



Теория вероятностей

Конспекты лекций и семинаров

ЛЕКТОР: Д.А. ШАБАНОВ

Конспекты вели Денис Беляков, Никита Попов и Алексей Хачиянц

НИУ ВШЭ, 2016-2017

Оглавление

Глава 1

Лекции

1.1 Лекция от 09.09.2016

1.1.1 Введение. Принцип устойчивости частоты

Чем занимается теория вероятностей? Она изучает *случайные* явления. Допустим, мы провели какой-либо эксперимент. Можем ли мы что-то заранее сказать о результате?

- Если да, то результат называют *детерминированным*. Пример такого эксперимента — выбрасывание кирпича из окна. Очевидно, что кирпич упадёт на землю¹ и результат предопределён. Такие задачи изучают в той же линейной алгебре или где-либо ещё, но не в теории вероятностей.
- А теперь предположим, что заранее сказать, каков будет результат, невозможно. Например, точно сказать, какой стороной упадёт подброшенная монетка, вряд ли получится. Тогда результат называют *недетерминированным*. Именно задачи с недетерминированным результатом и изучаются в теории вероятностей.

Небольшое историческое отступление — вообще говоря, теория вероятностей появилась в связи с изучением азартных игр наподобие рулетки ещё в средних веках. Но тогда она представляла собой скорее набор эмпирических фактов, чем полноценную науку. Теория вероятностей стала такой, какой она является сейчас, лишь в XX веке благодаря трудам А.Н. Колмогорова.

Хорошо, а как изучаются случайные процессы? Ну выпала решка, и что? На самом деле теория вероятностей не о единичных экспериментах, а об *асимптотике*. Это значит следующее: если проводить серию одинаковых экспериментов, то теория вероятностей поможет предсказать частоту, с которой будет появляться какой-либо ответ.

Теория вероятностей держится на крайне важном *принципе устойчивости частоты*. Перед тем, как ввести формальное определение, рассмотрим пару экспериментов, связанных с подбрасыванием монетки:

- В XVIII веке Жорж-Луи Леклерк де Бюффон провёл эксперимент, подбросив монетку 4040 раз. Из них в 2048 бросках выпал герб. В итоге частота составила около 0.506.
- В XIX веке пошли ещё дальше — Карл Пирсон подбросил монетку 24000 раз.² У него получилось так, что герб выпал 12012 раз. В итоге частота составила 0.5005.

¹Если его не запустили с первой космической скоростью, конечно.

²Оставим вопрос о том, как он не поленился проверить это, без ответа.

Отсюда видно, что эксперименты дают частоту, близкую к $1/2$.

Неформально говоря, принцип формулируется так: если мы проводим серию одинаковых экспериментов, то количество появлений одного определённого ответа при делении на число экспериментов сходится к некоторому числу $p \in [0, 1]$. Теперь можно ввести формальное определение.

Принцип устойчивости частоты. Пусть A — некоторое событие, а $v_n(A)$ — число экспериментов, в которых происходит событие A среди первых n . Тогда

$$\lim_{n \rightarrow \infty} \frac{v_n(A)}{n} = p, \quad p \in [0, 1].$$

Получаемое число p называют *вероятностью* события A и обозначают $P(A)$. Например, $P(\text{встретить живого динозавра на улице}) = 0$, так как они все вымерли.

1.1.2 Вероятностное пространство. Простейшие свойства вероятности

Именно после введения этого понятия Колмогоровым теория вероятностей перестала быть прежней. Введём определение для дискретного случая (общий оставим на потом):

Определение 1. Дискретным вероятностным пространством называется пара (Ω, P) , где Ω — множество элементарных исходов, а P — вероятность на Ω .

Множество элементарных исходов Ω — некоторое конечное или счётное множество. Элемент $\omega \in \Omega$ называют *элементарным исходом*. Полагается, что в случайном эксперименте обязательно получается один и только один элементарный исход.

Примеры множеств элементарных исходов:

1. $\Omega = \{O, P\}$ — бросок монеты.
2. $\Omega = \{\omega_1, \dots, \omega_6\}$, где $\omega_i = \{\text{выпало } i \text{ очков}\}$ — бросок игрального кубика.
3. $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$, где $\omega_i = \{\text{на данный момент горит } i \text{ зданий}\}$ — предсказание пожаров в городе.

Определение 2. Подмножество $A \subseteq \Omega$ называется *событием* на вероятностном пространстве (Ω, P) .

Пример события: пусть подбрасывают игральную кость, и $A = \{\text{выпало чётное число очков}\}$. Тогда $A = \{\omega_2, \omega_4, \omega_6\}$.

Определение 3. Отображение $P : \Omega \rightarrow [0, 1]$ называют *вероятностью*³, если $\sum_{\omega \in \Omega} P(\omega) =$

1. В случае счётного множества Ω данный ряд должен сходиться абсолютно.

Пусть у нас есть некоторое событие A на вероятностном пространстве (Ω, P) . Как посчитать его вероятность?

Определение 4. Вероятностью события $A \subseteq \Omega$ называют $P(A) = \sum_{\omega \in A} P(\omega)$.

³В общем случае вероятность ещё могут называть *вероятностной мерой*.

Определение 5. Пусть A — некое событие. Тогда *дополнением* к событию A называют событие $\bar{A} = \Omega \setminus A$.

Перед тем, как идти дальше, напомним определение дизъюнктного объединения.

Определение 6. Пусть есть множества A_1, A_2, \dots, A_n . Тогда *дизъюнктным объединением* множеств называют объединение попарно непересекающихся “копий” множеств:

$$\bigsqcup_{i=1}^n A_i = \bigcup_{i=1}^n \{(x, i) \mid x \in A_i\}.$$

В нашем случае полагается, что если пишут дизъюнктное объединение, то множества попарно не пересекаются.

Рассмотрим некоторые свойства вероятности.

Теорема 1 (Простейшие свойства вероятности). *Для любого дискретного вероятностного пространства (Ω, \mathbf{P}) выполняется следующее:*

1. $\mathbf{P}(\Omega) = 1$, $\mathbf{P}(\emptyset) = 0$.
2. *Конечная аддитивность:* $\mathbf{P}\left(\bigsqcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbf{P}(A_i)$.
3. $\mathbf{P}(A) + \mathbf{P}(\bar{A}) = 1$.
4. $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.
5. *Для любого набора событий A_1, A_2, \dots, A_n* $\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbf{P}(A_i)$.
6. *Счётная аддитивность:* $\mathbf{P}\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbf{P}(A_i)$.

Последнее свойство выполняется только для счётного Ω .

Доказательство. Распишем доказательства для каждого пункта по отдельности:

1. $\mathbf{P}(\Omega) = 1$ следует из определения вероятности, а $\mathbf{P}(\emptyset) = 0$ следует из определения вероятности события.
2. Случай с конечным множеством Ω очевиден. Тогда положим, что Ω счётно, то есть $\Omega = \{\omega_i \mid i \in \mathbb{N}\}$.

Пусть есть некоторое событие A . Тогда представим его вероятность в удобном для нас виде:

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\omega) = \sum_{i: \omega_i \in A} \mathbf{P}(\omega_i) = \lim_{N \rightarrow \infty} \sum_{\substack{i: \omega_i \in A \\ i < N}} \mathbf{P}(\omega_i).$$

Теперь распишем вероятность дизъюнктного объединения событий A_1, A_2, \dots, A_n :

$$\begin{aligned} \mathbf{P}\left(\bigsqcup_{i=1}^n A_i\right) &= \sum_{\omega \in \bigsqcup A_i} \mathbf{P}(\omega) = \lim_{N \rightarrow \infty} \left(\sum_{\substack{i: \omega_i \in \bigsqcup A_i \\ i < N}} \mathbf{P}(\omega_i) \right) = \lim_{N \rightarrow \infty} \sum_{i=1}^n \sum_{\substack{j: \omega_j \in A_i \\ j < N}} \mathbf{P}(\omega_j) = \\ &= \sum_{i=1}^n \lim_{N \rightarrow \infty} \sum_{\substack{j: \omega_j \in A_i \\ j < N}} \mathbf{P}(\omega_j) = \sum_{i=1}^n \mathbf{P}(A_i). \end{aligned}$$

3. Согласно второму пункту, $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$.
4. Так как $A \cup B = A \cup (B \setminus A)$ и $A \cap (B \setminus A) = \emptyset$, то $P(A \cup B) = P(A) + P(B \setminus A)$. Далее, заметим, что $B \setminus A = B \setminus (A \cap B)$. Тогда $P(A \cup B) = P(A) + P(B \setminus (A \cap B))$.
Рассмотрим второй член. Заметим, что $P(A \cap B) + P(B \setminus (A \cap B)) = P(B)$. Тогда получаем, что $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
5. Докажем это утверждение по индукции. База была доказана в пункте 4 (так как $P(A \cap B) \geq 0$). Теперь рассмотрим шаг индукции. Пусть утверждение верно для какого-то m . Тогда $P\left(\bigcup_{i=1}^{m+1} A_i\right) \leq P(A_{m+1}) + P\left(\bigcup_{i=1}^m A_i\right) \leq \sum_{i=1}^{m+1} P(A_i)$.
6. За доказательством этого пункта обращайтесь к учебнику матанализа на тему частичных сумм и абсолютной сходимости.⁴

□

1.1.3 Классическая модель. Примеры

Пусть (Ω, P) — некоторое конечное вероятностное пространство, при этом все элементарные исходы равновероятны. Тогда легко посчитать вероятность элементарного исхода: $P(\omega) = \frac{1}{|\Omega|}$ для всех $\omega \in \Omega$. Такую модель называют *классической*.

Как посчитать вероятность события в классической модели? Очень просто: $P(A) = \frac{|A|}{|\Omega|}$ для всех $A \subseteq \Omega$.

Рассмотрим некоторые примеры классических моделей.

1. Бросок монетки. В таком случае $\Omega = \{O, P\}$ и $P(O) = P(P) = \frac{1}{2}$.
2. Бросок двух монеток. С этой моделью связано одно заблуждение Д'Аламбера. Он рассуждал следующим образом: так как $\Omega = \{OO, PP, OP\}$, то $P(OO) = P(PP) = P(OP) = \frac{1}{3}$. Но это опровергается экспериментами. И как это исправить? Есть два варианта:
 - (а) Можно сказать, что модель не является классической и поправить вероятности: $P(OO) = P(PP) = \frac{1}{4}$, $P(OP) = \frac{1}{2}$.
 - (б) А можно просто изменить множество элементарных исходов. Начнём учитывать порядок выпадения: $\Omega = \{OO, PP, OP, PO\}$. Такая модель уже является классической.

Рассуждая в стиле Д'Аламбера, можно прийти к выводу, что вероятность встретить живого динозавра на улице равна $\frac{1}{2}$, ведь его можно либо встретить, либо не встретить.

3. Бросок n монет. В таком случае вероятностное пространство будет устроено следующим образом: $\Omega = \{\omega = (\omega_1, \omega_2, \dots, \omega_n) \mid \omega_i \in \{O, P\}\}$. Легко понять, что в данной модели 2^n элементарных исходов.

Примечание. Вероятностное пространство такого вида называют *симметрической схемой Бернулли*.

⁴На самом деле я попробую найти доказательство. Когда-нибудь. Но не сейчас. (А.Х.)

Но данная модель является классической только тогда, когда монетки “честные”, то есть которые падают орлом или решкой вверх равновероятно. Если же это не так, то вероятность элементарного исхода $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ задаётся следующей формулой: $P(\omega) = p^{\sum_{i=1}^n \omega_i} (1-p)^{n-\sum_{i=1}^n \omega_i}$.

4. Урновые схемы (размещение частиц по ячейкам). Пусть есть n различных шаров в ящике. Мы случайным образом вынимаем m шаров. Вопрос: каков размер множества элементарных исходов? Сначала приведём ответ, после чего докажем его.

Порядок? Возврат?	Упорядоченный набор	Неупорядоченный набор
С возвратом	n^m	$C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}$
Без возврата	$A_n^m = \frac{n!}{(n-m)!}$	$C_n^m = \frac{n!}{m!(n-m)!}$

Доказательство. Будем доказывать утверждения от верхнего левого против часовой стрелки.

- (a) Пусть набор упорядочен и можно возвращать. Тогда любой элемент набора можно получить n способами (так как все элементы можно вернуть). Отсюда получаем n^m .
- (b) Теперь положим, что набор упорядочен, но возвращать нельзя. Тогда первый элемент можно выбрать n способами, второй — $n-1$ способом и так далее до m -го элемента, который можно выбрать $n-m+1$ способом. По правилу умножения получаем $\frac{n!}{(n-m)!} = A_n^m$.
- (c) Рассмотрим случай, когда набор неупорядочен и возвращать нельзя. Тогда необходимо посчитать количество способов выбрать k шаров из m . Достаточно логично, что это равно $\frac{A_n^m}{m!} = \frac{n!}{m!(n-m)!} = C_n^m$, так как в последовательности нам не важен порядок.
- (d) Осталось рассмотреть последний случай — неупорядоченный набор с возвратом. В этом случае нам достаточно указать, сколько раз мы выбрали каждый шар. Как это сделать? Воспользуемся методом точек и перегородок. Пусть есть m точек и нужно распределить их по n группам. Для этого нужно использовать $n-1$ перегородку. Тогда задача сводится к нахождению количества способов выбрать m элементов из $n+m-1$. А это равно $C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}$.

