г) Постройте доверительный интервал для разницы долей $p_1 - p_2$.

4. Винни-Пух лично измерил количество мёда (в кг) на 100 деревьях и обнаружил, что $\bar{X} = 10$ и $\hat{\sigma}^2 = 4$. По мнению Кролика, состоятельная оценка для параметра α правильности мёда имеет вид $\hat{\alpha} = \bar{X} + \sqrt{\bar{X} + 6}$.

   а) «Халява, сэр!» Найдите точечную оценку параметра α;

   б) Найдите 95%-ый доверительный интервал для α, симметричный относительно $\hat{\alpha}$.

5. Фотографы Андрей и Белла независимо друг от друга пытаются фотографировать кадьяков. Андрею удаётся сфотографировать одного кадьяка в неделю с вероятностью 0.5, а Белле — с вероятностью p, независимо друг от друга и от прошлого. За 100 недель они вместе сфотографировали 130 кадьяков.

   а) Оцените p и постройте 95%-ый доверительный интервал для p;

   б) Оцените p и постройте 95%-ый доверительный интервал для p, если дополнительно известно, что один фотограф опередил другого на 10 фото.

Просто красивая задачка. Эту задачу **не нужно** решать на кр :)
Медведю Мишутке никак не удаётся заснуть в берлоге, и потому он подбрасывает правильную монетку n раз. Обозначим вероятность того, что ни разу не идёт двух решек подряд буквой q_n.

а) Найдите $2^8 q_8$ и **назовите** это число;

б) Найдите $\lim 2q_{n+1}/q_n$ и **назовите** это число.

8.3. 2016-2017

Главная мораль: байесовский подход — это всего лишь формула условной вероятности.

1. Задача о целебных лягушках :)

   У одного вида лягушек самки обладают целебными свойствами. Самцы и самки встречаются равновероятно. Неподалёку видны аж две лягушки данного вида, но издалека неясно кто.

   Определите вероятность того, что среди этих лягушек есть хотя бы одна целебная, в каждой из ситуаций:

   а) Самцы квакают, самки — нет, со стороны лягушек слышно кваканье, но не разобрать, одной лягушки или двух.

   б) Самцы и самки квакают по разному, но одинаково часто. Только что послышался отдельный квак одной из лягушек и это квак самца.

2. Яичный бой

   Саша и Маша играют в «яичный бой». Перед ними корзина яиц. В начале боя они берут по одному яйцу и бьют их острыми концами. Каждое яйцо в корзине обладает своей «силой», все силы — разные. Более сильное яйцо разбивает более слабое. Внешне яйца не отличимы. Сила яйца не убывает при ударах. Разбитое яйцо выбрасывают, побеждённый берёт новое, а победитель продолжает играть прежним.

   Какова вероятность того, что Маша победит в 11-ом раунде, если она уже победила 10 раундов подряд?

3. Классика жанра

   Перед нами определение бета-распределения $Beta(\alpha, \beta)$:

$$f(x) \propto \begin{cases} x^{\alpha-1}(1-x)^{\beta-1}, \text{ если } x \in [0;1] \\ 0, \text{ иначе.} \end{cases}$$

   Блондинка Анжелика хочет оценить неизвестную вероятность встретить динозавра, p. Она предполагает, что динозавры встречаются каждый день независимо от других с постоянной вероятностью. Априорно Анжелика считает, что неизвестное p имеет бета-распределение $Beta(2, 3)$. За 20 дней Анжелика 5 раз видела динозавра. Для краткости обозначим вектором y все имеющиеся наблюдения. Величина y_i — результат i-го дня: 1, если динозавр встретился, и 0 иначе.

   а) Чему, по-мнению Анжелики, равны априорные $\mathbb{E}(p)$, мода распределения p?

   б) Найдите апостериорное распределение $f(p|y)$.

   в) Найдите апостериорные ожидание $\mathbb{E}(p|y)$ и моду.

   г) Найдите условное распределение y_{21} с учётом имеющихся данных.

4. Рассмотрим следующий код в stan.

```
data {
  int<lower=1> N_x;
  int<lower=1> N_y;
  real y[N_y];
  real x[N_x];
}
parameters {
  real mu_x;
  real mu_y;
  real<lower=0> sigma_x;
  real<lower=0> sigma_y;
}
model {
  for (n_x in 1:N_x) {
    x[n_x] ~ normal(mu_x, sigma_x);
  }
  for (n_y in 1:N_y) {
    y[n_y] ~ normal(mu_y, sigma_y);
  }
  mu_x ~ normal(0, 100);
  mu_y ~ normal(0, 100);
7   sigma_y ~ exponential(50);
}
generated quantities {
  delta = mu_x - mu_y;
  ratio = sigma_x / sigma_y;
}
```

а) Выпишите предполагаемую модель для данных.

б) Выпишите априорное распределение.

в) Байесовский интервал для каких величин позволяет построить данный код?

г) Какие предпосылки мешают применить в данном случае классический доверительный интервал для разности математических ожиданий, основанный на F-распределении?

5. Просто красивая задача про выборку :)

Есть неизвестное количество чисел. Среди этих чисел одно число встречается строго больше 50% раз. Ведущий показывает числа исследователю Акану в некотором порядке. Когда все числа закончатся, ведущий скажет «всё». Задача Акана — определить, какое число встречается чаще всех. Проблема в том, что Акан так готовился к контрольной по теории вероятностей, что устал. И больше 10 чисел запомнить не в состоянии.

Предложите алгоритм, которой позволит Акану определить искомое число.

8.4. 2015-2016

Правила: 3 часа, всем можно пользоваться, интернетом тоже.
Все семь задач решать вовсе не обязательно, выбирайте любые пять! При самостоятельной работе можно всем пользоваться! :)

1. Случайные величины X_1, ..., X_n независимо и одинаково распределены с функцией плотности $f(x) = 2ax\exp(-ax^2)$ при $x > 0$.

   По 100 наблюдениям известно, что $\sum X_i = 170$, $\sum X_i^2 = 350$.

   а) Оцените параметр a методом максимального правдоподобия.
   б) Оцените дисперсию оценки $\hat{a}_{ML}$
   в) Постройте 95%-ый доверительный интервал для a с помощью оценки максимального правдоподобия
   г) Оцените параметр a методом моментов
   д) Оцените дисперсию оценки $\hat{a}_{MM}$
   е) Постройте 95%-ый доверительный интервал для a с помощью оценки метода моментов

2. Для того, чтобы люди давали правдивый ответ на деликатный вопрос (скажем, «Берёте ли Вы взятки?») при опросе используется рандомизация. Вопрос допускает всего два ответа «да» или «нет». Перед ответом респондент подбрасывает монетку, и только респондент видит результат подбрасывания. Если монетка выпадет «орлом», то респондент отвечает правду. Если «решкой», то респондент отвечает наоборот («да» вместо «нет» и «нет» вместо «да»).

   Монетка выпадает орлом с вероятностью 0.4. Из 500 опрошенных 300 ответили «да».

   а) Какова вероятность того, что человек берёт взятки, если он ответил «да» в анкете?
   б) Постройте оценку для доли людей берущих взятки
   в) Постройте 95%-ый доверительный интервал для доли людей берущих взятки

3. Винни-Пух хочет измерить высоту Большого дуба, d. Для этого Винни-Пух три раза в случайное время дня измерил длину тени Большого Дуба: 8.9, 13.2, 25.2.

   Предположим, что в дни измерений траектория движения Солнца проходила ровно через зенит :)

   а) Найдите функцию плотности длины тени
   б) Если возможно, постройте оценку метода моментов
   в) Если возможно, постройте оценку метода максимального правдоподобия
   г) Где живёт Винни-Пух и какого числа 2016 года он проводил измерения?

4. Встроенный в R набор данных morley содержит результаты 100 опытов Майкельсона и Морли. В 1887 году они проводили измерения скорости света, чтобы понять, зависит ли она от направления.

   а) Постройте 95%-ый доверительный интервал для скорости света
   б) Выпишите использованные формулы и алгоритм построения интервала
   в) Чётко сформулируйте все гипотезы при которых данный алгоритм даёт корректный результат
   г) Накрывает ли построенный доверительный интервал фактическую скорость света?

   Полезные команды: morley, help("morley"), mean, sd, qnorm, pnorm

5. Исследователь Вениамин дрожащей от волнения рукой рисует прямоугольники размера $a \times b$. Поскольку Вениамин очень волнуется прямоугольники де-факто выходят со случайными сторонами $a + u_i$ и $b + v_i$. Случайные ошибки u_i и v_i независимы и одинаково распределены $\mathcal{N}(0; 1)$.

   Вениамин нарисовал 400 прямоугольничков и посчитал очень аккуратно площадь каждого. Оказалась, что средняя площадь равна 1200 см^2, а выборочное стандартное отклонение площади — 50 см^2. Вениамин считает, что зная только площади прямоугольничков невозможно оценить оценить каждую из сторон.

   Если возможно, то оцените параметры a и b подходящим методом. Если невозможно, то докажите.

6. На поле $D4$ шахматной доски стоит конь. Ли Седоль переставляет коня наугад, выбирая каждый возможный ход равновероятно.

   Сколько в среднем пройдет ходов прежде чем Ли Седоль снова вернёт коня на $D4$?

7. В «Киллер» играли n человек. После окончания игры, когда были убиты все, кто может быть убит, встретились два игрока (возможно убитых) и оказалось, что один убил 5 человек, а другой — 7 человек.

   Оцените n подходящим методом.

8.5. 2013-2014

1. Дед Мазай подбирает зайцев. Предположим, что длина левого уха зайца имеет экспоненциальное распределение с плотностью $f(x) = a\exp(-ax)$ при $x \geqslant 0$. По 100 зайцам оказалось, что $\sum x_i = 2000$.

   а) Найдите оценку $\hat{a}$ методом моментов.

   б) Оцените стандартную ошибку $se(\hat{a})$.

   в) Постройте 90%-ый доверительный интервал для неизвестного a.

   г) На уровне значимости $\alpha = 0.05$ проверьте гипотезу H_0: $a = 15$ против $a > 15$. Найдите точное P-значение.

2. По совету Лисы Волк опустил в прорубь хвост и поймал 100 чудо-рыб. Веса рыбин независимы и имеют распределение Вейбулла, $f(x) = 2\exp(-x^2/a^2)\cdot x/a^2$ при $x \geqslant 0$. Известно, что $\sum x_i^2 = 120$.

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

   б) Оцените стандартную ошибку $se(\hat{a})$.

   в) Постройте 90%-ый доверительный интервал для неизвестного a.

   г) На уровне значимости $\alpha = 0.05$ проверьте гипотезу H_0: $a = 1$ против $a > 1$. Найдите точное P-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 приведениями	10	25	20
Дней без приведений	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.

   а) Оцените неизвестные параметры с помощью максимального правдоподобия.

   б) 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 P-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 $\sum x_i^2 = 25$.

   а) Постройте 90% доверительный интервал для σ^2, симметричный по вероятности.

   б) [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 номер[4]. В бою было подбиты 4 танка с номерами 2, 5, 7 и 12.

   а) Найдите оценку общего выпуска танков n с помощью метода максимального правдоподобия.

   б) Является ли оценка максимального правдоподобия несмещенной?

   в) Является ли максимум из номеров подбитых танков достаточной статистикой?

   г) Является ли максимум из номеров подбитых танков полной статистикой?

   д) Постройте с помощью оценки максимального правдоподобия несмещенную эффективную о ценку неизвестного n.

[4] В 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.

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.

   а) Найдите уровень значимости критерия.

   б) Найдите вероятность ошибки второго рода.

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.

   а) Оцените с помощью максимального правдоподобия параметр λ

   б) Оцените, сколько мне еще стоять в очереди.

9. Контрольная работа 4

9.1. 2018-2019

Комментарий: минимум писали 30 минут без официальных шпаргалок, основную часть 90 минут — с листом А4.

Минимум

1. Пусть $X_1, \ldots, X_n$ — случайная выборка из нормального распределения с неизвестным математическим ожиданием μ и известной дисперсией $\sigma^2 = 100$. Объем выборки $n = 25$. Для тестирования основной гипотезы $H_0 : \mu = (-1)$ против альтернативы $H_1 : \mu = 0$ Винни-Пух использует простой критерий. Если $\bar{X} \leqslant (-0.5)$, то Винни-Пух не отвергает гипотезу H_0, в противном случае Винни-Пух отвергает гипотезу H_0 в пользу гипотезы H_1.

   Найдите:

   а) вероятность ошибки 1-го рода;

   б) вероятность ошибки 2-го рода;

   в) мощность критерия.

2. Пусть $X_1, \ldots, X_n$ — случайная выборка из нормального распределения $\mathcal{N}(\mu, \sigma^2)$ с неизвестными параметрами μ и σ^2. Используя реализацию случайной выборки,

$$x_1 = -2, \quad x_2 = -1, \quad x_3 = 0, \quad x_4 = 1, \quad x_5 = 2$$

   постройте 90%-ый доверительный интервал для неизвестного параметра σ^2.

3. Пусть $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_m$ — независимые случайные выборки из нормальных распределений $\mathcal{N}(\mu_X, \sigma_X^2)$ и $\mathcal{N}(\mu_Y, \sigma_Y^2)$ соответственно. Уровень значимости $\alpha = 0.1$.

   Используя реализации случайных выборок

$$x_1 = -2, \quad x_2 = -1, \quad x_3 = 0, \quad x_4 = 1, \quad x_5 = 2,$$
$$y_1 = -2, \quad y_2 = 0, \quad y_3 = 2,$$

   проверьте гипотезу о равенстве дисперсий $\sigma_X^2 = \sigma_Y^2$ против гипотезы о неравенстве дисперсий.

4. Пусть $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_m$ — независимые случайные выборки из нормальных распределений $\mathcal{N}(\mu_X, \sigma^2)$ и $\mathcal{N}(\mu_Y, \sigma^2)$ соответственно.

   Используя реализации случайных выборок

$$x_1 = -2, \quad x_2 = -1, \quad x_3 = 0, \quad x_4 = 1, \quad x_5 = 2,$$
$$y_1 = -2, \quad y_2 = 0, \quad y_3 = 2,$$

   постройте 60%-ый доверительный интервал для разности математических ожиданий $\mu_X - \mu_Y$.

5. Пусть $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_m$ — две независимые случайные выборки из распределения Бернулли с неизвестными параметрами $p_X \in (0, 1)$ и $p_Y \in (0, 1)$. Известно, что $n = 100$, $m = 150$, $\sum_{i=1}^{n} x_i = 60$, $\sum_{i=1}^{n} y_i = 50$. На уровне значимости $\alpha = 0.05$ протестируйте гипотезу $H_0 : p_X = p_Y$ против альтернативной гипотезы $H_1 : p_X > p_Y$.

Основная часть

1. Пусть X — случайная выборка, содержащая одно наблюдение, из распределения с функцией плотности

$$f_X(x) = \begin{cases} \theta x^{\theta-1}, & \text{при } x \in [0;1] \\ 0 & , \text{ при } x \notin [0;1] \end{cases},$$

где $\theta > 0$ — неизвестный параметр распределения. Тестируется основная гипотеза $H_0 : \theta = 1$ против альтернативной гипотезы $H_1 : \theta = 2$.

   а) (7) С помощью леммы Неймана–Пирсона найдите наиболее мощный критерий, имеющий уровень значимости $\alpha = 0.05$.

   б) (3) Вычислите мощность полученного критерия.

2. На раскопках старинного замка Вася нашел древнюю монетку. Монетка ему показалась крайне необычной, и поэтому он решил провести серию из 100 подбрасываний. Результаты Вася записал в таблицу.

	Орел	Решка	Ребро
Количество раз	40	50	10

   а) (6) Постройте двусторонний 95%-ый доверительный интервал для вероятности выпадения ребра.

   б) (4) На уровне значимости 5% проверьте гипотезу о том, что ребро выпадает в одном случае из 100 против альтернативы, что чаще, и вычислите соответствующее P-значение.

   в) (10) С помощью теста отношения правдоподобия на уровне значимости 10% проверьте гипотезу о том, что орёл выпадает так же часто, как решка, и в три раза чаще, чем ребро.

   г) (5) С помощью критерия согласия Пирсона на уровне значимости 10% проверьте гипотезу из предыдущего пункта.

Подсказка: $\ln 2 \approx 0.7$, $\ln 3 \approx 1.1$, $\ln 5 \approx 1.6$, $\ln 7 \approx 1.95$, $\ln 11 \approx 2.4$.

3. Губка Боб и Патрик ловят медуз для коллекции. Число пойманных за i-ый день медуз имеет распределение Пуассона с неизвестным параметром λ. Уловы медуз в различные дни независимы. За прошедшие 100 дней они поймали 300 медуз.

   а) (5) С помощью асимптотических свойств оценок максимального правдоподобия постройте 95%-ый доверительный интервал для неизвестного параметра λ.

   б) (10) С помощью дельта-метода постройте приближенную 95%-ую интервальную оценку вероятности того, что за 101-ый день Губка Боб и Патрик не поймают ни одной медузы.

9.2. 2017-2018

Минимум

1. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из нормального распределения с неизвестным математическим ожиданием μ и неизвестной дисперсией $\sigma^2 = 9$. Объем выборки $n = 20$. Для тестирования основной гипотезы $H_0 : \mu = 0$ против альтернативной гипотезы $H_1 : \mu = 5$ вы используете критерий: если $\overline{X} \leqslant 2$, то вы не отвергаете гипотезу H_0, в противном случае вы отвергаете гипотезу H_0 в пользу гипотезы H_1. Найдите

   а) Вероятность ошибки 1-го рода

   б) Вероятносто ошибки 2-го рода

2. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из нормального распределения с неизвестными параметрами μ и σ^2. Используя реализацию случайной выборки: $x_1 = 2$, $x_3 = 8$, $x_3 = 5$, постройте 95%-ый доверительный интервал для неизвестного параметра σ^2.

3. Вася очень любит тестировать статистические гипотезы. В этот раз Вася собирается проверить утверждение о том, что его друг Пётр звонит Васе исключительно в то время, когда Вася ест. Для этого Вася трудится целый год и проводит серию из 365 испытаний. Результаты приведены в табилце ниже.

	Пётр не звонит	Пётр звонит
Вася ест	100	50
Вася не ест	125	90

На уровне значимости 10% протестируйте гипотезу о том, что Пётр звонит Васе независимо от момента приема пищи.

4. Вася Сидоров утверждает, что ходит в кино в четыре раза чаще, чем в спортзал, а в спортзал в четыре раза чаще, чем в театр. За последние полгода он 105 раз был в театре, 63 раза — в спортазе и 42 раза в кино. На уровне значимости 10% проверьте утверждение Васи.

Квантили χ^2 распределения с 1, 2 и 3 степенями свободы

	0.025	0.05	0.1	0.9	0.95	0.975
1	0.001	0.004	0.016	2.706	3.841	5.024
2	0.051	0.103	0.211	4.605	5.991	7.378
3	0.216	0.352	0.584	6.251	7.815	9.348

Задачи

При решении задач пять–семь используйте данные обследования Росстата за первый квартал 2018 года:

	Число наблюдений	Среднее (тыс. руб.)	Выборочное отклонение (тыс. руб.)
Врачи	40	136	55
Преподаватели	60	139	60

Распределение заработной платы работников любой отрасли хорошо описывается нормальным законом.

5. На уровне значимости 5% проверьте гипотезу о том, что средняя зарплата врача составляет 100 т.р., против альтернативы, что она больше 100 т.р. Вычислите минимальный уровень значимости, при котором основная гипотеза отвергается (P–значение).

6. На уровне значимости 10% проверьте гипотезу о том, что разброс в зарплатах врачей и преподавателей одинаков, против двухсторонней альтернативы.

7. На уровне значимости 10% проверьте гипотезу о том, что средняя зарплата врачей и преподавателей совпадают, против альтернативы, что у преподавателей зарплата выше:

   а) Считая объемы выборок достаточно большими

   б) Считая дисперсии одинаковыми

8. Время в часах безотказной работы микронаушника, величина X, подчиняется экспоненциальному (показательному) закону распределения с неизвестным параметром λ:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, \text{ при } x \geqslant 0 \\ 0, \text{ при } x < 0 \end{cases}$$

   По выборке из 100 независимых наблюдений $\bar{x} = 0.52$. С помощью асимптотических свойств оценок максимального правдоподобия постройте приближенный 95%-ый доверительный интервал:

   а) Для параметра λ

   б) Для вероятности того, что наушник проработает без сбоев весь тест — 45 минут

9. Приглашенный на Петербургский международный экономический форум Германом Грефом индийский мистик Садхгуру подарил Грефу древнюю шестигранную кость для принятия решений в сложных макроэкономических ситуациях. Служба безопасности Сбербанка провела серию из 100 испытаний и составила таблицу:

Грань	1	2	3	4	5	6
Число выпадений	10	10	15	15	25	25

   С помощью теста отношения правдоподобия на уровне значимости 5% проверьте гипотезу о том, что все грани равновероятны.

$$\ln(1/6) = -1.79,\ \ln(0.15) = -1.90,\ \ln(0.25) = -1.39,\ \ln(0.1) = -2.30$$

9.3. 2016-2017

I. Теоретический минимум

В пунктах 1, 3, 11 и 12 предполагается, что $X = (X_1, \ldots, X_n)$ и $Y = (Y_1, \ldots, Y_m)$ — две независимые случайные выборки из нормальных распределений $N(\mu_X, \sigma_X^2)$ и $N(\mu_Y, \sigma_Y^2)$ соответственно.

1. Приведите формулу статистики, при помощи которой можно проверить гипотезу $H_0\colon \sigma_X^2 = \sigma_Y^2$. Укажите распределение этой статистики при верной гипотезе H_0.
2. Приведите формулу информации Фишера о параметре θ, содержащейся в одном наблюдении случайной выборки.
3. Приведите формулу статистики, при помощи которой можно проверить гипотезу $H_0\colon \mu_X - \mu_Y = \Delta_0$ при условии, что дисперсии σ_X^2 и σ_Y^2 неизвестны, но равны между собой. Укажите распределение этой статистики при верной гипотезе H_0.
4. Дайте определение критической области.
5. Приведите формулу плотности нормального распределения $\mathcal{N}(\mu, \sigma^2)$.
6. Приведите формулы границ доверительного интервала с уровнем доверия $(1-\alpha)$, $\alpha \in (0;1)$, для вероятности появления успеха в случайной выборке $X = (X_1, \ldots, X_n)$ из распределения Бернулли с параметром $p \in (0;1)$.
7. Дайте определение несмещенной оценки $\hat{\theta}$ для неизвестного параметра $\theta \in \Theta$.
8. Дайте определение эффективной оценки $\hat{\theta}$ для неизвестного параметра $\theta \in \Theta$.
9. Приведите формулу выборочной дисперсии.
10. Приведите формулу выборочной функции распределения.
11. Приведите формулы границ доверительного интервала с уровнем доверия $(1-\alpha)$, $\alpha \in (0;1)$, для μ_X при условии, что дисперсия σ_X^2 известна.
12. Укажите распределение статистики $\frac{\overline{X}-\mu_X}{\sigma/\sqrt{n}}$.

II. Задачи

1. В ходе анкетирования ста сотрудников банка «Альфа» были получены ответы на вопрос о том, сколько времени они проводят на работе ежедневно. Среднее выборочное оказалось равным 9.5 часам, а выборочное стандартное отклонение 0.5 часа.

   а) На уровне значимости 5 % проверьте гипотезу о том, что сотрудники банка «Альфа» в среднем проводят на работе 10 часов, против альтернативной гипотезы о том, что сотрудники банка «Альфа» в среднем проводят на работе менее 10 часов.

   б) Найдите точное P-значение для наблюдаемой статистики из пункта (а).

   в) Сформулируйте предпосылки, которые были использованы вами для выполнения пункта (а).

   г) На уровне значимости 5 % проверьте гипотезу о $H_0\colon \sigma^2 = 0.3$.

2. Проверка сорока случайно выбранных лекций показала, что студент Халявин присутствовал только на 16 из них. На уровне значимости 5 % проверьте гипотезу о том, что Халявин посещает в среднем половину лекций.
3. В ходе анкетирования двадцати сотрудников банка «Альфа» были получены ответы на вопрос о том, сколько времени они проводят на работе ежедневно. Среднее выборочное оказалось равным 9.5 часам, а выборочное стандартное отклонение 0.5 часа. Аналогичные показатели для 25 сотрудников банка «Бета» составили 9.8 и 0.6 часа соответственно.

а) На уровне значимости 5 % проверьте гипотезу о равенстве математических ожиданий времени, проводимого на работе сотрудниками банков «Альфа» и «Бета».

б) Сформулируйте предпосылки, которые были использованы вами для выполнения пункта (а).

в) На уровне значимости 5 % проверьте гипотезу о равенстве дисперсий времени, проводимого на работе сотрудниками банков «Альфа» и «Бета».

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

Бутерброд	Маслом вниз	Маслом вверх
Число наблюдений	105	95

Можно ли утверждать, что бутерброд падает маслом вниз так же часто, как и маслом вверх? При ответе на вопрос используйте уровень значимости 5 %.

5. Пусть $X = (X_1, \ldots, X_{100})$ — случайная выборка из нормального распределения с математическим ожиданием μ и дисперсией ν. Оба параметра μ и ν неизвестны. Используя следующие данные $\sum_{i=1}^{100} x_i = 30$, $\sum_{i=1}^{100} x_i^2 = 146$ и $\sum_{i=1}^{100} x_i^3 = 122$ с помощью теста отношения правдоподобия проверьте гипотезу $H_0\colon \nu = 1$ на уровне значимости 5 %.

9.4. 2015-2016

1. Сформулируйте определения несмещённости, состоятельности и эффективности оценок.

2. На курсе учится 250 человек. Предположим, что число студентов, не явившихся на экзамен, хорошо описывается законом Пуассона.

   а) Методом максимального правдоподобия найдите оценку параметра распределения Пуассона.

   б) Проверьте выполнение свойств несмещенности, эффективности и состоятельности для данной оценки.

   в) Найдите оценку максимального правдоподобия для вероятности стопроцентной явки студентов на экзамен.

   г) Используя дельта-метод, постройте для этой вероятности асимптотический доверительный интервал.

3. Фармацевтическая компания выпустила новое лекарство от бессонницы, утверждая, что оно помогает 80% людей, страдающих бессонницей. Чтобы проверить утверждение компании, случайным образом выбираются 20 человек, страдающих бессонницей. Обозначим за Y количество человек из выборки, которым лекарство помогло. Основная гипотеза, H_0: $p = 0.8$, альтернативная гипотеза H_a: $p = 0.6$. Критическая область: $\{Y < 12\}$.

   а) В терминах этой задачи сформулируйте, что является ошибкой первого рода. Найдите уровень значимости, соответствующий заданной критической области.

   б) В терминах этой задачи сформулируйте, что является ошибкой второго рода. Найдите вероятность ошибки второго рода.

   в) Найдите такое значение c, что вероятность ошибки первого рода $\alpha \approx 0.1$ при критической области вида $\{Y < c\}$. Найдите соответствующее значение вероятности ошибки второго рода.

   г) Каким должен быть размер выборки, чтобы выборочная доля страдающих бессонницей отличалась от истинной вероятности не более, чем на 0.01 с вероятностью не менее, чем 0.95?

4. Вася Сидоров утверждает, что ходит в кино в два раза чаще, чем на лекции по статистике, на лекции по статистике в два раза чаще, чем в спортзал. За последние полгода он 10 раз был в спортзале, 1 раз — на лекциях по статистике и 39 раз в кино.

   При помощи критерия хи-квадрат Пирсона на уровне значимости 0.05 проверьте, правдоподобно ли Васино утверждение.

5. У Евдокла есть случайная выборка из экспоненциального распределения с неизвестным параметром λ в 50 наблюдений, X_1, X_2, ..., X_{50}. Оказалось, что $\bar{X} = 1.1$. Евдокл хочет проверить гипотезу о равенстве $\lambda = 1$ против альтернативной гипотезы о неравенстве $\lambda \neq 1$ на уровне значимости 0.1.

   Помогите Евдоклу и проверьте гипотезу с помощью критерия отношения правдоподобия.

   Пачка логарифмов: $\ln 50 \approx 3.9$, $\ln 55 \approx 4.0$, $\ln 11 \approx 2.4$, $\ln 60 \approx 4.1$, $\ln 12 \approx 2.5$

6. Американский демографический журнал опубликовал исследование, в котором утверждается, что посетители крупных торговых центров за одно посещение тратят в выходные дни больше, чем в будние. Наибольшие расходы приходятся на воскресенье в период с 4 до 6 часов вечера. Для двух независимых выборок посетителей средние расходы и выборочные стандартные отклонения расходов составили

	Выходные	Рабочие дни
Число наблюдений	21	19
Средние расходы ($)	78	67
Выборочное стандартное отклонение ($)	22	20

   а) Проверьте гипотезу о равенстве дисперсий расходов б) Предполагая, что дисперсии расходов одинаковы, проверьте гипотезу об отсутствии разницы в расходах в выходные и будние дни.

в) Сформулируйте все необходимые для проверки гипотез предыдущих пунктов предпосылки.

7. Винни Пух знает, что пчёлы и мёд бывают правильные и неправильные. По результатам 100 попыток добыть мёд Винни Пух составил таблицу сопряженности признаков.

	Мёд правильный	Мёд неправильный
Пчёлы правильные	12	36
Пчёлы неправильные	32	20

На уровне значимости 0.05 проверьте гипотезу о независимости характеристик пчёл и мёда.

9.5. 2014-2015

1. **Задача для первого потока.**

   Проверка 40 случайно выбранных лекций показала, что студент Халявин присутствовал только на 16 из них.

   а) Найдите 95% доверительный интервал для вероятности увидеть Халявина на лекции.

   б) На уровне значимости 5% проверьте гипотезу о том, что Халявин посещает в среднем половину лекций.

   в) Вычислите минимальный уровень значимости, при котором основная гипотеза отвергается (Р-значение).

1. **Задача 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.

   а) На уровне значимости 10% проверьте гипотезу, что в среднем вес упаковки составляет 55 г.

   б) 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.

2. **Задача 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».

2. **Задача 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.

   а) Можно ли при уровне значимости в 1% утверждать, что преподаватели предъявляют к студентам одинаковый уровень требований?

   б) Вычислите минимальный уровень значимости, при котором основная гипотеза отвергается (Р-значение).

3. Методом максимального правдоподобия найдите оценку параметра θ для выборки X_1, ..., X_n из распределения с функцией плотности

$$f(x) = \begin{cases} \frac{1}{\theta^2} x e^{-\frac{x}{\theta}}, \ x > 0 \\ 0, \ x \leqslant 0 \end{cases}$$

4. Пусть X_1, ..., 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%.

5. **Исследовательская задача.**

   Пусть X_1, ..., 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 $H_0: \ \mu = 0$.

а) Сравните статистики LR, LM и W между собой. Какая — наибольшая, какая — наименьшая?

б) Изменится ли упорядоченность статистик, если проверять гипотезу $H_0: \ \mu = \mu_0$?

Подсказка: $\frac{x}{1+x} \leqslant \ln(1+x) \leqslant x \quad \text{при } x > -1$

6. **Исследовательская задача.**

Величины X_1, ..., X_n независимы и одинаково распределены с функцией плотности

$$f(x) = \begin{cases} a^2 x e^{-ax}, \ x > 0 \\ 0, \ x \leqslant 0 \end{cases}$$

По выборке из 100 наблюдений оказалось, что $\sum x_i = 300$, $\sum x_i^2 = 1000$, $\sum x_i^3 = 3700$.

а) Найдите оценку неизвестного параметра a методом моментов

б) Используя дельта-метод или иначе оцените дисперсию полученной оценки a

в) Постройте 95%-ый доверительный интервал используя оценку метода моментов

9.6. 2009-2010

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) Определите вероятность попасть в «яблочко» при втором выстреле

   б) 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: $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 2$, $\mathbb{E}(X^2) = 2$, $\mathbb{E}(Y^2) = 8$, $\mathbb{E}(XY) = 1$.

   а) Найдите ковариацию и коэффициент корреляции величин X и Y.

   б) Определите, зависимы ли величины X и Y.

   в) 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.

   а) Постройте двусторонний доверительный интервал для математического ожидания величины X с уровнем доверия 0.90.

   б) Постройте двусторонний доверительный интервал для стандартного отклонения величины X с уровнем доверия 0.80.

   в) 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.

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

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

   в) 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.

Синий	76	45	57	65		
Зелёный	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 $x \geqslant 0$. По случайной выборке X_1, X_2, ..., X_n найдите оценку максимального правдоподобия параметра θ.

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

11. При расчете страхового тарифа страховая компания предполагает, что вероятность наступления страхового случая 0.005. По итогам прошедшего года из 10000 случайно выбранных договоров страховых случаев наблюдалось 67.

    а) Согласуются ли полученные данные с предположением страховой компании? Альтернативная гипотеза: вероятность страхового случая больше.

    б) Определить минимальный уровень значимости, при котором основная гипотеза отвергается.

10. Контрольная работа 4. ИП

10.1. 2018-2019

Ровно 189 лет назад, 1 июня 1830 британский учёный Джон Росс открыл северный магнитный полюс :)

1. Пусть y — стандартный нормальный n-мерный вектор. Он случайный, просто Джону Россу лень писать заглавные буквы для векторов :) Вектор a — неслучайный, но тоже гордый n-мерный.

   Пусть H — матрица, проецирующая любой вектор на $(n-1)$-мерное подпространство $a^{\perp}$, являющееся ортогональным дополнением к вектору a. То есть, для любого вектора v вектор Hv перпендикулярен вектору a.

   а) Найдите матрицу H, если $n = 3$ и $a = (1, 2, 2)'$.
   б) Найдите матрицу H для произвольного n и a;
   в) Найдите $\mathbb{E}(y)$ и $\mathrm{Var}(y)$;
   г) Найдите $\mathbb{E}(Hy)$ и $\mathrm{Var}(Hy)$;
   д) Укажите закон распределения $y'Hy$, где y' — это транспонированный вектор y.

2. Рассмотрим формулу, здорово упрощающую подсчёт критерия Пирсона:

$$\sum_{j=1}^{m} \frac{(X_j - np_j)^2}{np_j} + n = \sum_{j=1}^{m} \frac{X_j^2}{np_j}$$

   а) Докажите формулу.
   б) Нарисуйте картинку к этой формуле. На картинке подпишите прямой угол, катеты и гипотенузу. Явно запишите каждый вектор. Объясните, почему треугольник, действительно, прямоугольный.

3. На Земле короля Уильяма Джон Росс нашёл странную монетку. Он подбрасывает её n раз и обнаруживает, что она выпадает на орла, решку и ребро. Джон Росс проверяет гипотезу H_0 о том, что все три вероятности равны.

   Пусть $y = (y_1, y_2, y_3)'$ — количество выпадений орла, решки и ребра. Рассмотрим так же вектор $z = (z_1, z_2, z_3)'$, такой, что $z_i = (y_i - \mathbb{E}(y_i))/\sqrt{\mathbb{E}(y_i)}$. Джон Росс сознательно перепутал ожидание и дисперсию в классической формуле!

   Предположим, что гипотеза H_0 верна.

   а) Укажите закон распределения каждой величины y_i;
   б) Найдите вектор $\mathbb{E}(y)$ и матрицу $\mathrm{Var}(y)$;
   в) Найдите вектор $\mathbb{E}(z)$ и матрицу $\mathrm{Var}(z)$;
   г) Докажите, что матрица $H = \mathrm{Var}(z)$ является проектором на ортогональное дополнение к некоторому вектору a. Явно выпишите вектор a.
   д) Объясните, почему критерий Пирсона имеет хи-квадрат распределение с нужным числом степеней свободы.

4. На Земле короля Уильяма Джон Росс нашёл странную монетку. Он подбрасывает её n раз и обнаруживает, что она выпадает на орла, решку и ребро. Джон Росс проверяет гипотезу о том, что все три вероятности равны с помощью двух статистики: LR, отношения правдоподобия, и CP, критерия Пирсона.

   а) Найдите $\operatorname{plim}_{n\to\infty} \frac{LR}{CP}$;
   б) Обобщите решение на случай произвольного количества равновероятных граней у монетки.

5. Идея доказательства состоятельности ML оценки :)

   Пусть наблюдения y_1, ..., y_n независимы и одинаково распределены с функцией плотности, зависящей от параметра a. Истинное значение параметра обозначим буквой a_0. Оценку максимального правдоподобия обозначим $\hat{a}$.

   Рассмотрим отмасштабированную логарифмическую функцию правдоподобия $\ell_n(a) = \ell(a)/n$, и ожидаемую логарифмическую функцию правдоподобия[5], $\tilde{\ell}(a) = \mathbb{E}(\ell(a))$.

   а) Что больше, $\ln x$ или $x - 1$? Докажите!
   б) В какой точке находится максимум функции $\ell_n(a)$?
   в) В какой точке находится максимум функции $\tilde{\ell}(a)$?
   Подсказка: рассмотрите выражение $\tilde{\ell}(a) - \tilde{\ell}(a_0)$ и примените доказанное неравество :)
   г) К чему сходится $\ell_n(a)$ по вероятности?

10.2. 2017-2018

Напутствие в добрый путь:

1. Работа сдаётся только в виде запроса pull-request на гитхаб-репозиторий.
2. Имя файла должно быть вида ivanov_ivan_161_kr_4.Rmd.
3. Также фамилию и имя нужно указать в шапке документа в поле author :)
4. Если нужно, то установите пакеты tidyverse, maxLik, nycflights13.

1. Симулируем бурную деятельность! В качестве параметра k в задаче используй число букв в своей фамилии в именительном падеже :)

   Каждый день Василий съедает случайное количество булочек, которое распределено по Пуассону с параметром 10. Логарифм затрат в рублях на каждую булочку распределён нормально $N(2, 1)$. Андрей каждый день съедает биномиальное количество булочек $Bin(2k, 0.5)$. Затраты Андрей на каждую булочку распределены равномерно на отрезке $[2; 20]$.

   а) Сколько в среднем тратит Василий на булочки за день?
   б) Чему равна дисперсия дневных расходов Василия?
   в) Какова вероятность того, что за один день Василий потратит больше денег, чем Андрей?
   г) Какова условная вероятность того, что Василий за день съел больше булочек, чем Андрей, если известно, что Василий потратил больше денег?

2. Сражаемся с реальностью! В пакете nycflights13 встроен набор данных weather о погоде в разные дни в разных аэропортах.

   а) Постройте гистограмму переменной влажность, humid. У графика подпишите оси!

[5]Внимание: ожидание считается с помощью истинного a_0 от функции, в которую входит константа a.

б) Постройте диаграмму рассеяния переменных влажность и количество осадков, precip. У графика подпишите оси!

Посчитайте выборочное среднее и выборочную дисперсию влажности и количества осадков.

в) С помощью максимального правдоподобия оцените параметр μ, предположив, что наблюдения за влажностью имеют нормальное $N(\mu, 370)$-распределение и независимы. Постройте 95%-ый доверительный интервал для μ.

г) С помощью максимального правдоподобия оцените параметр σ^2, предположив, что наблюдения за влажностью имеют нормальное $N\left(60, \sigma^2\right)$-распределение и независимы. Постройте 95%-ый доверительный интервал для σ^2.

Если при численной оптимизации параметр σ^2 становится отрицательным, можно задать параметры по-другому, например, $\sigma^2 = \exp(\gamma)$.

11. Финальные экзамены

11.1. 2018-2019

Вопрос 1. Пусть $X_1, \ldots, X_n$ --- случайная выборка из распределения Пуассона с параметром $\lambda > 0$. Известно, что оценка максимального правдоподобия параметра λ равна $\bar{X}$. Чему равна оценка максимального правдоподобия для $1/\lambda$?

$\boxed{A}$ $\ln \bar{X}$

$\boxed{B}$ $\bar{X}/n$

$\boxed{C}$ $1/\bar{X}$

$\boxed{D}$ $\bar{X}$

$\boxed{E}$ $e^{\bar{X}}$

Вопрос 2. Время подготовки студента к экзаменам и по статистике, и макроэкономике, имеет нормальное распределение с неизвестными математическими ожиданиями и дисперсиями. По 10 наблюдениям Вениамин получил оценку стандартного отклонения времени подготовки к статистике равную 5 часам. Оценка стандартного отклонения времени подготовки к макроэкономике, рассчитанная по 20 наблюдениям, оказалась равной 2. Тестовая статистика при проверке гипотезы о равенстве дисперсий может быть равна

$\boxed{A}$ 0.16

$\boxed{B}$ 0.8

$\boxed{C}$ 0.4

$\boxed{D}$ 2.5

$\boxed{E}$ 12.5

Вопрос 3. Случайная выборка состоит из одного наблюдения X_1, которое имеет плотность распределения

$$f(x;\, \theta) = \begin{cases} \frac{1}{\theta^2} x e^{-x/\theta} & \text{при } x > 0, \\ 0 & \text{при } x \leqslant 0, \end{cases}$$

где $\theta > 0$. Чему равна оценка неизвестного параметра θ, найденная с помощью метода максимального правдоподобия?

$\boxed{A}$ X_1

$\boxed{B}$ $\ln X_1$

$\boxed{C}$ $X_1/2$

$\boxed{D}$ $\frac{X_1}{\ln X_1}$

$\boxed{E}$ $1/\ln X_1$

Вопрос 4. Последовательность оценок $\hat{\theta}_n$ называется состоятельной для параметра θ, если

$\boxed{A}$ $\hat{\theta}_n \xrightarrow{P} \theta$ при $n \to \infty$

$\boxed{B}$ $\operatorname{Var}(\hat{\theta}_n) = (\theta)^2/n$

$\boxed{C}$ $\mathbb{E}(\hat{\theta}_n) = \theta$

$\boxed{D}$ $\mathbb{E}((\hat{\theta}_n - \theta)^2) \to 0$ при $n \to \infty$

$\boxed{E}$ $\mathbb{E}((\hat{\theta}_n - \theta)^2) \leqslant \mathbb{E}((\tilde{\theta} - \theta)^2)$ для всех $\tilde{\theta} \in K$

Вопрос 5. Величина X принимает три значения 1, 2 и 3. По случайной выборке из ста наблюдений оказалось, что 1 выпало 40 раз, 2 --- 40 раз и 3 --- 20 раз. Карл хочет проверить гипотезу о том, что все три вероятности одинаковые. Значение критерия согласия Пирсона равно

$\boxed{A}$ 4 $\boxed{C}$ 6 $\boxed{E}$ 8

$\boxed{B}$ 7 $\boxed{D}$ 5

Вопрос 6. Пусть $X_1, \ldots, X_n$ --- случайная выборка из распределения Бернулли с параметром $p \in (0; 1)$. Чему равна информация Фишера о параметре p, заключенная в двух наблюдениях случайной выборки?

$\boxed{A}$ $\frac{2}{p(1-p)}$ $\boxed{C}$ $2p(1-p)$ $\boxed{E}$ $2(1-p)$

$\boxed{B}$ $2p$ $\boxed{D}$ $\frac{2}{p}$

Вопрос 7. Температура планеты Плюк и её спутника являются стандартными нормальными случайными величинами, имеющими совместное нормальное распределение. Ковариация между температурами равна 0.5. Найдите вероятность того, что на Плюке положительная температура, если на спутнике температура равна -1.

$\boxed{A}$ 0.739 $\boxed{C}$ 0.282 $\boxed{E}$ 0.718

$\boxed{B}$ 0.596 $\boxed{D}$ 0.114

Вопрос 8. Математическое ожидание оценки дисперсии $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ для выборки из распределения Пуассона с $\lambda = 3$, равняется

$\boxed{A}$ 1 $\boxed{C}$ 3 $\boxed{E}$ 9

$\boxed{B}$ $9/n$ $\boxed{D}$ $3/n$

Вопрос 9. Пусть $X_1, \ldots, X_n$ --- случайная выборка из распределения с плотностью распределения

$$f(x;\,\theta) = \begin{cases} \frac{2x}{\theta^2} & \text{при } x \in [0;\,\theta], \\ 0 & \text{при } x \notin [0;\,\theta], \end{cases}$$

где $\theta > 0$. Используя начальный момент 2-го порядка, при помощи метода моментов найдите оценку неизвестного параметра θ.

$\boxed{A}$ $\sqrt{\sum_{i=1}^{n} X_i^2}$ $\boxed{B}$ $\sqrt{\frac{2}{n}\sum_{i=1}^{n} X_i^2}$ $\boxed{D}$ $\sqrt{\frac{n}{2}\sum_{i=1}^{n} X_i^2}$

$\boxed{C}$ $\frac{3}{2}\bar{X}$ $\boxed{E}$ $\frac{2}{3}\bar{X}$

Вопрос 10. Даны выборки объёма n из равномерного на отрезке $[0, 1]$ распределения. Выборочный начальный момент второго порядка стремится по вероятности при $n \to \infty$ к

$\boxed{A}$ 1/3 $\boxed{C}$ 1/4 $\boxed{E}$ 1/12

$\boxed{B}$ 1 $\boxed{D}$ 1/2

Вопрос 11. Компоненты вектора $X = (X_1, X_2, X_3)$ имеют совместное нормальное распределение:

$$X \sim \mathcal{N}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 & 0.5 & 0.5 \\ 0.5 & 2 & 0.5 \\ 0.5 & 0.5 & 3 \end{bmatrix}\right).$$

Вероятность $\mathbb{P}(X_1 > X_2 + X_3)$ равна

$\boxed{A}$ 0.688 $\boxed{C}$ 0.215 $\boxed{E}$ 0.593

$\boxed{B}$ 0.369 $\boxed{D}$ 0.701

Вопрос 12. Случайные величины X и Y распределены нормально с неизвестным математическим ожиданием и неизвестной дисперсией. Для тестирования гипотезы о равенстве дисперсий выбирается 20 наблюдений случайной величины X и 30 наблюдений случайной величины Y. Какое распределение может иметь статистика, используемая в данном случае?

A $F_{20,30}$
B t_{48}
C χ^2_{48}
D $F_{29,19}$
E χ^2_{49}

Вопрос 13. Если функция правдоподобия пропорциональна $a^2(1-a)^6$, априорная плотность пропорциональна $\exp(-a)$, то апостериорная плотность параметра a пропорциональна

A $0.5a^2(1-a)^6 + 0.5\exp(-a)$
B $\frac{\exp(-a)}{a^2(1-a)^6}$
C $\frac{a^2(1-a)^6}{\exp(-a)}$
D $0.5a^2(1-a)^6 + 0.5\exp(a)$
E $\frac{a^2(1-a)^6}{\exp(a)}$

Вопрос 14. Величина X принимает три значения 1, 2 и 3. По случайной выборке из ста наблюдений оказалось, что 1 выпало 40 раз, 2 --- 40 раз и 3 --- 20 раз. Андрей Николаевич хочет проверить гипотезу о том, что все три вероятности одинаковые. Значение критерия согласия Колмогорова равно

A $2/5$
B $2/15$
C $3/4$
D $1/4$
E $3/5$

Вопрос 15. Пусть $X_1, \ldots, X_n$ --- случайная выборка и $\ell(\theta)$ --- её логарифмическая функция правдоподобия. Тестируется гипотеза $H_0 : \theta = 1$. Известно, что $\max_\theta \ell(\theta) = -10$, а $\ell(1) = -20$. Чему равно значение статистики отношения правдоподобия?

A 20
B 0
C 10
D -20
E -10

Вопрос 16. Максимальная ширина 90%-го симметричного по вероятности доверительного интервала для доли, построенного по выборке из 64 наблюдений, приблизительно равна

A 0.102
B 0.156
C 0.368
D 0.206
E 0.234

Вопрос 17. Требуется проверить гипотезу о равенстве математических ожиданий по двум нормальным независимым выборкам размером 33 и 16 наблюдений. Истинные дисперсии по обеим выборкам известны, совпадают и равны 196. Разница выборочных средних равна 1. Тестовая статистика может быть примерно равна

A -0.52
B 0.11
C 1.25
D -0.23
E 0.74

Вопрос 18. При проверке гипотезы о равенстве математических ожиданий оценок по статистике в двух группах против альтернативной гипотезы, что в первой группе оценки выше, оказалось, что выборочные средние равны. Тогда P-значение в этом тесте

A равно 0.25
B равно 0.5
C равно 0
D равно 1
E Недостаточно данных для ответа

Вопрос 19. Величина X принимает три значения 1, 2 и 3. По случайной выборке из ста наблюдений оказалось, что 1 выпало 40 раз, 2 --- 40 раз и 3 --- 20 раз. Карл хочет проверить гипотезу о том, что все три вероятности одинаковые. При верной H_0 критерий Пирсона имеет распределение

$\boxed{A}$ χ^2_1

$\boxed{B}$ χ^2_3

$\boxed{C}$ $\mathcal{N}(0;1)$

$\boxed{D}$ χ^2_{99}

$\boxed{E}$ χ^2_2

Вопрос 20. Вася считает, что контрольные по макроэкономике и статистике нравятся студентам с одинаковой вероятностью. Чтобы проверить эту гипотезу, он опросил по 100 случайных однокурсников после каждой контрольной и выяснил, что макроэкономика понравилась 30 студентам, а статистика --- 50. При проверке этой гипотезы, тестовая статистика может иметь распределение

$\boxed{A}$ t_{100}

$\boxed{B}$ $\mathcal{N}(0,1)$

$\boxed{C}$ t_{99}

$\boxed{D}$ t_{98}

$\boxed{E}$ t_{198}

Вопрос 21. Рассмотрим алгоритм Метрополиса-Гастингса для получения выборки параметра с апостериорной плотностью пропорциональной t^2. Предлагаемый переход из a в b задаётся правилом, $b = a + Z$, где $Z \sim \mathcal{N}(0;4)$. Вероятность одобрения перехода из точки 0.5 в точку 0.3 равна

$\boxed{A}$ 0.6

$\boxed{B}$ 1

$\boxed{C}$ 0.5

$\boxed{D}$ 0.64

$\boxed{E}$ 0.36

Вопрос 22. Оценка $\hat{\theta}_n$ называется эффективной оценкой параметра θ в классе оценок K, если

$\boxed{A}$ $\operatorname{Var}(\hat{\theta}_n) = (\theta)^2/n$

$\boxed{B}$ $\mathbb{E}((\hat{\theta}_n - \theta)^2) \to 0$ при $n \to \infty$

$\boxed{C}$ $\mathbb{E}((\hat{\theta}_n - \theta)^2) \leqslant \mathbb{E}((\tilde{\theta} - \theta)^2)$ для всех $\tilde{\theta} \in K$

$\boxed{D}$ $\mathbb{E}(\hat{\theta}_n) = \theta$

$\boxed{E}$ $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ при $n \to \infty$

Вопрос 23. Дана реализация выборки: -1, 1, 0, 2. Эмпирическая (выборочная) функция распределения в точке $x = 0.5$ принимает значение равное

$\boxed{A}$ 0.25

$\boxed{B}$ 0

$\boxed{C}$ 1

$\boxed{D}$ 0.8

$\boxed{E}$ 0.5

Вопрос 24. Длины катетов в сантиметрах прямоугольного треугольника являются модулями независимых стандартных нормальных случайных величин. Какую длину не превысит гипотенуза этого треугольника с вероятностью 0.95?

