
- (c) $\mathbb{E}[X]$,
- (d) $D[X]$.

Решение. (a) Известно, что если случайная величина X является абсолютно непрерывной, то для любого множества $B \subseteq \mathbb{R}$, для которого определена вероятность $\mathbb{P}(\{X \in B\})$, имеет место формула

$$\mathbb{P}(\{X \in B\}) = \int_B f_X(x) dx.$$

Поэтому $\mathbb{P}(\{X = 2\}) = \mathbb{P}(\{X \in [2; 2]\}) = \int_2^2 f_X(x) dx = 0$ и $\mathbb{P}(\{X \in [0; 2]\}) = \int_0^2 f_X(x) dx = \int_0^2 \left(-\frac{x^2}{36} + \frac{1}{4}\right) dx = -\frac{x^3}{3 \cdot 36} \Big|_{x=0}^{x=2} + \frac{2}{4} = -\frac{2}{27} + \frac{1}{2} = \frac{23}{54}$.

(b) Если $x < -3$, то

$$F_X(x) = \int_{-\infty}^x f_X(t) dt = \int_{-\infty}^x 0 dt = 0.$$

Если $-3 \leq x \leq 3$, то

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{-3} f_X(t) dt + \int_{-3}^x f_X(t) dt = \int_{-\infty}^{-3} 0 dt + \int_{-3}^x \left(-\frac{t^2}{36} + \frac{1}{4}\right) dt = \\ &= -\frac{t^3}{3 \cdot 36} \Big|_{t=-3}^{t=x} + \frac{x+3}{4} = -\frac{x^3}{108} + \frac{x}{4} + \frac{1}{2}. \end{aligned}$$

Если $x > 3$, то

$$F_X(x) = \int_{-\infty}^{-3} f_X(t) dt + \int_{-3}^3 f_X(t) dt + \int_3^x f_X(t) dt = \int_{-\infty}^{-3} 0 dt + \int_{-3}^3 \left(-\frac{t^2}{36} + \frac{1}{4}\right) dt + \int_3^x 0 dt = 1.$$

В итоге

$$F_X(x) = \begin{cases} 0 & \text{при } x < -3, \\ -\frac{x^3}{108} + \frac{x}{4} + \frac{1}{2} & \text{при } -3 \leq x \leq 3, \\ 1 & \text{при } x > 3. \end{cases}$$

$$(c) \mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-3}^3 x \left(-\frac{x^2}{36} + \frac{1}{4}\right) dx = 0.$$

(d)

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_{-3}^3 x^2 \left(-\frac{x^2}{36} + \frac{1}{4}\right) dx = \int_{-3}^3 \left(-\frac{x^4}{36} + \frac{x^2}{4}\right) dx = -\frac{x^5}{5 \cdot 36} \Big|_{x=-3}^{x=3} + \frac{x^3}{3 \cdot 4} \Big|_{x=-3}^{x=3} = \\ &= \frac{9}{5} \quad (46) \end{aligned}$$

Следовательно, $D[X] = \frac{9}{5}$. \square

5. Пусть $\Omega = \{1, \dots, 8\}$, $\mathbb{P}(\{1\}) = \dots = \mathbb{P}(\{8\}) = 1/8$, $X(\omega) = \cos(\pi\omega/4)$ и $Y(\omega) = \sin(\pi\omega/4)$.

- (a) Являются ли случайные величины X и Y независимыми? Ответ обоснуйте.
- (b) Постройте графики функций распределения $F_X(x)$ и $F_Y(x)$.
- (c) Постройте таблицу распределения случайной величины XY .
- (d) Найдите $\mathbb{E}[X]$, $\mathbb{E}[Y]$, $\mathbb{E}[XY]$ и $\text{cov}(X, Y)$.
- (e) Являются ли случайные величины X и Y некоррелированными? Ответ обоснуйте.
- (f) Постройте таблицу условного распределения случайной величины Y при условии $\{X = \sqrt{2}/2\}$.
- (g) Найдите $\mathbb{E}[Y|\{X = \sqrt{2}/2\}]$.

Решение. Для дальнейшего решения случайные величины X , Y и XY удобно задать табличным способом:

Ω	1	2	3	4	5	6	7	8
X	$\frac{\sqrt{2}}{2}$	0	$-\frac{\sqrt{2}}{2}$	-1	$-\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	1
Y	$\frac{\sqrt{2}}{2}$	1	$\frac{\sqrt{2}}{2}$	0	$-\frac{\sqrt{2}}{2}$	-1	$-\frac{\sqrt{2}}{2}$	0
XY	$\frac{1}{2}$	0	$-\frac{1}{2}$	0	$\frac{1}{2}$	0	$-\frac{1}{2}$	0

- (a) Случайные величины X и Y не являются независимыми, т.к., например, $0 = \mathbb{P}(\{X = 1\} \cap \{Y = 1\}) \neq \mathbb{P}(\{X = 1\}) \cdot \mathbb{P}(\{Y = 1\}) = \frac{1}{8} \cdot \frac{1}{8}$.
- (b) Таблицы распределения случайных величин X и Y имеют вид

X	-1	$-\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	1
\mathbb{P}_X	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$
Y	-1	$-\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	1
\mathbb{P}_Y	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$

Поэтому функции распределения равны

$$F_X(x) = \begin{cases} 0 & \text{при } x < -1, \\ \frac{1}{8} & \text{при } -1 \leq x < -\frac{\sqrt{2}}{2}, \\ \frac{3}{8} & \text{при } -\frac{\sqrt{2}}{2} \leq x < 0, \\ \frac{5}{8} & \text{при } 0 \leq x < \frac{\sqrt{2}}{2}, \\ \frac{7}{8} & \text{при } \frac{\sqrt{2}}{2} \leq x < 1, \\ 1 & \text{при } x > 1, \end{cases}$$

$$F_Y(x) = \begin{cases} 0 & \text{при } x < -1, \\ \frac{1}{8} & \text{при } -1 \leq x < -\frac{\sqrt{2}}{2}, \\ \frac{3}{8} & \text{при } -\frac{\sqrt{2}}{2} \leq x < 0, \\ \frac{5}{8} & \text{при } 0 \leq x < \frac{\sqrt{2}}{2}, \\ \frac{7}{8} & \text{при } \frac{\sqrt{2}}{2} \leq x < 1, \\ 1 & \text{при } x > 1. \end{cases}$$

(с) Таблица распределения случайной величины XY имеет вид

XY	$-\frac{1}{2}$	0	$\frac{1}{2}$
\mathbb{P}_{XY}	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

(d) $\mathbb{E}[X] = -1 \cdot \frac{1}{8} - \frac{\sqrt{2}}{2} \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} + \frac{\sqrt{2}}{2} \cdot \frac{1}{4} + 1 \cdot \frac{1}{8} = 0$, $\mathbb{E}[Y] = -1 \cdot \frac{1}{8} - \frac{\sqrt{2}}{2} \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} + \frac{\sqrt{2}}{2} \cdot \frac{1}{4} + 1 \cdot \frac{1}{8} = 0$, $\mathbb{E}[XY] = -\frac{1}{2} \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = 0$, $\text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0 - 0 \cdot 0 = 0$.

(е) Поскольку $\text{cov}(X, Y) = 0$, то случайные величины X и Y являются некоррелированными.

(f) В случае, когда $X = \frac{\sqrt{2}}{2}$, случайная величина Y принимает значения $-\frac{\sqrt{2}}{2}$ и $\frac{\sqrt{2}}{2}$. Находим условные вероятности

$$\mathbb{P}\left(\left\{Y = -\frac{\sqrt{2}}{2}\right\} \mid \left\{X = \frac{\sqrt{2}}{2}\right\}\right) = \frac{\mathbb{P}\left(\left\{Y = -\frac{\sqrt{2}}{2}\right\} \cap \left\{X = \frac{\sqrt{2}}{2}\right\}\right)}{\mathbb{P}\left(\left\{X = \frac{\sqrt{2}}{2}\right\}\right)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4},$$

$$\mathbb{P}\left(\left\{Y = \frac{\sqrt{2}}{2}\right\} \mid \left\{X = \frac{\sqrt{2}}{2}\right\}\right) = \frac{\mathbb{P}\left(\left\{Y = \frac{\sqrt{2}}{2}\right\} \cap \left\{X = \frac{\sqrt{2}}{2}\right\}\right)}{\mathbb{P}\left(\left\{X = \frac{\sqrt{2}}{2}\right\}\right)} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}.$$

Поэтому таблица условного распределения случайной величины Y при условии $\left\{X = \frac{\sqrt{2}}{2}\right\}$ имеет вид

Y	$-\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$
$\mathbb{P}_{Y \{X=\frac{\sqrt{2}}{2}\}}$	$\frac{1}{2}$	$\frac{1}{2}$

(g)

$$\begin{aligned} \mathbb{E}\left[Y \mid \left\{X = \frac{\sqrt{2}}{2}\right\}\right] &= -\frac{\sqrt{2}}{2} \cdot \mathbb{P}\left(\left\{Y = -\frac{\sqrt{2}}{2}\right\} \mid \left\{X = \frac{\sqrt{2}}{2}\right\}\right) + \frac{\sqrt{2}}{2} \cdot \mathbb{P}\left(\left\{Y = \frac{\sqrt{2}}{2}\right\} \mid \left\{X = \frac{\sqrt{2}}{2}\right\}\right) = \\ &= -\frac{\sqrt{2}}{2} \cdot \frac{1}{2} + \frac{\sqrt{2}}{2} \cdot \frac{1}{2} = 0 \quad (47) \end{aligned}$$

□

11.9 Билеты к зачёту

1. Билет 1

- (a) Аксиоматика Колмогорова. Случайные величины. Функция распределения случайной величины и ее основные свойства. Функция плотности.
- (b) Виды сходимости последовательности случайных величин.

2. Билет 2

- (a) Основные дискретные распределения: биномиальное, Пуассона, гипергеометрическое, отрицательное биномиальное. Примеры непрерывных распределений (равномерное, экспоненциальное).
- (b) Неравенство Маркова и неравенство Чебышёва. Закон больших чисел.

3. Билет 3

- (a) Понятие о случайном векторе. Совместное распределение нескольких случайных величин. Независимость случайных величин. Маргинальные распределения. Условное распределение.
- (b) Центральная предельная теорема.

4. Билет 4

- (a) Математическое ожидание и дисперсия случайной величины и их свойства. Распределение функции от случайных величин.
- (b) Многомерное нормальное распределение и его свойства.

5. Билет 5

- (a) Математическое ожидание и ковариационная матрица случайного вектора. Коэффициент корреляции и его свойства.
- (b) Определение и свойства Хи-квадрат распределения, распределения Стьюдента и Фишера. Их основные свойства.

6. Билет 6

- (a) Условное распределение и условное математическое ожидание.
- (b) Теорема Муавра – Лапласа.

11.10 Экзамен, 26.03.2014

Часть 1

A posse ad esse non valet consequentia

1. Условная вероятность $\mathbb{P}(A \mid B)$ для независимых событий равна

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) $\frac{\mathbb{P}(B)}{\mathbb{P}(A \cap B)}$	5) $\mathbb{P}(A)$	Ответ:
--	--	---	---	--------------------	--------

2. События A и B называются независимыми, если

1) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ 2) $\mathbb{P}(A) \cdot \mathbb{P}(B) = \mathbb{P}(A \cap B)$ 3) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ 4) $\mathbb{P}(A \cap B) = 0$ 5) нет верного	Ответ:
---	--------

3. Вероятность опечатки в одном символе равна 0.01. Событие A — в слове из 5 букв будет 2 опечатки. Вероятность $P(A)$ примерно равняется

1) 0.0001	2) 0.001	3) 0.0004	4) 0.004	5) 0.04	Ответ:
-----------	----------	-----------	----------	---------	--------

4. В урне 3 белых и 2 черных шара. Случайным образом вынимается один шар, пусть X — число вынутых черных шаров. Величина $\mathbb{E}(X)$ равняется

1) 1	2) 0.5	3) 2/3	4) 2/5	5) 1/5	Ответ:
------	--------	--------	--------	--------	--------

5. В урне 3 белых и 2 черных шара. Случайным образом вынимается один шар, пусть X — число вынутых черных шаров. Величина $\text{Var}(X)$ равняется

1) $6/25$	2) $1/25$	3) $2/5$	4) $2/3$	5) $2/25$	Ответ:
-----------	-----------	----------	----------	-----------	--------

6. В урне 3 белых и 2 черных шара. Случайным образом вынимается один шар и откладывается в сторону, затем вынимается еще один шар. Событие A — второй шар — черный. Вероятность $\mathbb{P}(A)$ равняется

1) $6/25$	2) $1/25$	3) $2/5$	4) $2/3$	5) $2/25$	Ответ:
-----------	-----------	----------	----------	-----------	--------

7. Если $f(x)$ — функция плотности, то $\int_{-\infty}^{+\infty} f(u) du$ равен

1) 0	2) 1	3) $\mathbb{E}(X)$	4) $\text{Var}(X)$	5) $F(x)$	Ответ:
------	------	--------------------	--------------------	-----------	--------

8. Если $f(x)$ — функция плотности, то $\int_{-\infty}^x f(u) du$ равен

1) 0	2) 1	3) $\mathbb{E}(X)$	4) $\text{Var}(X)$	5) $F(x)$	Ответ:
------	------	--------------------	--------------------	-----------	--------

9. Если случайная величина X нормальна $N(0, 1)$ и $F(x)$ — это ее функция распределения, то $F(4)$ примерно равняется

1) 0	2) 0.25	3) 0.5	4) 0.75	5) 1	Ответ:
------	---------	--------	---------	------	--------

10. Дисперсия $\text{Var}(X)$ считается по формуле

1) $\mathbb{E}^2(X)$ 2) $\mathbb{E}(X^2)$ 3) $\mathbb{E}(X^2) + \mathbb{E}^2(X)$ 4) $\mathbb{E}(X^2) - \mathbb{E}^2(X)$ 5) $\mathbb{E}^2(X) - \mathbb{E}(X^2)$	Ответ:
--	--------

11. Дисперсия разности случайных величин X и Y вычисляется по формуле

1) $\text{Var}(X - Y) = \text{Var}(X) - \text{Var}(Y)$ 2) $\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$ 3) $\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2 \text{Cov}(X, Y)$ 4) $\text{Var}(X - Y) = \text{Var}(X) - \text{Var}(Y) + 2 \text{Cov}(X, Y)$ 5) $\text{Var}(X - Y) = \text{Var}(X) - \text{Var}(Y) - 2 \text{Cov}(X, Y)$	Ответ:
---	--------

12. Известно, что $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 2$, $\text{Var}(X) = 4$, $\text{Var}(Y) = 9$, $\text{Corr}(X, Y) = 0.5$. Дисперсия $\text{Var}(2X + 3)$ равняется

1) 16	2) 8	3) 11	4) 4	5) 19	Ответ:
-------	------	-------	------	-------	--------

13. Известно, что $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 2$, $\text{Var}(X) = 4$, $\text{Var}(Y) = 9$, $\text{Corr}(X, Y) = 0.5$. Ковариация $\text{Cov}(X, Y)$ равняется

1) 0.5	2) 18	3) 3	4) 12	5) 0	Ответ:
--------	-------	------	-------	------	--------

14. Известно, что $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 2$, $\text{Var}(X) = 4$, $\text{Var}(Y) = 9$, $\text{Corr}(X, Y) = 0.5$. Корреляция $\text{Corr}(2X + 3, 1 - Y)$ равняется

1) 1	2) -1	3) -0.5	4) 0.5	5) 0	Ответ:
------	-------	---------	--------	------	--------

15. Совместная функция распределения $F(x, y)$ двух случайных величин X и Y это

1) $\mathbb{P}(X \leq x)/\mathbb{P}(Y \leq y)$ 2) $\mathbb{P}(X \leq x) \cdot \mathbb{P}(Y \leq y)$ 3) $\mathbb{P}(X \leq x \mid Y \leq y)$ 4) $\mathbb{P}(X \leq x, Y \leq y)$ 5) $\mathbb{P}(X \leq x) + \mathbb{P}(Y \leq y)$	Ответ:
--	--------

16. Если случайная величина X , имеющая функцию плотности $a(x)$, и случайная величина Y , имеющая функцию плотности $b(y)$, независимы, то для их совместной функции плотности $f(x, y)$ справедливо

1) $f(x, y) = a(x) + b(y)$ 2) $f(x, y) = a(x)/b(y)$ 3) $f(x, y) = a(x)b(y)/(a(x) + b(y))$ 4) $f(x, y) = a(x) \cdot b(y)$ 5) $f(x, y) = \mathbb{E}(a(X)b(Y))$	Ответ:
--	--------

17. Случайные величины X и Y независимы и стандартно нормально распределены. Тогда $Z = X - 2Y$ имеет распределение

1) $N(0,1)$	2) t_2	3) $N(0,5)$	4) $N(0,2)$	5) $U[0;2]$	Ответ:
-------------	----------	-------------	-------------	-------------	--------

18. $Z_1, Z_2, \dots, Z_n \sim N(0, 1)$. Тогда величина $\frac{Z_1\sqrt{n-2}}{\sqrt{\sum_{i=3}^n Z_i^2}}$ имеет распределение

1) $N(0, 1)$	2) t_n	3) $F_{1,n-2}$	4) χ_n^2	5) t_{n-2}	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 X_1, X_2, \dots, X_n независимы и равномерно распределены $U[-\sqrt{3}, \sqrt{3}]$ то при $n \rightarrow \infty$ величина \bar{X}_n стремится по распределению к

1) вырожденному с $\mathbb{P}(X = 0) = 1$ 2) $U[-\sqrt{3}, \sqrt{3}]$ 3) $U[0; 1]$ 4) $N(0, 1)$ 5) $N(0, 3)$	Ответ:
--	--------

21. 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:
--------------	--------------	-------------------	-----------------------	-----------------------	--------

22. Последовательность оценок $\hat{\theta}_1, \hat{\theta}_2, \dots$ называется состоятельной, если

1) $\mathbb{E}(\hat{\theta}_n) = \theta$ 2) $\text{Var}(\hat{\theta}_n) \rightarrow 0$ 3) $\mathbb{P}(\hat{\theta}_n - \theta > t) \rightarrow 0$ для всех $t > 0$ 4) $\mathbb{E}(\hat{\theta}_n) \rightarrow \theta$ 5) $\text{Var}(\hat{\theta}_n) \geq \text{Var}(\hat{\theta}_{n+1})$	Ответ:
---	--------

23. Величины X_1, \dots, X_5 равномерны на отрезке $[0; 2a]$. Известно, что $\sum_{i=1}^5 x_i = 25$. При использовании первого момента оценка методом моментов неизвестного a равна

1) 1	2) 5	3) 10	4) 20	5) нет верного ответа	Ответ:
------	------	-------	-------	-----------------------	--------

24. При построении доверительного интервала для дисперсии по выборке из n наблюдений при неизвестном ожидании 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:
--------------	--------------	-------------------	---------------	----------	--------

25. Из 100 случайно выбранных человек ровно 50 ответили, что предпочитают молочный шоколад темному. Реализация 90% доверительного интервала для предпочтения молочного шоколада равна:

1) [0.42;0.58]	2) [0.45;0.55]	3) [0.30;0.70]	4) [0.49;0.51]	5) [0.48;0.52]	Ответ:
----------------	----------------	----------------	----------------	----------------	--------

26. При построении доверительного интервала для отношения дисперсий по двум независимым нормальным выборкам из n наблюдений каждая, используется статистика, имеющая распределение

1) $F_{n-1,n-1}$	2) t_{n-1}	3) χ^2_{n-1}	4) χ^2_n	5) t_n	Ответ:
------------------	--------------	-------------------	---------------	----------	--------

27. Функция правдоподобия, построенная по случайной выборке X_1, \dots, X_n из распределения с функцией плотности $f(x) = (\theta + 1)x^\theta$ при $x \in [0; 1]$ имеет вид

1) $(\theta + 1)x^{n\theta}$	2) $\sum(\theta + 1)x_i^\theta$	3) $(\theta + 1)^{\sum x_i}$	4) $(\sum x_i)^\theta$	5) $(\theta + 1)^n \prod x_i^\theta$	Ответ:
------------------------------	---------------------------------	------------------------------	------------------------	--------------------------------------	--------

28. Если P -значение меньше уровня значимости α , то гипотеза $H_0: \mu = \mu_0$

1) отвергается 2) не отвергается 3) отвергается только если $H_a: \mu \neq \mu_0$ 4) отвергается только если $H_a: \mu < \mu_0$ 5) недостаточно информации	Ответ:
--	--------

29. Смещенной оценкой математического ожидания по выборке независимых, одинаково распределенных случайных величин X_1, X_2, X_3 является оценка

1) $(X_1 + X_2)/2$ 2) $(X_1 + X_2 + X_3)/3$ 3) $0.7X_1 + 0.2X_2 + 0.1X_3$ 4) $0.3X_1 + 0.3X_2 + 0.3X_3$ 5) $X_1 + X_2 - X_3$	Ответ:
--	--------

30. Ошибкой первого рода является

1) Принятие неверной гипотезы 2) Отвержение основной гипотезы, когда она верна 3) Отвержение альтернативной гипотезы, когда она верна 4) Отказ от принятия любого решения 5) Необходимость пересдачи ТВ и МС	Ответ:
--	--------

Часть 2

1. У тети Маши — двое детей, один старше другого. Предположим, что вероятности рождения мальчика и девочки равны и не зависят от дня недели, а пол первого и второго ребенка независимы.