□

1.1.4 Условная вероятность. Формула полной вероятности

Пусть (Ω, P) — дискретное вероятностное пространство.

Определение 7. Пусть $A \subseteq \Omega$ — некоторое событие и $B \subseteq \Omega$ — другое событие, причём $P(B) > 0$. Тогда *условной вероятностью события A при условии B* называют

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

Если $P(B) = 0$, то положим, что $P(A | B) = 0$ для любого события $A \subseteq \Omega$.

Условную вероятность можно воспринимать следующим образом: сузим множество элементарных исходов до B и посчитаем вероятность события A на полученном множестве.

Примечание. Если $P(B) > 0$, то $\tilde{P}(A) = P(A | B)$ тоже является вероятностью на Ω .

Определение 8. Пусть B_1, B_2, \dots, B_n — некоторые события на Ω такие, что $\bigsqcup_{i=1}^n B_i = \Omega$. Тогда этот набор событий называется (конечным) *разбиением* Ω .

Теперь докажем важную формулу:

Формула полной вероятности. Пусть B_1, B_2, \dots, B_n — разбиение Ω . Тогда для любого события $A \subseteq \Omega$ верно, что $P(A) = \sum_{i=1}^n P(A | B_i) P(B_i)$.

Доказательство. Так как $A \cap \Omega = A$ и $\bigsqcup_{i=1}^n B_i = \Omega$, то $P(A) = P\left(A \cap \left(\bigsqcup_{i=1}^n B_i\right)\right)$. Заметим, что $A \cap \left(\bigsqcup_{i=1}^n B_i\right) = \bigsqcup_{i=1}^n (A \cap B_i)$. Тогда $P(A) = \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A | B_i) P(B_i)$. \square

Заметим, что формула полной вероятности работает и в случае, когда $P(B_i) = 0$ для какого-то i .

1.2 Лекция от 16.09.2016

1.2.1 Классические задачи теории вероятностей

Задача 1 (Задача о сумасшедшей старушке). *Есть самолёт на n мест, в который садятся n пассажиров. Первой в него заходит безумная старушка, которая садится на случайное место. Каждый следующий пассажир действует по следующему правилу: садится на своё место, если оно свободно, и на случайное, если своё занято. С какой вероятностью*

1. *последний пассажир сядет на своё место?*
2. *предпоследний пассажир сядет на своё место?*
3. *и последний, и предпоследний пассажир сядут на свои места?*

Решение. У первого пункта есть элементарное решение. Пусть при некоторой рассадке пассажиров последний пассажир сел не на свое место (такую рассадку назовем неудачной). Тогда до прихода последнего пассажира его место было занято пассажиром A (A может быть и сумасшедшей старушкой). В момент прихода пассажира A перед ним стоит выбор — какое место занять. В рассматриваемой рассадке он занимает место последнего пассажира. Но с той же вероятностью он мог занять и место старушки, и в дальнейшем все пассажиры, включая последнего, займут свои собственные места. (Конечно, нужно еще пояснить, почему в момент прихода пассажира A старушкино место все еще свободно. Но это действительно так — нетрудно проследить, что пока старушкино место свободно, среди всех еще не вошедших пассажиров есть ровно один, чье место уже занято. Как только

очередной пассажир занимает старушкино место, все остальные будут садиться только на свои места.) Таким образом, каждой неудачной рассадке соответствует удачная, которая может случиться с той же вероятностью. Это говорит о том, что ровно в половине случаев рассадка будет неудачной.

Теперь рассмотрим формальное решение для первого пункта. Пусть $A = \{\text{последний сядет на своё место}\}$. Если⁵ $n = 2$, то $P(A) = \frac{1}{2}$. Теперь рассмотрим случай $n = 3$. Пусть $B_i = \{\text{бабушка села на } i\text{-е место}\}$. По формуле полной вероятности

$$P(A) = P(A | B_1) P(B_1) + P(A | B_2) P(B_2) + P(A | B_3) P(B_3) = \frac{1}{3} \left(1 + \frac{1}{2} + 0 \right) = \frac{1}{2}.$$

Намечается закономерность. Попробуем сформулировать гипотезу и доказать её:

Гипотеза. Для любого n вероятность того, что последний пассажир сядет на своё место, равна $\frac{1}{2}$.

Доказательство. По индукции. База ($n = 2$) была доказана ранее. Теперь предположим, что утверждение верно для всех $k < n$. Тогда докажем, что утверждение верно для $k = n$. Опять же, распишем вероятность по формуле полной вероятности:

$$P(A) = \sum_{i=1}^n P(A | B_i) P(B_i).$$

Заметим, что $P(B_i) = \frac{1}{n}$. Теперь посмотрим на значения условных вероятностей:

$$P(A | B_i) = \begin{cases} 1, & i = 1 \\ 0, & i = n \\ \frac{1}{2}, & 2 \leq i \leq n - 1 \end{cases}$$

Последнее условие на значение условной вероятности следует из шага индукции, так как i -й пассажир “становится” бабушкой. Тогда $P(A) = \frac{1}{n} \left(1 + \frac{1}{2} + \dots + \frac{1}{2} + 0 \right) = \frac{1}{2}$. \square

Перейдём ко второму пункту задачи. Пусть $C = \{\text{предпоследний пассажир сел на своё место}\}$. Тогда рассмотрим $P(C)$ в случае, когда $n = 3$. Он сможет сесть на своё место тогда и только тогда, когда бабушка не села на его место. Тогда $P(C) = \frac{2}{3}$. Попробуем доказать гипотезу, аналогичную случаю с последним пассажиром.

Гипотеза. Для любого n вероятность того, что предпоследний пассажир сядет на своё место, равна $\frac{2}{3}$.

Доказательство. Доказательство практически такое же, как и для последнего пассажира, с тем отличием, что значения условной вероятности будут несколько другие:

$$P(C | B_i) = \begin{cases} 1, & i = 1, n \\ 0, & i = n - 1 \\ \frac{2}{3}, & 2 \leq i \leq n - 2 \end{cases}$$

Тогда $P(C) = \frac{1}{n} \left(1 + \frac{2}{3} + \dots + \frac{2}{3} + 0 + 1 \right) = \frac{2}{3}$ \square

⁵Это уже вертолёт, скорее. (Д.А. Шабанов)

Аналогичными рассуждениями можно доказать, что вероятность того, что i -й с конца пассажир сядет на своё место, равна $\frac{i}{i+1}$.

Теперь приступим к третьему пункту. Докажем следующее утверждение:

Гипотеза. Вероятность этого события (обозначим его за D) равна $\frac{1}{3}$.

Доказательство. Опять же, по индукции. Базой служит случай $n = 3$. В таком случае условие выполнимо тогда и только тогда, когда бабушка сядет на своё место. Тогда вероятность равна $\frac{1}{3}$ и база верна.

Перейдём к шагу индукции. Рассуждаем абсолютно аналогично: по формуле полной вероятности $P(D) = \sum_{i=1}^n P(D | B_i) P(B_i)$. Значения условных вероятностей равны

$$P(D | B_i) = \begin{cases} 1, & i = 1 \\ 0, & i = n - 1, n \\ \frac{1}{3}, & 2 \leq i \leq n - 2 \end{cases}$$

$$\text{и } P(D) = \frac{1}{n} \left(1 + \frac{1}{3} + \dots + \frac{1}{3} + 0 + 0 \right) = \frac{1}{3}. \quad \square$$

Заметим, что $P(D) = P(A)P(C)$. Это ~~счастливое совпадение~~ неспроста. □

Задача 2 (Задача об удачливом студенте). *Студент начал готовиться к экзамену слишком поздно и выучил только k билетов из n . Студент решил схитрить и выбрать место в очереди такое, чтобы вероятность получить выученный билет была максимальной. Какое место ему выбрать?*

Решение. Пусть $A_s = \{\text{студент вытащил хороший билет, если он встал } s\text{-тым в очереди}\}$. Докажем следующую гипотезу:

Гипотеза. $P(A_s) = \frac{k}{n}$.

Доказательство. Введём разбиение B_0, B_1, \dots, B_k пространства Ω , где $B_i = \{\text{до студента взяли ровно } i \text{ хороших билетов}\}$. Тогда по формуле полной вероятности

$$P(A_s) = \sum_{i=0}^k P(A_s | B_i) P(B_i).$$

Посчитаем $P(B_i)$. Для определения количества успешных исходов нужно выбрать i билетов из k хороших и $s - i - 1$ из $n - k$ плохих. Тогда

$$P(B_i) = \frac{\binom{k}{i} \binom{n-k}{s-i-1}}{\binom{n}{s-1}}.$$

Теперь надо определить значение условной вероятности $P(A_s | B_i)$. Так как есть $n - s + 1$ невыбранный билет и $k - i$ из них изучены, то $P(A_s | B_i) = \frac{k-i}{n-s+1}$. Тогда

$$P(A) = \sum_{i=0}^k \frac{k-i}{n-s+1} \frac{\binom{k}{i} \binom{n-k}{s-i-1}}{\binom{n}{s-1}} = \frac{k}{n} \sum_{i=0}^k \frac{\binom{k-1}{i} \binom{n-k}{s-i-1}}{\binom{n-1}{s-1}}.$$

Теперь заметим, что сумма справа равна 1, так как она соответствует разбиению пространства в случае, когда всего задач $n - 1$. Тогда $P(A_s) = \frac{k}{n}$. □

В итоге как ни вставай — всё равно никакой разницы не будет.⁶ □

⁶Ибо по закону подлости попадётся невыученный билет.

1.2.2 Теорема Байеса

Теорема 2 (Байес). Пусть $\{B_1, B_2, \dots, B_n\}$ — разбиение Ω , причём $P(B_i) > 0$ для всех $i \in \{1, 2, \dots, n\}$. Тогда для любого события A такого, что $P(A) > 0$, выполняется

$$P(B_i | A) = \frac{P(A | B_i) P(B_i)}{\sum_{j=1}^n P(A | B_j) P(B_j)}$$

Доказательство. По определению условной вероятности $P(B_i | A) = \frac{P(A \cap B_i)}{P(A)} = \frac{P(A|B_i) P(B_i)}{P(A)}$. Тогда, применяя формулу полной вероятности для $P(A)$, получаем желаемое. \square

1.2.3 Независимость событий

Определение 9. События A и B на вероятностном пространстве (Ω, P) называются *независимыми*, если $P(A \cap B) = P(A)P(B)$. Иногда используется обозначение $A \perp B$.

Рассмотрим некоторые примеры независимых событий:

1. Задача про сумасшедшую бабушку: События $A = \{\text{последний сядет на своё место}\}$ и $B = \{\text{предпоследний сядет на своё место}\}$ независимы, как было доказано ранее.
2. Бросок игральной кости. Пусть $A = \{\text{выпало чётное число очков}\}$, а $B = \{\text{выпало число очков, кратное 3}\}$. Докажем, что они независимы. Заметим, что $A \cap B = \{\text{число очков кратно 6}\}$. Тогда $P(A \cap B) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = P(A)P(B)$.

Данное определение работает только для двух событий. Можно ли как-то его обобщить? Попробуем ввести аналогично:

Определение 10. События A_1, \dots, A_n называются *попарно независимыми*, если для любых i и j таких, что $i \neq j$, A_i независимо от A_j .

Однако обобщение обычно вводят по другому:

Определение 11. События A_1, A_2, \dots, A_n называются *независимыми по совокупности*, если для любого множества $\{i_1, i_2, \dots, i_k\} \subseteq \{1, 2, \dots, n\}$ выполняется, что $P(A_{i_1} \cap \dots \cap A_{i_k}) = \prod_{j=1}^k P(A_{i_j})$.

Примечание. Если говорят про независимость событий и не указывают тип, то обычно подразумевают независимость по совокупности.

Стоит заметить, что попарная независимость — гораздо более слабое условие, чем независимость по совокупности. Приведём пример. Пусть есть тетраэдр, грани которого покрашены следующим образом: первая грань покрашена в красный, вторая — в синий, третья — в зелёный, а четвёртая — во все три цвета сразу. Его подбрасывают. Введём три события: $A_K = \{\text{на нижней грани есть красный цвет}\}$, $A_S = \{\text{на нижней грани есть синий цвет}\}$ и $A_Z = \{\text{на нижней грани есть зелёный цвет}\}$. Очевидно, что вероятность любого события — $\frac{1}{2}$, а любой пары событий — $\frac{1}{4}$. Однако вероятность объединения всех трёх событий равна $\frac{1}{4}$, а не $\frac{1}{8}$. Следовательно, события попарно независимы, но не независимы по совокупности.

Упражнение. Приведите пример n событий таких, что любой набор из $n - 1$ события независим, а все n событий вместе зависимы.

Для независимости выполняются следующие свойства:

1. $A \perp A \iff P(A) = 0$ или $P(A) = 1 \iff$ для любого события B $A \perp B$.
2. $A \perp B \implies \bar{A} \perp B$.
3. Если A_1, A_2, \dots, A_n независимы в совокупности, то любой набор B_1, B_2, \dots, B_n такой, что $B_i = A_i$ или $B_i = \bar{A}_i$, тоже независим.

