Теория вероятности

Дима Трушин

Семинар 4

Случайные вектор

Пусть теперь на наш черный ящик, описываемый парой (Ω, P) мы хотим прицепить несколько измерительных устройств. Это означает, что мы имеем несколько случайных величин $\xi_1, \ldots, \xi_n \colon \Omega \to \mathbb{R}$. В этом случае удобно все эти случайные величины объединить в один вектор и рассмотреть $\bar\xi \colon \Omega \to \mathbb{R}^n$ по правилу $\omega \mapsto \bar\xi(\omega) = (\xi_1(\omega), \ldots, \xi_n(\omega))$. То есть для того, чтобы изучать несколько случайных величин вместе, достаточно изучить один случайный вектор. Тут возникает первый вопрос: а почему бы нам не изучать каждую из случайных величин отдельно? Основная проблема в том, что, изучая случайные величины отдельно, мы не можем ничего узнать о вероятностях их взаимодействия между собой. Например, если мы кидаем монетку и ξ дает 0 в случае орла и 1 в случае решки. Можно определить η, которая действует наоборот, дает 1 в случае орла и 0 в случае решки. Тогда у нас эти случайные величины не могут одновременно дать два нуля или две единицы. То есть между ними есть некое взаимодействие. Чтобы понять это взаимодействие их надо изучать одновременно, потому надо изучить случайный вектор (ξ, η).

Как и в случае одной случайной величины, случайный вектор $\bar\xi \colon \Omega \to \mathbb{R}^n$ дает новое распределение в пространстве векторов $\mathbb{R}^n$. А именно, для любого подмножества $A \subseteq \mathbb{R}^n$ рассмотрим вероятность попасть в это подмножество

$$P_{\bar\xi}(A) = P(\bar\xi \in A) = P(\{\omega \mid \bar\xi(\omega) \in A\}) = P(\bar\xi^{-1}(A))$$

Вероятность $P_{\bar\xi}$ называется совместным распределением $\xi_1, \ldots, \xi_n$.

Многомерная функция распределения Как и в одномерном случае, случайный вектор можно задавать некой функцией, называемой функцией распределения, но теперь это функция от нескольких переменных

$$F_{\bar\xi}(x_1, \ldots, x_n) = P_{\bar\xi}((-\infty, x_1] \times \ldots \times (-\infty, x_n]) = P(\xi_1 \leqslant x_1, \ldots, \xi_n \leqslant x_n)$$

Функция $F_{\bar\xi}(x_1, \ldots, x_n)$ называется функцией совместного распределения случайных величин $\xi_1, \ldots, \xi_n$. В этом случае можно восстановить вероятность попадания в произвольный прямоугольный параллелепипед $(a_1, b_1] \times \ldots \times (a_n, b_n]$ по формуле включения исключения. Я продемонстрирую это на случае двух переменных. Если

$$F_{(\xi,\eta)}(x_1, x_2) = P(\xi \leqslant x_1, \eta \leqslant x_2) = f(x_1, x_2)$$

Тогда

$$P(a_1 < \xi \leqslant b_1, a_2 < \eta \leqslant b_2) = f(b_1, b_2) - f(a_1, b_2) - f(b_1, a_2) + f(a_1, b_1)$$

И далее существует некая общая процедура (которая интересна только математикам), позволяющая продлить это определение на любые подмножества в $\mathbb{R}^n$.

Виды случайных векторов Как и в случае случайной величины, можно рассматривать два вида случайных векторов: дискретные и непрерывные, то есть все координаты либо одновременно дискретные, либо одновременно непрерывные, да еще и как-то могут между собой взаимодействовать. Бывают и более интересные смеси случайных векторов, когда одна координата дискретна, а другая непрерывна. Бывает еще более интересное сочетание, когда каждая координата является смесью и того и другого. К сожалению, чем сложнее устроены координаты у случайного вектора, тем сложнее с ним взаимодействовать и приходится страдать. Чтобы страдать поменьше и понять, как все это работает я сосредоточу свое внимание в начале на «чистых» случайных векторах (все координаты дискретны или все координаты непрерывны), а потом поговорю, что делать в случае, если еть координаты обоих сортов (такие штуки встречаются в приложениях).[1]

[1]Математики привыкли все это делать в самом общем виде и математической теории глубоко плевать, какие именно виды случайных величин участвуют в законах, она умеет работать со всеми одинаково, только для этого приходится писать формулы, которые вы не поймете, просто потому что я их вам не успею объяснить.

