Математика для Data Science. Теория вероятностей.
Решения задач

Содержание

Замечание. Вот этим цветом отмечены ссылки на страницы внутри этого файла.

Распределения и независимые случайные величины

Задача 1

Мы три раза подряд подбрасываем честную монету.

- Случайная величина Y – суммарное число выпавших решек
- Случайная величина Z – суммарное число выпавших орлов

1. Верно ли, что $Y = Z$?
2. Докажите, что $p_Y(a) = p_Z(a)$ для любого числа $a \in \mathbb{R}$. Тем самым, у Y и Z одинаковые функции вероятности. Можно записать это так: $p_Y = p_Z$.

Определение. Две случайные величины с совпадающими функциями вероятности называются *одинаково распределёнными.*

Подсказка. Равенство случайных величин — это то же, что равенство функций.

Решение.

1. Случайные величины Y и Z не равны, ведь при любом исходе ω выполнено $Y(\omega) \neq Z(\omega)$.
2. Пусть $a \in \mathbb{R}$. Тогда $p_Y(a) = P(Y = a)$ и $P_Z(a) = P(Z = a)$. Значения a, при которых эти вероятности не обращаются в ноль, это $\{0, 1, 2, 3\}$.

   Рассмотрим, например, случай $a = 0$. Тогда $p_Y(0) = P(Y = 0)$ — вероятность того, что на трёх честных монетках не выпало ни одной решки, то есть все три раза выпали орлы. Вероятность этого $\frac{1}{8}$.

   С другой стороны, $p_Z(0)$ — это вероятность того, что на трёх честных монетках не выпало ни одного орла, то есть все три раза выпали решки. Вероятность этого тоже $\frac{1}{8}$.

   Аналогичное рассуждение проводится для остальных значений a: вычисляем $p_Y(a)$, а затем в рассуждениях меняем орла и решку местами и получаем $p_Z(a) = p_Y(a)$.

Задача 2

Дана случайная величина X. Пусть она принимает ровно n различных значений: $x_1, \ldots, x_n$.

1. Докажите, что $\forall i : p_X(x_i) \geq 0$
2. Докажите, что $p_X(a) = 0$ для всех a не равных одному из $x_1, \ldots, x_n$.
3. Докажите, что $\sum\limits_{i=1}^{n} p_X(x_i) = 1$. Или в другой записи: $p_X(x_1) + p_X(x_2) + \cdots + p_X(x_n) = 1$.

То есть про функцию вероятности можно думать, как про суммарную массу 1, как-то раскиданную по n точкам на числовой прямой.

Поэтому можно задавать функцию вероятности случайной величины, просто перечислив значения в точках $x_1, \ldots, x_n$. Например, такие условия: $p_Y(25) = 0.33, p_Y(38) = 0.67$ однозначно задают функцию вероятности случайной величины Y, которая принимает ровно два значения: 25 и 38.

Комментарий. Если эта задача вам кажется подозрительно простой, то да, подвоха нет, это действительно простая задача – фактически, мы просим вас применить определение функции вероятности случайной величины.

Решение.

1. По определению $p_X(x_i) = P(X = x_i) \geq 0$ как и любая вероятность.
2. Если $a \notin \{x_1, \ldots, x_n\}$, то $p_X(a) = P(X = a) = 0$, ведь такое значение случайная величина X по условию не принимает.
3. $\sum\limits_{i=1}^{n} p_X(x_i) = \sum\limits_{i=1}^{n} P(X = x_i) = 1$, поскольку возможные значения величины X это только $x_1, \ldots, x_n$.

Независимые случайные величины

Задача 1

В конце прошлой недели мы поняли, что вычисление математического ожидания не перестановочно с умножением. То есть не всегда выполнено $E[X \cdot Y] = E[X] \cdot E[Y]$.

Задача. Докажите, что если X и Y независимы, то $E[X \cdot Y] = E[X] \cdot E[Y]$.

Подсказка. Введите обозначения для всех значений, которые могут принимать X и Y, и для вероятностей, соответствующих этим значениям.

Решение. Пусть случайная величина X принимает значения $x_1, \ldots, x_n$ с вероятностями $p_1, \ldots, p_n$ соответственно. Тогда по определению $E[X] = \sum\limits_{i=1}^{n} x_i p_i$.

Аналогично, пусть случайная величина Y принимает значения $y_1, \ldots, y_m$ с вероятностями $q_1, \ldots, q_m$. Тогда $E[Y] = \sum\limits_{j=1}^{m} y_j q_j$. Итак,

$$E[X] \cdot E[Y] = \left(\sum_{i=1}^{n} x_i p_i\right) \cdot \left(\sum_{j=1}^{m} y_j q_j\right).$$

Наконец, по определению $E[X \cdot Y] = \sum\limits_{i,j} x_i y_j P(X = x_i, Y = y_j)$, где суммирование ведётся по всем возможным парам $i \in \{1, \ldots, n\}$ и $j \in \{1, \ldots, m\}$. Далее, поскольку X и Y независимы, равенство продолжается так:

$$E[X \cdot Y] = \sum_{i,j} x_i y_j P(X = x_i) \cdot P(Y = y_j) = \sum_{i,j} x_i y_j p_i q_j = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i p_i y_j q_j = \left(\sum_{i=1}^{n} x_i p_i\right) \cdot \left(\sum_{j=1}^{m} y_j q_j\right) = E[X] \cdot E[Y].$$

Дополнительная задача

Обозначение. Аналогично обозначению события $X = a$ введём следующее обозначение. Событие "случайная величина X приняла значение в отрезке $[a, b]$" будем обозначать так: $X \in [a, b]$.

Докажите, что если случайные величины X и Y независимы, то для любых $a, b, c, d \in \mathbb{R}$ события $X \in [a, b]$ и $Y \in [c, d]$ независимы.

Подсказка. Вероятность $P(X \in [a, b], Y \in [c, d])$ надо выразить через вероятности вида $P(X = x_i, Y = y_j)$, а затем воспользоваться независимостью случайных величин X и Y.