1.2.4 Случайные величины в дискретных вероятностных пространствах

Определение 12. Пусть (Ω, P) — дискретное вероятностное пространство. Тогда любое отображение $\xi : \Omega \rightarrow \mathbb{R}$ называется *случайной величиной*.

Примечание. В литературе случайную величину часто сокращают до с.в.

Рассмотрим некоторые примеры случайных величин:

1. Широко распространённым примером случайной величины служит индикатор какого-либо события. Введём определение.

Определение 13. Пусть $A \in \Omega$ — событие. Тогда *индикатором* события A называют называется случайная величина, равная

$$I_A(\omega) = \begin{cases} 1, & \omega \in A; \\ 0, & \omega \notin A; \end{cases}$$

2. Бросок игральной кости. Тогда случайной величиной будет число выпавших очков.
3. Схема Бернулли: $\Omega = \{\omega = (\omega_1 \dots \omega_n), \omega_i \in \{0, 1\}\}$. В данном случае вводится случайная величина, равная $\xi(\omega) = \sum_{i=1}^n \omega_i$. Ещё её называют *числом успехов в схеме Бернулли*.

1.3 Лекция от 23.09.2016

1.3.1 Распределение случайной величины. Примеры распределений

Рассмотрим некоторое дискретное пространство (Ω, P) . Тогда случайная величина ξ на этом пространстве принимает не более, чем счетное множество значений. Пусть $X = (x_1, x_2, \dots, x_n, \dots)$ — множество значений случайной величины ξ . Введём событие $A_i = \{\omega : \xi(\omega) = x_i\}$. Его можно интерпретировать, как событие $\{\xi \text{ приняло значение } x_i\}$. Для удобства будем использовать обозначение $A_i := \{\xi = x_i\}$. Также введем обозначение p_i для вероятности события A_i : $p_i = P(A_i) = P(\xi = x_i)$.

Определение 14. Вместе множество значений $X = (x_1, x_2, \dots, x_n, \dots)$ и набор вероятностей $(p_1, p_2, \dots, p_n, \dots)$ образуют то, что называется *распределением случайной величины* ξ .

Понятно, что каждому числу x_i сопоставлено число p_i . Легко заметить, что A_i образуют разбиение пространства Ω . Тогда $\sum_{i=0}^{|X|} p_i = 1$.

Рассмотрим некоторые известные примеры распределений:

1. Распределение Бернулли.

Определение 15. Случайная величина ξ имеет *распределение Бернулли*, если она принимает всего два значения, 1 или 0, с заранее известными вероятностями p и $q \equiv 1 - p$. Обозначение: $\xi \sim \text{Bern}(p)$.

Легко понять, что $X = \{0, 1\}$, $P(\xi = 1) = p$, $P(\xi = 0) = q$. Принято говорить, что событие $\{\xi = 1\}$ соответствует “успеху”, а событие $\{\xi = 0\}$ — “неудаче”. Эти названия достаточно условные, и в зависимости от конкретной задачи могут быть заменены на противоположные.

2. Биномиальное распределение.

Определение 16. *Биномиальное распределение* в теории вероятностей — распределение количества “успехов” ξ в последовательности из n независимых случайных экспериментов, таких, что вероятность “успеха” в каждом из них постоянна и равна p . Обозначение: $\xi \sim \text{Bin}(n, p)$.

Пусть x_1, \dots, x_n — последовательность независимых случайных величин с одинаковым распределением Бернулли с параметром p . Тогда $\xi = \sum_{i=1}^n x_i$. Посчитаем $P(\xi = k)$. Для этого необходимо выбрать k исходов из n и сказать, что они успешны. Тогда $P(\xi = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

3. Пуассоновское распределение.

Определение 17. Говорят, что случайная величина ξ имеет *распределение Пуассона*, если она принимает любое значение⁷ $k \in \mathbb{N}$ с вероятностью $\frac{\lambda^k}{k!} e^{-\lambda}$, где $\lambda > 0$ — некоторый параметр. Обозначение: $\xi \sim \text{Pois}(\lambda)$.

Почему оно вводится именно так? Об этом будет рассказано позднее.

4. Геометрическое распределение.

Определение 18. Говорят, что случайная величина ξ имеет *геометрическое распределение*, если она принимает любое значение $k \in \mathbb{N}$ с вероятностью $p(1 - p)^{k-1}$. Обозначение: $\xi \sim \text{Geom}(p)$.

⁷Есть путаница, откуда считать — с нуля или же с единицы. Будем считать, что начинаем с единицы.

1.3.2 Независимость случайных величин

Далее множество значений случайной величины ξ на вероятностном пространстве (Ω, \mathcal{P}) будем обозначать как $\xi(\Omega)$.

Определение 19. Пусть ξ, η — случайные величины на вероятностном пространстве (Ω, \mathcal{P}) . Будем говорить, что эти случайные величины *независимы*, если для любых $a \in \xi(\Omega), b \in \eta(\Omega)$ события $\{\xi = a\}$ и $\{\eta = b\}$ независимы.

По определению независимых событий получаем, что

$$\mathcal{P}(\{\xi = a\} \cap \{\eta = b\}) = \mathcal{P}(\xi = a) \mathcal{P}(\eta = b).$$

Примечание. Для простоты левую часть часто обозначают за $\mathcal{P}(\xi = a, \eta = b)$.

Теперь обобщим это понятие на произвольное количество случайных величин:

Определение 20. Пусть ξ_1, \dots, ξ_n — случайные величины на вероятностном пространстве (Ω, \mathcal{P}) и известно, что ξ_i принимает значения $(a_1^{(i)}, a_2^{(i)}, \dots, a_n^{(i)})$. Тогда будем говорить, что случайные величины ξ_1, \dots, ξ_n *независимы (в совокупности)*, если $\forall j_1, \dots, j_n$ выполнено:

$$\mathcal{P}(\xi_1 = a_1^{(i)}, \dots, \xi_n = a_n^{(i)}) = \prod_{k=1}^n \mathcal{P}(\xi_k = a_{j_k}^{(k)}).$$

Упражнение. Пусть A, B — некоторые события над (Ω, \mathcal{P}) . Показать, что A и B независимы тогда и только тогда, когда их индикаторы I_A и I_B независимы.

Упражнение. Показать, что случайные величины ξ_1, \dots, ξ_n независимы тогда и только тогда, когда

$$\forall \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R} \quad \mathcal{P}(\xi_1 = \alpha_1, \dots, \xi_n = \alpha_n) = \prod_{i=1}^n \mathcal{P}(\xi_i = \alpha_i)$$

1.3.3 Математическое ожидание. Свойства математического ожидания

Опять же, зафиксируем некоторое дискретное вероятностное пространство (Ω, \mathcal{P}) .

Определение 21. Математическим ожиданием случайной величины ξ называют величину $E[\xi] = \sum_{\omega \in \Omega} \xi(\omega) \mathcal{P}(\omega)$.

Примечание. Если Ω счётно, то ряд $\sum_{\omega \in \Omega} \xi(\omega) \mathcal{P}(\omega)$ должен сходиться абсолютно; иначе сумма ряда не определена однозначно, так как порядок перебора ω не задан (см. теорему Римана о перестановке членов условно сходящегося ряда).

Смысл у математического ожидания случайной величины простой — его можно понимать как среднее значение этой случайной величины.

Рассмотрим некоторые свойства математического ожидания.

Теорема 3 (Простейшие свойства математического ожидания). Пусть ξ, η — некоторые случайные величины. Тогда выполняются следующие свойства:

1. Матожидание линейно: $\forall a, b \in \mathbb{R} \quad E[a\xi + b\eta] = a E[\xi] + b E[\eta]$;
2. Оно сохраняет относительный порядок: если $\xi \leq \eta$ (то есть для любого $\omega \in \Omega$ $\xi(\omega) \leq \eta(\omega)$), то $E[\xi] \leq E[\eta]$;
3. Модуль математического ожидания меньше математического ожидания модуля: $|E[\xi]| \leq E[|\xi|]$;
4. $E[\xi] = \sum_{a \in \xi(\Omega)} a P(\xi = a)$;
5. Для любой функции $\varphi(x)$ выполняется $E[\varphi(\xi)] = \sum_{a \in \xi(\Omega)} \varphi(a) P(\xi = a)$;
6. Если $P(\xi = c) = 1$ для некоторой константы c , то $E[\xi] = c$;
7. Если $\xi \geq 0$, то $E[\xi] \geq 0$;
8. Если $E[\xi] = 0$ и $\xi \geq 0$, то $P(\xi = 0) = 1$;
9. Если ξ и η независимы, то $E[\xi\eta] = E[\xi] E[\eta]$.

Доказательство. Докажем пункты по порядку:

1. $E[a\xi + b\eta] = \sum_{\omega \in \Omega} (a\xi + b\eta)(\omega) P(\omega) = \sum_{\omega \in \Omega} a\xi(\omega) P(\omega) + \sum_{\omega \in \Omega} b\eta(\omega) P(\omega) = a E[\xi] + b E[\eta]$.
2. $E[\xi] = \sum_{\omega \in \Omega} \xi(\omega) P(\omega) \leq \sum_{\omega \in \Omega} \eta(\omega) P(\omega) = E[\eta]$.
3. $E[-|\xi|] \leq E[\xi] \leq E[|\xi|] \implies -E[|\xi|] \leq E[\xi] \leq E[|\xi|] \implies |E[\xi]| \leq E[|\xi|]$
- 4.

$$\begin{aligned} E[\xi] &= \sum_{\omega \in \Omega} \xi(\omega) P(\omega) = \sum_{a \in \xi(\Omega)} \sum_{\omega: \xi(\omega)=a} \xi(\omega) P(\omega) = \sum_{a \in \xi(\Omega)} \left(a \sum_{\omega: \xi(\omega)=a} P(\omega) \right) = \\ &= \sum_{a \in \xi(\Omega)} a P(\xi = a) \end{aligned}$$

5.

$$\begin{aligned} E[\varphi(\xi)] &= \sum_{\omega \in \Omega} \varphi(\xi(\omega)) P(\omega) = \sum_{a \in \xi(\Omega)} \sum_{\omega: \xi(\omega)=a} \varphi(\xi(\omega)) P(\omega) = \\ &= \sum_{a \in \xi(\Omega)} \left(\varphi(a) \sum_{\omega: \xi(\omega)=a} P(\omega) \right) = \sum_{a \in \xi(\Omega)} \varphi(a) P(\xi = a) \end{aligned}$$

6.

$$\begin{aligned} E[\xi] &= \sum_{\omega \in \Omega} \xi(\omega) P(\omega) = \sum_{\omega: \xi(\omega)=c} \xi(\omega) P(\omega) + \sum_{\omega: \xi(\omega) \neq c} \xi(\omega) P(\omega) = c P(\xi = c) + \\ &+ \sum_{\omega: \xi(\omega) \neq c} \xi(\omega) \cdot 0 = c. \end{aligned}$$

$$7. \xi \geq 0 \implies E[\xi] \geq E[0] \geq 0.$$

$$8. E[\xi] = \sum_{\omega \in \Omega} \underbrace{\xi(\omega)}_{\geq 0} \underbrace{P(\omega)}_{\geq 0} = 0 \implies \forall \omega \in \Omega \xi(\omega) P(\omega) = 0. \text{ Тогда если } \xi(\omega) \neq 0, \text{ то } P(\omega) = 0.$$

$$\text{Тогда } P(\xi = 0) = 1.$$

Заметим, что если величина может быть и отрицательной, то это свойство не выполняется. Предположим, что ξ равновероятно принимает значения 1 и -1 . Тогда $E[\xi] = 0$, но $P(\xi = 0) = 0$.

9.

$$\begin{aligned} E[\xi\eta] &= \sum_{\omega \in \Omega} \xi(\omega)\eta(\omega) P(\omega) = \left\{ \begin{array}{l} \xi(\Omega) = (a_1, a_2, \dots) \\ \eta(\Omega) = (b_1, b_2, \dots) \end{array} \right\} = \sum_{i,j} \sum_{\substack{\xi(\omega)=a_i \\ \eta(\omega)=b_j}} \xi(\omega)\eta(\omega) P(\omega) = \\ &= \sum_{i,j} a_i b_j P(\xi = a_i, \eta = b_j) \end{aligned}$$

Так как $\xi \perp \eta$, то $P(\xi = a_i, \eta = b_j) = P(\xi = a_i) P(\eta = b_j)$ и сумма равна

$$\sum_{i,j} a_i b_j P(\xi = a_i) P(\eta = b_j) = \left(\sum_i a_i P(\xi = a_i) \right) \left(\sum_j b_j P(\eta = b_j) \right) = E[\xi] E[\eta]$$

□

Примеры использования математического ожидания:

1. Посчитаем математическое ожидание индикатора события A :

$$E[I_A] = \sum_{\omega \in \Omega} I_A(\omega) P(\omega) = \sum_{\omega \in A} P(\omega) = P(A).$$

2. Если мы рассмотрим классическую модель, т.е. такую модель, где все исходы равновероятны, то $E[\xi] = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \xi(\omega)$. Иначе говоря, математическое ожидание в классической модели равно среднему арифметическому возможных значений ξ .