$\boxed{A}$ 4.61

$\boxed{B}$ 5.99

$\boxed{C}$ 0.21

$\boxed{D}$ 0.1

$\boxed{E}$ 0.68

Вопрос 25. При построения доверительного интервала для разности математических ожиданий в двух нормальных независимых выборках размером m и n в случае равных известных дисперсий используется распределение

$\boxed{A}$ $F_{m-1,n-1}$

$\boxed{B}$ $F_{m,n}$

$\boxed{C}$ $\mathcal{N}(0,1)$

$\boxed{D}$ t_{m+n}

$\boxed{E}$ t_{m+n-2}

Вопрос 26. Вася считает, что контрольные по макроэкономике и статистике нравятся студентам с одинаковой вероятностью. Чтобы проверить эту гипотезу, он опросил по 100 случайных однокурсников после каждой контрольной и выяснил, что макроэкономика понравилась 30 студентам, а статистика --- 50. При расчётах Вася получил P-значение равное 0.0038. Это означает, что гипотеза

$\boxed{A}$ отвергается на уровне значимости 5%, но не отвергается на 1%

$\boxed{B}$ отвергается на уровне значимости 1%, но не отвергается на 5%

$\boxed{C}$ не отвергается на любом возможном уровне значимости

$\boxed{D}$ отвергается на любом возможном уровне значимости

$\boxed{E}$ отвергается на уровне значимости 1%

Вопрос 27. Для выборки 1, 2, 3, 4, 5 границы 95%-го доверительного интервала для математического ожидания равны

$\boxed{A}$ $[1.04, 4.96]$

$\boxed{B}$ $[3.08, 5.92]$

$\boxed{C}$ $[-4.02, 1, 02]$

$\boxed{D}$ $[1.54, 5.46]$

$\boxed{E}$ $[0.86, 5.14]$

Вопрос 28. Пусть t_n --- случайная величина, распределенная по Стьюденту с n степенями свободы. Предел $\lim\limits_{n\to\infty} \mathbb{P}\left(t_n^2 > 1\right)$ равен

$\boxed{A}$ 0.102

$\boxed{B}$ 0.841

$\boxed{C}$ 0.788

$\boxed{D}$ 0.253

$\boxed{E}$ 0.317

Вопрос 29. Истинное значение параметра θ равно 2, в случайной выборке 100 наблюдений, а информация Фишера о параметре θ, заключенная в одном наблюдении равна $I_1(\theta) = 9$. Распределение оценки максимального правдоподобия $\hat{\theta}$ похоже на

$\boxed{A}$ $\mathcal{N}(2,\ 1/900)$

$\boxed{B}$ $\mathcal{N}(2,\ 9)$

$\boxed{C}$ $\mathcal{N}(2,\ 1/3)$

$\boxed{D}$ $\mathcal{N}(2,\ 1/9)$

$\boxed{E}$ $\mathcal{N}(2,\ 1/30)$

Вопрос 30. При проверке гипотезы о равенстве дисперсии 5 по 11 наблюдениям за нормально распределенной случайной величиной, оказалось, что тестовая статистика равна 2. Несмещённая оценка дисперсии была равна

$\boxed{A}$ 5/11

$\boxed{B}$ 10

$\boxed{C}$ 5

$\boxed{D}$ 1

$\boxed{E}$ 5/10

11.2. 2017-2018

Вопрос 1. Дана случайная выборка из двух наблюдений, X_1 и X_2. Несмещённой и наиболее эффективной оценкой математического ожидания из предложенных является

$\boxed{A}$ $\frac{20}{20}X_1 + \frac{20}{20}X_2$

$\boxed{B}$ $\frac{8}{20}X_1 + \frac{12}{20}X_2$

$\boxed{C}$ $\frac{10}{20}X_1 + \frac{10}{20}X_2$

$\boxed{D}$ $\frac{1}{20}X_1 + \frac{19}{20}X_2$

$\boxed{E}$ $\frac{5}{20}X_1 + \frac{15}{20}X_2$

Вопрос 2. Величины X_1 и X_2 независимы и одинаково распределены. Оценка $\hat{\mu} = 3aX_1 + 4a^2X_2$ математического ожидания $\mu = \mathbb{E}(X_i)$ будет несмещённой при a равном

$\boxed{A}$ 1.2

$\boxed{B}$ 3

$\boxed{C}$ -1

$\boxed{D}$ 0

$\boxed{E}$ -3

Вопрос 3. Величины X_1, ..., X_5 представляют собой случайную выборку. Несмещённой оценкой дисперсии X_i является

$\boxed{A}$ $0.25\sum_{i=1}^{4}(X_i-\bar{X})^2$

$\boxed{B}$ $0.5\sum_{i=1}^{5}(X_i-\bar{X})^2$

$\boxed{C}$ $0.2\sum_{i=1}^{4}(X_i-\bar{X})^2$

$\boxed{D}$ $0.25\sum_{i=1}^{5}(X_i-\bar{X})^2$

$\boxed{E}$ $0.2\sum_{i=1}^{5}(X_i-\bar{X})^2$

Вопрос 4. Последовательность оценок $\hat{a}_n$ параметра a является состоятельной, если

$\boxed{A}$ $\mathbb{E}((\mathbb{E}(\hat{a}_n)-a)^2)\xrightarrow{P}\hat{a}_n$

$\boxed{B}$ $\mathbb{E}((\hat{a}_n-a)^2)\xrightarrow{P}\hat{a}_n$

$\boxed{C}$ $\hat{a}_n\xrightarrow{P}a$

$\boxed{D}$ $\mathbb{E}((\hat{a}_n-a)^2)\xrightarrow{P}0$

$\boxed{E}$ $\mathbb{E}((\hat{a}_n-a)^2)\xrightarrow{P}a$

Вопрос 5. Оценка $\hat{a}$ называется эффективной оценкой параметра a в классе оценок K, если

$\boxed{A}$ $\mathbb{E}(\hat{a}^2)\geqslant\mathbb{E}(\tilde{a}^2)$ для всех $\tilde{a}\in K$

$\boxed{B}$ $\mathbb{E}((\hat{a}-a)^2)\leqslant\mathbb{E}((\tilde{a}-a)^2)$ для всех $\tilde{a}\in K$

$\boxed{C}$ $\mathbb{E}((\hat{a}-\tilde{a})^2)\leqslant\mathbb{E}((\tilde{a}-a)^2)$ для всех $\tilde{a}\in K$

$\boxed{D}$ $\mathbb{E}((\hat{a}-a)^2)\geqslant\mathbb{E}((\tilde{a}-a)^2)$ для всех $\tilde{a}\in K$

$\boxed{E}$ $\mathbb{E}((\hat{a}-\tilde{a})^2)\geqslant\mathbb{E}((\tilde{a}-a)^2)$ для всех $\tilde{a}\in K$

Вопрос 6. Апостериорная функция плотности пропорциональна

$\boxed{A}$ Отношению функции правдоподобия к априорной плотности

$\boxed{B}$ Отношению априорной плотности к функции правдоподобия

$\boxed{C}$ Разности априорной плотности и правдоподобия

$\boxed{D}$ Произведению априорной плотности и правдоподобия

$\boxed{E}$ Сумме априорной плотности и правдоподобия

Вопрос 7. Алгоритм Метрополиса-Гастингса порождает

$\boxed{A}$ Независимую выборку из апостериорного закона распределения

$\boxed{B}$ Независимую выборку из смеси априорного и апостериорного законов

$\boxed{C}$ Независимую выборку из априорного закона распределения

$\boxed{D}$ Зависимую выборку из априорного закона распределения

$\boxed{E}$ Зависимую выборку из апостериорного закона распределения

Вопрос 8. В алгоритме Метрополиса-Гастингса был предложен переход из точки $\theta^{(0)}=4$ в точку $\theta^{(1)}_{prop}=5$. Априорное распределение θ равномерное. Известны значения функций правдоподобия, $f(data|\theta=4)=0.7$, $f(data|\theta=5)=0.8$. Вероятность одобрения перехода равна

$\boxed{A}$ $0.8/5$

$\boxed{B}$ 1

$\boxed{C}$ $28/40$

$\boxed{D}$ $4/5$

$\boxed{E}$ $7/8$

Вопрос 9. Есть выборка $X_1, X_2, \ldots, X_5$ и выборка Y_1, Y_2, Y_3, Y_4. Исследовательница Ирина проводит тест суммы рангов Вилкоксона. У выборки X_i сумма рангов равна 7. Сумма рангов для выборки Y_j равна

$\boxed{A}$ 2

$\boxed{B}$ 1

$\boxed{C}$ 38

$\boxed{D}$ 43

$\boxed{E}$ 45

Вопрос 10. Вероятность того, что в случайной выборке три наблюдения подряд попадут в верхний теоретический квартиль равна

$\boxed{A}$ 0.01

$\boxed{B}$ 0.05

$\boxed{C}$ $1/2$

$\boxed{D}$ $1/4$

$\boxed{E}$ $1/64$

Вопрос 11. Величины $X_1,\ldots,X_n$ — случайная выборка с распределением

x	-4	0	3
$\mathbb{P}(X_i=x)$	$3/4-\theta$	$1/4$	θ

Оценка неизвестного параметра θ, найденная с помощью первого начального момента, равна

[A] $\frac{\bar{X}+7}{3}$

[B] $\frac{\frac{1}{n}\sum_{i=1}^{n} X_i^2-12}{7}$

[C] $\frac{12-\frac{1}{n}\sum_{i=1}^{n} X_i^2}{7}$

[D] $\frac{7-\frac{1}{n}\sum_{i=1}^{n} X_i^2}{12}$

[E] $\frac{\bar{X}+3}{7}$

Вопрос 12. Случайная выборка состоит из одного наблюдения X_1, которое имеет плотность распределения

$$f(x;\,\theta) = \begin{cases} \frac{1}{\theta}x^{-1+\frac{1}{\theta}} & \text{при } x \in (0;\,1), \\ 0 & \text{при } x \notin (0;\,1). \end{cases}$$

Оценка параметра θ, найденная с помощью метода максимального правдоподобия, равна

[A] X_1

[B] $\ln X_1$

[C] $-\ln X_1$

[D] $-X_1$

[E] $\frac{1}{\ln X_1}$

Вопрос 13. Величины $X_1,\,\ldots,\,X_n$ — случайная выборка из распределения Бернулли с параметром $p \in (0;\,1)$. Оценка максимального правдоподобия параметра p равна $\bar{X}$. Оценка максимального правдоподобия для $\sqrt{p}$ равна

[A] $\frac{\sqrt{\sum_{i=1}^{n} X_i}}{n}$

[B] $\sqrt{\sum_{i=1}^{n} X_i}$

[C] $\frac{1}{n}\sum_{i=1}^{n} X_i$

[D] $\frac{1}{n}\sum_{i=1}^{n} \sqrt{X_i}$

[E] $\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i}$

Вопрос 14. Величины $X_1,\,\ldots,\,X_n$ — случайная выборка из распределения Бернулли с параметром $p \in (0;\,1)$. Информация Фишера о параметре p, заключенная в одном наблюдении, равна

[A] $\frac{1}{p(1-p)}$

[B] $\frac{1}{p}$

[C] $p(1-p)$

[D] $1-p$

[E] p

Вопрос 15. Известно истинное значение параметра, $\theta = 1$, и информация Фишера о параметре θ, заключенная в одном наблюдении случайной выборки, $I_1(\theta) = 8$. Оценка максимального правдоподобия $\hat{\theta}$ параметра θ, найденная по ста наблюдениям случайной выборки, имеет распределение, похожее на

[A] $\mathcal{N}(1,\,1/800)$

[B] $\mathcal{N}(1,\,1/\sqrt{800})$

[C] $\mathcal{N}(1,\,1/\sqrt{8})$

[D] $\mathcal{N}(1,\,1/8)$

[E] $\mathcal{N}(1,\,8)$

Вопрос 16. Дана реализация выборки: 1, 2, 0. Выборочный начальный момент второго порядка равен

[A] 2.5

[B] 5/3

[C] 1/3

[D] 1

[E] 3

Вопрос 17. Математическое ожидание выборочного среднего, построенного по выборке из равномерного распределения на отрезке $[0, 2]$, равно

[A] $1/\sqrt{n}$

[B] 1

[C] 0

[D] 2

[E] 1.5

Вопрос 18. Дана реализация выборки: 3, 2, 5, 4, 2. Выборочная функция распределения в точке $x = 2.5$ принимает значение

[A] 0.25

[B] 0.4

[C] 0.5

[D] 0.6

[E] 0.2

Вопрос 19. Экзамен принимают два преподавателя: Злой и Добрый. Злой поставил оценки 2, 3, 10, 8, 1. А Добрый — оценки 6, 4, 7, 9. Значение статистики критерия Вилкоксона о совпадении распределений оценок может быть равно

$\boxed{A}$ 25 $\boxed{C}$ 26 $\boxed{E}$ 24

$\boxed{B}$ 23 $\boxed{D}$ 22

Вопрос 20. Датчик случайных чисел выдал два значения псевдослучайных чисел 0.1 и 0.8. Вычислите значение критерия Колмогорова и проверьте гипотезу о соответствии распределения равномерному на $(0, 1)$. Критическое значение статистики Колмогорова считайте равным 0.776.

$\boxed{A}$ 0.1, H_0 отвергается $\boxed{C}$ 0.8, H_0 отвергается $\boxed{E}$ 0.4, H_0 не отвергается

$\boxed{B}$ 0.3, H_0 не отвергается $\boxed{D}$ 0.2, H_0 не отвергается

Вопрос 21. Случайные величины $X_1, \ldots, X_m$ — случайная выборка из нормального распределения. Величины $Y_1, \ldots, Y_n$ — независимая случайная выборка из нормального распределения. Для построения доверительного интервала для отношения дисперсий можно использовать статистику с распределением

$\boxed{A}$ $F_{m,n-2}$ $\boxed{C}$ χ^2_{m+n-2} $\boxed{E}$ t_{m+n-2}

$\boxed{B}$ $F_{m-1,n-1}$ $\boxed{D}$ $F_{m+1,n+1}$

Вопрос 22. При построения доверительного интервала для разности математических ожиданий в двух нормальных выборках размером m и n в случае равных неизвестных дисперсий используется распределение

$\boxed{A}$ $\mathcal{N}(0, m+n-2)$ $\boxed{C}$ t_{m+n} $\boxed{E}$ t_{m+n-2}

$\boxed{B}$ $F_{m,n}$ $\boxed{D}$ $F_{m-1,n-1}$

Вопрос 23. Для построения доверительного интервала для математического ожидания используется выборка из 100 наблюдений. Выборочное среднее равно 5. Дисперсия генеральной совокупности известна и равна 25. Минимальная длина 95%-доверительного интервала примерно равна

$\boxed{A}$ 5 $\boxed{C}$ 0.98 $\boxed{E}$ 10

$\boxed{B}$ 2.5 $\boxed{D}$ 1.96

Вопрос 24. При построении 90%-доверительного интервала для вероятности используется выборка из 25 наблюдений. Выборочная доля составляет 0.6. В симметричный доверительный интервал попадают значения

$\boxed{A}$ 0.6, 0.7, 0.85 $\boxed{C}$ 0.35, 0.5, 0.65 $\boxed{E}$ 0.7, 0.8, 0.9

$\boxed{B}$ 0.5, 0.6, 0.65 $\boxed{D}$ 0.8, 0.9, 1.0

Вопрос 25. При построении 90%-доверительного интервала для дисперсии используется выборка из 26 наблюдений. Несмещенная оценка дисперсии равна 100. Левая граница симметричного по вероятности доверительного интервала равна

$\boxed{A}$ 43.25 $\boxed{C}$ 8.16 $\boxed{E}$ 66.4

$\boxed{B}$ 106.32 $\boxed{D}$ 32.8

Вопрос 26. При проверке гипотезы о равенстве математических ожиданий оценок по математической статистике в двух группах, было получено P-значение 0.03. Тогда нулевая гипотеза

$\boxed{A}$ отвергается на уровне значимости 0.05 и на уровне значимости 0.01

$\boxed{B}$ не отвергается на уровне значимости 0.05 и отвергается на уровне значимости 0.01

$\boxed{C}$ не отвергается ни на уровне значимости 0.05, ни на уровне значимости 0.01

$\boxed{D}$ отвергается на уровне значимости 0.05 и не отвергается на уровне значимости 0.01

$\boxed{E}$ ответ зависит от альтернативной гипотезы

Вопрос 27. Имеются две случайных выборки $X_1, \ldots, X_{31}$ и $Y_1, \ldots, Y_{41}$ из нормальных распределений. Известно, что $\sum_{i=1}^{31}(X_i - \bar{X})^2 = 120$ и $\sum_{i=1}^{41}(Y_i - \bar{Y})^2 = 400$. При проверке гипотезы о равенстве дисперсий этих распределений значение тестовой статистики может быть равно

$\boxed{A}$ 0.3

$\boxed{B}$ 2.52

$\boxed{C}$ 2

$\boxed{D}$ 3.33

$\boxed{E}$ 2.5

Вопрос 28. Имеется выборка из одного наблюдения X_1. На основе этой выборки тестируется гипотеза H_0: $X_1 \sim U[0;2]$ против альтернативной гипотезы $X_1 \sim U[1,3]$. Используется критерий следующего вида: если $X_1 > a$, то H_0 отвергается. Минимальная вероятность ошибки первого рода достигается при a равном

$\boxed{A}$ 0.5

$\boxed{B}$ 1.5

$\boxed{C}$ 2

$\boxed{D}$ 1.9

$\boxed{E}$ 1

Вопрос 29. Исследовательница Алевтина подбросила кубик 12 раз и 12 раз на нём выпала шестёрка. Алевтина хочет проверить, выпадают ли все грани равновероятно, при помощи критерия χ^2 Пирсона. Значение тестовой статистики будет равно

$\boxed{A}$ 60

$\boxed{B}$ 12

$\boxed{C}$ 5

$\boxed{D}$ 50

$\boxed{E}$ 6

Вопрос 30. Исследовательница Глафира считает, что любовь к энергетическим напиткам и успешность сдачи экзамена по математической статистике должны быть как-то связаны. Опросив 200 своих однокурсников, она получила следующие результаты:

	пьёт энергетик	не пьёт энергетик
Успешно сдал	20	120
Завалил	40	20

Статистика χ^2 Пирсона для проверки независимости признаков с округлением до целых равна

$\boxed{A}$ 70

$\boxed{B}$ 55

$\boxed{C}$ 65

$\boxed{D}$ 35

$\boxed{E}$ 45

12. Ответы к минимумам

12.1. Контрольная работа 1 — Задачный минимум

1. а) 0.25
   б) 0.6
   в) зависимы
2. а) 0.5
   б) 0.75
   в) независимы
3. $\frac{4}{10\cdot 11\cdot 12\cdot 13}$
4. 0.5
5. 7/15

6. 0.028
7. $\frac{5}{7}$
8. а) 0.5
   б) 0.75
   в) 0
   г) 0.5
   д)
   е) функция плотности не существует
9. а) 0.5
   б) 0
   в) 0.5
   г) 0.5
   д) 0.5
10. а) 0.25
    б) 0.75
    в) 0
    г) 0.25
    д)
    е) функция плотности не существует
11. а) 0.25
    б) 0.25
    в) 0.75
    г) $11/16$
    д) 0.75
12. а) $\left(\frac{1}{4}\right)^4$
    б) $1-\left(\frac{1}{4}\right)^4$
    в) 0
    г) 3
    д) 0.75
    е) $2, 3$
13. а) $\left(\frac{3}{5}\right)^5$
    б) $1-\left(\frac{3}{5}\right)^5$
    в) 0
    г) 2
    д) 1.2
    е) 2
14. а) e^{-100}
    б) $1-e^{-100}$
    в) 0
    г) 100
    д) $99, 100$
15. а) e^{-101}
    б) $1-e^{-101}$
    в) 0
    г) 101
    д) $100, 101$
16. $1-\frac{8^5}{9^5}$
17. $\frac{8^5}{9^5}$
18. $1-e^{-3}$
19. e^{-6}
20. а) 0.5
    б) 0.5
    в) 0.25
    г) 0
21. а) 0.5
    б) 0.5
    в) $\frac{1}{3}$
    г) $\frac{1}{12}$
    д) 1
22. а) 2
    б) 0.25
    в) $\frac{3}{4}$
    г) 1
23. а) 2
    б) 0.5
    в) 0.5
    г) 0
    д) 0.8

12.2. Контрольная работа 2 — Задачный минимум

1. а) 0.5
   б) 0.3
   в) 0.2
   г) нет
   д) 0.3
   е)

x	-1	1
$\mathbb{P}(X = x)$	0.5	0.5

   ж) $F_X(x) = \begin{cases} 0, & \text{при } x < -1 \\ 0.5, & \text{при } x \in [-1; 1) \\ 1, & \text{при } x \geqslant 1 \end{cases}$

2. а) 0.5
   б) 0.4
   в) 0.2
   г) да
   д) 0.6
   е)

y	-1	0	1
$\mathbb{P}(Y = y)$	0.4	0.2	0.4

   ж) $F_Y(y) = \begin{cases} 0, & \text{при } y < -1 \\ 0.4, & \text{при } y \in [-1; 0) \\ 0.6, & \text{при } y \in [0; 1) \\ 1, & \text{при } y \geqslant 1 \end{cases}$

3. а) 0
   б) 1
   в) 1
   г) 0
   д) 0.6
   е) 0.6
   ж) 0
   з) 0
   и) 0
   к) да, являются некоррелированными, но нельзя утверждать, что являются независимыми

4. а) 0
   б) 1
   в) 1
   г) 0
   д) 0.8
   е) 0.8
   ж) 0
   з) 0
   и) 0
   к) да, являются некоррелированными, но нельзя утверждать, что являются независимыми

5. а) 0.25
   б) 0.2
   в) Обозначим $A = \{X = -1\}$

y	-1	0	1
$\mathbb{P}(Y = y\|A)$	0.4	0.2	0.4

   г) 0
   д) 0.8

6. а) 0.5
   б) 0.2
   в) Обозначим $A = \{X = 1\}$

y	-1	0	1
$\mathbb{P}(Y = y\|A)$	0.4	0.2	0.4

   г) 0
   д) 0.8

7. а) 0
   б) 36
   в) 9
   г) 60
   д) -4
   е) $\frac{-1}{3\sqrt{5}}$
   ж) $\begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix}$

8. а) -4
   б) 8
   в) 1
   г) 10
   д) -6
   е) $\frac{-1}{\sqrt{5}}$
   ж) $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

9. а) 0.3413
   б) 0.0228
   в) 0.1915

10. а) 0.6826
    б) 0.0228
    в) 0.1574

11. 0.4332

12. 0.8185

13. 0.4514
14. 0.5328
15. ≈ 0.8185
16. ≈ 0.9115
17. ≈ 0.6422
18. ≈ 0.9606
19. а) 0.125
    б) 0.5
    в) $f_X(x) = \begin{cases} x + \frac{1}{2}, & \text{при } x \in [0;1] \\ 0, & \text{при } x \notin [0;1] \end{cases}$
    г) $f_Y(y) = \begin{cases} y + \frac{1}{2}, & \text{при } y \in [0;1] \\ 0, & \text{при } y \notin [0;1] \end{cases}$
    д) нет
20. а) $\frac{1}{16}$
    б) $\frac{1}{2}$
    в) $f_X(x) = \begin{cases} 2x, & \text{при } x \in [0;1] \\ 0, & \text{при } x \notin [0;1] \end{cases}$
    г) $f_Y(y) = \begin{cases} 2y, & \text{при } y \in [0;1] \\ 0, & \text{при } y \notin [0;1] \end{cases}$
    д) да
21. а) $\frac{7}{12}$
    б) $\frac{7}{12}$
    в) $\frac{1}{3}$
    г) $-\frac{1}{144}$
    д) $-\frac{1}{11}$
22. а) $\frac{2}{3}$
    б) $\frac{2}{3}$
    в) $\frac{4}{9}$
    г) 0
    д) 0
23. а) $f_Y(y) = \begin{cases} y + \frac{1}{2}, & \text{при } y \in [0;1] \\ 0, & \text{при } y \notin [0;1] \end{cases}$
    б) $f_{X|Y}(x|\frac{1}{2}) = \begin{cases} x + \frac{1}{2}, & \text{при } x \in [0;1] \\ 0, & \text{при } x \notin [0;1] \end{cases}$
    в) $\frac{7}{12}$
    г) $\frac{11}{144}$
24. а) $f_Y(y) = \begin{cases} 2y, & \text{при } y \in [0;1] \\ 0, & \text{при } y \notin [0;1] \end{cases}$
    б) $f_{X|Y}(x|\frac{1}{2}) = \begin{cases} 2x, & \text{при } x \in [0;1] \\ 0, & \text{при } x \notin [0;1] \end{cases}$
    в) $\frac{2}{3}$
    г) $\frac{1}{18}$

12.3. Контрольная работа 3 — Задачный минимум

1. а) ≈ 0.15
   б) $U \sim \mathcal{N}(101, 29)$, $f(u) = \frac{1}{\sqrt{2\pi \cdot 29}} e^{-\frac{1}{2}\frac{(u-101)^2}{29}}$
   в) ≈ 0.02
2. а) 71.14
   б) $f(y|x=170) = \frac{1}{\sqrt{2\pi \cdot 20}} e^{-\frac{1}{2}\frac{(y-71.14)^2}{20}}$
   в) ≈ 0
3. а) 0.25
   б) 0.6875
   в) 0.91(6)
   г) 0.75
   д) -0.28125
4. а) $-1, 0, 1, 1$
   б) -1
   в) 1
   г) $f(x) = \begin{cases} 0, & x < -1 \\ 0.25, & -1 \leqslant x < 0 \\ 0.5, & 0 \leqslant x < 1 \\ 1, & x \geqslant 1 \end{cases}$
5. а) θ
   б) да
6. а) нет, оценка смещена
   б) $c = 2$
7. а) все оценки несмещенные
   б) $\hat{p}_3$ наиболее эффективная
8. да
9. да
10. $\hat{\theta}_{MM} = \sqrt{\frac{\sum_{i=1}^n (X_i - \overline{X})^2 \cdot 20}{n}}$
11. $\hat{\theta}_{MM} = \frac{1}{5}\left(6 - \frac{1}{n}\sum_{i=1}^n X_i^2\right)$, $\hat{\theta}_{MM} = 0.68$

12. $\hat{\theta}_{ML} = \frac{\sum_{i=1}^{n} x_i^2}{n}$

13. $\hat{p}_{ML} = \frac{\sum_{i=1}^{n} x_i}{n}$

14. да

15. $n_1 \approx 260$, $n_2 \approx 232$, $n_3 \approx 658$

12.4. Контрольная работа 4 — Задачный минимум

1. $\left[-1.6 - 1.65 \cdot \frac{2}{\sqrt{3}}; -1.6 + 1.65 \cdot \frac{2}{\sqrt{3}}\right]$

2. $\left[-1.6 - 2.92 \cdot \sqrt{\frac{18.33}{3}}; -1.6 + 2.92 \cdot \sqrt{\frac{18.33}{3}}\right]$

3. $\left[\frac{17.43 \cdot 2}{4.61}; \frac{17.43 \cdot 2}{0.21}\right]$

4. $\left[-1.6 - (-2.6) - 1.96 \cdot \sqrt{\frac{2}{3} + \frac{1}{2}}; -1.6 - (-2.6) + 1.96 \cdot \sqrt{\frac{2}{3} + \frac{1}{2}}\right]$

5. $\left[1.04 - (-0.37) - 3.18 \cdot \sqrt{3.02}\sqrt{\frac{1}{3} + \frac{1}{2}}; 1.04 - (-0.37) + 3.18 \cdot \sqrt{3.02}\sqrt{\frac{1}{3} + \frac{1}{2}}\right]$

6. $\left[0.45 - 1.96 \cdot \sqrt{\frac{0.45 \cdot 0.55}{100}}; 0.45 + 1.96 \cdot \sqrt{\frac{0.45 \cdot 0.55}{100}}\right]$

7. $\left[0.6 - 0.4 - 1.96 \cdot \sqrt{\frac{0.6 \cdot 0.4}{100} + \frac{0.4 \cdot 0.6}{200}}; 0.6 - 0.4 + 1.96 \cdot \sqrt{\frac{0.6 \cdot 0.4}{100} + \frac{0.4 \cdot 0.6}{200}}\right]$

8. $\left[2.5 - 1.96 \cdot \sqrt{\frac{1}{40}}; 2.5 + 1.96 \cdot \sqrt{\frac{1}{40}}\right]$

9. $\left[\frac{1}{0.52} - 1.96 \cdot \sqrt{\frac{1}{100 \cdot 0.52^2}}; \frac{1}{0.52} + 1.96 \cdot \sqrt{\frac{1}{100 \cdot 0.52^2}}\right]$

10. а) ≈ 0.02
    б) ≈ 0.02
    в) ≈ 0.98

11. 0.2

12. $z_{obs} \approx -1.39$, $z_{crit} = 1.28$, нет оснований отвергать H_0.

13. $t_{obs} \approx -0.65$, $t_{crit} = 1.89$, нет оснований отвергать H_0.

14. $z_{obs} \approx 0.93$, $z_{crit} = -1.65$, нет оснований отвергать H_0.

15. $t_{obs} \approx 0.89$, $t_{crit} = -2.35$, нет оснований отвергать H_0.

16. $F_{obs} \approx 95.37$, $F_{crit} = 199.5$, нет оснований отвергать H_0.

17. $z_{obs} \approx 2.04$, $z_{crit} = 1.65$, основная гипотеза отвергается.

18. $z_{obs} \approx 4.16$, $z_{crit} = 1.96$, основная гипотеза отвергается.

19. $\gamma_{obs} \approx 0.26$, $\gamma_{crit} = 5.99$, нет оснований отвергать H_0.

20. $\gamma_{obs} \approx 139.4$, $\gamma_{crit} = 3.84$, основная гипотеза отвергается.

21. $LR_{obs} \approx 5.5$, $LR_{crit} = 3.84$, основная гипотеза отвергается.

13. Решения контрольной номер 1

13.1. 2018-2019

1. Покажем, что все вероятности равны. Пусть K — номер, под которым идёт Вася. Тогда вероятность того, что Вася вытащит правильный билет равна вероятности того, что до него не вытащили правильный билет, помноженная на вероятность того, что он вытащит верный (первое слагаемое) плюс вероятность того, что до него уже вытащили один правильный билет, помноженная на вероятность того, что он вытащит правильный билет. По сути, это формула полной вероятности, где вероятность вытащить у Васи одинаковая при разных условиях, так как он не знает, кто какой билет вытащил, а вероятности условий (вытащили до него правильный билет или нет) разные.

   Вероятность того, что Вася вытащит правильный билет равна
   $$\frac{\text{количество оставшихся хороших билетов}}{25-(K-1)},$$
   так как до Васи тянули ещё $K-1$ студентов.

   Начало второго слагаемого домножается на два, так как неизвестно, какой из двух хороших билетов могли вытянуть до Васи. Следовательно, вероятность увеличивается в два раза.

   $$\begin{aligned}\mathbb{P}(\text{Васе повезёт}) &= \frac{C_{23}^{K-1}}{C_{25}^{K-1}} \cdot \frac{2}{25-K+1} + \frac{2\cdot C_{23}^{K-2}}{C_{25}^{K-1}} \cdot \frac{1}{25-K+1} \\ &= \frac{2(25-K)}{25\cdot 24} + \frac{2(K-1)}{25\cdot 24} \\ &= \frac{48}{25\cdot 24} = \frac{2}{25}\end{aligned}$$

   Устное решение:

   Посмотрим на задачу с точки зрения билетов :) Есть два счастливых билета: «Как зовут лектора?», «Какого цвета учебник?». Какова вероятность того, что билет А достанется Васе? Конечно, $1/25$, билету же всё равно кому доставаться :) Какова вероятность того, что билет Б достанется Васе? Конечно, $1/25$, билету же всё равно кому доставаться :) Итого: Вася вытянет счастливый билет с вероятностью $2/25$ вне зависимости от места в очереди.

2. а) Зададим совместное распределение. Оно определяет все пересечения событий, то есть когда события происходят одновременно. Тогда таблица имеет вид:

	$\xi=0$	$\xi=1$
$\eta=0$	$4/6$	$1/6$
$\eta=1$	$1/6$	0

   б) $\mathbb{P}(\xi=1)\cdot\mathbb{P}(\eta=1) = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}$
   $\mathbb{P}(\xi=1\cap\eta=1)=0$, следовательно, величины не являются независимыми.

   в) Зададим распределение $\xi+\eta$:

x	0	1	2
$\mathbb{P}(\eta+\xi=x)$	$4/6$	$2/6$	0

   г) Зададим совместное распределение величин ξ и $\xi+\eta$. Легко убедиться, что случайные величины одновременно принимают соответствующие значения со следующими вероятностями:

	$\xi=0$	$\xi=1$
$\xi+\eta=0$	$4/6$	0
$\xi+\eta=1$	$1/6$	$1/6$

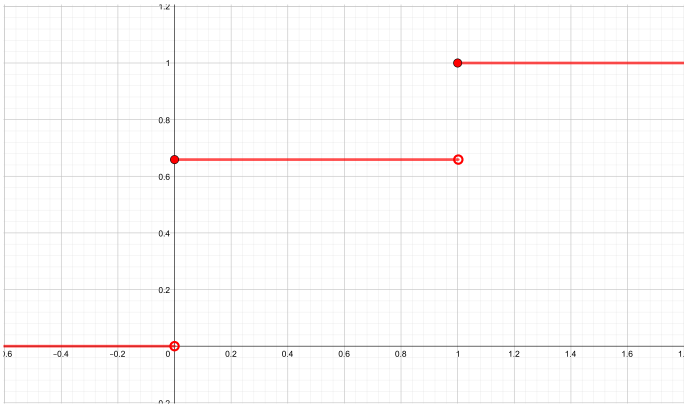


Рис. 4: Функция распределения $\eta + \xi$

Важно заметить, что когда, например, $\xi = 1$, то автоматически при $\xi+\eta = 1 \Rightarrow \eta = 0$ и вероятности легко находятся из частных распределений. Из этой таблицы по формуле условной вероятности легко найти распределение:

x	0	1
$\mathbb{P}(\xi = x \mid \xi + \eta = 1)$	1/2	1/2

3. а)

$$f_\xi(x) = \begin{cases} cx^2, x \in [0;1] \\ 0, \text{иначе} \end{cases}$$

$$\int_0^1 cx^2dx = \frac{1}{3}cx^3\Big|_0^1 = \frac{1}{3} = 1 \to c = 3$$

б) $\mathbb{P}(\xi = \frac{1}{2}) = 0$ по определению.

$$\mathbb{P}\left(\xi \in \left[0, \frac{1}{3}\right]\right) = \int_0^{\frac{1}{3}} 3x^2dx = x^3\Big|_0^{\frac{1}{3}} = \frac{1}{27}$$

в)

$$F_\xi(x) = \begin{cases} 0, x < 0 \\ x^3, x \in [0, 1] \\ 1, x > 1 \end{cases}$$

г)

$$\mathbb{P}\left(\xi \in \left[\frac{1}{3}, \frac{3}{2}\right]\right) = F(3/2) - F(1/3) = 1 - \frac{1}{27} = \frac{26}{27}$$

д)

$$\mathbb{P}\left(\xi \leqslant \frac{1}{2} \mid \xi \geqslant \frac{1}{3}\right) = \frac{\mathbb{P}(\xi \leqslant \frac{1}{2}, \xi \geqslant \frac{1}{3})}{\mathbb{P}(\xi \geqslant \frac{1}{3})} = \frac{(1/2)^3 - (1/3)^3}{1 - (1/3)^3}$$

е) Так как функция плотности — парабола ветвями вверх $[0, 1]$, на котором она определена, максимум будет достигаться на возрастающей части ветки, то есть в правом конце отрезка. Следовательно, мода равна 1.

Ожидание: $\mathbb{E}(\xi) = \int_0^1 x3x^2dx = \frac{3}{4}x^4|_0^1 = \frac{3}{4}$

Пусть q — медиана, тогда $\int_0^q 3x^2dx = x^3|_0^q = \frac{1}{2}$ и $q = \left(\frac{1}{2}\right)^{\frac{1}{3}}$

4. а) $n = 12, q = \frac{1}{3}, p = \frac{2}{3}$
$12 \cdot \frac{2}{3} - \frac{1}{3} \leqslant \mu_0 \leqslant 12 \cdot \frac{2}{3} + \frac{2}{3}$
$\frac{23}{3} \leqslant \mu_0 \leqslant \frac{26}{3}$
$7.6 \leqslant \mu_0 \leqslant 8.6 \to \mu_0 = 8$

б) $\mathbb{P}(\text{Семь судей проголосуют верно}) = \frac{12!}{7!5!}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^7 \approx 0.19$

в)

$\mathbb{P}(\text{Ровно 7 человек голосуют за обвинительный приговор}) = pC_{12}^7(2/3)^7(1/3)^5 + (1-p)C_{12}^7(1/3)^7(2/3)^5$

г) Решаем уравнение

$$pC_{12}^7(2/3)^7(1/3)^5 + (1-p)C_{12}^7(1/3)^7(2/3)^5 = 0.17$$

5. а) У каждого из 11 людей 11 вариантов, где выйти. А подходит нам 11! вариантов, так как нам нужно расставить 11 человек в очередь, в каком порядке им выходить. По классической формуле:

$$\mathbb{P}(\text{ровно один человек на каждом этаже}) = \frac{11!}{11^{11}}$$

б) Найдём условную вероятность того что все выйдут не выше 9-го, если известно, что они не вышли на со 2-го по 6-ой. Вероятность того, что все 11 человек выйдут на 7, 8 или 9, будет равна

$$\mathbb{P}(A \cap B) = \left(\frac{3}{11}\right)^{11}$$

Вероятность условия равна вероятности того, что никто не выйдет со 2-го по 6-ой этажи,

$$\mathbb{P}(B) = \left(\frac{6}{11}\right)^{11}$$

Тогда условная вероятность равна

$$\mathbb{P}(A|B) = \frac{\left(\frac{3}{11}\right)^{11}}{\left(\frac{6}{11}\right)^{11}} = \left(\frac{1}{2}\right)^{11} \approx 0.0005$$

13.2. 2017-2018

1. а) События называются независимыми, если $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

б) Запасёмся всеми нужными вероятностями:
$\mathbb{P}(A) = \frac{1}{2}$
$\mathbb{P}(B) = \frac{1}{3}$
$\mathbb{P}(C) = \frac{1}{2}$
$\mathbb{P}(A \cap C) = \frac{1}{3}$ — выпадет чётое число больше трёх
$\mathbb{P}(A \cap B) = \frac{1}{6}$ — выпадет чётное число, кратное трём
$\mathbb{P}(A \cap C) = \frac{1}{6}$ — выпадет число, большее трёх и кратное трём
Теперь можно проверять независимость:
$\mathbb{P}(A \cap C) \neq \mathbb{P}(A) \cdot \mathbb{P}(C) \Rightarrow$ не являются независимыми
$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \Rightarrow$ являются независимыми
$\mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C) \Rightarrow$ являются независимыми

2. а) Количество возможных вариантов ТМ: C_{10}^2, количество возможных вариантов ЗМ: C_{24}^2. Количество их возможных сочетаний: $C_{10}^2 \cdot C_{24}^2$, где $C_n^k = \frac{n!}{k!(n-k)!}$.

б) По классическому определению вероятностей, предполагая исходы равновероятными, искомая вероятность равна $\frac{C_{16}^2}{C_{24}^2}$.

в) По тому же принципу:

$$\frac{C_k^2}{C_{10}^2} = \frac{1}{15} \Rightarrow \frac{\frac{k!}{2!(k-2)!}}{\frac{10!}{2!\cdot 8!}} = \frac{1}{15} \Rightarrow \frac{(k-1)k}{2}\frac{2}{9\cdot 10} = \frac{1}{15}$$

Получаем квадратное уравнение вида $k^2 - k - 6 = 0$ с корнями -2 и 3. Так как k не может быть отрицательным, ответ 3.

3. а) Если эксперт отдаёт предпочтение Fit, то это можно интерпретировать как «успех» в схеме Бернулли. Так как ξ - количество успехов, $k \in [0; 4]$, $p = \frac{1}{3}$, то

$$\mathbb{P}(\xi = k) = C_n^k (p)^k (1-p)^{n-k}$$

Большинство означает, что либо три, либо четыре эксперта выбрали Fit.

$$\mathbb{P}(\xi = 3) = C_4^3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^1 = \frac{8}{81}$$

$$\mathbb{P}(\xi = 4) = C_4^4 \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^0 = \frac{1}{81}$$

$$\mathbb{P}(\xi > 2) = \frac{9}{81}$$

б) Аналогично:

$$\mathbb{P}(\xi = 0) = C_4^0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^4 = \frac{16}{81}$$

$$\mathbb{P}(\xi = 1) = C_4^1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^3 = \frac{32}{81}$$

$$\mathbb{P}(\xi = 2) = C_4^2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^2 = \frac{24}{81}$$

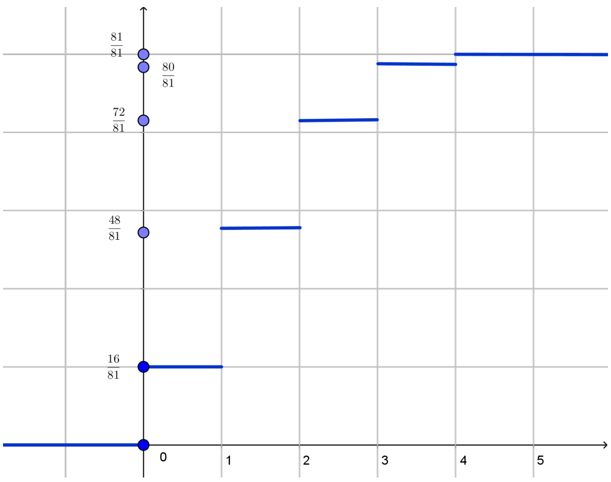


Рис. 5: Функция распределения

в) Все вероятности посчитаны, видим, что наибольшая достигается при $\xi = 1$.

г) $\mathbb{E}(X) = np = \frac{4}{3}$, $\text{Var}(X) = npq = \frac{8}{9}$

4. а) Так как указано, что цена сметаны распределена равномерно на отерзке $[250, 1000]$, максимальное значение цены — 1000, это и есть необходимая сумма.

   б) Вспомним, что функция распределения $F(x) = \mathbb{P}(X \leqslant x)$, нужно найти такой x, что $\mathbb{P}(X \leqslant x) = 0.9$:
$$0.9 = 1 - \exp(-x^2) \Rightarrow \exp(-x^2) = 0.1 \Rightarrow -x^2 = \ln(0.1) \Rightarrow x = \sqrt{-\ln(0.1)}$$

   в) Взяв производную от функции распределения списка без сметаны, получим функцию плотности:
$$f_X(x) = \begin{cases} 2x\exp(-x^2) & x \geqslant 0 \\ 0 & \text{иначе} \end{cases}$$
   Найдём математическое ожидание:
$$\int_0^{+\infty} 2x^2\exp(-x^2)dx = -x\exp(-x^2)\Big|_0^{+\infty} + \int_0^{+\infty}\exp(-x^2)dx = \frac{\sqrt{\pi}}{2}$$

   г) Математическое ожидание суммы случайных величин равно сумме математических ожиданий случайных влечин, если они существуют. Математическое ожидание от цены сметаны равно: $\frac{1000+250}{2} = 625$. Математическое ожидание списка без сметаны было найдено в предыдущем пункте, его осталось перевести в рубли. Получаем ответ: $625 + \frac{\sqrt{\pi}}{2} \cdot 1000$.

   д) Так как обе величины имеют абсолютно непрерывные распределения, вероятность попасть в конкретную точку равна нулю.

5. а) $\mathbb{P}(\text{детектор показал ложь и подозреваемый лжёт}) = 0.9 \cdot 0.1 + 0.1 \cdot 0.95 = 0.185$

   б) $\mathbb{P}(\text{невиновен}|\text{детектор показал ложь}) = \frac{0.9 \cdot 0.1}{0.185} = \frac{90}{185}$

   в) $\mathbb{P}(\text{эксперт точно выявит преступника}) = (0.9)^9 \cdot 0.95$

   г) $\mathbb{P}(\text{эксперт ошибочно выявит преступника}) = 9 \cdot 0.1 \cdot 0.9^8 \cdot 0.05$

13.3. 2016-2017

1. а) Возможны четыре равновероятные ситуации:
$$\mathbb{P}(\text{ММ}) = \mathbb{P}(\text{МД}) = \mathbb{P}(\text{ДМ}) = \mathbb{P}(\text{ДД}) = 1/4$$
   Посчитаем условную вероятность:
$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\text{МД, ДМ})}{\mathbb{P}(\text{ДМ, МД, ДД})} = \frac{2/4}{3/4} = \frac{2}{3}$$

   б) События A и B называются независимыми, если $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$
   В нашем случае: $\mathbb{P}(A \cap B) = \mathbb{P}(\text{МД, ДМ}) = 2/4$, $\mathbb{P}(A) \cdot \mathbb{P}(B) = 3/4 \cdot 3/4$.
   Следовательно, $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \cdot \mathbb{P}(B)$, значит, события A и B не являются независимыми.

2. Пусть событие A_i означает, что i-ый узел системы дал сбой, а событие B_N, что вся система дала сбой. В условии сказано, что $\mathbb{P}(A_i) = 10^{-6}$, а найти нужно такое максимальное $N \in \mathbb{N}$, при котором
$$\mathbb{P}(B_N) \leqslant \frac{1}{10^2}$$
$$\begin{aligned}\mathbb{P}(B_N) &= \mathbb{P}\left(\cup_{i=1}^n A_i\right) = 1 - \mathbb{P}((\cup_{i=1}^n A_i)^c) \\ &\overset{\text{ф-ла де Моргана}}{=} 1 - \mathbb{P}\left(\cap_{i=1}^N A_i^c\right) \overset{A_1,\ldots,A_N-\text{ независ.}}{=} 1 - \mathbb{P}(A_1^c) \cdot \ldots \cdot \mathbb{P}(A_N^c) \\ &= 1 - \left(1 - \frac{1}{10^6}\right)^N\end{aligned}$$

Чтобы найти такое максимальное $N \in \mathbb{N}$, надо решить следующее неравенство

$$\begin{aligned}
&1-\left(1-10^{-6}\right)^{N} \leqslant 10^{-2} \\
&1-10^{-2} \leqslant \left(1-10^{-6}\right)^{N} \\
&\ln\left(1-10^{-2}\right) \leqslant N \ln\left(1-10^{-6}\right) \\
&N \leqslant \frac{\ln\left(1-10^{-2}\right)}{\ln\left(1-10^{-6}\right)} \approx 10050.33
\end{aligned}$$

Значит, максимальное N равно 10050.

3. Введём обозначения для событий. Пусть A означает, что человек имеет заболевание лёгких, а B, что человек работал в шахте.

   В условии сказано, что $\mathbb{P}(B \mid A) = 0.22$, $\mathbb{P}(B \mid A^c) = 0.14$, $\mathbb{P}(A) = 0.04$.

   а) Нужно найти
   $$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)}$$
   Для этого с помощью формулы полной вероятности посчитаем
   $$\mathbb{P}(B) = \mathbb{P}(B \mid A)\,\mathbb{P}(A) + \mathbb{P}(B \mid A^c)\,\mathbb{P}(A^c) = 0.22 \cdot 0.04 + 0.14 \cdot 0.96 = 0.1432$$
   Осталось подставить значения:
   $$\mathbb{P}(A \mid B) = \frac{0.22 \cdot 0.04}{0.1432} \approx 0.0615$$

   б) Все необходимые значения для второго пункта у нас есть, осталось применить формулу условной вероятности:
   $$\begin{aligned}
   \mathbb{P}(A \mid B^c) &= \frac{\mathbb{P}(A \cap B^c)}{\mathbb{P}(B^c)} = \frac{\mathbb{P}(B^c \cap A)}{\mathbb{P}(A)} \cdot \frac{\mathbb{P}(A)}{\mathbb{P}(B^c)} = \mathbb{P}(B^c \mid A) \cdot \frac{\mathbb{P}(A)}{\mathbb{P}(B^c)} \\
   &= (1-\mathbb{P}(B \mid A)) \cdot \frac{\mathbb{P}(A)}{1-\mathbb{P}(B)} = (1-0.22) \cdot \frac{0.04}{1-0.1432} \approx 0.0364
   \end{aligned}$$

4. Введём индикатор события «Петя дал верный ответ на i-ый вопрос»:
   $$X_i = \begin{cases} 1, & \text{если на } i\text{-ый вопрос теста Петя дал верный ответ} \\ 0, & \text{иначе} \end{cases}$$
   Заметим, что $X_i \sim Be\,(p = 1/5)$, $X_1, \ldots, X_{17}$ – независимы, $X = X_1 + \ldots + X_{17}$ – общее число верных ответов, $X \sim Bin\,(n = 17, p = 1/5)$.

   а) Наибольшее вероятное число правильных ответов m_0 может быть нвйдено по формуле:

      1) если число $(n \cdot p - q)$ – не целое, где $q := 1 - p$, то
      $$m_0 = [np - q] + 1,$$
      2) если число $(n \cdot p - q)$ – целое, то наиболее вероятных значений m_0 два:
      $$m_0' = np - q \text{ и } m_0'' = np - q + 1$$

   Итак, поскольку $np - q = 17 \cdot \frac{1}{5} - \frac{4}{5} = 2.6$ – не целое, наиболее вероятное число верных ответов m_0 может быть найдено по формуле из пункта (1):
   $$m_0 = [np - q] + 1 = [2.6] + 1 = 3$$

б)

$$\mathbb{E}(X) = np = 17 \cdot \frac{1}{5} = 3.4$$

$$\text{Var}(X) = npq = 17 \cdot \frac{1}{5} \cdot \frac{4}{5} = 2.72$$

в)

$$\begin{aligned} \mathbb{P}(\text{у Пети «отлично»}) &= \mathbb{P}(X \geqslant 15) = \mathbb{P}(X = 15) + \mathbb{P}(X = 16) + \mathbb{P}(X = 17) \\ &= C_{17}^{15} \cdot \left(\frac{1}{5}\right)^{15} \cdot \left(\frac{4}{5}\right)^{2} + C_{17}^{16} \cdot \left(\frac{1}{5}\right)^{16} \cdot \left(\frac{4}{5}\right)^{1} + C_{17}^{17} \cdot \left(\frac{1}{5}\right)^{17} \cdot \left(\frac{4}{5}\right)^{0} \\ &= 136 \cdot \frac{16}{5^{17}} + 17 \cdot \frac{4}{5^{17}} + \frac{1}{5^{17}} \approx 2.94 \cdot 10^{-9} \end{aligned}$$

г) Рассмотрим первый вопрос теста. Петя может выбрать первый ответ с вероятностью $1/5$, и Вася может выбрать первый ответ с вероятностью $1/5$. Тогда они оба выберут одинаковый ответ с вероятностью $1/25$. Вариантов ответа в каждом вопросе 5, значит, вероятность совпадения ответа в одном вопросе равна $1/5$. Всего вопросов 17, тогда получаем

$$\mathbb{P}(\text{все ответы Пети и Васи совпадают}) = \left(\frac{1}{5}\right)^{17}$$

5. Введём случайную велчину η, которая означает число потенциальных покупателей, с которыми контактировал продавец оборудования. По условию задачи, η имеет таблицу распеределения:

y	1	2
$\mathbb{P}(\eta = y)$	$1/3$	$2/3$

Случайная величина ξ может принимать значения $0, 50000$ и 100000

а) Найдём $\mathbb{P}(\xi = 0)$. По формуле полной вероятности, имеем:

$$\begin{aligned} \mathbb{P}(\xi = 0) &= \mathbb{P}(\xi = 0 \mid \eta = 1) \cdot \mathbb{P}(\eta = 1) + \mathbb{P}(\xi = 0 \mid \eta = 2) \cdot \mathbb{P}(\eta = 2) \\ &= 0.9 \cdot \frac{1}{3} + 0.9 \cdot 0.9 \cdot \frac{2}{3} = 0.84 \end{aligned}$$

б) Найдём $\mathbb{P}(\xi = 50000)$ и $\mathbb{P}(\xi = 100000)$:

$$\begin{aligned} \mathbb{P}(\xi = 50000) &= \mathbb{P}(\xi = 50000 \mid \eta = 1) \cdot \mathbb{P}(\eta = 1) + \mathbb{P}(\xi = 50000 \mid \eta = 2) \cdot \mathbb{P}(\eta = 2) \\ &= 0.1 \cdot \frac{1}{3} + 2 \cdot 0.1 \cdot 0.9 \cdot \frac{2}{3} = 0.15(3) \end{aligned}$$

$$\begin{aligned} \mathbb{P}(\xi = 100000) &= \mathbb{P}(\xi = 100000 \mid \eta = 1) \cdot \mathbb{P}(\eta = 1) + \mathbb{P}(\xi = 100000 \mid \eta = 2) \cdot \mathbb{P}(\eta = 2) \\ &= 0 \cdot \frac{1}{3} + 0.1 \cdot 0.1 \cdot \frac{2}{3} = 0.00(6) \end{aligned}$$

Таблица распределения случайной величина ξ имеет вид:

x	0	50 000	100 000
$\mathbb{P}(\xi = x)$	0.84	0.15(3)	0.00(6)

Тогда функция распределения случайной величины ξ имеет вид:

$$F_\xi(x) = \begin{cases} 0 & \text{при } x < 0 \\ 0.84 & \text{при } 0 \leqslant x < 50\,000 \\ 0.84 + 0.15(3) & \text{при } 50\,000 \leqslant x < 100\,000 \\ 1 & \text{при } x > 100\,000 \end{cases}$$

Опр.: $F_\xi = \mathbb{P}(\xi \leqslant x), x \in \mathbb{R}$

в)

$$\mathbb{E}(X) = 0 \cdot 0.84 + 50\,000 \cdot 0.15(3) + 100\,000 \cdot 0.00(6) = 8\,333.(3)$$

$$\begin{aligned}\operatorname{Var}(X) = (0 - 8\,333.(3))^2 \cdot 0.84 + (50\,000 - 8\,333.(3))^2 \cdot 0.15(3) \\ + (100\,000 - 8\,333.(3))^2 \cdot 0.00(6) = 380\,555\,555.(5)\end{aligned}$$

6. а) $f_\xi(x) = \begin{cases} \frac{1}{b} & \text{при } x \in [0, b] \\ 0 & \text{при } x \notin [0, b] \end{cases}$

б) Известно, что если $\xi \sim U[a, b]$, то $\mathbb{E}(\xi) = \frac{a+b}{2}$. Стало быть, из уравнения $\mathbb{E}(\xi) = 1$ получаем $\frac{b}{2} = 1$, то есть $b = 2$.

