Условное распределение

октябрь–декабрь 2023 г.

Курс лекций «Теория вероятностей»

Оглавление

- Условное распределение дискретных случайных величин
- Условное среднее дискретных случайных величин
- Условное распределение непрерывных случайных величин
- Условное среднее непрерывных случайных величин
- Независимость случайных величин в среднем

Условное распределение дискретных случайных величин: определение

Пусть задан двумерный дискретный случайный вектор $[X, Y]'$, и

$$p_{ij} = P(X = x_i, Y = y_j), \quad i = 1, \dots, m, \dots, \quad j = 1, \dots, n, \dots$$

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

— его распределение. Обозначим

$$P(X=x_i) = p_{i\bullet} = \sum_j p_{ij}, \quad p_{\bullet j} = \sum_i p_{ij} = P(Y=y_j)$$

маржинальные распределения компонентов X и Y соответственно. Тогда

$$p_{ij} = P(X=x_i, Y=y_j)$$

$$P(Y = y_j \,|\, X = x_i) = \frac{p_{ij}}{p_{i\bullet}} = p_{j|i}.$$

$$P(X=x_i)$$

Определение
Набор $\{p_{j|i}, j = 1, \dots, n, \dots\}$ называется условным распределением случайной величины Y при условии $X = x_i$.

Условное распределение дискретных случайных величин: связь с независимостью

Предложение
Дискретные случайные величины X и Y независимые тогда и только тогда, когда условные распределения $\{p_{j|i}, j = 1, \ldots, n, \ldots\}$ величины Y при условии $X = x_i$ не зависят от x_i (и совпадают с маржинальным распределением $p_{\bullet j}$).

Доказательство.
Действительно, если $p_{j|i} = p_j$ (не зависят от i), то $p_{ij} = p_{i\bullet} \cdot p_j$, откуда $p_{\bullet j} = \sum_i p_{ij} = \sum_i p_{i\bullet} p_j = p_j$ и $p_{ij} = p_{i\bullet} \cdot p_{\bullet j}$, что и означает независимость. □

Аналогичное предложение справедливо и для условных распределений $p_{i|j}$ величины X при условии $Y = y_j$.

Условное распределение дискретных случайных величин: пример

Рассматриваются семьи с числом детей не более двух. X — число детей в семье, Y — число девочек в семье. Таблица совместного распределения:

$p_{ij} = p_{i\cdot} p_{\cdot j}$

$P(X = x_i, Y = y_j)$ $P(X = x_i)$ $P(Y = y_j)$

$X \setminus Y$	0	1	2	
$\boxed{0}$	1/4	0	0	1/4
$\boxed{1}$	1/4	1/4	0	1/2
$\boxed{2}$	1/16	1/8	1/16	1/4
	9/16	3/8	1/16	

1. Найти маржинальные распределения X и Y.
2. Являются ли X и Y независимыми?
3. Найти условные распределения $Y \mid X = x_i$, $X \mid Y = y_j$.
4. Пусть Z — число мальчиков в семье. Являются ли Y и Z независимыми?

$X - Y$

Условное распределение дискретных случайных величин: разбор примера

Маржинальные распределения:

	0	1	2
X	1/4	1/2	1/4
Y	9/16	3/8	1/16

Зависимость X и Y: $P(X = 1, Y = 2) = 0$, $P(X = 1) \cdot P(Y = 2) = 1/2 \cdot 1/16 = 1/32 \neq 0$.

Условное распределение дискретных случайных величин: разбор примера II

Условные распределения $Y \mid X = x_i$. Делим на маржинальное распределение X, в каждой строке теперь отдельное распределение:

$X \setminus Y$	0	1	2	
0	1/4	0	0	1/4
1	1/4	1/4	0	1/2
2	1/16	1/8	1/16	1/4

$X \setminus Y$	0	1	2
0	1	0	0
1	1/2	1/2	0
2	1/4	1/2	1/4

$(1/2)^2 \quad 2 \cdot \frac{1}{2} \cdot \frac{1}{2} \quad \left(\frac{1}{2}\right)^2$

Как видим, условные распределения $Y \mid X = x_i$ разные в разных строках, откуда еще раз следует зависимость X и Y.