- (a) Известно, что старший ребенок — мальчик. Какова вероятность того, что у тети Маши есть ребенок-девочка?
 - (b) Известно, что хотя бы один ребенок — мальчик. Какова вероятность того, что у тети Маши есть ребенок-девочка?
 - (c) На вопрос: «А правда ли, что у вас есть сын, родившийся в пятницу?» тетя Маша ответила: «Да». Какова вероятность того, что у тети Маши есть ребенок-девочка?
2. 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 $Z = X - Y$.
- (a) Найдите $\mathbb{P}(X > 3)$
 - (b) Найдите $Var(X)$ и $Cov(X, Y)$
 - (c) Найдите $Corr(X, Z)$
3. Маша подкидывает 300 игровых кубиков. Те, что выпали не на шестёрку, она перекидывает один раз. Обозначим буквой N количество шестёрок на всех кубиках после возможных перекидываний.
- (a) Найдите $\mathbb{E}(N)$, $Var(N)$
 - (b) Какова примерно вероятность того, величина N лежит от 50 до 70?
 - (c) Укажите любой интервал, в который величина N попадает с вероятностью 0.9

Математическая статистика

4. Карл Магнусен сыграл 100 партий в шахматы. Из них он 40 выиграл, 30 проиграл и 30 раз сыграл вничью. Используя метод максимального правдоподобия или критерий Пирсона проверьте гипотезу о том, что все три исхода равновероятны на уровне значимости 5%.
5. Случайные величины X_1, X_2, \dots, X_{100} независимы и имеют пуассоновское распределение с неизвестным параметром λ . Известно, что $\sum X_i = 150$.
- (a) С помощью метода максимального правдоподобия постройте оценку для λ и 95%-ый доверительный интервал.
 - (b) Предположим, что сумма X_i неизвестна, зато известно, что количество ненулевых X_i равно 20. С помощью метода максимального правдоподобия постройте оценку для λ и 95%-ый доверительный интервал.
 - (c) Являются ли полученные оценки несмещенными?
6. 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 если известно что $c \in (10; 30)$?
 - (c) Каким должно быть c , чтобы вероятность ошибки второго рода равнялась 0,15?

12.1 Контрольная работа 1. Базовый поток. 30.10.2014

Тест

1. Если $A \cap B = \emptyset$, то A и B — независимые события. Да. Нет.
2. Попарно независимые события независимы в совокупности. Да. Нет.
3. $P(A \cap B) \leq P(A|B)$. Да. Нет.
4. Для любого числа a и любой случайной величины X : $P(X = a) > 0$. Да. Нет.
5. Любая неотрицательная функция $f(x)$, такая что $\int_{-\infty}^{+\infty} f(x)dx = 1$, может быть функцией плотности некоторой случайной величины. Да. Нет.
6. Функция распределения имеет не более, чем счетное, число точек разрыва. Да. Нет.
7. Для того, чтобы существовало математическое ожидание абсолютно непрерывной случайной величины достаточно, чтобы $\int_{-\infty}^{+\infty} xf(x)dx < \infty$. Да. Нет.
8. В Пуассоновском распределении математическое ожидание всегда совпадает с дисперсией. Да. Нет.
9. Математическое ожидание суммы случайных величин всегда равно сумме их математических ожиданий. Да. Нет.
10. Дисперсия суммы случайных величин всегда равна сумме их дисперсий. Да. Нет.
11. Хорошо, что не было вопросов про корреляцию. Да. Нет.

Задачи

1. Вася забыл какую-то (какую?) формулу. Он помнит, что она начинается с $P(A|B) =$. Дальше была дробь, три буквы P со скобками после них и в сумме по две буквы A и B внутри этих скобок. Ещё там была вертикальная черта «|». Из этих элементов Вася случайным образом составляет формулу.
 - (a) С какой вероятностью Вася напишет правильную формулу?
 - (b) Напишите формулу, которую забыл Вася.

Примечание: Вася всё-таки успел сходить на пару лекций по теории вероятностей и помнит, что $P(A|B)$ и $P(B|A)$ — это не одно и то же, «|» должна стоять именно между буквами (то ли $A|B$, то ли $B|A$), а в скобках, которые идут после P , должно хоть что-то стоять. При этом формула должна иметь смысл, то есть $P(A|B)$ не должна выражаться через себя же, и дробь не должна быть сократимой.

2. Точка с координатами (ξ, η) бросается наудачу в треугольник с вершинами $(1, 0)$, $(0, 0)$, $(0, 1)$. Сформулируйте определение независимости двух событий и проверьте, будут ли события $A = \{\xi < 1/2\}$ и $B = \{\eta < 1/2\}$ независимыми?
3. На учениях три самолёта одновременно и независимо атакуют цель. Известно, что первый самолёт поражает цель с вероятностью 0.6, второй — 0.4, третий — 0.3. При разборе учений выяснилось, что цель была поражена только одним самолётом. Какова вероятность того, что это был первый самолёт?

4. Книга в 500 страниц содержит 400 опечаток. Предположим, что каждая из них независимо от остальных опечаток может с одинаковой вероятностью оказаться на любой странице книги.
- (a) Определите вероятность того, что на 13-й странице будет не менее двух опечаток, в явном виде и с помощью приближения Пуассона.
 - (b) Определите наиболее вероятное число, математическое ожидание и дисперсию числа опечаток на 13-ой странице.
 - (c) Является ли 13-ая страница более «несчастливой», чем все остальные (в том смысле, что на 13-ой странице ожидается большее количество очепяток, чем на любой другой)?

Подсказка. Можно считать, что опечатки «выбирают» любую из страниц для своего появления независимо друг от друга. Успех заключается в выборе 13-ой страницы. Вероятность успеха?

5. Вероятность того, что медицинский тест выявит наличие заболевания, когда оно действительно есть, называется чувствительностью теста. Специфичностью теста называется вероятность того, что тест покажет отсутствие заболевания, когда пациент здоров. Вероятность того, что пациент болен, когда тест показал наличие заболевания, называется прогностической силой теста. Предположим, что только 1 % всего населения страдает данным заболеванием. Чувствительность используемого теста равна 0.9, а специфичность — 0.95.

- (a) Какова вероятность того, что у случайно выбранного человека тест покажет наличие заболевания?
- (b) Какова прогностическая сила теста? Что нужно сделать, чтобы её повысить?

6. Функция плотности случайной величины X имеет вид:

$$f(x) = \begin{cases} 1.5(x-a)^2 & , x \in [0, a] \\ 1.5(x+a)^2 & , x \in [-a, 0] \\ 0 & , x \notin [-a, a] \end{cases}$$

- (a) Найдите константу a , вероятность попадания в отрезок $[1/2, 2]$, математическое ожидание X и дисперсию случайной величины X .
 - (b) Нарисуйте функцию распределения случайной величины X .
7. Вася случайным образом посещает лекции по ОВП (Очень Важному Предмету). С вероятностью 0.9 произвольно выбранная лекция полезна, и с вероятностью 0.7 она интересна. Полезность и интересность — независимые друг от друга и от номера лекции свойства. Всего Вася прослушал 30 лекций.
- (a) Определите математическое ожидание и дисперсию числа полезных лекций и числа интересных лекций, прослушанных Васей.
 - (b) Определите математическое ожидание числа бесполезных и неинтересных лекций, прослушанных Васей, и числа лекций, обладающих хотя бы одним из свойств (полезность, интересность).

8. Пусть $\mathbb{E}[X] = 1$, $\mathbb{E}[Y] = 2$, $\mathbb{E}[X^2] = 5$, $\mathbb{E}[XY] = -1$. Найдите:

- (a) $\mathbb{E}[2X + Y - 4]$

	$X = 1$	$X = 2$
$Y = -1$	0.1	0.2
$Y = 0$	0.2	0.3
$Y = 1$	0	0.2

- (b) $\text{Var}[X], \text{Var}[Y]$
- (c) $\text{Cov}(X, Y), \text{corr}(X, Y)$
- (d) $\text{Var}[X - Y - 1], \text{Var}[X + Y + 1]$
- (e) $\text{Cov}(X - Y - 1, X + Y + 1), \text{corr}(X - Y - 1, X + Y + 1)$

9. Совместное распределение случайных величин X и Y задано в виде таблицы:

- (a) Найти частные распределения Y и Y^2
- (b) Найти ковариацию случайных величин X и Y
- (c) Можно ли утверждать, что случайные величины зависимы?

10. Бонусная задача Какова вероятность того, что наугад выбранный ответ на этот вопрос окажется верным (искомую вероятность вычислить и записать!)?

- (a) 0.25
- (b) 0.5
- (c) 0.6
- (d) 0.25

12.2 Решение кр 1. Базовый поток

1. Внимательно читайте примечание! Всего 6 возможных ситуаций, только 1 — благоприятная. Требуемая вероятность равна $1/6$.

2. Два события A и B независимы, если: $P(AB) = P(A)P(B)$.

Проверим, независимы ли события $A = \{\xi < 1/2\}$ и $B = \{\eta < 1/2\}$:

$P(AB)$ ищется как отношение площади квадрата с вершинами в $(0, 0)$; $(0, 1/2)$; $(1/2, 1/2)$; $(1/2, 0)$ к площади данного треугольника, т.е.:

$$P(AB) = \frac{(1/2)^2}{1/2} = \frac{1}{2}$$

$P(A)$ ищется как отношение площади трапеции с вершинами в $(0, 0)$; $(0, 1)$; $(1/2, 1/2)$; $(1/2, 0)$ к площади данного треугольника, т.е.:

$$P(A) = \frac{(1/2) \cdot (3/2) \cdot (1/2)}{1/2} = \frac{3}{4}$$

$P(B)$ ищется как отношение площади трапеции с вершинами в $(0, 0)$; $(1, 0)$; $(1/2, 1/2)$; $(0, 1/2)$ к площади данного треугольника, т.е.:

$$P(B) = \frac{(1/2) \cdot (3/2) \cdot (1/2)}{1/2} = \frac{3}{4}$$

$$P(A) \cdot P(B) = \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16} \neq \frac{1}{2} = P(AB)$$

Итак, события A и B зависимы.

3. Пусть событие $A = \{ \text{Цель была поражена первым самолетом} \}$, событие $B = \{ \text{Цель была поражена только одним самолетом} \}$. Тогда событие $AB = \{ \text{Первый самолет поразил цель, второй и третий — промахнулись} \}$. По формуле условной вероятности:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{0.6 \cdot 0.6 \cdot 0.7}{0.6 \cdot 0.6 \cdot 0.7 + 0.4 \cdot 0.4 \cdot 0.7 + 0.4 \cdot 0.6 \cdot 0.3} = \frac{0.252}{0.436} \approx 0.578$$

4. Удобно рассуждать следующим образом: предположим, что каждая опечатка наугад (**с равными вероятностями и независимо от других опечаток**) выбирает, на какую страницу ей попасть.

- (а) Пусть X - число опечаток на 13 странице.

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$$

$P(X = 0) = \left(\frac{499}{500}\right)^{400}$ — каждая из 400 опечаток не должна попасть на 13 страницу.

$P(X = 1) = 400 \cdot \frac{1}{500} \cdot \left(\frac{499}{500}\right)^{399}$ — ровно одна опечатка (а есть 400 вариантов) должна попасть на 13 страницу, а остальные — мимо. Соответственно:

$$P(X \geq 2) = 1 - \left(\frac{499}{500}\right)^{400} - 400 \cdot \frac{1}{500} \cdot \left(\frac{499}{500}\right)^{399} \approx 0.1911357$$

Это если считать в явном виде. А если пользоваться приближением Пуассона:

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

неплохо бы вспомнить, что параметр λ это математическое ожидание X , поэтому расчеты здесь пока оставим до лучших времен.

- (b) Пусть X - число опечаток на 13 странице. Введем случайную величину

$$X_i = \begin{cases} 1 & \text{если } i\text{-ая опечатка попала на 13 страницу} \\ 0 & \text{если нет} \end{cases}$$

Тогда $X = \sum_{i=1}^{400} X_i$. Рассмотрим отдельно X_i :

X_i	1	0
$P(X_i = \cdot)$	$\frac{1}{500}$	$\frac{499}{500}$

Так как i -ая опечатка наугад выбирает одну страницу из 500 и это должна быть именно 13.

Тогда:

$$\mathbf{E}[X_i] = \frac{1}{500} = \mathbf{E}[X_i^2] \Rightarrow$$

$$\Rightarrow \text{Var}(X_i) = \mathbf{E}[X_i^2] - (\mathbf{E}[X_i])^2 = \frac{1}{500} - \left(\frac{1}{500}\right)^2 = \frac{499}{500^2}$$

Значит

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{400} X_i\right] = \sum_{i=1}^{400} \mathbf{E}[X_i] = \frac{400}{500} = 0.8$$

$$\text{Var}(X) = \text{Var}\left(\sum_{i=1}^{400} X_i\right) = \sum_{i=1}^{400} \text{Var}(X_i) = 400 \cdot \frac{499}{500^2} = 0.8 \cdot \frac{499}{500}$$

Теперь мы знаем, что $\lambda = \mathbb{E}[X] = 0.8$ поэтому можем вернуться к пункту (а):

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - \frac{0.8^0}{0!} e^{-0.8} - \frac{0.8^1}{1!} e^{-0.8} = 0.1912079$$

So close!

Осталось найти наиболее вероятное число опечаток на 13 странице:

$$P(X = k) = \frac{0.8^k}{k!} e^{-0.8} \rightarrow \max_k$$

Очевидно, что эта функция убывает по k , ведь с ростом k :

$k!$ растёт, а 0.8^k убывает. Значит наиболее вероятное число ошибок — $X = 0$

- (с) **Ох уж эти предрассудки!** 13-я страница точно такая же как и все остальные, ведь везде в решении можно просто заменить номер 13 на любой другой и ничего не изменится.

5. Перепишем условие задачи:

Чувствительность теста =

= $P(\text{Медицинский тест показывает наличие заболевания} | \text{Заболевание есть})$

Специфичность теста =

= $P(\text{Медицинский тест показывает отсутствие заболевания} | \text{Заболевания нет})$

Прогностическая сила теста =

= $P(\text{Заболевание есть} | \text{Медицинский тест показывает наличие заболевания})$

$P(\text{Заболевание есть}) = 0.01 \Rightarrow P(\text{Заболевания нет}) = 0.99$

По условию чувствительность теста равна 0.9, тогда из формулы условной вероятности:

$$\begin{aligned} & P(\text{Мед. тест показывает наличие заб-ия} | \text{Заб-ие есть}) = \\ & \frac{P(\text{Мед. тест показывает наличие заб-ия, Заб-ие есть})}{P(\text{Заб-ие есть})} \Rightarrow \\ & \Rightarrow P(\text{Мед. тест показывает наличие заб-ия, Заб-ие есть}) = 0.9 \cdot 0.01 = 0.009 \end{aligned}$$

При этом очевидно, что:

$$\begin{aligned} & P(\text{Заболевание есть}) = P(\text{Мед. тест показывает наличие заб-ия, Заб-ие есть}) + \\ & + P(\text{Мед. тест показывает отсутствие заб-ия, Заб-ие есть}) \Rightarrow \\ & \Rightarrow P(\text{Мед. тест показывает отсутствие заб-ия, Заб-ие есть}) = 0.01 - 0.009 = 0.001 \end{aligned}$$

По условию специфичность теста равна 0.95, тогда из формулы условной вероятности:

$$P(\text{Мед. тест показывает отсутствие заб-ия} | \text{Заб-ия нет}) =$$

$$\frac{P(\text{Мед. тест пок-ет отсутствие заб-ия, Заб-ия нет})}{P(\text{Заб-ия нет})} \Rightarrow$$

$$P(\text{Мед. тест пок-ет отсутствие заб-ия, Заб-ия нет}) = 0.95 \cdot 0.99 = 0.9405$$

При этом очевидно, что:

$$\begin{aligned} P(\text{Заб-ия нет}) &= P(\text{Мед. тест пок-ет наличие заб-ия, Заб-ия нет}) + \\ &+ P(\text{Мед. тест пок-ет отсутствие заб-ия, Заб-ия нет}) \Rightarrow \\ \Rightarrow P(\text{Мед. тест пок-ет наличие заб-ия, Заб-ия нет}) &= 0.99 - 0.9405 = 0.0495 \end{aligned}$$

Теперь мы готовы отвечать на заданные вопросы:

•

$$\begin{aligned} P(\text{Мед. тест пок-ет наличие заб-ия}) &= \\ &= P(\text{Мед. тест пок-ет наличие заб-ия, Заб-ия нет}) + \\ &+ P(\text{Мед. тест пок-ет наличие заб-ия, Заб-ие есть}) = 0.009 + 0.0495 = 0.0585 \end{aligned}$$

• Прогностическая сила теста:

$$\begin{aligned} P(\text{Заболевание есть} | \text{Медицинский тест показывает наличие заболевания}) &= \\ = \frac{P(\text{Мед. тест пок-ет наличие заб-ия, Заб-ие есть})}{P(\text{Мед. тест пок-ет наличие заб-ия})} &= \frac{0.009}{0.0585} \approx 0.154 \end{aligned}$$

Для того, чтобы повысить прогностическую силу теста, необходимо понизить $P(\text{Мед. тест пок-ет наличие заб-ия, Заб-ие есть})$, а для этого необходимо повысить специфичность теста.

6. • Должно выполняться условие нормировки:

$$\begin{aligned} \int_{-a}^0 1.5(x+a)^2 dx + \int_0^a 1.5(x-a)^2 dx &= 1 \\ 0.5(x+a)^3 \Big|_{-a}^0 + 0.5(x-a)^3 \Big|_0^a &= 1 \\ 0.5a^3 + 0.5a^3 &= 1 \\ a &= 1 \end{aligned}$$

Теперь легко понять, как выглядит функция распределения (смотри определение функции распределения):

$$F(x) = \begin{cases} 0, & x < -1 \\ 0.5(x+1)^3, & -1 \leq x < 0 \\ 1 + 0.5(x-1)^3, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

$$P\left(X \in \left[\frac{1}{2}, 2\right]\right) = F(2) - F\left(\frac{1}{2}\right) = 1 - 1 + 0.5^4 = 0.5^4$$

$$\mathbb{E}(X) = \int_{-1}^0 x \cdot 1.5(x+1)^2 dx + \int_0^1 x \cdot 1.5(x-1)^2 dx =$$

$$\begin{aligned}
&= 1.5 \int_{-1}^0 (x^3 + 2x^2 + x) dx + 1.5 \int_0^1 (x^3 - 2x^2 + x) dx = \\
&= \frac{3}{8}x^4|_{-1}^0 + x^3|_{-1}^0 + \frac{3}{4}x^2|_{-1}^0 + \frac{3}{8}x^4|_0^1 - x^3|_0^1 + \frac{3}{4}x^2|_0^1 = -\frac{3}{8} + 1 - \frac{3}{4} + \frac{3}{8} - 1 + \frac{3}{4} = 0
\end{aligned}$$

А можно было заметить, что функция плотности — четная функция, поэтому сразу $\mathbb{E}(X) = 0$

Вычислим $\mathbb{E}(X^2)$:

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_{-1}^0 x^2 \cdot 1.5(x+1)^2 dx + \int_0^1 x^2 \cdot 1.5(x-1)^2 dx = \\
&= 1.5 \int_{-1}^0 (x^4 + 2x^3 + x^2) dx + 1.5 \int_0^1 (x^4 - 2x^3 + x^2) dx = \\
&= \frac{3}{10}x^5|_{-1}^0 + \frac{3}{4}x^4|_{-1}^0 + \frac{1}{2}x^3|_{-1}^0 + \frac{3}{10}x^5|_0^1 - \frac{3}{4}x^4|_0^1 + \frac{1}{2}x^3|_0^1 = \frac{1}{10}
\end{aligned}$$

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 0.1$$

- Верим, что график $F(x)$, выписанной выше, вы построить можете :)

7. Пусть $A = \{\text{«Лекция полезна»}\}$, $B = \{\text{«Лекция интересна»}\}$. Заметим, что лекции вообще независимы друг от друга.

- (а) Пусть X_A — число полезных лекций, прослушанных Васей, X_B — число интересных лекций, прослушанных Васей. Введем случайную величину:

$$X_i = \begin{cases} 1 & \text{если } i\text{-ая лекция была полезна} \\ 0 & \text{если нет} \end{cases}$$

Тогда $X_A = \sum_{i=1}^{30} X_i$. Рассмотрим отдельно X_i :

X_i	1	0
$P(X_i = \cdot)$	0.9	0.1

Вероятность 0.9 дана. Тогда:

$$\begin{aligned}
\mathbf{E}[X_i] &= 0.9 = \mathbb{E}[X_i^2] \Rightarrow \\
\Rightarrow \mathbf{Var}(X_i) &= \mathbf{E}[X_i^2] - (\mathbf{E}[X_i])^2 = 0.9 - 0.9^2 = 0.09
\end{aligned}$$

Значит

$$\begin{aligned}
\mathbb{E}[X_A] &= \mathbb{E}\left[\sum_{i=1}^{30} X_i\right] = \sum_{i=1}^{30} \mathbf{E}[X_i] = 0.9 \cdot 30 = 27 \\
\mathbf{Var}(X_A) &= \mathbf{Var}\left(\sum_{i=1}^{30} X_i\right) = \sum_{i=1}^{30} \mathbf{Var}(X_i) = 0.09 \cdot 30 = 2.7
\end{aligned}$$

Аналогично для числа интересных лекций можем получить:

$$\mathbb{E}[X_B] = 0.7 \cdot 30 = 21$$

$$\mathbf{Var}(X_A) = 0.21 \cdot 30 = 6.3$$

- (b) Так как интересность и полезность — независимые свойства лекций, то:
 $P(\bar{A} \cap \bar{B}) = P(\bar{A}) \cdot P(\bar{B}) = 0.3 \cdot 0.1 = 0.03$, где \bar{A} значит «не A ». В свою очередь:
 $P(A \cup B) = P(A \cap \bar{B}) + P(B \cap \bar{A}) + P(A \cap B) = 1 - P(\bar{A}) \cdot P(\bar{B}) = 0.97$, где $(A \cup B)$ значит « A или B », а $(A \cap B)$ — « A и B ». Аналогично, путем введения бинарной случайной величины можем получить:

$$\mathbb{E}[X_{\bar{A} \cap \bar{B}}] = 0.03 \cdot 30 = 0.9$$

$$\mathbb{E}[X_{A \cup B}] = 0.97 \cdot 30 = 29.1$$

8. Дано: $\mathbb{E}[X] = 1$, $\mathbb{E}[Y] = 2$, $\mathbb{E}[X^2] = 5$, $\mathbb{E}[Y^2] = 8$, $\mathbb{E}[XY] = -1$.

Будем использовать только свойства математического ожидания, ковариации и дисперсии, и ничего больше. Ни-че-го.