в) Известно, что если $\xi \sim U[a, b]$, то $\operatorname{Var}(\xi) = \frac{(b-a)^2}{12}$. Значит, $\operatorname{Var}(\xi) = \frac{2^2}{12} = \frac{1}{3}$

г) Воспользуемся формулой $\mathbb{P}(\xi \in B) = \int_B f_\xi(x)dx$. Имеем:

$$\mathbb{P}(\xi > 1) = \mathbb{P}(\xi \in (1, +\infty)) = \int_1^{+\infty} f_\xi(x)dx = \int_1^2 \frac{1}{2}dx = \frac{1}{2}$$

д) Требуется найти такое минимальное число $q_{0.25}$, что $\int_{-\infty}^{q_{0.25}} f_\xi(x)dx = 0.25$. Итак:

$$\int_{-\infty}^{q_{0.25}} f_\xi(x)dx = 0.25 \Leftrightarrow \int_{-\infty}^{q_{0.25}} \frac{1}{2}dx = 0.25 \Leftrightarrow \frac{1/2}{q_{0.25}} = 0.25 \Leftrightarrow$$

$$q_{0.25} = 2 \cdot 0.25 = 0.5$$

е)

$$\begin{aligned}\mathbb{E}\left((\xi - \mathbb{E}(\xi))^{2017}\right) &= \int_{-\infty}^{+\infty} (x - \mathbb{E}(\xi))^{2017} \cdot f_\xi(x)dx = \int_{-\infty}^{+\infty} (x-1)^{2017} f_\xi(x)dx \\ &= \int_0^2 (x-1)^{2017} \cdot \frac{1}{2}dx = \frac{(x-1)^{2018}}{2018} \cdot \frac{1}{2}\Big|_{x=0}^{x=2} = 0\end{aligned}$$

ж) $F_\xi(x) = \begin{cases} 0 & \text{при } x < 0 \\ \frac{x}{2} & \text{при } 0 \leqslant x \leqslant 2 \\ 1 & \text{при } x > 2 \end{cases}$

з) Согласно условиям задачи, время до прихода 1-го поезда есть ξ; время до прихода 2-го поезда равно $\xi + b$; время до прихода 3-го (заветного) поезда есть $\xi + 2b$. Таким образом, Марья Ивановна в среднем ожидает «своего» поезда $\mathbb{E}(\xi + 2b) = 1 + 2b = 1 + 2 \cdot 2 = 5$ минут. При этом $\operatorname{Var}(\xi + 2b) = \operatorname{Var}(\xi) = 1/3$

к) Пусть τ – наименьший номер поезда без «подозрительных лиц». По условию задачи, таблица распределения случайной величины τ имеет вид:

t	1	2	3	4	...
$\mathbb{P}(\tau = t)$	$1/4$	$3/4 \cdot 1/4$	$(3/4)^2 \cdot 1/4$	$(3/4)^3 \cdot 1/4$	...

То есть случайная величина τ имеет геометрическое распределение с параметром $p = 1/4$ ($\tau \sim G(p = 1/4)$).

Несложно сообразить, что время ожидания Глафирой Петровной «своего» поезда составляет: $\eta := \xi + b(\tau - 1)$. Стало быть, $\mathbb{E}(\eta) = \mathbb{E}(\xi) + b \cdot (\mathbb{E}(\tau) - 1) = 1 + 2 \cdot (4 - 1) = 7$ минут.

Здесь мы воспользовались тем фактом, что если $\eta \sim G(p)$, то $\mathbb{E}(\eta) = 1/p$

и) Найдём теперь вероятность $\mathbb{P}(\eta \geqslant 5)$. Для нахождения искомой вероятности воспользуемся формулой полной вероятности:

$$\mathbb{P}(\eta \geqslant 5) = \mathbb{P}(\eta \geqslant 5, \tau < 3) + \mathbb{P}(\eta \geqslant 5, \tau = 3) + \mathbb{P}(\eta \geqslant 5, \tau > 3)$$

Если Глафира уехала на первом или втором поезде, то ждать больше 5 минут она не могла, то есть $\mathbb{P}(\eta \geqslant 5, \tau < 3) = 0$.

Если Глафира уехала на третьем поезде, то чтобы ждать больше пяти минут, ей нужно ждать первый поезд больше минуты, то есть $\mathbb{P}(\eta \geqslant 5, \tau = 3) = 0.5\,\mathbb{P}(\tau = 3)$.

Если Глафира уехала на четвертом поезде или позже, то она точно ждала больше 5 минут, $\mathbb{P}(\eta \geqslant 5, \tau > 3) = \mathbb{P}(\tau > 3)$.

$$\mathbb{P}(\eta \geqslant 5) = 0.5\,\mathbb{P}(\tau = 3) + \mathbb{P}(\tau > 3) = 0.5 \cdot (3/4)^2 \cdot (1/4) + (3/4)^3 = 63/128$$

7. Пусть ξ — случайная величина, обозначающая число остановок лифта. Представим её в виде суммы $\xi = \xi_2 + \ldots + \xi_{10}$, где ξ_i — индикатор того, что лифт остановился на i-ом этаже, то есть

$$\xi_i = \begin{cases} 1 & \text{если лифт остановился} \\ 0 & \text{иначе} \end{cases} \qquad \forall i = 2, \ldots, 10$$

Найдём соответсвующие вероятности:

$$\mathbb{P}(\xi_i = 0) = \left(\frac{8}{9}\right)^9$$

$$\mathbb{P}(\xi_i = 1) = 1 - \mathbb{P}(\xi = 0) = 1 - \left(\frac{8}{9}\right)^9$$

Тогда $\mathbb{E}(\xi_i) = \mathbb{P}(\xi_i = 0) \cdot 0 + \mathbb{P}(\xi_i = 1) \cdot 1 = 1 - \left(\frac{8}{9}\right)^9$, и в итоге получаем:

$$\mathbb{E}(\xi) = 9 \cdot \mathbb{E}(\xi_i) = 9 \cdot \left(1 - \left(\frac{8}{9}\right)^9\right)$$

13.4. 2015-2016

1. α) Найдём вероятности каждого события: $\mathbb{P}(A) = 1/2, \mathbb{P}(B) = 1/2, \mathbb{P}(C) = 1/2$.

   Проверим попарную независимость:

   - $\mathbb{P}(A \cap B) = 1/4, \mathbb{P}(A) \cdot \mathbb{P}(B) = 1/2 \cdot 1/2 = 1/4$
   - $\mathbb{P}(A \cap C) = 1/4, \mathbb{P}(A) \cdot \mathbb{P}(C) = 1/2 \cdot 1/2 = 1/4$
   - $\mathbb{P}(B \cap C) = 1/4, \mathbb{P}(B) \cdot \mathbb{P}(C) = 1/2 \cdot 1/2 = 1/4$

   Значит, события попарно независимы.

   β) События A_1, A_2, A_3 называются независимыми в совокупности, если $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3)$.

   В нашем случае: $\mathbb{P}(A \cap B \cap C) = 0, \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C) = (1/2)^3$, следовательно, события не являются независимыми в совокупности.

2. α) Воспользуемся формулой полной вероятности:

$$\begin{aligned}\mathbb{P}(\text{выпала «6»}) &= \mathbb{P}(\text{выпала «6»} \mid \text{взят белый кубик}) \cdot \mathbb{P}(\text{взят белый кубик}) \\ &+ \mathbb{P}(\text{выпала «6»} \mid \text{взят красный кубик}) \cdot \mathbb{P}(\text{взят красный кубик}) \\ &= \frac{1}{6} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{4}\end{aligned}$$

   β) Воспользуемся формулой условной вероятности и результатом предыдущего пункта:

$$\begin{aligned}\mathbb{P}(\text{взят красный кубик} \mid \text{выпала «6»}) &= \frac{\mathbb{P}(\text{взят красный кубик} \cap \text{выпала «6»})}{\mathbb{P}(\text{выпала «6»})} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{4}} = \frac{2}{3}\end{aligned}$$

3. α) Совместное распределение имеет вид:

	$\xi=1$	$\xi=2$	$\xi=3$	$\xi=4$	$\xi=5$	$\xi=6$
$\eta=1$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{1}{3}\cdot\frac{1}{6}*$
$\eta=2$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{1}{3}\cdot\frac{1}{6}*$
$\eta=3$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{1}{3}\cdot\frac{1}{6}*$
$\eta=4$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}*$	$\frac{1}{3}\cdot\frac{1}{6}*$
$\eta=5$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{1}{3}\cdot\frac{1}{6}*$
$\eta=6$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{1}{3}\cdot\frac{1}{6}$

β) $\mathbb{P}(\text{выиграет белый кубик}) = (6+5+4+3+2)\cdot\frac{2}{15}\cdot\frac{1}{6}+1\cdot\frac{1}{3}\cdot\frac{1}{6}=\frac{1}{2}$.
Значит, Пете безразлично, какой кубик брать.

γ) $F_\zeta(x)=\mathbb{P}(\zeta\leqslant x)$
Выпишем таблицу распределения случайной величины ζ:

x	1	2	3	4	5	6
$\mathbb{P}(\zeta=x)$	$\frac{2}{15}\cdot\frac{1}{6}$	$\frac{2}{15}\cdot\frac{1}{6}\cdot 3$	$\frac{2}{15}\cdot\frac{1}{6}\cdot 5$	$\frac{2}{15}\cdot\frac{1}{6}\cdot 7$	$\frac{2}{15}\cdot\frac{1}{6}\cdot 9$	$\frac{1}{3}\cdot\frac{1}{6}\cdot 6+\frac{2}{15}\cdot\frac{1}{6}\cdot 5$

Тогда функция распределения имеет вид:

$$F_\zeta(x)=\begin{cases}0 & x\leqslant 1\\ \frac{1}{45} & 1<x\leqslant 2\\ \frac{4}{45} & 2<x\leqslant 3\\ \frac{9}{45} & 3<x\leqslant 4\\ \frac{16}{45} & 4<x\leqslant 5\\ \frac{25}{45} & 5<x\leqslant 6\\ 1 & x>6\end{cases}$$

δ) $\mathbb{E}(\zeta)=\frac{2}{15}\cdot\frac{1}{6}\cdot 1+\frac{2}{15}\cdot\frac{1}{6}\cdot 3\cdot 2+\frac{2}{15}\cdot\frac{1}{6}\cdot 5\cdot 3+\frac{2}{15}\cdot\frac{1}{6}\cdot 7\cdot 4+\frac{2}{15}\cdot\frac{1}{6}\cdot 9\cdot 5+\frac{1}{3}\cdot\frac{1}{6}\cdot 6+\frac{2}{15}\cdot\frac{1}{6}\cdot 6=\frac{43}{9}\approx 4.8$

4. Пусть x — вероятность того, что мужчина честно любит петь в душе.

Распишем по формуле полной вероятности вероятность получить ответ «да»:

$$P(\text{ответ «Да»}) = 1\cdot\mathbb{P}(\text{выпала «6»}) + x\cdot(\mathbb{P}(\text{выпала «2»})+\mathbb{P}(\text{выпала «3»}) + \mathbb{P}(\text{выпала «4»})+\mathbb{P}(\text{выпала «5»})) = 1\cdot\frac{1}{6}+x\cdot\frac{4}{6}\Rightarrow x=\frac{3}{4}$$

Тогда истинный процент «певцов» составляет 75%

5. Предположим, что ваше имя — Студент (7 букв), а фамилия — Идеальный (9 букв).

α) $\mathbb{P}(\text{напишет фаимлию правильно}) = (0.9)^9$

β) $\mathbb{P}(\text{ровно 2 ошибки в имени}) = C_7^2\cdot 0.1^2\cdot 0.9^5$

γ) Наиболее вероятное число ошибок — 1

δ) $\mathbb{P}(\text{допустит хотя бы одну ошибку}) = 1-\mathbb{P}(\text{не допустит ни одной ошибки}) = 1-(0.9)^{16}$

6. α) Из условия $\int_0^1(cy^2+y)dy=1$ получаем, что $c=3/2$.

β) $F_Y(y)=\begin{cases}1 & y>1\\ \frac{y^3+y^2}{2} & 0<y\leqslant 1\\ 0 & y<0\end{cases}$

γ) $\mathbb{P}(Y<0.5)=\int_0^{0.5}\left(\frac{3}{2}y^2+y\right)dy=\frac{3}{16}$

δ) $F_Y(y) = 0.5 \Rightarrow y \approx 0.75$

ϵ) $\mathbb{P}(Y > 0.5 \mid Y \geqslant 0.25) = \frac{\mathbb{P}(Y>0.5)}{\mathbb{P}(Y\geqslant 0.25)} = \frac{1-\frac{3}{16}}{\int_{0.25}^{1}(\frac{3}{2}y^2+y)dy} = \frac{104}{123}$

7. α) $\mathbb{P}(\text{кисточка окажется на слоне}) = \frac{1}{1.5} = \frac{2}{3}$

β) $f_{\xi,\eta}(x,y) = \frac{1}{1.5}$

γ) $f_\xi(x) = \int_0^1 \frac{1}{1.5}dy = 1.5$

$f_\eta(y) = \int_0^{1.5} \frac{1}{1.5}dx = 1$

δ) Да, поскольку $f_\xi(x) \cdot f_\eta(y) = f_{\xi,\eta}(x,y)$

ϵ) $f_{\xi+\eta}(t) = \int_{-\infty}^{+\infty} f_\xi(u) f_\eta(t-u)du$

13.5. 2014-2015

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

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

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

   $\mathbb{P}(AB)$ ищется как отношение площади квадрата с вершинами в (0, 0), (0, 1/2), (1/2, 1/2), (1/2, 0) к площади данного треугольника, то есть:

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

   $\mathbb{P}(A)$ ищется как отношение площади трапеции с вершинами в (0, 0), (0, 1), (1/2, 1/2), (1/2, 0) к площади данного треугольника, то есть:

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

   $\mathbb{P}(B)$ ищется как отношение площади трапеции с вершинами в (0, 0), (1, 0), (1/2, 1/2), (0, 1/2) к площади данного треугольника, то есть:

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

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

   Получается, события A и B зависимы.

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

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

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.

$$\mathbb{P}(X \geqslant 2) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1)$$

$\mathbb{P}(X = 0) = \left(\frac{499}{500}\right)^{400}$ — каждая из 400 опечаток не должна попасть на 13 страницу.
$\mathbb{P}(X = 1) = 400 \cdot \frac{1}{500} \cdot \left(\frac{499}{500}\right)^{399}$ — ровно одна опечатка (а есть 400 вариантов) должна попасть на 13 страницу, а остальные — мимо. Соответственно:

$$\mathbb{P}(X \geqslant 2) = 1 - \left(\frac{499}{500}\right)^{400} - 400 \cdot \frac{1}{500} \cdot \left(\frac{499}{500}\right)^{399} \approx 0.19$$

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

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

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

б) 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 $X = \sum\limits_{i=1}^{400} X_i$. Рассмотрим отдельно X_i:

x	1	0
$\mathbb{P}(X = x)$	$\frac{1}{500}$	$\frac{499}{500}$

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

$$\mathbb{E}(X_i) = \frac{1}{500} = \mathbb{E}(X_i^2)$$
$$\operatorname{Var}(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = \frac{1}{500} - \left(\frac{1}{500}\right)^2 = \frac{499}{500^2}$$

Значит

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

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

$$\mathbb{P}(X \geqslant 2) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) = 1 - \frac{0.8^0}{0!} e^{-0.8} - \frac{0.8^1}{1!} e^{-0.8} \approx 0.19$$

Осталось найти наиболее вероятное число опечаток на 13 странице:

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

Очевидно, что эта функция убывает по k, ведь с ростом k:
$k!$ растет, а 0.8^k убывает. Значит наиболее вероятное число ошибок — $X = 0$

в) 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. Пусть событие A означает, что медицинский тест показал наличие заболевания. Событие B — заболевание на самом деле есть.

Перепишем условие задачи:

Чувствительность теста $= \mathbb{P}(A|B)$

Специфичность теста $= \mathbb{P}(A^c|B^c)$

Прогностическая сила теста $= \mathbb{P}(B|A)$

$\mathbb{P}(B) = 0.01 \Rightarrow \mathbb{P}(B^c) = 0.99$

По условию, чувствительность теста равна 0.9, тогда из формулы условной вероятности:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \Rightarrow \mathbb{P}(A \cap B) = 0.9 \cdot 0.01 = 0.009$$

При этом очевидно, что:

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B) \Rightarrow \mathbb{P}(A^c \cap B) = 0.01 - 0.009 = 0.001$$

По условию специфичность теста равна 0.95, тогда из формулы условной вероятности:

$$\mathbb{P}(A^c|B^c) = \frac{\mathbb{P}(A^c \cap B^c)}{\mathbb{P}(B^c)} \Rightarrow \mathbb{P}(A^c \cap B^c) = 0.95 \cdot 0.99 = 0.9405$$

При этом очевидно, что:

$$\mathbb{P}(B^c) = \mathbb{P}(A \cap B^c) + \mathbb{P}(A^c \cap B^c) \Rightarrow \mathbb{P}(A \cap B^c) = 0.99 - 0.9405 = 0.0495$$

Теперь мы готовы отвечать на заданные вопросы:

а)
$$\mathbb{P}(A) = \mathbb{P}(A \cap B^c) + \mathbb{P}(A \cap B) = 0.009 + 0.0495 = 0.0585$$

б) Прогностическая сила теста:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{0.009}{0.0585} \approx 0.154$$

Для того, чтобы повысить прогностическую силу теста, необходимо понизить $\mathbb{P}(A \cap B^c)$, а для этого необходимо повысить специфичность теста.

6. а) Должно выполняться условие нормировки:

$$\int_{-a}^{0} 1.5(x+a)^2 dx + \int_{0}^{a} 1.5(x-a)^2 dx = 1$$
$$0.5(x+a)^3\Big|_{-a}^{0} + 0.5(x-a)^3\Big|_{0}^{a} = 1$$
$$0.5a^3 + 0.5a^3 = 1 \Rightarrow a = 1$$

Теперь легко понять, как выглядит функция распределения (смотри определение функции распределения):

$$F(x)=\begin{cases}0, & x<1\\ 0.5(x+1)^3, & -1\leqslant x<0\\ 1+0.5(x-1)^3, & 0\leqslant x<1\\ 1, & x\geqslant 1\end{cases}$$

И с её помощью всё посчитать:

$$P\left(X\in\left[\frac{1}{2},2\right]\right)=F(2)-F\left(\frac{1}{2}\right)=1-1+0.5^4=0.5^4$$

$$\begin{aligned}\mathbb{E}(X)&=\int\limits_{-1}^{0}x\cdot 1.5(x+1)^2dx+\int\limits_{0}^{1}x\cdot 1.5(x-1)^2dx\\ &=1.5\int\limits_{-1}^{0}\left(x^3+2x^2+x\right)dx+1.5\int\limits_{0}^{1}\left(x^3-2x^2+x\right)dx\\ &=\frac{3}{8}x^4\Big|_{-1}^{0}+x^3\big|_{-1}^{0}+\frac{3}{4}x^2\Big|_{-1}^{0}+\frac{3}{8}x^4\Big|_{0}^{1}-x^3\big|_{0}^{1}+\frac{3}{4}x^2\Big|_{0}^{1}=-\frac{3}{8}+1-\frac{3}{4}+\frac{3}{8}-1+\frac{3}{4}=0\end{aligned}$$

А можно было заметить, что функция плотности — четная функция, поэтому сразу $\mathbb{E}(X)=0$.
Вычислим $\mathbb{E}\left(X^2\right)$:

$$\begin{aligned}\mathbb{E}(X^2)&=\int\limits_{-1}^{0}x^2\cdot 1.5(x+1)^2dx+\int\limits_{0}^{1}x^2\cdot 1.5(x-1)^2dx\\ &=1.5\int\limits_{-1}^{0}\left(x^4+2x^3+x^2\right)dx+1.5\int\limits_{0}^{1}\left(x^4-2x^3+x^2\right)dx\\ &=\frac{3}{10}x^5\Big|_{-1}^{0}+\frac{3}{4}x^4\Big|_{-1}^{0}+\frac{1}{2}x^3\Big|_{-1}^{0}+\frac{3}{10}x^5\Big|_{0}^{1}-\frac{3}{4}x^4\Big|_{0}^{1}+\frac{1}{2}x^3\Big|_{0}^{1}=\frac{1}{10}\end{aligned}$$

Тогда:

$$\mathrm{Var}(X)=\mathbb{E}(X^2)-(\mathbb{E}(X))^2=0.1$$

б) Верим, что график $F(x)$, выписанной выше, вы построить можете :)

7. Пусть $A=\{$«Лекция полезна»$\}$, $B=\{$«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 $X_A=\sum\limits_{i=1}^{30}X_i$. Рассмотрим отдельно X_i:

x	1	0
$\mathbb{P}(X=x)$	0.9	0.1

Вероятность 0.9 дана. Тогда:

$$
\begin{aligned}
\mathbb{E}(X_i) = 0.9 = \mathbb{E}(X_i^2) \Rightarrow \\
\operatorname{Var}(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = 0.9 - 0.9^2 = 0.09
\end{aligned}
$$

Значит

$$
\begin{aligned}
\mathbb{E}(X_A) &= \mathbb{E}\left(\sum_{i=1}^{30} X_i\right) = \sum_{i=1}^{30} \mathbb{E}(X_i) = 0.9 \cdot 30 = 27 \\
\operatorname{Var}(X_A) &= \operatorname{Var}\left(\sum_{i=1}^{30} X_i\right) = \sum_{i=1}^{30} \operatorname{Var}(X_i) = 0.09 \cdot 30 = 2.7
\end{aligned}
$$

Аналогично для числа интересных лекций можем получить:

$$
\begin{aligned}
\mathbb{E}(X_B) &= 0.7 \cdot 30 = 21 \\
\operatorname{Var}(X_A) &= 0.21 \cdot 30 = 6.3
\end{aligned}
$$

б) Так как интересность и полезность — независимые свойства лекций, то:

$$
\mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c) \cdot \mathbb{P}(B^c) = 0.3 \cdot 0.1 = 0.03,
$$

где A^c значит «не A». В свою очередь:

$$
\mathbb{P}(A \cup B) = \mathbb{P}(A \cap B^c) + \mathbb{P}(B \cap A^c) + \mathbb{P}(A \cap B) = 1 - \mathbb{P}(A^c) \cdot \mathbb{P}(B^c) = 0.97,
$$

где $(A \cup B)$ значит «A или B», а $(A \cap B)$ — «A и B». Аналогично, путем введения бинарной случайной величины можем получить:

$$
\begin{aligned}
\mathbb{E}(X_{A^c \cap B^c}) &= 0.03 \cdot 30 = 0.9 \\
\mathbb{E}(X_{A \cup B}) &= 0.97 \cdot 30 = 29.1
\end{aligned}
$$

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.

   - $\mathbb{E}(2X + Y - 4) = 2\,\mathbb{E}(X) + \mathbb{E}(Y) + \mathbb{E}(-4) = 2 + 2 - 4 = 0$
   - $\operatorname{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 5 - 1 = 4$
   - $\operatorname{Var}(Y) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = 8 - 4 = 4$
   - $\operatorname{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y) = -1 - 2 = -3$
   - $\operatorname{Corr}(X, Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(Y)}} = -\frac{3}{2 \cdot 2} = -0.75$
   - $\operatorname{Var}(X - Y - 1) = \operatorname{Var}(X) + \operatorname{Var}(Y) - 2\operatorname{Cov}(X, Y) = 4 + 4 - 2(-3) = 14$
   - $\operatorname{Var}(X + Y + 1) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X, Y) = 4 + 4 + 2(-3) = 2$
   -
   $$
   \begin{aligned}
   \operatorname{Cov}(X - Y - 1, X + Y + 1) &= \mathbb{E}((X - Y)(X + Y)) - \mathbb{E}(X - Y)\,\mathbb{E}(X + Y) \\
   &= \mathbb{E}\left(X^2 - Y^2\right) - (\mathbb{E}(X) - \mathbb{E}(Y))(\mathbb{E}(X) + \mathbb{E}(Y)) \\
   &= \mathbb{E}\left(X^2\right) - \mathbb{E}\left(Y^2\right) - \left((\mathbb{E}(X))^2 - (\mathbb{E}(Y))^2\right) \\
   &= \operatorname{Var}(X) - \operatorname{Var}(Y) = 0
   \end{aligned}
   $$
   - $\operatorname{Cov}(X - Y - 1, X + Y + 1) = 0 \Rightarrow \operatorname{Corr}(X - Y - 1, X + Y + 1) = 0$

9. Найдём частные распределения Y и Y^2:

	$X=1$	$X=2$	$\sum$
$Y=-1$	0.1	0.2	0.3
$Y=0$	0.2	0.3	0.5
$Y=1$	0	0.2	0.2
$\sum$	0.3	0.7	

y	-1	0	1
$\mathbb{P}(Y=y)$	0.3	0.5	0.2

Так как Y^2 может принимать только значения 0 или 1:

y^2	0	1
$\mathbb{P}(Y^2=y^2)$	0.5	0.5

А ковариация:

$$\begin{aligned}\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y) = ((-1)\cdot 1\cdot 0.1 + (-1)\cdot 2\cdot 0.2 + 1\cdot 2\cdot 0.2) \\ - (0.3\cdot 1 + 0.7\cdot 2)\cdot(0.3\cdot(-1) + 0.1\cdot 0.2) = 0.07\end{aligned}$$

Так как $\mathrm{Cov}(X,Y)\neq 0$ — величины зависимы

10. Бонусная 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

13.6. 2013-2014

1. Введём обозначения:

- $\mathbb{P}(\mathrm{B}|\mathrm{A}^c\cap \mathrm{M}^c) = 0.18$ — Вася пришёл, а девушки — нет
- $\mathbb{P}(\mathrm{B}|\mathrm{A}\cap \mathrm{M}) = 0.9$ — пришли и Вася, и девушки
- $\mathbb{P}(\mathrm{B}|\mathrm{A}^c\cap \mathrm{M}) = 0.54$ — Вася пришёл, если пришла только Маша
- $\mathbb{P}(\mathrm{B}|\mathrm{A}\cap \mathrm{M}^c) = 0.36$ — Вася пришёл, если пришла только Алёна
- $\mathbb{P}(\mathrm{M}) = 0.4$ — Маша пришла на лекцию
- $\mathbb{P}(\mathrm{A}) = 0.6$ — Алёна пришла на лекцию

а) Используя формулы Байеса и полной вероятности, получим:

$$\mathbb{P}(\mathrm{A}|\mathrm{B}) = \frac{\mathbb{P}(\mathrm{A}\cap\mathrm{B})}{\mathbb{P}(\mathrm{B})}$$

В числителе:

$$\begin{aligned}\mathbb{P}(\mathrm{B}|\mathrm{A})\cdot\mathbb{P}(\mathrm{A}) = P(\mathrm{B}|\mathrm{A}\cap\mathrm{M})\cdot\mathbb{P}(\mathrm{A})\cdot\mathbb{P}(\mathrm{M}) + \mathbb{P}(\mathrm{B}|\mathrm{A}\cap\mathrm{M}^c)\cdot\mathbb{P}(\mathrm{A}) = \cdot\,\mathbb{P}(\mathrm{M}^c) \\ = 0.9\cdot 0.4\cdot 0.6 + 0.36\cdot 0.6\cdot 0.6 = 0.3456\end{aligned}$$

А в знаменателе:

$$\begin{aligned}\mathbb{P}(\mathrm{B}) &= \mathbb{P}(\mathrm{B}|\mathrm{A}^c\cap\mathrm{M}^c)\cdot\mathbb{P}(\mathrm{A}^c\cap\mathrm{M}^c)+\mathbb{P}(\mathrm{B}|\mathrm{A}\cap\mathrm{M})\cdot\mathbb{P}(\mathrm{A}\cap\mathrm{M})+\mathbb{P}(\mathrm{B}|\mathrm{A}^c\cap\mathrm{M})\cdot\mathbb{P}(\mathrm{A}^c\cap\mathrm{M})\\ &\quad+\mathbb{P}(\mathrm{B}|\mathrm{A}\cap\mathrm{M}^c)\cdot\mathbb{P}(\mathrm{A}\cap\mathrm{M}^c)\\ &= 0.18\cdot 0.6\cdot 0.4+0.9\cdot 0.4\cdot 0.6+0.54\cdot 0.4\cdot 0.4+0.36\cdot 0.6\cdot 0.6=0.4752\end{aligned}$$

Ответ:

$$\mathbb{P}(\mathrm{A}|\mathrm{B})=\frac{\mathbb{P}(\mathrm{A}\cap\mathrm{B})}{\mathbb{P}(\mathrm{B})}=\frac{0.3456}{0.4752}=0.(72)$$

б) Необходимо найти

$$\mathbb{P}(\mathrm{M}|\mathrm{B})=\frac{\mathbb{P}(\mathrm{M}\cap\mathrm{B})}{\mathbb{P}(\mathrm{B})}$$

Знаменатель этой дроби посчитан в предыдущем пункте, посчитаем числитель:

$$\begin{aligned}\mathbb{P}(\mathrm{M}\cap\mathrm{B}) &= \mathbb{P}(\mathrm{B}|\mathrm{M})\cdot\mathbb{P}(\mathrm{M})\\ &= P(\mathrm{B}|\mathrm{M}\cap\mathrm{A})\cdot\mathbb{P}(\mathrm{A})\cdot\mathbb{P}(\mathrm{M})+\mathbb{P}(\mathrm{B}|\mathrm{A}^c\cap\mathrm{M})\cdot\mathbb{P}(\mathrm{A}^c)\cdot\mathbb{P}(\mathrm{M})\\ &= 0.9\cdot 0.4\cdot 0.6+0.54\cdot 0.4\cdot 0.4=0.3024\end{aligned}$$

Ответ:

$$\mathbb{P}(\mathrm{M}|\mathrm{B})=\frac{\mathbb{P}(\mathrm{M}\cap\mathrm{B})}{\mathbb{P}(\mathrm{B})}=\frac{0.3024}{0.4752}=0.(63)$$

Если Вася на лекции, вероятность застать на ней Алёну выше.

2. $\mathbb{P}(X=5)=C_{100}^5 0.002^5 0.998^{95}$,
   $\mathbb{E}(X)=0.2$,
   $\mathrm{Var}(X)=0.2\cdot 0.998$,
   наиболее вероятно событие $X=0$.

3. $c=1/2$, $\mathbb{P}(X\in[\ln 0.5,\ln 4])=5/8$, $\mathbb{E}(X)=0$, $\mathrm{Var}(X)=2$, $\mathbb{E}(X^{2k+1})=0$, $\mathbb{E}(X^{2k})=(2k)!$

4. а) $\mathbb{E}(Y-2X-3)=\mathbb{E}(Y)-2\,\mathbb{E}(X)-3=0$
   $\mathrm{Var}(Y-2X-3)=\mathrm{Var}(Y)+4\,\mathrm{Var}(X)-2\,\mathrm{Cov}(Y,2X)=16$
   $\mathrm{Cov}(X,Y)=\mathrm{Corr}(X,Y)\cdot\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}=6$

   б) $\mathrm{Corr}(Y-2X-3,X)=\frac{\mathrm{Cov}(Y,X)-2\,\mathrm{Var}(X)}{\sqrt{\mathrm{Var}(Y-2X-3)\cdot\mathrm{Var}(X)}}=-1.$

   в) Корреляция равна -1, значит, есть линейная взаимосвязь между переменными. Пусть $Y+aX=b$, тогда $\mathrm{Var}(Y+aX)=0$, $\mathbb{E}(Y)=-a+b=1$. Решая уравнения, находим, что $a=-2/3$, $b=1/3$.

5. а) Таблицы распределения имеют вид:

x	-1	0	1
$\mathbb{P}(X=x)$	0.3	0.3	0.4

y	-1	1
$\mathbb{P}(Y=y)$	0.5	0.5

   б)

$$\begin{aligned}\mathrm{Cov}(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\,\mathbb{E}(Y)&=(-1)\cdot(-1)\cdot 0.1+(-1)\cdot 0\cdot 0.2+\\ &+(-1)\cdot 1\cdot 0.2+1\cdot(-1)\cdot 0.2+1\cdot 0\cdot 0.1+1\cdot 1\cdot 0.1-0.1\cdot 0=-0.1\end{aligned}$$

   в) Да, поскольку если случайные величины независимы, то их коварицция равна нулю.

   г) Условное распределение:

x	-1	0	1
$\mathbb{P}(X=x\|Y=-1)$	0.2	0.4	0.4

   д) $\mathbb{E}(X|Y=-1)=-1\cdot 0.2+0\cdot 0.4+1\cdot 0.4=0.2$

13.7. 2012-2013

1. а) $\mathbb{P}(A) = 0.8 \cdot 0.3 + 0.7 \cdot 0.2 = 0.38$
   б) $\mathbb{P}(B) = 0.9$
   в) $\mathbb{P}(C|A) = \frac{0.3 \cdot 0.8}{0.38} = 0.632$
   г) $\mathbb{P}(C|D) = \frac{0.3 \cdot (0.9 \cdot 0.8 + 0.1 \cdot 0.2)}{0.9 \cdot 0.38 + 0.1 \cdot (1 - 0.38)} = 0.55$
2. Это была задачка-неберучка!
3. а) 1
   б) $\mathbb{E}(X) = 45/28 \approx 1.61$, $\mathbb{E}(X^2) = 93/35 \approx 2.66$, $\mathrm{Var}(X) = 291/3920 \approx 0.07$
   в) $37/56 \approx 0.66$
   г) $F(x) = \begin{cases} 0,\ x < 1 \\ \frac{x^3 - 1}{7},\ x \in [1;2] \\ 1,\ x > 1 \end{cases}$
4. а) $a = 0.1$
   б) $\mathbb{P}(X > -1) = 0.7$, $\mathbb{P}(X > Y) = 0.1$
   в) $\mathbb{E}(X) = -0.2$, $\mathbb{E}(X^2) = 2$
   г) $\mathrm{Corr}(X, Y) = 0.117$
5. а) Правильные: $\mathbb{E}(X) = 10$, $\mathrm{Var}(X) = 9$, неправильные: $\mathbb{E}(Y) = 9$, $\mathrm{Var}(Y) = 0.9$
   б) Наиболее вероятное число укусов равно математическому ожиданию
   в) Лучше идти к неправильным пчёлам, так как $\mathbb{P}(X \leqslant 2) < \mathbb{P}(Y \leqslant 2)$.

13.8. 2011-2012

1. $\mathbb{P}(A) = \frac{3 \cdot 2^3}{C_{10}^5} = \frac{2}{21} \approx 0.095$
2. $\mathbb{P}(A|B) = \frac{0.999 \cdot 0.01}{0.999 \cdot 0.01 + 0.001 \cdot 0.9} \approx 0.917$
3. а) $\mathbb{P}(A_1) = 0.079 + 0.209(0.209 + 0.337) + 0.375(0.375 + 0.337) + 0.337 \cdot 0.337 \approx 0.574$
   б) $\mathbb{P}(A_2) \approx 0.778$
4. а) $\mathbb{P}(X_v = 10) = 0.9^3 \cdot 0.3^3 \cdot 0.5^4$
   б) $\mathrm{Var}(X_m) = 0.9$, $\mathrm{Var}(X_d) = 2.1$, $\mathrm{Var}(X_v) = 0.27 + 0.63 + 1 = 1.9$
   $\mathrm{Corr}(X_v, X_d) = \frac{0.27}{\sqrt{1.9 \cdot 2.1}}$
   $\mathrm{Corr}(X_v, X_m) = \frac{0.63}{\sqrt{1.9 \cdot 0.9}}$
5. а) $c = 3$
   б) $F(x) = \begin{cases} 0, & x < 1 \\ 1 - x^{-3}, & x \geqslant 1 \end{cases}$
   в) $\mathbb{P}(0.5 < X < 1.5) = 1 - 1.5^{-3} = \frac{19}{27} \approx 0.70$
   г) Заметим, что $\mathbb{E}(X^a) = 3/(3 - a)$. Поэтому $\mathbb{E}(X) = 3/2$ и $\mathbb{E}(X^2) = 3$. Значит, $\mathrm{Var}(X) = 3/4$.
6. а) $F(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(1/X \leqslant y) = \mathbb{P}(X \geqslant 1/y) = \begin{cases} 0, y < 0 \\ y^3, y \in [0;1) \\ 1, y \geqslant 1 \end{cases}$

   $p(y) = \begin{cases} 3y^2, y \in [0;1] \\ 0, y \notin [0;1] \end{cases}$

б) $\mathbb{E}(X) = 3/2$, $\mathbb{E}(Y) = 3/4$, $\mathbb{E}(XY) = \mathbb{E}(1) = 1$, значит $\mathrm{Cov}(X, Y) = 1 - 9/8 = -1/8$
$\mathbb{E}(Y^2) = 3/5$, $\mathrm{Var}(Y) = 3/80$, $\mathrm{Corr}(X, Y) = -\sqrt{5}/3 \approx 0.75$

7. Функция плотности симметрична около нуля, поэтому: $\mathbb{E}\left((X - \mathbb{E}(X))^{2011}\right) = \mathbb{E}\left(X^{2011}\right) = 0$

8. а) $\mathbb{E}(U) = 5$, $\mathbb{E}(V) = -3$, $\mathrm{Var}(U) = 26$, $\mathrm{Var}(V) = 10$, $\mathrm{Cov}(U, V) = 0$

   б) Нет, даже нулевой ковариации недостаточно для того, чтобы говорить о независимости случайных величин.

9. а) $\mathbb{E}(X) = 80 \cdot 0.05 = 4$, $\mathrm{Var}(X) = 80 \cdot 0.05 \cdot 0.95 = 4 \cdot 0.95$

   б) $\mathbb{P}(X = 5) = C_{80}^5 0.05^5 0.95^{75}$

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

   г) $\triangle \leqslant \min\{p, np^2\} = \min\{0.05, 4 \cdot 0.05\} = 0.05$

10. а) $\mathbb{P}(X > 20) = \frac{80+80+50}{300} = 0.7$
$\mathbb{P}(X > 20 | X > Y) = \frac{80+50+50}{100+50+50} = 0.9$
$\mathbb{P}(X > Y | X > 20) = \frac{80+50+50}{80+80+50} = \frac{6}{7}$

    б) $\mathbb{E}(X) = 50$

    в) $F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{6} + \frac{4}{600}x, & x \in [0; 100) \\ 1, & x \geqslant 100 \end{cases}$ У функции два скачка высотой по $1/6$, в точках $x = 0$ и $x = 100$.
На остальных участках функция линейна.

    г) Нет, например, если $Y = 50$ мы можем быть уверены в том, что $X \notin [10; 90]$.

13.9. 2010-2011

1. p, всё равно

2. $\mathbb{P}(A) = 0.8$, $\mathbb{P}(B|A) = 0.84$

3. $a \geqslant 2\ln 10$

4. $\mathbb{E}(X) = 1.36$, $\mathrm{Var}(X) = 0.2$, $F(x) = \begin{cases} 0, \ x < 1 \\ 2/3, \ x \in [1; 2) \\ 35/36, \ x \in [2; 3) \\ 1, \ x \geqslant 3 \end{cases}$

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

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

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

13.10. 2008-2009 Демо-версия

1. $\mathbb{P}(A) = 2/15$

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

3.

4. $c = 0.1$, $\mathbb{P}(X > 30) = e^{-3}$, $\mathbb{P}(X > 45 \mid X > 15) = e^{-3}$, $\mathbb{E}(X) = 10$
5. $k^* = 2$, $\mathrm{Var}(X) = 1.875$, $\mathbb{E}(X) = 2.5$
6. $c = 0.2$, $\mathbb{P}(Y > -X) = 0.5$, $\mathbb{E}(XY) = 0$, $\mathrm{Corr}(X, Y) = -0.155$, $\mathbb{E}(Y|X > 0) = 1/4$
7. $\mathbb{E}(N) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \mathbb{E}(X_3) = 0.9 + 0.7 + 0.5 = 2.1$
8. $\mathbb{E}(X) \approx 1.7$, $\mathrm{Var}(X) \approx 1.08$

9-А.

9-Б. Рассмотрим совершенно конкурентный невольничий рынок начинающих певиц. Певицы в хорошем настроении продаются по V_1, в депрессии — по V_2. Получаем систему уравнений:

$$\begin{cases} V_1 = 0.75 + (0.5V_1 + 0.5V_2) \\ V_2 = \max_x \sqrt{x}V_1 + (1 - \sqrt{x})V_2 - x \end{cases}$$

Оптимизируем и получаем, $x^* = (V_1 - V_2)^2/4$. Из первого уравнения находим $(V_1 - V_2)/2 = 0.75$.

13.11. 2008-2009

1. а) $0.9 \cdot 0.9 \cdot 0.8$

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

   в) $0.9 + 0.9 + 0.8 = 2.6$

2. а) 9 (если взять 9 с вероятностью один)

   б) 4 (если взять 5 и 9 равновероятно)

3. а) $0.7 + 0.3 \cdot 0.5 = 0.85$

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

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

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

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

6. 0.2, $\frac{1}{e^4}$, $\frac{1}{e^4}$, 5

7. Любое разумное понимание «полугодовой» принимается. То есть подходят 182, 183, и если посчитаны только рабочие дни, и если взят пример марсианского теннисиста с указанием количествава дней в марсианском году и пр.

   И биномиальные и пуассоновские ответы принимаются.

   Для 182:

   а) $182 \cdot 0.00037 = 0.06734$

   б) $(1 - 0.00037)^1 82 \approx \exp(-0.06734)$

   в) $C_{182}^2 p^2 (1 - p)^{180} \approx 0.5 \exp(-0.06734) 0.06734^2$

8. а) $\mathbb{P}(Y < 2) = 1/4$

   б) два отрезка: на высоте 2/16 (от 0 до 4) и 1/16 (от 4 до 12)

   в) $\mathbb{E}(Y) = 5$, $\mathrm{Var}(Y) = 12.(3)$

   г) $\mathrm{Cov}(X, Y) = 3.(3)$

9-А. 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.

9-Б. X раскладывается в сумму индикаторов.

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

$\mathbb{E}(X) = 6 \cdot 9 \cdot 1/4 = 13.5$

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.

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

13.12. 2007-2008

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

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

   Итого: $p = p_1 \cdot p_2 = 2/15$

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

3. $p_a = \frac{1}{3}(0.9 + 0.5 + 0.3) = \frac{17}{30}$
   $p_b = \frac{1}{3}(0.9^2 + 0.5^2 + 0.3^2)/p_a = \frac{115}{170}$

4. 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) $1/\lambda = 10$

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

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

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

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

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

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

   $\mathbb{E}(N) = 28/15$, первая.

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

Получаем $p = n \cdot (0.5)^{n-1}$

9-Б. Будем считать координату одного за точку отсчета. На квадрате $[0; 1] \times [0; 1]$ нетрудно нарисовать нужное множество.

$p = 3/8$

13.13. 2006-2007

1. а) $\mathbb{P}(A) = 0.5, \mathbb{P}(B) = 1-\mathbb{P}(B^c) = 1-0.5^3 = \frac{7}{8}, \mathbb{P}(A\cap B) = 0.5\cdot\left(1-0.5^2\right) = \frac{3}{8}, \mathbb{P}(A\cap B) \neq \mathbb{P}(A)\,\mathbb{P}(B)$, события зависимы.

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

2. Пусть X — число правильных ответов.

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

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

   в) $\mathbb{E}(X) = 10\,\mathbb{E}(X_i) = \frac{10}{4}$
   $\mathrm{Var}(X) = 10\,\mathrm{Var}(X_i) = 10\cdot\frac{1}{4}\cdot\frac{3}{4}$

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

3. A — изделие браковано, B — изделие признано хорошим

   а) $\mathbb{P}(B) = 0.96\cdot 0.96 + 0.04\cdot 0.05$

   б) $\mathbb{P}(A|B) = \frac{0.04\cdot 0.05}{\mathbb{P}(B)}$

4. $\lambda = np = 4$
   $\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$
   $\mathbb{P}(X = 0) = e^{-4}$

5. а) Распределение имеет вид:

x	1	2	3	4
$\mathbb{P}(X = x)$	0.6	$(1-0.6)\cdot 0.5$	$(1-0.6)\cdot(1-0.5)\cdot 0.4$	$1-p_1-p_2-p_3$

   Упростим:

x	1	2	3	4
$\mathbb{P}(X = x)$	0.6	0.2	0.08	0.12

   б) $\mathbb{E}(X) = 1.7, \mathrm{Var}(X) \approx 1.08$

6. $\mathbb{P}(X \leqslant 0.5) = \frac{0.5}{2} = 0.25, \mathbb{E}(X) = \frac{0+2}{2} = 1$ (здравый смысл)
   $\mathrm{Var}(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2$
   $\mathbb{E}\left(X^2\right) = \int_0^2 t^2\cdot p(t)dt = \int_0^2 t^2\cdot 0.5dt = \frac{4}{3}$

7. $\mathbb{E}(X) = 10 = \frac{1}{\lambda}, \lambda = \frac{1}{10}, p(t) = \lambda e^{\lambda t}$ при $t > 0$
   $\mathbb{P}(X > 15) = \int_{15}^{\infty} p(t)dt = \ldots = e^{-\frac{3}{2}}$
   $\mathbb{P}(X > 25|X > 10) = \frac{\mathbb{P}(X>25)}{\mathbb{P}(X>10)} = \ldots = e^{-\frac{3}{2}}$

8. Функция распределения:
   $F_Y(t) = \mathbb{P}(Y \leqslant t) = \mathbb{P}(\max\{X_1, X_2\} \leqslant t) = \mathbb{P}(X_1 \leqslant t\cap X_2 \leqslant t) = \mathbb{P}(X_1 \leqslant t)\,\mathbb{P}(X_2 \leqslant t) = \frac{t+1}{2}\cdot t$ при $t \in [0;1]$.
   При $t > 1$ получаем, что $F_Y(t) = 1$ и при $t < 0$ получаем, что $F_Y(t) = 0$.
   $\mathbb{P}(\max\{X_1, X_2\} > 0.5) = 1 - \mathbb{P}(\max\{X_1, X_2\} \leqslant 0.5) = 1 - F(0.5) = \frac{5}{8}$

13.14. 2005-2006

1. $\mathbb{P}(A)=\frac{2\cdot 3!3!}{6!}=1/10$

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

3. $\mathbb{P}\left(X>\frac{1}{2}\right)=1/3, \mathbb{P}\left(X>\frac{1}{2}|Y<\frac{1}{2}\right)=1/4$

4. а)

x	2	3	4
$\mathbb{P}(X=x)$	1/6	1/3	1/2

   б) $\mathbb{E}(X)=3\frac{1}{3}$

5. $\mathbb{P}(A|B)=\frac{0.4\cdot 0.3}{0.4\cdot 0.3+0.6\cdot 0.7}$

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

   б) $\operatorname{Cov}(X,Y)=2.5, \operatorname{Cov}(6-X,3Y)=-7.5$

7. $x=-2, \operatorname{Var}(X)=3.1-0.49=2.61$

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

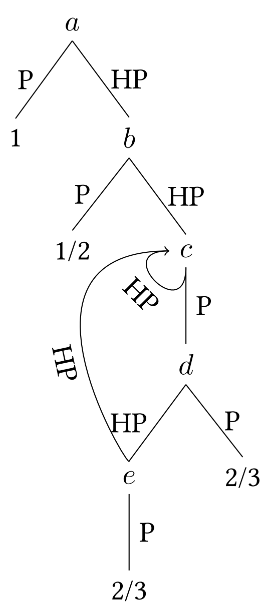
9. $\mathbb{P}(X=3\cap Y=5)=2/36, \mathbb{E}(X)=91/36, \operatorname{Var}(X)\approx 2.1$, заметим, что $X+Y=R_1+R_2$, поэтому $\mathbb{E}(X)+\mathbb{E}(Y)=7$, и $\mathbb{E}(3X-2Y)=3\,\mathbb{E}(X)-2\,\mathbb{E}(Y)=3\,\mathbb{E}(X)-2(7-\mathbb{E}(X))=5\,\mathbb{E}(X)-14$

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

14. Решения контрольной номер 1. ИП

14.1. 2018-2019

1. Воспользуемся методом первого шага. Дерево игры выглядит так (в вершинах — длительность сответствующей подыгры, а на концах ветвей — доля разбавленных пинт среди тех, что Джо помнит):



Если первая пинта разбавлена, то игра заканчивается (разбавлены 100% рома). Если первая пинта не разбавлена, то, если разбавлена вторая, игра заканчивается (50%). Если вторая не разбавлена, и третья тоже, то это равносильно тому, что не разбавлены только две (Джо не помнит больше трех). Если вторая не разбавлена, а третья разбавлена, возможны три случая: если четвёртая разбавлена, то игра заканчивается; если четвёртая не разбавлена, и пятая не разбавлена, то это эквивалентно тому, что не разбавлены только две; если четвёртая не разбавлена, а пятая разбавлена, то игра заканчивается. Из этих соображений получаем систему уравнений:

$$\begin{cases} a = \frac{1}{8} + \frac{7}{8}(b+1) \\ b = \frac{1}{8} + \frac{7}{8}(c+1) \\ c = \frac{1}{8}(d+1) + \frac{7}{8}(c+1) \\ d = \frac{1}{8} + \frac{7}{8}(e+1) \\ e = \frac{1}{8} + \frac{7}{8}(c+3) \end{cases}$$

Отсюда $a = 7514/192, b = 1046/24, c = 146/3, d = 122/3, e = 136/3$.

2. Пусть

$$Z_i = \begin{cases} 1, i\text{-я пинта разбавлена} \\ 0, \text{иначе} \end{cases}$$

Пусть с вероятностью x_n последовательность случайных величин (Z_n) длиной n не содержит двух единиц подряд и оканчивается нулём, а с вероятностью y_n — не содержит двух единиц подряд и оканчивается единицей.

Пусть на последнем месте (Z_n) стоит единица. Это может произойти с вероятностью $1/2$. При этом перед последней единицей может стоять любая последовательность длиной $n-1$, оканчивающаяся нулём. Значит, $y_n = 0.5x_{n-1}$.

Пусть на последнем месте (Z_n) стоит нуль. Это может произойти с вероятностью $1/2$. При этом перед последним нулём может стоять любая последовательность длиной $n-1$. Значит, $x_n = 0.5(x_{n-1} + y_{n-1})$.

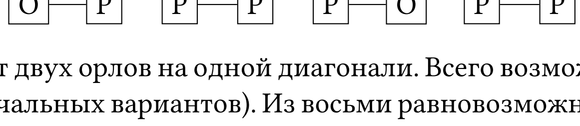
Таким образом, получена следующая разностная система:

$$\begin{cases} x_n = 0.5(x_{n-1} + y_{n-1}) \\ y_n = 0.5x_{n-1} \end{cases}$$

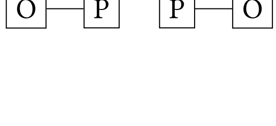
Более того, можно поставить задачу Коши. Так как (Z_2) может равновероятно иметь одну из 4 реализаций $(11, 10, 01, 00)$, из которых не содержат двух единиц подряд и оканчиваются на 0 две, то $x_2 = 1/2$. Аналогично, $y_2 = 1/4$. Решив задачу Коши, найдем формулы для x_n, y_n. Ответом будем число $x_{100} + y_{100}$.

3. а) Оптимальная стратегия для Али-Бабы состоит в чередовании открытия двух диагонально противоположных и двух соседних монет.

Изначально имеется только четыре варианта расположения монет:



На первом ходе Али делает двух орлов на одной диагонали. Всего возможно восемь случаев (по две диагонали в каждом из начальных вариантов). Из восьми равновозможных случаев два приводят к успеху. При этом, если успех не достигнут, можно получить в итоге только две комбинации орлов и решек (левую — в четырёх случаях, правую — в двух):



На втором ходе Али переворачивает две соседние монеты орлами вверх. При этом либо наступает успех (8 случаев из 24), либо неуспех, приводящий к единтсвенно возможной комбинации орлов и решек:

$$\begin{array}{ccc} \boxed{\text{О}} & - & \boxed{\text{О}} \\ | & & | \\ \boxed{\text{О}} & - & \boxed{\text{Р}} \end{array}$$

На третьей попытке открывается или диагональ с орлами, или диагональ с орлом и решкой. В о втором случае, очевидно, надо заменить решку на орла, и успех обеспечен. Но в первом случае оптимально перевернуть орла и сделать 2 решки. Это даст неуспех, но Али точно будет знать, что на четвёртой попытке монеты могут быть расположены только так:

$$\begin{array}{ccc} \boxed{\text{О}} & - & \boxed{\text{О}} \\ | & & | \\ \boxed{\text{Р}} & - & \boxed{\text{Р}} \end{array}$$

На четвёртом ходе в любом из четырёх вариантов (Али не знает, на каком ребре решки) нужно перевернуть обе открытые соседние монеты, тогда в двух случаях будет успех, а в двух других -- неуспех, при котором монеты могут быть расположены только так:

$$\begin{array}{ccc} \boxed{\text{О}} & - & \boxed{\text{Р}} \\ | & & | \\ \boxed{\text{Р}} & - & \boxed{\text{О}} \end{array}$$

Очевидно, что если на пятом ходе, открыв любую диагональ, Али перевернёт находящиеся на ней монеты, то он гарантирует себе успех.