Дискретные случайные векторы Для того, чтобы задать дискретный случайный вектор $\bar{\xi}$ в $\mathbb{R}^n$ надо задать следующие данные

1. Набор векторов $a_1, a_2, \dots, a_n, \dots \in \mathbb{R}^n$. Он может быть конечным или бесконечным. Это будут выделенные значения, которые разрешено принимать нашему случайному вектору. Они еще будут называться атомами.

2. Набор вероятностей $p_1, p_2, \dots, p_n, \dots \in [0, 1]$, которые приписываются соответствующему значению a_i. Они должны обладать свойствами

   (a) $0 < p_k \leqslant 1$.

   (b) $\sum_{k=1}^{\infty} p_k = 1$.

В этом случае $P(\bar{\xi} = a_i) = p_i$ и для любого числа $a \in \mathbb{R}$ не равного ни одному из a_i мы имеем $P(\bar{\xi} = a) = 0$. Для произвольного множества $A \subseteq \mathbb{R}^n$ его вероятность считается

$$P_{\bar{\xi}}(A) = \sum_{a_i \in A} p_i$$

то есть надо сложить вероятности всех атомов, попавших в множество A.

Непрерывные случайные векторы Для того, чтобы задать непрерывный случайный вектор $\bar{\xi}$, надо задать функцию $p\colon \mathbb{R}^n \to \mathbb{R}$, которая называется плотностью случайного вектора, со следующими свойствами

1. $0 \leqslant p(x_1, \dots, x_n)$.

2. $\int_{\mathbb{R}^n} p(x_1, \dots, x_n)\, dx_1 \dots dx_n = 1$

Тогда для произвольного подмножества $A \subseteq \mathbb{R}^n$ вероятность попасть в него равна

$$P_{\bar{\xi}}(A) = \int\limits_A p(x_1, \dots, x_n)\, dx_1 \dots dx_n$$

Примеры

1. Пусть мы бросаем монетку так, что вероятность выпадения орла есть p, а вероятность выпадения решки – q (при этом $p + q = 1$). Пусть ξ – случайная величина, которая равна 0, когда выпадает орел и 1, когда выпадает решка. Тогда это дискретная случайная величина с табличкой

$$\xi \sim \begin{cases} 0 & 1 \\ p & q \end{cases}$$

   Теперь рассмотрим случайную величину $\eta = 1 - \xi$, то есть у нее выпадает 0 на решку и 1 на орла. Тогда η задается табличкой

$$\eta \sim \begin{cases} 0 & 1 \\ q & p \end{cases}$$

   Давайте поймем, а какое же у них будет совместное распределение? Для случайного вектора (ξ, η) получим следующую таблицу

Значения (ξ, η)	Вероятность выпадения
$(0, 0)$	0
$(0, 1)$	p
$(1, 0)$	q
$(1, 1)$	0

2. Пусть теперь у нас есть квадрат $[0, 1]^2$. Рассмотрим функцию плотности $p(x, y)$ такую, что $p(x, y) = 1$ при условии $(x, y) \in [0, 1]^2$ и $p(x, y) = 0$ иначе. Это дает нам $\Omega = [0, 1]^2$ и вероятность $P(A) = \int_A p(x, y)\, dxdy$. Тогда вектор (x, y) – будет случайным вектором на квадрате. Это равномерно распределенный на квадрате вектор в том смысле, что вероятность попасть в подмножество $A \subseteq [0, 1]^2$ в точности равно площади этого подмножества A.

Восстановление распределений координат Предположим нам задан случайный вектор $\bar{\xi} = (\xi_1, \ldots, \xi_n)$ на пространстве $\mathbb{R}^n$ и мы знаем его распределение $P_{\bar{\xi}}$. Вопрос, а как найти распределения его координат ξ_i? Оказывается очень просто. Давайте проделаем на примере ξ_1. Нам надо для любого множества $A \subseteq \mathbb{R}$ посчитать вероятность $P(\xi_1 \in A)$. Но

$$P(\xi_1 \in A) = P(\xi_1 \in A, \xi_2 \in \mathbb{R}, \ldots, \xi_n \in \mathbb{R}) = P_{\bar{\xi}}(A \times \mathbb{R} \times \ldots \times \mathbb{R})$$