3. Пусть $\xi \sim \text{Bin}(n, p)$. Тогда

$$\begin{aligned} E[\xi] &= \sum_{k=0}^n k P(\xi = k) = \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} = \\ &= \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k} = np \sum_{t=0}^{n-1} \binom{n-1}{t} p^t (1-p)^{n-1-t}. \end{aligned}$$

1.3.4 Дисперсия. Ковариация. Их свойства

Определение 22. Пусть ξ — некоторая случайная величина над (Ω, P) . Тогда *дисперсией* ξ называется $D[\xi] = E[(\xi - E[\xi])^2]$.

Дисперсию случайной величины можно понимать как среднеквадратическое отклонение этой случайной величины от её среднего значения (математического ожидания).

Примечание. Вероятность незначительного отклонения величины от математического ожидания, то есть $P(E[\xi] - 2\sqrt{D[\xi]} \leq \xi \leq E[\xi] + 2\sqrt{D[\xi]})$, всегда велика! Это нам пока не пригодится, но на будущее стоит запомнить.

Определение 23. Пусть ξ и η — две случайные величины. Тогда *ковариацией* этих величин называется $\text{cov}(\xi, \eta) = E[(\xi - E[\xi])(\eta - E[\eta])]$.

Ковариацию стоит воспринимать как меру зависимости двух случайных величин.

Определение 24. Две случайные величины ξ и η называют *некоррелированными*, если $\text{cov}(\xi, \eta) = 0$.

Теорема 4 (Простейшие свойства дисперсии и ковариации). Пусть ξ, η и χ — некоторые случайные величины. Тогда выполняются следующие свойства:

1. Ковариация билинейна: $\text{cov}(\xi, a\eta + b\chi) = a \text{cov}(\xi, \eta) + b \text{cov}(\xi, \chi)$;
2. $D[\xi] = \text{cov}(\xi, \xi)$;
3. Для любого $c \in \mathbb{R}$ верно, что $D[c\xi] = c^2 D[\xi]$ и $D[\xi + c] = D[\xi]$;
4. Дисперсия неотрицательна: $D[\xi] \geq 0$. При этом $D[\xi] = 0 \iff P(\xi = E[\xi]) = 1$;
5. Если ξ и η некоррелированные, то $D[\xi + \eta] = D[\xi] + D[\eta]$;
6. Связь с матожиданием: $D[\xi] = E[\xi^2] - (E[\xi])^2$, $\text{cov}(\xi, \eta) = E[\xi\eta] - E[\xi]E[\eta]$

Доказательство. Докажем все пункты по порядку:

1. По определению ковариации $\text{cov}(\xi, a\eta + b\chi) = E[(\xi - E[\xi])(a\eta + b\chi - E[a\eta + b\chi])]$. Распишем вторую скобку:

$$(a\eta + b\chi) - E[a\eta + b\chi] = (a\eta + b\chi) - aE[\eta] - bE[\chi] = a(\eta - E[\eta]) + b(\chi - E[\chi]).$$

Тогда

$$\begin{aligned} \text{cov}(\xi, a\eta + b\chi) &= E[(\xi - E[\xi])(a(\eta - E[\eta]) + b(\chi - E[\chi]))] \\ &= E[a(\xi - E[\xi])(\eta - E[\eta]) + b(\xi - E[\xi])(\chi - E[\chi])] \\ &= aE[(\xi - E[\xi])(\eta - E[\eta])] + bE[(\xi - E[\xi])(\chi - E[\chi])] \\ &= a \text{cov}(\xi, \eta) + b \text{cov}(\xi, \chi). \end{aligned}$$

2. $\text{cov}(\xi, \xi) = E[(\xi - E[\xi])(\xi - E[\xi])] = E[(\xi - E[\xi])^2] = D[\xi]$.
3. Для произведения $D[c\xi] = \text{cov}(c\xi, c\xi) = c^2 \text{cov}(\xi, \xi) = c^2 D[\xi]$. Для суммы $D[\xi + c] = E[(\xi + c - E[\xi + c])^2] = E[(\xi + c - E[\xi] - c)^2] = E[(\xi - E[\xi])^2] = D[\xi]$.
4. По определению дисперсии, $D[\xi] = E[(\xi - E[\xi])^2]$, а $(\xi - E[\xi])^2 \geq 0$. Тогда $E[(\xi - E[\xi])^2] \geq 0$.
5. Воспользуемся выражением дисперсии через ковариацию и билинейностью ковариации: $D[\xi + \eta] = \text{cov}(\xi + \eta, \xi + \eta) = \text{cov}(\xi, \xi) + \text{cov}(\eta, \eta) + 2 \text{cov}(\xi, \eta)$. Так как ξ и η некоррелированные, то $\text{cov}(\xi, \eta) = 0$ и $D[\xi + \eta] = \text{cov}(\xi, \xi) + \text{cov}(\eta, \eta) = D[\xi] + D[\eta]$.

6. $\text{cov}(\xi, \eta) = E[(\xi - E[\xi])(\eta - E[\eta])] = E[\xi\eta - \xi E[\eta] - \eta E[\xi] + E[\xi]E[\eta]]$. По линейности матожидания $\text{cov}(\xi, \eta) = E[\xi\eta] - 2E[\xi]E[\eta] + E[\xi]E[\eta] = E[\xi\eta] - E[\xi]E[\eta]$. Для дисперсии воспользуемся тем, что $D[\xi] = \text{cov}(\xi, \xi)$.

□

Теорема 5. Если случайные величины независимы, то они некоррелированные. Обратное, вообще говоря, неверно.

Доказательство. Докажем, что из независимости следует некоррелируемость. Пусть ξ и η — независимые случайные величины. Тогда $\text{cov}(\xi, \eta) = E[\xi\eta] - E[\xi]E[\eta]$. Но по свойству матожидания из независимости случайных величин следует, что $E[\xi\eta] = E[\xi]E[\eta]$. Тогда $\text{cov}(\xi, \eta) = 0$ и эти величины некоррелированные.

Теперь покажем, что из некоррелированности не обязательно следует независимость. Пусть случайная величина ξ равновероятно принимает значения из множества $\{0, 1, -1\}$. Возьмем случайную величину $\eta = \xi^2$. По определению можно проверить, что величины η и ξ некоррелированные: $\text{cov}(\xi, \eta) = E[\xi\eta] - E[\xi]E[\eta] = E[\xi^3] - E[\xi]E[\xi^2] = 0 - 0 = 0$. Но ξ и η не являются независимыми, что проверяется опять же по определению: $P(\xi = 0, \eta = 0) = \frac{1}{3} \neq \frac{1}{9} = P(\xi = 0)P(\eta = 0)$. □

1.3.5 Неравенства Маркова и Чебышёва

Под конец лекции обсудим два неравенства, которые сами по себе являются весьма полезными.

Теорема 6 (Неравенство Маркова). Пусть $\xi \geq 0$ — неотрицательная случайная величина на (Ω, P) . Тогда для любого положительного a

$$P(\xi \geq a) \leq \frac{E[\xi]}{a}.$$

Доказательство. Как было доказано ранее, матожидание индикатора события равно вероятности события. Тогда $P(\xi \geq a) = E[I_{\xi \geq a}]$. Далее, заметим, что $I_{\xi \geq a} \leq \frac{\xi}{a} I_{\xi \geq a}$. Тогда

$$P(\xi \geq a) \leq E\left[\frac{\xi}{a} I_{\xi \geq a}\right] \leq E\left[\frac{\xi}{a}\right] = \frac{E[\xi]}{a}.$$

□

Теорема 7 (Неравенство Чебышёва). Пусть ξ — случайная величина на (Ω, P) такая, что $D[\xi] < \infty$. Тогда для любого положительного a

$$P(|\xi - E[\xi]| \geq a) \leq \frac{D[\xi]}{a^2}.$$

Доказательство. $P(|\xi - E[\xi]| \geq a) = P((\xi - E[\xi])^2 \geq a^2)$. По неравенству Маркова

$$P((\xi - E[\xi])^2 \geq a^2) \leq \frac{E[(\xi - E[\xi])^2]}{a^2} = \frac{D[\xi]}{a^2}.$$

□

1.4 Лекция от 30.09.2016

1.4.1 Сходимость случайных величин. Закон больших чисел.

Определение 25. Пусть $\{\xi_n, n \in \mathbb{N}\}$ — последовательность случайных величин. Тогда будем говорить, что ξ_n *сходится по вероятности* к некоторой случайной величине ξ , если для любого положительного α при $n \rightarrow \infty$ выполняется $P(|\xi_n - \xi| \geq \alpha) \rightarrow 0$.

Обозначение: $\xi_n \xrightarrow{P} \xi$.

Теорема 8 (Закон больших чисел). Пусть имеется последовательность попарно некоррелированных случайных величин $\{\xi_n, n \in \mathbb{N}\}$ и $\exists c$ такое, что $D[\xi_n] \leq c$ для любых $n \in \mathbb{N}$. Обозначим через $S_n = \xi_1 + \dots + \xi_n$ сумму первых n случайных величин. Тогда $\frac{S_n - E[S_n]}{n} \xrightarrow{P} 0$.

Доказательство. В силу того, что $\text{cov}(\xi_i, \xi_j) = 0 \forall i \neq j$, то $D[S_n] = D[\xi_1 + \dots + \xi_n] = \text{cov}(\xi_1 + \dots + \xi_n, \xi_1 + \dots + \xi_n) = \sum_{i=1}^n \text{cov}(\xi_i, \xi_i) = \sum_{i=1}^n D[\xi_i] \leq n \cdot c$.

Далее по свойству дисперсии $D[\frac{S_n - E[S_n]}{n}] = \frac{1}{n^2} D[S_n]$. А исходя из рассуждения выше, данную величину можно оценить сверху: $\frac{1}{n^2} D[S_n] \leq \frac{c}{n}$.

Наконец, для любых положительных α по неравенству Чебышёва $P\left(\left|\frac{S_n - E[S_n]}{n}\right| \geq \alpha\right) \leq \frac{D[\frac{S_n - E[S_n]}{n}]}{\alpha^2}$. Осталось только ограничить числитель уже имеющейся оценкой. $\frac{D[\frac{S_n - E[S_n]}{n}]}{\alpha^2} \leq \frac{c}{n\alpha^2}$. Несложно увидеть, что данная величина стремится к нулю при $n \rightarrow \infty$. \square

Следствие: Пусть $\{\xi_n, n \in \mathbb{N}\}$ — последовательность попарно независимых случайных величин, $E[\xi_n] = \alpha$ и $\exists c > 0$ такое, что $D[\xi_n] \leq c$. Тогда $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{P} \alpha$.

Доказательство. Из условия независимости случайных величин нам хватит следствия о том, что они являются некоррелированными. Из Закона больших чисел следует, что $\frac{S_n - E[S_n]}{n} \xrightarrow{P} 0$. $\frac{S_n}{n} - \alpha \xrightarrow{P} 0$ Также заметим, что в силу определения сходимости, $\frac{S_n}{n} \xrightarrow{P} \alpha$. \square

Давайте попробуем осознать, в чем заключается смысл Закона больших чисел. Пусть случайная величина задается индикатором происхождения события, $\xi_i = I\{\text{событие } A \text{ произошло в } i\text{-ом эксперименте}\}$, при этом все эксперименты проводятся независимо друг от друга (ξ_n является независимой случайной величиной). Обозначим за $\vartheta(A)$ частоту появления события A в n -м количестве экспериментов, $\vartheta_n(A) = \frac{\xi_1 + \dots + \xi_n}{n}$. Тогда по Закону больших чисел $\vartheta_n(A) \xrightarrow{P} E[\xi_i] = E[I\{\text{событие } A \text{ произошло в } i\text{-ом эксперименте}\}] = P(A)$. Выходит, что Закон больших чисел является теоретическим обоснованием принципа устойчивости частот.

Определение 26. Последовательность случайных $\{\xi_n, n \in \mathbb{N}\}$ сходится к случайной величине ξ с вероятностью 1 (или *почти наверное*⁸), если $P(\lim_{n \rightarrow \infty} \xi_n \rightarrow \xi) = P(\omega : \lim_{n \rightarrow \infty} \xi_n(\omega) = \xi(\omega)) = 1$. Обозначение: $\xi_n \xrightarrow{\text{п.н.}} \xi$ (сходится почти наверное).

Примечание. Если $\forall \omega \in \Omega, P(\omega) > 0$, то $\xi_n \xrightarrow{\text{п.н.}} \xi \Leftrightarrow \forall \omega \in \Omega, \xi_n(\omega) \rightarrow \xi(\omega)$.

Теорема 9. В дискретных вероятностных пространствах $\xi_n \xrightarrow{P} \xi \iff \xi_n \xrightarrow{\text{п.н.}} \xi$.

⁸Название “сходимость почти наверное” произошло от применяющегося в англоязычной литературе термина “almost sure convergence”.

Доказательство. Докажем по очереди в обе стороны:

- $\xi_n \xrightarrow{P} \xi \implies \xi_n \xrightarrow{\text{п.н.}} \xi$

Из сходимости *по вероятности* следует сходимость *почти наверное*. Для доказательства этого факта нужно показать, что $\forall \omega_0 \in \Omega : P(\omega_0) > 0$ выполняется сходимость $\xi_n(\omega_0) \rightarrow \xi(\omega_0)$. Пусть такой сходимости нет, $\xi_n(\omega_0) \not\rightarrow \xi(\omega_0)$. В этом случае из отрицания определения о сходимости *почти наверное* существуют такое положительное α_0 и подпоследовательность $\{\xi_{n_k}(\omega_0), k \in \mathbb{N}\}$, что $\forall k \in \mathbb{N}, P(|\xi_{n_k} - \xi| \geq \alpha_0) \geq \alpha_0$. Но тогда для любых натуральных k , $P(|\xi_{n_k} - \xi| \geq \alpha_0) \geq P(\omega_0) > 0$. Получаем противоречие с определением сходимости *по вероятности*. Значит, $\xi_n \xrightarrow{P} \xi \implies \xi_n \xrightarrow{\text{п.н.}} \xi$.