б) Как видим, в худшем случае потребовалось 5 попыток.

Источник: Кордемский, Математика изучает случайности.

4. Возьмём колоду, добавим в неё джокера. Разложим в открытую по окружности. Джокер означает место разрыва окружности для её выкладывания в обычную колоду. Замечаем, что джокер и четыре дамы разбивают окружность на пять случайных отрезко. В силу симметрии ожидаемые длины этих отрезков равны и равны по $48/5$. Значит первая дама попадается в среднем на $48/5 + 1$ месте.

5. Обозначим искомую вероятность быть в Неведении в момент t значком p_t.

$$p_{t+\Delta} = p_t(1 - \lambda\Delta - o(\Delta))$$

Отсюда получаем, что

$$\frac{p_{t+\Delta} - p_t}{\Delta} = -\lambda p_t + \frac{o(\Delta)}{\Delta}$$

Устремляем Δ к нулю и решаем получающееся дифференциальное уравнение с начальным условием $p_0 = 1$, так как изначально Ученик находится в Неведении.

Итого:

$$p_t = \exp(-\lambda t)$$

6.

$$\mathbb{P}(Y_4 \in [t; t+\Delta]) = C_{10}^1 \cdot \Delta \cdot C_9^3 t^3 (1-t)^6 + o(\Delta)$$

Читаем вслух:

а) Одна из десяти величин должна попасть в отрезок $[t; t+\Delta]$;

б) Три из девяти оставшихся должны оказаться меньше t;

в) Шесть из девяти оставших должны оказаться больше t;

Вероятностью попадания двух и более величин в отрезок длины Δ пренебрегаем!

14.2. 2017-2018

1. Обозначим вероятность того, что сыр достанется Белому за b, если игра начинается с его броска.

   а) Получаем уравнение
   $$b = \frac{1}{12} + \frac{11}{12} \cdot \frac{11}{12} b$$
   Пояснение: Как Белый может победить в исходной игре? Либо сразу выкинуть 6 с вероятностью $1/12$. Либо передать ход Серому $(11/12)$, получить ход снова $(11/12)$ и выиграть в продолжении игры. Продолжение игры по сути совпадает с исходной игрой.

   б) Игра продолжается до тех пор, пока кто-то не выкинет «6». Для нахождения среднего количества бросков воспользуемся методом первого шага.

   Обозначим среднее количество бросков нашей игры за S. Когда Белый бросает кубик, с вероятностью $\frac{1}{12}$ игра закончится за один бросок, а с вероятностью $\frac{11}{12}$ игра продолжится и ход перейдёт к Серому. Но та игра, которая начнётся, когда бросать будет Серый, ничем не отличается от предыдущей, поэтому среднее количество бросков в ней будет равно S. Однако мы попадём в эту игру, «потратив» один бросок. Таким образом мы получаем:
   $$S = \frac{1}{12} \cdot 1 + \frac{11}{12}(S+1)$$
   Получается, что $S = 12$, значит игра длится в среднем 12 бросков.

2.

3. Для того, чтобы выжить, мышам нужно ещё до начала игры договориться о стратегии, которая позволит им с наибольшей вероятностью открыть нужные сундуки. Если хотя бы две мыши выберут одинаковый сундук, то их в любом случае съедят. Поэтому одной из оптимальных стратегий будет ещё до начала игры мышам договориться и назвать левый сундук золотым, сундук посередине серебряным, а правый — платиновым. Каждый мышонок должен открыть тот сундук, в честь которого назван необходимый ему металл. Если внутри он обнаруживает свой металл, то он выбирает этот сундук, если внутри находится не тот металл, мышонок открывает тот сундук, на который указывает лежащий внутри предмет.

   Например, первым заходит Микки Маус. Он открывает золотой (левый) ящик. Если внутри лежит золото, то он выходит из комнаты. Если же внутри лежит, например, серебро, то Микки Маус открывает сундук посередине. Путём перебора можно посчитать, что в 4 случаях из 6 мыши смогут найти нужный металл, поэтому вероятность выигрыша при данной стратегии равна $\frac{2}{3}$.

4. а) Обозначим буквой X количество детей в случайной семье. Можно просуммировать ряд $\mathbb{E}(X) = 1 \cdot 0.5 + 2 \cdot 0.5^2 + 3 \cdot 0.5^3 + \ldots$. А можно воспользоваться методом первого шага и заметить, что либо первой рождается девочка и мышиная семья больше не заводит мышат, либо рождается мальчик и семья ситуация превратилась в первоначальную плюс один ребёнок.
   $$\mathbb{E}(X) = 0.5 \cdot 1 + 0.5 \cdot (1 + \mathbb{E}(X))$$
   Находим, $\mathbb{E}(X) = 2$.

   б) В мышиной стране много семей. Занумеруем семьи. Обозначим буквой X_i число детей в i-ой семье. Тогда доля D — это суммарное количество мальчиков делить на суммарное количество детей:

   Получаем
   $$\frac{(X_1 - 1) + (X_2 - 2) + \ldots + (X_n - 1)}{X_1 + X_2 + \ldots + X_n} = 1 - \frac{n}{\sum_{i=1}^{n} X_i}$$
   Основной смысл математического ожидания $\mathbb{E}(X_i)$ — это среднее выборочное при большом количестве повторений опыта.

   Следовательно, $\frac{\sum_i X_i}{n} \to \mathbb{E}(X_i) = 2$. И стало быть, при $n \to \infty$ получаем $D \to 1 - 1/2 = 1/2$.

Интуитивный аргумент немного опасен, но всё же приведём его: вероятности рождения мальчиков и девочек равны, поэтому и доля будет половина. Опасность аргумента состоит в том, что семьи могли бы в альтернативном условии рожать мышат по принципу: пока доля мальчиков в стране не превысит 50%.

в) Можно просто аккуратно суммировать ряд:

$$\mathbb{E}(Y) = 0.5 \cdot \frac{0}{1} + 0.5^2 \cdot \frac{1}{2} + 0.5^3 \cdot \frac{2}{3} + \ldots$$

Здесь мудрый студент может вспомнить ряд Тейлора для логарифма:

$$\ln(1 + r) = r - r^2/2 + r^3/3 - r^4/4 + \ldots$$

Стало быть,

$$\ln(1 - 0.5) = -0.5 - 0.5^2/2 - 0.5^3/3 - 0.5^4/4$$

И исходное ожидание равно:

$$\mathbb{E}(Y) = 1 + \ln 0.5 = 1 - \ln 2 \approx 0.31$$

г) Неожиданно лёгкий вопрос, $\mathbb{E}(X - 1) = 2 - 1 = 1$.

5. Благосостояние кота Василия, положившего один гурд на вклад, равно $m_t = 1 \cdot e^{rt}$, где r — процентная ставка, а t — прошедшее время.

Откуда взялась экспонента?

Допустим время дискретно, а процентная ставка равна 5% в год. Тогда сумма вклада каждый год домножается на 1.05 и эволюционирует по формуле:

$$m_t = 1 \cdot 1.05^t$$

Но любое число можно представить в виде экспоненты, $1.05 = e^{\ln 1.05}$.

Поэтому

$$m_t = e^{\ln 1.05 \cdot t} = e^{rt}$$

Момент закрытия вклада T равномерно распределён на отрезке от 0 до a, поэтому сумма, которую получит Василий, представима в виде $Z = e^Y$, где $Y \sim U[0; ra]$. По условию, a очень велико, поэтому ra тоже очень велико.

Вероятность того, что первая цифра будет равна 1, равна вероятности того, что доход Василия будет лежать в пределах от 1 до 2 гурдов, плюс вероятность того, что он лежит в пределах от 10 до 20 гурдов и т.д. Таким образом, можно представить эту вероятность, как:

$$\mathbb{P}(N = 1) = \mathbb{P}(e^Y \in [1; 2)) + \mathbb{P}(e^Y \in [10; 20)) + \ldots$$

Это выражение можно преобразовать таким образом:

$$\mathbb{P}(N = 1) = \mathbb{P}(Y \in [\ln 1; \ln 2)) + \mathbb{P}(Y \in [\ln 10; \ln 20)) + \ldots$$

Так как Y — равномерно распределённая величина, то $\mathbb{P}(Y \in [\ln 1; \ln 2)) = \frac{\ln 2 - \ln 1}{ra}$. Для последующих слагаемых вероятность рассчитывается таким же образом. Воспользовавшись свойством логарифма, можно заметить, что $\frac{\ln 20 - \ln 10}{ra} = \frac{\ln 2}{ra}$. Поэтому вероятность того, что на первом месте суммы вклада стоит единица, равна $n \cdot \frac{\ln 2}{ra}$, где n — количество слагаемых. Путём аналогичных рассуждений получаем, что вероятность того, что на первом месте стоит двойка, равна $n \cdot \frac{\ln 3 - \ln 2}{ra}$. Из-за того, что a велико, можно считать, что число слагаемых одинаково.

На первом месте обязательно будет находиться какая-то цифра, поэтому сумма вероятностей будет равна 1. Получаем:

$$\frac{n}{ra}\left(\ln\frac{2}{1}+\ln\frac{3}{2}+\ldots+\ln\frac{10}{9}\right)=1$$

Таким образом $\frac{n}{ra}=\frac{1}{\ln 10}$. Получается, что вероятность того, что на первом месте стоит единица, равна:

$$\mathbb{P}(N=1)=\frac{\ln 2}{\ln 10}$$

Закон распределения первой цифры выводится сложением соответствующих вероятностей.

14.3. 2016-2017

1. а) Для удобства занумеруем макаронины и выделим у каждой левый и правый конец. Взяли правый конец первой макаронины и подвязали случайной. Взяли свободный конец только что подвязанной макаронины и подвязали случайно. И так далее.

$$\frac{2n-2}{2n-1}\cdot\frac{2n-4}{2n-3}\cdot\ldots\cdot\frac{2}{3}\cdot 1$$

   б) Допустим, что при n макаронинах в среднем образуется e_n колец. После первого соединения задача сводится к меньшему числу макаронин, важно только учесть, образовалось ли кольцо при первом соединении:

$$e_n=\frac{1}{2n-1}(e_{n-1}+1)+\frac{2n-2}{2n-1}e_{n-1}=e_{n-1}+\frac{1}{2n-1}$$

   в) Количество коротких колец можно разбить в сумму, $X=Z_1+\ldots+Z_n$. Вероятность завязывания конкретной макаронины в кольцо равна $1/(2n-1)$: «левый конец» надо привазять именно к «правому». Значит, $\mathbb{E}(X)=n/(2n-1)$.

2. а) Рассмотрим обратную ситуацию: на планете есть точка, из которой связаться хотя бы с одним пепелацем нельзя. Такое возможно, если все, кроме одного, сели в одну полуокружность.

$$\mathbb{P}(\text{есть точка без связи})=n\cdot\left(\frac{1}{2}\right)^{n-1}\Rightarrow\mathbb{P}(\text{из любой точки есть связь})=1-n\cdot\left(\frac{1}{2}\right)^{n-1}$$

   б) Зафиксируем координату посадки первого пепелаца и возьмём её за точку отсчёта. Изобразим на плоскости возможные значения центральных углов между первым пепелацем и оставшимися и закрасим нужные участки. Получим $3/8$.

   в) Зафиксируем координату посадки первого пепелаца. Обозначим центральный угол между первым и вторым пепелацами α. Функция плотности имеет вид: $p(\alpha)=\frac{\sin(\alpha)}{2}$
   Итог: $\int_0^{\pi/2}p(\alpha)\frac{\alpha+\pi}{2\pi}d\alpha+\int_{\pi/2}^{\pi}p(\alpha)\frac{\pi-\alpha}{2\pi}d\alpha=\frac{\pi+2}{4\pi}$

3. Вспомним для начала, что площадь круга равна πr^2, а площадь сферы равна $4\pi r^2$. Составим из маленьких треугольничков многогранник очень похожий на сферу с единичным радиусом. Площадь этого многогранника будет примерно равна 4π. Проекция многогранника представляет собой примерно круг единичного радиуса. Проекция имеет два слоя. С учётом обоих слоёв площадь проекции равна 2π. Значит отношение площади проекции к площади многогранника равно $1/2$.

   От взаимного расположения треугольничков в пространстве ожидаемая площадь проекции не зависит в силу аддитивности математического ожидания.

   Ответ: 21 см^2.

4. а) Мысленно отметим на окружности три точки: места ударов Брюса Ли и точку, где схватился Чак Норрис. Можно считать, что эти три точки равномерно и независимо распределены по окружности. Следовательно, среднее расстояние между соседними точками равно $1/3$. Чак Норрис берёт два кусочка, слева и справа от своей точки. Значит ему в среднем достаётся $2/3$ окружности.

б) Объявим точку, где схватился Чак Норрис нулём. Координаты двух ударов изобразим на плоскости. Закрашиваем подходящий участок. Вероятность того, что кусок Брюса Ли длиннее, равна $1/4$.

5. Рассмотрим совершенно конкурентный невольничий рынок начинающих певиц. Певицы в хорошем настроении продаются по V_1, в депрессии — по V_2. Получаем систему уравнений:

$$\begin{cases} V_1 = 0.75 + (0.5V_1 + 0.5V_2) \\ V_2 = \max_x \sqrt{x}V_1 + (1-\sqrt{x})V_2 - x \end{cases}$$

Оптимизируем и получаем, $x^* = (V_1 - V_2)^2/4$. Из первого уравнения находим $(V_1 - V_2)/2 = 0.75$.

6. Да. Например, такая. До общения с Джульеттой подкидывать монетку до выпадения первого орла и запомнить число потребовавшихся подбрасываний. Пусть это будет число X. Открыть равновероятно левую или правую руку Джульетты. Если открытое число больше X, то сказать, что оно большее, иначе сказать, что меньшее.

7. Если γ — вероятность самостоятельного познания Истины, а α — передачи Истины отдельно в каждую из сторон, то
$$p = \gamma + (1-\gamma)p\alpha.$$
То есть $p = \gamma/(1 - \alpha(1-\gamma))$.

Для решения второго пункта наложим на Абу Али Хусейн ибн Абдуллах ибн аль-Хасан ибн Али ибн Сина обет молчания. Это не повлияет на вероятность постижения им Истины, однако превратит задачу в две уже решённых :) Получаем
$$q = \gamma + (1-\gamma)(2p\alpha - p^2\alpha^2)$$

14.4. 2015-2016

Индивидуальный тур

1. Сократ, эта — Η, η, дзета — Ζ, ζ, вега — нет, шо — ϸ, τ — тау, θ — тета, ξ — кси. Греческая буква шо, ϸ, была введена Александром Македонским и ныне вышла из употребления. По крайней мере, в греческом :) Заглавная примерно такая же, только её utf-код 03f7 не поддерживается шрифтом Linux Libertine.

2. Да. События независимы в совокупности, если для любого поднабора событий A_1, ..., A_k выполняется равенство $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_k) = \mathbb{P}(A_1) \cdot \ldots \cdot \mathbb{P}(A_k)$. Нет.

3. $1/4$, $2/3$, 15

4. 0, $0.8^5 \cdot 0.2$, $1 - 0.8^6$

5. 0.8^{10}, $C_{10}^3 0.2^3 0.8^7$, 2

6. $1/2$, $3/16$, $3/8$

7. а)
$$F_X(x) = \begin{cases} 0, \ x < 0 \\ x^3, \ x \in [0;1] \\ 1, \ x > 1 \end{cases}$$

б) $1/8$

в) $2^{-1/3}$

г) $56/63$

д)

$$F_Y(y) = \begin{cases} 0,\ y < 0 \\ 1 - 1/y^3,\ y > 0 \end{cases}$$

е)

$$f_Y(y) = \begin{cases} 0,\ y < 0 \\ 3y^{-4},\ y > 0 \end{cases}$$

Командный тур

1. Если отрублено 10 щупалец, значит либо был один удар породивший два новых щупальца, либо было два удара, породивших по одному новому, а все остальные удары не порождали новых щупалец.

   Искомая вероятность равна: $8 \cdot 0.5^9 \cdot 0.25^1 + C_8^2 0.5^8 0.25^2$.

   Вероятность вечного боя равна нулю. Достаточно доказать, что с вероятностью один за конечное время побеждается одноногий Кракен. А эта вероятность удовлетворяет уравнению: $p = \frac{1}{4}p + \frac{1}{4}p^2 + \frac{1}{2}1$. Единственный осмысленный корень у этого уравнения — 1.

   Замечаем, что на победу над k-шупальцевым Кракеном уходим в k раз больше ударов в среднем чем на победу на 1-щупальцевым. Отсюда:

   $$e_1 = 1 + 0.5 \cdot 0 + 0.25 \cdot e_1 + 0.25 \cdot 2e_1$$

   Решаем, получаем $e_1 = 4$ и $e_8 = 32$

2. Либо первая пинта разбавлена, либо первая неразбавлена, а вторая разбавлена, то есть

   $$0.25 + 0.75 \cdot 0.25 = 0.4375$$

   Рисуем граф.

   Составляем систему (индекс — количество выпитых неразбавленных пинт):

   $$\begin{cases} e_0 = \frac{1}{4} + \frac{3}{16}2 + \frac{9}{16}(2 + e_2) \\ e_2 = 1 + \frac{3}{4}e_2 + \frac{1}{4}e_0 \end{cases}$$

   Находим $e_0 = 64/7 \approx 9$

3. Для $t > 0$:

   $$\mathbb{P}(-\ln X \leqslant t) = \mathbb{P}(\ln X > -t) = \mathbb{P}(X > e^{-t}) = 1 - e^{-t}$$

   Итого,

   $$F_{-\ln X}(t) = \begin{cases} 0,\ t < 0 \\ 1 - e^{-t},\ t \geqslant 0 \end{cases}$$

   Из геометрических соображений легко найти $\mathbb{P}(XY < a)$ для $a \in (0; 1)$:

   $$\mathbb{P}(XY < a) = a + \int_a^1 \frac{a}{x}\, dx = a - a\ln a$$

   Переходим ко второму пункту, для $t > 0$:

   $$\mathbb{P}(-(\ln X + \ln Y) < t) = \mathbb{P}(XY > e^{-t}) = 1 - e^{-t} - te^{-t}$$

   Итого:

   $$F_{-\ln X - \ln Y}(t) = \begin{cases} 0,\ t < 0 \\ 1 - e^{-t} - te^{-t},\ t \geqslant 0 \end{cases}$$

После дифференциирования получаем функцию плотности для $S = -\ln X - \ln Y$:

$$f_S(s) = \begin{cases} 0, \; s < 0 \\ se^{-s}, \; s \geqslant 0 \end{cases}$$

Приближаемся к финальной вероятности:

$$\mathbb{P}(ZS > t) = \int_t^\infty \int_{t/s}^1 se^{-s}\, dz\, ds = \int_t^\infty (s-t)\cdot e^{-s}\, ds = \ldots = e^{-t}$$

Сравниваем результат с первым пунктом и приходим к выводу, что величина $(XY)^Z$ имеет равномерное распределение на $[0;1]$.

4. Если нанято n пиратов, то вероятность, того, что в конкретный день все работают равна $(364/365)^n$. Следовательно, ожидаемое количество праздничных дней равно $365(1-(364/365)^n)$.

   Решаем уравнение:
   $$1-(364/365)^n = 100/365$$
   Получаем:
   $$n = \frac{\ln 265 - \ln 365}{\ln 364 - \ln 365} \approx 117$$
   Ожидаемое количество рабочих пирато-дней равно: $365n(364/365)^n$.

   Получаем:
   $$n^* = 1/(\ln 365 - \ln 364) \approx 364$$

5. а) $\mathbb{P}(R_{100}) = 1/100$ (максимум из 100 величин должен плюхнуться на сотое место), $\mathbb{P}(B_{100}) = 1/3$ (максимум из трёх величин должен плюхнуться на второе место)

   б) $\mathbb{E}(X) = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{100} \approx \ln 100 \approx 4.6$. Т.к. $X = X_1 + X_2 + \ldots + X_{100}$ и $\mathbb{E}(X_i) = 1/i$.

   в) $\mathbb{P}(R_{99}|R_{100}) = 1/99$, $\mathbb{P}(R_{100}|B_{100}) = 3/101$

   Для проверки: $\mathbb{P}(R_{99} \cap R_{100}) = 98!/100!$ (100! — всего перестановок, 98! — первые 98 можно переставлять свободно, а в конце должны идти второй наибольшое и наибольшее). $\mathbb{P}(R_{100} \cap B_{100}) = 1/101$ (максимум из 101 числа плюхнется на 100ое место).

6. Если все пираты открывают первый и второй сундуки, то вероятность выигрыша равна нулю.

   Оптимальная стратегия (одна из). Три пирата заранее договариваются, о названиях сундуков. Они называют эти три сундука (ещё до игры) «рубиновым», «пиастровым» и «золотым». Генри Рубинов должен начать с открытия рубинового сундука, Френсис Пиастров — с пиастрового, Эдвард Золотов — с золотого. Далее каждый пират должен открыть тот сундук, на который указывает предмет, лежащий в первом открытом им сундуке. Например, если Генри Рубинов, открыв сначала рубиновый сундук обнаруживает там пиастры, он должен открывать пиастровый сундук.

   Вероятность победы при такой стратегии легко находится перебором 6 возможных вариантов и равна...Та-дам!!! 2/3.

14.5. 2014-2015

Часть 1

Не претендуя на единственность, решения претендуют на правильность!

1. а) $\mathbb{P}(\text{все арбузы спелые}) = 0.9^2 \cdot 0.7 = 0.567$

   б) A = {случайно выбранный арбуз — от тёти Маши}; B = {случайно выбранный арбуз оказался спелым}. Формула условной вероятности:

   $$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)} = \frac{2/3\cdot 0.9}{2/3\cdot 0.9 + 1/3 \cdot 0.7} = \frac{18}{25}$$

в) A = {второй и третий съеденные арбузы — от тёти Маши}; B = {все три арбуза — спелые}. Дает ли нам что-то о принадлежности арбузов к тёте Маше или тёте Оле то, что все арбузы — спелые? События независимы!

$$\mathbb{P}(A|B) = \mathbb{P}(A) = \frac{1}{3}$$

2. а) $\mathbb{E}(X) = \sum \mathbb{P}(X_i)X_i = 1.9$
$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 0 \cdot 0.1 + 1 \cdot 0.3 + 4 \cdot 0.2 + 9 \cdot 0.4 - 1.9^2 = 1.09$

б) 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
$\mathbb{P}(Y = y)$	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:

а) Да это же интеграл от функции плотности на всей числовой прямой! Ответ: единица!

б)

$$\mathbb{E}(X) = \int_0^2 xf(x)dx = \int_0^2 \frac{3}{8}x^3dx = \left.\frac{3}{32}x^4\right|_0^2 = \frac{3}{2}$$

$$\mathbb{E}\left(X^2\right) = \int_0^2 x^2f(x)dx = \int_0^2 \frac{3}{8}x^4dx = \left.\frac{3}{40}x^5\right|_0^2 = \frac{12}{5}$$

Формула дисперсии:

$$\text{Var}(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 = \frac{12}{5} - \frac{9}{4} = \frac{3}{20}$$

в)

$$\mathbb{P}(X > 1.5) = \int_{1.5}^2 f(x)dx = \int_{1.5}^2 \frac{3}{8}x^2dx = \left.\frac{1}{8}x^3\right|_{1.5}^2 = \frac{37}{64}$$

Вычислим вероятность условия:

$$\mathbb{P}(X > 1) = \int_1^2 f(x)dx = \int_1^2 \frac{3}{8}x^2dx = \left.\frac{1}{8}x^3\right|_1^2 = \frac{7}{8}$$

$$\mathbb{P}(X > 1.5|X > 1) = \frac{\mathbb{P}(X > 1.5)}{\mathbb{P}(X > 1)} = \frac{37/64}{7/8} = \frac{37}{56}$$

г) Должно выполниться следующее соотношение:

$$\int_{-\infty}^{+\infty} cxf(x)dx = 1$$

Применительно к нашей задаче:

$$\frac{3c}{8}\int_0^2 x^3dx = \left.\frac{3c}{32}x^4\right|_0^2 = \frac{3c}{2} = 1 \Rightarrow c = \frac{2}{3}$$

4. You have to learn the rules of the game. And then you have to play better than anyone else. (А. Эйнштейн)

   а)
   $$\operatorname{Var}(Z) = \mathbb{E}\left(Z^2\right) - (\mathbb{E}(Z))^2 = 15 - 9 = 6$$
   $$\operatorname{Var}(4 - 3Z) = 9\operatorname{Var}(Z) = 54$$
   $$\mathbb{E}\left(5 + 3Z - Z^2\right) = 5 + 3 \cdot (-3) - 15 = -19$$

   б)
   $$\operatorname{Var}(X \pm Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) \pm 2 \cdot \operatorname{Cov}(X, Y)$$

   Отсюда получаем:
   $$\operatorname{Var}(X + Y) - \operatorname{Var}(X - Y) = 4\operatorname{Cov}(X, Y) \Rightarrow \operatorname{Cov}(X, Y) = 2.5$$
   $$\operatorname{Cov}(6 - X, 3Y) = -3 \cdot 2.5 = -7.5$$

   в)
   $$\operatorname{Cov}(X, Y) = 2.5 \neq 0$$

   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 попасть[1].

   а) Пусть X — число опечаток на 13 странице.
   $$\mathbb{P}(X \geqslant 2) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1)$$
   $\mathbb{P}(X = 0) = \left(\frac{499}{500}\right)^{400}$ — каждая из 400 опечаток не доложна попасть на 13 страницу.
   $\mathbb{P}(X = 1) = 400 \cdot \frac{1}{500} \cdot \left(\frac{499}{500}\right)^{399}$ — ровно одна опечатка (а есть 400 вариантов) должна попасть на 13 страницу, а остальные — мимо. Соответственно:
   $$\mathbb{P}(X \geqslant 2) = 1 - \left(\frac{499}{500}\right)^{400} - 400 \cdot \frac{1}{500} \cdot \left(\frac{499}{500}\right)^{399} \approx 0.1911357$$
   Это если считать в явном виде. А если пользоваться приближением Пуассона:
   $$p(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$
   неплохо бы вспомнить, что параметр λ это математическое ожидание X, поэтому расчеты здесь пока оставим до лучших времен.

   б) 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 $X = \sum\limits_{i=1}^{400} X_i$. Рассмотрим отдельно X_i:

[1] Ну очень самостоятельные!

x	1	0
$\mathbb{P}(X=x)$	$\frac{1}{500}$	$\frac{499}{500}$

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

$$\mathbb{E}(X_i) = \frac{1}{500} = \mathbb{E}(X_i^2) \Rightarrow$$

$$\operatorname{Var}(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = \frac{1}{500} - \left(\frac{1}{500}\right)^2 = \frac{499}{500^2}$$

Значит

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{400} X_i\right) = \sum_{i=1}^{400} \mathbb{E}(X_i) = \frac{400}{500} = 0.8$$

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

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

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

So close!

Осталось найти наиболее вероятное число опечаток на 13 странице:

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

Очевидно, что эта функция убывает по k, ведь с ростом k:
$k!$ растет, а 0.8^k убывает. Значит наиболее вероятное число ошибок — $X = 0$

в) 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 $A = \{$«Лекция полезна»$\}$, $B = \{$«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 $X_A = \sum\limits_{i=1}^{30} X_i$. Рассмотрим отдельно X_i:

x	1	0
$\mathbb{P}(X=x)$	0.9	0.1

Вероятность 0.9 дана. Тогда:

$$\mathbb{E}(X_i) = 0.9 = \mathbb{E}(X_i^2) \Rightarrow$$

$$\mathrm{Var}(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = 0.9 - 0.9^2 = 0.09$$

Значит,

$$\mathbb{E}(X_A) = \mathbb{E}\left(\sum_{i=1}^{30} X_i\right) = \sum_{i=1}^{30} \mathbb{E}(X_i) = 0.9 \cdot 30 = 27$$

$$\mathrm{Var}(X_A) = \mathrm{Var}\left(\sum_{i=1}^{30} X_i\right) = \sum_{i=1}^{30} \mathrm{Var}(X_i) = 0.09 \cdot 30 = 2.7$$

Аналогично для числа интересных лекций можем получить:

$$\mathbb{E}(X_B) = 0.7 \cdot 30 = 21$$

$$\mathrm{Var}(X_B) = 0.21 \cdot 30 = 6.3$$

б) Так как интересность и полезность — независимые свойства лекций, то:
$\mathbb{P}(\bar{A} \cap \bar{B}) = \mathbb{P}(\bar{A}) \cdot \mathbb{P}(\bar{B}) = 0.3 \cdot 0.1 = 0.03$, где $\bar{A}$ значит «не A». В свою очередь:
$\mathbb{P}(A \cup B) = \mathbb{P}(A \cap \bar{B}) + \mathbb{P}(B \cap \bar{A}) + \mathbb{P}(A \cap B) = 1 - \mathbb{P}(\bar{A}) \cdot \mathbb{P}(\bar{B}) = 0.97$, где $(A \cup B)$ значит «A или 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. Будем пользоваться свойствами функций распределения и плотности. Для начала:

$$\lim_{x \to +\infty} F(x) = 1, \qquad \lim_{x \to -\infty} F(x) = 0,$$

$$\lim_{x \to +\infty} \left(\frac{ae^x}{1+e^x} + b\right) = a + b := 1$$

$$\lim_{x \to -\infty} \left(\frac{ae^x}{1+e^x} + b\right) = b := 0$$

Откуда сразу получаем

$$a = 1, b = 0 \Rightarrow F(x) = \frac{e^x}{1+e^x}$$

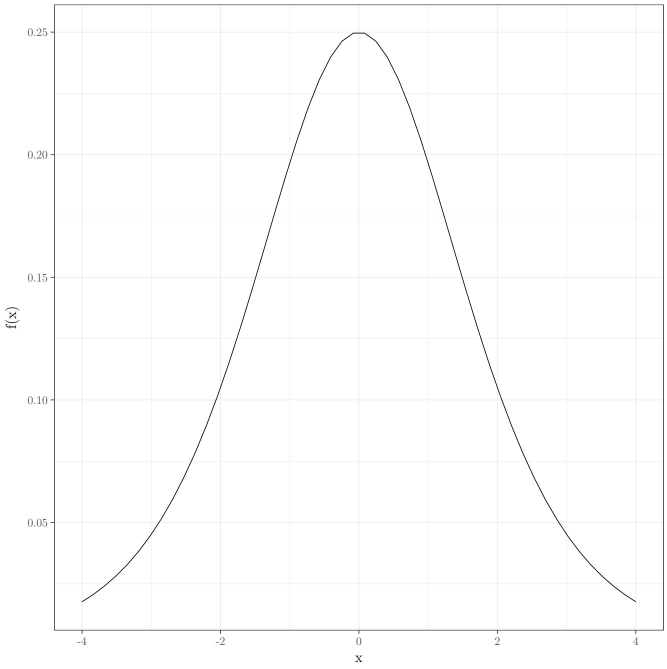
Для дальнейших развлечений нам понадобится функция плотности:

$$f(x) = F'(x) = \frac{e^x}{(1+e^x)^2}$$

Заметим, что она симметрична относительно нуля:

$$f(-x) = \frac{\frac{1}{e^x}}{\left(1 + \frac{1}{e^x}\right)^2} = \frac{e^x}{(1+e^x)^2} = f(x)$$

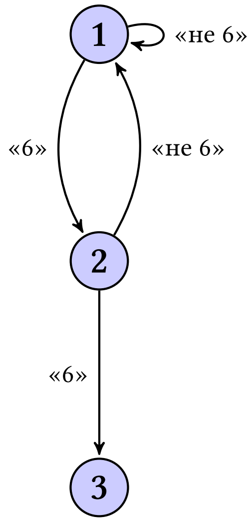
Из того этого следует, что **математическое ожидание, а так же мода и медиана равны нулю**. Более того, так как функция плотности симметрична относительно **нулевого** математического ожидания, **центральный и начальный моменты третьего порядка равны между собой и равны нулю.** Можно было выписать интегралы для математического ожидания и третьего начального момента и сослаться на нечетность функции.



Часть 2

— Это невозможно!
— Нет. Это необходимо.
© Interstellar

1. Алгоритм решения: рисуешь дерево $\to$ 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!

а) $\mathbb{P}(N = 1) = 0$ — невозможно за ход закончить игру.
$\mathbb{P}(N = 2) = \frac{1}{36}$
$\mathbb{P}(N = 3) = \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{5}{216}$

б) 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 математических 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 $\mathbb{E}\left(N^2\right)$:

$$\mathbb{E}(N^2) = \frac{5}{6}\cdot \mathbb{E}\left((N+1)^2\right) + \frac{1}{6}\cdot\frac{5}{6}\cdot \mathbb{E}\left((N+2)^2\right) + \frac{1}{6}\cdot\frac{1}{6}\cdot 2^2$$

Учитывая, что $\mathbb{E}(N) = 42$, получим: $\mathbb{E}(N^2) = 3414$.

в) Veni, vidi, vici

x_n	6
$\mathbb{P}(X_n = x_n)$	1

2. а) $\mathbb{P}(V = 1) = 1/30$, так как 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:

$$\mathbb{P}(M = 1) = \frac{26}{30}\cdot\frac{2}{29} + \frac{4}{30}\cdot\frac{1}{29} = \frac{56}{30\cdot 29} = \frac{28}{435}$$

$M = V$ 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):

$$\mathbb{P}(M = V) = \frac{1}{2}\cdot\frac{1}{29} = \frac{1}{58}$$

б)

$$\mathbb{E}(V) = \frac{0+1+\ldots+29}{30} = \frac{30\cdot 14 + 15}{30} = 14.5$$

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
$\mathbb{P}(X_i = x_i)$	$1/3$	$2/3$

Легко считается: $\mathbb{E}(X_i) = 1/3$, $\mathbb{E}(X_i^2) = 1/3$, $\mathrm{Var}(X_i) = 1/3 - 1/9 = 2/9$. Ясно, что $M = \sum_1^{28} X_i$. Тогда учитывая независимость X_i:

$$\mathbb{E}(M) = \frac{28}{3}$$

$$\mathrm{Var}(M) = \frac{56}{9}$$

3. Биномиальное распределение — *À l'abordage!*.

   Задача интерпретируется так: последний ход — это когда мы обратились к коробку, в котором нет спичек (то есть к одному коробку нужно обратиться $n + 1$ раз).

   а) Пусть ξ — это случайная величина, обозначающая число оставшихся спичек в непустом коробке перед последним ходом.

   Если $0 < k \leqslant n$, 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 $n - k$ спичек, то есть спички брались $2n - k$ раз. Таким образом, перед последним ходом произошло n успехов и $n - k$ неудач.

   $$\mathbb{P}(\xi = k) = C_{2n-k}^{n-k} = \left(\frac{1}{2}\right)^{n-k} \left(\frac{1}{2}\right)^{n} = C_{2n-k}^{n-k} \left(\frac{1}{2}\right)^{2n-k}$$

   Теперь нужно учесть, что на последнем ходе был выбран именно пустой коробок. Вероятность этого события — $1/2$, значит, искомая вероятность равна:

   $$\mathbb{P}(\text{в одном коробке осталось k спичек}) = C_{2n-k}^{n-k} \left(\frac{1}{2}\right)^{2n-k} \cdot \frac{1}{2} = C_{2n-k}^{n-k} \left(\frac{1}{2}\right)^{2n-k+1}$$

   б) Среднее спичек в другом коробке:

   $$\mathbb{E}(X) = \sum_{k=1}^{n} k \cdot C_{2n-k}^{n-k} \left(\frac{1}{2}\right)^{2n-k+1}$$

4. Для того чтобы количество упаковок, которые необходимо купить, равнялось 50, нужно чтобы ни одну из наклеек Покупатель не встретил дважды, поэтому:

   $$\mathbb{P}(X = 50) = 1 \cdot \frac{49}{50} \cdot \frac{48}{50} \cdot \dots \cdot \frac{1}{50} = \frac{49!}{50^{49}} \approx 3.4 \cdot 10^{-21}$$

   Dum spiro, spero![2]

   Теперь введём 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

[2]Надежда умирает последней!

необходимую наклейку и «прогулка» закончится. Введём случайную величину X_q равную количеству покупок в течение шага номер q. Тогда $X = \sum_{q=0}^{49} X_q$. Найдём 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 ещё нет, а $\frac{q}{n}$, соответственно — вероятность повториться. Вопрос теперь в том, как посчитать сумму:

$$\mathbb{E}(X_q) = \frac{n-q}{n}\left(1 + \frac{q}{n} \cdot 2 + \left(\frac{q}{n}\right)^2 \cdot 3 + \dots\right) = \frac{n-q}{n} \cdot \sum_{k=0}^{\infty} \left(\frac{q}{n}\right)^k (k+1)$$

Можем выписать в столбик несколько первых членов вышестоящей суммы:

$$\begin{array}{l} 1 \\ \left(\frac{q}{n}\right)^1 + \left(\frac{q}{n}\right)^1 \\ \left(\frac{q}{n}\right)^2 + \left(\frac{q}{n}\right)^2 + \left(\frac{q}{n}\right)^2 \\ \left(\frac{q}{n}\right)^3 + \left(\frac{q}{n}\right)^3 + \left(\frac{q}{n}\right)^3 + \left(\frac{q}{n}\right)^3 \\ \dots\dots\dots\dots\dots\dots\dots\dots \end{array}$$

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:

$$\sum_{k=0}^{\infty} \left(\frac{q}{n}\right)^k (k+1) = \frac{1}{1-\frac{q}{n}} + \frac{\frac{q}{n}}{1-\frac{q}{n}} + \frac{\left(\frac{q}{n}\right)^2}{1-\frac{q}{n}} + \frac{\left(\frac{q}{n}\right)^3}{1-\frac{q}{n}} + \dots =$$
$$= \frac{1}{1-\frac{q}{n}}\left(1 + \frac{q}{n} + \left(\frac{q}{n}\right)^2 + \left(\frac{q}{n}\right)^3 + \dots\right) = \frac{n}{n-q} \cdot \frac{n}{n-q} = \left(\frac{n}{n-q}\right)^2$$

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!

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{q=0}^{49} X_q\right) = \sum_{q=0}^{49} \mathbb{E}(X_q) = \frac{50}{50-0} + \frac{50}{50-1} + \dots + \frac{50}{50-49}$$
$$= 50\left(\frac{1}{50} + \frac{1}{49} + \dots + 1\right) \approx 50 \int_1^{50} \frac{1}{x} \mathbf{d}x = 50 \ln(50) \approx 195.5$$

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 $\mathbb{E}(X) = 1/p$ 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 $1/n$ она угадает свое место, значит, для каждого входящего его место будет свободно и он туда сядет.

   Ответ: $1/n$

   б) Будем искать вероятность того, что последний человек не сядет на свое место.

   Пусть $A_i = \{$Старушка села на место i-го$\}$, $B_{(i,j)} = \{i$-ый пассажир сел на место j-ого$\}$

$$P(n\text{-ый не сядет на свое место}) = \mathbb{P}(A_n) + P(A_{n-1})P(B_{(n-1,n)}) + \\ + P(A_{n-2})(P(B_{(n-2,n)}) + P(B_{(n-2,n-1)})P(B_{(n-1,n)}))) + \ldots$$

   Можем заметить, что:

   - $\mathbb{P}(A_i) = P(A_j) = \frac{1}{n} \ \forall\, i, j$
   - $\mathbb{P}(B_{(n-1,n)}) = \frac{1}{2}$, потому что $n-1$-ый выбирает из двух оставшихся мест
   - $\mathbb{P}(B_{(n-2,n)}) + \mathbb{P}(B_{(n-2,n-1)}]P[B_{(n-1,n)}) = \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{2}$
   -
$$\mathbb{P}(B_{(n-3,n)}) + \mathbb{P}(B_{(n-3,n-2)})(\mathbb{P}(B_{(n-2,n)}) + \mathbb{P}(B_{(n-2,n-1)})\,\mathbb{P}(B_{(n-1,n)})) \\ + \mathbb{P}(B_{(n-3,n-1)})\,\mathbb{P}(B_{(n-1,n)}) = \frac{1}{4} + \frac{1}{4}\left(\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2}\right) + \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{2}$$
   - 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:

$$\mathbb{P}[n\text{-ый не сядет на свое место}] = \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{2} + \frac{1}{n} \cdot \frac{1}{2} + \cdots = \frac{1}{n} + \frac{1}{2n}(n-2) = \frac{1}{2}$$

   Значит,

$$\mathbb{P}[n\text{-ый сядет на свое место}] = 1 - \frac{1}{2} = \frac{1}{2}$$

   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, чьё 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, $\ldots$, Y_n — сел ли n-ый на место (индикатор равен единице, если сел).

   Стало быть $\mathbb{E}(Y) = \mathbb{E}(Y_1) + \mathbb{E}(Y_2) + \ldots + \mathbb{E}(Y_n)$. $\mathbb{E}(Y_n) = \frac{1}{2}$.

   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 $3, 4, \ldots n$ 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 $2, 3, \ldots, (n-1)$, то есть $\frac{2}{3} \cdot \frac{n-2}{n+1}$ 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

$$\frac{1}{2}+\frac{2}{3}+\frac{3}{4}+\cdots+\frac{n-1}{n}+\frac{1}{n}$$

14.6. 2013-2014

Часть 1

1. а) Запишем все благоприятные исходы в таблицу:

Исход	Вероятность
ООО	$p^2 \cdot \frac{1}{2}$
ООН	$p^2 \cdot \frac{1}{2}$
ОНО	$p(1-p)\frac{1}{2}$
НОО	$(1-p)p\frac{1}{2}$

Нас устраивает любой из этих исходов, так что

$$\mathbb{P}(\text{жюри одобрит конкурсанта}) = p^2 \cdot \frac{1}{2} \cdot 2 + p(1-p)\frac{1}{2} \cdot 2 = p$$

б) Исходя из результата предыдущего пункта, получаем, что конкурсанту безразлично.

2. Введём обозначения:

- $\mathbb{P}(\mathrm{B}|\mathrm{A}^c \cap \mathrm{M}^c) = p$ — Вася пришёл, а девушки — нет.
- $\mathbb{P}(\mathrm{B}|\mathrm{A} \cap \mathrm{M}) = 5p$ — пришли и Вася, и девушки.
- $\mathbb{P}(\mathrm{B}|\mathrm{A}^c \cap \mathrm{M}) = 3p$ — Вася пришёл, если пришла только Маша.
- $\mathbb{P}(\mathrm{B}|\mathrm{A} \cap \mathrm{M}^c) = 2p$ — Вася пришёл, если пришла только Алёна.
- $\mathbb{P}(\mathrm{M}) = 0.6$ — Маша пришла на лекцию.
- $\mathbb{P}(\mathrm{A}) = 0.3$ — Алёна пришла на лекцию.

а) По теореме умножения:

$$\mathbb{P}(\mathrm{A}|\mathrm{B}) = \frac{\mathbb{P}(\mathrm{A} \cap \mathrm{B})}{\mathbb{P}(\mathrm{B})}$$

Выпишем числитель:

$$\begin{aligned}\mathbb{P}(\mathrm{B}|\mathrm{A}) \cdot \mathbb{P}(\mathrm{A}) &= P(\mathrm{B}|\mathrm{A} \cap \mathrm{M}) \cdot \mathbb{P}(\mathrm{A}) \cdot \mathbb{P}(\mathrm{M}) + \mathbb{P}(\mathrm{B}|\mathrm{A} \cap \mathrm{M}^c) \cdot \mathbb{P}(\mathrm{A}) = \cdot \mathbb{P}(\mathrm{M}^c) \\ &= 5p \cdot 0.6 \cdot 0.3 \cdot 0.6 + 2p \cdot 0.36 \cdot 0.4 \cdot 0.3 = 1.14p\end{aligned}$$

И знаменатель:

$$\begin{aligned}\mathbb{P}(\mathrm{B}|\mathrm{A}^c \cap \mathrm{M}^c) \cdot \mathbb{P}(\mathrm{A}^c \cap \mathrm{M}^c) + \mathbb{P}(\mathrm{B}|\mathrm{A} \cap \mathrm{M}) \cdot \mathbb{P}(\mathrm{A} \cap \mathrm{M}) + \mathbb{P}(\mathrm{B}|\mathrm{A}^c \cap \mathrm{M}) \cdot \mathbb{P}(\mathrm{A}^c \cap \mathrm{M}) + \\ + \mathbb{P}(\mathrm{B}|\mathrm{A} \cap \mathrm{M}^c) \cdot \mathbb{P}(\mathrm{A} \cap \mathrm{M}^c) = p \cdot 0.4 \cdot 0.7 + 5p \cdot 0.6 \cdot 0.3 \\ + 3p \cdot 0.6 \cdot 0.7 + 2p \cdot 0.4 \cdot 0.3 = 2.68p\end{aligned}$$

Ответ:

$$\mathbb{P}(\mathrm{A}|\mathrm{B}) = \frac{\mathbb{P}(\mathrm{A} \cap \mathrm{B})}{\mathbb{P}(\mathrm{B})} = \frac{1.14p}{2.68p} \approx 0.43$$

б) Теперь необходимо найти
$$\mathbb{P}(\mathrm{M}|\mathrm{B}) = \frac{\mathbb{P}(\mathrm{M}\cap \mathrm{B})}{\mathbb{P}(\mathrm{B})}$$
Знаменатель этой дроби посчитан в предыдущем пункте, посчитаем числитель:
$$\mathbb{P}(\mathrm{M}\cap \mathrm{B}) = \mathbb{P}(\mathrm{B}|\mathrm{M})\cdot \mathbb{P}(\mathrm{M}) = P(\mathrm{B}|\mathrm{M}\cap \mathrm{A})\cdot \mathbb{P}(\mathrm{A})\cdot \mathbb{P}(\mathrm{M}) \\ + \mathbb{P}(\mathrm{B}|\mathrm{A}^c\cap \mathrm{M})\cdot \mathbb{P}(\mathrm{A}^c)\cdot \mathbb{P}(\mathrm{M}) = 5p\cdot 0.6\cdot 0.3 + 3p\cdot 0.6\cdot 0.7 = 2.16p$$
Ответ:
$$\mathbb{P}(\mathrm{M}|\mathrm{B}) = \frac{\mathbb{P}(\mathrm{M}\cap \mathrm{B})}{\mathbb{P}(\mathrm{B})} = \frac{2.16p}{2.68p} \approx 0.8$$
Если Вася на лекции, вероятность застать на ней Машу выше.

в) $\mathbb{P}(\mathrm{B}) = 0.5, \mathbb{P}(\mathrm{B}) = 2.68p \Rightarrow p \approx 0.19$

3. а) Перед нами биномиальное распределение! Пусть X — случайная величина, число туристов, которые не выехали за границу. Тогда:
$$\mathbb{P}(X = 5) = C_{100}^5\cdot 0.002^5\cdot 0.998^{95}$$
б)
- $\mathbb{E}(X) = 2$
- $\mathrm{Var}(X) = 0.2\cdot 0.998$
- Наиболее вероятное число невыехавших — 0.

в) Пусть случайная величина S_i обозначает страховые выплаты, которые может получить один турист. Она может принимать значение 0, если турист выехал за границу и не обратился за медицинской помощью, 2000, когда он не выехал и 3000, когда турист выехал за границу и обратился за медицинской помощью. Тогда S_i распределена следующим образом:

s_i	0	2000	3000
$\mathbb{P}(S_i = s_i)$	$0.998\cdot 0.99$	0.002	$0.998\cdot 0.01$

- $\mathbb{E}(S_i) = 2000\cdot 0.002 + 3000\cdot 0.998\cdot 0.01 = 33.94 \Rightarrow \mathbb{E}(S) = 3394$
- $\mathbb{E}(S_i^2) = 2000^2\cdot 0.002 + 3000^2\cdot 0.998\cdot 0.01 = 97820$
- $\mathrm{Var}(S_i) = 97820 - 33.94^2 = 96668 \Rightarrow \mathrm{Var}(S) = 9666800$

г)

4. а) $\mathbb{E}(Y - 2X - 3) = \mathbb{E}(Y) - 2\,\mathbb{E}(X) - 3 = 0$
$\mathrm{Var}(Y - 2X - 3) = \mathrm{Var}(Y) + 4\,\mathrm{Var}(X) - 2\,\mathrm{Cov}(Y, 2X) = 16$
$\mathrm{Cov}(X, Y) = \mathrm{Corr}(X, Y)\cdot \sqrt{\mathrm{Var}(X)\cdot \mathrm{Var}(Y)} = 6$

б) $\mathrm{Corr}(Y - 2X - 3, X) = \frac{\mathrm{Cov}(Y,X) - 2\,\mathrm{Var}(X)}{\sqrt{\mathrm{Var}(Y-2X-3)\cdot \mathrm{Var}(X)}} = -1$, или проще: можно было заметить, что случайные величины линейно связаны.

в) Корреляция равна 1, значит, есть линейная взаимосвязь между переменными. Пусть $Y + aX = b$, тогда $\mathrm{Var}(Y + aX) = 0, \mathbb{E}(Y) = -a + b = 1$. Решая уравнения, находим, что $a = -2/3, b = 1/3$.

5. а) Частные распределения:

x	-1	0	1
$\mathbb{P}(X = x)$	0.3	0.3	0.4

y	-1	1
$\mathbb{P}(Y = y)$	0.5	0.5

б)
$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y) = (-1)\cdot(-1)\cdot 0.1 + (-1)\cdot 1\cdot 0.2 + 1\cdot(-1)\cdot 0.2 \\ + 1\cdot 1\cdot 0.2 - ((-1)\cdot 0.3 + 1\cdot 0.4)(-1\cdot 0.5 + 1\cdot 0.5) = -0.1$$

в) Да, так как $\mathrm{Cov}(X, Y) \neq 0$

г) Необходимо минимизировать дисперсию дохода:

$$\operatorname{Var}(\alpha X + (1-\alpha)Y) \to \min_{\alpha}$$

$$\begin{aligned}\operatorname{Var}(\alpha X + (1-\alpha)Y) &= \alpha^2 \operatorname{Var}(X) + (1-\alpha)^2 \operatorname{Var}(Y) + 2\alpha(1-\alpha)\operatorname{Cov}(X,Y) \\ &= 0.69\alpha^2 + (1-\alpha)^2 - 0.2\alpha(1-\alpha) \to \min_{\alpha}\end{aligned}$$

$$\frac{\partial \operatorname{Var}(\alpha X + (1-\alpha)Y)}{\partial \alpha} = 2\cdot 0.69\alpha - 2(1-\alpha) - 0.2 + 0.4\alpha = 0 \Rightarrow \alpha \approx 0.58$$

д) Условное распределение:

x	-1	0	1
$\mathbb{P}(X = x \mid Y = -1)$	0.2	0.4	0.4

е) $\mathbb{E}(X \mid Y = -1) = -1\cdot 0.2 + 1\cdot 0.4 = 0.2$

15. Решения контрольной работы 2

15.1. 2017-2018

7. а) Всем хватит места, если число явившихся на рейс пассажиров (X) не превысит 300, то есть нужно найти $\mathbb{P}(X \leqslant 300)$. Найдём матожидание и дисперсию случайной величины X:

$$\begin{aligned}\mathbb{E}(X) &= np = 330\cdot 0.9 = 297 \\ \operatorname{Var}(X) &= np(1-p) = 330\cdot 0.9\cdot 0.1 = 29.7\end{aligned}$$

Теперь посчитаем нужную вероятность:

$$\mathbb{P}(X \leqslant 300) = \mathbb{P}\left(\frac{X-297}{\sqrt{29.7}} \leqslant \frac{300-297}{\sqrt{29.7}}\right) = \mathbb{P}(\mathcal{N}(0,1) \leqslant 0.55) \approx 0.709$$

б) Вероятность переполнения не должна превышать 0.1:

$$\begin{aligned}&\mathbb{P}(X > 300) < 0.1 \\ &\mathbb{P}\left(\frac{X - 0.9\cdot n}{\sqrt{0.9\cdot 0.1\cdot n}} > \frac{300 - 0.9\cdot n}{\sqrt{0.9\cdot 0.1\cdot n}}\right) < 0.1 \\ &\frac{300 - 0.9\cdot n}{\sqrt{0.9\cdot 0.1\cdot n}} > 1.28 \\ &300 - 0.9n > 1.28\cdot 0.3\sqrt{n} \\ &n < 325.6\end{aligned}$$

8. а) Выпишем случайную величину X_i — цену акции после i-ого дня:

$$X_i = \begin{cases} 1.01, & p = 0.7 \\ 0.99, & p = 0.2999 \\ 0, & p = 0.0001 \end{cases}$$

Нужно посчитать ожидание цены акциии после 20 дней:

$$\mathbb{E}(X_1\cdot \ldots \cdot X_{20}) \overset{\text{незав-ть}}{=} \mathbb{E}(X_1)\cdot \ldots \cdot \mathbb{E}(X_{20}) = 1.004^{20} \approx 1.083$$

б) По ЗБЧ:

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mathbb{E}(X_i) = 1.004$$

в) Аналогично пункту (а):

$$\mathbb{E}(X_1 \cdot \ldots \cdot X_n) = (\mathbb{E}(X_1))^n = 1.004^n$$

И понятно, что $1.004^n \to_{n\to\infty} +\infty$.