Решение. Нам нужно доказать, что $P(X \in [a, b], Y \in [c, d]) = P(X \in [a, b]) \cdot P(Y \in [c, d])$.

Зафиксируем $a, b, c, d \in \mathbb{R}$ — произвольные числа. Пусть значения случайной величины X, попадающие в отрезок $[a, b]$ — это $x_1, \ldots, x_n$, а их вероятности равны соответственно $p_1, \ldots, p_n$. Аналогично обозначим значения Y, принадлежащие $[c, d]$ за $y_1, \ldots, y_m$ с вероятностями $q_1, \ldots, q_m$.

Тогда $P(X \in [a, b]) = P((X = x_1) \cup \cdots \cup (X = x_n)) = p_1 + \cdots + p_n$. Аналогично $P(Y \in [c, d]) = q_1 + \cdots + q_m$.

Осталось найти $P(((X = x_1) \cup \cdots \cup (X = x_n)) \cap ((Y = y_1) \cup \cdots \cup (Y = y_m)))$.

Вспомним свойство операций над множествами: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (это называется *дистрибутивностью*).

Много раз применив это свойство, получим $\Big((X = x_1) \cup \cdots \cup (X = x_n)\Big) \cap \Big((Y = y_1) \cup \cdots \cup (Y = y_m)\Big) =$
$= \Big((X = x_1) \cap (Y = y_1)\Big) \cup \cdots \cup \Big((X = x_1) \cap (Y = y_m)\Big) \cup \cdots \cup \Big((X = x_n) \cap (Y = y_1)\Big) \cup \cdots \cup \Big((X = x_n) \cap (Y = y_m)\Big)$.

Поскольку X и Y независимы, то вероятности вида $P((X = x_i) \cap (Y = y_j)) = P(X = x_i) \cdot P(Y = y_j) = p_i q_j$. Тогда искомая $P(X \in [a, b], Y \in [c, d]) = p_1 q_1 + \cdots + p_1 q_m + \cdots + p_n q_1 + p_n q_m$, а это как раз равно $P(X \in [a, b]) \cdot P(Y \in [c, d]) = (p_1 + \cdots + p_n)(q_1 + \cdots + q_m)$, ура!

Задача 2

Даны k совместно независимых случайных величин $X_1, \ldots, X_k$. Все эти величины имеют одинаковое распределение: $P(X_i = 0) = q, P(X_i = 1) = 1 - q$. Найдите распределение случайной величины $X_1 \cdot X_2 \cdot \ldots \cdot X_k$.

Решение.

Случайная величина $X_1 \cdot X_2 \cdot \ldots \cdot X_k$ равна единице только в том случае, когда все $X_i = 1$. Вероятность этого $P(X_1 = 1, \ldots, X_n = 1) = P(X_1 = 1) \cdot \ldots \cdot P(X_n = 1) = (1 - q)^n$.

Итак, $P(X_1 \cdot X_2 \cdot \ldots \cdot X_k = 1) = (1 - q)^n$ и
$P(X_1 \cdot X_2 \cdot \ldots \cdot X_k = 0) = 1 - (1 - q)^n$.

Дисперсия

Задача 2

1. Докажите, что для любой случайной величины X выполнено $Var[X] \geq 0$.
2. Докажите, что $Var[X] = 0$ если и только если X это постоянная случайная величина.

Подсказка. Здесь удобнее воспользоваться первым определением дисперсии: $Var[X] = E[(X - E[X])^2]$.

Решение.

1. Пусть случайная величина X принимает значения $x_1, \ldots, x_n$ с вероятностями $p_1, \ldots, p_n$ соответственно.

   По определению $Var[X] = E[(X - E[X])^2] = \sum\limits_{i=1}^{n} p_i(x_i - E[X])^2$. При этом $p_i \geq 0$ по определению вероятности, а $(x_i - E[X])^2 \geq 0$ как квадрат числа. Значит, мы суммируем неотрицательные величины и получаем тоже неотрицательную величину: $Var[X] \geq 0$.

2. Докажем сначала в одну сторону. Пусть X — постоянная случайная величина, $X = c$. Тогда $E[X] = c$ и $Var[X] = E[(X - E[X])^2] = E[(c - c)^2] = E[0] = 0$.

   Обратно, пусть $Var[X] = 0$. Пусть случайная величина X принимает значения $x_1, \ldots, x_n$ с ненулевыми вероятностями $p_1, \ldots, p_n$ соответственно.

   Тогда $Var[X] = E[(X - E[X])^2] = \sum\limits_{i=1}^{n} p_i(x_i - E[X])^2 = 0$. Поскольку мы суммируем неотрицательные выражения и получаем ноль, то каждое из слагаемых должно равняться нулю. Поскольку $p_i \neq 0$, то $x_i - E[X] = 0$. Обозначим константу $E[X]$ за c. Тогда $x_i = c$ для всех i. А значит, $X = c$.

Задача 3

1. Пусть X и Y это независимые случайные величины. Докажите, что $Var[X + Y] = Var[X] + Var[Y]$.
2. Приведите пример X и Y, таких что $Var[X + Y] \neq Var[X] + Var[Y]$.

Подсказка. Здесь пригодится задача про математическое ожидание произведения независимых случайных величин.

Решение.

1. Поскольку X и Y это независимые случайные величины, то по задаче прошлой недели $E[X \cdot Y] = E[X] \cdot E[Y]$. Тогда, пользуясь этим и линейностью математического ожидания, получаем

$$Var[X + Y] = E[(X + Y)^2] - E[X + Y]^2 = E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 =$$
$$= E[X^2] + 2E[XY] + E[Y^2] - E[X]^2 - 2E[X]E[Y] - E[Y]^2 = E[X^2] + \cancel{2E[X]E[Y]} + E[Y^2] - E[X]^2 - \cancel{2E[X]E[Y]} - E[Y]^2 =$$
$$= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 = Var[X] + Var[Y]$$

2. Рассмотрим $X = Y$. Тогда $Var[X+Y] = Var[2X] = 4Var[X]$. Но при этом $Var[X]+Var[Y] = 2Var[X]$. То есть в качестве X можно взять любую случайную величину с ненулевой дисперсией. Например, можно рассмотреть X с распределением Бернулли с вероятностью успеха $1 > p > 0$.