- $\mathbb{E}[2X + Y - 4] = 2\mathbb{E}[X] + \mathbb{E}[Y] + \mathbb{E}[-4] = 2 + 2 - 4 = 0$
- $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 5 - 1 = 4$
- $\text{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = 8 - 4 = 4$
- $\text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = -1 - 2 = -3$
- $\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = -\frac{3}{2 \cdot 2} = -0.75$
- $\text{Var}(X - Y - 1) = \text{Var}(X) + \text{Var}(Y) - 2\text{cov}(X, Y) = 4 + 4 - 2(-3) = 14$
- $\text{Var}(X + Y + 1) = \text{Var}(X) + \text{Var}(Y) + 2\text{cov}(X, Y) = 4 + 4 + 2(-3) = 2$
-

$$\begin{aligned} \text{cov}(X - Y - 1, X + Y + 1) &= \mathbb{E}[(X - Y)(X + Y)] - \mathbb{E}[X - Y]\mathbb{E}[X + Y] = \\ &= \mathbb{E}[X^2 - Y^2] - (\mathbb{E}[X] - \mathbb{E}[Y])(\mathbb{E}[X] + \mathbb{E}[Y]) = \mathbb{E}[X^2] - \mathbb{E}[Y^2] - ((\mathbb{E}[X])^2 - (\mathbb{E}[Y])^2) = \\ &= \text{Var}(X) - \text{Var}(Y) = 0 \quad (48) \end{aligned}$$

- $\text{cov}(X - Y - 1, X + Y + 1) = 0 \Rightarrow \text{cor}(X - Y - 1, X + Y + 1) = 0$

9.

	$X = 1$	$X = 2$
$Y = -1$	0,1	0,2
$Y = 0$	0,2	0,3
$Y = 1$	0	0,2

Найдем частные распределения Y и Y^2 :

	$X = 1$	$X = 2$	Σ
$Y = -1$	0,1	0,2	0.3
$Y = 0$	0,2	0,3	0.5
$Y = 1$	0	0,2	0.2
Σ	0,3	0,7	

Y	-1	0	1
$Pr(Y = \cdot)$	0.3	0.5	0.2

Так как Y^2 может принимать только значения 0 или 1:

Y^2	0	1
$Pr(Y^2 = \cdot)$	0.5	0.5

А ковариация:

$$\begin{aligned} cov(X, Y) = E[XY] - E[X]E[Y] &= ((-1) \cdot 1 \cdot 0,1 + (-1) \cdot 2 \cdot 0,2 + 1 \cdot 2 \cdot 0,2) - \\ &- (0,3 \cdot 1 + 0,7 \cdot 2) \cdot (0,3 \cdot (-1) + 0,1 \cdot 0,2) = 0,07 \quad (49) \end{aligned}$$

Так как $cov(X, Y) \neq 0$ — величины зависимы

усная задача Предположим, что правильный ответ 0.25. Но это невозможно, потому что вариантов ответа 0.25 — два (1 и 4), значит ответ 0.5 тоже был бы правильный. Предположим, что правильный 0.5. Тогда 0.25 тоже правильный — таких вариантов два из четырех, значит вероятность попасть в 0.25, выбрав ответ наугад, равна 0.5. Ответ 0.6, очевидно, неверен, потому что вероятность попасть в него равна 0.25.

Правильный ответ: 0

12.3 Праздник номер 1, i-поток, 30.10.2014

Часть 1

- Вася купил два арбуза у торговки тети Маши и один арбуз у торговки тети Оли. Арбузы у тети Маши спелые с вероятностью 90% (независимо друг от друга), арбузы у тети Оли спелые с вероятностью 70%.
 - Какова вероятность того, что все Васиные арбузы спелые?
 - Придя домой Вася выбрал случайным образом один из трех арбузов и разрезал его. Какова вероятность того, что это арбуз от тети Маши, если он оказался спелым?
 - Какова вероятность того, что второй и третий съеденные Васей арбузы были от тети Маши, если все три арбуза оказались спелыми?
- В большой большой стране живет очень большое количество n семей. Количества детей в разных семьях независимы. Количество детей в каждой семье — случайная величина с распределением заданным табличкой:

X_i	0	1	2	3
$\mathbb{P}()$	0.1	0.3	0.2	0.4

- Исследователь Афанасий выбирает одну семью из всех семей наугад, пусть X — число детей в этой семье. Найдите $E(X)$ и $\text{Var}(X)$.

- (b) Исследователь Бенедикт выбирает одного ребенка из всех детей наугад, пусть Y — число детей в семье этого ребёнка. Как распределена величина Y ? Что больше, $\mathbb{E}(Y)$ или $\mathbb{E}(X)$?

3. Функция плотности случайной величины X имеет вид

$$f(x) = \begin{cases} \frac{3}{8}x^2, & \text{если } x \in [0; 2] \\ 0, & \text{иначе} \end{cases}$$

- (a) Не производя вычислений найдите $\int_{-\infty}^{+\infty} f(x) dx$
(b) Найдите $\mathbb{E}(X)$, $\mathbb{E}(X^2)$ и дисперсию $\text{Var}(X)$
(c) Найдите $\mathbb{P}(X > 1.5)$, $\mathbb{P}(X > 1.5 \mid X > 1)$
(d) При каком c функция $g(x) = cx f(x)$ будет функцией плотности некоторой случайной величины?
4. Известно, что $\mathbb{E}(Z) = -3$, $\mathbb{E}(Z^2) = 15$, $\text{Var}(X + Y) = 20$ и $\text{Var}(X - Y) = 10$.
- (a) Найдите $\text{Var}(Z)$, $\text{Var}(4 - 3Z)$ и $\mathbb{E}(5 + 3Z - Z^2)$
(b) Найдите $\text{Cov}(X, Y)$ и $\text{Cov}(6 - X, 3Y)$
(c) Можно ли утверждать, что случайные величины X и Y независимы?

5. Листая сборник задач по теории вероятностей Вася наткнулся на задачу:

Какова вероятность того, что наугад выбранный ответ на этот вопрос окажется верным?
1) 0.25 2) 0.5 3) 0.6 4) 0.25

Чему же равна вероятность выбора верного ответа?

6. Книга в 500 страниц содержит 400 опечаток. Предположим, что каждая из них независимо от остальных опечаток может с одинаковой вероятностью оказаться на любой странице книги.
- (a) Определите вероятность того, что на 13-й странице будет не менее двух опечаток, в явном виде и с помощью приближения Пуассона.
(b) Определите наиболее вероятное число, математическое ожидание и дисперсию числа опечаток на 13-ой странице.
(c) Является ли 13-ая страница более «несчастливой», чем все остальные (в том смысле, что на 13-ой странице ожидается большее количество очепяток, чем на любой другой)?
7. Вася случайным образом посещает лекции по ОВП (Очень Важному Предмету). С вероятностью 0.9 произвольно выбранная лекция полезна, и с вероятностью 0.7 она интересна. Полезность и интересность — независимые друг от друга и от номера лекции свойства. Всего Вася прослушал 30 лекций.
- (a) Определите математическое ожидание и дисперсию числа полезных лекций, прослушанных Васей
(b) Определите математическое ожидание числа одновременно бесполезных и неинтересных лекций, прослушанных Васей, и математическое ожидание числа лекций, обладающих хотя бы одним из свойств (полезность, интересность)

8. Функция распределения случайной величины X задана следующей формулой:

$$F(x) = \frac{ae^x}{1 + e^x} + b$$

Определите: константы a и b , математическое ожидание и третий начальный момент случайной величины X , медиану и моду распределения.

Часть 2

1. Маша подкидывает кубик до тех пор, пока два последних броска в сумме не дадут⁶ 12. Обозначим случайные величины: N — количество бросков, а S — сумма набранных за всю игру очков.

(a) Найдите $\mathbb{P}(N = 2)$, $\mathbb{P}(N = 3)$

(b) Найдите $\mathbb{E}(N)$, $\mathbb{E}(S)$, $\mathbb{E}(N^2)$

(c) Пусть X_N — результат последнего броска. Как распределена случайная величина X_N ?

2. В столовую пришли 30 студентов и встали в очередь в случайном порядке. Среди них есть Вовочка и Машенька. Пусть V — это количество человек в очереди перед Вовочкой, а $M \geq 0$ — количество человек между Вовочкой и Машенькой.

(a) Найдите $\mathbb{P}(V = 1)$, $\mathbb{P}(M = 1)$, $\mathbb{P}(M = V)$

(b) Найдите $\mathbb{E}(V)$, $\mathbb{E}(M)$, $\text{Var}(M)$

3. Польский математик Стефан Банах имел привычку носить в каждом из двух карманов пальто по коробку спичек. Всякий раз, когда ему хотелось закурить трубку, он выбирал наугад один из коробков и доставал из него спичку. Первоначально в каждом коробке было по n спичек. Но когда-то наступает момент, когда выбранный наугад коробок оказывается пустым.

(a) Какова вероятность того, что в другом коробке в этот момент осталось ровно k спичек?

(b) Каково среднее количество спичек в другом коробке?

4. Производитель чудо-юдо-йогуртов наклеивает на каждую упаковку одну из 50 случайно выбираемых наклеек. Покупатель собравший все виды наклеек получает приз от производителя. Пусть X — это количество упаковок йогурта, которое нужно купить, чтобы собрать все наклейки.

Найдите $\mathbb{P}(X = 50)$, $\mathbb{E}(X)$, $\text{Var}(X)$

Hint: $\ln(50) \approx 3.91$, а $\sum_{i=1}^n \frac{1}{i} \approx \int_1^n \frac{1}{x} dx$:)

5. В самолете n мест и все билеты проданы. Первой в очереди на посадку стоит Сумасшедшая Старушка. Сумасшедшая Старушка несмотря на билет садиться на случайно выбираемое место. Каждый оставшийся пассажир садится на своё место, если оно свободно и на случайное выбираемое место, если его место уже кем-то занято.

(a) Какова вероятность того, что все пассажиры сядут на свои места?

(b) Какова вероятность того, что последний пассажир сядет на своё место?

(c) Чему примерно равно среднее количество пассажиров севших на свои места?

⁶Изначально вместо 12 задумывалось число 10, но опечатка была замечена поздно, поэтому решение приводится для 12.

12.4 Праздник номер 1 по теории вероятностей, i-поток. Решение

Часть 1

Не претендуя на единственность, решения претендуют на правильность!

1. (a) $P(\cdot) = 0.9^2 \cdot 0.7 = 0.567$
(b) $A = \{\text{случайно выбранный арбуз — от тети Маши}\}; B = \{\text{случайно выбранный арбуз оказался спелым}\}$. Формула условной вероятности:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{2/3 \cdot 0.9}{2/3 \cdot 0.9 + 1/3 \cdot 0.7} = \frac{18}{25}$$

- (c) $A = \{\text{второй и третий съеденные арбузы — от тети Маши}\}; B = \{\text{все три арбуза — спелые}\}$. Дает ли нам что-то о принадлежности арбузов к тете Маше или тете Оле то, что все арбузы — спелые? События независимы!

$$P(A|B) = P(A) = \frac{1}{3}$$

2. (a) $\mathbb{E}(X) = \sum P(X_i)X_i = 1.9$
 $\mathbb{V}ar(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 0 \cdot 0.1 + 1 \cdot 0.3 + 4 \cdot 0.2 + 9 \cdot 0.4 - 1.9^2 = 1.09$
(b) Раз ребенок выбран, значит, в его семье дети есть! Всего детей $n\mathbb{E}(X) = 1.9n$. Семей с одним ребенком — $0.3n$, значит, детей из семей с одним ребенком — $0.3n$. Аналогично, детей из семей с двумя детьми — $0.4n$; детей из семей с тремя детьми — $1.2n$.

Теперь легко построить закон распределения случайной величины Y :

Y	1	2	3
$P(Y)$	3/19	4/19	12/19

$$\mathbb{E}(Y) = \frac{3}{19} + \frac{8}{19} + \frac{36}{19} = \frac{47}{19} > \mathbb{E}(X)$$

3. Любителям (или нелюбителям) интегралов:

- (a) Да это же интеграл от функции плотности на всей числовой прямой! Ответ: единица!
(b)

$$\mathbb{E}(X) = \int_0^2 xf(x)dx = \int_0^2 \frac{3}{8}x^3dx = \frac{3}{32}x^4 \Big|_0^2 = \frac{3}{2}$$

$$\mathbb{E}(X^2) = \int_0^2 x^2 f(x)dx = \int_0^2 \frac{3}{8}x^4dx = \frac{3}{40}x^5 \Big|_0^2 = \frac{12}{5}$$

Формула дисперсии:

$$\mathbb{V}ar(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{12}{5} - \frac{9}{4} = \frac{3}{20}$$

(с)

$$P(X > 1.5) = \int_{1.5}^2 f(x)dx = \int_{1.5}^2 \frac{3}{8}x^2dx = \frac{1}{8}x^3|_{1.5}^2 = \frac{37}{64}$$

Вычислим вероятность условия:

$$P(X > 1) = \int_1^2 f(x)dx = \int_1^2 \frac{3}{8}x^2dx = \frac{1}{8}x^3|_1^2 = \frac{7}{8}$$

$$P(X > 1.5|X > 1) = \frac{P(X > 1.5)}{P(X > 1)} = \frac{37/64}{7/8} = \frac{37}{56}$$

(d) Должно выполняться следующее соотношение:

$$\int_{-\infty}^{+\infty} cx f(x)dx = 1$$

Применительно к нашей задаче:

$$\frac{3c}{8} \int_0^2 x^3 dx = \frac{3c}{32} x^4|_0^2 = \frac{3c}{2} = 1 \Rightarrow c = \frac{2}{3}$$

4. You have to learn the rules of the game. And then you have to play better than anyone else. (А. Эйнштейн)

(a)

$$\mathbb{V}ar(Z) = \mathbb{E}(Z^2) - (\mathbb{E}(Z))^2 = 15 - 9 = 6$$

$$\mathbb{V}ar(4 - 3Z) = 9\mathbb{V}ar(Z) = 54$$

$$\mathbb{E}(5 + 3Z - Z^2) = 5 + 3 \cdot (-3) - 15 = -19$$

(b)

$$\mathbb{V}ar(X \pm Y) = \mathbb{V}ar(X) + \mathbb{V}ar(Y) \pm 2 \cdot \mathbb{C}ov(X, Y)$$

Отсюда получаем:

$$\mathbb{V}ar(X + Y) - \mathbb{V}ar(X - Y) = 4\mathbb{C}ov(X, Y) \Rightarrow \mathbb{C}ov(X, Y) = 2.5$$

$$\mathbb{C}ov(6 - X, 3Y) = -3 \cdot 2.5 = -7.5$$

(с)

$$\mathbb{C}ov(X, Y) = 2.5 \neq 0$$

Случайные величины действительно независимы.

5. В условии не сказано сколько ответов являются верными. Предположим, что правильный ответ 0.25. Но это невозможно, потому что вариантов ответа 0.25 — два (1 и 4), значит ответ 0.5 тоже был бы правильный. Предположим, что правильный 0.5. Тогда 0.25 тоже правильный — таких вариантов два из четырех, значит вероятность попасть в 0.25, выбрав ответ наугад, равна 0.5. Ответ 0.6, очевидно, неверен, потому что вероятность попасть в него равна 0.25.

Правильный ответ: 0

6. Удобно рассуждать следующим образом: предположим, что каждая опечатка наугад (**с равными вероятностями и независимо от других опечаток**) выбирает, на какую страницу ей попасть¹.

(а) Пусть X - число опечаток на 13 странице.

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$$

$P(X = 0) = \left(\frac{499}{500}\right)^{400}$ — каждая из 400 опечаток не должна попасть на 13 страницу.
 $P(X = 1) = 400 \cdot \frac{1}{500} \cdot \left(\frac{499}{500}\right)^{399}$ — ровно одна опечатка (а есть 400 вариантов) должна попасть на 13 страницу, а остальные — мимо. Соответственно:

$$P(X \geq 2) = 1 - \left(\frac{499}{500}\right)^{400} - 400 \cdot \frac{1}{500} \cdot \left(\frac{499}{500}\right)^{399} \approx 0.1911357$$

Это если считать в явном виде. А если пользоваться приближением Пуассона:

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

неплохо бы вспомнить, что параметр λ это математическое ожидание X , поэтому расчеты здесь пока оставим до лучших времен.

(b) Пусть X — число опечаток на 13 странице. Введем случайную величину

$$X_i = \begin{cases} 1 & \text{если } i\text{-ая опечатка попала на 13 страницу} \\ 0 & \text{если нет} \end{cases}$$

Тогда $X = \sum_{i=1}^{400} X_i$. Рассмотрим отдельно X_i :

X_i	1	0
$P(X_i = \cdot)$	$\frac{1}{500}$	$\frac{499}{500}$

Так как i -ая опечатка наугад выбирает одну страницу из 500 и это должна быть именно 13.

Тогда:

$$\mathbf{E}[X_i] = \frac{1}{500} = \mathbf{E}[X_i^2] \Rightarrow$$

$$\Rightarrow \mathbf{Var}(X_i) = \mathbf{E}[X_i^2] - (\mathbf{E}[X_i])^2 = \frac{1}{500} - \left(\frac{1}{500}\right)^2 = \frac{499}{500^2}$$

Значит

$$\mathbf{E}[X] = \mathbf{E}\left[\sum_{i=1}^{400} X_i\right] = \sum_{i=1}^{400} \mathbf{E}[X_i] = \frac{400}{500} = 0.8$$

$$\mathbf{Var}(X) = \mathbf{Var}\left(\sum_{i=1}^{400} X_i\right) = \sum_{i=1}^{400} \mathbf{Var}(X_i) = 400 \cdot \frac{499}{500^2} = 0.8 \cdot \frac{499}{500}$$

Теперь мы знаем, что $\lambda = \mathbf{E}[X] = 0.8$ поэтому можем вернуться к пункту (а):

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - \frac{0.8^0}{0!} e^{-0.8} - \frac{0.8^1}{1!} e^{-0.8} = 0.1912079 \quad \boxed{\text{So close!}}$$

¹Ну очень самостоятельные!

Осталось найти наиболее вероятное число опечаток на 13 странице:

$$P(X = k) = \frac{0.8^k}{k!} e^{-0.8} \rightarrow \max_k$$

Очевидно, что эта функция убывает по k , ведь с ростом k :

$k!$ растет, а 0.8^k убывает. Значит наиболее вероятное число ошибок — $X = 0$

- (с) **Ох уж эти предрассудки!** 13-я страница точно такая же как и все остальные, ведь везде в решении можно просто заменить номер 13 на любой другой и ничего не изменится.



7. Пусть $A = \{\text{«Лекция полезна»}\}$, $B = \{\text{«Лекция интересна»}\}$. Заметим, что лекции вообще независимы друг от друга.

(а) Пусть X_A — число полезных лекций, прослушанных Васей, X_B — число интересных лекций, прослушанных Васей. Введем случайную величину:

$$X_i = \begin{cases} 1 & \text{если } i\text{-ая лекция была полезна} \\ 0 & \text{если нет} \end{cases}$$

Тогда $X_A = \sum_{i=1}^{30} X_i$. Рассмотрим отдельно X_i :

X_i	1	0
$P(X_i = \cdot)$	0.9	0.1

Вероятность 0.9 дана. Тогда:

$$\begin{aligned} \mathbf{E}[X_i] &= 0.9 = \mathbf{E}[X_i^2] \Rightarrow \\ \Rightarrow \mathbf{Var}(X_i) &= \mathbf{E}[X_i^2] - (\mathbf{E}[X_i])^2 = 0.9 - 0.9^2 = 0.09 \end{aligned}$$

Значит

$$\begin{aligned} \mathbf{E}[X_A] &= \mathbf{E}\left[\sum_{i=1}^{30} X_i\right] = \sum_{i=1}^{30} \mathbf{E}[X_i] = 0.9 \cdot 30 = 27 \\ \mathbf{Var}(X_A) &= \mathbf{Var}\left(\sum_{i=1}^{30} X_i\right) = \sum_{i=1}^{30} \mathbf{Var}(X_i) = 0.09 \cdot 30 = 2.7 \end{aligned}$$

Аналогично для числа интересных лекций можем получить:

$$\begin{aligned} \mathbf{E}[X_B] &= 0.7 \cdot 30 = 21 \\ \mathbf{Var}(X_B) &= 0.21 \cdot 30 = 6.3 \end{aligned}$$

(b) Так как интересность и полезность — независимые свойства лекций, то:
 $P(\bar{A} \cap \bar{B}) = P(\bar{A}) \cdot P(\bar{B}) = 0.3 \cdot 0.1 = 0.03$, где \bar{A} значит «не A ». В свою очередь:
 $P(A \cup B) = P(A \cap \bar{B}) + P(B \cap \bar{A}) + P(A \cap B) = 1 - P(\bar{A}) \cdot P(\bar{B}) = 0.97$, где $(A \cup B)$ значит « A или B ». Аналогично, путем введения бинарной случайной величины можем получить:

$$\begin{aligned} \mathbf{E}[X_{\bar{A} \cap \bar{B}}] &= 0.03 \cdot 30 = 0.9 \\ \mathbf{E}[X_{A \cup B}] &= 0.97 \cdot 30 = 29.1 \end{aligned}$$