г)

$$\mathbb{P}(\text{разорения}) = 1 - \mathbb{P}(X_1 > 0, \ldots, X_n > 0) = 1 - \prod_{i=1}^{n} \mathbb{P}(X_i > 0)$$
$$= 1 - (1 - 0.0001)^n \to_{n\to\infty} 1$$

15.2. 2016-2017

1. а) $\mathbb{E}(2\xi - \eta + 1) = 2\,\mathbb{E}(\xi) - \mathbb{E}(\eta) + 1 = 2 \cdot 1 - (-2) + 1 = 5$
   $\operatorname{Cov}(\xi, \eta) = \mathbb{E}(\xi\eta) - \mathbb{E}(\xi)\,\mathbb{E}(\eta) = -1 - \cdot 1 \cdot (-2) = 1$
   $\operatorname{Corr}(\xi, \eta) = \frac{\operatorname{Cov}(\xi,\eta)}{\sqrt{\operatorname{Var}(X)\cdot \operatorname{Var}(Y)}} = \frac{1}{\sqrt{1\cdot(8-(-2)^2)}} = \frac{1}{2}$
   $\operatorname{Var}(2\xi - \eta + 1) = 4\operatorname{Var}(\xi) + \operatorname{Var}(\eta) - 2\operatorname{Cov}(2\xi, \eta) = 4 \cdot 1 + 4 - 4 \cdot 1 = 4$

   б) $\operatorname{Cov}(\xi + \eta, \xi + 1) = \operatorname{Cov}(\xi) + \operatorname{Cov}(\xi, 1) + \operatorname{Cov}(\eta, \xi) + \operatorname{Cov}(\eta, 1) = 1 + 1 = 2$
   $\operatorname{Corr}(\xi + \eta, \xi + 1) = \frac{\operatorname{Cov}(\xi+\eta,\xi+1)}{\sqrt{\operatorname{Var}(\xi+\eta)\cdot\operatorname{Var}(\xi+1)}} = \frac{2}{\sqrt{(1+4+2\cdot 1)\cdot 1}} = \frac{2}{\sqrt{7}}$
   $\operatorname{Corr}(\xi + \eta - 24, 365 - \xi - \eta) = -1$
   $\operatorname{Cov}(2016 \cdot \xi, 2017) = 0$

2. а) Частные распределения:

x	-1	0	2
$\mathbb{P}(\xi = x)$	0.3	0.4	0.3

y	-1	1
$\mathbb{P}(\eta = y)$	0.5	0.5

   $\mathbb{E}(\xi) = -1 \cdot 0.3 + 0 \cdot 0.4 + 2 \cdot 0.3 = 0.3$
   $\mathbb{E}(\xi^2) = (-1)^2 \cdot 0.3 + 0^2 \cdot 0.4 + 2^2 \cdot 0.3 = 1.5$
   $\operatorname{Var}(\xi) = \mathbb{E}(\xi^2) - (\mathbb{E}(\xi))^2 = 1.5 - 0.3^2 = 1.41$
   $\mathbb{E}(\eta) = -1 \cdot 0.5 + 1 \cdot 0.5 = 0$
   $\mathbb{E}(\eta^2) = (-1)^2 \cdot 0.5 + 1^2 \cdot 0.5 = 1$
   $\operatorname{Var}(\eta) = \mathbb{E}(\eta^2) - (\mathbb{E}(\eta))^2 = 1 - 0^2 = 1$

   б)

xy	-2	-1	0	1	2
$\mathbb{P}(\xi \cdot \eta = xy)$	0.2	0.2	0.4	0.1	0.1

   $\mathbb{E}(\xi \cdot \eta) = (-2) \cdot 0.2 + (-1) \cdot 0.2 + 0 \cdot 0.4 + 1 \cdot 0.1 + 2 \cdot 0.1 = -0.3$
   $\operatorname{Cov}(\xi, \eta) = \mathbb{E}(\xi \cdot \eta) - \mathbb{E}(\xi) \cdot \mathbb{E}(\eta) = -0.3 - 0.3 \cdot 0 = -0.3$

   в) Пусть случайная величина X принимает значения $a_1, \ldots, a_m$, случайная веилчина Y принимает значения $b_1, \ldots, b_n$. Тогда случйаня величина X и Y называются независимыми, если $\forall i = 1, \ldots, m \quad \forall j = 1, \ldots, n : \mathbb{P}(X = a_i \cap Y = b_j) = P(X = a_i) \cdot P(Y = b_j)$

   г) Заметим, что $\mathbb{P}(\xi = -1 \cap \eta = -1) = 0.1$, $\mathbb{P}(\xi = -1) = 0.3$ и $\mathbb{P}(\eta = -1) = 0.5$.
   Тогда поскольку $\mathbb{P}(\xi = -1 \cap \eta = -1) \neq \mathbb{P}(\xi = -1) \cdot \mathbb{P}(\eta = -1)$, случайные величины ξ и η не являются независимыми.

д) $\mathbb{P}(\xi=-1\cap\eta=1)=\frac{\mathbb{P}(\xi=-1\cap\eta=1)}{\mathbb{P}(\eta=1)}=\frac{0.2}{0.5}=\frac{2}{5}$
$\mathbb{P}(\xi=0\cap\eta=1)=\frac{\mathbb{P}(\xi=0\cap\eta=1)}{\mathbb{P}(\eta=1)}=\frac{0.2}{0.5}=\frac{2}{5}$
$\mathbb{P}(\xi=2\cap\eta=1)=\frac{\mathbb{P}(\xi=2\cap\eta=1)}{\mathbb{P}(\eta=1)}=\frac{0.1}{0.5}=\frac{1}{5}$
Следовательно, условное распределение случайной величины ξ при условии $\{\eta=1\}$ может быть описано следующей таблицей:

x	-1	0	2
$\mathbb{P}(\xi=x)$	$2/5$	$2/5$	$1/5$

е) $\mathbb{E}(\xi\mid\eta=1)=-1\cdot\frac{2}{5}+0\cdot\frac{2}{5}+2\cdot\frac{1}{5}=0$

ж) $\mathbb{E}(\pi)=\mathbb{E}(0.5\xi+0.5\eta)=0.5\,\mathbb{E}(\xi)+0.5\,\mathbb{E}(\eta)=0.15$

$$\begin{aligned}\operatorname{Var}(\pi)&=\operatorname{Var}(0.5\xi+0.5\eta)=\operatorname{Var}(0.5\xi)+\operatorname{Var}(0.5\eta)+2\operatorname{Cov}(0.5\xi,0.5\eta)\\&=0.25\operatorname{Var}(\xi)+0.25\operatorname{Var}(\eta)+2\cdot0.5\cdot0.5\operatorname{Cov}(\xi,\eta)\\&=0.25\cdot1.41+0.25\cdot1+2\cdot0.5\cdot0.5\cdot(-0.3)=0.4525\end{aligned}$$

з)

$$\begin{aligned}\operatorname{Var}(\pi(\alpha))&=\operatorname{Var}(\alpha\xi+(1-\alpha)\eta)=\alpha^2\operatorname{Var}(\xi)+(1-\alpha)^2\operatorname{Var}(\eta)\\&+2\alpha(1-\alpha)\operatorname{Cov}(\xi,\eta)=1.41\cdot\alpha^2+1\cdot(1-\alpha)^2+2\alpha(1-\alpha)\cdot(-0.3)\\&=1.41\cdot\alpha^2+(1-\alpha)^2-0.6\cdot(\alpha-\alpha^2)\to\min_\alpha\\\frac{\partial}{\partial\alpha}\operatorname{Var}(\pi(\alpha))&=2\cdot1.41\cdot\alpha-2(1-\alpha)-0.6\cdot(1-2\alpha)\\&=2.82\cdot\alpha-2+2\alpha-0.6+1.2\cdot\alpha=6.02\cdot\alpha-2.6=0\\\alpha&=\frac{2.6}{6.02}=0.4319\end{aligned}$$

3. а) Для любой неотрицательной случайной величины X и любого числа $\lambda>0$ справедлива оценка: $\mathbb{P}(X>\lambda)\leqslant\frac{\mathbb{E}(X)}{\lambda}$
Пусть случайная величина ξ_i означает число посетителей сайта за i-ый день. По условию, $\xi_i\sim Pois(\lambda=250)$. Известно, что если $\xi\sim Pois(\lambda)$, то $\mathbb{E}(\xi)=\operatorname{Var}(\xi)=\lambda$.
Имеем:

$$\mathbb{P}(\xi_i>500)\leqslant\mathbb{P}(\xi_i\geqslant500)\leqslant\frac{\mathbb{E}(\xi_i)}{500}=\frac{250}{500}=\frac{1}{2}$$

б) Для любой случайной величины X с конечным $\mathbb{E}(X)$ и любого положительного числа $\epsilon>0$ имеет место неравенство: $\mathbb{P}(X-\mathbb{E}(X)\geqslant\epsilon)\leqslant\frac{\operatorname{Var}(X)}{\epsilon^2}$
Обозначим $\bar{\xi}_n:=\frac{1}{n}(\xi_1+\ldots+\xi_n)$ – среднее число посетителей сайта за n дней. Тогда

$$\mathbb{E}\left(\bar{\xi}_n\right)=\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n\xi_i\right)=\frac{1}{n}\sum_{i=1}^n\mathbb{E}(\xi_i)=\frac{1}{n}\cdot n\cdot\lambda=\lambda=250$$

$$\operatorname{Var}\left(\bar{\xi}_n\right)=\operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^n\xi_i\right)=\frac{1}{n^2}\sum_{i=1}^n\operatorname{Var}(\xi_i)=\frac{n\cdot\lambda}{n^2}=\frac{\lambda}{n}=\frac{250}{n}$$

Оценим вероятность

$$\mathbb{P}\left(\left|\bar{\xi}_n-250\right|>10\right)\leqslant\frac{\operatorname{Var}\left(\bar{\xi}_n\right)}{100}=\frac{250}{100\cdot n}$$

Следовательно, $1-\frac{250}{100\cdot n}\leqslant\mathbb{P}\left(\left|\bar{\xi}_n-250\right|\leqslant10\right)$.
Найдём наименьшее целое n, при котором левая часть неравенства ограничена снизу $0.99\leqslant1-\frac{250}{100\cdot n}$.

Имеем:

$$0.99 \leqslant 1 - \frac{250}{100 \cdot n} \Leftrightarrow \frac{250}{100 \cdot n} \leqslant 0.01 \Leftrightarrow n \geqslant \frac{250}{100 \cdot 0.01} \Leftrightarrow n \geqslant 250$$

Значит, $n = 250$ – наименьшее число дней, при котором с вероятностью не менее 99% среднее число посетителей будет отличаться от 250 не более чем на 10.

в) Требуется найти наименьшее целое n, при котором $\mathbb{P}\left(\left|\bar{\xi}_n - 250\right| \leqslant 10\right) = 0.99$

Имеем:

$$\mathbb{P}(|\bar{\xi}_n - 250| \leqslant 10) = 0.99 \Leftrightarrow \mathbb{P}(-10 \leqslant \bar{\xi}_n - 250 \leqslant 10) = 0.99$$
$$\mathbb{P}(-10n \leqslant S_n - 250n \leqslant 10n) = 0.99$$
$$\mathbb{E}(S_n) = \mathbb{E}(\xi_1 + \ldots + \xi_n) = \mathbb{E}(\xi_1) + \ldots + \mathbb{E}(\xi_n) = 250 \cdot n$$
$$\operatorname{Var}(S_n) = \operatorname{Var}(\xi_1 + \ldots + \xi_n) = \operatorname{Var}(\xi_1) + \ldots + \operatorname{Var}(\xi_n) = 250 \cdot n$$
$$\mathbb{P}\left(\frac{-10n}{\sqrt{250n}} \leqslant \frac{S_n - \mathbb{E}(S_n)}{\sqrt{\operatorname{Var}(S_n)}} \leqslant \frac{10n}{\sqrt{250n}}\right) = 0.99$$
$$2\Phi\left(\frac{10n}{\sqrt{250n}}\right) - 1 = 0.99$$
$$\Phi\left(\frac{10n}{\sqrt{250n}}\right) = \frac{1 + 0.99}{2}$$
$$\frac{10n}{\sqrt{250n}} = 2.58$$
$$\sqrt{n} = 2.58 \cdot \frac{\sqrt{250}}{10}$$
$$n = 16.641$$

Следовательно, наименьшее целое n, есть $n = 17$.

г) Пусть $X_1, X_2, \ldots, X_n, \ldots$ – последовательность независимых случайных величин с одинаковыми конечными математическими ожиданимяи и фиксированными конечными дисперсиями. Тогда $\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(X_i)$ при $n \to \infty$.

В нашем случае случаные величины $\xi_1^2, \xi_2^2, \ldots, \xi_n^2, \ldots$ – независимы, $\mathbb{E}(\xi_1^2) = \ldots = \mathbb{E}(\xi_n^2) = \ldots < +\infty$ и $\operatorname{Var}(\xi_1^2) = \ldots = \operatorname{Var}(\xi_n^2) = \ldots < +\infty$. Поэтому в соответствии с ЗБЧ имеем:

$$\frac{\xi_1^2 + \ldots + \xi_n^2}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(\xi_i^2) = \operatorname{Var}(\xi_i) + \mathbb{E}(\xi_i)^2 = \lambda + \lambda^2 = \lambda(\lambda + 1) = 250 \cdot 251 = 62750$$

4. а) Пусть $X_1, X_2, \ldots, X_n, \ldots$ – последовательность независимых, одинаково распределённых случайных величин с $0 < \operatorname{Var}(X_i) < \infty$, $i \in \mathbb{N}$. Тогда для любого (борелевского) множества $B \subseteq R$ имеет место

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\operatorname{Var}(S_n)}} \in B\right) = \int_B \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt,$$

где $S_n := X_1, \ldots, X_n$, $n \in \mathbb{N}$

б) Введём случайную величину

$$X_i = \begin{cases} 1, & \text{если на i-ом шаге Винни-Пух пошёл направо} \\ -1, & \text{если пошёл налево} \end{cases} \qquad i = 1, \ldots, n;$$

Тогда $S_n := X_1 + \ldots + X_n$ означает местоположение Винни-Пуха в n-ую минуту его блужданий по прямой.

$\mathbb{E}(X_i) = -1 \cdot 1/2 + 1 \cdot 1/2 = 0,$

$\mathbb{E}(X_i^2) = (-1)^2 \cdot 1/2 + (1)^2 \cdot 1/2 = 1,$

$\operatorname{Var}(X_i) = \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2 = 1,$

$\mathbb{E}(S_n) = \mathbb{E}(X_1 + \ldots + X_n) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n) = 0,$
$\text{Var}(S_n) = \text{Var}(X_1 + \ldots + X_n) = \text{Var}(X_1) + \ldots + \text{Var}(X_n) = n$

$$\mathbb{P}(S_n \in (-\infty, -5]) = \mathbb{P}(S_n \leqslant -5) = \mathbb{P}\left(\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\text{Var}(S_n)}} \leqslant \frac{-5-0}{\sqrt{n}}\right)$$
$$\overset{n=60}{=} \mathbb{P}\left(\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\text{Var}(S_n)}} \leqslant -0.6454\right) \approx \int_{-\infty}^{-0.6454} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$
$$= \Phi(-0.6454) = 1 - \Phi(0.6454) \approx 0.2593$$

в) Для любых $n \in \mathbb{N}$ и всех $x \in \mathbb{R}$ имеет место оценка:

$$\left|F_{S_n^*}(x) - \Phi(x)\right| \leqslant 0.48 \cdot \frac{\mathbb{E}(|\xi_i - \mathbb{E}\,\xi_i|^3)}{\text{Var}^{3/2}(\xi_i) \cdot \sqrt{n}},$$

где $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt,\quad S_n^* = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{\text{Var}(S_n)}},\quad S_n = \xi_1 + \ldots + \xi_n$

В нашем случае:

$$\mathbb{P}\left(\frac{S_{60} - \mathbb{E}(S_{60})}{\sqrt{\text{Var}(S_{60})}} \leqslant -0.6454\right) = \mathbb{P}(S_{60}^* \leqslant -0.6454) = F_{S_{60}^*}(-0.6454)$$

Согласно неравенству Берри-Эссеена, погрешность $|F_{S_{60}^*}(-0.6454) - \Phi(-0.6454)|$ оценивается сверху величиной

$$0.48 \cdot \frac{\mathbb{E}(|X_i - \mathbb{E}(X_i)|^3)}{\text{Var}(X_i)^{3/2} \cdot \sqrt{n}} = 0.48 \cdot \frac{\mathbb{E}(|X_i|^3)}{1 \cdot \sqrt{60}} = \frac{0.48}{\sqrt{60}} \approx 0.062$$

5. а) Сначала найдём плотность распределения случайной величины X. Пусть $x \leqslant 0$, тогда $f_X(X) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = 0.$

Пусть $x > 0$, тогда

$$f_X(X) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \int_0^{+\infty} 0.005e^{-0.05x-0.1y}dy$$
$$= 0.005e^{-0.05x} \int_0^{+\infty} e^{-0.1y}dy = 0.005e^{-0.05x} \cdot \left(-10e^{-0.1y}\right)\Big|_{y=0}^{y=+\infty} = 0.05e^{-0.05x}$$

Таким образом, имеем:

$$f_X(x) = \begin{cases} 0.05e^{-0.05x} & \text{при } x > 0 \\ 0 & \text{при } x \leqslant 0 \end{cases}$$

То есть $X \sim Exp(\lambda = 0.05)$ – случайная величина X имеет показательное распределение с параметром $\lambda = 0.05$.

Теперь найдём плотность распределения случайной величины Y.

Пусть $y \leqslant 0$, тогда $f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx = 0.$

Пусть $y > 0$, тогда

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx = \int_0^{+\infty} 0.005e^{-0.05x-0.1y}dx$$
$$= 0.005e^{-0.1y} \int_0^{+\infty} e^{-0.05x}dx = 0.005e^{-0.1y} \cdot \left(-20e^{-0.05x}\right)\Big|_{x=0}^{x=+\infty} = 0.1e^{-0.1y}$$

Таким образом, имеем:

$$f_Y(y) = \begin{cases} 0.1e^{-0.1y} & \text{при } y > 0 \\ 0 & \text{при } y \leqslant 0 \end{cases}$$

То есть $Y \sim Exp(\lambda = 0.1)$ – случайная величина Y имеет показательное распределение с параметром $\lambda = 0.1$.

б) Поскольку для любых точек $x, y \in \mathbb{R}$ справедливо равенство $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$, случайные величины X и Y являются независимыми.

в) Найдём вероятность $\mathbb{P}(Y > 5)$:

$$\mathbb{P}(Y > 5) = \int_5^{+\infty} f_Y(y)dy = \int_5^{+\infty} 0.1e^{-0.1y}dy = 0.1 \cdot (-10e^{-0.1x})\Big|_{y=5}^{y=+\infty} = e^{-0.5} \approx 0.6065$$

г) Требуется найти условную вероятность $\mathbb{P}(Y > 8 \mid Y \geqslant 3)$. Для этого предварительно найдём вероятности $\mathbb{P}(Y > 8)$ и $\mathbb{P}(y \geqslant 3)$:

$$\mathbb{P}(Y > 8) = \int_8^{+\infty} f_Y(y)dy = \int_8^{+\infty} 0.1e^{-0.1y}dy = 0.1 \cdot (-10e^{-0.1x})\Big|_{y=8}^{y=+\infty} = e^{-0.8}$$

$$\mathbb{P}(Y \geqslant 3) = \int_3^{+\infty} f_Y(y)dy = \int_3^{+\infty} 0.1e^{-0.1y}dy = 0.1 \cdot (-10e^{-0.1x})\Big|_{y=3}^{y=+\infty} = e^{-0.3}$$

Теперь находим требуемую условную веростность:

$$\mathbb{P}(Y > 8 \mid Y \geqslant 3) = \frac{\mathbb{P}(Y > 8 \cap Y \geqslant 3)}{\mathbb{P}(Y \geqslant 3)} = \frac{\mathbb{P}(Y > 8)}{\mathbb{P}(Y \geqslant 3)} = \frac{e^{-0.8}}{e^{-0.3}} = e^{-0.5} \approx 0.6065$$

д) Сначала найдём условную плотность распределения случайной величины X при условии $Y = y$:

$$\begin{aligned} f_{X|Y}(x \mid y) &= \begin{cases} \frac{f_{X|Y}(x,y)}{f_Y(y)} & \text{при } f_Y(y) \\ 0 & \text{иначе} \end{cases} \\ &= \begin{cases} \frac{0.005e^{-0.05x-0.1y}}{0.1e^{-0.1y}} & \text{при } x > 0, \quad y > 0 \\ 0 & \text{иначе} \end{cases} \\ &= \begin{cases} 0.05e^{-0.05x} & \text{при } x > 0, \quad y > 0 \\ 0 & \text{иначе} \end{cases} \\ &= \begin{cases} f_X(x) & \text{при } y > 0 \\ 0 & \text{при } y \leqslant 0 \end{cases} \end{aligned}$$

Теперь находим условное математическое ожидание

$$\mathbb{E}(X \mid Y = 5) = \int_{-\infty}^{+\infty} x f_{X|Y}(x \mid 5)dx = \int_{-\infty}^{+\infty} x f_X(x)dx = \mathbb{E}(X) = \frac{1}{0.05} = 20$$

Здесь мы воспользовались известным фактом, что если $X \sim Exp(\lambda)$, то $\mathbb{E}(X) = \frac{1}{\lambda}$

е) Требуется найти вероятность $\mathbb{P}(X - Y > 2)$. Для этого введём множества $B := \{(x,y) \in \mathbb{R} : y < x - 2\}$ и $C := \{(x,y) \in \mathbb{R} : y < x - 2, x > 0, y > 0\}$.
Заметим, что искомая вероятность $\mathbb{P}(X - Y > 2)$ может быть записана в виде

$$\mathbb{P}(X - Y > 2) = \mathbb{P}((X,Y) \in B) = \int\int_B f_{X,Y}(x,y)dxdy = \int\int_C f_{X,Y}(x,y)dxdy$$

Стало быть, искомая вероятность

$$
\begin{aligned}
\mathbb{P}(X-Y>2) &= \int\int_C f_{X,Y}(x,y)dxdy = \int_2^{+\infty}\left[\int_0^{x-2} f_{X,Y}(x,y)dy\right]dx \\
&= \int_2^{+\infty}\left[\int_0^{x-2} 0.005e^{-0.05x-0.1y}dy\right]dx \\
&= \int_2^{+\infty}\left[0.005e^{-0.05x}\cdot(-10e^{-0.1y})\Big|_{y=0}^{y=x-2}\right]dx \\
&= \int_2^{+\infty}\left[0.005e^{-0.05x}\cdot\left(1-e^{-0.1(x-2)}\right)\right]dx = \int_2^{+\infty}0.005e^{-0.05x}dx \\
&- \int_2^{+\infty}0.005e^{-0.05x-0.1x+0.2}dx = 0.05\cdot\left(-\frac{1}{0.05}e^{-0.05x}\right)\Big|_{x=2}^{x=+\infty} \\
&- e^{0.02}\cdot 0.05\cdot\left(\frac{1}{0.15}e^{-0.15x}\right)\Big|_{x=2}^{x=+\infty} = e^{-0.1}-\frac{1}{3}e^{-0.1} = \frac{2}{3}e^{-0.1} \approx 0.6032
\end{aligned}
$$

6. Для решения задачи воспользуемся хорошо известными соотношениями:

$$
\begin{aligned}
&\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = 1 \\
&\int_{-\infty}^{+\infty}x\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \mu \\
&\int_{-\infty}^{+\infty}x^2\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \mu^2+\sigma^2
\end{aligned}
$$

а) Указанная в задании функция f_X является плотностью распределения, так как она удовлетворяет двум условиям: f_X является неотрицательной и интеграл от функции f_X в пределах от $-\infty$ до $+\infty$ равен единице:

$$
\int_{-\infty}^{+\infty}f_X(x)dx = \frac{1}{2}\cdot\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-1)^2}{2}}dx + \frac{1}{2}\cdot\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{(x+1)^2}{2}}dx = 1
$$

б) $\mathbb{E}(X) = \int_{-\infty}^{+\infty}xf_X(x)dx = \frac{1}{2}\cdot\int_{-\infty}^{+\infty}x\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-1)^2}{2}}dx + \frac{1}{2}\cdot\int_{-\infty}^{+\infty}x\frac{1}{\sqrt{2\pi}}e^{-\frac{(x+1)^2}{2}}dx = 1-1=0$

в)

$$
\begin{aligned}
\mathbb{E}(X^2) &= \int_{-\infty}^{+\infty}x^2f_X(x)dx = \frac{1}{2}\cdot\int_{-\infty}^{+\infty}x^2\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-1)^2}{2}}dx + \frac{1}{2}\cdot\int_{-\infty}^{+\infty}x^2\frac{1}{\sqrt{2\pi}}e^{-\frac{(x+1)^2}{2}}dx \\
&= \frac{1}{2}(1^2+1^2+(-1)^2+1^2) = 2
\end{aligned}
$$

г) $\mathrm{Var}(X) = \mathbb{E}(X^2)-(\mathbb{E}(X))^2 = 2-0^2 = 2$

15.3. 2015-2016

1. а) $\mathbb{E}(\xi_1\cdot\xi_2) = \int_0^1\int_0^1 xyf(x,y)\,dx\,dy = \int_0^1\int_0^1 \frac{1}{2}\cdot x^2y+\frac{3}{2}\cdot xy^2\,dx\,dy = \int_0^1\frac{y}{6}+\frac{3y^2}{4}\,dy = \frac{1}{3}$

б) $f_{\xi_1|\xi_2}(x|y) = \frac{f_{\xi_1,\xi_2}(x,y)}{f_{\xi_2}(y)} = \frac{0.5x+1.5y}{0.25+1.5y}$, при $y\in(0,1)$

в)

$$
\begin{aligned}
\mathbb{E}(\xi_1|\xi_2=y) &= \int_0^1 xf_{\xi_1|\xi_2}(x|y)dx = \int_0^1 x\frac{0.5x+1,5y}{0.25+1.5y}dx \\
&= \frac{1}{0.25+1,5y}\left(\frac{0.5x^3}{3}+\frac{1.5yx^2}{2}\right)\Big|_0^1 = \frac{1/6+3/4y}{0.25+1.5y}
\end{aligned}
$$

г) Для того, чтобы функция являлась совместной плотностью для пары случайных величин, должно выполнятся следующее:

$$\int_\Omega kxf(x,y)\,dx\,dy = 1$$

Вычислим, чему равняется левая часть:

$$1 = \int_\Omega kxf(x,y)\,dx\,dy = \int_0^1\int_0^1 kx\left(\frac{x+3y}{2}\right)dx\,dy = \int_0^1 \frac{k}{6}+\frac{3ky}{4}\,dy = \frac{k}{6}+\frac{3k}{8} \Rightarrow$$

$$k = \frac{24}{13}$$

2. Заметим, что $\xi+\eta = 10$, тогда $\operatorname{Cov}(\xi,\eta) = \operatorname{Cov}(\xi, 10-\xi) = -\operatorname{Var}(\xi)$.

Представим случайную величину ξ в виде суммы случайных величин $\xi = \xi_1 + \ldots + \xi_{10}$, где

$$\xi_i = \begin{cases} 1, & \text{если у студента есть хотя бы один незачёт}, p = 0.2 \\ 0, & \text{иначе}, p = 0.8 \end{cases} \qquad i = 1,\ldots,10$$

Поскольку результаты каждого из стуентов независимы, $\operatorname{Var}(\xi) = 10\operatorname{Var}(\xi_1)$

$$\operatorname{Cov}(\xi,\eta) = -10(1^2\cdot 0.2 - \big(1\cdot 0.2)^2\big) = -1.6$$

Так как случайные величины ξ и η связаны соотношением $\xi = 10-\eta$, $\operatorname{Corr}(\xi,\eta) = -1$.

Подставив в $\operatorname{Cov}(\xi-\eta,\xi)$ выражение $\eta = 10-\xi$, получим:

$$\operatorname{Cov}(\xi-\eta,\xi) = 2\operatorname{Cov}(\xi,\xi) = 2\cdot 0.16 = 0.32$$

Случайные величины $\xi-\eta$ и ξ не являются независиыми.

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.

$$\mathbb{E}\,R = \alpha + (1-\alpha) = 1 \;\forall\; \alpha\in[0,1]$$

Находим дисперсию:

$$\operatorname{Var}(R) = \alpha^2\cdot 4 + (1-\alpha)^2\cdot 9 - 6\alpha(1-\alpha) = 19\alpha^2 - 24\alpha + 9 \to \min_\alpha$$

Теперь, найдем оптимальное α:

$$\alpha = \frac{24}{38}$$

Финальные цифры:

$$\begin{cases} \operatorname{Var}(R)^P = 4 \Rightarrow \sigma_P = 2 \\ \operatorname{Var}(R)^V = 1.75 \Rightarrow \sigma_V \approx 1.32 \\ \operatorname{Var}(R)^M = \frac{27}{19} \Rightarrow \sigma_M \approx 1.19 \end{cases}$$

4. а) Пусть S количество мальчиков, тогда используя неравенство Маркова получаем:

$$\mathbb{P}(S \geqslant 750) \leqslant \frac{\mathbb{E}(S)}{750} = \frac{2}{3}$$

б) 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:

$$\mathbb{P}(|\bar{X} - 0.5| \geqslant 0.25) \leqslant \frac{\operatorname{Var}(\bar{X})}{0.25^2} = \frac{1/4000}{0.25^2} = 0.004$$

в) Вероятность из предыдущего пункта можно записать в таком виде:

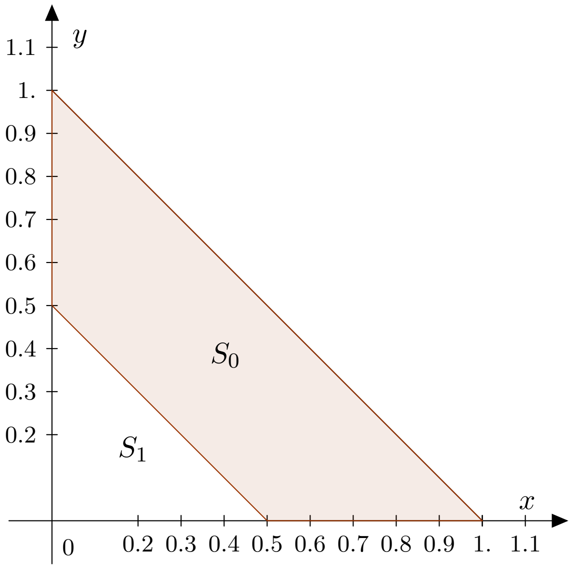
$$\mathbb{P}\left(\left|\bar{X}-0.5\right| \geqslant 0.25\right)=\mathbb{P}\left(\bar{X} \geqslant 0.75\right)+\mathbb{P}\left(\bar{X} \leqslant 0.25\right)=2\,\mathbb{P}\left(\bar{X} \geqslant 0.75\right)=$$
$$=2\,\mathbb{P}(\mathcal{N}(0;1) \geqslant 0.25\sqrt{4000})=2\,\mathbb{P}(\mathcal{N}(0;1) \geqslant 15.8)=1.3\cdot 10^{-56} \approx 0$$

5. Пусть случайная величина S — это валютный курс через полгода. Заметим, что $S=100+\delta_1+\ldots+\delta_{171}$. Тогда по ЦПТ $S \sim \mathcal{N}(142.75, 203.0625)$. Теперь можно искать нужную вероятность:

$$\mathbb{P}(S>250)=\mathbb{P}\left(\frac{S-142.75}{\sqrt{203.0625}}>\frac{250-142.75}{\sqrt{203.0625}}\right)=\mathbb{P}(\mathcal{N}(0,1)>7.6) \approx 0$$

15.4. 2014-2015

1. а) Так как (X,Y) имеют совместное равномерное распределение, $\mathbb{P}\left\{X+Y>1/2\right\}$ можно рассчитать как отношение соответствующих площадей:



Соответственно:

$$\mathbb{P}\left(X+Y>\frac{1}{2}\right)=\frac{S_0}{S_0+S_1}=\frac{0.5-S_1}{0.5}=\frac{\frac{1}{2}-\frac{1}{8}}{\frac{1}{2}}=\frac{3}{4}$$

б)

$$f_Y(y)=\int\limits_0^{1-y} f_{XY}(x,y)dx$$

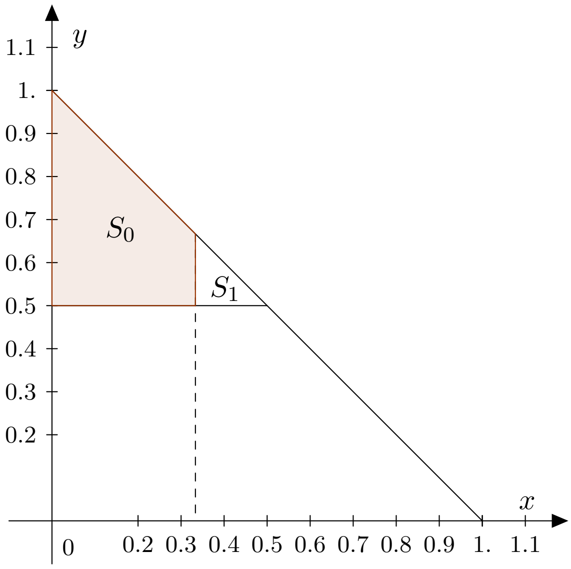
Поэтому, нам сначала нужно найти $f_{XY}(x,y)$, которая для равномерного распределения должна быть константой. Это можно сделать из условия:

$$\int\limits_0^1\int\limits_0^{1-x} f_{XY}(x,y)dxdy=\int\limits_0^1\int\limits_0^{1-x} Cdxdy=1 \Rightarrow$$
$$\int\limits_0^1\int\limits_0^{1-x} Cdxdy=\int\limits_0^1 C(1-x)dx=\left(Cx-C\frac{x^2}{2}\right)\Big|_0^1=\frac{C}{2}=1$$

Откуда имеем $f_{XY}(x,y)=C=2$. Теперь можем найти плотность распределения расходов Васи:

$$f_Y(y)=\int\limits_0^{1-y} 2dx=2(1-y)$$

в) В данном случае площади немного другие, но смысл тот же:



$$\mathbb{P}\left(X < \frac{1}{3} \middle| Y > \frac{1}{2}\right) = \frac{S_0}{S_0 + S_1} = \frac{\frac{1}{8} - S_1}{\frac{1}{8}} = \frac{\frac{1}{8} - \frac{1}{72}}{\frac{1}{8}} = \frac{8}{9}$$

г) При $Y = 1/2$, X распределен равномерно от 0 до $1/2$, поэтому его плотность равна

$$f_X(x) = \frac{1}{\frac{1}{2} - 0} = 2$$

Соответственно, условное математическое ожидание:

$$\mathbb{E}\left(X \middle| Y = \frac{1}{2}\right) = \frac{1}{4}$$

д) $\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$, а маргинальную функцию плотности для X мы можем найти так же, как искали для Y, и получим $f_X(x) = 2(1-x)$. Отсюда:

$$\mathbb{E}(X) = \int_0^1 2x(1-x)dx = \left(x^2 - \frac{2}{3}x^3\right)\Big|_0^1 = \frac{1}{3}$$

е) Если вспомнить формулу для корреляции:

$$\rho_{XY} = \frac{\operatorname{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)}{\sigma_X \sigma_Y}$$

то станет более-менее очевидно, что надо найти $\mathbb{E}(XY)$ и дисперсии X и Y.

$$\mathbb{E}(XY) = \int_0^1 \int_0^{1-x} 2xy dx dy = \int_0^1 2x dx \int_0^{1-x} y dy = \int_0^1 x(x^2 - 2x + 1)dx =$$
$$= \left(\frac{x^4}{4} - \frac{2}{3}x^3 + \frac{x^2}{2}\right)\Big|_0^1 = \frac{3}{4} - \frac{2}{3} = \frac{1}{12}$$

Соответственно:

$$\operatorname{Cov}(X,Y) = \frac{1}{12} - \frac{1}{3} \cdot \frac{1}{3} = -\frac{1}{36}$$

Найдем теперь дисперсии X и Y (они будут одинаковыми, как и математические ождания, в силу симметрии):

$$\mathbb{E}\left(X^2\right) = \int\limits_0^1 2x^2(1-x)dx = \left(\frac{2}{3}x^2 - \frac{x_4}{2}\right)\Big|_0^1 = \frac{1}{6}$$

Поэтому:

$$\mathrm{Var}(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$$

Теперь наконец-то можем найти корреляцию:

$$\rho_{XY} = -\frac{\frac{1}{36}}{\sqrt{\frac{1}{18}}\sqrt{\frac{1}{18}}} = -\frac{1}{2}$$

2. а) Закон больших чисел гласит, что $\bar{X} \to \mathbb{E}(X)$ при $n \to \infty$. Проверим его выполнение в данном случае:

$$\mathbb{E}(X_n) = \frac{1}{2n}(-\sqrt{n}) + \left(1 - \frac{1}{n}\right) \cdot 0 + \frac{1}{2n}\sqrt{n} = 0$$

$$\lim_{n\to\infty} \bar{X} = \lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = 0$$

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.

б) Неравенство Чебышева:

$$\mathbb{P}(|X - \mathbb{E}(X)| \geqslant \varepsilon) \leqslant \frac{\mathrm{Var}(X)}{\varepsilon^2}$$

Соответственно, искомую вероятность можем оценить следующим образом:

$$\mathbb{P}\left(\left|\bar{X}\right| \leqslant 1\right) = 1 - \mathbb{P}\left(\left|\bar{X}\right| \geqslant 1\right) \Rightarrow \mathbb{P}\left(\left|\bar{X}\right| \leqslant 1\right) \geqslant 1 - \frac{\mathrm{Var}\left(\bar{X}\right)}{1}$$

$$\mathrm{Var}\left(\bar{X}\right) = \mathrm{Var}\left(\frac{\sum\limits_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}\, X_i$$

В свою очередь:

$$\mathbb{E}\left(X_i^2\right) = 2 \cdot \frac{1}{2n} \cdot n + \left(1 - \frac{1}{n}\right) \cdot 0 = 1 \Rightarrow \mathrm{Var}(X_i) = 1 \Rightarrow \mathrm{Var}\left(\bar{X}\right) = \frac{1}{n}$$

Поэтому:

$$\mathbb{P}\left(\left|\bar{X}\right| \leqslant 1\right) \geqslant 1 - \frac{1}{n}$$

в)

$$1 - \frac{1}{n} = 0.9 \Rightarrow n = 10$$

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: $\mathbb{E}\, X = 60 \cdot 5000 = 3 \cdot 10^5$; $\mathrm{Var}\, X = 60 \cdot 2000^2 = 2.4 \cdot 10^8$; $\sigma_X = \sqrt{2.4} \cdot 10^4 \approx 1.55 \cdot 10^4$

Теперь по ЦПТ:

$$\mathbb{P}(R \geqslant X) = 0.95$$
$$\mathbb{P}\left(\frac{X - \mathbb{E}\,X}{\sigma_X} \leqslant \frac{R - \mathbb{E}\,X}{\sigma_X}\right) = 0.95$$
$$\mathbb{P}\left(Z \leqslant \frac{R - 3 \cdot 10^5}{1.55 \cdot 10^4}\right) = 0.95$$

Слева функция распределения; подставляя 95-% квантиль стандартного нормального распределения, получаем:

$$\frac{R - 3 \cdot 10^5}{1.55 \cdot 10^4} = 1.64 \Rightarrow R = 325420$$

4. а) По предельной теореме Муавра-Лапласа:

$$\frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \sim \mathcal{N}(0, 1)$$

$$\mathbb{P}\left(\frac{|\hat{p} - p|}{\sqrt{p(1-p)/n}} \leqslant \frac{0.1}{\sqrt{p(1-p)/n}}\right) \geqslant 0.99$$

$$\mathbb{P}\left(|Z| \leqslant \frac{0.1}{\sqrt{p(1-p)/n}}\right) \geqslant 0.99$$

Из симметричности стандартного нормального распределения и зная его 99.5-% квантиль, равный приблизительно 2.58, получаем:

$$\frac{0.1}{\sqrt{p(1-p)/n}} \geqslant 2.58$$
$$\frac{\sqrt{n}}{\sqrt{p(1-p)}} \geqslant \frac{2.58}{0.1}$$
$$\sqrt{n} \geqslant \frac{2.58}{0.1}\sqrt{p(1-p)}$$
$$n \geqslant 665.64 \cdot p(1-p)$$

С помощью неравенства Чебышева:

$$\mathbb{P}\left(|\hat{p} - p| \leqslant 0.1\right) \geqslant 0.99$$

$$\mathbb{P}\left(|\hat{p} - p| \geqslant 0.1\right) \leqslant 0.01$$

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!

б) По предельной теореме Муавра-Лапласа:

$$\mathbb{P}\left(\frac{|\hat{p} - p|}{\sqrt{p(1-p)/n}} \leqslant \frac{\varepsilon}{\sqrt{p(1-p)/1000}}\right) \geqslant 0.99$$

$$\mathbb{P}\left(|Z| \leqslant \frac{\varepsilon}{\sqrt{p(1-p)/1000}}\right) \geqslant 0.99$$

Аналогично пункту 1:

$$\frac{\varepsilon}{\sqrt{p(1-p)/1000}} \geqslant 2.58$$

$$\varepsilon \geqslant 0.082\sqrt{p(1-p)}$$

С помощью неравенства Чебышева:

$$\mathbb{P}\left(|\hat{p}-p| \leqslant \varepsilon\right) \geqslant 0.99$$

$$\mathbb{P}\left(|\hat{p}-p| \geqslant \varepsilon\right) \leqslant 0.01$$

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.

15.5. 2013-2014

1. а)

$$\mathbb{P}\left(Y < X^2\right) = \int_0^1 \int_0^{x^2} (x+y)dydx = \int_0^1 \left(xy\frac{y^2}{2}\right)\Big|_0^{x^2} dx = \int_0^1 \left(x^3 + \frac{x^4}{2}\right) dx =$$
$$= \frac{x^4}{4} + \frac{x^5}{10}\Big|_0^1 = 0.35$$

б) $f_X(x) = \int_0^1 (x+y)dy = xy + \frac{y^2}{2}\Big|_0^1 = x + 0.5$

$\mathbb{E}(X) = \int_0^1 f_X(x)\cdot xdx = \int_0^1 (x^2 + 0.5x)dx = 7/12$

в) $f_{X|Y=2}(x) = \frac{f_{X,Y}(x,2)}{f_Y(2)} = \frac{x+2}{2.5}$

2. а) Из условия находим, что $X \sim \mathcal{N}(0,9)$, тогда

$$\mathbb{P}(X > 1) = \mathbb{P}\left(\frac{X-0}{3} > \frac{1-0}{3}\right) = \mathbb{P}\left(\mathcal{N}(0,1) > \frac{1}{3}\right) = 0.37$$

б) Подготовимся: $\mathbb{E}(2X+Y) = 0$, $\mathrm{Var}(2X+Y) = 4\,\mathrm{Var}(X) + \mathrm{Var}(Y) + 4\,\mathrm{Cov}(X,Y) = 36$

$$\mathbb{P}(2X+Y > 3) = \mathbb{P}\left(\frac{2X+Y-0}{6} > \frac{3-0}{6}\right) = \mathbb{P}(\mathcal{N}(0,1) > 0.5) = 0.31$$

в) $\mathbb{P}(2X+Y > 3|X=1) = \mathbb{P}(Y>1) = \mathbb{P}\left(\frac{Y-0}{2} > \frac{1}{2}\right) = \mathbb{P}(\mathcal{N}(0,1) > 0.5) = 0.31$

г) Заметим, что $\frac{X^2}{9} + \frac{Y^2}{4} \sim \chi^2_2$, и тогда по таблице находим, что $\mathbb{P}\left(\frac{X^2}{9} + \frac{Y^2}{4} > 12\right) = 0.0025$

д) Совместная функция плотности имеет вид:

$$f(x,y) = \frac{1}{2\pi}\cdot\frac{1}{2\cdot 3\sqrt{1-\left(-\frac{1}{6}\right)^2}}\cdot\exp\left(-\frac{1}{2}\cdot\frac{1}{4\cdot 9}\left(1-\left(-\frac{1}{6}\right)^2\right)\left(9x^2+2xy+4y^2\right)\right)$$

3. а) Заметим, что $\frac{X_1}{\sqrt{\frac{X_3^2+X_4^2+X_5^2}{3}}} \sim t_3$. По таблице находим искомую вероятность: 0.15.

   б)

   в) Заметим, что $\frac{X_1^2}{X_2^2+X_3^2} \sim F_{1,2}$. Нужное значение находим в таблице: 0.95.

4. а) По неравенству Чебышёва: $\mathbb{P}(|X_i - 0.5| \geqslant 0.3) \leqslant \frac{1/12}{9/100} = \frac{25}{27}$

   б) По неравенству Маркова: $\mathbb{P}(X_i \geqslant 0.8) \leqslant \frac{5}{8}$

   в) $\mathbb{E}\left(\bar{X}\right) = \frac{1}{2}$, $\mathrm{Var}\left(\bar{X}\right) = \frac{1}{36\cdot 12}$, $\bar{X} \sim \mathcal{N}(\frac{1}{2}, \frac{1}{36\cdot 12})$

   $\mathbb{P}\left(\bar{X} > 0.8\right) = \mathbb{P}\left(\frac{\bar{X}-\frac{1}{2}}{\sqrt{\frac{1}{36\cdot 12}}} \geqslant \frac{0.8-0.5}{\frac{1}{36\cdot 12}}\right) = \mathbb{P}(\mathcal{N}(0,1) \geqslant 6.235) \approx 0$

   г) $\mathbb{P}(|X_i - 0.5| \geqslant 0.3) = 1 - \mathbb{P}(|X_i - 0.5| \geqslant 0.3) = 1 - \mathbb{P}(-0.3 \leqslant X_i - 0.5 \leqslant 0.3) = 0.4$

   д) Нужно воспользоваться неравенством Берри-Ессеена.

$$\mathbb{E}\left(|X_1 - 0.5|^3\right) = \int_0^1 |x_1 - 0.5|^3 \cdot 1dx = 2\int_{0.5}^1 (x_1 - 0.5)^3 dx = \frac{1}{2^5}$$

   е) $\mathbb{P}\left(\bar{X} - 0.5 > 0.3\right) = \frac{25}{27n} \to_{n\to\infty} 0$

5. $\mathbb{P}\left(\hat{p} \leqslant 0.25\right) = \mathbb{P}\left(\frac{\hat{p}-0.2}{\sqrt{\frac{0.2\cdot 0.8}{n}}} \leqslant \frac{0.25}{\sqrt{\frac{0.2\cdot 0.8}{n}}}\right) = 0.99$

   По таблице: $\frac{0.25}{\sqrt{\frac{0.2\cdot 0.8}{n}}} = 2.33 \Rightarrow n = 348$

6. а) $3, 4, 5, 7, 8, 9$

   б) $F(x) = \begin{cases} 0 & x < 3 \\ 1/6 & 3 < x \leqslant 4 \\ 2/6 & 4 < x \leqslant 5 \\ 3/6 & 5 < x \leqslant 7 \\ 4/6 & 7 < x \leqslant 8 \\ 5/6 & 8 < x \leqslant 9 \\ 1 & x > 9 \end{cases}$

   в) $\bar{X} = 6$, $\widehat{\mathrm{Var}}(X) = 28$

15.6. 2012-2013

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

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

2. а) $\mathbb{P}(X > 0.5) = \mathbb{P}(Z > 0.13) \approx 0.45$, $\sigma_X \approx 0.38$

   б) $\mathbb{P}(X + Y > 0.5) = \mathbb{P}(Z > -0.65) \approx 0.74$, $\sigma_{X+Y} \approx 0.17$, $\mathbb{E}(X+Y) = 0.61$

   в) $X = 0.25$ при нормировке даёт $\tilde{X} = -0.53$. Получаем:

$$\mathbb{E}\left(\tilde{Y} \mid \tilde{X} = -0.53\right) = 0.47$$
$$\mathrm{Var}\left(\tilde{Y} \mid \tilde{X} = -0.53\right) = 0.19$$

   Значит $\mathbb{E}\left(Y \mid \tilde{X} = -0.53\right) = 0.34$, $\mathrm{Var}\left(Y \mid \tilde{X} = -0.53\right) = 0.027$.

г) $\mathbb{P}\left(Y > 1/3 \mid \tilde{X} = -0.53\right) = \mathbb{P}(Z > -0.04) = 0.52$

д) Ноль

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

$$\mathbb{P}(|\hat{p} - 0.5| \leqslant 0.01) \geqslant 1 - \frac{\mathrm{Var}(\hat{p})}{0.01^2} = \ldots = 0.375$$

Используя нормальную аппроксимацию:

$$\mathbb{P}(|\hat{p} - 0.5| \leqslant 0.01) = \mathbb{P}(|Z| \leqslant 1.26) \approx 0.79$$

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

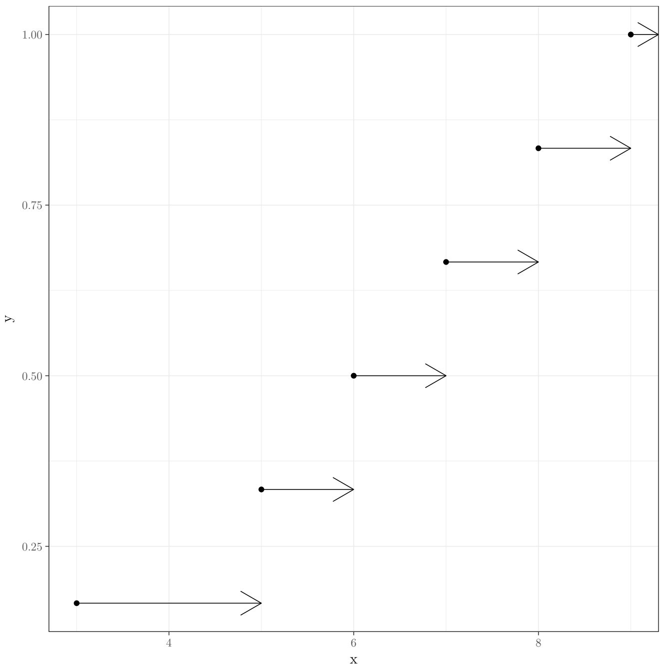
$$\mathbb{P}(120N > 1\,080\,000) = \mathbb{P}(N > 9000) = \mathbb{P}\left(Z > \frac{9000 - 0.3n}{\sqrt{0.21n}}\right) = 0.99$$

Из таблицы находим, что

$$\frac{9000 - 0.3n}{\sqrt{0.21n}} = -2.33$$

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

5. Вариационный ряд: 3, 5, 6, 7, 8, 9. $\bar{X} \approx 6.3$, $\frac{\sum(X_i - \bar{X})^2}{n-1} \approx 4.7$, $\frac{\sum(X_i - \bar{X})^2}{n} \approx 3.9$



15.7. 2011-2012

1. а) $\mathbb{P}(X + Y > 1) = 4/5$. Здесь нужно брать интеграл...

   б) $\mathbb{E}(X) = 13/20 = 0.65$, $\mathbb{E}(XY) = 2/5 = 0.4$, $\mathrm{Cov}(X, Y) = -9/400 = -0.0225$

   в) Нет, так как функция плотности не раскладывается в произведение $h(x) \cdot g(y)$.

   г) Да, так как функция плотности симметрична по x и y

2. а) Заметим, что величина $|X_i|$ распределена равномерно на $[0; b]$, поэтому $\mathbb{E}(|X_i|) = b/2$ и $\mathrm{Var}(|X_i|) = b^2/12$. Значит $\mathbb{E}(\hat{b}) = cb$ и для несмещённости $c = 1$.

б) Находим MSE через b и c:

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

Отсюда $c = \frac{6}{7}$.