Задача 4

Пусть c это произвольное действительное число. Выразите дисперсию случайной величины cX через дисперсию случайной величины X.

Подсказка. Пригодится то, что константы можно выносить за знак математического ожидания. Это мы доказали в первом пункте этой задачи.

Решение. Воспользуемся определением дисперсии и линейностью математического ожидания:

$$Var[cX] = E[(cX)^2] - (E[cX])^2 = E[c^2X^2] - (cE[X])^2 = c^2E[X^2] - c^2E[X]^2 = c^2(E[X^2] - E[X]^2) = c^2 \cdot Var[X].$$

Биномиальное распределение и стандартное отклонение

Задача 1

В вашей фирме работают n разработчиков. Каждый рабочий день каждый разработчик приходит в офис с вероятностью p, и с вероятностью $(1-p)$ остаётся работать из дома. Разработчики приходят работать независимо друг от друга.

1. Какова вероятность, что сегодня конкретные k разработчиков будут работать из офиса, а все остальные – из дома? Например, если разработчиков зовут Аня, Гриша, Вика, Петя и Илья, то какова вероятность, что из разработчиков в офис придут только Аня и Петя?

2. Какова вероятность, что сегодня в офис придут ровно k разработчиков? Имена пришедших не уточняются – важно только чтобы пришло ровно k.

3. Найдите распределение случайной величины S, где S – число разработчиков, которые пришли сегодня в офис.

Подсказка. Здесь полезно вспомнить про биномиальные коэффициенты и их комбинаторный смысл.

Решение. Пронумеруем всех разработчиков числами от 1 до n и введём случайные величины: X_i равно 1, если i-ый разработчик пришёл сегодня в офис, и 0 иначе. Тогда X_i имеет распределение Бернулли с вероятностью успеха p. А значит, поскольку $S = X_1 + \ldots + X_n$, то S имеет биномиальное распределение $Bin(n, p)$.

1. Пусть конкретные k разработчиков — это $X_1, \ldots, X_k$. Тогда вероятность того, что именно они будут работать из офиса равна
$$P(X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_n = 0).$$
Поскольку разработчики приходят работать независимо друг от друга, то эта вероятность равна произведению
$$P(X_1 = 1) \cdot \ldots \cdot P(X_k = 1) \cdot P(X_{k+1} = 0) \cdot \ldots \cdot P(X_n = 0) = \underbrace{p \cdot \ldots \cdot p}_{\text{k}} \cdot \underbrace{(1-p) \cdot \ldots \cdot (1-p)}_{\text{n-k}} = p^k(1-p)^{n-k}.$$
То есть в случае, когда разработчиков зовут Аня, Гриша, Вика, Петя и Илья, то вероятность, что из разработчиков в офис придут только Аня и Петя равна $p^2(1-p)^{5-2} = p^2(1-p)^3$.

2. Если теперь важно только чтобы пришло ровно k разработчиков, то в предыдущем пункте можно было выбрать не разработчиков с номерами $\{1, \ldots, k\}$, а любое другое подмножество множества $\{1, \ldots, n\}$, состоящее из k элементов. Как мы помним, способов сделать этот выбор всего $\binom{n}{k}$. Значит, чтобы найти искомую вероятность, мы должны сложить $\binom{n}{k}$ вероятностей $p^k(1-p)^{n-k}$. Мы получим вероятность, равную $\binom{n}{k}p^k(1-p)^{n-k}$.

3. В предыдущем пункте мы показали, что $p_S(k) = P(S = k) = \binom{n}{k} p^k (1-p)^{n-k}$, где $k \in \mathbb{Z}$ и $0 \leq k \leq n$. Для всех остальных $a \notin \{0, \ldots, n\}$ функция вероятности равна нулю: $p_S(a) = 0$.

Задача 2

У вас та же фирма, что и в прошлой задаче. Пусть $X_i = 1$ если i-ый разработчик пришёл, и $X_i = 0$ в противном случае. Тогда $S := X_1 + \cdots + X_n$ это количество пришедших разработчиков.

1. Найдите $E[S]$.

2. Найдите $Var[S]$.

Подсказка. Вспомните, что происходит с математическим ожиданием и дисперсией, если складывать случайные величины или умножать их на число.

Решение.

1. По линейности математического ожидания $E[S] = E[X_1] + \ldots + E[X_n]$. А поскольку X_i имеет распределение Бернулли с вероятностью успеха p, то $E[X_i] = p$. Итого $E[S] = \underbrace{p + \ldots + p}_{\text{n}} = np$.

2. Так как случайные величины X_i независимы, то по второй задаче этого урока $Var[S] = Var[X_1] + \ldots + Var[X_n]$. Дисперсию бернуллиевской случайной величины мы уже считали (например, во втором пункте этой задачи). Итак, поскольку $Var[X_i] = p(1-p)$, то $Var[S] = np(1-p)$.

Задача 3

Теперь пусть разработчики либо все одновременно приходят (с вероятностью p), либо все одновременно не приходят (с вероятностью $1-p$). Обозначим число пришедших разработчиков за T.

1. Найдите $E[T]$.

2. Найдите $Var[T]$.

Подсказка. Вспомните, что происходит с математическим ожиданием и дисперсией, если складывать случайные величины или умножать их на число.

Решение.
Если разработчики либо все одновременно приходят (с вероятностью p), либо все одновременно не приходят (с вероятностью $1-p$), то число пришедших разработчиков может равняться либо 0, либо n. По-другому это можно записать так: $T = nY$, где Y имеет распределение Бернулли с вероятностью успеха p.