Условное распределение дискретных случайных величин: разбор примера III

Условные распределения $X \mid Y = y_i$. Делим на маржинальное распределение Y, в каждом столбце теперь отдельное распределение:

$X \setminus Y$	0	1	2
0	1/4	0	0
1	1/4	1/4	0
2	1/16	1/8	1/16
	9/16	3/8	1/16

$X \setminus Y$	0	1	2
0	4/9	0	0
1	4/9	2/3	0
2	1/9	1/3	1

Как видим, условные распределения $X \mid Y = y_j$ разные в разных столбцах, откуда еще раз следует зависимость X и Y.

Условное распределение дискретных случайных величин: разбор примера IV

Совместное распределение Y и $Z = X - Y$:

$X \setminus Y$	0	1	2
0	1/4	0	0
1	1/4	1/4	0
2	1/16	1/8	1/16

$Z \setminus Y$	0	1	2
0	1/4	1/4	1/16
1	1/4	1/8	0
2	1/16	0	0

1/16

1/16

Y и Z зависимые (только потому, что рассматриваются семьи не более чем с двумя детьми, так что чем больше девочек в семье, тем вероятность (условная) того, что меньше мальчиков, выше).

Условное матожидание дискретных случайных величин

Определение
Величина

$$E(Y \mid X = x_i) = \sum_j y_j p_{j|i}$$

называется условным математическим ожиданием Y при условии $X = x_i$.

Условное матожидание дискретных случайных величин: разбор примера V

Найдем условные матожидания $E(Y \mid X = x_i)$ и $E(X \mid Y = y_j)$. Условные распределения $Y \mid X = x_i$ (по строкам) и $X \mid Y = y_j$ (по столбцам)

$X \setminus Y$	0	1	2	$E(Y \mid X = x_i)$
0	1	0	0	0
1	1/2	1/2	0	1/2
2	1/4	1/2	1/4	1

$X \setminus Y$	0	1	2
0	4/9	0	0
1	4/9	2/3	0
2	1/9	1/3	1
$E(X \mid Y = y_j)$	2/3	4/3	2

Условное матожидание дискретных случайных величин как случайная величина

Определение
Случайная величина $E(Y \mid X)$, принимающая значения $E(Y \mid X = x_i)$ при $\omega \in \{X = x_i\}$ называется условным математическим ожиданием Y при условии X.

Заметим, что условное математическое ожидание $E(Y \mid X)$ является функцией условия, т. е. некоторой функцией $\varphi(X)$.

Условное матожидание дискретных случайных величин: разбор примера VI

Запишем условные матожидания $E(Y \mid X)$ и $E(X \mid Y)$ как случайные величины.

$$E(Y \mid X) = \begin{cases} 0 & X = 0, \\ 1/2 & X = 1, \\ 1 & X = 2 \end{cases} \implies E(Y \mid X) = X/2$$

$$E(X \mid Y) = \begin{cases} 2/3 & Y = 0, \\ 4/3 & Y = 1, \\ 2 & Y = 2 \end{cases} \implies E(X \mid Y) = \frac{2(1+Y)}{3}$$

Формулу, как в правой колонке, подобрать не всегда удается, но в виде перечисления условное матожидание задать можно всегда.

Условное матожидание дискретных случайных величин: пример

Пусть X, Y — независимые случайные величины, имеющие геометрическое распределение:

$$P(X = k) = P(Y = k) = q^{k-1}p, \quad q = 1 - p, \quad k = 1, 2, \dots$$

Найдем условное распределение

$$P(X = k \mid X + Y = n), \quad k = 1, 2, \dots, n - 1,$$

и условное среднее

$$E(X \mid X + Y = n).$$

Условное матожидание дискретных случайных величин: разбор примера

Найдем сначала вероятность

$$P(X+Y=n)=\sum_{i=1}^{n-1}P(X=i,Y=n-i)=\sum_{i=1}^{n-1}q^{i-1}p\cdot q^{n-i-1}p=(n-1)q^{n-2}p^2.$$

Тогда для $k=1,\ldots,n-1$

$$P(X=k\,|\,X+Y=n)=\frac{P(X=k,X+Y=n)}{P(X+Y=n)}=\frac{P(X=k,Y=n-k)}{P(X+Y=n)}=\frac{q^{k-1}p\cdot q^{n-k-1}p}{(n-1)q^{n-2}p^2}=\frac{1}{n-1},$$