С другой стороны может возникнуть желание задать другой вопрос, а как восстановить совместное распределение по распределениям координат? Тут ответ очень простой – никак. Вообще говоря совместное распределение содержит больше информации, чем просто распределение каждой координаты, оно еще и кодирует, как координаты взаимодействуют. Если мы знаем распределение координат, то прежде чем восстанавливать совместное распределение, нам нужна какая-то информация об этом самом взаимодействии между координатами. В общем виде есть только один случай, когда это возможно – при условии независимости координат. Что это такое и как восстанавливать совместное распределение в этом случае я расскажу чуть позже. А пока продемонстрирую вышесказанное на примере дискретного и непрерывного распределений.

Восстановление координат в дискретном случае Если случайный вектор $\bar{\xi}$ задавался данными

$$\xi \sim \begin{cases} a_1 & a_2 & \ldots & a_n & \ldots & \in \mathbb{R}^n \\ p_1 & p_2 & \ldots & p_n & \ldots \end{cases}$$

Обозначим, через $(a_i)_1$ первую координату вектора a_i. Тогда значения случайной величины ξ_1 будут все числа вида $(a_i)_1$, то есть все первые координаты, которые нам встретились в таблице. А вероятность считается по формуле

$$P(\xi_1 = a) = \sum_{(a_i)_1 = a} p_i$$

То есть мы находим все векторы a_i, у которых первая координата равна a. Если таких нет, то вероятность полагается нулем. Если есть, то мы складываем все их вероятности вместе – это и будет вероятность попасть в a.

Восстановление координат в непрерывном случае Пусть случайный вектор $\bar{\xi}$ задан плотностью $p(x_1, \ldots, x_n)$. Тогда посчитаем функцию распределения ξ_1 по определению

$$F_{\xi_1}(x_1) = \int\limits_{(-\infty, x_1] \times \mathbb{R} \times \ldots \times \mathbb{R}} p(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = \int\limits_{-\infty}^{x_1} \left(\int\limits_{\mathbb{R}^{n-1}} p(x_1, x_2, \ldots, x_n)\, dx_2 \ldots dx_n \right) dx_1$$

Но плотность вычисляется, как производная от функции распределения, то есть

$$p_{\xi_1}(x_1) = F_{\bar{\xi}}(x_1)' = \int\limits_{\mathbb{R}^{n-1}} p(x_1, x_2, \ldots, x_n)\, dx_2 \ldots dx_n$$

Таким образом, чтобы получить плотность для первой координаты, надо проинтегрировать совместную плотность по всем остальным координатам вдоль всей прямой.

Например, в случае двумерного случайного вектора (ξ, η) с плотностью $p_{\bar{\xi}}(x, y)$, плотности самих ξ и η получаются по формулам

$$p_\xi(x) = \int\limits_{\mathbb{R}} p(x, y)\, dy \quad \text{и} \quad p_\eta(y) = \int\limits_{\mathbb{R}} p(x, y)\, dx$$

Независимость Очень часто встречается понятие о независимых измерениях. Мы думаем о них, как об измерениях, которые никак друг на друга не влияют, никоим образом. Например, мы можем взять и независимо подбросить монетку, сначала один раз, а потом другой. Или еще лучше, взять двух разных людей и каждый из них независимо от другого подбросит монетку. Мы интуитивно чувствуем, что это должно значить, но хорошо бы иметь математическое выражение для подобного явления, чтобы им как-то пользоваться. Но для начала я хочу привести пример зависимых измерений. Скажем пример, когда ξ дает 0 в случае выпадения орла и 1 в случае выпадения решки, а η наоборот. Тут мы интуитивно понимаем, что если мы знаем ответ для ξ, то и для η ответ восстанавливается однозначно, нет никакой независимости между этими случайными величинами.