8. Будем пользоваться свойствами функций распределения и плотности. Для начала:

$$\lim_{x \rightarrow +\infty} F(x) = 1, \quad \lim_{x \rightarrow -\infty} F(x) = 0,$$

$$\lim_{x \rightarrow +\infty} \left(\frac{ae^x}{1 + e^x} + b \right) = a + b := 1$$

$$\lim_{x \rightarrow -\infty} \left(\frac{ae^x}{1 + e^x} + b \right) = b := 0$$

Откуда сразу получаем

$$a = 1, b = 0 \Rightarrow F(x) = \frac{e^x}{1 + e^x}$$

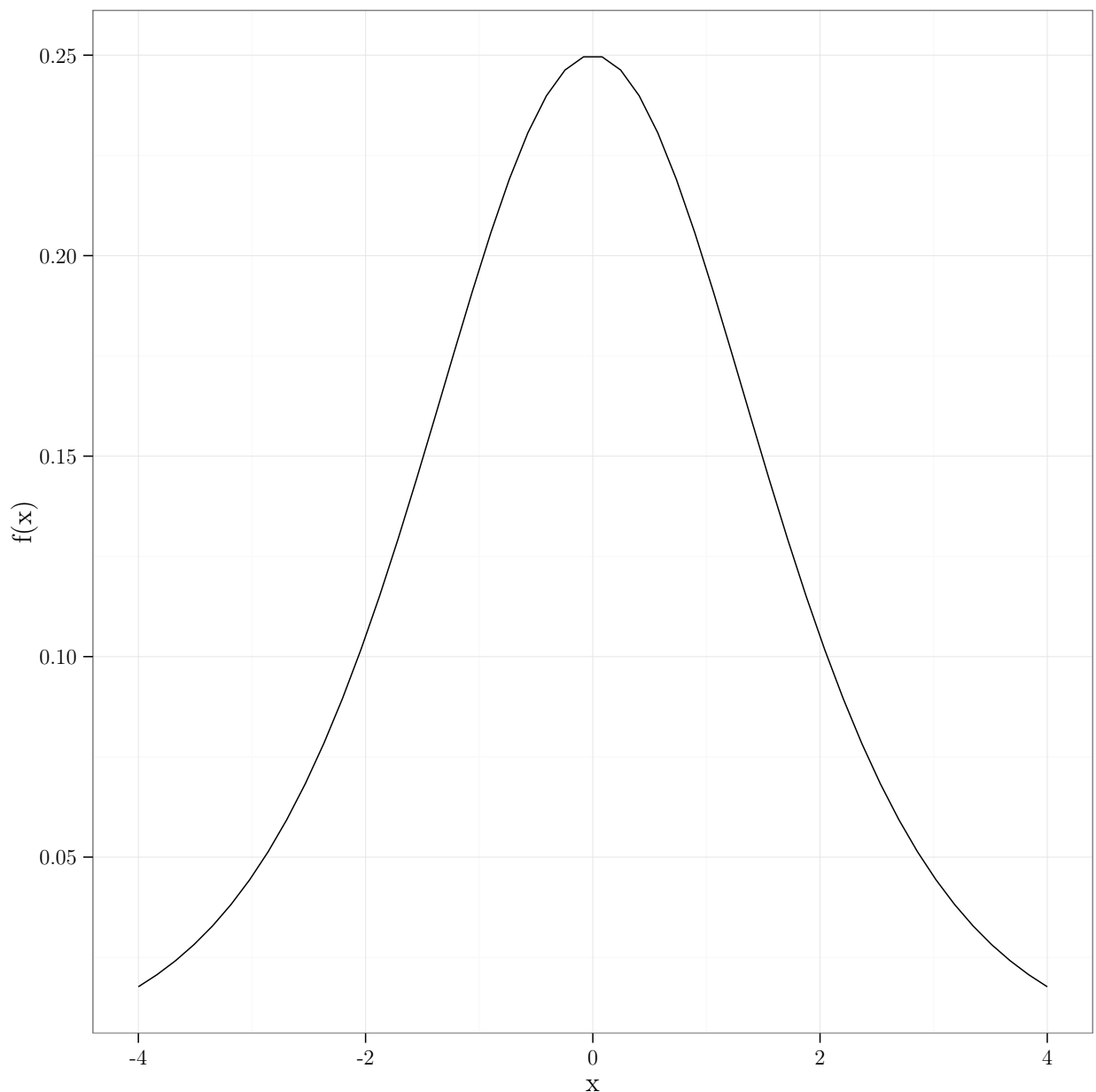
Для дальнейших развлечений нам понадобится функция плотности:

$$f(x) = F'(x) = \frac{e^x}{(1 + e^x)^2}$$

Заметим, что она симметрична относительно нуля:

$$f(-x) = \frac{\frac{1}{e^x}}{(1 + \frac{1}{e^x})^2} = \frac{e^x}{(1 + e^x)^2} = f(x)$$

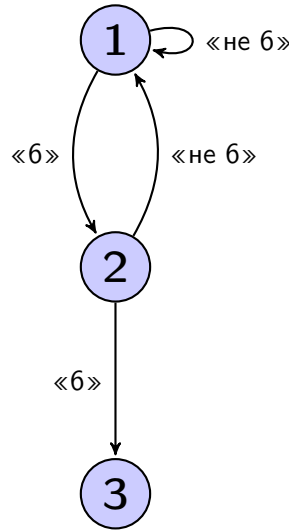
Из того этого следует, что **математическое ожидание, а так же мода и медиана равны нулю**. Более того, так как функция плотности симметрична относительно **нулевого** математического ожидания, **центральный и начальный моменты третьего порядка равны между собой и равны нулю**. Можно было выписать интегралы для математического ожидания и третьего начального момента и сослаться на нечетность функции.



Часть 2

— Это невозможно!

1. Алгоритм решения: рисуешь дерево \rightarrow PROFIT



Комментарии к построению дерева: состояние 1 — начальное, состояние 3 — конец игры, когда выпало две «шестерки» подряд. Заметим, что выпадение любой «нешестерки» в процессе игры приводит нас к состоянию, эквивалентному начальному.

Вероятность выпадения «шестерки» равна $1/6$, «нешестерки» — $5/6$.

Теперь мы готовы оседлать коня!

(a) $P(N = 1) = 0$ — невозможно за ход закончить игру.

$$P(N = 2) = \frac{1}{36}$$

$$P(N = 3) = \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{5}{216}$$

(b) А теперь будет видна вся сила рисования дерева:

Пусть \mathbb{E}_1 — число ходов, за которое мы ожидаем закончить игру, если игра начинается в состоянии 1, \mathbb{E}_2 — число ходов, за которое мы ожидаем закончить игру, если игра начинается в состоянии 2.

Получим два уравнения:
$$\begin{cases} \mathbb{E}_2 = \frac{1}{6} \cdot 1 + \frac{5}{6}(\mathbb{E}_1 + 1) \\ \mathbb{E}_1 = \frac{5}{6}(\mathbb{E}_1 + 1) + \frac{1}{6}(\mathbb{E}_2 + 1) \end{cases}$$

Решив эту систему, получим, что $\mathbb{E}_1 = 42$. А ведь это и есть $\mathbb{E}(N)$.

Аналогична логика для оставшихся мат. ожиданий.

Найдем математическое ожидание суммы набранных очков. Ясно, что если выпадает «не 6», то мы ждем 3 очка. Тогда переопределив \mathbb{E}_1 и \mathbb{E}_2 следующим образом: пусть \mathbb{E}_1 — число набранных очков, которое мы ожидаем получить за игру, если игра начинается в состоянии 1, \mathbb{E}_2 — число набранных очков, которое мы ожидаем получить за игру, если игра начинается в состоянии 2.

Новые два уравнения:
$$\begin{cases} \mathbb{E}_2 = \frac{1}{6} \cdot 6 + \frac{5}{6}(\mathbb{E}_1 + 3) \\ \mathbb{E}_1 = \frac{5}{6}(\mathbb{E}_1 + 3) + \frac{1}{6}(\mathbb{E}_2 + 6) \end{cases}$$

Решаем и получаем: $\mathbb{E}(S) = \mathbb{E}_1 = 147$

А можно было сделать еще круче! Выше показано, что $\mathbb{E}(N) = 42$. А сколько мы ждем очков за 1 ход? 3.5! Тогда $\mathbb{E}(S) = \mathbb{E}(N) \cdot 3.5 = 147$

Применяя схожую логику для $\mathbb{E}(N^2)$:

$$\mathbb{E}(N^2) = \frac{5}{6} \cdot \mathbb{E}((N+1)^2) + \frac{1}{6} \cdot \frac{5}{6} \cdot \mathbb{E}((N+2)^2) + \frac{1}{6} \cdot \frac{1}{6} \cdot 2^2$$

Учитывая, что $\mathbb{E}(N) = 42$, получим: $\mathbb{E}(N^2) = 3414$.

(с) Veni, vidi, vici

X_n	6
$P(X_n)$	1

2. (а) $P(V = 1) = 1/30$, т.к. именно этому равна вероятность того, что Вовочка стоит ровно вторым в очереди;

$M = 1$ значит, что между Машенькой и Вовочкой ровно один человек в очереди. Если Вовочка находится от 3 (включительно) до 28 позиции в очереди, то для Машеньки есть две благоприятные позиции для события $M = 1$ (например, если Вовочка стоит на 15 месте, то благоприятные позиции для Машеньки — стоять либо 13-ой, либо 17-ой). Если же Вовочка стоит на других позициях в очереди, то для Машеньки существует ровно одна благоприятная позиция:

$$P(M = 1) = \frac{26}{30} \cdot \frac{2}{29} + \frac{4}{30} \cdot \frac{1}{29} = \frac{56}{30 \cdot 29} = \frac{28}{435}$$

$M = V$ произойдет только, если Машенька стоит за Вовочкой. При этом для Машеньки существует только одна благоприятная позиция и только в том случае, что Вовочка стоит до 15 позиции (включительно):

$$P(M = V) = \frac{1}{2} \cdot \frac{1}{29} = \frac{1}{58}$$

(b)

$$\mathbb{E}(V) = \frac{0 + 1 + \dots + 29}{30} = \frac{30 \cdot 14 + 15}{30} = 14.5$$

Для $\mathbb{E}(M)$ можно решить в лоб, и получится красивая сумма, а можно вот так:

Сначала случайно кинем Вовочку и Машеньку на две из 30 позиций в очереди. Образуется три отрезка: точки между Вовочкой и Машенькой и два крайних отрезка (может быть, отрезок из 0 точек). Затем будем закидывать в очередь на оставшиеся позиции случайно 28 оставшихся людей (назовем их «пропавшими»). Т.к. все броски были случайны (или из соображений симметрии, как хотите), вероятность попасть в отрезок между Машенькой и Вовочкой для «пропавшего» равна $1/3$, вне отрезка — соответственно $2/3$, и независима от остальных бросков (!).

Введем случайную величину X_i для i -го «пропавшего», которая равна 1, если он попал в отрезок между Машенькой и Вовочкой, 0, если не попал:

X_i	1	0
$P(X_i)$	1/3	2/3

Легко считается: $\mathbb{E}(X_i) = 1/3$, $\mathbb{E}(X_i^2) = 1/3$, $\text{Var}(X_i) = 1/3 - 1/9 = 2/9$. Ясно, что $M = \sum_1^{28} X_i$. Тогда учитывая независимость X_i :

$$\mathbb{E}(M) = \frac{28}{3}$$

$$\mathbb{V}ar(M) = \frac{56}{9}$$

3. Биномиальное распределение — *À l'abordage!*

Задача интерпретируется так: последний ход — это когда мы обратились к коробку, в котором нет спичек (т.е. к одному коробку нужно обратиться $n + 1$ раз).

- (а) Если $0 < k \leq n$, будем считать успехом — попадание в коробок, к которому мы на последнем ходу игры (пустому коробку) обратились. До этого момента из него было вытащено n спичек, а из другого $2n - k$ спичек, т.е. всего в игре было $2n - k + 1$ шагов. Успехов — $n + 1$ (вытащено n спичек, и на последнем ходу мы к нему обратились). По формуле Бернулли получаем следующее (X — случайная величина, показывающая сколько спичек осталось в коробке, к которому мы не обратились на последнем ходу игры):

$$P(X = k) = C_{2n-k+1}^{n+1} \left(\frac{1}{2}\right)^{2n-k+1}$$

Если $k = 0$, то мы вытащили все спички из обоих коробков к последнему ходу, и нам без разницы к какому коробку мы обратимся на последнем шагу, т.е.:

$$P(X = 0) = 2C_{2n+1}^{n+1} \left(\frac{1}{2}\right)^{2n+1}$$

- (b) Среднее спичек в другом коробке:

$$\mathbb{E}(X) = \sum_{k=1}^n k \cdot C_{2n-k+1}^{n+1} \left(\frac{1}{2}\right)^{2n-k+1}$$

4. Для того чтобы количество упаковок, которые необходимо купить, равнялось 50, нужно чтобы ни одну из наклеек Покупатель не встретил дважды, поэтому:

$$P(X = 50) = 1 \cdot \frac{49}{50} \cdot \frac{48}{50} \cdot \dots \cdot \frac{1}{50} = \frac{49!}{50!} \approx 3.4 \cdot 10^{-21}$$

2

Теперь введем понятие «шаг». Переход на новый шаг происходит в тот момент, когда покупатель получил наклейку, которой у него раньше не было. Начинаем с шага 0, когда нет ни одной наклейки, и шагать будем до 49, потому что в момент перехода на шаг 50 Покупатель получит последнюю необходимую наклейку и «прогулка» закончится. Введем случайную величину X_q равную количеству покупок в течение шага номер q . Тогда $X = \sum_{q=0}^{49} X_q$. Найдем математическое ожидание X_q :

$$\mathbb{E}[X_q] = \frac{n-q}{n} \cdot 1 + \frac{q}{n} \cdot \frac{n-q}{n} \cdot 2 + \left(\frac{q}{n}\right)^2 \cdot \frac{n-q}{n} \cdot 3 + \dots$$

здесь $\frac{n-q}{n}$ — это вероятность найти наклейку, которой еще нет, а $\frac{q}{n}$, соответственно — вероятность повториться. Вопрос теперь в том, как посчитать сумму:

$$\mathbb{E}[X_q] = \frac{n-q}{n} \left(1 + \frac{q}{n} \cdot 2 + \left(\frac{q}{n}\right)^2 \cdot 3 + \dots\right) = \frac{n-q}{n} \cdot \sum_{k=0}^{\infty} \left(\frac{q}{n}\right)^k (k+1)$$

²Надежда умирает последней!

Можем выписать в столбик несколько первых членов вышестоящей суммы:

$$\begin{array}{c}
 1 \\
 \left(\frac{q}{n}\right)^1 + \left(\frac{q}{n}\right)^1 \\
 \left(\frac{q}{n}\right)^2 + \left(\frac{q}{n}\right)^2 + \left(\frac{q}{n}\right)^2 \\
 \left(\frac{q}{n}\right)^3 + \left(\frac{q}{n}\right)^3 + \left(\frac{q}{n}\right)^3 + \left(\frac{q}{n}\right)^3 \\
 \dots\dots\dots
 \end{array}$$

Достаточно! Можем скомпоновать всю сумму другим способом, а именно — по столбцам. Заметим, что сумма элементов в каждом столбце это сумма бесконечно убывающей геометрической прогрессии с одним и тем же знаменателем $\frac{q}{n}$ и различными первыми членами. Соответственно:

$$\begin{aligned}
 \sum_{k=0}^{\infty} \left(\frac{q}{n}\right)^k (k+1) &= \frac{1}{1-\frac{q}{n}} + \frac{\frac{q}{n}}{1-\frac{q}{n}} + \frac{\left(\frac{q}{n}\right)^2}{1-\frac{q}{n}} + \frac{\left(\frac{q}{n}\right)^3}{1-\frac{q}{n}} + \dots = \\
 &= \frac{1}{1-\frac{q}{n}} \left(1 + \frac{q}{n} + \left(\frac{q}{n}\right)^2 + \left(\frac{q}{n}\right)^3 + \dots\right) = \frac{n}{n-q} \cdot \frac{n}{n-q} = \left(\frac{n}{n-q}\right)^2
 \end{aligned}$$

Таким образом, получаем, что:

$$\mathbb{E}[X_q] = \frac{n-q}{n} \cdot \left(\frac{n}{n-q}\right)^2 = \frac{n}{n-q}$$

и это верно для любого q !

$$\begin{aligned}
 \mathbb{E}[X] &= \mathbb{E}\left[\sum_{q=0}^{49} X_q\right] = \sum_{q=0}^{49} \mathbb{E}[X_q] = \frac{50}{50-0} + \frac{50}{50-1} + \dots + \frac{50}{50-49} = 50 \left(\frac{1}{50} + \frac{1}{49} + \dots + 1\right) \approx \\
 &\approx 50 \int_1^{50} \frac{1}{x} dx = 50 \ln(50) \approx 195.5
 \end{aligned}$$

А теперь ещё одно решение:

Величины X_q независимы (но по разному распределены). Если долго пришлось ждать i -го шага, это ничего не говорит о j -ом шаге. Величины X_q имеют известный закон распределения — это число опытов до первого успеха при заданной вероятности успеха. Это геометрическое распределение, математическое ожидание которого равно $\frac{1}{p}$, а дисперсия: $\frac{1-p}{p^2}$, где p — вероятность успеха.

А те, кто забыл, могут **проще решить** методом первого шага: Если X — число опытов до успеха при вероятности успеха p , то

$$\mathbb{E}[X] = p \cdot 1 + (1-p) \cdot \mathbb{E}[X+1]$$

Откуда $\mathbb{E}[X] = \frac{1}{p}$ и дело в шляпе :) Аналогично:

$$\mathbb{E}[X^2] = p \cdot 1^2 + (1 - p) \cdot \mathbb{E}[(X + 1)^2]$$

и решая, находим $\mathbb{E}[X^2]$.

5. (а) Необходимое и достаточное условие — старушка не должна занять чужое место. С вероятностью $\frac{1}{n}$ она угадает свое место, значит, для каждого входящего его место будет свободно и он туда сядет.

Ответ: $\frac{1}{n}$

- (б) Будем искать вероятность того, что последний человек не сядет на свое место.

Пусть $A_i = \{\text{Старушка села на место } i\text{-го}\}$, $B_{(i,j)} = \{i\text{-ый пассажир сел на место } j\text{-ого}\}$

$$\begin{aligned} P[\text{n-ый не сядет на свое место}] &= P(A_n) + P[A_{n-1}]P[B_{(n-1,n)}] + \\ &+ P[A_{n-2}](P[B_{(n-2,n)}] + P[B_{(n-2,n-1)}]P[B_{(n-1,n)}]) + \dots \end{aligned}$$

Можем заметить, что:

$$\checkmark P[A_i] = P[A_j] = \frac{1}{n} \quad \forall i, j$$

$$\checkmark P[B_{(n-1,n)}] = \frac{1}{2}, \text{ потому что } n-1\text{-ый выбирает из двух оставшихся мест}$$

$$\checkmark P[B_{(n-2,n)}] + P[B_{(n-2,n-1)}]P[B_{(n-1,n)}] = \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{2}$$

✓

$$\begin{aligned} &P[B_{(n-3,n)}] + P[B_{(n-3,n-2)}](P[B_{(n-2,n)}] + P[B_{(n-2,n-1)}]P[B_{(n-1,n)}]) + \\ &+ P[B_{(n-3,n-1)}]P[B_{(n-1,n)}] = \frac{1}{4} + \frac{1}{4} \left(\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} \right) + \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{2} \quad (50) \end{aligned}$$

- ✓ И так далее до того момента, пока старушка не сядет на место первого человека, который заходит после нее, — всего $n - 2$ вариантов.

Таким образом мы получаем сумму:

$$P[\text{n-ый не сядет на свое место}] = \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{2} + \frac{1}{n} \cdot \frac{1}{2} + \dots = \frac{1}{n} + \frac{1}{2n}(n-2) = \frac{1}{2}$$

Значит вероятность $P[\text{n-ый сядет на свое место}] = 1 - \frac{1}{2} = \frac{1}{2}$

А вот ещё один вариант решения:

Метод математической индукции: допустим что это утверждение доказано для одного, двух и так далее до k человек. Рассмотрим $k + 1$ человека. Когда последний сядет на своё место? Если старушка сядет на своё место, а вероятность этого равна $\frac{1}{k+1}$ или, с вероятностью $\frac{1}{2}$ (по индукции), если старушка сядет на любое место кроме своего и последнего, то есть $\frac{1}{2} \cdot \frac{k-1}{k+1}$. В этом случае тот пассажир, чье место она заняла, становится старушкой, и мы получаем задачу при меньшем k . Складывая эти две дроби, получаем $\frac{1}{2}$.

Чтобы найти среднее число пассажиров, разобьем эту величину в сумму индикаторов: Y_1 — сел ли первый на место, \dots , Y_n — сел ли n -ый на место (индикатор равен единице, если сел).

Стало быть $E(Y) = E(Y_1) + E(Y_2) + \dots + E(Y_n)$. $E(Y_n) = \frac{1}{2}$.

Почти аналогично можем рассуждать для предпоследнего:

База индукции: если пассажиров трое ($n = 3$ включая старушку), то для предпоследнего вероятность сесть на своё место равна $\frac{2}{3}$.

Шаг индукции: допустим что для $3, 4, \dots, n$ пассажиров эта вероятность равна $\frac{2}{3}$. Рассмотрим случай $(n + 1)$ -го пассажира. Предпоследний сядет на своё место, если:

- старушка сядет на своё место или на место последнего $\frac{2}{n+1}$
- в $\frac{2}{3}$ тех случаев, когда старушка сядет на место $2, 3, \dots, (n-1)$, т.е. $\frac{2}{3} \cdot \frac{n-2}{n+1}$ складываем, получаем $\frac{2}{3}$. То есть по индукции вероятность того, что предпоследний сядет на своё место равна $\frac{2}{3}$

И по аналогии можно увидеть, что вероятность того, что k -ый с конца пассажир сядет на своё место равна $k/(k + 1)$

Если у нас n пассажиров включая СС, то среднее количество севших на свои места (раскладывая с конца) равно

$$\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \dots + \frac{n-1}{n} + \frac{1}{n}$$

Разбалловка

Часть 1

1. (a) 2 балла
(b) 4 балла
(c) 4 балла
2. (a) мат. ожидание — 3 балла, дисперсия — 3 балла
(b) распределение — 3 балла, знак неравенства — 1 балл
3. (a) 2 балла
(b) все по 1 баллу
(c) каждая вероятность — 1.5 балла
(d) 2 балла
4. (a) все по 1 баллу
(b) первая ковариация — 3 балла, вторая — 2 балла
(c) 2 балла
5. 10 или 0
6. (a) явный вид 2, приближение Пуассона 2
(b) наиболее вероятное число 2, математическое ожидание 1, дисперсия 1
(c) 2
7. (a) математическое ожидание 2.5, дисперсия 2.5
(b) для бесполезных и неинтересных 2.5, с хотя бы одним из свойств 2.5

8. (a) константы a и b 5
- (b) математическое ожидание 2
- (c) третий начальный момент, медиана и мода – по 1

Часть 2

1. (a) каждая вероятность — 1 балл
- (b) каждое мат. ожидание — 2 балла
- (c) 1 балл
2. (a) каждая вероятность — 1.5 балла
- (b) первое мат. ожидание — 1.5 балла, последние два мат. ожидания — по 2 балла
3. (a) 6 баллов
- (b) 4 балла
4. (a) $\mathbb{P}(X = 50) - 2$
- (b) $\mathbb{E}[X] - 4$
- (c) $\text{Var}(X) - 4$
5. (a) 2
- (b) 4
- (c) 4

12.5 Контрольная работа 2. Базовый поток. 15.12.2014

1. Ежемесячные расходы студенческой семьи Маши и Васи хорошо описываются случайным вектором (X, Y) , (X — расходы Маши, Y — расходы Васи), имеющим равномерное распределение в треугольнике, задаваемом ограничениями $\{0 \leq X, 0 \leq Y, X + Y \leq 1\}$.