- $\xi_n \xrightarrow{\text{п.н.}} \xi \implies \xi_n \xrightarrow{P} \xi$

Из сходимости *почти наверное* следует, что $P(\lim_{n \rightarrow \infty} \xi_n = \xi) = 1$. Следовательно, выполнено и $\xi_n(\omega) \rightarrow \xi(\omega)$ для $\omega \in \Omega : P(\omega) > 0$. Теперь нужно проверить, выполняется ли условие сходимости *по вероятности*: $P(|\xi_n - \xi| \geq \alpha) \rightarrow 0$. Пусть $\Omega = \{\omega_i, i \in \mathbb{N}\}$. Теперь для любого положительного δ выберем номер, от него зависящий, такой что $\sum_{i > N} P(\omega_i) \leq \delta$. Тогда

$$\begin{aligned} P(|\xi_n - \xi| \geq \alpha) &= \sum_{\substack{i: \\ |\xi_n(\omega_i) - \xi(\omega_i)| \geq \alpha}} P(\omega_i) = \sum_{\substack{i: \\ i \leq N, \\ |\xi_n(\omega_i) - \xi(\omega_i)| \geq \alpha}} P(\omega_i) + \\ &+ \sum_{\substack{i: \\ i > N, \\ |\xi_n(\omega_i) - \xi(\omega_i)| \geq \alpha}} P(\omega_i) \leq \delta + \sum_{\substack{i: \\ i \leq N, \\ |\xi_n(\omega_i) - \xi(\omega_i)| \geq \alpha, \\ Pr(\omega_i) > 0}} P(\omega_i). \end{aligned}$$

Теперь нужно перейти к пределу при $n \rightarrow \infty$. Но чтобы избежать ситуации, когда предела не существует, перейдем к верхнему пределу для тех же n . Также заметим, что так как $\xi_n(\omega_i) \rightarrow \xi(\omega_i)$, начиная с некоторого n , ω_i перестает входить в сумму:

$$\lim_{n \rightarrow \infty} P(|\xi_n - \xi| \geq \alpha) \leq \delta + 0.$$

А в силу произвольности $\delta > 0$ можно говорить о том, что этот предел стремится к нулю.

□

Вопрос 1: Как оценить скорость сходимости $\frac{S_n}{n} - a$ к нулю?

Вопрос 2: Как оценить скорость сходимости $P(|\frac{S_n}{n} - a| \geq \varepsilon)$ к нулю?

1.4.2 Схема испытаний Бернулли

Сначала введём вспомогательное определение:

Определение 27. Случайные величины ξ и η называют *одинаково распределёнными*, если

$$\begin{cases} \xi(\Omega) = \eta(\Omega); \\ \forall x \in \xi(\Omega), P[\xi = x] = P(\eta = x). \end{cases}$$

Говоря простым языком, у этих случайных величин совпадают возможные значения и их вероятности.

Обозначение: $\xi \stackrel{d}{=} \eta$.

Описание: Схема испытаний Бернули подразумевает много однородных экспериментов, в каждом из которых возможно два исхода — “успех” и “неудача”; при этом нас интересует распределение числа успехов.

Математическая модель: Пусть $\{\xi_i\}$ — независимые *одинаково распределённые* случайные величины (например, последовательные броски монеты); $\xi_i(\Omega) = \{0; 1\}$; пусть также $P(\xi_i = 1) = p$ — вероятность успеха каждого отдельного эксперимента. Введём также обозначение $S_n = \sum_{i=1}^n \xi_i$ — случайную величину, обозначающую количество успехов.

Распределение S_n кажется очевидным: $S_n \sim \text{Bin}(n, p)$, т.е. $P(S_n = k) = C_n^k p^k (1-p)^{n-k}$.

Упражнение. Докажите, что распределение S_n и вправду биномиально.

Упражнение. В качестве упражнения, читателю предлагается оценить, сколько нужно попыток, чтобы $\frac{\text{число успехов}}{\text{число попыток}}$ было примерно равно p , т.е. как быстро наблюдаемая вероятность сходится к реальной. Для этого предлагается оценить следующую вероятность:

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) = ?$$

Кстати говоря, p совершенно не обязана быть константой и может являться функцией от n : $p(n)$.

Тогда, вообще говоря, есть несколько отдельных случаев:

- $np(n) \xrightarrow{n \rightarrow \infty} \lambda > 0$.

Теорема 10 (Пуассон). Пусть $np(n) \xrightarrow{n \rightarrow \infty} \lambda > 0$. Тогда

$$\forall k \in \mathbb{Z}_+, P(S_n = k) \xrightarrow{n \rightarrow \infty} e^{-\lambda} \frac{\lambda^k}{k!}.$$

Доказательство.

$$\begin{aligned} C_n^k p^k (1-p)^{n-k} &= \frac{n(n-1)(n-2) \dots (n-k+1)}{k!} p(n)^k (1-p(n))^{n-k} \sim \\ &\sim \frac{n^k}{k!} \left(\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^k \left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^{n-k} = \\ &= \frac{1}{k!} (\lambda + o(1))^k \left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^{n-k} \xrightarrow{n \rightarrow \infty} \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

□

Следствие: $P(S_n = k) \xrightarrow{d} P(\xi = k)$, $\xi \sim \text{Pois}(\lambda)$ — сходимость по распределению.

- $np(n) \xrightarrow{n \rightarrow \infty} \infty$.

Сделаем ещё одно отступление в сторону и рассмотрим величину $\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{D}[S_n]}}$. Легко понять, что

$$\mathbb{E}[S_n] = np.$$

Несложно также показать, что

$$\mathbb{D}[S_n] = \sum \mathbb{D}[\xi_i] + \sum \text{cov}[x_i, x_j] = \sum \mathbb{D}[\xi_i] = n \mathbb{D}[\xi_1] = n (\mathbb{E}[\xi_1^2] - (\mathbb{E}[\xi_1])^2) = npq.$$

Теорема 11 (Муавр-Лаплас). Пусть p — константа. Введём обозначение

$$P_n(a, b) := P \left(a < \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathbb{D}[S_n]}} < b \right).$$

Тогда

$$\sup_{a < b} \left| P_n(a, b) - \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \right| \xrightarrow{n \rightarrow \infty} 0$$

Доказательство. Без доказательства. □

Глава 2

Семинары

2.1 Семинар от 09.09.2016

Перед тем, как начать решать задачи, кратко опишем вероятностное пространство для броска n -гранного кубика: $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, $\omega_i = \{\text{выпало число } i\}$, $P(\omega_i) = \frac{1}{n}$ для всех i .

Задача 1. Пусть бросаются n -гранный и m -гранный кубики. Какова вероятность P того, что выпадет одно чётное и одно нечётное число?

Решение. В данной задаче есть два случая:

1. На первом выпало чётное число очков, а на втором — нечётное. Количество чётных чисел от 1 до n равно $\lfloor \frac{n}{2} \rfloor$, а нечётных чисел от 1 до m — $\lceil \frac{m}{2} \rceil$. Тогда есть $\lceil \frac{m}{2} \rceil \lfloor \frac{n}{2} \rfloor$ успешных исходов.
2. На первом выпало нечётное число очков, на втором — чётное. Аналогичными рассуждениями получаем $\lfloor \frac{m}{2} \rfloor \lceil \frac{n}{2} \rceil$ успешных исходов.

Всего же исходов mn . Следовательно,

$$P = \frac{\lceil \frac{m}{2} \rceil \lfloor \frac{n}{2} \rfloor + \lfloor \frac{m}{2} \rfloor \lceil \frac{n}{2} \rceil}{mn}.$$

□

Задача 2. Пусть бросаются два n -гранных кубика. Какова вероятность $P(i)$ того, что суммарно выпадет $2 \leq i \leq 2n$ очков?

Решение. В данной задаче есть два случая:

1. $i \leq n + 1$. Представим i в следующем виде: $i = k + (i - k)$, где $1 \leq k \leq i - 1$. Такое ограничение сверху на k объясняется тем, что иначе $i - k$ будет меньше 1, а при броске кубика не может выпасть меньше 1 очка. Ограничение снизу объясняется аналогично. Тогда есть $i - 1$ подходящий случай.
2. $n + 2 \leq i \leq 2n$. Опять же, представим i в виде $i = k + (i - k)$. Теперь определим границы для k . Очевидно, что $k \leq n$. Так как $i - k \leq n$, то $k \geq i - n$. Тогда получаем $i - n \leq k \leq n$. Тогда есть $n - (i - n) + 1 = 2n - i + 1$ подходящий случай.

Так как всего есть n^2 разных вариантов того, сколько очков выпадет на кубиках, то получаем, что

$$P(i) = \begin{cases} \frac{i-1}{n^2} & i \leq n+1, \\ \frac{2n-i+1}{n^2} & n+2 \leq i \leq 2n. \end{cases} \quad \square$$

Примечание. Если нарисовать график функции $P(i)$, то он будет выглядеть, как треугольник с вершиной в точке $(n+1, \frac{1}{n})$. Такой график называют *треугольным распределением*.

Перейдём от кубиков к монеткам.

Задача 3. Пусть последовательно бросают n монет (полагается, что $\Omega = \{O, P\}$). Какова вероятность P того, что не выпадет последовательно

1. орёл и решка?
2. два орла?

Решение. Как рассказывалось ранее, в такой модели есть 2^n элементарных исходов. Посчитаем количество успешных исходов в каждом случае:

1. В таком случае легко понять, что будут допустимы только последовательности вида $\underbrace{PP \dots P}_{k \text{ раз}} \underbrace{OO \dots O}_{n-k \text{ раз}}$, где $0 \leq k \leq n$. Тогда есть $n+1$ подходящий исход.
2. Пусть f_n — количество последовательностей длины n , в которых нет двух орлов подряд. Как посчитать f_n ? Попробуем выразить рекурсивно. Если при последнем броске выпал орёл, то при предпоследнем обязательно выпала решка. То, что идёт до решки, явно угадать невозможно. Но нам известно, что это последовательность размера $n-2$ и в ней нет двух орлов подряд. Тогда их f_{n-2} вариантов. Если же выпала решка, то есть f_{n-1} вариант. Отсюда получаем, что $f_n = f_{n-1} + f_{n-2}$. Так как $f_1 = 2$, а $f_2 = 3$ (допускаются ОР, РО, РР), то $f_n = F_{n+2}$, где F_n — n -е число Фибоначчи.

Ответ: 1) $P = \frac{n+1}{2^n}$, 2) $P = \frac{F_{n+2}}{2^n}$. □

Перед тем, как идти дальше, сделаем небольшое отступление. Во втором пункте последней задачи нам повезло, что последовательность совпала с последовательностью чисел Фибоначчи. А что делать, если не удаётся угадать последовательность? В таком случае можно воспользоваться общим методом решения. Рассмотрим его на примере из последней задачи:

$$\begin{aligned} f_n &= f_{n-1} + f_{n-2} \\ f_1 &= 2 \\ f_2 &= 3 \end{aligned}$$

Выпишем *характеристическое уравнение* для данного рекуррентного соотношения: $\lambda^2 - \lambda - 1 = 0$. Находим его корни — в данном случае они равны $\frac{1 \pm \sqrt{5}}{2}$. Тогда для каких-то констант a_1 и a_2 верно, что

$$f_n = a_1 \left(\frac{1 + \sqrt{5}}{2} \right)^n + a_2 \left(\frac{1 - \sqrt{5}}{2} \right)^n.$$

Данные константы можно определить по начальным условиям.

Задача 4 (Парадокс дней рождения). В группе 27 студентов. Считаем их дни рождения случайными и равновероятными. Какова вероятность P того, что хотя бы у двух студентов совпадают дни рождения?

Решение. В данной задаче гораздо проще посчитать вероятность дополнения, то есть вероятность того, что у всех 27 студентов будут разные дни рождения. Так как порядок дней рождения важен, то эта вероятность равна $\frac{A_{365}^{27}}{365^{27}}$. В итоге получаем, что

$$P = 1 - \frac{A_{365}^{27}}{365^{27}}.$$

Хорошо, ответ получен. Но по нему сложно сказать, много ли это или мало. Попробуем посчитать его приближенно: $P = 1 - \left(1 - \frac{1}{365}\right) \left(1 - \frac{2}{365}\right) \dots \left(1 - \frac{26}{365}\right)$. Так как $1 + x \approx e^x$, то $P \approx 1 - e^{-\frac{1+2+\dots+26}{365}} \approx 1 - e^{-1.04} \approx 0,66$. Как видно, вероятность достаточно велика. \square

Сделаем небольшое теоретическое отступление. Вспомним формулу включений-исключений: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Попробуем придумать аналогичную формулу для трёх событий:

$$\begin{aligned} P(A \cup B \cup C) &= P(A \cup B) + P(C) - P((A \cup B) \cap C) \\ &= P(A \cup B) + P(C) - P((A \cap C) \cup (B \cap C)) \\ &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \end{aligned}$$

Уже видна некоторая закономерность. Сформулируем обобщение.