3. а) $\mathbb{E}(\hat{\mu}_1) = 6\mu/6 = \mu$, несмещённая

б) $\mathbb{E}(\hat{\mu}_2) = \alpha\mu + \beta\mu$ и $\mathrm{Var}(\hat{\mu}_2) = \alpha^2 \frac{\sigma^2}{3} + \beta^2 \frac{2\sigma^2}{3}$. Для несмещённости необходимо условие $\alpha + \beta = 1$. Для минимизации дисперсии получаем уравнение

$$\alpha - 2(1-\alpha) = 0$$

Отсюда оценка имеет вид $\frac{2}{3}\bar{X} + \frac{1}{3}\bar{Y}$

4. а) $S = X_1 + X_2 + X_3$, слагаемых мало, использовать нормальное распределение некорректно. Можно использовать неравенство Чебышева, $\mathbb{E}(S) = 27$, $\mathrm{Var}(S) = 27$, поэтому

$$\mathbb{P}(S \in [20; 34]) = \mathbb{P}(|S - \mathbb{E}(X)| \leqslant 7) \geqslant 1 - \frac{27}{7^2} = \frac{22}{49}$$

б) Используем неравенство Маркова:

$$\mathbb{P}(X_1 \geqslant 12) \leqslant \mathbb{E}(X_1)/12 = 9/12 = 0.75$$

в) Если $S = X_1 + \ldots + X_{50}$, то можно считать, что $S \sim \mathcal{N}(450; 450)$, поэтому

$$\mathbb{P}(S \in [430; 470]) \approx \mathbb{P}(N(0;1) \in [-0.94; +0.94]) \approx 0.6528$$

5. а) Если $Y = X_1 + X_2$, то $\mathbb{E}(Y) = 3$ и $\mathrm{Var}(Y) = 1 + 9 - 2 = 8$, значит $\mathbb{P}(Y > 1) = \mathbb{P}(\mathcal{N}(0,1) > -2/\sqrt{8}) \approx \mathbb{P}(\mathcal{N}(0,1) > -0.71) \approx 0.7602$

б) Находим $\mathrm{Cov}(X_1, Y) = 1 - 1 = 0$. Итого: вектор имеет совместное нормальное распределение с

$$(X_1, Y) \sim \mathcal{N}\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}; \begin{pmatrix} 1 & 0 \\ 0 & 8 \end{pmatrix} \right)$$

в) Стандартизируем величины. То есть мы хотим представить их в виде:

$$\begin{cases} X_1 = 1 + aZ_1 + bZ_2 \\ X_2 = 2 + cZ_2 \end{cases}$$

Единица и двойка — это математические ожидания X_1 и X_2. Мы хотим, чтобы величины Z_1 и Z_2 были $\mathcal{N}(0,1)$ и независимы. Получаем систему:

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

Одно из решений этой системы: $c = 3$, $b = -1/3$, $a = 2\sqrt{2}/3$

Используя это разложение получаем:

$$(X_1 \mid X_2 = 2) \sim \left(1 + \frac{2\sqrt{2}}{3} Z_1 - \frac{1}{3} Z_2 \mid 2 + 3Z_2 = 2 \right) \sim$$

$$\sim \left(1 + \frac{2\sqrt{2}}{3} Z_1 - \frac{1}{3} Z_2 \mid Z_2 = 0 \right) \sim \left(1 + \frac{2\sqrt{2}}{3} Z_1 \right) \sim \mathcal{N}(1; 8/9)$$

Еще возможные решения: выделить полный квадрат в совместной функции плотности, готовая формула, etc

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

   б) У нас задано неравенство:

$$\mathbb{P}(|\hat{p} - p| > 0.02) < 0.1$$

   Делим внутри вероятности на $\sqrt{\operatorname{Var}(\hat{p})}$:

$$\mathbb{P}\left(|\mathcal{N}(0;1)| > 0.02\sqrt{4n}\right) < 0.1$$

   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. а) $\mathbb{E}(X_i) = (1+10)/2 = 5.5$, $\mathbb{E}\left(X_1^2\right) = \frac{1}{10}\frac{10\cdot 11\cdot 21}{6} = 77/2$, $\operatorname{Var}(X_i) = 33/4 = \sigma^2$. Можно найти $\operatorname{Cov}(X_1, X_2)$ по готовой формуле, но мы пойдем другим путем. Заметим, что сумма номеров всех вариантов — это константа, поэтому $\operatorname{Cov}(X_1, X_1 + \ldots + X_{40}) = 0$. Значит, $\operatorname{Var}(X_1) + 39\operatorname{Cov}(X_1, X_2) = 0$. В итоге получаем $\operatorname{Cov}(X_1, X_2) = -\sigma^2/39$

   б) $\mathbb{E}(\bar{X}) = 11/2$, $\operatorname{Var}(\bar{X}) = 4\frac{1}{52}$

   в) Да, являются, так как и X_1 и X_2 — это номер случайно выбираемого варианта

   г) Нет, если известно чему равно X_1, то шансы получить такой же X_2 падают

8. Если мы наняли n работников, то ожидаемое количество рабочих человеко-дней равно:

$$\mathbb{E}(X) = 365 \cdot n \cdot \left(\frac{364}{365}\right)^n$$

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

15.8. 2010-2011

1. Перед нами функция плотности двумерного нормального распределения!

   Поэтому: $\mathbb{E}(X) = 0$, $\operatorname{Var}(Y) = 1$, $\operatorname{Cov}(X, Y) = \rho$

2. С помощью таблицы находим, что $\mathbb{P}(X > \sqrt{2}) = 1 - \mathbb{P}(X < 1.14) \approx 0.13$

   Заметим, что $X^2 + Y^2 \sim \chi^2_2$, и находим искомую вероятность по таблице: $\mathbb{P}(X^2 + Y^2 > 4) \approx 0.87$

3. Вспомним важные формулы:

$$\mathbb{E}(X|Y=y) = \mathbb{E}(X) + \operatorname{Cov}(X,Y)\cdot\operatorname{Var}^{-1}(Y)\cdot(y - \mu_y)$$

$$\operatorname{Var}(X|Y=y) = \operatorname{Var}(X) - \operatorname{Cov}(X,Y)\operatorname{Var}^{-1}(Y)\cdot\operatorname{Cov}(Y,X)$$

   Применив их, получим: $\mathbb{E}(X|Y=0) = \frac{22}{9}$, $\operatorname{Var}(X|Y=0) = \frac{32}{9}$. Тогда:

$$\mathbb{P}(X > 0|Y=0) = \mathbb{P}\left(\mathcal{N}(0,1) > \frac{0 - \frac{22}{9}}{\sqrt{\frac{32}{9}}}\right) \approx 0.9$$

   Далее, найдём дисперсию портфеля и минимизиреум её по α:

$$\operatorname{Var}(\alpha X + (1-\alpha)Y) = \alpha^2\operatorname{Var}(X) + (1-\alpha)^2\operatorname{Var}(Y) + 2\alpha(1-\alpha)\operatorname{Cov}(X,Y) =$$
$$= 4\alpha^2 + 9(1-\alpha)^2 - 4\alpha(1-\alpha) = 17\alpha^2 - 22\alpha + 5 \to \min_\alpha$$

$$\alpha = \frac{11}{17}$$

   Нельзя, так как из $\operatorname{Cov}(X+Y, 7X - 2Y) = 0$ не следует независимость $X+Y$ и $7X - 2Y$.

4. Сначала подготовимся и найдём дисперсию случайной величины X:

$$\mathbb{E}(X) = \int_1^\infty \frac{3}{x^4} \cdot x dx = \left. \frac{3x^{-2}}{-2} \right|_1^\infty = \frac{3}{2}$$

$$\mathbb{E}(X^2) = \int_1^\infty \frac{3}{x^4} \cdot x^2 dx = \left. \frac{3x^{-1}}{-1} \right|_1^\infty = 3$$

$$\mathrm{Var}(X) = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4}$$

Перепишем исходное неравенство в виде: $\mathbb{P}(|\bar{X} - \mathbb{E}(X)| < 0.1) \geqslant 1 - 0.02$.

$$\frac{\mathrm{Var}(\overline{X})}{0.1^2} \leqslant 0.02 \Rightarrow \frac{\mathrm{Var}(X)}{n} \leqslant 0.0002 \Rightarrow n \geqslant 3750$$

5. Нужно найти $\mathbb{P}(\hat{p} > \frac{60}{90})$. Воспользуемся теоремой Муавра-Лапласа:

$$\mathbb{P}\left(\hat{p} > \frac{60}{90}\right) = \mathbb{P}\left(\frac{\hat{p} - 0.8}{\frac{0.8 \cdot 0.2}{90}} > \frac{2/3 - 0.8}{\frac{0.8 \cdot 0.2}{90}}\right) = \mathbb{P}(\mathcal{N}(0,1) > -3.16) \approx 1$$

Найдём объём выборки:

$$\mathbb{P}\left(\frac{\hat{p} - 0.8}{\frac{0.8 \cdot 0.2}{n}} > \frac{0.02}{\frac{0.8 \cdot 0.2}{n}}\right) = 0.95 \Rightarrow \frac{0.02}{\frac{0.8 \cdot 0.2}{n}} = 1.65 \Rightarrow n = 33$$

6.

7. Необоходимо решить следующую систему уравнений:

$$\begin{cases} D(\overline{X}_S) = \sum_{l=1}^{L} \left(\frac{w_l^2 \sigma_l^2}{n_l}\right) \to \min \\ n = n_1 + n_2 + n_3 \end{cases}$$

Выпишем лагранжиан:

$$L = \sum_{l=1}^{L} \frac{w_l^2 \sigma_l^2}{n_l} - \lambda(n_1 + n_2 + n_3 - n)$$

$$\frac{\partial L}{\partial n_l} = 0 \Rightarrow n_l^o = \sqrt{\frac{w_l^2 \sigma_l^2}{-\lambda}} \Rightarrow \sum_{l=1}^{L} w_l \sigma_l = \sqrt{-\lambda} n \Rightarrow \sqrt{-\lambda} = \frac{\sum_{l=1}^{L} w_l \sigma_l}{n}$$

Тогда находим объём выборки каждой группы по формуле: $n^o = \frac{w_l \sigma_l}{\sum_{k=1}^{L} w_k \sigma_k} n$

- $n^o_{\text{недорогие}} = 0.255n$
- $n^o_{\text{средние}} = 0.532n$
- $n^o_{\text{дорогие}} = 0.213n$

8. Выборочное среднее: -0.2; выборочная дисперсия: 70.98;

вариационный ряд: $-5.6, -3.2, -0.8, 1.1, 2.9, 4.4$.

9. Для проверки свойства несмещённости найдём математические ожидания оценок:

$$\mathbb{E}(T_1) = \mathbb{E}(2\overline{X}) = 2\,\mathbb{E}(X_1) = \theta$$

$$\mathbb{E}(T_2) = \mathbb{E}((n+1)X_{(1)}) = (n+1)\frac{\theta}{2}$$

Несмещённой является только оценка T_1.

Для проверки оставшихся свойств посчитаем дисперсию оценок:

$$\operatorname{Var}(T_1) = 4\operatorname{Var}(\overline{X}) = \frac{\theta^2}{3n} \to_{n\to\infty} 0$$

$$\operatorname{Var}(T_2) = (n+1)^2 \cdot \frac{\theta^2}{12}$$

Оценка T_1 является более эффективной, и она состоятельна. T_2 не является состоятельной оценкой.

10. По формуле Байеса:

$$\mathbb{P}(X = k|X+Y = k) = \frac{\mathbb{P}(X = k \cap X+Y = n)}{\mathbb{P}(X+Y = n)} = \frac{\mathbb{P}(X+Y = k|X = k)\,\mathbb{P}(X = k)}{\mathbb{P}(X+Y = n)}$$

Находим числитель:

$$\mathbb{P}(X+Y = n|X = k)\cdot\mathbb{P}(X = k) = \mathbb{P}(Y = n-k)\cdot\mathbb{P}(X = k) = p\cdot(1-p)^{n-k-1}\cdot p\cdot p^{k-1} = p^2\cdot(1-p)^{n-2}$$

И знаменатель:

$$\mathbb{P}(X+Y = n) = \sum_{i=1}^{n}\mathbb{P}(X = i \cap Y = n-i) = \sum_{i=1}^{n} p\cdot(1-p)^{i-1}\cdot p\cdot(1-p)^{n-i-1}$$
$$= \sum_{i=1}^{n} p^2(1-p)^n = np^2(1-p)^n$$

Итого:

$$\mathbb{P}(X = k|X+Y = k) = \frac{p^2\cdot(1-p)^{n-2}}{np^2(1-p)^n} = \frac{1}{n(1-p)^2}$$

Второе выражение:

$$\mathbb{P}(Y = k|X = Y) = \frac{\mathbb{P}(Y = k \cap X = Y)}{\mathbb{P}(X = Y)} = \mathbb{P}(Y = k) = p(1-p)^{k-1}$$

15.9. 2009-2010

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

2. По таблице: $\mathbb{P}(X < \sqrt{3}) = 0.9582$

   Заметим, что $X^2+Y^2 \sim \chi^2_2$. По таблице находим искомую вероятность: $\mathbb{P}(X^2+Y^2 < 6) = 0.95$

3. а) $\mathbb{P}(|X-16| > 4) \leqslant 0.75$

   б) $\mathbb{P}(|X-16| > 4) = 1-\mathbb{P}(-4 < X-16 < 4) = 1-\mathbb{P}(12 < X < 20) = \frac{1}{3}$

   в) $\mathbb{P}(|X-16| > 4) = 1-\mathbb{P}\left(\frac{-4}{\sqrt{12}} < \frac{X-16}{\sqrt{12}} < \frac{4}{\sqrt{12}}\right) = 2\,\mathbb{P}\left(\mathcal{N}(0,1) < \frac{4}{\sqrt{12}}\right) = 0.75$

4. $\mathbb{E}(2X+Y) = 2$, $\operatorname{Var}(2X+Y) = 3$

   $\mathbb{P}(2X+Y > 1) = 1-\mathbb{P}\left(\frac{2X+Y-2}{\sqrt{3}} < \frac{1-2}{\sqrt{3}}\right) = 1-\mathbb{P}\left(\mathcal{N}(0,1) < \frac{-1}{\sqrt{3}}\right) = 0.72$

   $\mathbb{P}(2X+Y > 1 \mid Y = 2) = \mathbb{P}(2X > -1) = \mathbb{P}\left(\frac{X+1}{1} > \frac{-0.5+1}{1}\right) = 1-\mathbb{P}(\mathcal{N}(0,1) < -0.5) = 0.31$

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

15.10. 2008-2009 Демо-версия

1. $\mathbb{P}(Y > 2X) = \int_0^{0.5} \int_{2x}^{1} (x+y)dydx = 5/24$
   $f_X(x) = \int_0^1 (x+y)dy = x + \frac{1}{2} \Rightarrow \mathbb{E}(X) = \int_0^1 \left(x + \frac{1}{2}\right) xdx = \frac{7}{12}$
2. а) $\mathbb{E}(X_1 + 3X_2) = -1$, $\text{Var}(X_1 + 3X_2) = 207$
   $\mathbb{P}(X_1 + 3X_2 > 20) = \mathbb{P}\left(\frac{X_1+3X_2+1}{\sqrt{207}} > \frac{20+1}{\sqrt{207}}\right) = \mathbb{P}(\mathcal{N}(0,1) > 1.46) \approx 0.07$
   б) $\mathbb{E}(X_1|X_2 = 0) = 2 - 4.5 \cdot \frac{1}{25}(0+1) = 1.82$, $\text{Var}(X_1|X_2 = 0) = 9 - (-4.5) \cdot \frac{1}{25} \cdot (-4.5) = 8.19$
   $X_1|X_2 = 0 \sim \mathcal{N}(1.82, 8.19)$
3. а) $\mathbb{E}(X) = 1000$, $\text{Var}(X) = 10^6$
   б) X — случайная величина, сумма выплат по одному контракту, $X \sim \exp(0.001)$
   $\mathbb{P}(1000X > 110000) = \mathbb{P}(X > 100) = \int_{110}^{+\infty} 0.001 \exp(-0.001x)dx \approx 0.9$
4. $\mathbb{P}\left(\frac{(X-30)^2}{\text{Var}(X)} < 3\right) = \mathbb{P}(|X - 30| < \sqrt{3\,\text{Var}(X)}) \geqslant 1 - \frac{\text{Var}(X)}{\sqrt{3\,\text{Var}(X)}} = \frac{2}{3}$
   $\mathbb{P}(|X - 30| < \sqrt{3\,\text{Var}(X)}) = \mathbb{P}(X < \sqrt{3 \cdot 900} + 30) = \int_0^{82} \frac{1}{30} \exp\left(-\frac{1}{30}x\right) dx \approx 0.94$
5. $a \approx 1.28$

9-Б. Подразумевая под точками концы гирлянды, докажем следующее утверждение.

Бросим $2n \geqslant 4$ точек $X_1, X_2, \ldots, X_{2n}$ случайным образом на отрезок $[0;1]$. Пусть для $1 \leqslant i \leqslant n$ J_i — это отрезок с концами X_{2i-1} и X_{2i}. Тогда вероятность того, что найдётся такой отрезок J_i, который пересекает все другие отрезки, равна $2/3$ и не зависит от n.

Доказательство. Бросим $2n + 1$ точек на окружность, тогда $2n$ точек образуют пары, а оставшуюся обозначим X и будем считать её и началом, и концом отрезка. Каждому получившемуся отрезку присвоим уникальное имя. Далее, будем двигаться от точки X по часовой стрелке до тех пор, пока не встретим одно и то же имя дважды, например «а». После этого пойдём в обратную сторону, и будем идти, пока не встретим какое-нибудь другое имя дважды, например, «б». Теперь посмотрим на получившуюся последовательность между «б» и «а» на концах пути, читая её по часовой стрелке от «б» до «а». Нас интересует взаимное расположение X, второй «а» и второй «б». Зная, что «а» стоит после X, выпишем все возможные случаи, где может стоять «б»:

а) перед X

б) между X и «а»

в) после «а»

Покажем, что во втором и третьем случае отрезок «б» пересекает все остальные, а в первом такого отрезка вообще нет. Попутно заметим, что появление каждого и случаев равновероятно.

Действительно, если «б» стоит после X, и отрезок соответствующий этому имени, не пересекает какой-нибудь другой отрезок «в», то последовательность выглядела бы как «бввXб» или «бXввб», что противоречит описанному построению. Если «б» стоит перед X и отрезок «в» пересекает оба отрезка «а» и «б», то мы снова приходим в противоречие с построением. В итоге, получаем, что искомая вероятность равна $2/3$.

15.11. 2008-2009

1. а) $\int_0^1 \int_x^1 p(x,y)dydx = 5/12$
   $\int_0^1 \int_0^1 y \cdot p(x,y)dydx = 13/24$
   б) Нет, так как совместная функция плотности не разлагается в произведение индивидуальных
2. а) $\mathbb{P}(|X_1 + X_2 + ... + X_7| > 14) \leqslant \frac{7}{14^2} = \frac{1}{28}$
   $\mathbb{P}(X_1^2 + \ldots + X_7^2 > 14) = \mathbb{P}(X_1^2 + \ldots + X_7^2 - 7 > 7) = \mathbb{P}(|X_1^2 + \ldots + X_7^2 - 7| > 7) \leqslant \frac{2 \cdot 7}{7^2} = \frac{2}{7}$

б) $\mathbb{P}(|X_1+\ldots+X_7|>14)=\mathbb{P}(|\mathcal{N}(0;1)|>14/\sqrt{7})=\mathbb{P}(|\mathcal{N}(0;1)|>5.29)\approx 0$
$\mathbb{P}(X_1^2+X_2^2+\ldots+X_7^2>14)\approx 0.05$

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

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

б) Нормальное, причем $\mathcal{N}(1.4;8)$, корреляция равна $-1/3$

4. $\beta=\frac{1}{3}a^2$
$\mathbb{E}(XY)=\frac{3}{4}a^2$
$\mathbb{E}\left(Y^2\right)=3a^2$
$\hat{\beta}_1=\frac{4}{9}XY$
$\hat{\beta}_2=\frac{1}{9}Y^2$
Так как обе оценки несмещенные вместо сравнения дисперсий можно сравнить квадраты ожиданий
$\frac{16}{81}\mathbb{E}\left(X^2Y^2\right)$ vs $\frac{1}{81}\mathbb{E}\left(Y^4\right)$
...
$16a^4$ vs $\frac{81}{5}a^4$
Дисперсия васиной оценки меньше.

5. Заметим, что Пуассоновская величина с положительной вероятностью принимает значение ноль, значит бывает, что монстры дохнут 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.
$\mathbb{E}(S)=160$, $\operatorname{Var}(S)=160$
$\mathbb{P}(S>200)=\mathbb{P}\left(\frac{S-160}{\sqrt{160}}>3.16\right)\approx 0$

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

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

7. $\operatorname{Var}(X)=5\cdot 0.1\cdot 0.9+5\cdot 0.5\cdot 0.5=1.7$
$\operatorname{Var}(Y)=3\cdot 0.1\cdot 0.9+7\cdot 0.5\cdot 0.5=2.02$
Пусть Z — число правильных ответов на вопросы с 3-го по 7-ой (у Пети и у Васи)

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

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

Аналогично все остальные ковариации равны нулю.

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

8. Любые совпадения с курсом экономической и социальный статистики случайны и непреднамеренны.

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

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

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

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

Принцип кота Матроскина[6] (aka бюджетное ограничение): $4n_1+16n_2+25n_3=7000$

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

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

Некоторые маньяки наизусть знают:

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

9-А. Замечание: неудачные переноски считаются, так как иначе решение тривиально — пробовать нести по 1000 шашек.

а) Так как p небольшая будем считать, что $\ln(1-p) \approx -p$. Уже страшно, да?

б) 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.

в) Начинаем:
$s(1) = 1$, $e(1) = 1/(1-p)$
$s(n) = \arg\min_a (1/(1-p)^a + e(n-a))$, $e(n) = \min_a (1/(1-p)^a + e(n-a))$
Замечаем, что поначалу (где-то до $1/p$ шашек) все идет хорошо, а затем плохо…

г) Ищем упрощенное решение вида $s(n) = s$.
Ожидаемое число переносок равно $\frac{1000}{s} \frac{1}{(1-p)^s}$
Минимизируем по s. Получаем: $s = -1/\ln(1-p) \approx 1/p$.

д) Для тех кому интересно, точный график (10000 шашек, p=0.01):

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

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

9-Б. а) Если сейчас 0 долларов, то брать 1 доллар.
Назовем ситуацию, «шоколадной» если можно выиграть без риска. То есть если игр осталось больше, чем недостающее количество денег.

б) Если игрок не в шоколаде, то оптимальным будет рисковать на первом ходе.
Почему? Получение одного доллара можно перенести на попозже.

в) 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 $d \leqslant k$, то брать по доллару.
Если $d > k$, то с риском попробовать захапнуть $1 + d - k$ долларов.

15.12. 2007-2008 Демо-версия

1. $c = 0.4$

$\mathbb{P}(Y > X) = \mathbb{P}(Y = 1, X = -1) + \mathbb{P}(Y = 2, X = -1) + \mathbb{P}(Y = 2, X = 1) = 0.7$

$\mathbb{E}(XY) = 0.1 - 0.4 - 0.4 - 0.1 + 0.1 + 0.2 = -0.5$

$\mathbb{E}(X|Y > 0) = -1 \cdot \frac{0.6}{0.8} + 1 \cdot \frac{0.2}{0.8} = -0.5$

Случайны величины X и Y не являются независимыми.

2. а) Найдём распредление случайной величины $Z = X_1 + X_2$:

$$\mathbb{E}(Z) = -1, \operatorname{Var}(Z) = \operatorname{Var}(X_1) + \operatorname{Var}(X_2) + 2\operatorname{Cov}(X_1, X_2) = 37$$

Получили, что $Z \sim \mathcal{N}(-1, 37)$, тогда

$$\mathbb{P}(Z > 0) = \mathbb{P}\left(\frac{Z+1}{\sqrt{37}} > \frac{0+1}{\sqrt{37}}\right) = 0.4364$$

б) $\mathbb{E}(X_1|X_2 = -1) = -2 - 4 \cdot \frac{1}{36} \cdot (-1-1) = -\frac{16}{9}$
$\operatorname{Var}(X_1|X_2 = -1) = 9 - (-4) \cdot \frac{1}{36} \cdot (-4) = \frac{77}{9}$
$X_1|X_2 = -1 \sim \mathcal{N}(-\frac{16}{9}, \frac{77}{9})$

3. $c = 6$

$\mathbb{P}(3Y > X) = \int_0^1 \int_0^{3y} 6(x-y)dxdy = 3$

$f_X(x) = \int_0^x 6(x-y)dy = 3x^2 \Rightarrow \mathbb{E}(X) = \int_0^1 3x^3 dx = 0.75$

4. Введём следующие случайные величины:

$X = \begin{cases} 1 & \text{в субботу не будет дождя}, p = 0.5 \\ 0 & \text{иначе}, p = 0.5 \end{cases}$ $Y = \begin{cases} 1 & \text{в воскресенье не будет дождя }, p = 0.7 \\ 0 & \text{иначе }, p = 0.3 \end{cases}$

Найдем их математические ожидания и дисперсии: $\mathbb{E}(X) = 0.5$, $\operatorname{Var}(X) = 0.25$, $\mathbb{E}(Y) = 0.3$, $\operatorname{Var}(Y) = 0.21$.

В условии дана кореелляция X и Y, найдём ковариацию: $\operatorname{Cov}(X, Y) = r \cdot 0.5\sqrt{0.21}$. По определению, $\operatorname{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$, откуда можно найти $\mathbb{E}(XY)$: $\mathbb{E}(XY) = r \cdot 0.5\sqrt{0.21} + 0.5 \cdot 0.7$.

Заметим, что $\mathbb{E}(XY)$ — это и есть искомая вероятность, потому что при подсчёте совместного математического ожидания в сумме будет только одно слагаемое, в котором $X = 1$ и $Y = 1$, остальные же будут равны нулю.

5. Пусть X — случайная величина, обозначающая количество проданных книг. Будем считать, что продажи каждой книги — независимые события.

$\mathbb{E}(50 + 5X) = 100$, $\operatorname{Var}(50 + 5X) = 25$

6. Пусть X — случайная величина, обозначающая изменение цены акции за день, а S — финальную стоимость акции.

а) $\mathbb{E}(S) = \mathbb{E}(1000 + 60X) = 1000 + 60(0.5 \cdot 3 + 0.5 \cdot 5) = 1240$
$\operatorname{Var}(S) = \operatorname{Var}(1000 + 60X) = 3600(0.5 \cdot 9 + 0.5 \cdot 25 - 16) = 3600$

б) $\mathbb{P}(S > 900) = \mathbb{P}\left(\frac{S-1240}{60} > \frac{900-1240}{60}\right) = 1 - \mathbb{P}(\mathcal{N}(0,1) < -\frac{17}{3}) \approx 1$

7. $\mathbb{P}\left(|\hat{p} - 0.6| < 0.01\right) = 0.99 \Rightarrow \mathbb{P}\left(\frac{|\hat{p}-0.6|}{\sqrt{\frac{0.6 \cdot 0.4}{n}}} < \frac{0.01}{\sqrt{\frac{0.6 \cdot 0.4}{n}}}\right) = 0.99$

По таблице: $\frac{0.01}{\sqrt{\frac{0.6 \cdot 0.4}{n}}} = 2.57 \Rightarrow n = 62$

8. а) $\mathbb{P}(-2 < \mathcal{N}(0,1) < 2) = 0.9544$, $1 - \frac{1}{4} < \mathbb{P}(-2\sigma < X - \mu < 2\sigma) < 1$

б) $\mathbb{P}(8 < X < 12) = 0.2$, $1 - \frac{20^2}{12} < \mathbb{P}(-2 < X - \mathbb{E}(X) < 2) < 1$

в) $\mathbb{P}(-1 < X < 3) = \int_{-1}^3 e^{-x}dx \overset{x \geq 0}{=} 1 - e^{-3}$, $1 - \frac{1}{4} < \mathbb{P}(-2 < X - \mathbb{E}(X) < 2) < 1$

9-А. Если убийц нечетное число, то в живых останется только один убийца. Если убийц чётное число, то в живых либо не останется никого, либо мирные граждане.

Следовательно, если кроме гостя в городе нечётное число убийц u, то шансов у гостя никаких нет. Если гость мирный, то в живых останется в финале один из убийц, если гость — убийца, то в финале убийц не останется.

Если кроме гостя в городе чётное число убийц u, и гость будет новым нечётным убийцей, то в финале останется один убийца, и вероятность выжить для гостя равна $1/(u+1)$.

Рассмотрим случай чётного числа убийц и мирного гостя. Замечаем, что прочие мирные лишь отдаляют по времени разборки и не влияют на вероятность выжить гостя. Поэтому уберём остальных мирных жителей.

Чтобы гость выжил, все встречи должны быть между убийцами. Следовательно, вероятность выжить равна:

$$\frac{u}{u-1}\frac{u-1}{u-2}\cdot\frac{u-2}{u-3}\frac{u-3}{u-4}\cdot\ldots\cdot\frac{2}{3}\frac{1}{2}=\frac{1}{u+1}$$

От стратегии гостя ничего не зависит. И вероятность выжить гостя либо 0, либо $1/(u+1)$ в зависимости от чётности числа убийц.

9-Б. Подразумевая под точками концы гирлянды, докажем следующее утверждение.

Бросим $2n \geqslant 4$ точек $X_1, X_2, \ldots, X_{2n}$ случайным образом на отрезок $[0;1]$. Пусть для $1 \leqslant i \leqslant n$ J_i — это отрезок с концами X_{2i-1} и X_{2i}. Тогда вероятность того, что найдётся такой отрезок J_i, который пересекает все другие отрезки, равна $2/3$ и не зависит от n.

Доказательство. Бросим $2n+1$ точек на окружность, тогда $2n$ точек образуют пары, а оставшуюся обозначим X и будем считать её и началом, и концом отрезка. Каждому получившемуся отрезку присвоим уникальное имя. Далее, будем двигаться от точки X по часовой стрелке до тех пор, пока не встретим одно и то же имя дважды, например «а». После этого пойдём в обратную сторону, и будем идти, пока не встретим какое-нибудь другое имя дважды, например, «б». Теперь посмотрим на получившуюся последовательность между «б» и «а» на концах пути, читая её по часовой стрелке от «б» до «а». Нас интересует взаимное расположение X, второй «а» и второй «б». Зная, что «а» стоит после X, выпишем все возможные случаи, где может стоять «б»:

i. перед X
ii. между X и «а»
iii. после «а»

Покажем, что во втором и третьем случае отрезок «б» пересекает все остальные, а в первом такого отрезка вообще нет. Попутно заметим, что появление каждого и случаев равновероятно.

Действительно, если «б» стоит после X, и отрезок соответствующий этому имени, не пересекает какой-нибудь другой отрезок «в», то последовательность выглядела бы как «бввXб» или «бXввб», что противоречит описанному построению. Если «б» стоит перед X и отрезок «в» пересекает оба отрезка «а» и «б», то мы снова приходим в противоречие с построением. В итоге, получаем, что искомая вероятность равна $2/3$.

15.13. 2007-2008

1. $c = 0.2$, далее $\mathbb{P}(Y > -X) = 0.5$ и $\mathbb{E}(X \cdot Y) = 0.1$

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

   $\mathbb{E}(Y|X>0) = 0.25$

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

   б) $p(x_1|0) \sim \exp\left(-\frac{1}{2}\begin{pmatrix} x_1-2 & 0+1 \end{pmatrix}\begin{pmatrix} 9 & -4.5 \\ -4.5 & 25 \end{pmatrix}^{-1}\begin{pmatrix} x_1-2 \\ 0+1 \end{pmatrix}\right)$

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

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

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

   Есть страшные люди, которые наизусть помнят, что:

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

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

3. $\mathbb{P}(Y > 2X) = \int_0^1 \int_0^{y/2} (x+y)dxdy = \frac{5}{24}$
   $\mathbb{E}(X) = \int_0^1 \int_0^1 x(x+y)dxdy = \frac{7}{12}$
   Зависимы

4. Рассмотрим $X = 8 -$ (Васин бал) и $Y =$ (Петин бал) $- 6$
   $\operatorname{Corr}(X,Y) = -0.7$ (т.к. при линейном преобразовании может поменяться только знак корреляции)
   $\operatorname{Var}(X) = \frac{1}{2}\left(1 - \frac{1}{2}\right)$
   $\operatorname{Var}(Y) = \frac{1}{3}\left(1 - \frac{1}{3}\right)$
   Интересующая нас величина - это $\mathbb{P}(X = 1 \cap Y = 1) = \mathbb{E}(XY) = \operatorname{Cov}(X,Y) + \mathbb{E}(X)\,\mathbb{E}(Y)$

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

   key point: $\operatorname{Cov} = -\frac{7\sqrt{2}}{60}$

5. $\frac{37}{40} = 0.925$

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

   Пусть N — число подъемов акции.

   а)
   $$\mathbb{P}(100 \cdot 1.08^N \cdot 0.95^{64-N} > 110) = \mathbb{P}(N\ln(1.08) + (64-N)\ln(0.95) > \ln(1.1)) = \\ = \mathbb{P}\left(N > \frac{\ln(1.1) - 64\ln(0.95)}{\ln(1.08) - \ln(0.95)}\right)$$

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

   б) $\mathbb{E}(N\ln(1.08) + (100-N)\ln(0.95))$
   На этот раз $\mathbb{E}(N) = 50$ и $\mathbb{E}(\ln(P_{100})) = 1.28$

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

   $p = 0.8 \cdot 0.51 \approx 0.4$

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

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

8. а) $\mathbb{P}(X^2 > 2.56\operatorname{Var}(X)) = \mathbb{P}(|X - 0| > 1.6\sigma) \leqslant \frac{\operatorname{Var} X}{2.56\operatorname{Var}(X)} = \frac{100}{256} \approx 0.4$

   б) $\mathbb{P}(X^2 > 2.56\operatorname{Var}(X)) = \mathbb{P}(|Z| > 1.6) = 0.11$

9-А. б) Искомая вероятность равна $\mathbb{P}(A) = f(k+1, n-k)/f(k, n-k)$, где
   $f(a,b) = \int_0^1 x^a(1-x)^b dx$

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

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

   Подставляем, и получаем: $\mathbb{P}(A) = \frac{k+1}{n+2}$

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

9-Б. 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 $F_iM_i = 0$. M, F — общее число мальчиков и девочек соответственно.

   Запасаемся простыми фактами:

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

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

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

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

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

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

Заметим, что $X_i = X_i + M_i F_i = M_i F$. Таким образом,

$X = MF = nF - F^2$

$Y_i = F - F_i - X_i$

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

Далее берём математическое ожидание (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) ...

15.14. 2006-2007

1. $c = 0.3$, $\mathbb{P}(Y > -X) = 0.5$, $\mathbb{E}(XY^2) = 0.5$, $\mathbb{E}(Y|X > 0) = \frac{0.1}{0.4} = 0.25$
2. $\mathbb{E}(Y) = -1$, $\text{Var}(Y) = 207$, $\mathbb{P}(Y > 20) = \mathbb{P}(Z > \frac{21}{\sqrt{207}}) = \mathbb{P}(Z > 1.46) = 0.07$
3. $\mathbb{P}(Y > 2X) = \int_0^1 \int_0^{y/2} (x+y)dxdy = \frac{5}{24}$

   $\mathbb{E}(X) = \int_0^1 \int_0^1 x(x+y)dxdy = \frac{7}{12}$
4. 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 $n_i = \frac{w_i \sigma_i}{\sum w_j \sigma_j}$

   $\text{Var}(\bar{X}^s) = 14.4$
5. а) $(2\theta - 0.2)(1.2 - 3\theta) \to \max$

   $\hat{\theta} = 0.25$

   б) $2.4 - 7\hat{\theta} = 1$, $\hat{\theta} = 0.2$
6. $\mathbb{P}(\bar{X} > 0.33) = \mathbb{P}\left(\frac{\bar{X} - 0.3}{\sqrt{\frac{0.3 \cdot 0.7}{200}}} > \frac{0.33 - 0.3}{\sqrt{\frac{0.3 \cdot 0.7}{200}}}\right) = \mathbb{P}(Z > 1.03) = 0.15$
7. $\bar{X} = 1$, $\hat{\sigma}^2 = 1$

   $\mathbb{P}(\hat{\sigma}^2 > 3\sigma^2) = \mathbb{P}\left(2\frac{\hat{\sigma}^2}{\sigma^2} > 6\right) = \mathbb{P}(\chi_2^2 > 6) = 0.05$
8. а) $\mathbb{P}(X^2 > 4\,\text{Var}(X)) = \mathbb{P}(|X - 0| > 2\sigma) \leqslant \frac{\text{Var}(X)}{4\,\text{Var}(X)} = \frac{1}{4}$

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

   б) Да

   в) $\text{Var}(\bar{X}) = \frac{\theta^2}{n}$;

   г) Да: несмещенность и предел дисперсии равный нулю

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

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

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

 в) $\lim \mathrm{Var}(\hat{\theta}_n) = 0$, оценка несмещённая, следовательно, состоятельная.

11-А. 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, $e_0 = (1 + \frac{1}{100})^{99} = 2.718\ldots$

11-Б. $p_0 = 0.1, p_1 = 0.7 \cdot 0.1;$

$p_n = \mathbb{P}($в первый день Петя познакомился с одной девушкой$)p_{n-1} + \mathbb{P}($в первый день Петя познакомился с дву

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

15.15. 2005-2006

1. $\mathbb{P}(X = 0) = e^{-0.003} \approx 0.997$

2. $c = 0.3, \mathbb{P}(Y > -X) = 0.4, \mathbb{E}(X \cdot Y^2) = 0.5, \mathbb{E}(Y|X > 0) = 1/3$

3. $\mathbb{P}(X_1 + 3X_2 > 20) = \mathbb{P}\left(\frac{X_1+3X_2+1}{\sqrt{207}} > \frac{20+1}{\sqrt{207}}\right) = \mathbb{P}(\mathcal{N}(0,1) > 1.46) = 0.0721$

 $\mathbb{E}(X_1 + 3X_2) = -1, \mathrm{Var}(X_1 + 3X_2) = 9 + 9 \cdot 25 + 6 \cdot (-4.5) = 207$

4. $\mathbb{P}(Y > X) = \int_0^1 \int_x^1 (x + y) dy dx = 0.5$

 $\mathbb{E}(X) = \int_0^1 (x + 0.5) x dx = 7/12$

5. $\mathbb{P}(\hat{p} < 0.21) = \mathbb{P}\left(\frac{\hat{p}-0.8}{\sqrt{\frac{0.8(1-0.8)}{200}}} < \frac{0.21-0.8}{\sqrt{\frac{0.8(1-0.8)}{200}}}\right) \approx 0$

6.

7. а) 0.5118

 б) 7/30

 в) $-e^{-\frac{1}{20\cdot 23}} + e^{-\frac{1}{20\cdot 16}} \approx 0.13$

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

9. $\mathbb{E}(X) = -c\sqrt{\frac{\pi}{2}}, \mathrm{Var}(X) = c\frac{5}{4}\frac{\pi}{2} - c^2\frac{\pi}{2}$

10.

$$\mathbb{P}(X=k|X+Y=50)=\frac{\mathbb{P}(X+Y=50|X=k)\,\mathbb{P}(X=k)}{\mathbb{P}(X+Y=50)}=\frac{\mathbb{P}(Y=50-k)\,\mathbb{P}(X=k)}{\mathbb{P}(X+Y=50)}=$$

$$\frac{\left(\frac{e^{-15}15^{50-k}}{(50-k)!}\right)\left(\frac{e^{-5}5^k}{k!}\right)}{\frac{e^{-(5+15)}(5+15)^{50}}{50!}}=C_{50}^k\frac{5^k\cdot 15^{50-k}}{(5+15)^{50}}=C_{50}^k\left(\frac{5}{5+15}\right)^k\left(\frac{15}{5+15}\right)^{50-k}$$

Получили биномиальное распределение с параметрами $p=1/4$, $n=50$.

15.16. 2004-2005

1. $\mathbb{P}\left(|X-\mathbb{E}(X)|>2\sqrt{\mathrm{Var}(X)}\right)=\mathbb{P}\left(\frac{|X-\mathbb{E}(X)|}{\sqrt{\mathrm{Var}(X)}}>2\right)=2\,\mathbb{P}(\mathcal{N}(0,1)>2)\approx 0.05$
2. $\mu=-1$, $\sigma^2=9$
3. $C_{200}^4\left(\frac{1}{130}\right)^4\left(\frac{129}{130}\right)^{196}$
4. $C_{20}^5 0.52^{15}0.48^9$
5. $\mathbb{P}(|X|>15)\leqslant\frac{16}{15^2}$
6. Подготовимся: $\mathbb{E}(2X-Y)=-5$, $\mathrm{Var}(2X-Y)=16+\mathrm{Var}(Y)$

   $\mathbb{P}(Z>0)=0.9\Rightarrow\mathbb{P}\left(\frac{2X-Y+5}{\sqrt{16+\mathrm{Var}(Y)}}>\frac{5}{\sqrt{16+\mathrm{Var}(Y)}}\right)\Rightarrow\frac{5}{\sqrt{16+\mathrm{Var}(Y)}}=1.28$

   $\mathrm{Var}(Y)=0.74$
7. $\mathbb{E}(X)=\mathrm{Var}(X)=\lambda=0.09$

   $$\mathbb{P}(|X-0.09|>0.18)=1-\mathbb{P}(-0.18<X-0.09<0.18)\overset{X\geqslant 0}{=}1-\mathbb{P}(X=0)=1-e^{-0.09}$$
8. Пусть X — случайная величина, число страховых случаев, $X\sim Bin(n=1000,p=0.05)$. S — размер резерва.

   Тогда условие можно записать в виде: $\mathbb{P}(1500X\leqslant S)=0.95$

   $$\mathbb{P}\left(\frac{X-50}{\sqrt{1000\cdot 0.05\cdot 0.95}}\leqslant\frac{\frac{S}{1500}-50}{\sqrt{1000\cdot 0.05\cdot 0.95}}\right)=0.95\Rightarrow\frac{\frac{S}{1500}-50}{\sqrt{1000\cdot 0.05\cdot 0.95}}\approx 1.65\Rightarrow S\approx 92058$$
9. а) Да

   б) Нет

   в) 1356
10. Пусть X — случайная величина, число сочинённых песенок в день, когда Винни-Пуха кусает пчела, $X\sim Bin(n,p)$.

    Из данной в условии выборки находим $\overline{X}=19/18$, поскольку число наблюдений достаточно велико, $\mathbb{E}(X)=np=36p=19/18$, откуда получаем $p=19/(18\cdot 36)$ и $\mathrm{Var}(X)=np(1-p)\approx 1$.

    Нормальная аппроксимация: $X\sim\mathcal{N}(19/18,1)$.

    $$\mathbb{P}(X>1)=\mathbb{P}\left(X-\frac{19}{18}>-\frac{1}{18}\right)\approx 0.52$$

11. а) Заметим, что величину X_t можно представить в виде:

$$X_t = A_t \cdot X_{t-1} = A_t \cdot A_{t-1} \cdot X_{t-1} = \ldots = A_t \cdot A_{t-1} \cdot \ldots \cdot A_2 \cdot X_1$$

Тогда и предел тоже можно переписать:

$$\lim_{n\to\infty} \frac{\ln X_n}{n} = \lim_{n\to\infty} \frac{\ln A_n + \ldots + \ln A_2 + \ln X_1}{n} \overset{X_1=1}{=} \lim_{n\to\infty} \frac{\ln A_n + \ldots + \ln A_2}{n} \overset{\text{ЗБЧ}}{=} \mathbb{E}(\ln A_1)$$

Осталось найти математическое ожидание $\ln A_1$:

$$\mathbb{E}(\ln A_1) = \int_0^{2a} \frac{1}{2a} \cdot \ln x dx = \ln(2a) - 1$$

б) Из неравенста $\ln(2a) - 1 > 0$ получаем, что темп роста будет положительным при $a > e/2$.

16. Решения промежуточных экзаменов

16.1. 2017-2018

Здесь табличка с ответами

16.2. 2016-2017

17. Решения контрольной номер 3

17.1. 2018-2019

17.2. 2017-2018

5. а) $L(X_1, \ldots, X_n, \mu) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\sum_{i=1}^n (X_i-\mu)^2}$

б) $\hat{\mu}_{ML} = \bar{X}$

в) $\mathbb{E}(\hat{\mu}_{ML}) = \mathbb{E}(\bar{X}) = \mu \Rightarrow$ оценка несмещённая
$\operatorname{plim} \hat{\mu}_{ML} = \operatorname{plim} \bar{X} = \mu \Rightarrow$ оценка состоятельная

г) $I(\mu) = n$

д) $\operatorname{Var}(\theta) \geqslant \frac{1}{I(\theta)}$

е) $\operatorname{Var}(\hat{\mu}_{ML}) = \frac{1}{n}$, так как неравенство Рао-Крамера выполнено как равенство, оценка является эффективной.

ж) $\theta = \mathbb{E}\left(X^2\right) = \operatorname{Var}(X) + \mu^2 = 1 + \mu^2$. Тогда в силу инвариантности оценок максимального правдоподобия: $\hat{\theta}_{ML} = 1 + \hat{\mu}^2$.

з) $\mathbb{E}(\hat{\theta}_{ML}) = 1 + \mathbb{E}(\hat{\mu}^2) = 1 + \mathbb{E}((\bar{X})^2)$
Пользуясь соотношением $\mathbb{E}((\bar{X})^2) = \operatorname{Var}(\bar{X}) + (\mathbb{E}(\bar{X}))^2$, получим: $\mathbb{E}(\hat{\theta}_{ML}) = 1 + \frac{1}{n} + \mu^2$, то есть оценка смещена.
Однако, $\lim_{n\to\infty}\left(1 + \frac{1}{n} + \mu^2\right) = 1 + \mu^2$, значит, оценка асимптотически несмещена.

и) $\hat{\theta}_{ML} \approx 1 + \mu^2 + 2\mu(\hat{\mu} - \mu)$
$\operatorname{Var}(\hat{\theta}_{ML}) \approx 4\mu^2 \operatorname{Var}(\hat{\mu}) = \frac{4\mu^2}{n}$

к) Так как $\hat{\theta}_{ML}$ асимптотически несмещена, то для проверки состоятельности достаточно показать, что $\operatorname{Var}(\hat{\theta}_{ML}) = \frac{4\mu^2}{n} \to_{n\to\infty} 0$.

6. а) $\mathbb{E}(X_1) = \int_0^\theta \frac{2}{\theta^2}(\theta - x)x dx = \frac{\theta}{3}$
$\frac{\hat{\theta}_{MM}}{3} = \bar{X} \Rightarrow \hat{\theta}_{MM} = 3\bar{X}$

б) Оценка $\hat{\theta}$ состоятельна. если $\operatorname{plim}\hat{\theta}_n = \theta$.
$\operatorname{plim}\hat{\theta}_{MM} = \operatorname{plim} 3\bar{X} = 3\,\mathbb{E}(X_1) = \theta \Rightarrow$ оценка состоятельна.

7. а) $\mathbb{E}\left(\frac{X_1+X_2+X_3}{3}\right) = \frac{1}{3}\cdot 3\,\mathbb{E}(X_1) = 132.5$
$\operatorname{Var}\left(\frac{X_1+X_2+X_3}{3}\right) = \frac{1}{9}\operatorname{Var}(X_1+X_2+X_3) = \frac{1}{9}(\operatorname{Var}(X_1)+\operatorname{Var}(X_2)+\operatorname{Var}(X_3)+2\operatorname{Cov}(X_1,X_2)+2\operatorname{Cov}(X_1,X_3)+2\operatorname{Cov}(X_2,X_3)) = \frac{1}{9}(3\operatorname{Var}(X_1)+6\operatorname{Cov}(X_1,X_2))$
$\operatorname{Var}(X_1) = \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2 = \frac{1}{4}\cdot 30^2 - \frac{1}{4}\cdot 500^2 - 132.5^2 = 45168.75$
$\operatorname{Cov}(X_1, X_1+\ldots+X_4 = \operatorname{Var}(X_1) + 3\operatorname{Cov}(X_1,X_2) = 0 \Rightarrow \operatorname{Cov}(X_1,X_2) = -\frac{45168.75}{3} = -15056.25$
$\operatorname{Var}\left(\frac{X_1+X_2+X_3}{3}\right) = 5018.75$

б) Вовочке удастся войти в метро, учитывая, что стоимость проезда по тройке составляет 35 рублей, только если одной из выбранных карт будет карта с 500 рублями на счету.
Всего Вовочка может выбрать три карты C_4^3 способами. Если одна из выбранных карт — карта с 500 рублями, то выбрать оставшиеся две карты можно C_3^2 способами.
Тогда вероятность того, что одной из выбранных карт будет карта с 500 рублями равна $\frac{C_3^2}{C_4^3} = \frac{3}{4}$

8. $\Delta_i = X_i - Y_i \sim \mathcal{N}(\mu_x - \mu_y, \sigma^2)$
$\bar{X} = 297.5$, $\bar{Y} = 247.5$, $\bar{\Delta} = \bar{X} - \bar{Y} = 50$
$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\Delta_i - \bar{\Delta})^2 = 18266.(6)$.
Критическое значение — $t_{0.975,3} = 3.182$ и доверительный интервал имеет вид:

$$50 - 3.182\sqrt{\frac{18266.(6)}{4}} < \mu_x - \mu_y < 50 + 3.182\sqrt{\frac{18266.(6)}{4}}$$

Так как 0 входит в доверительный интервал, нельзя отвергнуть предположение о равенстве расхожов.

9. а) $0.7 - 1.96\sqrt{\frac{0.7\cdot 0.3}{60}} < p < 0.7 + 1.96\sqrt{\frac{0.7\cdot 0.3}{60}}$
б) Да, так как 0.7667 входит в доверительный интервал.
в) $\mathbb{P}(|p-\hat{p}| \leqslant 0.01) = 0.95$
$\mathbb{P}\left(\frac{|0.7-p|}{\sqrt{\frac{0.7\cdot 0.3}{n}}} < \frac{0.01}{\sqrt{\frac{0.7\cdot 0.3}{n}}}\right) = 0.95$
$\frac{0.01}{\sqrt{\frac{0.7\cdot 0.3}{n}}} = 1.96 \Rightarrow n \approx 8068$

17.3. 2016-2017

1. а) $-2, 1, 4, 7, 10$
б) 4
в) $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = 18$
г) $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = 22.5$
д) $\frac{1}{n}\sum_{i=1}^{n} X_i^2 = 34$
е) $F_n(x) = \begin{cases} 0, & x < -2 \\ \frac{1}{5}, & -2 \leqslant x < 1 \\ \frac{2}{5}, & 1 \leqslant x < 4 \\ \frac{3}{5}, & 4 \leqslant x < 7 \\ \frac{4}{5}, & 7 \leqslant x < 10 \\ 1, & x \geqslant 1 \end{cases}$

2. $\mathbb{E}(X_1 + X_2) = 2 \cdot 11000 = 22000$

   $\operatorname{Var}(X_1 + X_2) = \operatorname{Var}(X_1) + \operatorname{Var}(X_2) + 2\operatorname{Cov}(X_1, X_2) = \operatorname{Var}(X_1) + \operatorname{Var}(X_2) - \frac{2\operatorname{Var}(X_1)}{N-1} = 2 \cdot 3000 - \frac{2\cdot 3000}{3-1} = 3000$

3. а) Необходимо решить следующую задачу:

$$\begin{cases} \frac{0.4^2\cdot 10^2}{n_1} + \frac{0.5^2\cdot 30^2}{n_2} + \frac{0.1^2\cdot 60^2}{n_3} \to \min_{n_1,n_2,n_3} \\ 150n_1 + 300n_2 + 600n_3 \leqslant 30000 \end{cases}$$

Выпишем функцию Лагранжа и найдём её частные производные по n_1, n_2 и n_3:

$$L(n_1, n_2, n_3, \lambda) = \frac{0.4^2 \cdot 10^2}{n_1} + \frac{0.5^2 \cdot 30^2}{n_2} + \frac{0.1^2 \cdot 60^2}{n_3} + \lambda(150n_1 + 300n_2 + 600n_3 - 30000)$$

$$\frac{\partial L}{\partial n_1} = -\frac{0.4^2 \cdot 10^2}{n_1^2} + 150\lambda \quad \Rightarrow \quad 150\lambda = \frac{0.4^2 \cdot 10^2}{n_1^2}$$

$$\frac{\partial L}{\partial n_2} = -\frac{0.5^2 \cdot 30^2}{n_2^2} + 300\lambda \quad \Rightarrow \quad 150\lambda = \frac{0.5^2 \cdot 30^2}{2n_2^2}$$

$$\frac{\partial L}{\partial n_2} = -\frac{0.1^2 \cdot 60^2}{n_3^2} + 600\lambda \quad \Rightarrow \quad 150\lambda = \frac{0.1^2 \cdot 60^2}{4n_3^2}$$

Выразим n_2 и n_3 через n_1:

$$\frac{0.4 \cdot 10}{n_1} = \frac{0.5 \cdot 30}{\sqrt{2}n_2} \Rightarrow n_2 = \frac{15n_1}{4\sqrt{2}}$$

$$\frac{0.4 \cdot 10}{n_1} = \frac{0.1 \cdot 60}{2n_3} \Rightarrow n_3 = \frac{6n_1}{8}$$

Подставим всё в бюджетное ограничение:

$$150n_1 + 300 \cdot \frac{15n_1}{4\sqrt{2}} + 600 \cdot \frac{6n_1}{8} = 30000$$

Откуда получаем: $n_1 = 21.5 \approx 22$, $n_2 \approx 57$, $n_3 \approx 16$.