Найдите:

- (a) Вероятность того, что совокупные расходы превысят половину бюджета, $\mathbb{P}(X + Y > 1/2)$
 - (b) Плотность распределения расходов Васи.
 - (c) Вероятность того, что Машины расходы составили менее трети бюджета, если известно, что Вася израсходовал более половины семейного бюджета.
 - (d) Условную плотность распределения и условное математическое ожидание расходов Маши, при условии, что Вася израсходовал половину бюджета.
 - (e) Математическое ожидание условного математического ожидания расходов Маши, $\mathbb{E}(\mathbb{E}(X|Y))$
 - (f) Коэффициент корреляции расходов Маши и Васи
2. Задана последовательность независимых случайных величин X_1, X_2, \dots

$$X_n \quad -\sqrt{n} \quad 0 \quad \sqrt{n}$$

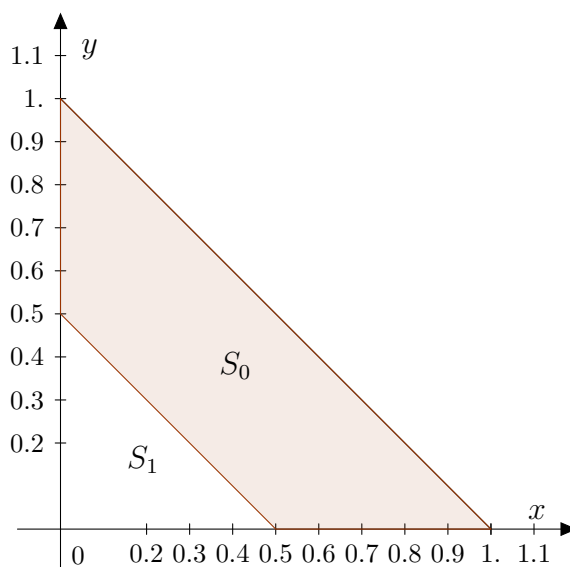
$$\mathbb{P}(X_n = \cdot) \quad 1/2n \quad 1 - 1/n \quad 1/2n$$

- (a) Сформулируйте закон больших чисел. Выполняется ли для данной последовательности закон больших чисел?
- (b) Запишите неравенство Чебышёва. Оцените вероятность того, что модуль среднего значения по n наблюдениям не превысит 1, $\mathbb{P}(|\bar{X}_n| \leq 1)$
- (c) Сколько членов последовательности необходимо взять, чтобы вероятность того, что модуль среднего значения не превысит 1, была не менее 0.9, $\mathbb{P}(|\bar{X}_n| \leq 1) \geq 0.9$
3. Размер выплат каждому клиенту банка — случайная величина с математическим ожиданием, равным 5000 ед. и среднеквадратическим отклонением, равным 2000 ед. Выплаты отдельным клиентам независимы. Сколько должно быть наличных денег в банке, чтобы с вероятностью 0.95 денег хватило на обслуживание 60 клиентов?
4. Рекламная компания хочет оценить вероятность p , с которой адресная реклама приводит к заявке. С этой целью она рассылает n рекламных проспектов. Обозначим за \hat{p} отношение числа поданных заявок к числу разосланных проспектов n . С помощью теоремы Муавра–Лапласа и неравенства Чебышёва определите:
- (a) Сколько нужно разослать рекламных проспектов, для того чтобы \hat{p} отличалось от истинной вероятности p не более, чем на 0.1 с вероятностью не меньшей 0.99
- (b) С какой точностью ε удастся оценить p с вероятностью 0.99, если разослана 1000 проспектов, т.е. $\mathbb{P}(|\hat{p} - p| \leq \varepsilon) \geq 0.99$?

12.6 Решение. Контрольная работа 2. Базовый поток. 15.12.2014

Вариант 1

Задача 1 1.1 Так как (X, Y) имеют совместное равномерное распределение, $\mathbb{P}\{X + Y > \frac{1}{2}\}$ можно рассчитать как отношение соответствующих площадей:



Соответственно:

$$\mathbb{P}\left\{X + Y > \frac{1}{2}\right\} = \frac{S_0}{S_0 + S_1} = \frac{0.5 - S_1}{0.5} = \frac{\frac{1}{2} - \frac{1}{8}}{\frac{1}{2}} = \frac{3}{4}$$

1.2

$$f_Y(y) = \int_0^{1-y} f_{XY}(x, y) dx$$

Поэтому, нам сначала нужно найти $f_{XY}(x, y)$, которая для равномерного распределения должна быть константой. Это можно сделать из условия:

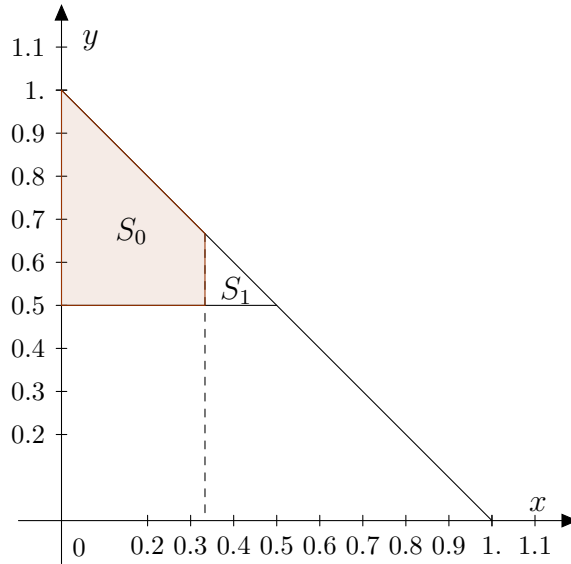
$$\int_0^1 \int_0^{1-x} f_{XY}(x, y) dx dy = \int_0^1 \int_0^{1-x} C dx dy = 1 \Rightarrow$$

$$\int_0^1 \int_0^{1-x} C dx dy = \int_0^1 C(1-x) dx = \left(Cx - C\frac{x^2}{2} \right) \Big|_0^1 = \frac{C}{2} = 1$$

Откуда имеем $f_{XY}(x, y) = C = 2$. Теперь можем найти плотность распределения расходов Васи:

$$f_Y(y) = \int_0^{1-y} 2 dx = 2(1-y)$$

1.3 В данном случае площади немного другие, но смысл тот же:



$$\mathbb{P}\left(X < \frac{1}{3} \mid Y > \frac{1}{2}\right) = \frac{S_0}{S_0 + S_1} = \frac{\frac{1}{8} - S_1}{\frac{1}{8}} = \frac{\frac{1}{8} - \frac{1}{72}}{\frac{1}{8}} = \frac{8}{9}$$

1.4 При $Y = \frac{1}{2}$, X распределен равномерно от 0 до $\frac{1}{2}$, поэтому его плотность равна

$$f_X(x) = \frac{1}{\frac{1}{2} - 0} = 2$$

Соответственно, условное математическое ожидание:

$$\mathbb{E}\left[X \mid Y = \frac{1}{2}\right] = \frac{1}{4}$$

1.5 $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$, а маргинальную функцию плотности для X мы можем найти так же, как искали для Y , и получим $f_X(x) = 2(1-x)$. Отсюда:

$$\mathbb{E}[X] = \int_0^1 2x(1-x) dx = \left(x^2 - \frac{2}{3}x^3 \right) \Big|_0^1 = \frac{1}{3}$$

1.6 Если вспомнить формулу для корреляции:

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]}{\sigma_X \sigma_Y}$$

то станет более менее очевидно, что надо найти $\mathbb{E}[XY]$ и дисперсии X и Y .

$$\begin{aligned}\mathbb{E}[XY] &= \int_0^1 \int_0^{1-x} 2xy dx dy = \int_0^1 2x dx \int_0^{1-x} y dy = \int_0^1 x(x^2 - 2x + 1) dx = \\ &= \left(\frac{x^4}{4} - \frac{2}{3}x^3 + \frac{x^2}{2} \right) \Big|_0^1 = \frac{3}{4} - \frac{2}{3} = \frac{1}{12}\end{aligned}$$

Соответственно:

$$\text{Cov}(X, Y) = \frac{1}{12} - \frac{1}{3} \cdot \frac{1}{3} = -\frac{1}{36}$$

Найдем теперь дисперсии X и Y (они будут одинаковыми, как и математические ожидания, в силу симметрии):

$$\mathbb{E}[X^2] = \int_0^1 2x^2(1-x) dx = \left(\frac{2}{3}x^2 - \frac{x^4}{4} \right) \Big|_0^1 = \frac{1}{6}$$

Поэтому:

$$\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$$

Теперь наконец-то можем найти корреляцию:

$$\rho_{XY} = -\frac{\frac{1}{36}}{\sqrt{\frac{1}{18}}\sqrt{\frac{1}{18}}} = -\frac{1}{2}$$

Задача 2 2.1 Закон больших чисел гласит что $\bar{X} \rightarrow \mathbb{E}[X]$ при $n \rightarrow \infty$. Проверим его выполнение в данном случае:

$$\mathbb{E}[X_n] = \frac{1}{2n}(-\sqrt{n}) + \left(1 - \frac{1}{n}\right) \cdot 0 + \frac{1}{2n}\sqrt{n} = 0$$

$$\lim_{n \rightarrow \infty} \bar{X} = \lim_{n \rightarrow \infty} \frac{X_1 + \dots + X_n}{n} = 0$$

так как числитель ограничен, а знаменатель бесконечно возрастает. Видим, что ЗБЧ в данном случае, конечно, выполняется.

Как вариант, можно было сказать, что дисперсия ограничена, и из этого также следует выполнение ЗБЧ.

2.2 Неравенство Чебышева:

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\text{Var}(X)}{\varepsilon^2}$$

Соответственно, искомую вероятность можем оценить следующим образом:

$$\mathbb{P}(|\bar{X}| \leq 1) = 1 - \mathbb{P}(|\bar{X}| \geq 1) \Rightarrow \mathbb{P}(|\bar{X}| \leq 1) \geq 1 - \frac{\text{Var}[\bar{X}]}{1}$$

$$\text{Var}[\bar{X}] = \text{Var}\left[\frac{\sum_{i=1}^n X_i}{n}\right] = \frac{1}{n^2} \sum_{i=1}^n \text{Var} X_i$$

В свою очередь:

$$\mathbb{E}[X_i^2] = 2 \cdot \frac{1}{2n} \cdot n + \left(1 - \frac{1}{n}\right) \cdot 0 = 1 \Rightarrow \text{Var}[X_i] = 1 \Rightarrow \text{Var}[\bar{X}] = \frac{1}{n}$$

Поэтому:

$$\mathbb{P}(|\bar{X}| \leq 21 \geq 1 - \frac{1}{n}$$

2.3

$$1 - \frac{1}{n} = 0.9 \Rightarrow n = 10$$

Задача 3 Обозначим за R — необходимое количество наличных денег в банке. Пусть X — случайная величина, показывающее размер суммарной выплаты 60 (n — достаточное большое для применения ЦПТ) клиентам. Ясно, что т.к. выплаты отдельным клиентам независимы: $\mathbb{E}X = 60 \cdot 5000 = 3 \cdot 10^5$; $\text{Var} X = 60 \cdot 2000^2 = 2.4 \cdot 10^8$; $\sigma_X = \sqrt{2.4 \cdot 10^4} \approx 1.55 \cdot 10^4$

Теперь по ЦПТ:

$$\begin{aligned} \mathbb{P}(R \geq X) &= 0.95 \\ \mathbb{P}\left(\frac{X - \mathbb{E}X}{\sigma_X} \leq \frac{R - \mathbb{E}X}{\sigma_X}\right) &= 0.95 \\ \mathbb{P}\left(Z \leq \frac{R - 3 \cdot 10^5}{1.55 \cdot 10^4}\right) &= 0.95 \end{aligned}$$

Слева функция распределения; подставляя 95-го квантиль стандартного нормального распределения, получаем:

$$\begin{aligned} \frac{R - 3 \cdot 10^5}{1.55 \cdot 10^4} &= 1.64 \\ R &= 325420 \end{aligned}$$

Задача 4 4.1 По предельной теореме Муавра-Лапласа:

$$\begin{aligned} \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} &\sim N(0, 1) \\ \mathbb{P}\left(\frac{|\hat{p} - p|}{\sqrt{p(1-p)/n}} \leq \frac{0.1}{\sqrt{p(1-p)/n}}\right) &\geq 0.99 \\ \mathbb{P}\left(|Z| \leq \frac{0.1}{\sqrt{p(1-p)/n}}\right) &\geq 0.99 \end{aligned}$$

Из симметричности стандартного нормального распределения и зная его 99.5-го квантиль, равный приблизительно 2.58, получаем:

$$\begin{aligned} \frac{0.1}{\sqrt{p(1-p)/n}} &\geq 2.58 \\ \frac{\sqrt{n}}{\sqrt{p(1-p)}} &\geq \frac{2.58}{0.1} \end{aligned}$$

$$\sqrt{n} \geq \frac{2.58}{0.1} \sqrt{p(1-p)}$$

$$n \geq 665.64 \cdot p(1-p)$$

С помощью неравенства Чебышева:

$$\mathbb{P}(|\hat{p} - p| \leq 0.1) \geq 0.99$$

$$\mathbb{P}(|\hat{p} - p| \geq 0.1) \leq 0.01$$

Теперь просто смотрим на неравенство Чебышева и на строчку выше, на неравенство Чебышева и на строчку выше...

$$\frac{p(1-p)/n}{0.1^2} = 0.01$$

$$n = 10^4 p(1-p)$$

Принимаются оба ответа!

4.2 По предельной теореме Муавра-Лапласа:

$$\mathbb{P}\left(\frac{|\hat{p} - p|}{\sqrt{p(1-p)/n}} \leq \frac{\varepsilon}{\sqrt{p(1-p)/1000}}\right) \geq 0.99$$

$$\mathbb{P}\left(|Z| \leq \frac{\varepsilon}{\sqrt{p(1-p)/1000}}\right) \geq 0.99$$

Аналогично пункту 1:

$$\frac{\varepsilon}{\sqrt{p(1-p)/1000}} \geq 2.58$$

$$\varepsilon \geq 0.082 \sqrt{p(1-p)}$$

С помощью неравенства Чебышева:

$$\mathbb{P}(|\hat{p} - p| \leq \varepsilon) \geq 0.99$$

$$\mathbb{P}(|\hat{p} - p| \geq \varepsilon) \leq 0.01$$

Аналогично пункту 1:

$$\frac{p(1-p)/1000}{\varepsilon^2} = 0.01$$

$$\varepsilon^2 = \frac{p(1-p)}{10}$$

$$\varepsilon = \sqrt{\frac{p(1-p)}{10}} \approx 0.316 \sqrt{p(1-p)}$$

Нужно было показать, как мастерство владения теоремой Муавра-Лапласа, так и неравенством Чебышева.

12.7 Праздник номер 2, i-поток, 15.12.2014

Time: 180 min

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

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

2. 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. В каких пределах лежит вероятность того, что возраст случайно выбираемого женатого мужчины отклоняется от своего математического ожидания больше чем на 10 лет?

3. На окружности единичной длины случайным образом равномерно и независимо друг от друга выбирают две дуги: длины 0.3 и длины 0.4.

(a) Найдите функцию распределения длины пересечения этих отрезков

(b) Найдите среднюю длину пересечения

4. Совместная функция плотности величин X и Y имеет вид

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

(a) [1] Найдите $\mathbb{P}(X + Y > 1)$

(b) [6] Найдите $\text{Cov}(X, Y)$

(c) [1] Являются ли величины X и Y независимыми?

(d) [1] Являются ли величины X и Y одинаково распределенными?

5. Изначально цена акций компании «Пумперникель» равна $X_0 = 1000$ рублей. Каждый последующий день в течение 100 дней цена равновероятно может вырасти на 2 рубля или упасть на 1 рубль. Обозначим цену акции через n дней как X_n .

(a) Чему равны математическое ожидание и дисперсия изменения цены за отдельный день?

(b) Найдите $\mathbb{E}(X_n)$, $\text{Var}(X_n)$, $\text{Cov}(X_n, X_k)$

(c) Сформулируйте центральную предельную теорему

(d) Примерно найдите вероятность $\mathbb{P}(X_{100} > 1060)$

(e) Биржевой игрок Вениамин утверждает, что через 100 дней с вероятностью 95% цена акций «Пумперникель» не опустится ниже a . Чему равно a ?

6. Сэр Фрэнсис Гальтон — учёный XIX-XX веков, один из основоположников как генетики, так и статистики — изучал, среди всего прочего, связь между ростом детей и родителей. Он исследовал данные о росте 928 индивидов. Обозначим X_1 — рост случайного человека, а X_2 — среднее арифметическое роста его отца и матери. По результатам исследования Гальтона:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N \left[\begin{pmatrix} 68.1 \\ 68.3 \end{pmatrix} ; \begin{pmatrix} 6.3 & 2.1 \\ 2.1 & 3.2 \end{pmatrix} \right]$$

- (a) Обратите внимание на то, что дисперсия роста детей выше дисперсии среднего роста родителей. С чем это может быть связано? Учтите, что рост детей измерялся уже по достижении зрелости, так что разброс не должен быть связан с возрастными различиями.
- (b) Рассчитайте корреляцию между X_1 и X_2
- (c) Один дюйм примерно равен 2.54 сантиметра. Пусть X'_1 и X'_2 — это те же X_1 и X_2 , только измеренные в сантиметрах. Найдите вектор математических ожиданий и ковариационную матрицу вектора $X' = (X'_1, X'_2)$.
- (d) Определите, каков ожидаемый рост и дисперсия роста человека, средний рост родителей которого составляет 72 дюйма?
- (e) Найдите вероятность того, что рост человека превысит 68 дюймов, если средний рост его родителей равен 72 дюймам. Подсказка: используйте предыдущий пункт и нормальность распределения!

7. Звонки поступают в пожарную часть согласно пуассоновскому потоку в среднем 2 раза в час. Предположим, что после получения звонка пожарная часть занята тушением пожара случайное время равномерно распределённое от получаса до часа. В это время звонки перенаправляются в соседнюю пожарную часть.

Пожарная часть только-только начала работать и готова принимать звонки.

- (a) Какова вероятность того, что за ближайший час не поступит звонков?
- (b) Какова вероятность того, что за ближайший час не будет перенаправленных звонков?
- (c) Найдите закон распределения количества звонков до первого перенаправленного звонка.

8. Судьба Дон-Жуана

У Дон-Жуана n знакомых девушек, и их всех зовут по-разному. Он пишет им n писем, но по рассеянности раскладывает их в конверты наугад. Случайная величина X обозначает количество девушек, получивших письма, адресованные лично им.

- (a) Найдите $\mathbb{E}(X)$, $\text{Var}(X)$
- (b) Какова при большом n вероятность того, что хотя бы одна девушка получит письмо, адресованное ей?

9. Игла Бюффона

Плоскость разлинована параллельными линиями через каждый сантиметр. Случайным образом на эту плоскость бросается иголка длины $a < 1$.

- (a) Какова вероятность того, что иголка пересечёт какую-нибудь линию?
- (b) Предложите вероятностный способ оценки числа π

12.8 Решение. Праздник номер 2, i-поток

Задача 1 Пусть X — оценка Пети, Y — оценка Васи. Тогда исходя из условия:

X	7	8
$\mathbb{P}(\cdot)$	$\frac{1}{2}$	$\frac{1}{2}$

Y	7	8
$\mathbb{P}(\cdot)$	$\frac{2}{3}$	$\frac{1}{3}$

Теперь можем найти математические ожидания и дисперсии:

$$\begin{aligned}\mathbb{E}[X] &= \frac{15}{2}, \quad \mathbb{E}[Y] = \frac{22}{3} \Rightarrow \\ \text{Var}[X] &= \frac{1}{2} \cdot 64 + \frac{1}{2} \cdot 49 - \left(\frac{15}{2}\right)^2 = \frac{1}{4} \\ \text{Var}[Y] &= \frac{1}{3} \cdot 64 + \frac{2}{3} \cdot 49 - \left(\frac{22}{3}\right)^2 = \frac{2}{9}\end{aligned}$$

Теперь из формулы для корреляции мы можем найти $\mathbb{E}[XY]$:

$$0.7 = \frac{\mathbb{E}[XY] - \frac{15}{2} \cdot \frac{22}{3}}{\sqrt{\frac{1}{4}} \sqrt{\frac{2}{9}}} \Rightarrow \mathbb{E}[XY] \approx 55.16499$$

Можем составить табличку для совместного распределения, немножко подумав:

		Y		
		7	8	
X	7	p	$\frac{1}{2} - p$	$\frac{1}{2}$
	8	$\frac{2}{3} - p$	$p - \frac{1}{6}$	$\frac{1}{2}$
		$\frac{2}{3}$	$\frac{1}{3}$	

$$\mathbb{E}[XY] = 49p + 56 \left(\frac{1}{2} - p + \frac{2}{3} - p \right) + 64 \left(p - \frac{1}{6} \right) = 55.16499$$

$$\Rightarrow p \approx 0.4983233 \Rightarrow \mathbb{P}(X = Y) = p + p - \frac{1}{6} = 0.8299799$$

Задача 2 Пусть M — возраст мужа, F — возраст жены. Тогда из условия:

$$\text{Var}(M - F) = \text{Var}(M) + \text{Var}(F) - 2 \text{Cov}(M, F) = 4 \Rightarrow \text{Cov}(M, F) = \frac{37}{2}$$

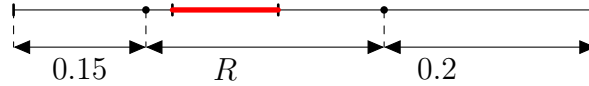
$$\rho_{MF} = \frac{\frac{37}{2}}{5 \cdot 4} = \frac{37}{40}$$

Согласно неравенству Чебышева:

$$\mathbb{P}(|M - \mathbb{E}[M]| \geq 10) \leq \frac{\text{Var}(M)}{10^2} = 0.25$$

Поэтому данная вероятность лежит в пределах $(0, 25]$.