Теорема 12 (Общая формула включений-исключений). Пусть A_1, A_2, \dots, A_n — некоторые события на Ω , а $S_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})$. Тогда

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{k=1}^n (-1)^{k-1} S_k.$$

Доказательство. По индукции. База ($n = 2$) была доказана ранее. Теперь предположим, что утверждение верно для какого-то n . Докажем, что из этого следует, что утверждение верно и для $n + 1$:

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n \cup A_{n+1}) &= P(A_1 \cup A_2 \cup \dots \cup A_n) + \\ &\quad + P(A_{n+1}) - P((A_1 \cup A_2 \cup \dots \cup A_n) \cap A_{n+1}) \end{aligned}$$

Так как $(A_1 \cup A_2 \cup \dots \cup A_n) \cap A_{n+1} = \bigcup_{i=1}^n (A_i \cap A_{n+1})$, то, пользуясь предположением индукции и рассуждениями, аналогичными доказательству для трёх множеств, получаем желаемое. \square

Примечание. Важное следствие из этой формулы: если A_1, A_2, \dots, A_n — некоторые события на Ω , то по закону де Моргана получаем, что $P\left(\bigcap_{i=1}^n \overline{A_i}\right) = P\left(\overline{\bigcup_{i=1}^n A_i}\right) = 1 - P\left(\bigcup_{i=1}^n A_i\right) = 1 - \sum_{k=1}^n (-1)^{k-1} S_k$. Если положить $S_0 = 1$, то эту формулу можно записать в виде

$$P\left(\bigcap_{i=1}^n \overline{A_i}\right) = \sum_{k=0}^n (-1)^k S_k.$$

Задача 5. Пусть мы раскидали n шаров по m ящикам. Какова вероятность P того, что ни один ящик не пуст? Рассмотрите случаи, когда шары различимы и неразличимы.

Решение. Начнём со случая различимых шаров. В данном случае элементарным исходом будет $\omega = (\omega_1, \omega_2, \dots, \omega_m)$, где ω_i — количество шаров в i -м ящике. В таком случае $|\Omega| = \binom{n+m-1}{m-1}$ (схема выбора неупорядоченных наборов с возвратом). Теперь посчитаем количество подходящих исходов. Так как ни один ящик не пуст, то в каждом из них есть хотя бы по одному шару. Тогда нужно посчитать количество способов раскидать $n-m$ шаров по m ящикам. Это можно сделать $\binom{n-1}{m-1}$ способом. Отсюда получаем, что вероятность равна

$$P = \frac{\binom{n-1}{m-1}}{\binom{n+m-1}{m-1}}.$$

Теперь предположим, что шары неразличимы. Рассмотрим событие $A_i = \{i\text{-й ящик пуст}\}$. Чему равна вероятность такого события? Для каждого из n шаров есть $m-1$ подходящий ящик. Тогда $P(A_i) = \frac{(m-1)^n}{m^n} = \left(1 - \frac{1}{m}\right)^n$. Пересечение $A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}$ означает, что k ящиков с номерами i_1, i_2, \dots, i_k пусты. Тогда $P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \left(1 - \frac{k}{m}\right)^n$. Заметим, что событие “ни один ящик не пуст” равно $\overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_m}$. Пользуясь формулой включений-исключений, получаем, что вероятность равна

$$P = \sum_{k=0}^m (-1)^k \binom{m}{k} \left(1 - \frac{k}{m}\right)^n.$$

□

Задача 6. Алиса и Боб случайно подбрасывают n монет. Какова вероятность P того, что число орлов у Алисы будет строго больше, чем у Боба? Каков будет ответ на этот вопрос, если Алиса подбросила $n+1$ монету?

Решение. Для начала посмотрим, чему равна вероятность того, что число орлов у Алисы равно числу орлов у Боба. Если у Алисы выпало k орлов, что достигается в $\binom{n}{k}$ случаев, то у Боба тоже должно выпасть k орлов. Тогда достаточно логично, что число успешных исходов равно $\sum_{k=0}^n \binom{n}{k}^2$. Как это упростить? Воспользуемся тем, что $\binom{n}{k} = \binom{n}{n-k}$. Теперь представим себе следующую ситуацию: пусть есть строка, содержащая $2n$ символов. Из первых n нужно выбрать k символов, из вторых n нужно выбрать $(n-k)$. Это можно сделать $\binom{n}{k} \binom{n}{n-k}$ способами. Если просуммировать эти числа по k от 0 до n , то легко заметить, что это то же самое, что и посчитать количество способов выбрать n символов из $2n$. Тогда

$$\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}.$$

Всего исходов 4^n (по 2^n на Алису и на Боба). Тогда вероятность равна $\frac{\binom{2n}{n}}{4^n}$.

Теперь рассмотрим вероятность из условия. Из-за симметричности она равна вероятности того, что у Алисы будет строго меньше орлов, чем у Боба. Тогда получаем, что $2P + \frac{\binom{2n}{n}}{4^n} = 1$. Отсюда $P = \frac{1}{2} - \frac{\binom{2n}{n}}{2^{2n+1}}$.

Теперь перейдём ко второму пункту. Его мы решим двумя способами — стандартным и “олимпиадным”. Начнём со стандартного. Если у Алисы уже было больше орлов, чем у

Боба, то что бы у неё не выпало, то ситуация не изменится. Если же было так, что у неё столько же орлов, сколько у Боба, то ей необходимо, чтобы выпал орёл. Тогда искомая вероятность равна

$$P' = P + \frac{1}{2} \frac{\binom{2n}{n}}{4^n} = \frac{1}{2}.$$

Теперь рассмотрим “олимпиадный” способ решения. Заметим, что вероятность того, что у Алисы будет больше орлов, чем у Боба, равна вероятности того, что у неё будет больше решек (из-за симметрии). При этом вероятность того, что у неё будет больше орлов, равна вероятности того, что у неё будет не больше решек, чем у Боба (пусть у неё на одного орла больше, тогда число решек у них совпадает). Отсюда сразу получаем, что $P' = \frac{1}{2}$. \square

А сейчас мы посмотрим, почему стоит быть осторожным с азартными играми.

Задача 7. Пусть есть 52 карты, и игроку выдают 5 случайных карт. Найдите вероятности получения различных наборов из покера.

Решение. Начнём с того, что заметим, что выбрать 5 карт из 52 мы можем $\binom{52}{5}$ способами. Теперь достаточно найти количество подходящих исходов.

1. Royal Flush — туз, король, дама, валет и десятка одной масти. Есть лишь 4 подходящих исхода.
2. Straight Flush — пять последовательных по достоинству карт одной масти (начиная не с туза). Так как первую карту можно выбрать 8 способами (от пятёрки до короля), то есть $9 \cdot 4 = 36$ успешных исходов.
3. Four Of A Kind — четыре карты одного достоинства. Выберем достоинство (это можно сделать 13) способами и последнюю карту (это можно сделать 48 способами). Тогда есть $13 \cdot 48$ подходящих комбинаций.
4. Full House — три карты одного достоинства и две карты другого достоинства. Выберем первое достоинство (13 вариантов) и выберем 3 карты из 4 подходящих ($\binom{4}{3}$ способов). Теперь выберем второе достоинство (12 вариантов) и 2 карты из 4 ($\binom{4}{2}$ вариантов). Тогда всего есть $13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}$ вариантов.
5. Flush — пять карт одной масти. Всего выбрать пять карт одной масти можно $\binom{13}{5}$ способами. Но в таком случае мы ещё учитываем Straight Flush и Royal Flush. Тогда есть $4(\binom{13}{5} - 10)$ вариантов.
6. Straight — пять последовательных карт (не одной масти). Всего пять последовательных карт можно выбрать $10 \cdot 4^5$ способами (сначала выбираем старшую карту, после чего масть для каждой). Но в таком случае учитывается Straight Flush и Royal Flush. Тогда есть $10(4^5 - 4)$ подходящих наборов.
7. Three Of A Kind — три карты одного достоинства. Сначала выберем достоинство (13 вариантов), после чего выберем 3 карты из 4 ($\binom{4}{3}$ вариантов). После чего выберем два разных достоинства (иначе будет Full House), что можно сделать $\binom{12}{2}$ способами, и масти для двух карт (4^2 способа). Итого — $13 \cdot \binom{4}{3} \cdot \binom{12}{2} \cdot 4^2$ варианта.

8. Two Pair — две пары карт одного достоинства. Выберем достоинства и масти для двух пар ($\binom{13}{2}$ варианта для достоинств, по $\binom{4}{2}$ для выбора 2-х карт каждого достоинства). Осталось выбрать последнюю карту — это можно сделать $52 - 8 = 44$ способами. Итого $\binom{13}{2} \binom{4}{2}^2 \cdot 44$ исхода.
9. One Pair — одна пара карт одного достоинства. Выберем достоинство и 2 карты из 4 ($\binom{13}{1} \cdot \binom{4}{2}$ вариантов). Теперь выберем три разных достоинства и масти для карт ($\binom{12}{3} \cdot 4^3$ варианта). Итого $\binom{13}{1} \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot 4^3$ исходов.
10. High Card — ничего из вышеперечисленного. Выберем пять разных достоинств, не идущих подряд ($\binom{13}{5} - 10$ вариантов) и выберем масти для каждой карты так, чтобы они не совпадали ($4^5 - 4$ варианта). Итого $(\binom{13}{5} - 10)(4^5 - 4)$ вариантов.

Теперь приближенно посчитаем вероятность каждого из наборов:

Тип	Вероятность
Royal Flush	$4/2598960 \approx 0,00015\%$
Straight Flush	$36/2598960 \approx 0,0014\%$
Four Of A Kind	$624/2598960 \approx 0,024\%$
Full House	$3744/2598960 \approx 0,15\%$
Flush	$5108/2598960 \approx 0,2\%$
Straight	$10200/2598960 \approx 0,39\%$
Three Of A Kind	$54912/2598960 \approx 2,11\%$
Two Pair	$123552/2598960 \approx 4,75\%$
One Pair	$1098240/2598960 \approx 42,26\%$
High Card	$1302540/2598960 \approx 50,12\%$

Как видно из таблицы, получить что-то лучше, чем одну пару, уже не так просто. \square

2.2 Семинар от 16.09.2016

Начнём с разбора домашнего задания.

Задача 1. n шаров раскладывают по N ящикам. Найдите вероятность того, что для каждого $i = 1, 2, \dots, N$ в i -м ящике лежит n_i шаров, где $n_1 + n_2 + \dots + n_N = n$, если

1. шары различимы,

2. шары неразличимы.

Решение. Начнём со случая различимых шаров. В первый ящик необходимо выбрать n_1 шаров из n , что можно сделать $\binom{n}{n_1}$ способами. Для второго ящика надо выбрать n_2 шаров из $n - n_1$, что даёт $\binom{n-n_1}{n_2}$. Рассуждая аналогично, получаем, что всего есть $\binom{n}{n_1} \binom{n-n_1}{n_2} \dots \binom{n-n_1-\dots-n_{N-1}}{n_N} = \frac{n!}{n_1!(n-1)!} \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \dots \frac{(n-n_1-\dots-n_{N-1})!}{n_N!(n-n_1-\dots-n_{N-1}-n_N)!} = \frac{n!}{n_1!n_2!\dots n_N!}$ успешных исходов. Всего же исходов N^n . Тогда искомая вероятность равна

$$P = \frac{n!}{N^n n_1! n_2! \dots n_N!}.$$

Теперь рассмотрим случай, когда шары неразличимы. Заметим, что тогда есть лишь один подходящий случай. Всего же случаев $\binom{n+N-1}{N-1}$. Тогда вероятность равна

$$P = \frac{1}{\binom{n+N-1}{N-1}}.$$

□

Примечание. Заметим, что число успешных исходов в случае различимых шаров, равное $\frac{n!}{n_1!n_2!\dots n_N!}$, принято называть *мультиномиальным коэффициентом*. Его можно получить, рассматривая полином $(x_1 + x_2 + \dots + x_N)^n$.

Задача 2.

- 1) Случайно бросаются три N -гранных кубика, на гранях которых написаны числа от 1 до N . Найдите вероятность события $A_i = \{\text{сумма чисел, выпавших на кубиках, равна } i\}$, $i \leq N + 2$.
- 2)* Случайно бросаются три N -гранных кубика, на гранях которых написаны числа от 1 до N . Найдите вероятность события $A_i = \{\text{сумма чисел, выпавших на кубиках, равна } i\}$, $i = 2, \dots, 2N$.
- 3)** Случайно бросаются k различных N -гранных кубиков, на гранях которых написаны от 1 до N . Найдите вероятность события $A_i = \{\text{сумма чисел, выпавших на кубиках, равна } i\}$, $i = k, \dots, kN$.

Решение пункта (1). Пусть $i = k_1 + k_2 + k_3$, и $1 \leq k_1, k_2, k_3 \leq N$. Как посчитать число подходящих наборов? Воспользуемся методом точек и перегородок. Пусть есть i точек и нужно расставить 2 перегородки по $i - 1$ допустимой позиции. Тогда есть $\binom{i-1}{2}$ допустимых наборов. Отсюда получаем, что искомая вероятность равна

$$P[i] = \frac{(i-1)(i-2)}{2N^3}.$$

□

Примечание. Пункты (2) и (3) на данный момент слишком сложны. Их адекватное решение будет рассказано ближе к концу курса.