1. Тогда $E[Y] = p$ и $Var[Y] = p(1-p)$. Значит, поскольку n — константа, $E[T] = E[nY] = nE[Y] = np$

2. $Var[T] = Var[nY] = n^2 Var[Y] = n^2 p(1-p)$.

Дополнительная задача

1. Зачем вообще искать средний квадрат отклонения? Давайте лучше искать само среднее отклонение, а не его квадрат. Найдите $E[X - E[X]]$.

2. Пусть мы приняли определение из прошлого шага $H[X] := E[|X - E[X]|]$. Мы хотим использовать H в наших вычислениях, а не дисперсию. Докажите, что даже для независимых X и Y не всегда выполнено $H[X + Y] = H[X] + H[Y]$.

Как мы помним, для независимых X и Y выполнено $Var[X+Y] = Var[X] + Var[Y]$. Так что дисперсия ещё и этим удобнее для вычислений, чем H (а не только тем, что H использует модуль, а Var нет).

Подсказка. В обоих пунктах пригодится линейность математического ожидания.

Решение.

1. Воспользуемся линейностью математического ожидания: $E[X - E[X]] = E[X] - E[E[X]]$. Поскольку $E[X]$ — это просто число, то его математическое ожидание равно ему же: $E[E[X]] = E[X]$. А тогда, продолжая начатое вычисление, получаем $E[X - E[X]] = E[X] - E[X] = 0$.

2. Пусть

$$X = \begin{cases} 1 \text{ с вероятностью } 0.5 \\ 0 \text{ с вероятностью } 0.5 \end{cases}$$

$$Y = \begin{cases} -1 \text{ с вероятностью } 0.5 \\ 0 \text{ с вероятностью } 0.5 \end{cases}$$

При этом пусть X и Y независимы.

Найдём $H[X] = E[|X - E[X]|]$. Для начала по определению посчитаем $E[X] = 1 \cdot 0.5 + 0 \cdot 0.5 = 0.5$. Следовательно, $H[X] = E[|X - 0.5|]$. При этом $X - 0.5$ равен либо 0.5, либо -0.5. Значит, $|X - 0.5|$ всегда равен 0.5, как и его математическое ожидание. Итак, мы нашли $H[X] = 0.5$.

Теперь аналогично разберёмся с $H[Y] = E[|Y - E[Y]|]$. Начнём с $E[Y] = (-1) \cdot 0.5 + 0 \cdot 0.5 = -0.5$. Тогда $Y - E[Y] = Y - (-0.5) = Y + 0.5$ и эта случайная величина может принимать два значения: -0.5 и 0.5. А её модуль всегда равен 0.5, а значит $H[Y] = 0.5$.

Итак, мы получили, что $H[X] + H[Y] = 0.5 + 0.5 = 1$.

Наконец найдём $H[X+Y] = E[|X + Y - E[X+Y]|]$. Из линейности математического ожидания следует, что $E[X+Y] = E[X] + E[Y] = 0.5 + (-0.5) = 0$. То есть $H[X+Y] = E[|X+Y|]$.

Далее, поскольку X и Y независимы, то

$$X + Y = \begin{cases} 0 \text{ с вероятностью } 0.5 \\ 1 \text{ с вероятностью } 0.25 \\ -1 \text{ с вероятностью } 0.25 \end{cases}$$

Тогда $|X+Y|$ равен 0 с вероятностью 0.5 и равен 1 с вероятностью 0.5. А значит $H[X+Y] = E[|X+Y|] = 0 \cdot 0.5 + 1 \cdot 0.5 = 0.5$.

Итого $H[X+Y] = 0.5 \neq H[X] + H[Y] = 1$.

Ряды

Задача 1

Докажите, что для любого $\beta \neq 1$ и любого $t \in \mathbb{N}$ выполнено $1 + \beta + \beta^2 + \dots + \beta^{t-2} + \beta^{t-1} = \frac{1-\beta^t}{1-\beta}$.
Комментарий. Формулу из пункта 2 мы уже доказывали в курсе матана, когда говорили про градиентный спуск с моментом.

Решение. Умножим обе части равенства на $(1-\beta)$. Тогда надо доказать $(1+\beta+\beta^2+\dots+\beta^{t-2}+\beta^{t-1})(1-\beta) = 1-\beta^t$.
Раскроем скобки в левой части: $(1+\beta+\beta^2+\dots+\beta^{t-2}+\beta^{t-1})(1-\beta) =$
$= 1+\beta+\beta^2+\dots+\beta^{t-2}+\beta^{t-1}-\beta-\beta^2-\beta^3-\dots-\beta^{t-1}-\beta^t$. Все β^i, где $i \in \{1, 2, \dots, t-1\}$ сократятся. Останется как раз $1-\beta^t$.

Задача 2

Дан ряд $1 + \beta + \beta^2 + \beta^3 + \dots = \sum\limits_{n=0}^{\infty} \beta^n$. Для каждого β найдите сумму ряда или докажите, что ряд расходится.

Решение.

- При $\beta = 1$ ряд расходится, потому что последовательность его частичных сумм образует ряд натуральных чисел и предела не имеет.
- А теперь воспользуемся предыдущими пунктами этой задачи: по первому пункту при $\beta = -1$ ряд расходится.

  По второму пункту при $\beta \neq 1$ выполнено $S_t = 1 + \beta + \dots + \beta^{t-1} = \frac{1-\beta^t}{1-\beta}$.
- Тогда если $|\beta| < 1$, то $\lim\limits_{t\to\infty} \beta^t = 0$ и тогда предел $\lim\limits_{t\to\infty} S_t = \frac{1}{1-\beta}$ равен сумме ряда.
- Если же $|\beta| > 1$, то $\lim\limits_{t\to\infty} \beta^t = +\infty$ или $-\infty$, а тогда $\lim\limits_{t\to\infty} S_t$ не конечен и, следовательно, ряд расходится.