Задача 3 (а) Пусть X — длина пересечения. Требуется найти $\mathbb{P}(X \leq x)$. Рассмотрим расположение центров отрезков относительно друг друга:



Заметим, что $X = 0.15 - (R - 0.2) = 0.35 - R$. Однако, $R \sim U[0, \frac{1}{2}]$. Можно представить, что эти точки бросают на окружность по очереди, тогда расстояние между ними не может превышать $\frac{1}{2}$ и распределено равномерно. Имеем:

$$\mathbb{P}(X \leq x) = \mathbb{P}(0.35 - R \leq x) = \mathbb{P}(R \geq 0.35 - x) = \frac{0.5 - 0.35 + x}{0.5} = 0.3 + 2x$$

Строго говоря, длина пересечения не может быть больше 0.3 по понятным причинам, поэтому функция распределения выглядит так:

$$F(x) = \begin{cases} 0.3 + 2x, & x < 0.3 \\ 1, & x \geq 0.3 \end{cases}$$

(b) Зная функцию распределения, можем найти плотность:

$$f_X(x) = \begin{cases} 2, & x < 0.3 \\ 0, & \text{else} \end{cases}$$

Соответственно:

$$\mathbb{E}[X] = \int_0^{0.3} 2x dx = x^2|_0^{0.3} = 0.09$$

Задача 4 (а)

$$\mathbb{P}(X + Y > 1) = \iint_{\mathcal{G}} 2(x^3 + y^3) dx dy = \int_0^1 dx \int_{1-x}^1 2(x^3 + y^3) dy = \dots = 0.8$$

(b) Для того, чтобы найти ковариацию, придется найти :

$$\mathbb{E}[X] = \int_0^1 dx \int_0^1 2x(x^3 + y^3) dy = \dots = 0.65$$

$$\mathbb{E}[Y] = \int_0^1 dy \int_0^1 2y(x^3 + y^3) dx = \dots = 0.65$$

$$\mathbb{E}[XY] = \int_0^1 dx \int_0^1 2xy(x^3 + y^3) dy = \dots = 0.4$$

Ну а потом все просто: $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = -0.0225$

(с) Проверить свойство функций плотности $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$, а если $\text{Cov}(X, Y) \neq 0$, то сразу зависимы. Здесь — зависимы.

(d) Да, являются.

Задача 5 Для этой задачи аккуратно нужно писать $(\Delta X)_j$, но так как приращения цен независимы (см. биномиальная модель рынка в интернете) будем писать ΔX .

(a)

$$\begin{aligned}\mathbb{E}(\Delta X) &= 2 \cdot \frac{1}{2} - 1 \cdot \frac{1}{2} = \frac{1}{2} \\ \mathbb{E}(\Delta X)^2 &= 4 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{5}{2} \\ \text{Var}(\Delta X) &= \frac{5}{2} - \left(\frac{1}{2}\right)^2 = \frac{10-1}{4} = \frac{9}{4}\end{aligned}$$

(b)

$$\begin{aligned}\mathbb{E}(X_n) &= \mathbb{E}(1000 + n\Delta X) = 1000 + n\mathbb{E}\Delta X = 1000 + 0.5n \\ \text{Var}(X_n) &= n \text{Var}(\Delta X) = n \frac{9}{4} \\ \text{Cov}(X_n, X_k) &= \text{Cov}(X_k, X_k + (n-k)\Delta X)\end{aligned}$$

Так как цена в момент k никак не связана с последующими случайными блужданиями, то:

$$\text{Cov}(X_n, X_k) = \text{Cov}(X_k, X_k + (n-k)\Delta X) = \text{Cov}(X_k, X_k) + 0 = \text{Var } X_k = k \frac{9}{4}$$

(c) Заметим, что для самих цен акций ЦПТ мы формулировать не можем ввиду того, что случайные величины коррелированы, не i.i.d. Зато мы можем сформулировать ее для независимых приращений!

$(\Delta X)_1, \dots, (\Delta X)_k$ — последовательность i.i.d. случайных величин (при большом k) с конечными математическим ожиданием $1/2$ и стандартным отклонением $3/2$. Пусть также $S_k = \sum_{i=1}^k (\Delta X)_j$. Тогда формулировка ЦПТ:

$$\frac{S_k - k/2}{3/2\sqrt{k}} \sim N(0, 1)$$

(d) Воспользуемся результатом ЦПТ!

$$\begin{aligned}\mathbb{P}(X_{100} > 1060) &= \mathbb{P}(1000 + S_{100} > 1060) = \mathbb{P}(S_{100} > 60) = \\ &= \mathbb{P}\left(\frac{S_{100} - 50}{3/2 \cdot 10} > \frac{60 - 50}{3/2 \cdot 10}\right) = \mathbb{P}\left(Z > \frac{2}{3}\right) = 1 - \text{pnorm}(2/3) \approx 0.25\end{aligned}$$

(e)

$$\begin{aligned}\mathbb{P}(1000 + S_{100} > a) &= 0.95 \\ \mathbb{P}(S_{100} < a - 1000) &= 0.05\end{aligned}$$

Используя ЦПТ:

$$\begin{aligned}a - 1000 &= \text{qnorm}(0.05, \text{mean} = 50, \text{sd} = 15) \\ a &\approx 1025.33\end{aligned}$$

Задача 6

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N \left(\begin{pmatrix} 68.1 \\ 68.3 \end{pmatrix}; \begin{pmatrix} 6.3 & 2.1 \\ 2.1 & 3.2 \end{pmatrix} \right)$$

(a) Смотреть здесь

(b)

$$\text{corr}(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\sigma_{X_1} \sigma_{X_2}} = \frac{2.1}{\sqrt{6.3 \cdot 3.2}} \approx 0.47$$

(c)

$$\begin{pmatrix} X'_1 \\ X'_2 \end{pmatrix} \sim N \left(\begin{pmatrix} 68.1 \cdot 2.54 \\ 68.3 \cdot 2.54 \end{pmatrix}; \begin{pmatrix} 6.3 \cdot 2.54^2 & 2.1 \cdot 2.54^2 \\ 2.1 \cdot 2.54^2 & 3.2 \cdot 2.54^2 \end{pmatrix} \right)$$
$$\begin{pmatrix} X'_1 \\ X'_2 \end{pmatrix} \sim N \left(\begin{pmatrix} 172.97 \\ 173.48 \end{pmatrix}; \begin{pmatrix} 40.65 & 13.55 \\ 13.55 & 20.61 \end{pmatrix} \right)$$

(d) Универсальный способ:

Всегда можно представить X_1 и X_2 в следующем виде (из соображений ЛНЗ для X_1 две стандартные нормальные независимые Z_1 и Z_2):

$$X_1 = \mathbb{E}X_1 + aZ_1 + bZ_2$$

$$X_2 = \mathbb{E}X_2 + cZ_1$$

$$X_1 = 68.1 + aZ_1 + bZ_2$$

$$X_2 = 68.3 + cZ_1$$

Зная ковариационную матрицу, легко найти коэффициенты a, b, c :

$$\text{Var } X_1 = a^2 + b^2 = 6.3$$

$$\text{Var } X_2 = c^2 = 3.2$$

$$\text{Cov}(X_1, X_2) = ac \text{Var } Z_1 = ac = 2.1$$

Получаем $c = 1.79$, $a = 1.17$, $b = 2.22$. Итак:

$$X_1 = 68.1 + 1.17Z_1 + 2.22Z_2$$

$$X_2 = 68.3 + 1.79Z_1$$

Теперь легко получаем:

$$\begin{aligned} \mathbb{E}(X_1|X_2 = 72) &= \mathbb{E}(68.1 + 1.17Z_1 + 2.22Z_2|68.3 + 1.79Z_1 = 72) = \\ &= \mathbb{E}(68.1 + 1.17Z_1 + 2.22Z_2|Z_1 = 2.07) = \\ &= \mathbb{E}(68.1 + 1.17 \cdot 2.07 + 2.22Z_2) = 70.52 \end{aligned}$$

Это кстати отражает довольно известный факт, что дети высоких родителей будут тоже высокими, но менее высокими: некий mean reversion (regression to the mean — узнаете в эконометрике).

$$\text{Var}(X_1|X_2 = 72) = \text{Var}(68.1 + 1.17 \cdot 2.07 + 2.22Z_2) = 2.22^2 = 4.93$$

- (е) Используем стандартизацию и факт о том, что условное распределение нормальных — нормальное:

$$\begin{aligned}\mathbb{P}(X_1 > 68 | X_2 = 72) &= \mathbb{P}\left(\frac{(X_1 | X_2 = 72) - \mathbb{E}(X_1 | X_2 = 72)}{\sqrt{\text{Var}(X_1 | X_2 = 72)}} > \frac{68 - 70.52}{\sqrt{4.93}}\right) = \\ &= \mathbb{P}(Z > -1.135) = 1 - \text{pnorm}(-1.135) = 0.87\end{aligned}$$

Задача 7

- (а)

$$X \sim \text{Pois}(\lambda = 2)$$

Тогда:

$$\mathbb{P}(X = 0) = e^{-\lambda} = e^{-2} \approx 0.135$$

- (b) Учитывая закон распределения времени на тушение пожара $Z \sim U[1/2, 1]$ можно однозначно сказать, что за ближайший час может не быть перенаправленных звонков только при $X \leq 2$. Пусть Y_1 — время между 1-ым и 2-ым звонком. Оно имеет экспоненциальное распределение $Y \sim \exp(1/2)$. Итак, искомая вероятность:

$$\mathbb{P}(\cdot) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)\mathbb{P}(Y_1 > Z_1)$$

Сложность в нахождении вероятности $\mathbb{P}(Y_1 > Z_1)$. Предоставим путь в лоб:

Совместная функция плотности случайных величин Y_1 и Z_1 , очевидно:

$$p(y, z) = p(y)p(z) = 2e^{-2y} \cdot 2 = 4e^{-2y}$$

Тогда:

$$\begin{aligned}\mathbb{P}(Y_1 > Z_1) &= \int_{1/2}^1 \int_z^\infty 4e^{-2y} dy dz = - \int_{1/2}^1 2e^{-2y} \Big|_z^\infty dz = \int_{1/2}^1 2e^{-2z} dz = \\ &= -e^{-2z} \Big|_{1/2}^1 = e^{-1} - e^{-2}\end{aligned}$$

Можно и без двойных интегралов, но более мозгоемко. Благодаря memoryless property пуассоновского потока:

$$\mathbb{P}(Y_1 > Z_1) = F_{Y_1}(1) - F_{Y_1}(1/2) = 1 - e^{-2} - (1 - e^{-1}) = e^{-1} - e^{-2}$$

Итак:

$$\mathbb{P}(\cdot) = e^{-2} + 2e^{-2} + 2e^{-2}(e^{-1} - e^{-2}) \approx 0.47$$

- (с) Благодаря memoryless property, вероятность перенаправления для каждого звонка, начиная со 2, равна $e^{-1} - e^{-2}$. Обозначим за Q — количество звонков до первого перенаправленного звонка.

$$Q - 1 \sim \text{Geom}(e^{-1} - e^{-2})$$

Задача 8 (а) Обозначим: $X_i = \begin{cases} 1, & \text{если } i\text{-ая девушка получила свое письмо} \\ 0, & \text{иначе} \end{cases}$

Так как письма раскладывались рандомно, то: $\mathbb{P}(X_i = 1) = 1/n$. Тогда:

$$\mathbb{E}X_i = \frac{1}{n}$$

$$\mathbb{E}X_i^2 = \frac{1}{n}$$

$$\text{Var } X_i = \frac{1}{n} - \frac{1}{n^2}$$

Легко найти:

$$\mathbb{E}X = \mathbb{E}(X_1 + \dots + X_n) = n\mathbb{E}X_i = 1$$

Для поиска дисперсии понадобятся ковариации, так как ясно, что если одна из девушек получила свое письмо, значит, она не отняла чье-то письмо, а следовательно, повышает вероятность получения своего письма для других девушек.

$$\text{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - (\mathbb{E}X_i)(\mathbb{E}X_j) = \mathbb{P}(X_i = 1, X_j = 1) - \frac{1}{n^2}$$

Чтобы убедиться в $\mathbb{E}(X_i X_j) = \mathbb{P}(X_i = 1, X_j = 1)$ можно нарисовать табличку совместного распределения этих случайных величин. Теперь если i -ая девушка получила свое письмо, то для второй девушки остается $n - 1$ писем и одно благоприятное письмо, следовательно:

$$\mathbb{P}(X_i = 1, X_j = 1) = \frac{1}{n(n-1)}$$

Итак:

$$\text{Cov}(X_i, X_j) = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{n - n + 1}{n^2(n-1)} = \frac{1}{n^2(n-1)}$$

Остался последний шаг:

$$\begin{aligned}\text{Var } X &= \text{Var}(X_1 + \dots + X_n) = n \text{Var } X_i + 2C_n^2 \text{Cov}(X_i, X_j) = \\ &= 1 - \frac{1}{n} + 2 \frac{n(n-1)}{2} \cdot \frac{1}{n^2(n-1)} = 1\end{aligned}$$

(b) Т.к. $\mathbb{P}(X_i = 1) = 1/n$, то:

$$\mathbb{P}(\cdot) = \left(1 - \frac{1}{n}\right)^n$$

Первый замечательный предел!

$$\mathbb{P}(\cdot) = \frac{1}{e}$$

Задача 9 [Решение внизу страницы](#)

12.9 Пересдача за 1-ый семестр??

Вопрос 1 ♣ Случайным образом выбирается семья с двумя детьми. Событие A — в семье старший ребенок — мальчик, событие B — в семье только один из детей — мальчик, событие C — в семье хотя бы один из детей — мальчик. Вероятность $\mathbb{P}(C)$ равна

☐ A 1/4

☐ C 1

☐ E 2/3

☐ B 1/2

☒ 3/4

☐ F Нет верного ответа.

Вопрос 2 ♣ Случайным образом выбирается семья с двумя детьми. Событие A — в семье старший ребенок — мальчик, событие B — в семье только один из детей — мальчик, событие C — в семье хотя бы один из детей — мальчик. Вероятность $\mathbb{P}(A|C)$ равна

☐ A 3/4

☐ C 1/4

☒ 2/3

☐ B 1/2

☐ D 1

☐ F Нет верного ответа.

Вопрос 3 ♣ Имеется три монетки. Две «правильных» и одна — с «орлами» по обеим сторонам. Вася выбирает одну монетку наугад и подкидывает ее один раз. Вероятность того, что выпадет орел равна

- ☐ A $1/3$
☐ C $1/2$
☐ E $2/5$
☐ B $3/5$
☒ 2/3
☐ F Нет верного ответа.

Вопрос 4 ♣ Случайным образом выбирается семья с двумя детьми. Событие A — в семье старший ребенок — мальчик, событие B — в семье только один из детей — мальчик, событие C — в семье хотя бы один из детей — мальчик.

- ☐ A Любые два события из A, B, C — зависимы
☐ B События A, B, C — независимы попарно, но зависимы в совокупности
☒ A и B — независимы, A и C — зависимы, B и C — зависимы
☐ D События A, B, C — независимы в совокупности
☐ E $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$
☐ F Нет верного ответа.

Вопрос 5 ♣ Случайным образом выбирается семья с двумя детьми. Событие A — в семье старший ребенок — мальчик, событие B — в семье только один из детей — мальчик, событие C — в семье хотя бы один из детей — мальчик. Вероятность $\mathbb{P}(A \cup C)$ равна

- ☐ A 1
☐ C $1/2$
☐ E $2/3$
☐ B $3/8$
☒ 3/4
☐ F Нет верного ответа.

Вопрос 6 ♣ Имеется три монетки. Две «правильных» и одна — с «орлами» по обеим сторонам. Вася выбирает одну монетку наугад и подкидывает ее один раз. Вероятность того, что была выбрана неправильная монетка, если выпал орел, равна

- ☐ A $2/3$
☐ C $3/5$
☐ E $1/3$
☒ 1/2
☐ D $3/2$
☐ F Нет верного ответа.

Вопрос 7 ♣ Вася бросает 7 правильных игральных кубиков. Наиболее вероятное количество выпавших шестёрок равно

- ☒ 1
☐ C 2
☐ E 0
☐ B $6/7$
☐ D $7/6$
☐ F Нет верного ответа.

Вопрос 8 ♣ Вася бросает 7 правильных игральных кубиков. Пусть величина X — сумма очков, выпавших на первых двух кубиках, а величина Y — сумма очков, выпавших на следующих пяти кубиках. Ковариация $\text{Cov}(X, Y)$ равна

- ☐ A $2/5$
☐ C $-2/5$
☐ E 0.5
☒ 0
☐ D 1
☐ F Нет верного ответа.

Вопрос 9 ♣ Вася бросает 7 правильных игральных кубиков. Дисперсия суммы выпавших очков равна

- ☒ $7 \cdot \frac{35}{12}$
☐ C $7 \cdot \frac{35}{36}$
☐ E 7
☐ B $7/6$
☐ D $35/36$
☐ F Нет верного ответа.

Вопрос 10 ♣ Вася бросает 7 правильных игральных кубиков. Математическое ожидание суммы выпавших очков равно

- ☐ A 30 ☐ C 42 ☒ 24.5
☐ B 21 ☐ D $7/6$ ☐ F Нет верного ответа.

Вопрос 11 ♣ Вася бросает 7 правильных игральных кубиков. Вероятность того, что ровно на пяти из кубиков выпадет шестёрка равна

- ☐ A $\frac{525}{12} \left(\frac{1}{6}\right)^7$ ☐ C $\left(\frac{1}{6}\right)^7$ ☐ E $\left(\frac{1}{6}\right)^5$
☐ B $\frac{7}{12} \left(\frac{1}{6}\right)^5$ ☒ 525 $\left(\frac{1}{6}\right)^7$ ☐ F Нет верного ответа.

Вопрос 12 ♣ Число изюминок в булочке — случайная величина, имеющая распределение Пуассона. Известно, что в среднем каждая булочка содержит 13 изюминок. Вероятность того, что в случайно выбранной булочке окажется только одна изюминка равна:

- ☐ A $e^{13}/13!$ ☐ C e^{-13} ☐ E $1/13$
☐ B $e^{-13}/13$ ☒ $13e^{-13}$ ☐ F Нет верного ответа.

В вопросах 13-16 совместное распределение пары величин X и Y задано таблицей:

	$Y = -1$	$Y = 0$	$Y = 1$
$X = -1$	$1/4$	0	$1/4$
$X = 1$	$1/6$	$1/6$	$1/6$

Вопрос 13 ♣ Математическое ожидание случайной величины X при условии, что $Y = -1$ равно

- ☐ A $-1/12$ ☐ C $1/10$ ☒ $-1/5$
☐ B $-1/3$ ☐ D 0 ☐ F Нет верного ответа.

Вопрос 14 ♣ Дисперсия случайной величины Y равна

- ☐ A $5/12$ ☐ C $1/2$ ☐ E $12/5$
☐ B $1/3$ ☒ $5/6$ ☐ F Нет верного ответа.

Вопрос 15 ♣ Ковариация, $\text{Cov}(X, Y)$, равна

- ☒ 0 ☐ C 0.5 ☐ E -1
☐ B 1 ☐ D -0.5 ☐ F Нет верного ответа.

Вопрос 16 ♣ Вероятность того, что $X = 1$ при условии, что $Y < 0$ равна

- ☐ A $1/12$ ☐ C $1/6$ ☐ E $1/3$
☒ $2/5$ ☐ D $5/12$ ☐ F Нет верного ответа.