Задача 3. Пусть выбирается произвольная перестановка из S_n . Какова вероятность того, что 1 и 2 будут лежать в одном цикле?

Решение. Воспользуемся тем фактом, что каждая перестановка однозначно представима в виде композиции циклов. Пусть цикл, содержащий 1 и 2, состоит из $k + 1$ элемента ($1 \leq k \leq n - 1$). Позицию для 2 можно выбрать k способами. Далее будем заполнять цикл. Есть $(n-2)(n-3) \dots (n-k)$ вариантов его заполнения. Остальное же мы можем заполнять, как хотим. Следовательно, итого есть $k(n-2)(n-3) \dots (n-k)(n-k-1)! = k(n-2)!$ допустимых перестановок с нужным циклом размера k . Тогда, суммируя по k от 1 до $n-1$, получаем $(n-2)!(1 + 2 + \dots + (n-1)) = \frac{1}{2}n!$. Но всего перестановок $n!$. Тогда вероятность равна $1/2$. □

Задача 4. Пусть в группе 25 студентов. Считаем, что дни рождения равновероятны и случайны. Найдите вероятность того, что найдётся ровно одна пара студентов такая, что

- дни рождения у них совпадают

- у всех других студентов дни рождения не совпадают с днём рождения данной пары студентов

Доказательство. Для начала посчитаем вероятность дополнения к событию. В данном случае дополнением является событие “нет такой пары студентов, что их дни рождения совпадают и у остальных они другие”.

Рассмотрим событие $A_i = \{\text{дни рождения в день } i \text{ совпадают только у двух человек}\}$. Его вероятность равна $P(A_i) = \frac{\binom{25}{2} \cdot 364^{23}}{365^{25}}$. Теперь посчитаем вероятность объединения k событий $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ равна

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \frac{\binom{25}{2} \binom{23}{2} \dots \binom{25-2(k-1)}{2} (365-k)^{25-2k}}{365^{25}}.$$

Теперь посчитаем вероятность дополнения. Она равна $P\left(\bigcap_{i=1}^{365} \overline{A_i}\right)$. Её можно посчитать с помощью формулы включений-исключений. Теперь введём функцию $\alpha(x, y)$, где x — количество студентов, а y — количество дней. Данная функция равна вероятности того, что “среди x студентов нет такой пары студентов, что их дни рождения совпадают, и среди других студентов они другие и находятся среди y выбранных дней”. Она считается аналогично.

Теперь посчитаем вероятность из условия. Для этого выберем двух человек из 25, выберем им день рождения. После чего посчитаем вероятность дополнения для 23 студентов и 364 дней и умножим на 364^{23} . Тогда ответ равен

$$P = \frac{\binom{25}{2} \cdot 365 \cdot \alpha(23, 364) \cdot 364^{23}}{365^{25}}.$$

□

Теперь рассмотрим несколько классических задач на условную вероятность.

Задача 5 (Парадокс Монти-Холла). *Вы участвуете в игре, в которой надо выбрать одну дверь из трёх. За одной из них автомобиль, а за другими — козы. Вы выбрали первую дверь. Ведущий открыл третью дверь, за которой стоит коза. Ведущий предлагает изменить выбор с первой двери на вторую. Стоит ли это делать?*

Решение. Рассмотрим два решения — элементарное и через теорему Байеса. Начнём со второго.

Пусть $C_i = \{\text{машина стоит за } i\text{-й дверью}\}$. Очевидно, что $P(C_i) = 1/3$. Теперь введём событие $H = \{\text{ведущий открывает третью дверь}\}$. Так как ведущий не желает открывать дверь с автомобилем, то условные вероятности будут равны

$$\begin{aligned} P(H | C_1) &= 1/2 \\ P(H | C_2) &= 1 \\ P(H | C_3) &= 0 \end{aligned}$$

Теперь посчитаем вероятность $P(C_2 | H)$. По теореме Байеса она равна

$$P(C_2 | H) = \frac{P(H | C_2) P(C_2)}{P(H | C_1) P(C_1) + P(H | C_2) P(C_2) + P(H | C_3) P(C_3)} = \frac{1 \cdot 1/3}{1/2 \cdot 1/3 + 1 \cdot 1/3 + 0 \cdot 1/3} = \frac{2}{3}$$

Если же не менять дверь, то вероятность не изменится и будет равна $1/3$. Поэтому выгоднее изменить выбор двери.

Теперь рассмотрим элементарное решение. Так как вероятность того, что за дверью будет машина, равна $1/3$, то на вторую и третью дверь вместе приходится $2/3$. Так как после открытия третьей двери оказалось, что за ней стояла коза, то вероятность “переходит” второй двери. Тогда вероятность того, что за первой дверью будет машина, равна $1/3$, а за второй — $2/3$. Выбор очевиден. \square

Задача 6 (Задача о поручике Ржевском). *Поручик Ржевский пришёл в казино и решил поиграть на деньги. Сначала у него есть 8 рублей, и он хочет выйти из казино с 256 рублями. Ему предлагают две тактики:*

- *Каждый раз идти ва-банк.*
- *Каждый раз играть на 1 рубль.*

Какую тактику выбрать поручику, если вероятность выигрыша составляет: (a) $1/4$, (b) $3/4$?

Решение. В случае первой тактики всё просто — он не имеет права проиграть. Поэтому вероятность того, что он уйдёт с желаемой суммой, равна p^5 .

Со второй тактикой дела обстоят интереснее. Введём событие $p_l = \{\text{поручик получил желаемое, изначально имея } l \text{ рублей}\}$. По формуле полной вероятности получим рекурсивную формулу. Вместе с начальными условиями получаем систему

$$\begin{aligned} p_l &= (1-p)p_{l-1} + pp_{l+1} \\ p_0 &= 0 \\ p_{256} &= 1 \end{aligned}$$

Выпишем характеристическое уравнение: $p\lambda^2 - \lambda + (1-p) = 0$. Корни этого уравнения имеют вид

$$\frac{1 \pm \sqrt{1-4p+4p^2}}{2p} = \frac{1 \pm (1-2p)}{2p} = \left[\frac{1}{p}, \frac{1-p}{p} \right]$$

Тогда получаем, что $p_l = a_1 + a_2 \left(\frac{1-p}{p}\right)^l$. Теперь определим значение констант. Так как $p_0 = 0$, то $0 = a_1 + a_2$ и $a_1 = -a_2 = -a$. Теперь рассмотрим p_{256} :

$$1 = -a + a \left(\frac{1-p}{p}\right)^{256} \implies a = \frac{1}{\left(\frac{1-p}{p}\right)^{256} - 1}$$

Отсюда получаем, что $p_l = \frac{\left(\frac{1-p}{p}\right)^l - 1}{\left(\frac{1-p}{p}\right)^{256} - 1}$ и вероятность успеха второй стратегии равна p_8 .

Теперь осталось посчитать. Сначала посчитаем для $p = 1/4$:

$$\begin{aligned} P_I &= \left(\frac{1}{4}\right)^5 = \frac{1}{1024} \\ P_{II} &= \frac{\left(\frac{1-1/4}{1/4}\right)^8 - 1}{\left(\frac{1-1/4}{1/4}\right)^{256} - 1} = \frac{3^8 - 1}{3^{256} - 1} \approx \frac{1}{3^{248}} \end{aligned}$$

В таком случае шанс выйти из казино по своей воле выше, если каждый раз играть ва-банк.

Теперь же посчитаем для $p = 3/4$:

$$P_I = \left(\frac{3}{4}\right)^5 = \frac{243}{1024}$$

$$P_{II} = \frac{\left(\frac{1-3/4}{3/4}\right)^8 - 1}{\left(\frac{1-3/4}{3/4}\right)^{256} - 1} = \frac{1 - (1/3)^8}{1 - (1/3)^{256}} \approx \frac{6560}{6561}$$

В таком же случае гораздо безопаснее каждый раз играть на один рубль. \square

Задача 7 (Задача о контрольной работе). Пусть студенты A, B, C пишут контрольную. Студент A решает любую задачу с вероятностью $3/4$, студент B — с вероятностью $1/2$, а студент C — с вероятностью $1/4$. В контрольной работе 4 задачи. Преподаватель получает анонимную работу, в которой решено 3 задачи. Кому эта работа скорее всего принадлежит?

Решение. Пусть $D = \{\text{автор решил 3 задачи из 4}\}$. Теперь введём ещё три события: $D_A = \{\text{автор — студент } A\}$, $D_B = \{\text{автор — студент } B\}$, $D_C = \{\text{автор — студент } C\}$. Очевидно, что $P(D_A) = P(D_B) = P(D_C) = \frac{1}{3}$. Найдём условную вероятность события D для разных условий:

$$P(D | D_A) = \left(\frac{3}{4}\right)^3 \cdot \binom{4}{1} \cdot \frac{1}{4} = \frac{27}{64}$$

$$P(D | D_B) = \left(\frac{1}{2}\right)^3 \cdot \binom{4}{1} \cdot \frac{1}{2} = \frac{1}{4} = \frac{16}{64}$$

$$P(D | D_C) = \left(\frac{1}{4}\right)^3 \cdot \binom{4}{1} \cdot \frac{3}{4} = \frac{3}{64}$$

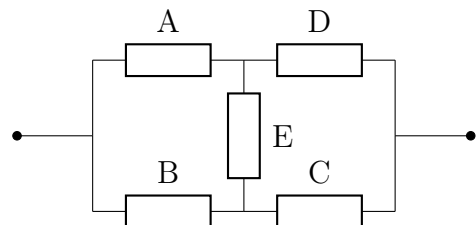
Теперь воспользуемся теоремой Байеса:

$$P(D_A | D) = \frac{P(D | D_A) P(D_A)}{P(D | D_A) P(D_A) + P(D | D_B) P(D_B) + P(D | D_C) P(D_C)}$$

$$= \frac{\frac{27}{64}}{\frac{27}{64} + \frac{16}{64} + \frac{3}{64}} = \frac{27}{46}$$

Аналогично получаем, что $P(D_B | D) = \frac{16}{46}$ и $P(D_C | D) = \frac{3}{46}$. Следовательно, эту работу, скорее всего, сдал студент A . \square

Задача 8. Пусть пять приборов соединены в схему. Каждый из них пропускает ток с вероятностью p . Какова вероятность того, что схема пропускает ток? Какова вероятность того, что есть ток, но при этом E сломан?



Решение. Для начал посмотрим, по каким путям может пройти ток:

- Ток может пойти по AD — тогда вероятность того, что ток будет, равна p^2 .
- Если D сломан, то ток может пойти по AEC . Вероятность такого случая равна $p^3(1 - p)$.

- Если A сломан, то ток может пойти по BC . Вероятность этого равна $p^2(1-p)$.
- Если же сломаны и A , и C , то ток пойдёт по BED . Вероятность такого равна $p^3(1-p)^2$.

Тогда итоговая вероятность равна сумме:

$$P(\text{в цепи есть ток}) = p^2 + p^3(1-p) + p^2(1-p) + p^3(1-p)^2.$$

Теперь ответим на второй вопрос. Воспользуемся определением условной вероятности:

$$P(E \text{ не проводит ток} \mid \text{в цепи есть ток}) = \frac{P(E \text{ не проводит ток и в цепи есть ток})}{P(\text{в цепи есть ток})}$$

Вероятность сверху посчитать несложно — достаточно рассмотреть допустимые пути. Тогда ответ равен

$$P(E \text{ не проводит ток} \mid \text{в цепи есть ток}) = \frac{p^2(1-p) + p^2(1-p)^2}{p^2 + p^3(1-p) + p^2(1-p) + p^3(1-p)^2}. \quad \square$$

2.3 Семинар от 23.09.2016

Как обычно, начнём с разбора домашнего задания.

Задача 1. В ящике N различных шаров, из которых ровно M белых. Последовательно вынимают $n \leq N$ шаров. Пусть событие A_k означает, что k -й по счёту вынутый шар — белый, а событие B_m — что всего вынули $m \leq M$ белых шаров.

Найдите $P(A_k \mid B_m)$, если (а) шары вынимаются без возвращения, (б) с возвращением.