Задача 3

1. Докажите, что ряд $\frac{1}{1} + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{16} + \dots$ расходится.
2. Докажите, что ряд $\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \dots = \sum\limits_{n=1}^{\infty} \frac{1}{n}$ расходится.

Ряд из Пункта 2 называется *гармоническим.*

Комментарий. По аналогии с определениями пределами последовательностей можно сказать, что $\sum\limits_{n=1}^{\infty} \frac{1}{n} = +\infty$, ведь последовательность частичных сумм этого ряда сходится к $+\infty$. При этом такой ряд мы всё равно называем расходящимся.

Подсказка. Второй пункт следует из первого.

Решение.

1. Посчитаем частичные суммы этого ряда: $S_1 = 1, S_2 = 1.5, S_4 = 2, S_8 = 2.5$ и так далее: $S_{2^k} = 1 + 0.5 \cdot k$. Но $\lim\limits_{k\to\infty} (1 + 0.5 \cdot k) = +\infty$, то есть подпоследовательность $\{S_{2^k}\}$ не имеет конечного предела. А тогда конечного предела не имеет и последовательность частичных сумм $\{S_n\}$. То есть ряд расходится.
2. n-ый член гармонического ряда больше n-ого члена ряда из предыдущего пункта. Поэтому предел частичных сумм гармонического ряда тоже равен $+\infty$, то есть ряд расходится.

Задача 4

Даны два сходящихся ряда: $\sum\limits_{n=1}^{\infty} a_n$ и $\sum\limits_{n=1}^{\infty} b_n$.

1. Докажите, что ряд $\sum\limits_{n=1}^{\infty} (a_n + b_n)$ сходится
2. Докажите, что ряд $\sum\limits_{n=1}^{\infty} (a_n - b_n)$ сходится
3. Докажите, что ряд $\sum\limits_{n=1}^{\infty} (ca_n)$ сходится для любого $c \in \mathbb{R}$

Естественно, в доказательстве можно использовать уже доказанные утверждения из курса матана.

Подсказка. Здесь пригодятся свойства операции суммирования и взятия предела.

Решение. Обозначим суммы рядов так: $A := \sum\limits_{n=1}^{\infty} a_n = \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N} a_n\right)$ и $B := \sum\limits_{n=1}^{\infty} b_n = \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N} b_n\right)$.

1. $\sum\limits_{n=1}^{\infty}(a_n+b_n) = \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N}(a_n+b_n)\right) = \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N} a_n\right) + \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N} b_n\right) = A+B.$
2. $\sum\limits_{n=1}^{\infty}(a_n-b_n) = \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N}(a_n-b_n)\right) = \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N} a_n\right) - \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N} b_n\right) = A-B.$
3. $\sum\limits_{n=1}^{\infty} ca_n = \lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N} ca_n\right) = c\cdot\lim\limits_{N\to\infty}\left(\sum\limits_{n=1}^{N} a_n\right) = cA.$

Абсолютно сходящиеся ряды

Задача 1

Давайте потренируемся применять теорему с прошлого шага.
Пусть

- $a_1 = \frac{1}{1} - \frac{1}{2}$
- $a_2 = \frac{1}{3} - \frac{1}{4}$,
- $a_3 = \frac{1}{5} - \frac{1}{6}$,
- $a_4 = \frac{1}{7} - \frac{1}{8}$,
- и так далее

Ясно, что все $a_i \geq 0$. Давайте докажем, что ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится.

Задача. Докажите, что последовательность частичных сумм ряда неубывающая и ограничена сверху числом 1.

Подсказка. Попробуйте по-разному расставить скобки в формуле для частичных сумм S_k.

Решение.

$$S_k = \left(\frac{1}{1}-\frac{1}{2}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\left(\frac{1}{5}-\frac{1}{6}\right)+\ldots+\left(\frac{1}{2k-1}-\frac{1}{2k}\right) = \frac{1}{1\cdot 2}+\frac{1}{3\cdot 4}+\frac{1}{5\cdot 6}+\ldots+\frac{1}{(2k-1)2k}$$

Каждое из слагаемых положительно, поэтому последовательность частичных сумм $S_1, S_2, S_3, \ldots$ возрастает.
С другой стороны,

$$\begin{aligned} S_k = \frac{1}{1} - \left(\frac{1}{2}-\frac{1}{3}\right) - \left(\frac{1}{4}-\frac{1}{5}\right) - \left(\frac{1}{6}-\frac{1}{7}\right) - \ldots - \left(\frac{1}{2k-2}-\frac{1}{2k-1}\right) - \frac{1}{2k} = \\ = 1 - \frac{1}{2\cdot 3} - \frac{1}{4\cdot 5} - \frac{1}{6\cdot 7} - \ldots - \frac{1}{(2k-2)(2k-1)} - \frac{1}{2k} < 1 \end{aligned}$$

Итак, мы доказали, что последовательность $S_1, S_2, S_3, S_4, \ldots$ возрастает и ограничена сверху числом 1, а к этому мы и хотели прийти!

Задача 2

В этой задаче мы докажем утверждения, ради которых и вводили теорему из пред-предыдущего шага. А эти утверждения помогут нам доказать некоторые свойства абсолютно сходящихся рядов.

Даны два ряда $\sum\limits_{n=1}^{\infty} a_n$ и $\sum\limits_{n=1}^{\infty} b_n$, при этом $0 \leq a_n \leq b_n$ для всех n. В таком случае говорят, что ряд $\sum\limits_{n=1}^{\infty} b_n$ *мажорирует* ряд $\sum\limits_{n=1}^{\infty} a_n$.

1. Докажите, что если $\sum\limits_{n=1}^{\infty} b_n$ сходится, то $\sum\limits_{n=1}^{\infty} a_n$ сходится. Приведите пример, когда обратное утверждение неверно.

2. Докажите, что если $\sum\limits_{n=1}^{\infty} a_n$ расходится, то $\sum\limits_{n=1}^{\infty} b_n$ расходится. Приведите пример, когда обратное утверждение неверно.