б) $\operatorname{Var}(\bar{X}_S) = \sum_{l=1}^{L} \frac{w_l^2 \cdot \sigma_l^2}{n_l} = \frac{0.4^2\cdot 10^2}{22} + \frac{0.5^2\cdot 30^2}{57} + \frac{0.1^2\cdot 60^2}{16} \approx 6.92$

4. $\hat{p} = \frac{8000}{12300000} = \frac{2}{3075}$, $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \approx 7.27 \cdot 10^{-6}$, $z_{\frac{\alpha}{2}} = 1.96$

   $\frac{2}{3075} - 1.96 \cdot 7.27 \cdot 10^{-6} < p < \frac{2}{3075} + 1.96 \cdot 7.27 \cdot 10^{-6}$

   $0.00064 < p < 0.00066$

   Поскольку 0 не входит в доверительный интервал, утверждать, что доля статистически не отличается от нуля нельзя.

5. а) i. $\bar{Y} = 43$, $\hat{\sigma}_Y^2 = 32.5$, $t_{0.005,4} = 4.6$

   $43 - 4.6 \cdot \sqrt{\frac{32.5}{5}} < \mu < 43 + 4.6 \cdot \sqrt{\frac{32.5}{5}}$

   $31.27 < \mu < 54.72$

   ii. $\chi^2_{0.95,4} = 9.49$, $\chi^2_{0.05,4} = 0.71$

   $\frac{32.5\cdot 4}{9.49} < \sigma^2 < \frac{32.5\cdot 4}{0.71}$

   $13.7 < \sigma^2 < 183$

   б) i. $X_1, \ldots, X_{n_X} \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y_1, \ldots, Y_{n_Y} \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, $\sigma_X^2 = \sigma_Y^2 = \sigma_0^2$, выборки независимы

   ii. $\bar{Y} - \bar{X} = 43 - 37 = 6$

   $\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{n_X}(X_i - \bar{X})^2 + \sum_{i=1}^{n_Y}(Y_i - \bar{Y})^2}{n_X + n_Y - 2} = \frac{680+130}{5+5-2} = 101.25$

   $t_{0.95,8} = 1.86$

   $6 - 1.86\sqrt{101.25}\sqrt{\frac{1}{5} + \frac{1}{5}} < \mu_Y - \mu_X < 6 + 1.86\sqrt{101.25}\sqrt{\frac{1}{5} + \frac{1}{5}}$

   $-5.83 < \mu_Y - \mu_X < 17.83$

iii. Да, так как ноль входит в доверительный интервал.

6. а) Выборочный второй начальный момент: $\frac{1}{n}\sum_{i=1}^{n} X_i^2$.
Теоретический второй начальный момент: $\mathbb{E}\left(X^2\right) = \mathrm{Var}(X) + (\mathbb{E}\,X)^2 = \theta$
$\hat{\theta}_{MM} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$

б) $\mathbb{E}(\hat{\theta}_{MM}) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i^2) = \theta$ — оценка несмещённая.

в) $\mathrm{Var}(\hat{\theta}_{MM}) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i^2) = \frac{3\theta^2-\theta^2}{n} \underset{n\to\infty}{\to} 0$ — оценка состоятельная ($\mathbb{E}(X^4) = 3\theta^2$).

г)

$$L(x,\theta) = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\theta}}\exp\left(-\frac{1}{2}\frac{x_i^2}{\theta}\right) = \frac{1}{(\sqrt{2\pi\theta})^n}\exp\left(-\frac{1}{2\theta}\sum_{i=1}^{n}x_i^2\right)$$
$$l(x,\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\theta - \frac{1}{2\theta}\sum_{i=1}^{n}x_i^2$$
$$\frac{\partial l}{\partial\theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2$$
$$\hat{\theta}_{ML} = \frac{\sum_{i=1}^{n}x_i^2}{n}$$

д)

$$\frac{\partial^2 l}{\partial\theta^2} = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n}x_i^2$$
$$-\mathbb{E}\left(\frac{\partial^2 l}{\partial\theta^2}\right) = -\frac{n}{2\theta^2} + \frac{1}{\theta^3}\cdot n\theta = \frac{n}{2\theta^2}$$
$$I(\theta) = \frac{n}{2\theta^2}$$

е) $\mathrm{Var}(\hat{\theta}) \geqslant \frac{1}{I(\theta)}$

ж) $\mathrm{Var}(\hat{\theta}_{ML}) = \mathrm{Var}\left(\frac{\sum_{i=1}^{n}x_i^2}{n}\right) = \frac{1}{n^2}\cdot n\,\mathrm{Var}(X_1^2) = \frac{1}{n}(\mathbb{E}(X_1^4) - \mathbb{E}(X_1^2)^2) = \frac{2\theta^2}{n}$
Так как $\mathrm{Var}(\hat{\theta}_{ML}) = \frac{1}{I(\theta)}$, $\hat{\theta}_{ML}$ — эффективная оценка.

7. а) Вспомним, что для распределения Пуассона $\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$

$$L(x,\lambda) = \prod_{i=1}^{n} e^{-\lambda}\frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda}\lambda^{\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}\frac{1}{x_i!}$$
$$l(x,\lambda) = -n\lambda + \ln\lambda\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\ln x_i!$$
$$\frac{\partial l}{\partial\lambda} = -n + \frac{1}{\lambda}\sum_{i=1}^{n}x_i$$
$$\hat{\lambda}_{ML} = \bar{X}$$

Значение по выборке: $\bar{X} = 14.5$

б) см. предыдущий пункт

в) $\hat{\sigma}^2 = \sqrt{\lambda_{ML}} = \sqrt{14.5}$

г) $\mathbb{P}(X=0) = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-\lambda} \Rightarrow \mathbb{P}(\widehat{X=0}) = e^{-\hat{\lambda}} = e^{-\bar{X}}$

д) $\left[14.5 - 1.96\sqrt{\frac{14.5}{6}}; 14.5 + 1.96\sqrt{\frac{14.5}{6}}\right]$, где 1.96 — критическое значение $\mathcal{N}(0;1)$. Конечно, этот результат верен только при больших n. Мы усиленно делаем вид, что $n = 6$ велико. Полученный нами интервал может быть довольно далёк от 95%-го.

е) В данном случае: $g(\hat{\lambda}) = e^{-\hat{\lambda}}$, $g'(\hat{\lambda}) = -e^{-\hat{\lambda}}$. И доверительный интервал имеет вид:

$$\left[e^{-\bar{X}} - 1.96\sqrt{\frac{e^{-2\bar{X}}\bar{X}}{n}}; e^{-\bar{X}} + 1.96\sqrt{\frac{e^{-2\bar{X}}\bar{X}}{n}}\right]$$
$$\left[e^{-14.5} - 1.96\sqrt{\frac{e^{-29}14.5}{6}}; e^{-14.5} + 1.96\sqrt{\frac{e^{-29}14.5}{6}}\right]$$

Снова отметим, что наш интервал может на самом деле быть далеко не 95%-ым, так наше $n = 6$ мало для серьёзного применения метода максимального правдоподобия.

17.4. 2015-2016

1. Пусть случайная величина S – это сумма поглощённых калорий

s	650	800	950
$\mathbb{P}(S = s)$	1/3	1/3	1/3

Тогда

$$\mathbb{E}(S) = \frac{1}{3}\cdot 650 + \frac{1}{3}\cdot 800 + \frac{1}{3}\cdot 950 = 800$$
$$\mathrm{Var}(S) = \frac{1}{3}(650-800)^2 + \frac{1}{3}(800-800)^2 + \frac{1}{3}(950-800)^2 = 15000$$

2. Вариационный ряд: $4, 6, 11$; медиана: 6; выборочное среднее: 7; несмещённая оценка дисперсии: 13

3. Фунуция плотности двумерного нормального распределения имеет вид:

$$f(x,y) = \frac{1}{2\pi}\cdot\frac{1}{\sigma_x\sigma_y\sqrt{1-\rho^2}}$$
$$\cdot\exp\left\{-\frac{1}{2}\frac{1}{\sigma_x^2\sigma_y^2(1-\rho^2)}\left[\sigma_y^2(x-\mu_x)^2 - 2\rho\sigma_x\sigma_y(x-\mu_x)(y-\mu_y) + \sigma_x^2(y-\mu_y)^2\right]\right\}$$

Откуда: $\mu_X = 1$, $\mu_Y = 0$, $\sigma_X = 1$, $\sigma_Y = 1$, $\rho = 0.2$

Обратите внимание, при скобке с x коэффициент σ_y^2, там потом делится на их произведение!

4. а) $X \sim \mathcal{N}(178, 49)$

$$\mathbb{P}(X > 185) = 1 - \mathbb{P}(X < 185) = 1 - \mathbb{P}\left(\frac{X-178}{7} < \frac{185-178}{7}\right)$$
$$= 1 - 0.8413 = 0.1587$$

б) Нет, так как $\mathrm{Cov}(X,Y) = 5.6 \neq 0$

в) $Y \mid X \sim \mathcal{N}\left(\mu_Y + \rho\sigma_Y\cdot\frac{X-\mu_X}{\sigma_X}; \sigma_Y^2\left(1-\rho^2\right)\right)$
$Y \mid X = 185 \sim \mathcal{N}(42.8; 0.36)$
$\mathbb{P}(Y < 42 \mid X = 185) = \mathbb{P}\left(\frac{Y-42.8}{0.6} < \frac{42-42.8}{0.6} \mid X = 185\right) \approx 0.09$

5. а) $\mathbb{E}(X) = \frac{0+2\theta}{2}\mid_{\hat{\theta}} = \bar{X}$, $\hat{\theta}_{MM} = \bar{X}$

б) $\forall\theta \in \Theta : \mathbb{E}\left(\hat{\theta}\right) = \theta \Rightarrow \hat{\theta}$ – несмещённая.

$\forall\theta \in \Theta, \forall\epsilon > 0 : \mathbb{P}\left(|\widehat{\theta}_n - \theta| > \epsilon\right) \to 0 \Rightarrow \widehat{\theta}_n$ – состоятельная.

$\forall\theta \in \Theta : I_n^{-1}(\theta) = \mathrm{Var}\left(\hat{\theta}\right) \Rightarrow \hat{\theta}$ – эффективная.

в) $\mathbb{E}(\theta) = \mathbb{E}(\bar{X}) = \mathbb{E}(X_1) = \theta \Rightarrow \hat{\theta}$ – несмещённая оценка

$\mathrm{Var}\left(\hat{\theta_n}\right) = \mathrm{Var}\left(\bar{X}\right) = \frac{\mathrm{Var}(X_1)}{n} = \frac{4\theta^2}{12\cdot n} \underset{n\to\infty}{\to} 0$; из условий $\mathbb{E}\left(\widehat{\theta}_n\right) = \theta$ и $\mathrm{Var}\left(\widehat{\theta}_n\right) \underset{n\to\infty}{\to} 0$ следует, что $\widehat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ при $n \to \infty$, т.е. $\widehat{\theta}_n$ является состоятельной.

г)

$$\begin{aligned} F_{X_{(n)}} &= \mathbb{P}(\max(X_1, \ldots, X_n) \leqslant x) = \mathbb{P}(X_1 \leqslant x) \cdot \ldots \cdot \mathbb{P}(X_n \leqslant x) = (\mathbb{P}(X_1 \leqslant x))^n \\ &= \begin{cases} 0 & \text{при } x < 0 \\ \left(\frac{x}{2\theta}\right)^n & \text{при } x \in [0, 2\theta] \\ 1 & \text{при } x > 2\theta \end{cases} \end{aligned}$$

$$f_{X_{(n)}}(x) = \begin{cases} 0 & \text{при } x < 0 \\ \frac{nx^{n-1}}{2^n\theta^n} & \text{при } x \in [0, 2\theta] \\ 0 & \text{при } x > 2\theta \end{cases}$$

$$\begin{aligned} \mathbb{E}(X_{(n)}) &= \int_{-\infty}^{+\infty} x \cdot f_{X_{(n)}}(x)dx = \int_0^{2\theta} x \cdot \frac{nx^{n-1}}{2^n\theta^n}dx = \frac{n}{2^n\theta^n} \cdot \frac{x^{n+1}}{n+1}\Bigg|_{x=0}^{x=2\theta} \\ &= \frac{n}{2^n\theta^n} \cdot \frac{2^{n+1} \cdot \theta^{n+1}}{n+1} = \frac{n2\theta}{n+1} \end{aligned}$$

Следовательно, $\mathbb{E}\left(\frac{n+1}{2n} \cdot X_{(n)}\right) = \theta$, а значит, $\tilde{\theta} = \frac{n+1}{2n} \cdot X_{(n)}$ – несмещённая оценка вида $c \cdot X_{(n)}$

д) $\mathrm{Var}\left(\tilde{\theta}\right) = \frac{(n+1)^2}{4n^2}\mathrm{Var}(X_{(n)})$

$$\begin{aligned} \mathbb{E}\left(X_{(n)}^2\right) &= \int_{-\infty}^{+\infty} x^2 f_{X_{(n)}}(x)dx = \int_0^{2\theta} x^2\frac{nx^{n-1}}{2^n\theta^n}dx = \frac{n}{2^n\theta^n}\int_0^{2\theta} x^{n+1}dx \\ &= \frac{n}{2^n\theta^n} \cdot \frac{x^{n+2}}{n+2}\Bigg|_{x=0}^{x=2\theta} = \frac{n}{2^n\theta^n} \cdot \frac{2^{n+2} \cdot \theta^{n+2}}{n+2} = \frac{n \cdot 4 \cdot \theta^2}{n+2} \end{aligned}$$

$$\mathrm{Var}(X_{(n)}) = \mathbb{E}\left(X_{(n)}^2\right) - (\mathbb{E}(X_{(n)}))^2 = \frac{4n\theta^2}{n+2} - \frac{4n^2 \cdot \theta^2}{(n+1)^2} = 4n\theta^2\left(\frac{1}{n+2} - \frac{n}{(n+1)^2}\right)$$

$$\mathrm{Var}\left(\tilde{\theta}\right) = \frac{(n+1)^2}{4n^2}\mathrm{Var}(X_{(n)}) = \frac{(n+1)^2}{4n^2} \cdot 4n\theta^2\left(\frac{n^2+2n+1-n^2-2n}{(n+2)(n+1)^2}\right) = \frac{\theta^2}{n(n+2)}$$

Оценка $\tilde{\theta}_n$ является состоятельной, так как $\mathbb{E}\left(\tilde{\theta}_n\right) = \theta$ и $\mathrm{Var}\left(\tilde{\theta}_n\right) = \frac{\theta^2}{n(n+2)} \underset{n\to\infty}{\to} 0$

е) Поскольку $\mathrm{Var}\left(\widehat{\theta}_n\right) = \frac{\theta^2}{3n}$, $\mathrm{Var}\left(\tilde{\theta}_n\right) = \frac{\theta^2}{n(n+2)}$ при достаточно большом n $\mathrm{Var}\left(\tilde{\theta}_n\right) < \mathrm{Var}\left(\widehat{\theta}_n\right)$. Значит, при таких n оценка $\tilde{\theta}_n$ будет более эффективной по сравнению с оценкой $\widehat{\theta}_n$.

6. а) $X_i \sim Bin(n = 10, p)$

б) $L(p) = \prod_{i=1}^n C_{10}^{x_i}p^{x_i}(1-p)^{10-x_i}$

в) $\ln L(p) = \sum_{i=1}^n \ln C_{10}^{x_i} + \sum_{i=1}^n x_i \ln p + \sum_{i=1}^n (10 - x_i)\ln(1-p) \to \max_p$

$\frac{\partial \ln L}{\partial p} = \frac{\sum_{i=1}^n x_i}{p} - \frac{\sum_{i=1}^n(10-x_i)}{1-p}\ |_{p=\hat{p}} = 0 \Rightarrow \hat{p} = \frac{\bar{X}}{n} = \frac{\sum_{i=1}^n x_i}{10n}$

$\frac{\partial^2 \ln L}{\partial p^2} = -\frac{\sum_{i=1}^n x_i}{p^2} - \frac{\sum_{i=1}^n(10-x_i)}{(1-p)^2}$

г) $I(p) = -\mathbb{E}\left(\frac{\partial^2 \ln L}{\partial p^2}\right) = \mathbb{E}\left(\frac{\sum_{i=1}^n x_i}{p^2} + \frac{\sum_{i=1}^n (10-x_i)}{(1-p)^2}\right) = \frac{10np}{p^2} + \frac{10n-10np}{(1-p)^2} = \frac{10n}{p(1-p)}$

$i(p) = \frac{I(p)}{n} = \frac{10}{p(1-p)}$

д) $\operatorname{Var}(T) \geqslant \frac{1}{ni(T)}$

е) $\operatorname{Var}(\hat{p}_{ML}) = \operatorname{Var}\left(\frac{\sum_{i=1}^n x_i}{10n}\right) = \frac{1}{(10n)^2} n \operatorname{Var}(X_i) = \frac{1}{100n} 10p(1-p) = \frac{p(1-p)}{10n}$

$\frac{p(1-p)}{10n} = \frac{1}{\frac{10n}{p(1-p)}} \Rightarrow$ да

ж) $\mathbb{E}(X_i) = 10p \Rightarrow \widehat{\mathbb{E}(X_i)} = 10\hat{p}_{ML} = \bar{X}$

$\operatorname{Var}(X_i) = 10p(1-p) \Rightarrow \widehat{\operatorname{Var}(X_i)} = \bar{X}\left(1 - \frac{\bar{X}}{10}\right)$

з) $\hat{p} = \frac{3+4+0+2+6}{10 \cdot 5} = 0.3$

7. $L(x,\theta) = \prod_{i=1}^n (1+\theta)x_i^\theta = (1+\theta)^n \prod_{i=1}^n x_i^\theta \to \max_\theta$

$\ln L(x,\theta) = n\ln(1+\theta) + \theta\sum_{i=1}^n \ln x_i \to \max_\theta$

$\frac{\partial \ln L}{\partial \theta} = \frac{n}{1+\theta} + \sum_{i=1}^n \ln x_i \,|_{\theta=\hat{\theta}} = 0 \Rightarrow \hat{\theta}_{ML} = -\frac{1}{\sum_{i=1}^n \ln x_i} - 1$

8. $\bar{X} - 1.96\frac{7}{\sqrt{20}} < \mu < \bar{X} + 1.96\frac{7}{\sqrt{20}}$

17.5. 2014-2015

1. Пусть случайная величина S – это сумма поглощённых калорий

s	650	800	950
$\mathbb{P}(S=s)$	1/3	1/3	1/3

Тогда

$$\mathbb{E}(S) = \frac{1}{3}\cdot 650 + \frac{1}{3}\cdot 800 + \frac{1}{3}\cdot 950 = 800$$
$$\operatorname{Var}(S) = \frac{1}{3}(650-800)^2 + \frac{1}{3}(800-800)^2 + \frac{1}{3}(950-800)^2 = 15000$$

2. Ответ — решение оптимизационной задачи:

$$\begin{cases} \operatorname{Var}(\bar{X}_S) = \frac{0.3^2\cdot 10^2}{n_1} + \frac{0.6^2\cdot 30^2}{n_2} + \frac{0.1^2\cdot 60^2}{n_3} \to min_{n_1,n_2,n_3} \\ 150\cdot n_1 + 300\cdot n_2 + 600\cdot n_3 \leqslant 15000 \end{cases}$$

3. а) $\mathbb{E}(\mu_1) = \mathbb{E}\left(\frac{X_1+X_2}{2}\right) = \frac{1}{2}(\mu+\mu) = \mu \Rightarrow$ несмещённая

$\mathbb{E}(\mu_2) = \mathbb{E}\left(\frac{X_1}{4} + \frac{X_2+\ldots+X_{n-1}}{2n-4} + \frac{X_n}{4}\right) = \frac{1}{4}\mu + \frac{n-2}{2n-4}\mu + \frac{1}{4}\mu = \mu \Rightarrow$ несмещённая

$\mathbb{E}(\bar{X}) = \mathbb{E}\left(\frac{X_1+\ldots+X_n}{n}\right) = \mu \Rightarrow$ несмещённая

б) $\operatorname{Var}(\mu_1) = \operatorname{Var}\left(\frac{X_1+X_2}{2}\right) = \frac{1}{4}2\operatorname{Var}(X_1) = \frac{\sigma^2}{2}$

$\operatorname{Var}(\mu_2) = \operatorname{Var}\left(\frac{X_1}{4} + \frac{X_2+\ldots+X_{n-1}}{2n-4} + \frac{X_n}{4}\right) = \frac{\sigma^2}{16} + \frac{(n-2)\sigma^2}{(2n-4)^2} + \frac{\sigma^2}{16} = \sigma^2\left(\frac{1}{8} + \frac{1}{2(2n-4)}\right)$

$\operatorname{Var}(\mu_3) = \operatorname{Var}\left(\frac{X_1+\ldots+X_n}{n}\right) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$

4. а) $X \sim \mathcal{N}(1;1)$

$\mathbb{P}(X > 1) = 0.5$, так как нормальное распределение симметрично относительно своего математического ожидания.

б) $X \sim \mathcal{N}(1;1), 2X \sim \mathcal{N}(2;4), Y \sim \mathcal{N}(2,4) \Rightarrow 2X+Y \sim \mathcal{N}(4,4)$
$\mathbb{P}(2X+Y>2) = 1-\mathbb{P}(2X+Y<2) = 1-\mathbb{P}\left(\frac{2X+Y-4}{2} < \frac{2-4}{2}\right) \approx 1-0.16 = 0.84$

в) $Y \mid X \sim \mathcal{N}\left(\mu_Y + \rho\sigma_Y \cdot \frac{X-\mu_X}{\sigma_X}; \sigma_Y^2\left(1-\rho^2\right)\right), Y \mid X=2 \sim \mathcal{N}(1.5,3)$
$\mathbb{E}(2X+Y \mid X=2) = 2\,\mathbb{E}(X \mid X=2) + \mathbb{E}(Y \mid X=2) = 4+1.5 = 5.5$

5. а) $X_1^2 + X_2^2 \sim \chi_2^2, \mathbb{P}(X_1^2+X_2^2 > 6) \approx 0.05$

б) $\mathbb{P}\left(\frac{X_1^2}{X_2^2+X_3^2} > 9.25\right) = \mathbb{P}\left(\frac{\frac{X_1^2}{1}}{\frac{X_2^2+X_3^2}{2}} > 18.5\right) \approx 0.05, \frac{\frac{X_1^2}{1}}{\frac{X_2^2+X_3^2}{2}} \sim F_{1,2}$

6. а)

$$F_{X_{(n)}} = \mathbb{P}(\max(X_1,\ldots,X_n) \leqslant x) = \mathbb{P}(X_1 \leqslant x)\cdot\ldots\cdot\mathbb{P}(X_n \leqslant x) = (\mathbb{P}(X_1 \leqslant x))^n$$
$$= \begin{cases} 0 & \text{при } x<0 \\ \left(\frac{x}{\theta}\right)^n & \text{при } x \in [0,\theta] \\ 1 & \text{при } x>\theta \end{cases}$$

$$f_{X_{(n)}}(x) = \begin{cases} 0 & \text{при } x<0 \\ \frac{nx^{n-1}}{\theta^n} & \text{при } x \in [0,\theta] \\ 0 & \text{при } x>\theta \end{cases}$$

$$\mathbb{E}(X_{(n)}) = \int_{-\infty}^{+\infty} x\cdot f_{X_{(n)}}(x)dx = \int_0^\theta x\cdot\frac{nx^{n-1}}{\theta^n}dx = \frac{n}{\theta^n}\cdot\frac{x^{n+1}}{n+1}\bigg|_{x=0}^{x=\theta}$$
$$= \frac{n}{\theta^n}\cdot\frac{\cdot\theta^{n+1}}{n+1} = \frac{n\theta}{n+1}$$

Следовательно, $\mathbb{E}\left(\frac{n+1}{n}\cdot X_{(n)}\right) = \theta$, а значит, $\hat{\theta} = \frac{n+1}{n}\cdot X_{(n)}$ – несмещённая оценка вида $c\cdot X_{(n)}$.

б) $\mathrm{Var}\left(\hat{\theta}\right) = \frac{(n+1)^2}{n^2}\mathrm{Var}(X_{(n)})$

$$\mathbb{E}\left(X_{(n)}^2\right) = \int_{-\infty}^{+\infty} x^2 f_{X_{(n)}}(x)dx = \int_0^\theta x^2\frac{nx^{n-1}}{\theta^n}dx = \frac{n}{\theta^n}\int_0^\theta x^{n+1}dx$$
$$= \frac{n}{\theta^n}\cdot\frac{x^{n+2}}{n+2}\bigg|_{x=0}^{x=\theta} = \frac{n}{\theta^n}\cdot\frac{\theta^{n+2}}{n+2} = \frac{n\cdot\theta^2}{n+2}$$

$$\mathrm{Var}(X_{(n)}) = \mathbb{E}\left(X_{(n)}^2\right) - (\mathbb{E}(X_{(n)}))^2 = \frac{n\theta^2}{n+2} - \frac{n^2\cdot\theta^2}{(n+1)^2} = n\theta^2\left(\frac{1}{n+2} - \frac{n}{(n+1)^2}\right)$$

$$\mathrm{Var}\left(\tilde{\theta}\right) = \frac{(n+1)^2}{n^2}\mathrm{Var}(X_{(n)}) = \frac{(n+1)^2}{n^2}\cdot n\theta^2\left(\frac{n^2+2n+1-n^2-2n}{(n+2)(n+1)^2}\right) = \frac{\theta^2}{n(n+2)}$$

Оценка $\hat{\theta}_n$ является состоятельной, так как $\mathbb{E}(\hat{\theta}_n) = \theta$ и $\mathrm{Var}(\hat{\theta}_n) = \frac{\theta^2}{n(n+2)} \underset{n\to\infty}{\to} 0$

в) $\mathbb{E}(X_1) = \frac{\theta}{2}\big|_{\theta=\hat{\theta}_{MM}} = \bar{X} \Rightarrow \hat{\theta}_{MM} = 2\bar{X}$

г) $\mathrm{Var}\left(2\bar{X}\right) = \frac{4}{n^2}n\,\mathrm{Var}(X_1) = \frac{4\theta}{12n} = \frac{\theta}{3n} > \mathrm{Var}\left(\hat{\theta}_n\right)$

7. а) $L(x,p) = \prod_{i=1}^n \mathbb{P}(X_i = x_i) = p^{\sum_{i=1}^n (x_i-1)}(1-p)^n$
$\ln L(x,p) = \sum_{i=1}^n (x_i-1)\ln p + n\ln(1-p)$
$\frac{\partial \ln L}{\partial p} = \frac{\sum_{i=1}^n (x_i-1)}{p} - \frac{n}{1-p} = 0 \Rightarrow \hat{p} = \frac{\sum_{i=1}^n x_i - n}{\sum_{i=1}^n x_i}$

б) $\mathbb{E}(X) = \frac{n}{p} \Rightarrow \hat{\mathbb{E}}(X) = \frac{n\sum_{i=1}^n x_i}{\sum_{i=1}^n x_i - n}$

17.6. 2011-2012

1. а) $L(x,\lambda)=\prod_{i=1}^{n}\lambda e^{-\lambda x_i}=\lambda^n e^{-\lambda\sum_{i=1}^{n}x_i}$
$\ln L(x,\lambda)=n\ln\lambda-\lambda\sum_{i=1}^{n}x_i\to\max_\lambda$
$\frac{\partial\ln L}{\partial\lambda}=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i\,|_{\lambda=\hat\lambda}=0\Rightarrow\hat\lambda_{ML}=\frac{1}{\bar X}$
$\frac{\partial^2\ln L}{\partial\lambda^2}=-\frac{n}{\lambda^2}\,|_{\lambda=\hat\lambda}<0$

   б) $\hat a=\bar X$

   в) $\mathbb{E}(\hat a)=\mathbb{E}(\bar X)=\frac{1}{\lambda}\Rightarrow$ несмещённая

   г) $\mathrm{Var}(\hat a)=\mathrm{Var}(\bar X)=\frac{1}{n}n\,\mathrm{Var}(X_i)=\frac{1}{n\lambda^2}\to_{n-\to\infty}0\Rightarrow$ состоятельная

   д)

   е) $\mathbb{E}(X)=\frac{1}{\lambda}\,|_{\lambda=\hat\lambda_{MM}}=\bar X\Rightarrow\lambda_{MM}=\frac{1}{\bar X}$

2. а) Истинное стандартное отклонение неизвестно, поэтому используем распределение Стьюдента:
$\bar X-t_{0.025,99}\frac{\hat\sigma}{\sqrt n}<\mu<\bar X+t_{0.025,99}\frac{\hat\sigma}{\sqrt n}$
$9.5-1.98\frac{0.5}{\sqrt{100}}<\mu<9.5-1.98\frac{0.5}{\sqrt{100}}$
$9.4<\mu<9.6$

   б) Значение 10 не лежит в доверительном интервале, значит, гипотеза отвергается
$t_{obs}=\frac{9.5-10}{0.5/\sqrt{100}}=10\Rightarrow$ p-value ≈ 0

3. а) $\frac{\hat\sigma^2_\alpha}{\hat\sigma^2_\beta}\sim F_{n_\alpha-1,n_\beta-1}$
$F_{obs}=\frac{0.5}{0.6}\approx 0.83$, p-value$/2\approx 0.35\Rightarrow$ на любом разумном уровне значимости оснований отвергать H_0 нет

   б) $\hat\sigma^2_0=\frac{\hat\sigma^2_\alpha(n_\alpha-1)+\hat\sigma^2_\beta(n_\beta-1)}{n_\alpha+n_\beta-2}=\frac{0.25\cdot19+0.36\cdot24}{20+25-2}=0.31$
$\frac{\bar X-\bar Y}{\hat\sigma_0\sqrt{\frac{1}{n_\alpha}+\frac{1}{n_\beta}}}\sim t_{n_\alpha+n_\beta-2}$
$t_{obs}=\frac{9.5-9.8}{0.56\sqrt{\frac{1}{20}+\frac{1}{25}}}=-1.79$, p-value$/2=0.04$

17.7. 2010-2011

1. а) Введём обозначения для событий:
   - A — заболеть
   - $\bar A$ — не заболеть
   - B — привиться
   - $\bar B$ — не привиться

   Тогда вероятность заболеть можно получить по формуле полной вероятности:

$$\mathbb{P}(A)=\mathbb{P}(A|B)\cdot\mathbb{P}(B)+\mathbb{P}(A|\bar B)\cdot\mathbb{P}(\bar B)=0.15\cdot0.1+0.2\cdot0.9=0.195$$

   Значит, во время эпидемии заболевает 19.5% людей.

   б) Пусть всё население составляет n человек. Тогда доля заболевших — $0.195n$. Доля привитых заболевших — $0.015n$. Значит, среди заболевших

$$\frac{0.015n}{0.195n}\cdot 100\%\approx 7.7\%$$

   привитых.

2. а)

$$\mathbb{P}(X > 4) = \mathbb{P}\left(\frac{X-3}{\sqrt{25}} > \frac{4-3}{\sqrt{25}}\right) = \mathbb{P}\left(\mathcal{N}(0,1) > 0.2\right) \approx 0.42$$

$$\mathbb{P}(4 < X \leqslant 5) = \mathbb{P}\left(\frac{4-3}{\sqrt{25}} < \frac{X-3}{\sqrt{25}} \leqslant \frac{5-3}{\sqrt{25}}\right)$$
$$= \mathbb{P}(0.2 < \mathcal{N}(0,1) \leqslant 0.4) \approx 0.076$$

б) Найдём математическое ожидание и дисперсию случайной величины $S = X - 2Y$:

$$\mathbb{E}(S) = \mathbb{E}(X) - 2\,\mathbb{E}(Y) = 3 - 2 \cdot 1 = 1$$
$$\operatorname{Var}(S) = \operatorname{Var}(S) + 4\operatorname{Var}(Y) - 4\operatorname{Cov}(X,Y) = 57$$

Теперь можно посчитать вероятность:

$$\mathbb{P}(X - 2Y < 4) = \mathbb{P}\left(\frac{S-1}{\sqrt{57}} < \frac{4-1}{\sqrt{57}}\right) = \mathbb{P}(\mathcal{N}(0,1) < 0.4) \approx 0.66$$

в)

3. а) Сначала найдём среднее и несмещённую оценку дисперсии по выборке Y:

$$\bar{Y} = \frac{480}{12}40, \quad \hat{\sigma}_Y^2 = \frac{358.3 \cdot 12}{11} \approx 391$$

Доверительный интервал для математического ожидания при неизвестной дисперсии имеет вид:

$$\bar{Y} - t_{\frac{\alpha}{2},n-1}\frac{\hat{\sigma}_Y}{\sqrt{n}} < \mu_Y < \bar{Y} + t_{\frac{\alpha}{2},n-1}\frac{\hat{\sigma}_Y}{\sqrt{n}}$$

Критическое значение $t_{0.95,11} \approx 1.8$:

$$40 - 1.8 \cdot \sqrt{\frac{391}{12}} < \mu_Y < 40 + 1.8 \cdot \sqrt{\frac{391}{12}}$$

б) Найдём также среднее и несмещённую оценку дисперсии по выборке X:

$$\bar{X} = \frac{540}{15} = 36, \quad \hat{\sigma}_X^2 = \frac{410.264 \cdot 15}{14} \approx 440$$

И посчитаем оценку общей дисперсии:

$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{n_X}\left(X_i - \bar{X}\right)^2 + \sum_{i=1}^{n_Y}\left(Y_i - \bar{Y}\right)^2}{n_X + n_Y - 2} = \frac{410.264 \cdot 15 + 358.3 \cdot 12}{15 + 12 - 2} \approx 418$$

Тогда, проверяя гипотезу $H_0 : \mu_X = \mu_Y$ против альтернативной $H_a : \mu_X < \mu_Y$, получим следующие наблюдаемое и критическое значения статистики:

$$t_{obs} = \frac{\mu_X - \mu_Y}{\sqrt{\sigma_0^2\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}} = \frac{36 - 40}{\sqrt{418 \cdot \left(\frac{1}{15} + \frac{1}{12}\right)}} \approx -0.51$$
$$t_{crit,0.05,25} \approx -1.7$$

Так как $t_{obs} > t_{crit}$, нет оснований отвергать H_0.

в) Выпишем вариационный ряд по двум выборкам, выделяя наблюдения, относящиеся к Юго-Западному округу, и присвоим им ранги:

Наблюдение	8.4	**14.4**	15.2	15.6	**18.0**	19.1	21.2	**22.0**	**23.9**	**26.6**
Ранг	1	2	3	4	5	6	7	8	9	10
Наблюдение	28.2	28.3	**32.0**	33.8	34.5	38.2	**43.3**	44.1	45.0	**46.7**
Ранг	11	12	13	14	15	16	17	18	19	20
Наблюдение	**54.8**	56.0	**64.0**	**65.1**	68.2	**69.2**	84.2			
Ранг	21	22	23	24	25	26	27			

Теперь посчитаем сумму рангов по выборке меньшего размера, то есть по Юго-Западному округу:

$$T = \sum_{i=1}^{n_Y} R_i = 2 + 5 + 8 + 9 + 10 + 13 + 17 + 20 + 21 + 23 + 24 + 26 = 178$$

Осталось найти значение наблюдаемой статистики и критическое значение:

$$\gamma_{obs} = \frac{T - \frac{1}{2}n_X(n_X + n_Y + 1)}{\sqrt{\frac{1}{12}n_X n_Y(n_X + n_Y)}} = \frac{178 - \frac{1}{2} \cdot 12(12 + 15 + 1)}{\sqrt{\frac{1}{12} \cdot 12 \cdot 15 \cdot (12 + 15)}} \approx 0.5$$
$$\gamma_{crit} = 1.96$$

Так как $\gamma_{obs} < |\gamma_{crit}|$, нет оснований отвергать H_0.

4. Будем проверять гипотезу о том, что вероятность падения маслом вверх равна 0.5:

$$\begin{cases} H_0 : p = 0.5 \\ H_a : p \neq 0.5 \end{cases}$$

Найдём наблюдаемое и критическое значения статистики:

$$z_{obs} = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} = \frac{\frac{105}{200} - \frac{1}{2}}{\sqrt{\frac{\frac{105}{200} \cdot \frac{95}{200}}{200}}} \approx 0.71$$
$$z_{crit} = 1.96$$

Так как $z_{obs} < |z_{crit}|$, нет оснований отвергать H_0.

5. а)

$$\bar{X} = \mathbb{E}(X_1)|_{\mu_1 = \hat{\mu}_1} = \hat{\mu}_1$$
$$S^2 = \mathrm{Var}(X_1)|_{\mu_1 = \hat{\mu}_1, \mu_2 = \hat{\mu}_2} = \hat{\mu}_2 - \hat{\mu}_1^2 \Rightarrow \hat{\mu}_2 = \bar{X^2}$$

б) Найдём максимум логарифмической функции правдоподобия по параметру θ:

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i - \theta)^2\right)$$
$$\ell = \frac{n}{2}\ln 2\pi - \frac{1}{2}\sum_{i=1}^{n}(x_i - \theta)^2$$
$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^{n}(x_i - \theta)\bigg|_{\theta = \hat{\theta}} = 0$$
$$\hat{\theta} = \bar{X}$$

Для проверки эффективности достаточно показать, что неравенство Рао-Крамера для полученной оценки выполняется как равенство.

Найдём информацию Фишера:

$$\frac{\partial^2 \ell}{\partial \theta^2} = -n$$
$$I(\theta) = n$$

и дисперсию оценки $\hat{\theta}$:

$$\operatorname{Var}\left(\hat{\theta}\right) = \frac{1}{n}$$

Подставив полученные значения в неравенство Рао-Крамера

$$\operatorname{Var}\left(\hat{\theta}\right) \geqslant \frac{1}{I(\theta)},$$

получим равенство, значит, найденная оценка является эффективной.

17.8. 2009-2010

6. Пусть $S_i = X_i Y_i$. Замечаем, что $\mathbb{E}(S_i) = \mathbb{E}(X_i)\,\mathbb{E}(Y_i) = ab$, $\mathbb{E}(S_i) = \mathbb{E}(X_i^2)\,\mathbb{E}(Y_i^2) = (a^2+1)(b^2+1)$. Отсюда получаем систему
$$\begin{cases} \hat{a}\hat{b} = \frac{\sum S_i}{n} \\ (\hat{a}^2+1)(\hat{b}^2+1) = \frac{\sum S_i^2}{n} \end{cases}$$
Для заданных чисел решением будет $\hat{a} = 2$ и $\hat{b} = 18$, или наоборот.

17.9. 2008-2009

17.10. 2007-2008. Демо

Заметим, что $\hat{a}_n \geqslant a$.
$\mathbb{P}(|\hat{a}_n - a| > \varepsilon) = \mathbb{P}(\hat{a}_n a > \varepsilon) = \mathbb{P}(\hat{a}_n > a+\varepsilon) = \mathbb{P}(\min\{X_1, X_2, \ldots, X_n\} > a+\varepsilon) =$
$= \mathbb{P}(X_1 > a+\varepsilon \cap X_2 > a+\varepsilon \cap \ldots) = \mathbb{P}(X_1 > a+\varepsilon)\cdot\mathbb{P}(X_2 > a+\varepsilon)\cdot\ldots = \left(\int_{a+\varepsilon}^{\infty} e^{a-t}dt\right)^n = (e^{-\varepsilon})^n = e^{-n\varepsilon}$
$\lim_{n\to\infty} e^{-n\varepsilon} = 0$
Оценка смещена при любых n, хотя смещение с ростом n убывает

17.11. 2007-2008

Да, Нет, Да, Нет, Нет, Да, Нет, Да, Да, Нет

1. а) $[13.61; 14.39]$
   б) Отвергается ($t_{obs} = -2.12$, $t_{crit} = -1.29$)
   в) $P_{value} \approx 0.017$
2. 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%, $Z_{critical} = 1.96$. Значит, гипотеза H_0 не отвергается.
3. а) $\hat{a} = \frac{5-\bar{X}}{10}$
   б) Да, является

4. $L = -\frac{n}{2}\ln(a) - \frac{na}{8} - \frac{\sum X_i^2}{8a} + c$
   $L' = 0$ равносильно $\hat{a}^2 + 4\hat{a} + 4 = 4 + \frac{\sum X_i^2}{n}$
   $\hat{a} = -2 + \sqrt{4 + \frac{\sum X_i^2}{n}}$
5. $F_{29.39} = \frac{32}{20} = 1.6$
   $F_{crit} = 1.74$
   Гипотеза о том, что дисперсия одинакова не отвергается.
6. $\chi^2_{observed} = 1.15$
   $\chi^2_{2,5\%} = 5.99$
   Правдоподобно.
7. а) $p = 0.7 \cdot 0.8 + 0.3 \cdot 0.7 = 0.77$
   б) $p = \frac{0.7 \cdot 0.8 \cdot 0.2}{0.7 \cdot 0.8 \cdot 0.2 + 0.7 \cdot 0.2 \cdot 0.3} = \frac{8}{11}$
8. а) Заметим, что $\hat{a}_n \leqslant a$.

$$\mathbb{P}(|\hat{a}_n - a| > \varepsilon) = \mathbb{P}(-(\hat{a}_n - a) > \varepsilon) = \mathbb{P}(\hat{a}_n < a - \varepsilon) = \mathbb{P}(\max\{X_1, X_2, \ldots, X_n\} < a - \varepsilon) =$$
$$= \mathbb{P}(X_1 < a - \varepsilon \cap X_2 < a - \varepsilon \cap \ldots) = \mathbb{P}(X_1 < a - \varepsilon) \cdot \mathbb{P}(X_2 < a - \varepsilon) \cdot \ldots = (1 - \varepsilon)^n$$

   $\lim_{n\to\infty} (1 - \varepsilon)^n = 0$
   б) Нет, не является ни при каких n, хотя смещение с ростом n убывает.

9-А. Да, http://en.wikipedia.org/wiki/Simpson's_paradox

9-Б. а) Пусть истинные веса слитков равны x, y и z.
   Назовем оценку буквой $\hat{x}$, $\hat{x} = aX + bY + cZ$.
   Несмещённость: $\mathbb{E}(\hat{x}) = a\,\mathbb{E}(X) + b\,\mathbb{E}(Y) + c\,\mathbb{E}(Z) = ax + by + c(x + y) = x$
   $a + c = 1, b + c = 0$
   $\hat{x} = (1 - c)X + (-c)Y + cZ$
   Эффективность: $\mathrm{Var}(\hat{x}) = ((1 - c)^2 + c^2 + c^2) \cdot \sigma^2 = (3c^2 - 2c + 1)\sigma^2$
   Чтобы минимизировать дисперсию, нужно выбрать $c = 1/3$.
   То есть $\hat{x} = \frac{2}{3}X - \frac{1}{3}Y + \frac{1}{3}Z$.
   б) $\mathrm{Var}(\hat{x}) = \frac{2}{3}\sigma^2$
   $\mathrm{Var}\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{2}\sigma^2$
   Усреднение двух взвешиваний первого слитка лучше.

17.12. 2006-2007

Верный ответ на первый вопрос теста — нет! Некоррелированные одномерные нормальные распределения могут быть не нормальными в совокупности.

1. а) $\mathbb{P}(X = -1) = \frac{5}{12} \cdot \frac{4}{11} = \frac{5}{33}$
   $\mathbb{E}(X) = -1 \cdot \frac{5}{33} + 0 \cdot 2 \cdot \frac{5}{12} \cdot \frac{7}{11} + 1 \cdot \frac{7}{12} \cdot \frac{6}{11} = \frac{1}{6}$
   $\mathbb{E}(X^2) = (-1)^2 \cdot \frac{5}{33} + 1^2 \cdot \frac{7}{12} \cdot \frac{6}{11} = \frac{31}{66}$
   $\mathrm{Var}(X) = \frac{175}{396}$
   б) $F(x) = \begin{cases} 0, & x < -1 \\ \frac{5}{33}, & -1 \leqslant x < 0 \\ \frac{15}{22}, & 0 \leqslant x < 1 \\ 1, & x \geqslant 1 \end{cases}$

2. а) $c = 1, \mathbb{P}(0.5 < X < 2) = 0.75, 0.5$

   б) $\mathbb{E}(X) = \frac{2}{3}, \operatorname{Var}(X) = \frac{1}{18}, \operatorname{Cov}(X, -X) = -\frac{1}{18}, \operatorname{Corr}(2X, 3X) = 1$

   в) $f(x) = \begin{cases} 0, & t < 0, t \geqslant 1 \\ 2t, & 0 \leqslant t < 1 \end{cases}$

3. а) $\operatorname{Corr}(X, Y) = -\frac{1}{3}$

   б) $\alpha = \frac{11}{17}$

   в) Нет

   г) Да

4. а) $\hat{p} = \frac{1}{20}$
   $\left[\frac{1}{20} - 1.65 \cdot \sqrt{\frac{\frac{1}{20}\cdot\frac{19}{20}}{40}}; \frac{1}{20} + 1.65 \cdot \sqrt{\frac{\frac{1}{20}\cdot\frac{19}{20}}{40}}\right]$

   б) $\mathbb{P}(|\hat{p} - p| \leqslant 0.1) = 0.9 \Rightarrow \mathbb{P}\left(\frac{|\hat{p}-p|}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \leqslant \frac{0.1}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}\right) = 0.9 \Rightarrow \frac{0.1}{\sqrt{\frac{\frac{1\cdot 19}{20^2}}{n}}} = 1.65 \Rightarrow n \approx 13$

5. а) $\gamma_{obs} = \frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2} \approx 1.32, \gamma_{crit,0.95} \approx 1.64$, оснований отвергать H_0 нет

   б) $X_1, \ldots, X_n \sim \mathcal{N}(\mu_X, \sigma_X^2), Y_1, \ldots, Y_n \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ — независимые выборки

   в) $\left[17.51 - 1.66 \cdot 59.4\sqrt{\frac{1}{40} + \frac{1}{60}}; 17.51 + 1.66 \cdot 59.4\sqrt{\frac{1}{40} + \frac{1}{60}}\right]$
   $[-2.61; 37.64]$
   $\hat{\sigma}_0^2 = \frac{\hat{\sigma}_X^2(n_X-1)+\hat{\sigma}_Y^2(n_Y-1)}{n_X+n_Y-2} \approx 59.4$

   г) Оснований считать новую методику более эффективной нет, так как 0 входит в доверительный интервал.

6. $\hat{p}_1 = \frac{57}{159}, \hat{p}_2 = \frac{48}{159}, \hat{p}_3 = \frac{54}{159}$
   $Q_{obs} = \sum_{i=1}^n \frac{(n_i - n\cdot p_i)^2}{n\cdot p_i} = \frac{42}{53}$. $Q_{crit} = 3.84$. Оснований отвергать нулевую гипотезу нет.

7. $\gamma_{obs} = \sum_{i=1}^s \sum_{j=1}^m \frac{\left(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{n}\right)^2}{\frac{n_{i\cdot} n_{\cdot j}}{n}} \approx 1.15, \gamma_{crit} = 3.84$, оснований отвергать H_0 нет.

8. а)

$$L(x_1, \ldots, x_n, \theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2}\cdot\frac{x_i^2}{\theta}} = \frac{1}{(\sqrt{2\pi\theta})^n} e^{-\frac{1}{2\theta}\sum_{i=1}^n x_i^2}$$

$$l(x_1, \ldots, x_n, \theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\theta - \frac{1}{2\theta}\sum_{i=1}^n x_i^2 \to \max_\theta$$

$$\frac{\partial l}{\partial \theta} = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n x_i^2$$

$$\hat{\theta}_{ML} = \frac{\sum_{i=1}^n x_i^2}{n}$$

   б) $\mathbb{E}(\hat{\theta}_{ML}) = \mathbb{E}\left(\frac{\sum_{i=1}^n x_i^2}{n}\right) = \frac{1}{n}\cdot n\,\mathbb{E}(x_1^2) = \theta$, оценка несмещённая.
   $\operatorname{Var}(\hat{\theta}_{ML}) = \operatorname{Var}\left(\frac{\sum_{i=1}^n x_i^2}{n}\right) = \frac{1}{n^2}\cdot n\operatorname{Var}(x_1^2) = \frac{3\theta^2 - \theta^2}{n} \to_{n\to\infty} 0$, оценка состоятельная.
   $\frac{\partial^2 l}{\partial\theta^2} = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^n x_i^2$
   $-\mathbb{E}\left(\frac{\partial^2 l}{\partial\theta^2}\right) = -\frac{n}{2\theta^2} + \frac{1}{\theta^3}\cdot n\theta = \frac{n}{2\theta^2}$
   $\operatorname{Var}(\hat{\theta}_{ML}) = \frac{2\theta^2}{n} = \frac{1}{\frac{n}{2\theta^2}} = I(\theta)$, оценка эффективная.

9. а) O1P: выбрали H_a, но верна H_0, то есть $X \sim [0, 100]$, но при этом $X \geqslant c$.
O2P: выбрали H_0, но верна H_a, то есть $X \sim [50, 150]$, но при этом $X > c$.

б) $\alpha = \begin{cases} 1, & c < 0 \\ \frac{100-c}{100}, & 0 \leqslant c \leqslant 100 \\ 0, & c > 100 \end{cases}$

$\beta = \begin{cases} 0, & c < 50 \\ \frac{c-50}{100}, & 50 \leqslant c \leqslant 150 \\ 1, & c > 150 \end{cases}$

10. Считаем через сумму рангов, $T = 2 + 4 + 6 + 8 + 10 = 30$.
$\mathbb{E}(T) = \frac{1}{2}(n_x(n_x + n_y + 1)) = 37.5$.
$\mathrm{Var}(T) = \frac{1}{12}(n_x n_y(n_x + n_y)) = \frac{75}{6}$
$\gamma_{obs} = \frac{30-37.5}{\sqrt{\frac{75}{6}}} \approx -2.12$, $\gamma_{crit,0.05} = -1.65$, основная гипотеза отвергается.
Можно было бы считать через U-статистику, у неё было бы другое ожидание, $n_x n_y/2$, но после стандартизации снова вышло бы -2.12.

17.13. 2005-2006

Нет, Нет, Да, Да, Да, Нет, Нет, Нет, Да, Нет, Нет, Да, Нет, Да, Нет, Да

1. а)
б) $\mathbb{E}(\hat{\theta}) = 1 \cdot \mathbb{P}(X < 3) + 0 \cdot \mathbb{P}(X \geqslant 3) = \theta$, да является
в) $\mathbb{E}\left(\left(\hat{\theta} - \theta\right)^2\right) = \mathbb{E}\left(\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2\right) \overset{\hat{\theta}^2 = \hat{\theta}}{=} \theta - 2\theta^2 + \theta^2 = \theta - \theta^2$

2. а) $L = (k+1)^n (x_1 \cdot x_2 \cdot \ldots \cdot x_n)^k$
$l = \ln L = n \ln(k+1) + k(\sum \ln x_i)$
$\frac{\partial l}{\partial k} = \frac{n}{k+1} + \sum \ln x_i$
$\frac{n}{\hat{k}+1} + \sum \ln x_i = 0$
$\hat{k} = -\left(1 + \frac{n}{\sum \ln x_i}\right)$
б) $\mathbb{E}(X_i) = \int t \cdot p(t) dt = \int_0^1 (k+1) t^{k+1} = \frac{k+1}{k+2}$
$\frac{\hat{k}+1}{\hat{k}+2} = \bar{X}$
$\hat{k} = \frac{2\bar{X}-1}{1-\bar{X}}$

3. $C = \sum \frac{(X_{i,j} - n\hat{p}_{i,j})^2}{n\hat{p}_{i,j}} \sim \chi^2_{(r-1)(c-1)}$
$C \sim \chi^2_1$
$C = 35$
Если $\alpha = 0.1$, то $C_{crit} = 2.706$.
Вывод: H_0 (гипотеза о независимости признаков) отвергается.

4. а) Число наблюдений велико, используем нормальное распределение.
$\mathbb{P}\left(-1,65 < \frac{\bar{X} - \bar{Y} - \triangle}{\sqrt{\frac{\hat{\sigma}_x^2}{40} + \frac{\hat{\sigma}_y^2}{50}}} < 1,65\right) = 0.9$
$\triangle \in 4 \pm 1.65\sqrt{\frac{49}{40} + \frac{64}{50}}$
$\triangle \in [1.4; 6.6]$

б) Используем результат предыдущего пункта: H_0 отвергается, так как число 0 не входит в доверительный интервал.