т. е. получили равномерное распределение на точках $1,2,\ldots,n-1$. Поэтому $E(X\,|\,X+Y=n)=\frac{1}{n-1}\sum_{i=1}^{n-1}i=n/2$, и значит,

$$E(X\,|\,X+Y)=\frac{X+Y}{2}.$$

Условное распределение непрерывных случайных величин: определение

Пусть $U = [X', Y']'$ — случайный вектор, $X \in \mathbb{R}^m$, $Y \in \mathbb{R}^n$, $f(x, y)$ — совместная плотность распределения.

Определение
Определение. Функция

$$f(y \mid x) = \frac{f(x, y)}{\int_{\mathbb{R}^n} f(x, y) dy} = \frac{f(x, y)}{f_X(x)}$$

называется плотностью условного распределения величины Y при условии $X = x$.

В определении x фиксировано, $f(y \mid x)$ трактуется как функция y.
Она определяется для тех x, для которых $f_X(x) > 0$.

Условное распределение непрерывных случайных величин: содержательный смысл

Содержательный смысл:

$$P(Y \in dV_y \,|\, X \in dV_x) = \frac{P(X \in dV_x, Y \in dV_y)}{P(X \in dV_x)} \approx \frac{f(x,y)|dV_x||dV_y|}{f_X(x)|dV_x|} = \frac{f(x,y)|dV_y|}{f_X(x)},$$

что и оправдывает терминологию определения.

В дальнейшем для простоты будем рассматривать ситуацию $m = n = 1$, хотя результаты с очевидной поправкой на размерность верны и в многомерном случае.

Условное распределение непрерывных случайных величин: связь с независимостью

Предложение
Случайные величины X и Y, имеющие непрерывное совместное распределение, независимые тогда и только тогда, когда условные плотности $f(y \mid x)$ величины Y при условии $X = x$ не зависят от x (и совпадают с маргинальной плотностью $f_Y(y)$).

Доказательство.
Действительно, если $f(y \mid x) = f(y)$ (не зависят от x), то $f(x, y) = f_X(x)f(y)$, откуда $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx = \int_{-\infty}^{+\infty} f_X(x)f(y)dx = f(y)$ и $f(x, y) = f_X(x)f_Y(y)$, что и означает независимость. $\square$

Аналогичное предложение справедливо и для условных распределений $f(x \mid y)$ величины X при условии $Y = y$.

Условное распределение непрерывных случайных величин: пример

На отрезке $[0, 1]$ случайным образом выбирается точка X, затем на отрезке $[0, X]$ случайным образом выбирается точка Y. Найти $\mathrm{Cov}(X, Y)$.

Фраза «случайным образом выбирается точка ...» означает, что соответствующая случайная величина имеет равномерное распределение на соответствующем отрезке.

Условное распределение непрерывных случайных величин: разбор примера

Плотность распределения X есть

$$f_X(x) = \begin{cases} 1, & x \in [0, 1], \\ 0, & x \notin [0, 1]. \end{cases}$$

Если $x \in [0, 1]$, то условная плотность величины Y при условии $X = x$, есть

$$f_{Y|X}(y \mid x) = \begin{cases} 1/x, & y \in [0, x], \\ 0, & x \notin [0, x]. \end{cases}$$

Условное распределение непрерывных случайных величин: разбор примера II

Обозначим через D треугольник $\{0 \leqslant y \leqslant x \leqslant 1\}$. Тогда совместная плотность $f(x, y)$ есть

$$f(x, y) = f_X(x) f_{Y|X}(y \mid x) = \begin{cases} 1/x, & (x, y) \in D, \\ 0, & (x, y) \notin D. \end{cases}$$

Поэтому

$$E(XY) = \iint_{\mathbb{R}^2} xyf(x, y)dxdy = \iint_D ydxdy = \int_0^1 \left(\int_0^x ydy \right) dx = \int_0^1 x^2/2dx = x^3/6 \Big|_0^1 = \frac{1}{6}.$$

Плотность маржинального распределения Y есть

$$f_Y(y) = \int_{\mathbb{R}} f(x, y)dx = \int_y^1 \frac{dx}{x} = -\ln y, \ 0 < y \leqslant 1,$$

поэтому $E(Y) = -\int_0^1 y \ln y dy = 1/4$. Кроме того $E(X) = 1/2$. Поэтому $\mathrm{Cov}(X, Y) = \frac{1}{6} - \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{24}$.