Подсказка. Здесь тоже пригодится пройденная ранее теорема.

Решение. Обозначим частичные суммы рядов так: $A_n = a_1 + \cdots + a_n$ и $B_n = b_1 + \cdots + b_n$. Поскольку $0 \leq a_n \leq b_n$ для всех n, то $A_n \leq B_n$ и последовательности $\{A_n\}$ и $\{B_n\}$ не убывают.

1. Если $\sum\limits_{n=1}^{\infty} b_n$ сходится, то у последовательности $\{B_n\}$ есть предел. Значит, $\{B_n\}$ ограничена: $\exists C : B_n < C$ для всех $n \in \mathbb{N}$. Но тогда и $A_n < C$ для всех натуральных n. А мы уже доказали, что у любой неубывающей ограниченной сверху последовательности есть предел. Итак, последовательность $\{A_n\}$ сходится, а значит сходится и ряд $\sum\limits_{n=1}^{\infty} a_n$.

   Пример, когда обратное не верно: $\sum\limits_{n=1}^{\infty} a_n = \sum\limits_{n=1}^{\infty} 0 = 0$ — ряд сходится, и $\sum\limits_{n=1}^{\infty} b_n = \sum\limits_{n=1}^{\infty} 1 = +\infty$ — ряд расходится.

2. Если $\sum\limits_{n=1}^{\infty} a_n$ расходится, то $\lim\limits_{n\to\infty} A_n = +\infty$. То есть по определению $\forall C \in \mathbb{R}\ \exists N \in \mathbb{N}\ \forall n \geq N : A_n > C$. Но тогда и $\forall n \geq N$ выполнено $B_n > C$, то есть $\lim\limits_{n\to\infty} B_n = +\infty$ и ряд $\sum\limits_{n=1}^{\infty} b_n$ расходится.

   Пример, когда обратное не верно, такой же, как в прошлом пункте.

Задача 3

Докажем теорему:

Теорема. Любой абсолютно сходящийся ряд является сходящимся.

Пусть ряд $\sum\limits_{n=1}^{\infty} a_n$ абсолютно сходится, то есть ряд $\sum\limits_{n=1}^{\infty} |a_n|$ сходится. Давайте в три этапа докажем, что тогда $\sum\limits_{n=1}^{\infty} a_n$ сходится.

1. Докажите, что ряд $\sum\limits_{n=1}^{\infty} 2|a_n|$ неотрицателен и сходится.

2. Докажите, что ряд $\sum\limits_{n=1}^{\infty} (a_n + |a_n|)$ неотрицателен и мажорируется рядом $\sum\limits_{n=1}^{\infty} 2|a_n|$.

3. Докажите, что ряд $\sum\limits_{n=1}^{\infty} a_n = \sum\limits_{n=1}^{\infty} (a_n + |a_n|) - \sum\limits_{n=1}^{\infty} |a_n|$ сходится

Тем самым мы доказали, что из абсолютной сходимости ряда следует сходимость ряда.

Подсказка. Пригодятся четвёртая задача урока про ряды и предыдущая задача.

Решение.

1. $2|a_n| \geq 0$, поэтому ряд неотрицателен. По третьему пункту второй задачи урока про ряды $\sum\limits_{n=1}^{\infty} 2|a_n| = 2\sum\limits_{n=1}^{\infty} |a_n|$, то есть ряд сходится.

2. Найдём, чему равно $a_n + |a_n|$. Если $a_n \geq 0$, то $a_n + |a_n| = a_n + a_n = 2a_n \geq 0$. Если $a_n < 0$, то $a_n + |a_n| = a_n - a_n = 0$. Значит, ряд $\sum\limits_{n=1}^{\infty}(a_n + |a_n|)$ неотрицателен. Кроме того, из нашего разбора случаев следует, что $a_n + |a_n| \leq 2|a_n|$, то есть действительно ряд $\sum\limits_{n=1}^{\infty}(a_n + |a_n|)$ мажорируется рядом $\sum\limits_{n=1}^{\infty} 2|a_n|$.

3. Из двух предыдущих пунктов и предыдущей задачи следует, что ряд $\sum\limits_{n=1}^{\infty}(a_n + |a_n|)$ сходится, так как сходится мажорирующий его ряд. А тогда из второго пункта четвёртой задачи следует, что ряд $\sum\limits_{n=1}^{\infty} a_n = \sum\limits_{n=1}^{\infty}(a_n + |a_n|) - \sum\limits_{n=1}^{\infty} |a_n|$ сходится как разность двух сходящихся рядов.

Дополнительная задача

На этом шаге мы докажем теорему с пред-предыдущего шага:

Теорема. Если ряд абсолютно сходится к сумме S, то любой ряд, полученный из него перестановкой слагаемых, тоже абсолютно сходится к той же сумме S.

Пусть дан абсолютно сходящийся ряд $a_1+a_2+a_3+\dots = \sum\limits_{n=1}^{\infty} a_n$. Обозначим его сумму за S, а его частичные суммы за $\{S_n\}$.

Обозначим за $b_1 + b_2 + b_3 + \dots = \sum\limits_{n=1}^{\infty} b_n$ ряд, полученный из ряда $\sum\limits_{n=1}^{\infty} a_n$ при помощи любой перестановки. Частичные суммы этого ряда будем обозначать за $\{\tilde{S}_n\}$.

В первых трёх пунктах давайте считать, что все $a_i \geq 0$. А в 4-ом и 5-ом пунктах мы поймём, что делать в случае, когда a_i могут быть отрицательными.

1. Пусть все $a_i \geq 0$. Докажите, что для любого i найдётся j, такой что $S_i \leq \tilde{S}_j$. Другими словами, для любой частичной суммы первого ряда найдётся не меньшая (то есть такая же или бóльшая) частичная сумма второго ряда.