в) $Z = 2.505$ и $P_{value} = 0.0114$

5. а) $\chi^2_9 = \frac{9\hat{\sigma}^2}{\sigma^2} \in [4.17; 14.69]$
   $\sigma^2 \in [8822.3; 31080]$
   $\sigma \in [93.9; 176.3]$

   б) $\mathbb{P}(|\hat{\sigma}^2 - \sigma^2| < 0.4\sigma^2) = \mathbb{P}(0.6 < \frac{\hat{\sigma}^2}{\sigma^2} < 1.4) = \mathbb{P}(11.4 < \chi^2_{19} < 26.6) \approx 0.8$

6. а) $W_1 = 2 + 4 + 6 + 8 = 20$ или $W_2 = 1 + 3 + 5 + 7 + 9 + 10 = 35$
   $U_1 = 10$ или $U_2 = 14$
   $Z_1 = -0.43 = -Z_2$
   Вывод: H_0 (гипотеза об отсутствии сдвига между законами распределения) не отвергается

   б) Нет, т.к. наблюдения не являются парными.

7. $\mathbb{P}(-2.13 < t_4 < 2.13) = 0.9$
   $\mu \in 1560 \pm 2.13 \cdot \sqrt{\frac{670^2}{5}}$
   $\mu \in [921.8; 2198.2]$

8. а)

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

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

   Построение оставлено читателю в качестве самостоятельного упражнения :)

9. $\mathbb{P}(\sqrt{X^2 + Y^2} > 2.45) = \mathbb{P}(X^2 + Y^2 > 2.45^2) = \mathbb{P}(\chi^2_2 > 6) = 0.05$

10. а) 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) $\mathbb{P}(B|C) = \frac{\mathbb{P}(B)}{\mathbb{P}(C)} = \frac{0.7}{0.79} = \frac{70}{79}$

11-А. 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 (!!!).

То есть $AA_1 = AA_2 = \ldots$ и $Aa_1 = Aa_2 = \ldots$.

Вероятность получить 'A' 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: $p_1 = AA_0 + 0.5 \cdot Aa_0$

$AA_1 = p_1^2 = 0.36$, $Aa_1 = 2p_1(1 - p_1) = 0.48$.

$p_2 = AA_1 + 0.5 \cdot Aa_1 = p_1^2 + p_1(1 - p_1) = p_1$

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

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

18. Решения контрольной номер 3. ИП

18.1. 2018-2019

1. а) Пары ортогональных величин: (X_1, X_2); (X_2, X_3); (X_1, X_3); (S_2, X_3).

б) В качестве примера можно взять величину $X_2 - X_1$, поскольку

$$\begin{aligned}\operatorname{Cov}(X_2 - X_1, S_3) &= \operatorname{Cov}(X_2 - X_1, X_1 + X_2 + X_3)\\ &= \operatorname{Var}(X_1) - \operatorname{Var}(X_1) = 0\end{aligned}$$

в) Как известно из линейной алгебры, если вектор X_1^- — проекция X_1 на S_3, а $X_1^\perp$ — соответствующая ортогональная составляющая, то $X_1^\perp = X_1 - X_1^- = X_1 - \alpha S_3$, причем

$$\langle X_1^\perp, S_3\rangle = \operatorname{Cov}(X_1^\perp, S_3) = \operatorname{Cov}(X_1^\perp, X_1 - \alpha S_3) = 0,$$

откуда

$$\alpha = \frac{\operatorname{Cov}(X_1, S_3)}{\operatorname{Var}(S_3)}.$$

Тогда искомая проекция равна

$$\frac{\operatorname{Cov}(X_1, S_3)}{S_3} \cdot S_3$$

г) Искомая проекция равна ортогональной составляющей из предыдущего пункта:

$$X_1^\perp = X_1 - \frac{\operatorname{Cov}(X_1, S_3)}{S_3} \cdot S_3$$

д) Пусть соответствующие проекции равны $X_1^- = \alpha S_3$, $X_2^- = \beta S_3$. Тогда

$$\begin{aligned}\operatorname{Cov}(X_1^-, X_2^-) &= \operatorname{Cov}(X_1 - \alpha S_3, X_2 - \beta S_3)\\ &= -\alpha \operatorname{Var}(X_1) - \beta \operatorname{Var}(X_2) + 3\alpha\beta \operatorname{Var}(X_1)\end{aligned}$$

$$\begin{aligned}\operatorname{Var}(X_1^-) &= \operatorname{Var}(X_1 - \alpha S_3)\\ &= \operatorname{Var}(X_1) + 3\alpha^2 \operatorname{Var}(X_1) - 2\alpha \operatorname{Var}(X_1)\end{aligned}$$

Заметим, что дисперсии величин X_1 и X_2 равны, поэтому

$$\begin{aligned}\operatorname{Var}(X_2^-) &= \operatorname{Var}(X_2 - \beta S_3)\\ &= \operatorname{Var}(X_1) + 3\beta^2 \operatorname{Var}(X_1) - 2\beta \operatorname{Var}(X_1)\end{aligned}$$

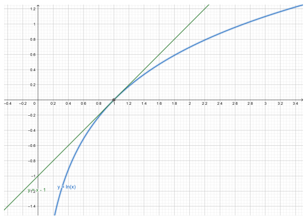
Таким образом, искомая частная корреляция равна

$$\frac{(3\alpha\beta - \alpha - \beta)\operatorname{Var}(X_1)}{\operatorname{Var}(X_1)\sqrt{(1 - 2\alpha + 3\alpha^2)(1 - 2\beta + 3\beta^2)}}$$

Вспоминаем, что

$$\alpha = \frac{\operatorname{Cov}(X_1, S_3)}{\operatorname{Var}(S_3)} = \frac{\operatorname{Var}(X_1)}{3\operatorname{Var}(X_1)} = \frac{1}{3} = \beta,$$

откуда искомая частная корреляция равна $-1/2$.



2. а) Как видно по графику, $x - 1 \geqslant \ln x$.

б) По свойству из предыдущего пункта,

$$\frac{q(x)}{p(x)} - 1 \geqslant \ln\left(\frac{q(x)}{p(x)}\right)$$
$$g(x) - p(x) \geqslant p(x)\ln q(x) - p(x)\ln p(x)$$
$$-p(x)\ln q(x) \geqslant -p(x)\ln p(x) + p(x) - q(x)$$

Интегрируем:

$$-\int_{-\infty}^{+\infty} p(x)\ln q(x)\ dx \geqslant -\int_{-\infty}^{+\infty} p(x)\ln p(x)\ dx + \int_{-\infty}^{+\infty} p(x)\ dx - \int_{-\infty}^{+\infty} q(x)\ dx$$

Поскольку $p(x), q(x)$ – функции плотности, интегрирование каждой из них дает единицу, а значит, последние два слагаемых в сумме равны нулю:

$$-\int_{-\infty}^{+\infty} p(x)\ln q(x)\ dx \geqslant -\int_{-\infty}^{+\infty} p(x)\ln p(x)\ dx$$

в) Функция плотности для величины $X \sim \mathcal{N}(\mu;\sigma^2)$ есть

$$q(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$
$$\ln q(x) = -\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{(x-\mu)^2}{2\sigma^2}$$
$$H(q) = -\int_{-\infty}^{+\infty} q(x)\ln q(x)\ dx$$

$$\begin{aligned}
H(q) &= \int_{-\infty}^{+\infty} q(x)\left(\frac{1}{2}\ln 2\pi + \ln\sigma + \frac{(x-\mu)^2}{2\sigma^2}\right)\ dx \\
&= \frac{1}{2}\ln 2\pi + \ln\sigma + \int_{-\infty}^{+\infty} q(x)\frac{(x-\mu)^2}{2\sigma^2}\ dx \\
&= \frac{1}{2}\ln 2\pi + \ln\sigma + \int_{-\infty}^{+\infty} \frac{\left(q(x)x^2 - 2x\mu q(x) + \mu^2 q(x)\right)}{2\sigma^2}\ dx \\
&= \frac{1}{2}\ln 2\pi + \ln\sigma + \frac{1}{2\sigma^2}\left(\mathbb{E}(X^2) - 2\mu\,\mathbb{E}(X) + \mu^2\right) \\
&= \frac{1}{2}\ln 2\pi + \ln\sigma + \frac{1}{2\sigma^2}\left(\sigma^2 + \mu^2 - 2\mu^2 + \mu^2\right) \\
&= \frac{1}{2}\ln 2\pi + \ln\sigma + \frac{1}{2}
\end{aligned}$$

г) Пусть $Y \sim \mathcal{N}(\mu_N;\sigma_N^2)$, ее функция плотности равна $q(x)$. Тогда, аналогично пункту в), получаем:

$$\begin{aligned}
CE_p(q) &= \frac{1}{2}\ln 2\pi + \ln\sigma_N + \int_{-\infty}^{+\infty} \frac{\left(p(x)x^2 - 2x\mu_N p(x) + \mu_N^2 p(x)\right)}{2\sigma_N^2}\ dx \\
&= \frac{1}{2}\ln 2\pi + \ln\sigma_N + \frac{1}{2\sigma_N^2}\left(\sigma^2 + \mu^2 - 2\mu\mu_N + \mu_N^2\right)
\end{aligned}$$

3. а)

$$L = \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2}$$
$$\ln L = 2\ln\lambda - \lambda(x_1 + x_2)$$
$$(\ln L)'_\lambda = \frac{2}{\hat\lambda} - (x_1 + x_2) = 0$$
$$\hat\lambda = \frac{2}{x_1 + x_2} = \frac{1}{15}$$

б) Пусть X – время обслуживания одного клиента, тогда

$$\mathbb{P}(X > 30) = \int_{30}^{+\infty} \lambda e^{-\lambda x}\, dx \Rightarrow$$
$$L = \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2} \prod_{i=3}^{10} \int_{30}^{+\infty} \lambda e^{-\lambda x_i}\, dx_i$$
$$\int_{30}^{+\infty} \lambda e^{-\lambda x}\, dx = -e^{-\lambda x}|_{30}^{+\infty} = e^{-30\lambda}$$
$$L = \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2} \prod_{i=3}^{10} e^{-30\lambda}$$
$$\ln L = 2\ln\lambda - \lambda(x_1 + x_2) - 240\lambda$$
$$(\ln L)'_\lambda = \frac{2}{\hat\lambda} - (x_1 + x_2) - 240 = 0$$
$$\hat\lambda = \frac{1}{135}$$

а)

$$\mathbb{P}(X < 30) = 1 - \int_{30}^{+\infty} \lambda e^{-\lambda x}\, dx = 1 - e^{-30\lambda} \Rightarrow$$
$$L = \prod_{i=1}^{2} (1 - e^{-30\lambda}) \prod_{i=3}^{10} e^{-30\lambda}$$
$$\ln L = 2\ln(1 - e^{-30\lambda}) - 240\lambda$$
$$(\ln L)'_\lambda = \frac{2 \cdot 30 e^{-30\hat\lambda}}{1 - e^{-30\hat\lambda}} - 240 = 0$$
$$\hat\lambda \approx 0,0074$$

г)

$$L = \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2} \prod_{i=3}^{10} e^{-20\lambda}$$
$$\ln L = 2\ln\lambda - \lambda(x_1 + x_2) - 160\lambda$$
$$\hat\lambda = \frac{1}{95}$$

4. У всех одинаковые нулевые шансы стать Самым Главным, что можно доказать методом математической индукции. Пусть всего метеорологов двое. Тогда, если у них монеты упали одной стороной, они оба выбывают, если разными, то игра продолжается. Если метеорологов трое и у всех монеты выпали одной стороной, то все выходят из игры; если у одного из них выпала решка, а у остальных – орлы, то первый выбывает, получаем ситуацию для двух и так далее.

5. а) Обозначим количества орлов в соответствующие дни за S_1, S_2, S_3, а количества бросков – за n_1, n_2, n_3. По своим данным Анна может построить оценку $\hat{p}_a = (S_1 + S_2)/(n_1 + n_2)$, а Белла – $\hat{p}_b = (S_2 + S_3)/(n_2 + n_3)$. Наша цель – минимизировать по α дисперсию взвешенной оценки:

$$\text{Var}(\hat{p}) = \text{Var}(\alpha\hat{p}_a + (1-\alpha)\hat{p}_b)$$

Или явно:

$$p(1-p)\left(\frac{\alpha}{n_1+n_2} + \frac{(1-\alpha)^2}{n_2+n_3} + \frac{2\alpha(1-\alpha)n_2}{(n_1+n_2)(n_2+n_3)}\right)$$

$$\alpha^* = \frac{n_1}{n_1+n_3}$$

Оценки: $\hat{p}_a = 0.4$, $\hat{p}_b = 0.5$, $\alpha = 0.4$, $\hat{p} = 0.46$.

б) Подставим найденную α^* в формулу для дисперсии и получим

$$\frac{n_1n_2 + n_2n_3 + n_1n_3}{(n_1+n_2)(n_1+n_3)(n_2+n_3)}p(1-p)$$

При наших данных $\hat{\text{Var}}(\hat{p}) \approx 0.0458$.

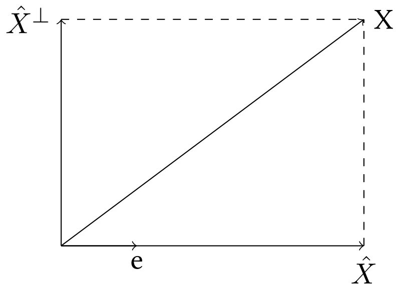
в) Используя стандартную формулу, получим:

$$[\hat{p} - 1.96se(\hat{p}); \hat{p} + 1.96se(\hat{p})],$$

где $\hat{p} = 0.46$ и $se(\hat{p}) \approx 0.214$.

18.2. 2017-2018

1а) - в) См. картинку :)



г) $\hat{X} = e \cdot \bar{X}$
$\|\hat{X}\| = \sqrt{n} \cdot \bar{X}$
$\hat{X}^\perp = X - e \cdot \bar{X} = (X_1 - \bar{X}, \ldots, X_n - \bar{X})$
$\|\hat{X}^\perp\| = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}$

д) $\|X\|^2 = \|\hat{X}^\perp\|^2 + \|\hat{X}\|^2$
$\sum_{i=1}^n X_i^2 = \sum_{i=1}^n (X_i - \bar{X})^2 + n\bar{X}^2$

е) t-статистика для построения доверительного интервала для μ имеет вид:

$$t = \frac{\bar{X} - \mu}{\sqrt{\bar{\sigma}^2/n}} = \frac{\bar{X} - \mu}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2/(n \cdot (n-1))}}$$
$$= \sqrt{n-1} \cdot \frac{\sqrt{n} \cdot (\bar{X} - \mu)}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}} = \sqrt{n-1} \cdot \frac{\|\hat{X}\| - \sqrt{n} \cdot \mu}{\|\hat{X}^\perp\|}$$

Заметим, что $\operatorname{ctg}\alpha$ есть отношение прилежащего катета к противолежащему, таким образом, нужный нам угол α образуется между векторами X и $\hat{X}$. Зметим однако, что в нашем случае

$$t = \sqrt{n-1} \cdot \operatorname{ctg}\alpha = \frac{\|\hat{X}\|}{\|\hat{X}^{\perp}\|},$$

то есть наша статистика подойдёт только для проверки гипотезе о равенстве матожидания нулю.
Замечание. $t = \sqrt{n-1} \cdot \operatorname{ctg}\alpha$ будет t-статистикой только в том случае, если X_i будут н.о.р.с.в. с нормальным распределением, о чём в условие сказано не было.

2. Выпишем функцию правдоподобия для выборки из трёх видов, два из которых совпадают. Первый медведепришелец будет нового вида с вероятностью 1. Вероятность, что вид второго пойманного медведепришельца совпадёт с первмым, составляет $1/n$. После этого нужно поймать медведепришельца нового вида – это произойдёт с вероятностью $(n-1)/n$, и ещё одного нового вида – вероятность этого $(n-2)/n$. Поскольку медведепришелец, вид которого встречается дважды, мог встретить на любой из трёх позиций, функцию правдоподобия необходимо домножить на C_3^1. Таким образом, функция правдоподобия имеет вид:

$$L(n) = C_3^1 \cdot 1 \cdot \frac{1}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n}, n \geqslant 3.$$

Максимизируя её, внутри области определения получаем $\hat{n} = 5$.

Так как количество медведей велико и все они встречаются равновероятно, то $p_1 = p_2 = p_3 = 1/n$. Так же из выборки известно, что число видов космомедведей не меньше трёх. Потому $\hat{n} \geqslant 3$.

Найдите хитрую ошибку в предложенном решении:
$L(n) = \left(\frac{1}{n}\right)^2 \cdot \frac{1}{n} \cdot \frac{1}{n} = n^{-4}$
$\frac{\partial L(n)}{\partial n} = -4 \cdot n^{-5} = 0$
Данное уравнение не имеет решений при конечных n, но заметим, что при всех $n \geqslant 3$ выполняется $\frac{\partial L(n)}{\partial n} = -4 \cdot n^{-5} < 0$, таким образом максимальное значение находится в граничных точках.
$\lim\limits_{n\to\infty} \frac{1}{n^{-4}} = 0 < \frac{1}{3^{-4}}$
Таким образом получаем, что $\hat{n} = 3$.

3. а) $L(p_1, p_2) = p_1^{150} \cdot p_2^{100} \cdot (1 - p_1 - p_2)^{50}$
$\ell(p_1, p_2) = 150 \ln p_1 + 100 \ln p_2 + 50 \ln(1 - p_1 - p_2)$

$$\begin{cases} \frac{\partial \ell(p_1,p_2)}{\partial p_1} = \frac{150}{p_1} - \frac{50}{1-p_1-p_2} = 0 \\ \frac{\partial \ell(p_1,p_2)}{\partial p_2} = \frac{100}{p_2} - \frac{50}{1-p_1-p_2} = 0 \end{cases}$$

Откуда получаем:

$$\begin{cases} \hat{p}_1 = 1/2 \\ \hat{p}_2 = 1/3 \end{cases}$$

б) Найдём, какие значения должны стоять в теоретической ковариационной матрице. Заметим, что случайная величина найти кальмаромедведя (X) или двурога (Y) есть бернулевская случайная величина с параметром $p_{1+2} = p_1 + p_2$ и дисперсией $p_{1+2} \cdot (1 - p_{1+2})$, но тогда:
$\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2 \cdot \operatorname{Cov}(X,Y)$
$\operatorname{Cov}(X,Y) = \frac{1}{2} \cdot (\operatorname{Var}(X+Y) - \operatorname{Var}(X) - \operatorname{Var}(Y)) = \frac{1}{2} \cdot ((p_1+p_2) \cdot (1-p_1-p_2) - p_1 \cdot (1-p_1) - p_2 \cdot (1-p_2)) = -p_1 \cdot p_2$
Тогда подставляя в теоретическую ковариационную матрицу оценки параметров и домнажая всё на $1/300$, так как $\hat{p}_1$ и $\hat{p}_2$ являются средними, получим:

$$\operatorname{Var}(\hat{p}) = \frac{1}{n} \begin{pmatrix} \hat{p}_1 \cdot (1 - \hat{p}_1) & -\hat{p}_1 \cdot \hat{p}_2 \\ -\hat{p}_1 \cdot \hat{p}_2 & \hat{p}_1 \cdot (1 - \hat{p}_1) \end{pmatrix} = \frac{1}{300} \begin{pmatrix} 1/4 & -1/6 \\ -1/6 & 2/9 \end{pmatrix}$$

в) Для начала, найдём теоретическую дисперсию $\mathrm{Var}(X-Y)$.

$$\mathrm{Var}(X-Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\cdot\mathrm{Cov}(X,Y) = p_1\cdot(1-p_1) + p_2\cdot(1-p_2) + 2\cdot p_1\cdot p_2$$

Тогда подставляя оценки для p_1 и p_2 и учитывая, что это оценки среднего, получим оценку:

$$\mathrm{Var}(\hat{p}_1-\hat{p}_2) = 1/300\cdot(1/4+2/9+2\cdot 1/6) = 29/(36\cdot 300)$$

г) Так как выборка достаточно велика, то статистика $\hat{p}_1-\hat{p}_2$, являясь средним, будет иметь примерно нормальное распределение, и тогда:
$\frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\mathrm{Var}(\hat{p}_1-\hat{p}_2)}} \sim \mathcal{N}(0,1)$
$\hat{p}_1-\hat{p}_2-z_{1-\frac{\alpha}{2}}\cdot\sqrt{\mathrm{Var}(\hat{p}_1-\hat{p}_2)} \leqslant p_1-p_2 \leqslant \hat{p}_1-\hat{p}_2-z_{\frac{\alpha}{2}}\cdot\sqrt{\mathrm{Var}(\hat{p}_1-\hat{p}_2)}$

4. а) $\hat{\alpha} = \bar{X} + \sqrt{\bar{X}+6} = 10 + \sqrt{10+6} = 14$

б) Так как $\bar{X}$ сходится по распределению к нормальному распределению и $\hat{\alpha} = g(\bar{X})$, где $g(\bar{X})$ гладкая по $\bar{X}$ функция при $\bar{X} \geqslant 0$, а также $\bar{X}$ сходится по вероятности к матожиданию, то можно абсолютно спокойно применить дельта-метод. Тогда:

$$(\alpha - g(\bar{X})) \sim N(0;\sigma^2(g'(\mathbb{E}(X_1)))^2/n)$$

Но так как $\hat{\alpha}$ является состоятельной оценкой, то можно заменить $g'(\mathbb{E}(X_1))$ на $g'(\bar{X})$:

$$g'(\bar{X}) = 1 + \frac{1}{2\cdot\sqrt{\bar{X}+6}} = 1 + \frac{1}{2\cdot 4} = \frac{9}{8} = 1.125$$

и тогда можно построить ассимтотический доверительный интервал:

$$\begin{aligned}
\hat{\alpha} - z_{97.5\%}\cdot\sqrt{\sigma^2\cdot(g'(\bar{X}))^2/n} \leqslant &\alpha \leqslant \hat{\alpha} - z_{2.5\%}\cdot\sqrt{\sigma^2\cdot(g'(\bar{X}))^2/n}\\
16 - 1.96\cdot 2\cdot 9/(8\cdot 10) \leqslant &\alpha \leqslant 16 + 1.96\cdot 2\cdot 9/(8\cdot 10)\\
13.559 \leqslant &\alpha \leqslant 14.441
\end{aligned}$$

5. а) Так как не известно точно, кто сколько фотографий сделал, и так как метод оценки не указан, то воспользуемся методом моментов для построения оценки.

$$\begin{aligned}
N &= \mathbb{E}(\text{«фото Андрея»}) + \mathbb{E}(\text{«фото Беллы»})\\
130 &= 100\cdot 0.5 + p\cdot 100\\
\hat{p} &= 0.8
\end{aligned}$$

Так как выборка достаточно велика, то $\frac{\hat{p}-p}{\sqrt{\hat{p}\cdot(1-\hat{p})/W}} \sim \mathcal{N}(0,1)$

$$\begin{aligned}
\hat{p} - z_{97.5\%}\sqrt{\hat{p}\cdot(1-\hat{p})/W} \leqslant &p \leqslant \hat{p} - z_{2.5\%}\sqrt{\hat{p}\cdot(1-\hat{p})/W}\\
0.8 - 1.96\cdot\sqrt{0.8\cdot 0.2/100} \leqslant &p \leqslant 0.8 + 1.96\cdot\sqrt{0.8\cdot 0.2/100}\\
0.72 \leqslant &p \leqslant 0.88
\end{aligned}$$

б) Так как неизвестно, кто больше снимков сделал, то рассмотрим два случая: Андрей сделал 60 фото и Белла — 70 фото, Андрей сделал 70 фото и Белла — 60 фото. В каждом случае при помощи метода максимального правдоподобия оценим вероятность p, после чего сравним значения функции правдоподобия с оценёнными параметрами для каждого случая.

$$\begin{aligned}
L(p) &= C_{100}^{60} \cdot 0.5^{60} \cdot 0.5^{40} \cdot C_{100}^{70} \cdot p^{70} \cdot (1-p)^{30} \\
\ell(p) &= const + 70 \ln p + 30 \ln(1-p) \\
\frac{\partial \ell(p)}{\partial p} &= \frac{70}{p} - \frac{30}{1-p} = 0 \\
\hat{p}_1 &= 0.7
\end{aligned}$$

Аналогично для второго случая получим оценку: $\hat{p}_2 = 0.6$.
Для простоты, будем сравнивать логарифмическии функции правдоподобия $\ell_1(p_1)$ и $\ell_2(p_2)$ и тогда получим:

$$\begin{aligned}
\ell_1(p_1) &= const + 70 \ln 0.7 + 30 \ln 0.3 \approx const - 70 \cdot 0.357 - 30 \cdot 1.204 = const - 61.11 \\
\ell_2(p_2) &= const + 60 \ln 0.6 + 40 \ln 0.4 \approx const - 60 \cdot 0.511 - 40 \cdot 0.916 = const - 67.3
\end{aligned}$$

Так как $-67.3 < -61.11$, то более вероятно, что $\hat{p} = 0.7$
Тогда анологично предыдущему пункту получим доверительный интервал:

$$0.7 - 1.96 \cdot \sqrt{0.7 \cdot 0.3/100} \leqslant p \leqslant 0.7 + 1.96 \cdot \sqrt{0.7 \cdot 0.3/100}$$
$$0.61 \leqslant p \leqslant 0.79$$

19. Решения контрольной номер 4

19.1. 2018-2019

19.2. 2017-2018

1. Проверяем следующую гипотезу:
$$\begin{cases} H_0 : \mu_D = 100 \\ H_a : \mu_D > 100 \end{cases}$$
Считаем наблюдаемое значение статистики:
$$t_{obs} = \frac{\bar{X} - \mu_D}{\frac{\sigma_D}{\sqrt{n_D}}} = \frac{136 - 100}{\frac{55}{\sqrt{40}}} \approx 4.14$$
При верной H_0 t-статистика имеет распределение t_{40-1}, значит, $t_{crit} \approx 1.68$. Поскольку $t_{crit} > t_{obs}$, основная гипотеза отвергается, $p - value \approx 0$.

2. Проверяем следующую гипотезу:
$$\begin{cases} H_0 : \sigma_D^2 = \sigma_T^2 \\ H_a : \sigma_D^2 \neq \sigma_T^2 \end{cases}$$
Считаем наблюдаемое значение статистики:
$$F_{obs} = \frac{\hat{\sigma}_D^2}{\hat{\sigma}_T^2} = \frac{55^2}{60^2} \approx 0.84$$
При верной H_0 F-статистика имеет распределение $F_{40-1,60-1}$. Находим критические значения: $F_{left} \approx 0.6$, $F_{right} \approx 1.6$. Поскольку $F_{left} < F_{obs} < F_{right}$, нет оснований отвергать H_0.

3. Проверяем гипотезу

$$\begin{cases} H_0 : \mu_D = \mu_T \\ H_a : \mu_D < \mu_T \end{cases}$$

Когда n_D, n_T велики,

$$\frac{\bar{D} - \bar{T} - (\mu_D - \mu_T)}{\sqrt{\frac{\sigma_D^2}{n_D} + \frac{\sigma_T^2}{n_T}}} \overset{H_0}{\sim} \mathcal{N}(0, 1)$$

Считаем наблюдаемое значение статистики:

$$z_{obs} = \frac{136 - 139}{\sqrt{\frac{3025}{40} + \frac{3600}{60}}} \approx -0.25$$

По таблице находим $z_{crit} = -1.28$. Так как $z_{crit} < z_{obs}$, нет оснований отвергать H_0.

б) Когда считаем дисперсии одинаковыми, то:

$$\hat{\sigma}_0^2 = \frac{\hat{\sigma}_D^2(n_D - 1) + \hat{\sigma}_T^2(n_T - 1)}{n_D + n_T - 2} = \frac{3025 \cdot 39 + 3600 \cdot 59}{30 + 60 - 2} \approx 3371$$

и

$$\frac{\bar{D} - \bar{T} - (\mu_D - \mu_T)}{\hat{\sigma}_0^2 \sqrt{\frac{1}{n_D} + \frac{1}{n_T}}} \overset{H_0}{\sim} t_{n_D + n_T - 2}$$

Считаем наблюдаемое значение статистики:

$$t_{obs} = \frac{136 - 139}{\sqrt{3371} \sqrt{\frac{1}{40} + \frac{1}{60}}} \approx -0.25$$

По таблице находим критическое значение: $t_{crit} \approx -1.29$. Поскольку $t_{crit} < t_{obs}$, нет оснований отвергать H_0.

4. а) Сначала найдём оценку максимального правдоподобия параметра λ:

$$L = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

$$\ell = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda}\Big|_{\lambda = \hat{\lambda}} = 0$$

$$\frac{\partial^2 \ell}{\partial \lambda^2} = -\frac{n}{\lambda^2}$$

$$\hat{\lambda} = \frac{1}{0.52}$$

Так как

$$\frac{\hat{\lambda} - \lambda}{\sqrt{\frac{1}{I(\lambda)}}} \overset{as}{\sim} \mathcal{N}(0, 1),$$

доверительный интервал имеет вид

$$\frac{1}{0.52} - 1.96 \frac{1}{\frac{10}{0.52}} < \lambda < \frac{1}{0.52} + 1.96 \frac{1}{\frac{10}{0.52}}$$

б) Найдём вероятность того, что наушник проработает без сбоев 45 минут:

$$g(\lambda) = \mathbb{P}(X > 0.75) = 1 - F(0.75) = e^{-0.75\lambda}$$

Тогда

$$g(\hat{\lambda}) = e^{-0.75/0.52}$$
$$g'(\hat{\lambda}) = -0.75e^{-0.75/0.52}$$

И доверительный интервал имеет вид:

$$e^{-0.75/0.52} - 1.96 \cdot \frac{0.75 \cdot 0.52}{10} \cdot e^{-0.75/0.52} < g(\lambda) < e^{-0.75/0.52} + 1.96 \cdot \frac{0.75 \cdot 0.52}{10} \cdot e^{-0.75/0.52}$$

5. Выпишем функцию правдоподобия:

$$L = p_1^{10} \cdot p_2^{10} \cdot p_3^{15} \cdot p_4^{15} \cdot p_5^{25} \cdot (1 - p_1 - p_2 - p_3 - p_4 - p_5)^{25}$$
$$\ell = 10 \ln p_1 + 10 \ln p_2 + 15 \ln p_3 + 15 \ln p_4 + 25 \ln p_5 + 25 \ln(1 - p_1 - p_2 - p_3 - p_4 - p_5)$$

Максимизируя логарифмическую функцию правдоподобия по всем параметрам, получим следующие оценки для неограниченной модели:

$$\hat{p}_1 = \hat{p}_2 = 0.1$$
$$\hat{p}_3 = \hat{p}_4 = 0.15$$
$$\hat{p}_5 = 0.25$$

Подставив найденные значения в логарифмическую функцию правдоподобия, получим

$$\ell_{UR} \approx -172$$

В ограниченной модели $p_1 = \ldots = p_6 = 1/6$, и значение функции правдоподобия будет

$$\ell_R \approx -179$$

Теперь можно посчитать наблюдаемое значение:

$$LR = 2(\ell_{UR} - \ell_R) = 2(-172 - (-179)) = 14$$

Критическое значение $\chi_{0.95,5} =\approx 11 < 14$, значит, основная гипотеза отвергается.

19.3. 2016-2017

1. а) $t_{obs} = \frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}} = \frac{9.5-10}{0.5/\sqrt{100}} = -10$
   В таблице для $t_{0.975;100-1}$ находим $-t_{crit} = -1.66$.
   Поскольку $t_{obs} < -t_{crit}$, основная гипотеза отвергается.
   б) $p - value \approx 0$
   в) $X_1, \ldots, X_{100} \sim \mathcal{N}(\mu, \sigma^2)$
   г) $\gamma_{obs} = \frac{\hat{\sigma}^2}{\sigma_0^2}(n-1) = \frac{0.5^2}{0.3}(100-1) = 82.5$
   В таблице находим нужные значения $\chi^2_{0.975;100-1}$, $\gamma_{crit,r} = 128$ и $\chi^2_{0.025;100-1}$, $\gamma_{crit,l} = 73$.
   Так $\gamma_{crit,l} < \gamma_{obs} < \gamma_{crit,r}$, нет оснований отвергать H_0.

2. $\hat{p} = \frac{16}{40} = 0.4$, проверять будем двустороннюю гипотезу.
   $z_{obs} = \frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} = \frac{0.4-0.5}{\sqrt{0.4 \cdot 0.6/40}} \approx -1.3$
   В таблице нормального распределения находим значение $z_{0.975}$, $z_{crit} = 1.96$.
   Так как $|z_{obs}| < z_{crit}$, нет оснований отвергать H_0.

3. а) $\hat\sigma_0^2 = \frac{\hat\sigma^2_\alpha(n_\alpha-1)+\hat\sigma^2_\beta(n_\beta-1)}{n_\alpha+n_\beta-2} = \frac{0.25\cdot19+0.36\cdot24}{20+25-2} = 0.31$
$\frac{\bar X-\bar Y}{\hat\sigma_0\sqrt{\frac{1}{n_\alpha}+\frac{1}{n_\beta}}} \sim t_{n_\alpha+n_\beta-2}$
$t_{obs} = \frac{9.5-9.8}{0.56\sqrt{\frac{1}{20}+\frac{1}{25}}} = -1.79$
$t_{crit} = 2.02$
Поскольку $|t_{obs}| < t_{crit}$, нет оснований отвергать H_0.
б) Выборки независимы, дисперсии неизвестны, но равны, $X_1,\ldots,X_{n_\alpha} \sim \mathcal{N}\left(\mu_X,\sigma^2_X\right)$, $X_1,\ldots,Y_{n_\beta} \sim \mathcal{N}\left(\mu_Y,\sigma^2_Y\right)$.
в) $\frac{\hat\sigma^2_\alpha}{\hat\sigma^2_\beta} \sim F_{n_{\alpha-1},n_{\beta-1}}$
$F_{obs} = \frac{0.5^2}{0.6^2} \approx 0.69$
$F_{crit,0.975} = 2.35$, $F_{crit,0.025} = 0.41$
Поскольку $F_{crit,0.025} < F_{obs} < F_{crit,0.975}$, нет оснований отвергать H_0.

4. $H_0 : p = 0.5$, где p — вероятность того, что бутерброд упадёт маслом вниз.
$\hat p = \frac{105}{200} = 0.525$
$z_{obs} = \frac{\hat p - p}{\sqrt{\hat p(1-\hat p)/n}} = \frac{0.525-0.5}{\sqrt{0.525\cdot0.475/200}} \approx 0.7$
$z_{crit} = 1.96$
Так как $z_{obs} < z_{crit}$, нет оснований отвергать H_0.

5. $LR \sim \chi^2_1$, так как в основной гипотезе одно уравнение. Выпишем функцию правдоподобия и найдём $\hat\mu_{ML}$ и $\hat\nu_{ML}$.

$$L = \prod_{i=1}^{100} \frac{1}{\sqrt{2\pi\nu}} \exp\left(-\frac{1}{2}\frac{(x_i-\mu)^2}{\nu}\right) = \frac{1}{(\sqrt{2\pi\nu})^{100}} \exp\left(-\frac{1}{2\nu}\sum_{i=1}^{100}(x_i-\mu)^2\right)$$

$$\ell = -\frac{100}{2}\ln(2\pi) - \frac{100}{2}\ln\nu - \frac{1}{2\nu}\sum_{i=1}^{100}(x_i-\mu)^2$$

$$\frac{\partial\ell}{\partial\mu} = \frac{1}{\nu}\sum_{i=1}^{100}(x_i-\mu) \Rightarrow \hat\mu_{ML} = \frac{\sum_{i=1}^{100} x_i}{100} = 0.3$$

$$\frac{\partial\ell}{\partial\nu} = -\frac{100}{2\nu} + \frac{1}{2\nu^2}\sum_{i=1}^{100}(x_i-\mu)^2 \Rightarrow \hat\nu_{ML} = 1.37$$

Тогда $LR = 2(\ell(\hat\mu_{ML},\hat\nu_{ML}) - \ell(\hat\mu_{ML},\nu=1))$ имеет вид:

$$Q_{obs} = 2\left(-50\ln(2\pi) - 50\ln 1.37 - \frac{1}{2\cdot1.37}\cdot137 + 50\ln(2\pi) + 0 + \frac{1}{2}\cdot137\right) \approx 5.5$$

Из таблицы: $Q_{crit} = 3.84$. Поскольку $Q_{obs} > Q_{crit}$, основная гипотеза отвергается.

19.4. 2015-2016

2. а)

$$L(x,\lambda) = \prod_{i=1}^{250} e^{-\lambda}\frac{\lambda^{x_i}}{x_i!} = e^{-250\lambda}\lambda^{\sum_{i=1}^{250}x_i}\prod_{i=1}^{250}\frac{1}{x_i!}$$

$$\ell(x,\lambda) = -250\lambda + \ln\lambda\sum_{i=1}^{250}x_i - \sum_{i=1}^{250}\ln x_i!$$

$$\frac{\partial\ell}{\partial\lambda} = -250 + \frac{1}{\lambda}\sum_{i=1}^{250}x_i$$

$$\hat\lambda_{ML} = \bar X$$

б) $\mathbb{E}\left(\hat{\lambda}_{ML}\right) = \mathbb{E}\left(\bar{X}\right) = \lambda \Rightarrow$ оценка несмещённая.

$\operatorname{Var}\left(\hat{\lambda}_{ML}\right) = \operatorname{Var}\left(\bar{X}\right) = \frac{1}{n^2} \cdot n \operatorname{Var}(X_1) = \frac{\lambda}{n} \to_{n\to\infty} 0 \Rightarrow$ оценка состоятельная.

$\frac{\partial^2 \ell}{\partial \lambda^2} = -\frac{1}{\lambda^2}\sum_{i=1}^n x_i$, $I(\lambda) = -\mathbb{E}\left(-\frac{1}{\lambda^2}\sum_{i=1}^n x_i\right) = \frac{n}{\lambda}$. Так как $\operatorname{Var}\left(\hat{\lambda}_{ML}\right) = \frac{1}{I(\lambda)}$, оценка является эффективной.

в) $\mathbb{P}(X=0) = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-\lambda} \Rightarrow \widehat{\mathbb{P}(X=0)} = e^{-\hat{\lambda}} = e^{-\bar{X}}$

г) В данном случае: $g\left(\hat{\lambda}\right) = e^{-\hat{\lambda}}$, $g'\left(\hat{\lambda}\right) = -e^{-\hat{\lambda}}$. И доверительный интервал имеет вид:

$$\left[e^{-\bar{X}} - 1.96\sqrt{\frac{e^{-2\bar{X}}\bar{X}}{n}}; e^{-\bar{X}} + 1.96\sqrt{\frac{e^{-2\bar{X}}\bar{X}}{n}}\right]$$

3. а) $\hat{p} \overset{as.}{\sim} \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$

О1Р: лекарство помогает в 80% случаев, но в данной выборке оно помогло менее чем 12 людям.

$\alpha = \mathbb{P}(\text{О1Р}) = \mathbb{P}\left(\hat{p} < \frac{12}{20} \middle| p = 0.8\right) = \mathbb{P}\left(\frac{\hat{p}-0.8}{\sqrt{\frac{0.8\cdot 0.2}{20}}} < \frac{\frac{12}{20}-0.8}{\sqrt{\frac{0.8\cdot 0.2}{20}}}\right) = 0.0125$

б) О2Р: лекарство помогает в 60% случаев, но $Y \geqslant 12$.
$\hat{p} \sim \mathcal{N}\left(0.6, \frac{0.6\cdot 0.4}{20}\right)$
$\beta = \mathbb{P}\left(\hat{p} \geqslant \frac{12}{20}\right) = \frac{1}{2}$

в) $\mathbb{P}(Z < a) = 0.1$, из таблицы находим, что $a = -1.28$.

$$a = \frac{\frac{c}{20} - 0.8}{\sqrt{\frac{0.8\cdot 0.2}{20}}} = -1.28 \Rightarrow c \approx 13.7$$

г) $\mathbb{P}(|\hat{p} - p| \leqslant 0.01) \geqslant 0.95$, будем считать, что $p = 0.6$.

$$\mathbb{P}(|\hat{p} - p| \leqslant 0.01) = \mathbb{P}(-0.01 \leqslant \hat{p} - p \leqslant 0.01) = \mathbb{P}\left(-\frac{0.01}{\sqrt{\frac{0.6\cdot 0.4}{n}}} \leqslant Z \leqslant \frac{0.01}{\sqrt{\frac{0.6\cdot 0.4}{n}}}\right) = 0.95$$

Из таблицы находим
$$\frac{0.01}{\sqrt{\frac{0.6\cdot 0.4}{n}}} = 1.96 \Rightarrow n = \frac{0.6\cdot 0.4\cdot 1.96^2}{0.01^2}$$

4. $H_0 : p_{\text{с}} = \frac{1}{7}, p_{\text{л}} = \frac{2}{7}, p_{\text{к}} = \frac{4}{7}$

$Q = \sum_{i=1}^{s=3} \frac{(\nu_i - np_i)^2}{np_i} \sim \chi^2_{s-k-1} = \chi^2_2$

$Q_{obs} = \frac{\left(10 - 50\frac{1}{7}\right)^2}{50\frac{1}{7}} + \frac{\left(1 - 50\frac{2}{7}\right)^2}{50\frac{2}{7}} + \frac{\left(39 - 50\frac{4}{7}\right)^2}{50\frac{4}{7}} = 17.29$

$Q_{crit} = 5.99$, $Q_{crit} < Q_{obs} \Rightarrow$ гипотеза отвергается

5. $LR \sim \chi^2_1$, так как основная гипотеза содержит одно уравнение

$L(x, \lambda) = \prod_{i=1}^{n=50} \lambda e^{-\lambda x} = \lambda^{50} e^{-\lambda \sum_{i=1}^{n=50} x_i}$

$\ln L(x, \lambda) = 50 \ln \lambda - \lambda \sum_{i=1}^{n=50} x_i \to \max_\lambda$

$\frac{\partial \ln L}{\partial \lambda} = \frac{50}{\lambda} - \sum_{i=1}^{n=50} x_i \mid_{\lambda = \hat{\lambda}} = 0 \Rightarrow \hat{\lambda}_{ML} = \frac{1}{\bar{X}} = \frac{10}{11}$

При верной $H_0 : \lambda = 1$, тогда $\ln L(\lambda = 1) = 50 \ln 1 - 1 \cdot 1.1 \cdot 50 = -55$

При верной $H_1 : \lambda = \lambda_{ML}$, тогда $\ln L\left(\lambda = \frac{10}{11}\right) = 50 \ln \frac{10}{11} - \frac{10}{11} \cdot 50 \cdot 1.1 = -54.77$

$LR_{obs} = 2(\ln L(H_1) - \ln(H_0)) = 2(-54.77 - (-55)) = 0.46$

$LR_{crit} = 2.71$, $LR_{crit} > LR_{obs} \Rightarrow$ оснований отвергать H_0 нет

6. Будем проверять гипотезы на уровне значимости 0.05

   а) $\hat{\sigma}^2_{\text{в}} = 484$, $\hat{\sigma}^2_{\text{р}} = 400$
   $\frac{\hat{\sigma}^2_{\text{в}}}{\hat{\sigma}^2_{\text{р}}} \sim F_{21-1,19-1}$
   $F_{obs} = \frac{484}{400} = 1.21$, $F_{crit,left} = 0.4$, $F_{crit,right} = 2.6 \Rightarrow$ оснований отвергать H_0 нет

   б) $\hat{\sigma}^2_0 = \frac{484\cdot(21-1)+400\cdot(19-1)}{21+19-2} \approx 444$
   $t_{obs} = \frac{78-67}{\sqrt{444}\sqrt{\frac{1}{21}+\frac{1}{19}}} \approx 1.8$
   $t_{crit} \sim t_{21+19-2} = t_{38}$, $t_{crit} = \pm 2.02 \Rightarrow$ нет оснований отвергать H_0

7. $\gamma = \sum_{i=1}^{s}\sum_{j=1}^{m} \frac{\left(n_{ij} - \frac{n_{i\cdot}n_{\cdot j}}{n}\right)^2}{\frac{n_{i\cdot}n_{\cdot j}}{n}} \sim \chi^2_{(s-1)(m-1)}$

   $\gamma_{obs} = \frac{\left(12-\frac{44\cdot 48}{100}\right)^2}{\frac{44\cdot 48}{100}} + \frac{\left(36-\frac{56\cdot 48}{100}\right)^2}{\frac{56\cdot 48}{100}} + \frac{\left(32-\frac{44\cdot 52}{100}\right)^2}{\frac{44\cdot 52}{100}} + \frac{\left(20-\frac{50\cdot 52}{100}\right)^2}{\frac{50\cdot 52}{100}} \approx 12$

   $\gamma_{crit} = 3.84 \Rightarrow$ гипотеза отвергается

19.5. 2014-2015

1. **Задача для первого потока.**

   а)

$$-Z_{\frac{\alpha}{2}} < \frac{\hat{p}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} < Z_{\frac{\alpha}{2}}$$
$$-1.96 < \frac{0.4-p}{\sqrt{\frac{0.4\cdot 0.6}{40}}} < 1.96$$
$$-0.56 < p < 1.36$$

   б) H_0 не отвергается, так как

$$-0.56 < 0.5 < 1.36$$

   в) H_0 отвергается, если $p = 0.5$ не лежит в построенном доверительном интервале:

$$0.5 \geqslant 0.5Z_{\frac{\alpha}{2}} + 0.4$$
$$0.2 \geqslant Z_{\frac{\alpha}{2}} \Rightarrow \alpha = 0.0456$$

1. **Задача для второго потока.**

   а) При верной $H_0 : \mu = 55$. Находим $t_{crit} = 1.721$ и $t_{obs} = \frac{51-55}{0.45} = -8.9$, H_0 отвергается.

   б) При верной $H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$. Находим $F_{left} = 0.5$, $F_{right} = 2.07$ и $F_{obs} = 2/3$, H_0 не отвергается.

2. **Задача для первого потока.**

   При верной $H_0 : \mu_\alpha = \mu_\beta$. Находим $t_{crit} = 2.048$ и $t_{obs} = \frac{9.5-9.8}{0.216} = -1.4$, H_0 не отвергается.

2. **Задача для второго потока.**

   а)

$$\hat{\sigma}_1^2 = 0.2\cdot 0.8 = 0.16$$
$$\hat{\sigma}_2^2 = 0.17\cdot 0.83 = 0.1411$$

   При верной $H_0 : \mu_1 = \mu_2$. Находим $Z_{crit} = 2.58$ и $t_{obs} = \frac{0.2-0.17}{0.06} = 0.5$, H_0 не отвергается.

б)

$$p_{value} = 1 - F(0.5) \approx 1 - 0.6915 = 0.3085$$

3.

$$L(x,\theta) = \theta^{-2n} x^n \exp\left(-\frac{1}{\theta}\sum_i x_i\right)$$

$$\ell = \ln(L) = -2n\ln(\theta) + n\ln(x) - \frac{1}{\theta}\sum_i x_i$$

$$\ell'_\theta = -\frac{2n}{\theta} + \frac{1}{\theta^2}\sum_i x_i$$

$$-2n\frac{1}{\hat\theta}\sum_i x_i = 0$$

$$\hat\theta_{ML} = \frac{1}{2}\bar X$$

4. Пусть $\ell = \ln(L)$ — логарифмическая функция правдоподобия.

$$\ell = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\nu) - \frac{1}{2\nu}\sum_i (x_i - \bar x)^2$$

Известно, что для выборки из нормального распределения $\hat\mu_{ML} = \bar x = 0.3$, $\hat\nu_{ML} = S^2 = 1.37$, поэтому

$$\ell_{UR} = -50\ln(2\pi) - 50\ln(1.37) - \frac{137}{2\cdot 1.37}$$

$$\ell_R = -50\ln(2\pi) - \frac{137}{2}$$

$$LR_{obs} = 2\left(-50\ln(2\pi) - 50\ln(1.37) - \frac{137}{2\cdot 1.37} + 50\ln(2\pi) + \frac{137}{2}\right) \approx 12$$

При верной H_0 $LR \sim \chi^2_1 \Rightarrow LR_{crit} \approx 3.8$, основная гипотеза отвергается.

5. **Исследовательская задача.**

а) Пусть $\ell = \ln(L)$ — логарифмическая функция правдоподобия. Воспользуемся также тем, что для выборки из нормального распределения $\hat\mu_{ML} = \bar X$, $\hat\nu_{ML} = S^2$.

$$\ell = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\nu) - \frac{1}{2\nu}\sum_{i=1}^n (x_i - \bar x)^2$$

$$\operatorname{Var}\left(\ell'(\hat\mu)\right) = \frac{1}{\nu^2}\operatorname{Var}\left(\sum_{i=1}^n (X_i - \bar X)\right) = \frac{n}{\nu^2}\operatorname{Var}(X_1) = \frac{n}{\nu}$$

$$LM = \frac{(\ell'(\hat\mu) - \ell'(0))^2}{\operatorname{Var}(\ell'(\hat\mu))} = \frac{\left(\frac{1}{\nu}\sum_{i=1}^n (X_i - \bar X) - \frac{1}{\nu}\sum_{i=1}^n X_i\right)^2}{\frac{n}{\nu}} = \frac{n}{\nu}\bar X^2$$

$$W = \frac{(\hat\mu - 0)^2}{\operatorname{Var}(\hat\mu)} = \frac{\left(\frac{\sum_{i=1}^n X_i}{n}\right)^2}{\frac{\nu}{n}} = \frac{n}{\nu}\bar X^2$$

$$LR = 2\left(-\frac{1}{2\nu}\sum_{i=1}^n (X_i - \bar X) + \frac{1}{2\nu}\sum_{i=1}^n X_i\right) =$$

$$= \frac{1}{\nu}\left(\sum_{i=1}^n X_i^2 - \sum_{i=1}^n X_i^2 + 2\bar X\cdot\sum_{i=1}^n X_i - n\bar X^2\right) =$$

$$= \frac{1}{\nu}\left(2\bar X\cdot n\cdot\bar X - n\bar X^2\right) = \frac{n}{\nu}\bar X^2$$

Как видим, $LR = LM = W$.

б)

6. **Исследовательская задача.**

а)

$$\mathbb{E}(X) = \int_0^{+\infty} a^2 x^2 e^{-ax} dx = -ax^2 e^{-ax}\Big|_0^{+\infty} + 2\int_0^{+\infty} axe^{-ax} dx =$$
$$= -2xe^{-ax}\Big|_0^{+\infty} + 2\int_0^{+\infty} e^{-ax} dx = -\frac{2}{a} e^{-ax}\Big|_0^{+\infty} = \frac{2}{a}$$
$$\frac{2}{\hat{a}_{MM}} = \frac{\sum_{i=1}^n x_i}{n} \Rightarrow \hat{a}_{MM} = \frac{2}{3}$$

б)

$$\hat{a}_{MM} = \frac{2}{\bar{X}}$$

Разложим $\hat{a}_{MM}$ по формуле Тейлора в окрестности точки $\bar{X} = 3$:

$$\hat{a}_{MM} \approx \frac{2}{3} - \frac{2}{9}\left(\bar{X} - 3\right)$$
$$\widehat{\text{Var}}(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} = \frac{\sum_{i=1}^n X_i^2 - 2\bar{X}\sum_{i=1}^n X_i + n\bar{X}^2}{n}$$
$$= \frac{1000 - 6 \cdot 300 + 100 \cdot 9}{100} = 1$$
$$\text{Var}\left(\hat{a}_{MM}\right) \approx \frac{4}{81}\text{Var}\left(\bar{X}\right) = \frac{4 \cdot 100}{81 \cdot 100^2}\text{Var}(X) \Rightarrow \widehat{\text{Var}}\left(\hat{a}_{MM}\right) = \frac{4}{8100}$$

в)

$$-t_{\frac{\alpha}{2};n-1} < \frac{\bar{X} - \mu}{\frac{\hat{\sigma}}{\sqrt{n}}} < t_{\frac{\alpha}{2};n-1}$$
$$-1.98 < \frac{3 - \mu}{\frac{1}{10}} < 1.98$$

Воспользуемся тем, что $\mu = \frac{2}{a}$:

$$-1.98 < \frac{3 - \frac{2}{a}}{\frac{1}{10}} < 1.98$$
$$-0.198 < 3 - \frac{2}{a} < 0.198$$
$$0.63 < a < 0.71$$

19.6. 2009-2010

20. Решения финальных экзаменов

20.1. 2018-2019

CACAE ACCBA BDEBA DDBEB ECEBC EAEAD

20.2. 2017-2018

Здесь табличка с ответами