Пусть у нас есть две случайные величины ξ и η на одном и том же вероятностном пространстве (Ω, P), то есть это два отображения $\xi, \eta\colon \Omega \to \mathbb{R}$. Они называются независимыми и пишут $\xi \perp\!\!\!\perp \eta$, если для любых множеств $A \subseteq \mathbb{R}$ и $B \subseteq \mathbb{R}$ события $\xi^{-1}(A)$ и $\eta^{-1}(B)$ независимы. То есть событие «ξ попало в A» не зависимо с событием «η попало в B» для любых двух подмножеств. Напомню, что независимость событий означает, что выполнено

$$P(\xi^{-1}(A) \cap \eta^{-1}(B)) = P(\xi^{-1}(A))P(\eta^{-1}(B))$$

или в текстовой форме

$$P(\text{«}\xi \text{ попала в } A\text{» и «}\eta \text{ попала в } B\text{»}) = P(\text{«}\xi \text{ попала в } A\text{»})P(\text{«}\eta \text{ попала в } B\text{»})$$

Давайте я немного поясню это определение. Множества вида $\xi^{-1}(A) \subseteq \Omega$, когда A пробегает все подмножества в $\mathbb{R}$, – это все возможные события, которые можно выразить в терминах ξ. Аналогично, $\eta^{-1}(B) \subseteq \Omega$, когда B пробегает все возможные подмножества в $\mathbb{R}$, – это все возможные события, которые можно выразить в терминах η. Потому определение означает, что мы хотим, чтобы любое событие которое мы только можем выразить в терминах одной случайной величины было независимым с любым другим событием, которое можно выразить в терминах другой случайной величины.

Независимые индикаторы Пусть (Ω, P) – вероятностное пространство описывающее какой-то случайный черный ящик. Фиксируем два произвольных события $A, B \subseteq \Omega$ и рассмотрим две индикаторные случайные величины χ_A и χ_B. Напомню, что $\chi_A(\omega)$ равна 1, если $\omega \in A$, и 0 иначе. Тогда с помощью χ_A можно выразить четыре события: $\varnothing$, A, $\bar{A}$ и Ω. Действительно, $A = \chi_A^{-1}(1)$ и $\bar{A} = \chi_A^{-1}(0)$, $\varnothing = \chi_A^{-1}(\varnothing)$ и $\Omega = \chi_A^{-1}(\mathbb{R})$. Ничего другого выразить через нее нельзя. Аналогично и для χ_B. Для того, чтобы проверить их независимость, надо проверить, что любое событие из списка $\{\varnothing, A, \bar{A}, \Omega\}$ независимо с любым событием из списка $\{\varnothing, B, \bar{B}, \Omega\}$. Так как $\varnothing$ и Ω не зависят с любым событием, то остается проверить, что любое событие из $\{A, \bar{A}\}$ не зависит с любым событием из $\{B, \bar{B}\}$. В частности отсюда следует, что события A и B независимы. Наоборот, если события A и B независимы, то и события A и $\bar{B}$ независимы (я сейчас это покажу). А значит и все остальные пары событий независимы. То есть независимость индикаторных случайных величин χ_A и χ_B равносильная независимости самих событий A и B. Теперь давайте я проверю независимость A и $\bar{B}$, если A и B были независимыми. Действительно,

$$P(A\cap\bar{B}) = P(A\cap(\Omega\setminus B)) = P(A\setminus(A\cap B)) = P(A) - P(A\cap B) = P(A) - P(A)P(B) = P(A)(1-P(B)) = P(A)P(\bar{B})$$

Независимость в дискретном случае Пусть заданы две дискретные случайные величины ξ и η. Они будут независимы тогда и только тогда, когда

$$P(\xi = a, \eta = b) = P(\xi = a)P(\eta = b), \text{ для любых } a, b \in \mathbb{R}$$

Если же нам выдали две дискретные случайные величины

$$\xi \sim \begin{cases} a_1 & \dots & a_n & \dots \\ p_1 & \dots & p_n & \dots \end{cases} \quad \text{и} \quad \eta \sim \begin{cases} b_1 & \dots & b_n & \dots \\ q_1 & \dots & q_n & \dots \end{cases}$$

Тогда случайный вектор (ξ, η) имеет распределение

$$(\xi, \eta) \sim \begin{cases} (a_1, b_1) & \dots & (a_n, b_m) & \dots \\ p_1 q_1 & \dots & p_n q_m & \dots \end{cases}$$

Независимость в непрерывном случае Пусть заданы две непрерывные случайные величины ξ и η с совместной плотностью $p_{\xi,\eta}(x, y)$ и каждая имеет плотность $p_\xi(x)$ и $p_\eta(y)$ соответственно. Тогда эти случайные величины независимы тогда и только тогда, когда $p_{\xi,\eta}(x, y) = p_\xi(x)p_\eta(y)$. Это дает способ проверки на независимость и способ построения вектора из двух независимых случайных величин. Кратко правило можно резюмировать следующим образом. Если у вас есть две плотности независимых случайных величин, то для построения совместной плотности, их надо перемножить, подставляя в каждую плотностью свою собственную переменную.