2. Пусть все $a_i \geq 0$. Докажите, что для любого i найдётся j, такой что $\tilde{S}_i \leq S_j$. Другими словами, для любой частичной суммы второго ряда найдётся не меньшая (то есть такая же или бóльшая) частичная сумма первого ряда.

3. Пусть все $a_i \geq 0$. Докажите, что $\lim\limits_{n\to\infty} \tilde{S}_n$ существует и равен $S := \lim\limits_{n\to\infty} S_n$.

4. Теперь пусть a_i могут быть какими угодно. Докажите, что ряд, составленный только из неотрицательных членов ряда $\sum\limits_{n=1}^{\infty} a_n$ абсолютно сходится. Аналогично, докажите, что ряд, составленный только из отрицательных членов ряда $\sum\limits_{n=1}^{\infty} a_n$ абсолютно сходится.

5. Выведите из пункта 4, что $\lim\limits_{n\to\infty} \tilde{S}_n$ существует и равен $S := \lim\limits_{n\to\infty} S_n$.

Решение.

1. Мы хотим для любого i найти j, такой что $S_i \leq \tilde{S}_j$. Достаточно будет взять j таким, что в $\tilde{S}_j$ присутствуют все слагаемые из S_i. При этом кроме $a_1, \dots, a_i$ в формуле для $\tilde{S}_j$ могут также быть другие слагаемые, а поскольку все они неотрицательны, то $S_i \leq \tilde{S}_j$

2. Аналогично предыдущему пункту достаточно взять j таким, что в S_j присутствуют все слагаемые из $\tilde{S}_i$. А тогда $\tilde{S}_i \le S_j$.

3. Из второго пункта следует, что для любого n найдётся j такой, что $\tilde{S}_n \le S_j = \sum\limits_{k=1}^{j} a_k \le \sum\limits_{k=1}^{\infty} a_k = S$. Значит, последовательность $\tilde{S}_n$ ограничена. Кроме того, она не убывает. А значит, как мы обсуждали ранее, у последовательности $\tilde{S}_n$ есть предел. Обозначим его за $\tilde{S}$. Поскольку $\tilde{S}_n \le S$ для любого n, то и $\tilde{S} \le S$.

   С другой стороны, из первого пункта следует, что для любого n существует j такой, что $S_n \le \tilde{S}_j \le \tilde{S}$. Итак, $S_n \le \tilde{S}$ для любого n. А тогда и $S \le \tilde{S}$.

   В последних двух параграфах мы доказали, что $\tilde{S} \le S$ и $S \le \tilde{S}$. Значит, $\tilde{S} = S$.

4. Введём обозначения:

$$a_n^+ = \begin{cases} a_n, \text{ если } a_n \ge 0 \\ 0, \text{ если } a_n < 0 \end{cases}$$

$$a_n^- = \begin{cases} a_n, \text{ если } a_n < 0 \\ 0, \text{ если } a_n \ge 0 \end{cases}$$

   Тогда $\sum\limits_{n=1}^{\infty} a_n^+$ — это ряд, составленный только из неотрицательных членов ряда $\sum\limits_{n=1}^{\infty} a_n$. И также $\sum\limits_{n=1}^{\infty} a_n^-$ — это ряд, составленный только из отрицательных членов ряда $\sum\limits_{n=1}^{\infty} a_n$.

   Для любого натурального N выполнено

$$\sum_{n=1}^{N} a_n^+ + \sum_{n=1}^{N} a_n^- = \sum_{n=1}^{N} a_n$$

$$\sum_{n=1}^{N} a_n^+ - \sum_{n=1}^{N} a_n^- = \sum_{n=1}^{N} |a_n|$$

   А значит, частичные суммы рядов можно выразить так:

$$\sum_{n=1}^{N} a_n^+ = \frac{1}{2}\left(\sum_{n=1}^{N} a_n + \sum_{n=1}^{N} |a_n|\right)$$

$$\sum_{n=1}^{N} a_n^- = \frac{1}{2}\left(\sum_{n=1}^{N} a_n - \sum_{n=1}^{N} |a_n|\right)$$

   Правые части написанных равенств имеют предел при $N \to \infty$, ведь по условию ряд $\sum\limits_{n=1}^{\infty} a_n$ абсолютно сходится. А тогда и левые части равенств имеют предел при $N \to \infty$, то есть ряды $\sum\limits_{n=1}^{\infty} a_n^+$ и $\sum\limits_{n=1}^{\infty} a_n^-$ сходятся.

   Наконец, абсолютная сходимость этих рядов следует из того, что они знакопостоянны: $\sum\limits_{n=1}^{\infty} |a_n^+| = \sum\limits_{n=1}^{\infty} a_n^+$ и $\sum\limits_{n=1}^{\infty} |a_n^-| = -\sum\limits_{n=1}^{\infty} a_n^-$.

5. Напомним, ряд $\sum\limits_{n=1}^{\infty} b_n$ получен из ряда $\sum\limits_{n=1}^{\infty} a_n$ перестановкой. Как и в предыдущем пункте введём обозначения b_n^+ и b_n^-.

Поскольку $\sum\limits_{n=1}^{\infty} a_n^+$ абсолютно сходится и $a_n^+ \geq 0$, то по третьему пункту этой задачи ряд $\sum\limits_{n=1}^{\infty} b_n^+$ сходится. Снова по третьему пункту этой задачи получаем, что ряд $\sum\limits_{n=1}^{\infty} b_n^+$ сходится к числу $\sum\limits_{n=1}^{\infty} a_n^+$.

Далее, ряд $\left(-\sum\limits_{n=1}^{\infty} a_n^-\right)$ тоже абсолютно сходится и его члены неотрицательны. Тогда аналогично предыдущему рассуждению получаем, что ряд $\left(-\sum\limits_{n=1}^{\infty} b_n^-\right)$ сходится к числу $\left(-\sum\limits_{n=1}^{\infty} a_n^-\right)$. А тогда и ряд $\sum\limits_{n=1}^{\infty} b_n^-$ сходится к числу $\sum\limits_{n=1}^{\infty} a_n^-$.