Решение. Начнём со случая, когда нельзя возвращать шары. По определению условной вероятности $P(A_k \mid B_m) = \frac{P(A_k \cap B_m)}{P(B_m)}$. Для начала посчитаем $P(B_m)$. Как это сделать? Зафиксируем набор из n шаров, в котором первые m шаров белые. Какова вероятность того, что выпадет такой набор? Она равна

$$\frac{M}{N} \cdot \frac{M-1}{N-1} \cdot \dots \cdot \frac{M-m+1}{N-m+1} \cdot \frac{N-M}{N-m} \cdot \frac{N-M-1}{N-m-1} \cdot \dots \cdot \frac{N-M-(n-m)+1}{N-n+1}.$$

Теперь заметим, что если переставить числители местами, то получится вероятность того, что выпадет какой-то другой набор из n шаров, среди которых m белых. Тогда вероятность того, что выпадет хоть какой-то набор, подходящий под это условие, равна

$$P(B_m) = \binom{n}{m} \frac{\frac{M!}{(M-m)!} \frac{(N-M)!}{(N-M-(n-m))!}}{\frac{N!}{(N-n)!}} = \frac{\binom{n}{m} \binom{N-n}{M-m}}{\binom{N}{M}}.$$

Теперь перейдём к числителю. Как посчитать $P(A_k \cap B_m)$? В принципе, точно так же, как и $P(B_m)$. Однако, в данном случае зафиксирована k -я позиция, поэтому нужно лишь выбрать $m-1$ позицию из $n-1$ для белых шаров. Тогда

$$P(A_k \cap B_m) = \frac{\binom{n-1}{m-1} \binom{N-n}{M-m}}{\binom{N}{M}}.$$

Отсюда получаем, что $P(A_k | B_m) = \frac{\binom{n-1}{m-1}}{\binom{n}{m}} = \frac{m}{n}$.

Переходим к случаю (б). Опять же, посчитаем $P[B_m]$ и $P[A_k \cap B_m]$. Рассуждения о перестановке так же имеют место, поэтому:

$$P(B_m) = \binom{n}{m} \frac{M^m (N-M)^{n-m}}{N^n}$$

$$P(A_k \cap B_m) = \binom{n-1}{m-1} \frac{M^m (N-M)^{n-m}}{N^n}$$

Подставляя полученные значения в формулу условной вероятности, получаем, что

$$P(A_k | B_m) = \frac{\binom{n-1}{m-1}}{\binom{n}{m}} = \frac{m}{n}.$$

Ответ: $\frac{m}{n}$ в обоих случаях. □

Задача 2. Ящик содержит a белых и b чёрных шаров (все шары различимы). Наудачу извлекается шар. Он возвращается обратно, и, кроме того, добавляется c шаров одного с ним цвета. Далее, подобная процедура повторяется снова. Пусть событие A_k означает, что на k -м шаге извлечён белый шар. Найдите

- (а) вероятность того, что при первых $n = n_1 + n_2$ извлечениях попало n_1 белых и n_2 чёрных шаров;
- (б) вероятность события A_k ;
- (в) условную вероятность $P(A_m | A_k)$ при $m > k$;
- (г) условную вероятность $P(A_m | A_k)$ при $m < k$;

Решение. Рассмотрим ситуацию, когда последовательно выпало n_1 белых и n_2 чёрных шаров. Какова вероятность такого события? Она равна

$$\frac{a}{a+b} \cdot \frac{a+c}{a+b+c} \cdots \frac{a+(n_1-1)c}{a+b+(n_1-1)c} \cdot \frac{b}{a+b+n_1c} \cdot \frac{b+c}{a+b+(n_1+1)c} \cdots \frac{b+(n_2-1)c}{a+b+(n-1)c}.$$

Теперь переставим числители так, чтобы числители вида $a+x$ и $b+x$ были отсортированы по возрастанию. Тогда эта вероятность будет соответствовать какому-то другому набору из n_1 белых и n_2 чёрных. Пусть $P(a, b, n_1, n_2)$ — вероятность того, что из a белых и b чёрных при первых $n = n_1 + n_2$ извлечениях попало n_1 белых и n_2 чёрных шаров. Тогда

$$P(a, b, n_1, n_2) = \binom{n}{n_1} \frac{a(a+c) \cdots (a+(n_1-1)c) b(b+c) \cdots (b+(n_2-1)c)}{(a+b)(a+b+c) \cdots (a+b+(n-1)c)}.$$

Теперь посчитаем $P(A_k)$. Пусть $C_{ki} = \{\text{до } k\text{-ой процедуры вытащили ровно } i \text{ белых шаров}\}$. Очевидно, что эти события образуют разбиение вероятностного пространства.

Тогда по формуле полной вероятности $P(A_k) = \sum_{i=0}^{k-1} P(A_k \cap C_{ki})$. Заметим, что $P(A_k \cap C_{ki})$ совпадает с $P(a, b, i+1, k-i-1)$ с тем лишь отличием, что в данном случае нужно выбрать i позиций из $k-1$:

$$P[A_k \cap C_{ki}] = \binom{k-1}{i} \frac{a(a+c) \cdots (a+ic) b(b+c) \cdots (b+(k-i)c)}{(a+b)(a+b+c) \cdots (a+b+(n-1)c)}.$$

Тогда

$$P(A_k) = \frac{a}{a+b} \sum_{i=0}^{k-1} \binom{k-1}{i} \frac{(a+c) \dots (a+ic)b(b+c) \dots (b+(k-i)c)}{(a+b+c) \dots (a+b+(n-1)c)}.$$

Теперь заметим, что элемент суммы есть ни что иное, как $P(a+c, b, i, k-i-1)$. Но

$$\sum_{i=0}^{k-1} P(a+c, b, i, k-i-1) = 1,$$

так как эта сумма соответствует вероятности вытащить любой набор. Отсюда следует, что

$$P(A_k) = \frac{a}{a+b}.$$

Перейдём к третьему (да и четвёртому тоже) пункту. По определению условной вероятности: $P(A_k | A_m) = \frac{P(A_k \cap A_m)}{P(A_m)}$. Как посчитать числитель? Точно так же, как и во втором случае. Пропустив аналогичные выкладки, выпишем ответ:

$$P(A_k \cap A_m) = \frac{a(a+c)}{(a+b)(a+b+c)}.$$

Тогда получаем, что $P(A_k | A_m) = \frac{a+c}{a+b+c}$. □

Задача 3. Пусть A, B, C — попарно независимые равновероятные события, причём $A \cap B \cap C = \emptyset$. Найти максимально возможное значение $P(A)$.

Решение. Начнём с того, что заметим следующее: $P(A) \geq P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C))$. Так как $A \cap B \cap C = \emptyset$, то $(A \cap B) \cap (A \cap C) = \emptyset$. Следовательно, $P[A] \geq P[A \cap B] + P[A \cap C] = 2(P[A])^2$ и $P(A) \leq 1/2$.

Приведём пример, когда выполняется условие, причём $P(A) = 1/2$. Рассмотрим классическую модель $\Omega = \{1, 2, 3, 4\}$ и события: $A = \{1, 2\}$, $B = \{2, 3\}$, $C = \{3, 4\}$. Легко понять, что данные события удовлетворяют условию и $P(A) = 1/2$. □

Задача 4. Игроки A и B играют в теннис. При розыгрыше на подаче A игрок A выигрывает с вероятностью p_1 , а при розыгрыше на подаче B — с вероятностью p_2 , все розыгрыши независимы. Игрок A подаёт первым, а выигрывает тот, кто первым наберёт n очков. Существует два варианта правил перехода подачи:

(а) поочерёдная;

(б) игрок подаёт до тех пор, пока не проиграет розыгрыш.

Покажите, что вероятность выигрыша A не зависит от правил перехода подачи, и вычислите её.

Решение. Будем считать, что всего было проведено $2n-1$ розыгрышей. При таком количестве один игрок гарантированно наберёт не менее n очков, а второй — гарантированно меньше. Теперь опишем вероятностное пространство. Элементарные исходы будут иметь

вид $\omega = (a_1, a_2, \dots, a_{2n-1})$, где $a_i \in \{0, 1\}$ (0 соответствует проигрышу, 1 — победе). Согласно этой схеме исход будет подходящим, если в наборе будет не меньше n единиц.

Начнём с пункта (а).

Поймём, как посчитать вероятность какого-либо элементарного случая. Пусть $n = 4$ и мы хотим найти вероятность элементарного исхода 0110110. Она равна $(1 - p_1)p_2p_1(1 - p_2)p_1p_2(1 - p_1) = p_1^2p_2^2(1 - p_1)^2(1 - p_2)^1$. Отсюда получаем закономерность: вероятность элементарного исхода $a_1a_2 \dots a_{2n-1}$ равна

$$p_1^{\sum_{i=0}^{n-1} a_{2i+1}} p_2^{\sum_{i=1}^{n-1} a_{2i}} (1 - p_1)^{n - \sum_{i=0}^{n-1} a_{2i+1}} (1 - p_2)^{n-1 - \sum_{i=1}^{n-1} a_{2i}}.$$

Пусть $k_1 = \sum_{i=0}^{n-1} a_{2i+1}$ — количество единиц на нечётных местах, а $k_2 = \sum_{i=1}^{n-1} a_{2i}$ — на чётных.

Тогда вероятность того, что A выиграл, будет равна

$$P = \sum_{\substack{k_1, k_2 \\ k_1 + k_2 \geq n \\ k_1, k_2 \leq n}} \binom{n}{k_1} \binom{n-1}{k_2} p_1^{k_1} p_2^{k_2} (1 - p_1)^{n-k_1} (1 - p_2)^{n-1-k_2}.$$

Теперь перейдём к пункту (б).

Докажем следующее: между данными методами подачи есть биекция, т.е. игре с поочерёдной подачей можно сопоставить игру с подачей до проигрыша. Рассмотрим частный случай: поочерёдно вышел исход 0110110. Тогда можно “раскидать” партии так:

Подаёт первый: 0110

Подаёт второй: 101

Биекция будет иметь вид 0101101. Как её построить? Разбиваем исход на подачи первого и второго игрока, тем самым получая строки длиной n и $n - 1$. После этого строим по ним новую строку по следующему алгоритму:

Алгоритм 1 Построение исхода в случае подачи до проигрыша по исходу в случае поочерёдной подачи

- 1: Начинаем со строки длины n (строки для первого игрока);
 - 2: Копируем строку посимвольно до тех пор, пока не попадём на 0;
 - 3: Переходим на другую строку;
 - 4: Повторяем два предыдущих шага до тех пор, пока не перенесём все символы.
-

Алгоритм построения исхода при поочерёдной подаче по исходу при подаче до проигрыша будет почти аналогичен. \square

Перейдём к задачам на тему математического ожидания и дисперсии.

Задача 5. Бросили два N -гранных кубика. Пусть ξ — сумма выпавших очков. Найдите $E[\xi]$ и $D[\xi]$.

Решение. Пусть $E[\xi_1]$ — матожидание количество очков, выпавших на первом кубике. Посчитать его несложно: $E[\xi_1] = \sum_{i=1}^N \frac{i}{N} = \frac{N+1}{2}$. Заметим, что $E[\xi] = E[\xi_1 + \xi_2] = E[\xi_1] + E[\xi_2]$. Тогда

$$E[\xi] = \frac{N+1}{2} + \frac{N+1}{2} = N+1.$$

Дисперсию будем считать по следующей формуле: $D[\xi_1] = E[\xi_1^2] - (E[\xi_1])^2$. Посчитаем первый член:

$$E[\xi_1^2] = \sum_{i=1}^N \frac{i^2}{N} = \frac{N(2N+1)(N+1)}{6N} = \frac{(2N+1)(N+1)}{6}.$$

Отсюда получаем, что

$$D[\xi_1] = \frac{N+1}{2} \left(\frac{2N+1}{3} - \frac{N+1}{2} \right) = \frac{(N+1)(N-1)}{12} = \frac{N^2-1}{12}.$$

Заметим, что ξ_1 и ξ_2 независимы (ведь кубики тоже независимы). Тогда $D[\xi] = D[\xi_1] + D[\xi_2]$ и

$$D[\xi] = \frac{N^2-1}{6}. \quad \square$$

Задача 6. Пусть выбрана случайная перестановка $\sigma \in S_n$. Введём случайную величину ξ , равная количеству стационарных точек (чисел i таких, что $\sigma(i) = i$). Найдите $E[\xi]$ и $D[\xi]$.

Решение. Введём событие $A_i = \{\sigma(i) = i\}$. Тогда $\xi = \sum_{i=1}^n I_{A_i}$. По свойству линейности:

$$E[\xi] = E\left[\sum_{i=1}^n I_{A_i}\right] = \sum_{i=1}^n E[I_{A_i}] = \sum_{i=1}^n P(A_i)$$

Так как $P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$, то $E[\xi] = 1$.

Теперь перейдём к подсчёту дисперсии. Распишем дисперсию через ковариации:

$$D[\xi] = \text{cov}(\xi, \xi) = \text{cov}\left(\sum_{i=1}^n I_{A_i}, \sum_{i=1}^n I_{A_i}\right) = \sum_{i=1}^n \sum_{j=1}^n \text{cov}(I_{A_i}, I_{A_j}).$$

Посчитаем $\text{cov}(I_{A_i}, I_{A_j})$. По свойству ковариации она равна

$$E[I_{A_i} I_{A_j}] - E[I_{A_i}] E[I_{A_j}] = E[I_{A_i \cap A_j}] - E[I_{A_i}] E[I_{A_j}] = P[A_i \cap A_j] - P[A_i] P[A_j].$$

Возникают два случая:

(а) $i = j$. Тогда $\text{cov}(I_{A_i}, I_{A_j}) = \frac{1}{n} - \frac{1}{n^2}$.

(б) $i \neq j$. Тогда $P[A_i \cap A_j] = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$ и $\text{cov}(I_{A_i}, I_{A_j}) = \frac{1}{n^2(n-1)}$.

Отсюда получаем, что

$$D[\xi] = \frac{n(n-1)}{2} \frac{1}{n^2(n-1)} + n \left(\frac{1}{n} - \frac{1}{n^2} \right) = 1. \quad \square$$