В вопросах 17-19 функция распределения случайной величины X имеет вид

$$F(x) = \begin{cases} 0, & \text{если } x < 0 \\ cx^2, & \text{если } x \in [0; 1] \\ 1, & \text{если } x > 1 \end{cases}$$

Вопрос 17 ♣ Вероятность того, что величина X примет значение из интервала $[0.5, 1.5]$ равна

- | | | |
|--------------------------------|--------------------------------|--|
| <input type="checkbox"/> $3/4$ | <input type="checkbox"/> $3/2$ | <input type="checkbox"/> $2/3$ |
| <input type="checkbox"/> $1/2$ | <input type="checkbox"/> 1 | <input type="checkbox"/> Нет верного ответа. |

Вопрос 18 ♣ Математическое ожидание $\mathbb{E}(X)$ равно

- | | | |
|--------------------------------|--------------------------------|--|
| <input type="checkbox"/> $2/3$ | <input type="checkbox"/> $1/2$ | <input type="checkbox"/> $1/4$ |
| <input type="checkbox"/> $3/4$ | <input type="checkbox"/> 2 | <input type="checkbox"/> Нет верного ответа. |

Вопрос 19 ♣ Константа c равна

- | | | |
|--------------------------------|--------------------------------|--|
| <input type="checkbox"/> $2/3$ | <input type="checkbox"/> 1.5 | <input checked="" type="checkbox"/> 1 |
| <input type="checkbox"/> 0.5 | <input type="checkbox"/> 2 | <input type="checkbox"/> Нет верного ответа. |

В вопросах 20-23 совместная функция плотности пары X и Y имеет вид

$$f(x, y) = \begin{cases} cx^2y^2, & \text{если } x \in [0; 1], y \in [0; 1] \\ 0, & \text{иначе} \end{cases}$$

Вопрос 20 ♣ Условная функция плотности $f_{X|Y=2}(x)$ равна

- | | |
|--|---|
| <input checked="" type="checkbox"/> не определена | <input type="checkbox"/> $f_{X Y=2}(x) = \begin{cases} 9x^2 & \text{если } x \in [0; 1] \\ 0, & \text{иначе} \end{cases}$ |
| <input type="checkbox"/> $f_{X Y=2}(x) = \begin{cases} 36x^2 & \text{если } x \in [0; 1] \\ 0, & \text{иначе} \end{cases}$ | <input type="checkbox"/> $f_{X Y=2}(x) = \begin{cases} x^2 & \text{если } x \in [0; 1] \\ 0, & \text{иначе} \end{cases}$ |
| <input type="checkbox"/> $f_{X Y=2}(x) = \begin{cases} 3x^2 & \text{если } x \in [0; 1] \\ 0, & \text{иначе} \end{cases}$ | <input type="checkbox"/> Нет верного ответа. |

Вопрос 21 ♣ Константа c равна

- | | | |
|--------------------------------|---|--|
| <input type="checkbox"/> $1/4$ | <input type="checkbox"/> 1 | <input type="checkbox"/> 2 |
| <input type="checkbox"/> $1/2$ | <input checked="" type="checkbox"/> 9 | <input type="checkbox"/> Нет верного ответа. |

Вопрос 22 ♣ Математическое ожидание $\mathbb{E}(X/Y)$ равно

- | | | |
|---|------------------------------|--|
| <input checked="" type="checkbox"/> $9/8$ | <input type="checkbox"/> 2 | <input type="checkbox"/> $1/2$ |
| <input type="checkbox"/> 1 | <input type="checkbox"/> 3 | <input type="checkbox"/> Нет верного ответа. |

Вопрос 23 ♣ Вероятность $\mathbb{P}(X < 0.5, Y < 0.5)$ равна

- | | | |
|---------------------------------|--|--|
| <input type="checkbox"/> $1/16$ | <input type="checkbox"/> $1/4$ | <input type="checkbox"/> $1/8$ |
| <input type="checkbox"/> $9/16$ | <input checked="" type="checkbox"/> $1/64$ | <input type="checkbox"/> Нет верного ответа. |

В вопросах 24-25 известно, что $\mathbb{E}(X) = 1$, $\text{Var}(X) = 1$, $\mathbb{E}(Y) = 4$, $\text{Var}(Y) = 9$, $\text{Cov}(X, Y) = -3$

Вопрос 24 ♣ Корреляция $\text{Corr}(2X + 3, 4Y - 5)$ равна

- | | | |
|-----------------------------------|--|--|
| <input type="checkbox"/> A $-1/8$ | <input type="checkbox"/> C $1/3$ | <input type="checkbox"/> E $1/6$ |
| <input type="checkbox"/> B 1 | <input checked="" type="checkbox"/> D -1 | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 25 ♣ Ковариация $\text{Cov}(2X - Y, X + 3Y)$ равна

- | | | |
|----------------------------------|---|--|
| <input type="checkbox"/> A 22 | <input checked="" type="checkbox"/> D -40 | <input type="checkbox"/> E 40 |
| <input type="checkbox"/> B -18 | <input type="checkbox"/> C 18 | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 26 ♣ Пусть X_1, X_2, \dots, X_n — последовательность независимых одинаково распределенных случайных величин, $\mathbb{E}(X_i) = \mu$ и $\text{Var}(X_i) = \sigma^2$. Следующее утверждение в общем случае **НЕ ВЕРНО**:

- | |
|---|
| <input type="checkbox"/> A $\frac{\bar{X}_n - \mu}{\sqrt{n}\sigma} \xrightarrow{P} 0$ при $n \rightarrow \infty$ |
| <input checked="" type="checkbox"/> B $\frac{X_n - \mu}{\sigma} \xrightarrow{F} N(0; 1)$ при $n \rightarrow \infty$ |
| <input type="checkbox"/> C $\bar{X}_n \xrightarrow{P} \mu$ при $n \rightarrow \infty$ |
| <input type="checkbox"/> D $\bar{X}_n - \mu \xrightarrow{F} 0$ при $n \rightarrow \infty$ |
| <input type="checkbox"/> E $\lim_{n \rightarrow \infty} \text{Var}(\bar{X}_n) = 0$ |
| <input type="checkbox"/> F $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{F} N(0, 1)$ при $n \rightarrow \infty$ |

Вопрос 27 ♣ Пусть случайные величины X и Y — независимы, тогда **НЕ ВЕРНЫМ** является утверждение

- | | |
|--|---|
| <input type="checkbox"/> A $\mathbb{E}(X Y) = \mathbb{E}(X)$ | <input type="checkbox"/> D $\text{Cov}(X, Y) = 0$ |
| <input type="checkbox"/> B $\mathbb{P}(X < a, Y < b) = \mathbb{P}(X < a)\mathbb{P}(Y < b)$ | <input type="checkbox"/> E $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ |
| <input type="checkbox"/> C $\mathbb{P}(X < a Y < b) = \mathbb{P}(X < a)$ | <input checked="" type="checkbox"/> F $\text{Var}(X - Y) < \text{Var}(X) + \text{Var}(Y)$ |

Вопрос 28 ♣ Если $\mathbb{E}(X) = 0$, то, согласно неравенству Чебышева, $\mathbb{P}(|X| \leq 5\sqrt{\text{Var}(X)})$ лежит в интервале

- | | | |
|---------------------------------------|---|--|
| <input type="checkbox"/> A $[0; 0.2]$ | <input checked="" type="checkbox"/> D $[0.96; 1]$ | <input type="checkbox"/> E $[0.8; 1]$ |
| <input type="checkbox"/> B $[0.5; 1]$ | <input type="checkbox"/> C $[0; 0.04]$ | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 29 ♣ Случайная величина X имеет функцию плотности $f(x) = \frac{1}{3\sqrt{2\pi}} \exp\left(-\frac{(x-1)^2}{18}\right)$. Следующее утверждение **НЕ ВЕРНО**

- | | | |
|--|--|--|
| <input type="checkbox"/> A $\text{Var}(X) = 9$ | <input type="checkbox"/> C $\mathbb{P}(X > 1) = 0.5$ | <input type="checkbox"/> E $\mathbb{E}(X) = 1$ |
| <input checked="" type="checkbox"/> B Случайная величина X дискретна | <input type="checkbox"/> D $\mathbb{P}(X < 0) > 0$ | <input type="checkbox"/> F $\mathbb{P}(X = 0) = 0$ |

Вопрос 30 ♣ Пусть X_1, X_2, \dots, X_n — последовательность независимых одинаково распределенных случайных величин, $E(X_i) = 3$ и $\text{Var}(X_i) = 9$. Следующая величина имеет асимптотически стандартное нормальное распределение

☐ A $\frac{\bar{X}_n - 3}{3\sqrt{n}}$

☒ $\sqrt{n} \frac{\bar{X}_n - 3}{3}$

☐ E $\frac{\bar{X}_n - 3}{3}$

☐ B $\sqrt{n}(\bar{X} - 3)$

☐ D $\frac{X_n - 3}{3}$

☐ F Нет верного ответа.

12.10 Контрольная номер 3

...

12.11 Контрольная номер 4, 05.06.2015

Задача 1 (для первого потока).

Проверка 40 случайно выбранных лекций показала, что студент Халявин присутствовал только на 16 из них.

1. Найдите 95% доверительный интервал для вероятности увидеть Халявина на лекции.
2. На уровне значимости 5% проверьте гипотезу о том, что Халявин посещает в среднем половину лекций.
3. Вычислите минимальный уровень значимости, при котором основная гипотеза отвергается (Р-значение).

Задача 1 (для второго потока).

Вес упаковки с лекарством является нормальной случайной величиной. Взвешивание 20 упаковок показало, что выборочное среднее равно 51 г., а несмещенная оценка дисперсии равна 4.

1. На уровне значимости 10% проверьте гипотезу, что в среднем вес упаковки составляет 55 г.
2. Контрольное взвешивание 30 упаковок такого же лекарства другого производителя показало, что несмещенная оценка дисперсии веса равна 6. На уровне значимости 10% проверьте гипотезу о равенстве дисперсий веса упаковки двух производителей.

Задача 2 (для первого потока).

В ходе анкетирования 15 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. Аналогичные показатели для 12 сотрудников банка «Бета» составили 9.8 и 0.6 часа соответственно.

Считая распределение времени нормальным, на уровне значимости 5% проверьте гипотезу о том, что сотрудники банка «Альфа» в среднем проводят на работе столько же времени, сколько и сотрудники банка «Бета».

Задача 2 (для второго потока).

Экзамен принимают два преподавателя, случайным образом выбирая студентов. По выборке из 85 и 100 наблюдений, выборочные доли не сдавших экзамен студентов составили соответственно 0.2 и 0.17.

1. Можно ли при уровне значимости в 1% утверждать, что преподаватели предъявляют к студентам одинаковый уровень требований?
2. Вычислите минимальный уровень значимости, при котором основная гипотеза отвергается (Р-значение).

Задача 3 (общая).

Методом максимального правдоподобия найдите оценку параметра θ для выборки X_1, \dots, X_n из распределения с функцией плотности

$$f(x) = \begin{cases} \frac{1}{\theta^2} x e^{-\frac{x}{\theta}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Задача 4 (общая).

Пусть X_1, \dots, X_{100} — случайная выборка из нормального распределения с математическим ожиданием μ и дисперсией ν , где μ и ν — неизвестные параметры. По 100 наблюдениям $\sum x_i = 30, \sum x_i^2 = 146, \sum x_i^3 = 122$.

При помощи теста отношения правдоподобия протестируйте гипотезу $H_0 : \nu = 1$ на уровне значимости 5%.

Задача 5 (исследовательская).

Пусть X_1, \dots, X_n — случайная выборка из нормального распределения с математическим ожиданием μ и дисперсией ν , где μ и ν — неизвестные параметры. Рассмотрим три классических теста, отношения правдоподобия, LR , множителей Лагранжа, LM и Вальда, W , для тестирования гипотезы $H_0 : \mu = 0$.

1. Сравните статистики LR , LM и W между собой. Какая — наибольшая, какая — наименьшая?
2. Изменится ли упорядоченность статистик, если проверять гипотезу $H_0 : \mu = \mu_0$?

Подсказка: $\frac{x}{1+x} \leq \ln(1+x) \leq x$ при $x > -1$

Задача 6 (исследовательская).

Величины X_1, \dots, X_n независимы и одинаково распределены с функцией плотности

$$f(x) = \begin{cases} a^2 x e^{-ax}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

По выборке из 100 наблюдений оказалось, что $\sum x_i = 300, \sum x_i^2 = 1000, \sum x_i^3 = 3700$.

1. Найдите оценку неизвестного параметра a методом моментов
2. Используя дельта-метод или иначе оцените дисперсию полученной оценки a
3. Постройте 95%-ый доверительный интервал используя оценку метода моментов

12.12 Экзамен, 15.06.2015

Можно пользоваться простым калькулятором. В каждом вопросе единственный верный ответ.
Ни пуха, ни пера!

Вопрос 31 ♣ Пусть X_1, X_2, \dots, X_{11} — выборка из распределения с математическим ожиданием μ и стандартным отклонением σ . Известно, что $\sum_{i=1}^{11} x_i = 33$, $\sum_{i=1}^{11} x_i^2 = 100$. Несмещенная оценка μ принимает значение

- | | | |
|-----------------------------------|---------------------------------------|--|
| <input type="checkbox"/> A 100/11 | <input checked="" type="checkbox"/> 3 | <input type="checkbox"/> E 0.33 |
| <input type="checkbox"/> B 10 | <input type="checkbox"/> D 3.3 | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 32 ♣ Если X_i независимы, $\mathbb{E}(X_i) = \mu$ и $\text{Var}(X_i) = \sigma^2$, то математическое ожидание величины $Y = \sum_{i=1}^n (X_i - \bar{X})^2$ равно

- | | | |
|---|---|--|
| <input type="checkbox"/> A μ | <input type="checkbox"/> C σ^2/n | <input type="checkbox"/> E $\hat{\sigma}^2$ |
| <input checked="" type="checkbox"/> $(n-1)\sigma^2$ | <input type="checkbox"/> D σ^2 | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 33 ♣ Геродот Геликарнасский проверяет гипотезу $H_0: \mu = 2$. Лог-функция правдоподобия имеет вид $\ell(\mu, \nu) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \nu - \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\nu}$. Оценка максимального правдоподобия для ν при предположении, что H_0 верна, равна

- | | | |
|--|--|---|
| <input type="checkbox"/> A $\frac{\sum x_i^2 - 4 \sum x_i}{n} + 2$ | <input type="checkbox"/> C $\frac{\sum x_i^2 - 4 \sum x_i + 2}{n}$ | <input checked="" type="checkbox"/> $\frac{\sum x_i^2 - 4 \sum x_i}{n} + 4$ |
| <input type="checkbox"/> B $\frac{\sum x_i^2 - 4 \sum x_i}{n}$ | <input type="checkbox"/> D $\frac{\sum x_i^2 - 4 \sum x_i + 4}{n}$ | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 34 ♣ Пусть X_1, X_2, \dots, X_n — случайная выборка размера 36 из нормального распределения $N(\mu, 9)$. Для тестирования основной гипотезы $H_0: \mu = 0$ против альтернативной $H_a: \mu = -2$ вы используете критерий: если $\bar{X} \geq -1$, то вы не отвергаете гипотезу H_0 , в противном случае вы отвергаете гипотезу H_0 в пользу гипотезы H_a . Мощность критерия равна

- | | | |
|---------------------------------|---------------------------------|--|
| <input type="checkbox"/> A 0.58 | <input type="checkbox"/> C 0.85 | <input checked="" type="checkbox"/> 0.98 |
| <input type="checkbox"/> B 0.87 | <input type="checkbox"/> D 0.78 | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 35 ♣ Имеется случайная выборка размера n из нормального распределения. При проверке гипотезы о равенстве математического ожидания заданному значению при известной дисперсии используется статистика, имеющая распределение

- | | | |
|---|---|--|
| <input type="checkbox"/> A χ_{n-1}^2 | <input checked="" type="checkbox"/> $N(0, 1)$ | <input type="checkbox"/> E t_{n-1} |
| <input type="checkbox"/> B χ_n^2 | <input type="checkbox"/> D t_n | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 36 ♣ Ковариационная матрица вектора $X = (X_1, X_2)$ имеет вид $\begin{pmatrix} 10 & 3 \\ 3 & 8 \end{pmatrix}$. Дисперсия разности элементов вектора, $\text{Var}(X_1 - X_2)$, равняется

- | | | |
|-------------------------------|-------------------------------|--|
| <input type="checkbox"/> A 6 | <input type="checkbox"/> C 18 | <input checked="" type="checkbox"/> 12 |
| <input type="checkbox"/> B 15 | <input type="checkbox"/> D 2 | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 37 ♣ Проверая гипотезу о равенстве дисперсий в двух выборках (размером в 3 и 5 наблюдений), Анаксимандр Милетский получил значение тестовой статистики 10. Если оценка дисперсии по одной из выборок равна 8, то другая оценка дисперсии может быть равна

- ☐ A $3/4$
☒ B 80
 ☐ C $4/3$
☐ D 4
 ☐ E 25
 ☐ F Нет верного ответа.

Вопрос 38 ♣ Пусть X_1, \dots, X_{2n} — выборка объема $2n$ из некоторого распределения. Какая из нижеперечисленных оценок математического ожидания имеет наименьшую дисперсию?

- ☐ A $\frac{1}{n} \sum_{i=n+1}^{2n} X_i$
☒ B $\frac{1}{2n} \sum_{i=1}^{2n} X_i$
☐ C $\frac{1}{n} \sum_{i=1}^n X_i$
☐ D X_1
☐ E $\frac{X_1 + X_2}{2}$
☐ F Нет верного ответа.

Вопрос 39 ♣ Величины Z_1, Z_2, \dots, Z_n независимы и нормальны $N(0, 1)$. Случайная величина $Z_1^2 + Z_4^2$ имеет распределение

- ☐ A χ_3^2
☒ B χ_2^2
☐ C t_2
☐ D χ_4^2
☐ E χ_1^2
☐ F Нет верного ответа.

Вопрос 40 ♣ Имеется случайная выборка размера n из нормального распределения. При проверке гипотезы о равенстве дисперсии заданному значению при неизвестном математическом ожидании используется статистика, имеющая распределение

- ☐ A $N(0, 1)$
☐ C t_n
☐ E χ_n^2
☒ B χ_{n-1}^2
☐ D t_{n-1}
☐ F Нет верного ответа.

Вопрос 41 ♣ Производитель мороженого попросил оценить по 10-бальной шкале два вида мороженого: с кусочками шоколада и с орешками. Было опрошено 5 человек.

	Евламий	Аристарх	Капитолина	Аграфена	Эвридика
С крошкой	10	6	7	5	4
С орехами	9	8	8	7	6

Вычислите модуль значения статистики теста знаков. Используя нормальную аппроксимацию, проверьте на уровне значимости 0.05 гипотезу об отсутствии предпочтения мороженого с орешками против альтернативы, что мороженое с орешками вкуснее.

- ☐ A 1.29, H_0 отвергается
 ☐ C 1.65, H_0 отвергается
 ☐ E 1.96, H_0 отвергается
☒ B 1.34, H_0 не отвергается
 ☐ D 1.29, H_0 не отвергается
 ☐ F Нет верного ответа.

Вопрос 42 ♣ Вероятностью ошибки второго рода называется

- ☐ A Вероятность принять неверную гипотезу
☒ B Вероятность отвергнуть альтернативную гипотезу, когда она верна
☐ C Единица минус вероятность отвергнуть основную гипотезу, когда она верна
☐ D Вероятность отвергнуть основную гипотезу, когда она верна
☐ E Единица минус вероятность отвергнуть альтернативную гипотезу, когда она верна
☐ F Нет верного ответа.

Вопрос 43 ♣ Величины Z_1, Z_2, \dots, Z_n независимы и нормальны $N(0, 1)$. Случайная величина $\frac{2Z_1^2}{Z_2^2 + Z_7^2}$ имеет распределение

☐ A $F_{2,7}$

☐ C $F_{7,2}$

☐ E $F_{1,7}$

☐ B t_2

☒ D $F_{1,2}$

☐ F Нет верного ответа.

Вопрос 44 ♣ Зулус Чака каСензангакона проверяет гипотезу о равенстве математических ожиданий в двух нормальных выборках небольших размеров n_1 и n_2 . Если дисперсии неизвестны, но равны, то тестовая статистика имеет распределение

☒ A $t_{n_1+n_2-2}$

☐ C $\chi_{n_1+n_2-1}^2$

☐ E F_{n_1, n_2}

☐ B $t_{n_1+n_2-1}$

☐ D $t_{n_1+n_2}$

☐ F Нет верного ответа.

Вопрос 45 ♣ Пусть X_1, \dots, X_n — выборка объема n из равномерного на $[a, b]$ распределения. Оценка $X_1 + X_2$ параметра $c = a + b$ является

☐ A асимптотически несмещенной и состоятельной

☐ D смещенной и состоятельной

☒ B несмещенной и несостоятельной

☐ E смещенной и несостоятельной

☐ C несмещенной и состоятельной

☐ F Нет верного ответа.

Вопрос 46 ♣ Пусть X_1, \dots, X_n — выборка объема n из равномерного на $[0, \theta]$ распределения. Оценка параметра θ методом моментов по k -му моменту имеет вид:

☒ A $\sqrt[k]{(k+1)\bar{X}^k}$

☐ C $\sqrt[k]{k\bar{X}^k}$

☐ E $\sqrt[k]{(k+1)\bar{X}^k}$

☐ B $\sqrt[k+1]{(k+1)\bar{X}^k}$

☐ D $\sqrt[k]{k\bar{X}^k}$

☐ F Нет верного ответа.

Вопрос 47 ♣ Среди 100 случайно выбранных ацтеков 20 платят дань Кулуакану, а 80 — Аскапоцалько. Соответственно, оценка доли ацтеков, платящих дань Кулуакану, равна $\hat{p} = 0.2$. Разумная оценка стандартного отклонения случайной величины \hat{p} равна

☐ A 0.016

☒ B 0.04

☐ E 1.6

☐ B 0.16

☐ D 0.4

☐ F Нет верного ответа.

Вопрос 48 ♣ По случайной выборке из 100 наблюдений было оценено выборочное среднее $\bar{X} = 20$ и несмещенная оценка дисперсии $\hat{\sigma}^2 = 25$. В рамках проверки гипотезы $H_0: \mu = 15$ против альтернативной гипотезы $H_a: \mu > 15$ можно сделать следующее заключение

☐ A Гипотеза H_0 отвергается на уровне значимости 10%, но не на уровне значимости 5%

☐ B Гипотеза H_0 отвергается на уровне значимости 20%, но не на уровне значимости 10%

☐ C Гипотеза H_0 не отвергается на любом разумном уровне значимости

☒ D Гипотеза H_0 отвергается на любом разумном уровне значимости

☐ E Гипотеза H_0 отвергается на уровне значимости 5%, но не на уровне значимости 1%

☐ F Нет верного ответа.

Вопрос 49 ♣ Пусть X_1, X_2, \dots, X_{11} — выборка из распределения с математическим ожиданием μ и стандартным отклонением σ . Известно, что $\sum_{i=1}^{11} x_i = 33$, $\sum_{i=1}^{11} x_i^2 = 100$. Несмещенная оценка дисперсии принимает значение

- ☐ A 1/11 ☒ 1/10 ☐ E 11/100
☐ B 10 ☐ D 100/11 ☐ F Нет верного ответа.

Вопрос 50 ♣ Николай Коперник подбросил бутерброд 200 раз. Бутерброд упал маслом вниз 95 раз, а маслом вверх — 105 раз. Значение критерия χ^2 Пирсона для проверки гипотезы о равной вероятности данных событий равно

- ☐ A 0.75 ☐ C 7.5 ☐ E 0.25
☐ B 2.5 ☒ 0.5 ☐ F Нет верного ответа.

Вопрос 51 ♣ Если Р-значение (P-value) больше уровня значимости α , то гипотеза $H_0 : \sigma = 1$

- ☒ Не отвергается ☐ D Отвергается
☐ B Отвергается, только если $H_a : \sigma \neq 1$ ☐ E Отвергается, только если $H_a : \sigma < 1$
☐ C Отвергается, только если $H_a : \sigma > 1$ ☐ F Нет верного ответа.

Вопрос 52 ♣ Функция правдоподобия, построенная по случайной выборке X_1, \dots, X_n из распределения с функцией плотности $f(x) = (\theta + 1)x^\theta$ при $x \in [0; 1]$ имеет вид

- ☐ A $(\theta + 1)^{\sum x_i}$ ☐ D $(\sum x_i)^\theta$
☐ B $(\theta + 1)x^{n\theta}$ ☒ $(\theta + 1)^n \prod x_i^\theta$
☐ C $\sum(\theta + 1)x_i^\theta$ ☐ F Нет верного ответа.

Вопрос 53 ♣ Последовательность оценок $\hat{\theta}_1, \hat{\theta}_2, \dots$ называется состоятельной, если

- ☒ $\mathbb{P}(|\hat{\theta}_n - \theta| > t) \rightarrow 0$ для всех $t > 0$ ☐ D $\mathbb{E}(\hat{\theta}_n) = \theta$
☐ B $\text{Var}(\hat{\theta}_n) \geq \text{Var}(\hat{\theta}_{n+1})$ ☐ E $\mathbb{E}(\hat{\theta}_n) \rightarrow \theta$
☐ C $\text{Var}(\hat{\theta}_n) \rightarrow 0$ ☐ F Нет верного ответа.