Условное матожидание непрерывных случайных величин: определение

Определение
Если $[X', Y']'$ — непрерывный случайный вектор, то функция

$$\varphi(x) = E(Y \mid X = x) = \int_{\mathbb{R}^n} y f_{Y|X}(y \mid x) dy$$

называется условным математическим Y при условии $X = x$.

Если случайная величина X вырожденная (т. е. $P(X = x) = 1$), то по определению $E(Y \mid X = x) = E(Y)$.

Определение
Функция $\varphi(x)$, $x \in \mathbb{R}$ называется функцией регрессии Y на X.

Определение
Случайная величина $\varphi(X)$ называется условным математическим ожиданием Y при условии X, обозначается $E(Y \mid X)$.

Условное матожидание непрерывных случайных величин: связь с независимостью

Если X и Y независимы, то условное распределение совпадает с безусловным и, значит, условное математическое ожидание совпадает с безусловным.

Действительно, пусть, например, вектор $[X, Y]'$ непрерывный и компоненты X, Y непрерывны. Тогда $f(x, y) = f_X(x) f_Y(y)$, $f_{Y|X}(y \mid x) = f_Y(x)$ и

$$\varphi(x) = E(Y \mid X = x) = \int_{\mathbb{R}} y f_{Y|X}(y \mid x) dy = \int_{\mathbb{R}} y f_Y(y) dy = E(Y).$$

Свойства условного математического ожидания

Пусть X, Y — случайные величины.

1. Линейность: $E(aY_1 + bY_2 \mid X) = aE(Y_1 \mid X) + bE(Y_2 \mid X)$ для любых $a, b \in \mathbb{R}$
2. Правило повторного условного математического ожидания (телескопическое свойство, chain rule): если $g : \mathbb{R} \to \mathbb{R}$ — (неслучайная) функция, то

   $$E(E(Y \mid X) \mid g(X)) = E(Y \mid g(X)).$$

   В частности, $E(E(Y \mid X)) = E(Y)$.
3. Пусть $g : \mathbb{R} \to \mathbb{R}$ — (неслучайная) функция. Тогда

   $$E(g(X)Y \mid X) = g(X)E(Y \mid X).$$

4. Если X и Y независимы, то $E(Y \mid X) = E(Y)$.

Свойства условного математического ожидания II

5. Оптимизационное свойство условного математического ожидания

Предложение
Пусть Y — скалярная случайная величина и $f(\cdot)$ — числовая измеримая функция. Тогда минимум среднеквадратичного отклонения $E(Y-f(X))^2$ достигается при $f(\cdot)=\varphi(\cdot)$, т. е. функция $\varphi(\cdot)$ является решением задачи

$$E(Y-f(X))^2 \to \min_{f(\cdot)}.$$

Свойства условного математического ожидания III

Доказательство.
Докажем, что для любой функции $f(\cdot)$ выполнено неравенство

$$E((Y - f(X))^2 \mid X) \geqslant E((Y - \varphi(X))^2 \mid X). \tag{1}$$

Имеем

$$\begin{aligned} E((Y - f(X))^2 \mid X) &= E((Y - \varphi(X) + \varphi(X) - f(X))^2 \mid X) \\ &= E((Y - \varphi(X))^2 \mid X) + E((\varphi(X) - f(X))^2 \mid X) + 2E((Y - \varphi(X))(\varphi(X) - f(X)) \mid X) \\ &= E((Y - \varphi(X))^2 \mid X) + (\varphi(X) - f(X))^2 + 2(\varphi(X) - f(X))E(Y - \varphi(X) \mid X). \end{aligned}$$

Последнее слагаемое равно 0:

$$E(Y - \varphi(X) \mid X) = E(Y \mid X) - E(\varphi(X) \mid X) = \varphi(X) - \varphi(X) = 0.$$