Итак, мы получаем

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_n^+ + \sum_{n=1}^{\infty} a_n^- = \sum_{n=1}^{\infty} b_n^+ + \sum_{n=1}^{\infty} b_n^- = \sum_{n=1}^{\infty} b_n$$

А это и есть то, что мы хотели доказать: что пределы частичных сумм этих рядов равны.

Счётное пространство исходов

Задача 1

Мы бросаем честную монетку, пока не выпадет орёл. Обозначим орла за h и решку за t. У нас будут такие исходы и соответствующие вероятности:

1. $P(h) = \frac{1}{2}$
2. $P(th) = \frac{1}{4}$
3. $P(tth) = \frac{1}{8}$
4. $P(ttth) = \frac{1}{16}$
5. ...

Задача.

1. Найдите вероятность события "число бросков больше 2."
2. Найдите вероятность события "число бросков делится на 3."

Подсказка. Воспользуйтесь формулой для геометрической прогрессии.

Решение. Обозначим число бросков за N. Число бросков может быть любым натуральным числом.

1. $P(N > 3) = 1 - P(N = 1) - P(N = 2) - P(N = 3) = 1 - P(h) - P(th) - P(tth) = 1 - \frac{1}{2} - \frac{1}{4} - \frac{1}{8} = \frac{1}{8} = 0.125.$
2. Будем действовать аналогично предыдущему пункту:

$$P(N = 3) + P(N = 6) + \cdots + P(N = 3k) + \cdots = P(tth) + P(ttttth) + \cdots + P(\underbrace{t \ldots t}_{3k-1} h) + \cdots =$$

$$= \frac{1}{2^3} + \frac{1}{2^6} + \cdots + \frac{1}{2^{3k}} + \cdots = \sum_{k=1}^{\infty} \frac{1}{2^{3k}} = \sum_{k=1}^{\infty} \left(\frac{1}{8}\right)^k = \frac{1}{1 - \frac{1}{8}} - 1 = \frac{1}{7}$$

Задача 2

В случае конечного количества исходов мы определяли пространства с равновероятными исходами.

Докажите, что не существует вероятностного пространства со счётным количеством исходов, такого что все исходы равновероятны.

Подсказка. Проблема возникнет со свойством вероятности, про которое мы говорили в определении вероятностного пространства.

Решение. Докажем утверждение от противного. Пусть (Ω, F, P) — вероятностное пространство, $\Omega = (\omega_1, \omega_2, \dots)$ счётно и все исходы равновероятны. Обозначим вероятность исхода так: $P(\omega_i) = p$, где $i \in \mathbb{N}$.

По определению сумма вероятностей всех элементарных исходов должна равняться единице: $\sum\limits_{i=1}^{\infty} P(\omega_i) = 1$. С другой стороны, $\sum\limits_{i=1}^{\infty} P(\omega_i) = \sum\limits_{i=1}^{\infty} p = p \sum\limits_{i=1}^{\infty} 1$. Но этот ряд расходится, а значит, сумма вероятностей не может равняться 1. Мы получили противоречие с нашим изначальным предположением. Значит, мы доказали, что не существует вероятностного пространства со счётным количеством исходов, такого что все исходы равновероятны.

Задача 3

Мы работаем с тем же вероятностным пространством, что и раньше. Мы подбрасываем монетку до первого орла и получаем такие исходы и соответствующие вероятности:

1. $P(h) = \frac{1}{2}$
2. $P(th) = \frac{1}{4}$
3. $P(tth) = \frac{1}{8}$
4. $P(ttth) = \frac{1}{16}$
5. …

Задача. Определим случайную величину X так: $X(\underbrace{ttt\dots t}_{k}h) = \frac{1}{3^{k+1}}$. Найдите $E[X]$.

Подсказка. Здесь пригодится вторая задача из урока про ряды.

Решение. Пронумеруем элементарные исходы ω_k, как в условии: ω_k — это выпадение $\underbrace{t\dots t}_{k-1}h$. Обозначим через P_k вероятность k-ого исхода:

$$P_k = P(w_k) = \frac{1}{2^k}.$$

Также обозначим через x_k значение случайной величины X на k-ом исходе.
$X(\underbrace{ttt\dots t}_{k-1}h) = x_k = \frac{1}{3^k}$. Тогда по определению

$$\sum_{k=1}^{\infty} x_k P_k = \sum_{k=1}^{\infty} \frac{1}{3^k} \cdot \frac{1}{2^k} = \sum_{k=1}^{\infty} \left(\frac{1}{6}\right)^k = \frac{1}{1-\frac{1}{6}} - 1 = \frac{1}{5}.$$

Все члены ряда положительны, поэтому он сходится абсолютно и, значит, математическое ожидание случайной величины X определено: $E[X] = \frac{1}{5}$.

Задача 4

Мы работаем в том же вероятностном пространстве, что и в предыдущей задаче. Определим ещё одну случайную величину Y так: $Y(\underbrace{ttt\dots\ t}_{k}h) = (-1)^{k+1}$.

Найдите $E[Y]$.

Подсказка. Здесь пригодится вторая задача из урока про ряды.

Решение. Как и в предыдущем решении, пронумеруем элементарные исходы $\omega_k : \omega_k$ — это выпадение $\underbrace{t\dots t}_{k-1}h$. Обозначим через P_k вероятность k-ого исхода:

$$P_k = P(w_k) = \frac{1}{2^k}.$$

Также обозначим через y_k значение случайной величины Y на k-ом исходе.
$Y(\underbrace{ttt\dots t}_{k-1}h) = y_k = (-1)^k$. Тогда

$$\sum_{k=1}^{\infty} y_k P_k = \sum_{k=1}^{\infty}(-1)^k \cdot \frac{1}{2^k} = \sum_{k=1}^{\infty}\left(\frac{-1}{2}\right)^k = \frac{1}{1-\frac{-1}{2}} - 1 = -\frac{1}{3}.$$