Вопрос 54 ♣ Каждое утро в 8:00 Иван Андреевич Крылов, либо завтракает, либо уже позавтракал. В это же время кухарка либо заглядывает к Крылову, либо нет. По таблице сопряженности вычислите статистику χ^2 Пирсона для тестирования гипотезы о том, что визиты кухарки не зависят от того, позавтракал ли уже Крылов или нет.

Время 8:00	кухарка заходит	кухарка не заходит
Крылов завтракает	200	40
Крылов уже позавтракал	25	100

- ☐ A 79 ☒ 139 ☐ E 39
☐ B 179 ☐ D 100 ☐ F Нет верного ответа.

Вопрос 55 ♣ Геродот Геликарнасский проверяет гипотезу $H_0 : \mu = 0, \sigma^2 = 1$ с помощью LR статистики теста отношения правдоподобия. При подстановке оценок метода максимального правдоподобия в лог-функцию правдоподобия он получил $\ell = -177$, а при подстановке $\mu = 0$ и $\sigma = 1$ оказалось, что $\ell = -211$. Найдите значение LR статистики и укажите её закон распределения при верной H_0

☐ A $LR = 34, \chi^2_2$

☒ B $LR = 68, \chi^2_2$

☐ E $LR = \ln 68, \chi^2_{n-2}$

☐ B $LR = 34, \chi^2_{n-1}$

☐ D $LR = \ln 34, \chi^2_{n-2}$

☐ F Нет верного ответа.

Вопрос 56 ♣ Пусть $\hat{\sigma}_1^2$ — несмещенная оценка дисперсии, полученная по первой выборке размером n_1 , $\hat{\sigma}_2^2$ — несмещенная оценка дисперсии, полученная по второй выборке, с меньшим размером n_2 . Тогда статистика $\frac{\hat{\sigma}_1^2/n_1}{\hat{\sigma}_2^2/n_2}$ имеет распределение

☐ A $t_{n_1+n_2-1}$

☐ C F_{n_1-1, n_2-1}

☐ E $N(0; 1)$

☐ B F_{n_1, n_2}

☐ D $\chi^2_{n_1+n_2}$

☒ F Нет верного ответа.

Вопрос 57 ♣ На основе случайной выборки, содержащей одно наблюдение X_1 , тестируется гипотеза $H_0 : X_1 \sim U[0; 1]$ против альтернативной гипотезы $H_a : X_1 \sim U[0.5; 1.5]$. Рассматривается критерий: если $X_1 > 0.8$, то гипотеза H_0 отвергается в пользу гипотезы H_a . Вероятность ошибки 2-го рода для этого критерия равна:

☐ A 0.5

☒ B 0.3

☐ E 0.1

☐ B 0.2

☐ D 0.4

☐ F Нет верного ответа.

Вопрос 58 ♣ Бессмертный гений поэзии Ли Бо оценивает математическое ожидание по выборке размера n из нормального распределения. Он построил оценку метода моментов, $\hat{\mu}_{MM}$, и оценку максимального правдоподобия, $\hat{\mu}_{ML}$. Про эти оценки можно утверждать, что

☐ A они не равны, и не сближаются при $n \rightarrow \infty$

☐ D они не равны, но сближаются при $n \rightarrow \infty$

☐ B $\hat{\mu}_{MM} > \hat{\mu}_{ML}$

☐ E $\hat{\mu}_{MM} < \hat{\mu}_{ML}$

☒ F они равны

☐ F Нет верного ответа.

Вопрос 59 ♣ Все условия регулярности для применения метода максимального правдоподобия выполнены. Вторая производная лог-функции правдоподобия равна $\ell''(\theta) = -100$. Дисперсия несмещенной эффективной оценки для параметра θ равна

☐ A 100

☐ C 1

☒ B 0.01

☐ B 0.1

☐ D 10

☐ F Нет верного ответа.

Вопрос 60 ♣ Адтек Монтесума Илуикамина хочет оценить параметр a методом максимального правдоподобия по выборке из неотрицательного распределения с функцией плотности $f(x) = \frac{1}{2}a^3x^2e^{-ax}$ при $x \geq 0$. Для этой цели ему достаточно максимизировать функцию

☐ A $3n \sum \ln a_i - a \sum \ln x_i$

☐ C $3n \prod \ln a - ax^n$

☐ E $3n \ln a - an \ln x_i$

☒ B $3n \ln a - a \sum x_i$

☐ D $3n \ln a - a \prod \ln x_i$

☐ F Нет верного ответа.

Вопрос 61 ♣ У пяти случайно выбранных студентов первого потока результаты за контрольную по статистике оказались равны 82, 47, 20, 43 и 73. У четырёх случайно выбранных студентов второго потока — 68, 83, 60 и 52. Вычислите статистику Вилкоксона и проверьте гипотезу H_0 об однородности результатов студентов двух потоков. Критические значения статистики Вилкоксона равны $T_L = 12$ и $T_R = 28$.

- | | | |
|---|--|--|
| <input type="checkbox"/> A 65.75, H_0 отвергается | <input checked="" type="checkbox"/> B 24, H_0 не отвергается | <input type="checkbox"/> E 12.75, H_0 не отвергается |
| <input type="checkbox"/> B 53, H_0 отвергается | <input type="checkbox"/> D 20, H_0 не отвергается | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 62 ♣ Вероятность ошибки первого рода, α , и вероятность ошибки второго рода, β , всегда связаны соотношением

- | | | |
|--|--|---|
| <input type="checkbox"/> A $\alpha + \beta \geq 1$ | <input type="checkbox"/> C $\alpha \geq \beta$ | <input type="checkbox"/> E $\alpha + \beta = 1$ |
| <input type="checkbox"/> B $\alpha + \beta \leq 1$ | <input type="checkbox"/> D $\alpha \leq \beta$ | <input checked="" type="checkbox"/> F Нет верного ответа. |

Вопрос 63 ♣ Датчик случайных чисел выдал следующие значения псевдо случайной величины: 0.78, 0.48. Вычислите значение критерия Колмогорова и проверьте гипотезу H_0 о соответствии распределения равномерному на $[0; 1]$. Критическое значение статистики Колмогорова для уровня значимости 0.1 и двух наблюдений равно 0.776.

- | | | |
|--|--|---|
| <input type="checkbox"/> A 1.26, H_0 отвергается | <input type="checkbox"/> C 0.78, H_0 отвергается | <input type="checkbox"/> E 0.37, H_0 не отвергается |
| <input checked="" type="checkbox"/> B 0.48, H_0 не отвергается | <input type="checkbox"/> D 0.3, H_0 не отвергается | <input type="checkbox"/> F Нет верного ответа. |

Вопрос 64 ♣ Величины Z_1, Z_2, \dots, Z_n независимы и нормальны $N(0, 1)$. Случайная величина $\frac{Z_1 \sqrt{n-3}}{\sqrt{\sum_{i=4}^n Z_i^2}}$ имеет распределение

- | | | |
|---|---|--|
| <input checked="" type="checkbox"/> A t_{n-3} | <input type="checkbox"/> C χ_{n-4}^2 | <input type="checkbox"/> E $N(0, 1)$ |
| <input type="checkbox"/> B t_{n-1} | <input type="checkbox"/> D $F_{1, n-2}$ | <input type="checkbox"/> F Нет верного ответа. |

13 2015-2016

13.1 Контрольная номер 1, базовый поток

13.2 Праздник номер 1, исследователи, индивидуальный тур

13.3 Регата, исследователи, командный тур

13.4 Контрольная номер 2, поток Арктика, 12.12.2015

Продолжительность: 1 час 20 минут

1. Функция плотности случайного вектора $\xi = (\xi_1, \xi_2)^T$ имеет вид

$$f(x, y) = \begin{cases} 0.5x + 1.5y, & \text{если } 0 < x < 1, 0 < y < 1 \\ 0, & \text{иначе} \end{cases}$$

Найдите:

- Математическое ожидание $\mathbb{E}(\xi_1 \cdot \xi_2)$
- Условную плотность распределения $f_{\xi_1|\xi_2}(x|y)$

- (с) Условное математическое ожидание $\mathbb{E}(\xi_1 | \xi_2 = y)$
- (d) Константу k , такую, что функция $h(x, y) = kx \cdot f(x, y)$ будет являться совместной функцией плотности некоторой пары случайных величин
2. На курсе учатся очень много студентов. Вероятность того, что случайно выбранный студент по результатам рубежного контроля имеет хотя бы один незачет равна 0.2. Пусть ξ и η — число студентов с незачетами и без незачетов в случайной группе из 10 студентов. Найдите $\text{Cov}(\xi, \eta)$, $\text{Corr}(\xi, \eta)$, $\text{Cov}(\xi - \eta, \xi)$. Являются ли случайные величины $\xi - \eta$ и ξ независимыми?
3. Доходности акций компаний А и В — случайные величины ξ и η . Известно, что $\mathbb{E}(\xi) = 1$, $E(\eta) = 1$, $\text{Var}(\xi) = 4$, $\text{Var}(\eta) = 9$, $\text{Corr}(\xi, \eta) = -0.5$. Петя принимает решение потратить свой рубль на акции компании А, Вася — 50 копеек на акции компании А и 50 копеек на акции компании В, а Маша принимает решение вложить свой рубль в портфель $R = \alpha\xi + (1 - \alpha)\eta$, ($0 \leq \alpha \leq 1$), обладающий минимальным риском. Найдите α , ожидаемые доходности и риски портфелей Пети, Васи и Маши.
4. Будем считать, что рождение мальчика и девочки равновероятны.
- (а) Оцените с помощью неравенства Маркова вероятность того, что среди тысячи новорожденных младенцев, мальчиков будет более 75%.
- (б) Оцените с помощью неравенства Чебышёва вероятность того, что доля мальчиков среди тысячи новорожденных младенцев будет отличаться от 0.5 более, чем на 0.25
- (с) С помощью теоремы Муавра-Лапласа вычислите вероятность из предыдущего пункта.
5. Сейчас валютный курс племени «Мумба» составляет 100 оболов за один рубль. Изменение курса за один день — случайная величина δ_i с законом распределения:

δ_i	-1	0	2
$\mathbb{P}(\cdot)$	0.25	0.5	0.25

Найдите вероятность того, что через полгода (171 день) рубль будет стоить более 250 оболов, если ежедневные изменения курса происходят независимо друг от друга.

6. Бонусная задача

Число посетителей, зашедших в магазин в течении дня — пуассоновская случайная величина с параметром λ . Каждый из посетителей совершает покупку с вероятностью p , не зависимо от других посетителей. Найдите математическое ожидание числа человек, совершивших покупку.

13.5 Контрольная номер 2, поток Риччи, 12.12.2015

Продолжительность: 1 час 20 минут

1. Функция плотности случайного вектора $\xi = (\xi_1, \xi_2)^T$ имеет вид

$$f(x, y) = \begin{cases} 0.5x + 1.5y, & \text{если } 0 < x < 1, 0 < y < 1 \\ 0, & \text{иначе} \end{cases}$$

Найдите:

- (a) Математическое ожидание $\mathbb{E}(\xi_1 \cdot \xi_2)$
- (b) Условную плотность распределения $f_{\xi_1|\xi_2}(x|y)$
- (c) Условное математическое ожидание $\mathbb{E}(\xi_1|\xi_2 = y)$
- (d) Константу k , такую, что функция $h(x, y) = kx \cdot f(x, y)$ будет являться совместной функцией плотности некоторой пары случайных величин
2. На курсе учится очень много студентов. Вероятность того, что случайно выбранный студент получит «отлично» за контрольную равна 0.2, «хорошо» — 0.3. Вероятности остальных результатов неизвестны. Пусть ξ и η — число отличников и хорошистов в случайной группе из 10 студентов. Найдите $\text{Cov}(\xi, \eta)$, $\text{Corr}(\xi, \eta)$, $\text{Cov}(\xi - \eta, \xi)$. Являются ли случайные величины $\xi - \eta$ и ξ независимыми?
3. Доходности акций компаний А и В — случайные величины ξ и η . Известно, что $\mathbb{E}(\xi) = 1$, $E(\eta) = 1$, $\text{Var}(\xi) = 4$, $\text{Var}(\eta) = 9$, $\text{Corr}(\xi, \eta) = -0.5$. Петя принимает решение потратить свой рубль на акции компании А, Вася — 50 копеек на акции компании А и 50 копеек на акции компании В, а Маша принимает решение вложить свой рубль в портфель $R = \alpha\xi + (1 - \alpha)\eta$, ($0 \leq \alpha \leq 1$), обладающий минимальным риском. Найдите α , ожидаемые доходности и риски портфелей Пети, Васи и Маши.
4. Будем считать, что рождение мальчика и девочки равновероятны.
- (a) С помощью закона больших чисел определите в каком городе, большом или маленьком, больше случается таких дней, когда рождается более 75% мальчиков.
- (b) Оцените с помощью неравенства Маркова вероятность того, что среди тысячи новорожденных младенцев мальчиков будет более 75%.
- (c) Оцените с помощью неравенства Чебышёва вероятность того, что доля мальчиков среди тысячи новорожденных младенцев будет отличаться от 0.5 более, чем на 0.25
- (d) С помощью теоремы Муавра-Лапласа вычислите вероятность из предыдущего пункта.
5. Сейчас валютный курс племени «Мумба» составляет 100 оболов за один рубль. Процентное изменение курса за один день — случайная величина δ_i с законом распределения:

δ_i	-1%	1%
$\mathbb{P}(\cdot)$	0.25	0.75

Найдите вероятность того, что через полгода (171 день) рубль будет стоить более 271 оболоа, если ежедневные изменения курса происходят независимо друг от друга.

6. Величины X_1, X_2, \dots независимы и равновероятно принимают значения -1 и 3 .

(a) Найдите $\text{plim}_{n \rightarrow \infty} \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$

(b) С помощью дельта-метода найдите примерный закон распределения $\frac{\sum_{i=1}^{100} (X_i - \bar{X})^2}{100}$

13.6 Контрольная номер 2, поток Риччи, 12.12.2015, решение

Решение: Аршак Минасян

1. a) [2 pts]

$$\begin{aligned}\mathbb{E}(\xi_1 \cdot \xi_2) &= \int_0^1 \int_0^1 xy f(x, y) dx dy = \int_0^1 \int_0^1 \frac{1}{2} \cdot x^2 y + \frac{3}{2} \cdot xy^2 dx dy = \\ &= \int_0^1 \frac{y}{6} + \frac{3y^2}{4} dy = \frac{1}{3}\end{aligned}$$

b) [3 pts]

$$f_{\xi_1|\xi_2}(x|y) = \frac{f_{\xi_1, \xi_2}(x, y)}{f_{\xi_2}(y)} = \frac{0.5x + 1.5y}{0.25 + 1.5y}, \text{ при } y \in (0, 1)$$

c) [3 pts]

$$\mathbb{E}(\xi_1|\xi_2 = y) = \int_0^1 x f_{\xi_1|\xi_2}(x|y) dx = \frac{1}{0.5 + 1.5y} \times \left(\frac{x^2}{4} + 1.5xy \right) \Big|_0^1 = \frac{0.25 + 1.5y}{0.5 + 1.5y}$$

d) [2 pts] Для того, чтобы функция являлась совместной плотностью для пары случайных величин, должно выполняться следующее:

$$\int_{\Omega} kx f(x, y) dx dy = 1$$

Вычислим чему равняется левая часть:

$$\begin{aligned}1 &= \int_{\Omega} kx f(x, y) dx dy = \int_0^1 \int_0^1 kx \left(\frac{x + 3y}{2} \right) dx dy = \int_0^1 \frac{k}{6} + \frac{3ky}{4} dy = \frac{k}{6} + \frac{3k}{8} \Rightarrow \\ k &= \frac{24}{13}\end{aligned}$$

2. При расчёте ковариации применим разложение случайной величины в сумму простых случайных величин!

[4 pts]:

$$\begin{aligned}\text{Cov}(\xi, \eta) &= \text{Cov}(\xi_1 + \dots + \xi_{10}, \eta_1 + \dots + \eta_{10}) = \\ &= 10 \text{Cov}(\xi_1, \eta_1) = 10(0 - 0.2 \cdot 0.3) = -0.6\end{aligned}$$

[1 pt]:

$$\text{Var}(\xi) = 10 \cdot 0.2 \cdot 0.8 = 1.6$$

[1 pt]:

$$\text{Var}(\eta) = 10 \cdot 0.3 \cdot 0.7 = 2.1$$

[2 pts]:

$$\text{Corr}(\xi, \eta) = -0.6 / \sqrt{1.6 \cdot 2.1} \approx -0.33$$

[1 pt]:

$$\text{Cov}(\xi - \eta, \xi) = \text{Var}(\xi) - \text{Cov}(\xi, \eta) \neq 0,$$

[1 pt]: Следовательно ξ и η зависимы.

3. Найдем ожидаемую доходность и риск портфеля $R = \alpha\xi + (1 - \alpha)\eta$ для любого α , тогда при $\alpha = 1$ получим результаты Пети, при $\alpha = 0.5$ — результаты Васи. [2 pts]

$$\mathbb{E}R = \alpha + (1 - \alpha) = 1 \quad \forall \alpha \in [0, 1]$$

Находим дисперсию: [3 pts]

$$\text{Var}(R) = \alpha^2 \cdot 4 + (1 - \alpha)^2 \cdot 9 - 6\alpha(1 - \alpha) = 19\alpha^2 - 24\alpha + 9 \rightarrow \min_{\alpha} \Rightarrow$$

Теперь, найдем оптимальное α : [2 pts]

$$\alpha = \frac{24}{38}$$

Финальные цифры: [3 pts]

$$\begin{cases} \text{Var}(R)^P = 4 \Rightarrow \sigma_P = 2 \\ \text{Var}(R)^V = 1.75 \Rightarrow \sigma_V \approx 1.32 \\ \text{Var}(R)^M = \frac{27}{19} \Rightarrow \sigma_M \approx 1.19 \end{cases}$$

4. (a) По ЗБЧ имеем: [2 pts]

$$\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow \mathbb{E}(\xi_1) = \frac{1}{2},$$

Поэтому, чем больше n (количество жителей в городе), тем меньше таких дней, когда количество мальчиков больше 75%.

- (b) Пусть S количество мальчиков, тогда используя **неравенство Маркова** получаем: [2 pts]

$$\mathbb{P}(S \geq 750) \leq \frac{\mathbb{E}(S)}{750} = \frac{2}{3}$$

- (c) Пусть, теперь, \bar{X} доля мальчиков, то есть, $\bar{X} = \sum_{i=1}^n X_i/n$, где

$$X_i = \begin{cases} 1, & \text{если } i\text{-ый ребёнок — мальчик} \\ 0, & \text{иначе} \end{cases}$$

тогда используя **неравенство Чебышева** получаем: [3 pts]

$$\mathbb{P}(|\bar{X} - 0.5| \geq 0.25) \leq \frac{\text{Var}(\bar{X})}{0.25^2} = \frac{1/4000}{0.25^2} = 0.004$$

- (d) Вероятность из предыдущего пункта можно записать в таком виде: [3 pts]

$$\begin{aligned} \mathbb{P}(|\bar{X} - 0.5| \geq 0.25) &= \mathbb{P}(\bar{X} \geq 0.75) + \mathbb{P}(\bar{X} \leq 0.25) = 2\mathbb{P}(\bar{X} \geq 0.75) = \\ &= 2\mathbb{P}(\mathcal{N}(0; 1) \geq 0.25\sqrt{4000}) = 2\mathbb{P}(\mathcal{N}(0; 1) \geq 15.8) = 1.3 \cdot 10^{-56} \approx 0 \end{aligned}$$

5. Ищем вероятность

$$100 \cdot X_1 \cdot \dots \cdot X_{171} \geq 271$$

Здесь X_i принимают значения 0.99 или 1.01 с вероятностями 0.25 и 0.75

Берем логарифм:

$$\sum \log X_i \geq 1$$

Исходим из худшего случая, когда на калькуляторе нет логарифма, тогда неплохо знать, что $\log(1 + \alpha) \sim \alpha$, поэтому можно считать, что $\log X_i$ принимает значения -0.01 и 0.01 .

Значит $\mathbb{E}(\log X_i) = 1/200$, $\text{Var}(\log X_i) = 1/100 - 1/200^2 \approx 1/100$.

Поэтому сумма $S \sim \mathcal{N}(171/200; 171/100)$ и

$$\mathbb{P}(S \geq 1) = \mathbb{P}(\mathcal{N}(0; 1) \geq 0.11) \approx 0.46$$

Разбалловка: идея что проценты — это умножение [2 pts]

переход к логарифмам [2 pts]

переход к вероятности для суммы [2 pts]

расчёт ожидания и дисперсии суммы [2 pts]

стандартизация и таблица [2 pts]

6. (a) Используем ЗБЧ [4 pts]:

$$\text{plim}_{n \rightarrow \infty} \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} = \text{plim}_{n \rightarrow \infty} \frac{\sum X_i^2}{n} - \text{plim}_{n \rightarrow \infty} \bar{X} \cdot \bar{X} = \mathbb{E}(X_1^2) - (\mathbb{E}(X_1))^2 = \text{Var}(X_1) = 4$$

(b) Обозначим, $Y_i = X_i^2$, тогда наше выражение можно записать в виде: [6 pts]

$$Q = \bar{Y} - (\bar{X})^2$$

Причём, $\text{plim } \bar{Y} = 5$, $\text{plim } \bar{X} = 1$.

Согласно дельта-методу заменяем его на линейную аппроксимацию в окрестности предела:

$$Q \approx 4 + (\bar{Y} - 5) + 2(\bar{X} - 1)$$

Стало быть, при больших n :

$$\mathbb{E}(Q) \approx 4$$

$$\begin{aligned} \text{Var}(Q) &\approx \text{Var}(\bar{Y}) + 4 \text{Var}(\bar{X}) + 4 \text{Cov}(\bar{Y}, \bar{X}) = \\ &= \frac{1}{n} (\text{Var}(Y_1) + 4 \text{Var}(X_1) + 4 \text{Cov}(X_1, Y_1)) = \\ &= 0.01 \cdot (16 + 4 \cdot 4 + 4 \cdot 8) = 0.64 \quad (51) \end{aligned}$$

Итого, $Q \approx \mathcal{N}(4; 0.64)$