Лекции по предмету
Теория вероятностей

Денис Беляков
Никита Попов
Алексей Хачиянц

2016/2017 учебный год

1 Организационные моменты

Формат отчётности:

1. Две контрольных работы.
2. Два коллоквиума.
3. Домашние задания. В среднем на каждом семинаре будут выдавать по 2-3 задачи для самостоятельного решения, которые будет нужно сдавать ассистентам.
4. Письменный экзамен — "расширенная КР"

Итоговая оценка считается следующий образом:

$$O_{\text{итог}} = 0.3 \cdot O_{\text{экз}} + 0.1 \cdot O_{\text{ДЗ}} + 0.3 \cdot O_{\text{КР}} + 0.3 \cdot O_{\text{коллок}}$$

Округление оценки арифметическое.
На данный момент автоматов не предусмотрено.

2 Лекция от 09.09.2016

2.1 Введение

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 *детерменированным*. Пример такого эксперимента — выбрасывание кирпича из окна. Очевидно, что кирпич упадёт на землю[1] и результат предопределён. Такие задачи изучают в той же линейной алгебре или где-либо ещё, но не в теории вероятностей.

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

- А теперь предположим, что заранее сказать, каков будет результат, невозможно. Например, точно сказать, какой стороной упадёт подброшенная монетка, вряд ли получится. Тогда результат называют *недетерменированным*. Именно задачи с недетерменированным 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.

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*

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

Получаемое число p называют *вероятностью* события A и обозначают $P(A)$. Например, P(“встретить живого динозавра”) $= 0$, так как они все вымерли.

2.2 Вероятностное пространство

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:

[2] Оставим вопрос о том, как он не поленился провернуть это, без ответа.

1. $\Omega = \{\text{О}, \text{Р}\}$ — бросок монеты.
2. $\Omega = \{\omega_1, \dots, \omega_6\}$, где $\omega_i =$ "выпало i очков" — бросок игрального кубика.
3. $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$, где $\omega_i =$ "на данный момент горит i 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 $A = \{$выпало чётное число очков$\}$. Тогда $A = \{\omega_2, \omega_4, \omega_6\}$.

Определение 3. Отображение $P : \Omega \to [0, 1]$ 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

$$P(A) = \sum_{\omega \in A} P(\omega)$$

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

Перед тем, как идти дальше, напомню 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 простейшие свойства вероятности.

Лемма. *Для любого дискретного вероятностного пространства* (Ω, P) *выполняется следующее:*

1. $P(\Omega) = 1,\ P(\varnothing) = 0.$
2. *Конечная аддитивность:*
$$P\left(\bigsqcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

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

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

Последнее свойство выполняется только для счётного Ω.

Доказательство. Докажем каждый пункт леммы:

1. $P(\Omega) = 1$ следует из определения вероятности, а $P(\varnothing) = 0$ следует из определения вероятности события.
2. Случай с конечным множеством Ω очевиден. 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:
$$P(A) = \sum_{\omega \in A} P(\omega) = \sum_{i\,:\,\omega_i \in A} P(\omega_i) = \lim_{N \to \infty} \sum_{\substack{i\,:\,\omega_i \in A \\ i < N}} P(\omega_i)$$
   Теперь распишем вероятность дизъюнктного объединения событий $A_1, A_2, \ldots, A_n$:
$$P\left(\bigsqcup_{i=1}^{n} A_i\right) = \sum_{\omega \in \bigsqcup A_i} P(\omega) = \lim_{N \to \infty}\left(\sum_{\substack{i\,:\,\omega_i \in \bigsqcup A_i \\ i < N}} P(\omega_i)\right) = \lim_{N \to \infty} \sum_{i=1}^{n} \sum_{\substack{j\,:\,\omega_j \in A_i \\ j < N}} P(\omega_j) =$$
$$= \sum_{i=1}^{n} \lim_{N \to \infty} \sum_{\substack{j\,:\,\omega_j \in A_i \\ j < N}} P(\omega_j) = \sum_{i=1}^{n} P(A_i)$$
3. Согласно пункту 2, $1 = P(\Omega) = P(A \cup \overline{A}) = P(A) + P(\overline{A})$.
4. Так как $A \cup B = A \cup (B \setminus A)$ и $A \cap (B \setminus A) = \varnothing$, то
$$P(A \cup B) = P(A) + P(B \setminus A)$$
   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.[4]

□

2.3 Классическая модель

Пусть (Ω, P) — некоторое конечное вероятностное пространство, при этом все элементарные исходы равновероятны. Тогда легко посчитать вероятность элементарного исхода:

$$P(\omega) = \frac{1}{|\Omega|} \text{ для всех } \omega \in \Omega$$

Такую модель называют *классической*.

Как посчитать вероятность события в классической модели? Очень просто:

$$P(A) = \frac{|A|}{|\Omega|} \text{ для всех } A \subseteq \Omega$$

Рассмотрим некоторые примеры классических моделей.

1. Бросок монетки. В таком случае $\Omega = \{\text{О}, \text{Р}\}$ и $P(\text{О}) = P(\text{Р}) = \frac{1}{2}$.

2. Бросок двух монеток. С этой моделью связано одно заблуждение Д'Аламбера. Он рассуждал следующим образом: так как $\Omega = \{\text{ОО}, \text{РР}, \text{ОР}\}$, то $P(\text{ОО}) = P(\text{РР}) = P(\text{ОР}) = \frac{1}{3}$. Но это опровергается экспериментами. И как это исправить? Есть два варианта:

   - Можно сказать, что модель не является классической и поправить вероятности: $P(\text{ОО}) = P(\text{РР}) = \frac{1}{4}$, $P(\text{ОР}) = \frac{1}{2}$.
   - А можно просто изменить множество элементарных исходов. Начнём учитывать порядок выпадения: $\Omega = \{\text{ОО}, \text{РР}, \text{ОР}, \text{РО}\}$. 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:

   $$\Omega = \{\omega = (\omega_1, \omega_2, \ldots, \omega_n) \mid \omega_i \in \{\text{О}, \text{Р}\}\}$$

   Легко понять, что в данной модели 2^n элементарных исходов. Замечание: вероятностное пространство такого вида называют *симметрической схемой Бернулли*.

[4] На самом деле я попробую найти доказательство. Когда-нибудь.

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:

$$P(\omega) = p^{\sum\limits_{i=1}^{n} \omega_i} (1-p)^{n - \sum\limits_{i=1}^{n} \omega_i}$$

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	$C_{n+m-1}^m = \dfrac{(n+m-1)!}{m!(n-1)!}$
Без возврата	$A_n^m = \dfrac{n!}{(n-m)!}$	$C_n^m = \dfrac{n!}{m!(n-m)!}$

Доказательство. (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 $C_{n+m-1}^m = \frac{(n+m-1)!}{m!(n-1)!}$.

□

2.4 Условная вероятность

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

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

Если $P(B) = 0$, то положим, что $P(A \mid B) = 0$ 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 1. Если $P(B) > 0$, то $\tilde{P}(A) = P(A \mid B)$ тоже является вероятностью на Ω.

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

Теперь докажем важную формулу:

Формула полной вероятности. *Пусть $B_1, B_2, \ldots, B_n$ — разбиение Ω. Тогда для любого события $A \subseteq \Omega$ верно, что*

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

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

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i)$$

$\square$

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