Второе слагаемое неотрицательное, и неравенство доказано. Остается взять математическое ожидание обеих частей неравенства (1). $\square$

Свойства условного математического ожидания III

6. Пусть $g\colon \mathbb{R} \to \mathbb{R}$ — (неслучайная) числовая функция. Тогда $\mathrm{Cov}(g(X), Y - E(Y \mid X)) = 0$.
   Действительно, так как $E(Y - E(Y \mid X)) = 0$, то
   $$\begin{aligned}\mathrm{Cov}(g(X), Y - E(Y \mid X)) &= E(g(X)(Y - E(Y \mid X))) \\ &= E(E(g(X)(Y - E(Y \mid X)) \mid X)) = E(g(X)E((Y - E(Y \mid X)) \mid X)) = 0.\end{aligned}$$

В приложениях случайную величину X (или ее обобщение) трактуют как имеющуюся в наличии информацию. Тогда условное ожидание $E(Y \mid X)$ можно интерпретировать, как максимально доступную «информацию» о случайной величине Y. Оправданием этому служат свойства 5 и 6:

- Минимум среднеквадратичной ошибки
- То, что осталось в Y, а именно $Y - E(Y \mid X)$, уже не коррелирует с X

Условное математическое ожидание: пример

Пусть совместная плотность случайных величин X и Y есть

$$f(x, y) = \begin{cases} \dfrac{x + y}{3}, & \text{если } 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 2, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Найти $E(Y \mid X)$, $E(X \mid Y)$.

Условное математическое ожидание: разбор примера

Маржинальные плотности $f_X(x)$, $f_Y(y)$ равны соответственно

$$f_X(x) = \int_{\mathbb{R}} f(x,y)dy = \int_0^2 \frac{x+y}{3}dy = \frac{xy + y^2/2}{3}\Big|_0^2 = \frac{2(x+1)}{3}, \quad x \in [0,1],$$

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)dx = \int_0^1 \frac{x+y}{3}dx = \frac{x^2/2 + yx}{3}\Big|_0^1 = \frac{2y+1}{6}, \quad y \in [0,2].$$

Условные плотности $f_{Y|X}(y \mid x)$, $f_{X|Y}(x \mid y)$ равны соответственно

$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)} = \frac{x+y}{2(x+1)}, \quad y \in [0,2], \text{определена при } x \in [0,1],$$

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} = \frac{2(x+y)}{2y+1}, \quad x \in [0,1], \text{определена при } y \in [0,2].$$

Условное математическое ожидание: разбор примера II

Условные матожидания равны

$$E(Y \mid X = x) = \int_0^2 y \frac{x+y}{2(x+1)} dy = \left(\frac{xy^2/2 + y^3/3}{2(x+1)} \right) \Big|_0^2 = \frac{2x + 8/3}{2(x+1)}, \text{опр. при } x \in [0,1],$$

$$E(X \mid Y = y) = \int_0^1 x \frac{2(x+y)}{2y+1} dx = \left(\frac{2(x^3/3 + yx^2/2)}{2y+1} \right) \Big|_0^1 = \frac{2/3 + y}{2y+1}, \text{опр. при } y \in [0,2].$$

Соответственно,

$$E(Y \mid X) = \frac{6X + 8}{6(X+1)}, \quad E(X \mid Y) = \frac{2 + 3Y}{6Y + 3}$$

Независимость в среднем

Определение
Говорят, что случайная величина Y не зависит от X в среднем, если условное матожидание $E(Y \mid X)$ не зависит от X.

В отличие от обычной независимости и некоррелированности свойство независимости в среднем несимметричное. Так, в примере $X \sim N(0, 1)$, $Y = X^2$ выполняется равенство $E(Y \mid X) = X^2$, т. е. Y зависит от X в среднем, но $E(X \mid Y) \equiv 0$ (при фиксированном $Y = y$ величина X принимает два значения: $-\sqrt{y}$ и $\sqrt{y}$ с равными вероятностями), т. е. X не зависит от Y в среднем.

Независимость в среднем часто используется при выводе свойств оценок коэффициентов в регрессионных моделях.