Матожидание и дисперсия в случае независимых случайных величин Пусть ξ и η – две независимые случайные величины, тогда

1. $\mathbb{E}(\xi\eta) = \mathbb{E}(\xi)\mathbb{E}(\eta)$.
2. $\mathbb{D}(\xi + \eta) = \mathbb{D}(\xi) + \mathbb{D}(\eta)$.

Второе выводится из первого, пользуясь определением. А первое является предельной версией определения. Давайте дам один полезный пример. Если $\xi = \chi_A$ и $\eta = \chi_B$ – индикаторные случайные величины. Тогда $\mathbb{E}(\chi_A) = P(A)$, $\mathbb{E}(\chi_B) = P(B)$, $\mathbb{E}(\chi_A\chi_B) = \mathbb{E}(\chi_{A\cap B}) = P(A \cap B)$. Таким образом формула превращается в определение независимости $P(A \cap B) = P(A)P(B)$. Общая формула получается переходом от этой простейшей формулы к интегралу с помощью сумм и предельных переходов. Но это уже отдельная история для любителей тонкостей математики. Для нас более полезно понимать идею, почему это так работает, и уметь всем этим пользоваться.

Замечание о большом количестве случайных величин В случае, когда у нас есть несколько событий $A_1, \ldots, A_n \subseteq \Omega$, то они называются независимыми, если

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \ldots P(A_n)$$

Напомню, что события могут быть попарно независимыми, но при все вместе они зависимы, то есть это понятие не сводится к проверки независимости в каждой паре. Тогда на независимость случайных величин это определение распространяется так. Если есть случайные величины $\xi_1, \ldots, \xi_n \colon \Omega \to \mathbb{R}$, то для любых подмножеств $A_1, \ldots, A_n \subseteq \mathbb{R}$, мы хотим, чтобы события $\xi_1^{-1}(A_1), \ldots, \xi_n^{-1}(A_n) \subseteq \Omega$ были в совокупности независимы. Как и раньше, интуитивно это означает, что все события, которые только можно выразить в терминах наших случайных величин по отдельности не зависят в совокупности.

В дискретном случае, если $\xi_1, \ldots, \xi_n$ – дискретные случайные величины, то они независимы тогда и только тогда, когда

$$P(\xi_1 = a_1, \ldots, \xi_n = a_n) = P(\xi_1 = a_1) \ldots P(\xi_n = a_n), \text{ для любых } a_1, \ldots, a_n \in \mathbb{R}$$

В непрерывном случае, если $\xi_1, \ldots, \xi_n$ – непрерывные случайные величины, то они независимы тогда и только тогда, когда $p_{\bar{\xi}}(x_1, \ldots, x_n) = p_{\xi_1}(x_1) \ldots p_{\xi_n}(x_n)$.

С помощью функций распределения можно проверить на независимость любые наборы из случайных величин любого вида. Случайные величины $\xi_1, \ldots, \xi_n$ независимы тогда и только тогда, когда для всех $x_1, \ldots, x_n$

$$F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \ldots F_{\xi_n}(x_n)$$

напомню, что функция распределения – это $F_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n) = P(\xi_1 \leqslant x_1, \ldots, \xi_n \leqslant x_n)$.

Построение независимых случайных величин Пусть на случайном черном ящике (Ω, P) задан набор независимых в совокупности случайных величин $\xi_1, \ldots, \xi_n \colon \Omega \to \mathbb{R}$. Давайте разобьем их на две не пересекающиеся части $\xi_1, \ldots, \xi_m$ и $\xi_{m+1}, \ldots, \xi_n$. Теперь пусть $f \colon \mathbb{R}^m \to \mathbb{R}$ и $g \colon \mathbb{R}^{n-m} \to \mathbb{R}$ – две произвольные функции. Тогда случайные величины $f(\xi_1, \ldots, \xi_m)$ и $g(\xi_{m+1}, \ldots, \xi_n)$ будут независимыми. Аналогичный трюк работает, если мы разобьем исходное множество случайных величин $\xi_1, \ldots, \xi_n$ на произвольное число не пересекающихся подмножеств, элементы каждого подмножества подставим в свою функцию, тогда полученный набор случайных величин так же окажется независимым в совокупности.