Независимость в среднем, независимость и некоррелированность

Предложение
Пусть X, Y — случайные величины. Тогда

1. *из независимости X и Y следует независимость в среднем (как Y от X, так и X от Y)*
2. *из независимости в среднем Y от X (или X от Y) следует некоррелированность X и Y*

Доказательство.

1. Это свойство 4 условного матожидания.
2. Пусть $E(Y \mid X)$ не зависит от X (и значит $E(Y \mid X) = E(Y)$). Тогда $E(XY) = E(E(XY \mid X)) = E(XE(Y \mid X)) = E(XE(Y)) = E(X)E(Y)$, откуда и получается, что $\mathrm{Cov}(X, Y) = 0$.

□

Определение
Величина $E((Y - E(Y \mid X))^2 \mid X)$ называется условной дисперсией. Обозначается $\mathrm{Var}(Y \mid X)$.

Аналогично обычной дисперсии выполняется равенство
$\mathrm{Var}(Y \mid X) = E(Y^2 \mid X) - (E(Y \mid X))^2$.

Предложение (формула условной дисперсии)
Выполняется равенство $\mathrm{Var}(Y) = E(\mathrm{Var}(Y \mid X)) + \mathrm{Var}(E(Y \mid X))$.

Доказательство.

$$E(\mathrm{Var}(Y \mid X)) + \mathrm{Var}(E(Y \mid X)) = E\big(E(Y^2 \mid X) - (E(Y \mid X))^2\big) + E((E(Y \mid X))^2) - (E(E(Y \mid X)))^2$$
$$= E(E(Y^2 \mid X)) - E((E(Y \mid X))^2) + E(E(Y \mid X)^2) - (E(E(Y \mid X)))^2 = E(Y^2) - (E(Y))^2 = \mathrm{Var}(Y)$$

$\square$

Литература

[1] Sh. Ross. *A First Course in Probability*. 9th ed. London: Pearson, 2014.

[2] Б. В. Гнеденко. *Курс теории вероятностей*. М.: Издательство УРСС, 2005.

[3] С. В. Головань, П. К. Катышев, А. А. Пересецкий. *Сборник задач по курсу теории вероятностей*. М.: Издательский дом «Дело» РАНХиГС, 2021.

Приложение: геометрическая интерпретация условного математического ожидания i

Условное математическое ожидание допускает следующую геометрическую интерпретацию.

Пусть $(\Omega, \mathcal{F}, P)$ — некоторое вероятностное пространство. Обозначим через L^2 множество всех случайных величин $\{\xi\}$ на этом пространстве, таких что $E(\xi^2) < +\infty$, т. е. имеющих вторые моменты. Нетрудно проверить, что относительно обычного сложения случайных величин и умножения их на число это множество является линейным пространством. Отметим, что если пространство элементарных исходов Ω бесконечно, то пространство L^2 является бесконечномерным линейным пространством.

Приложение: геометрическая интерпретация условного математического ожидания ii

Введем в этом пространстве скалярное произведение:

$$(\xi, \eta) = E(\xi \cdot \eta) \tag{2}$$

Можно проверить, что (2) — скалярное произведение, т. е. симметричная билинейная положительно определенная функция (с точностью до равенства 0 почти наверное). Таким образом, пространство L^2 становится евклидовым (или гильбертовым) пространством.

Как известно, в любом евклидовом пространстве E существует операция ортогонального проецирования на (замкнутое) подпространство. А именно,

Приложение: геометрическая интерпретация условного математического ожидания iii

пусть G — (замкнутое) линейное подпространство, и пусть x — произвольный вектор из E. Тогда существуют вектор $y \in G$ и вектор $z \in G^{\perp}$, такие что

$$x = y + z, \tag{3}$$

и разложение (3) единственно. Запись $z \in G^{\perp}$ означает, что вектор z ортогонален любому вектору из подпространства G. Вектор y в разложении (3) называется ортогональной проекцией вектора x на подпространство G. При этом длина вектора z представляет наименьшее

Приложение: геометрическая интерпретация условного математического ожидания iv

расстояние от вектора x до любого вектора в подпространстве G. Иными словами, вектор z есть решение следующей задачи:

$$(x - w, x - w) \to \min_{w \in G}.$$

Пусть $X, Y \in L^2$ — случайные величины. Обозначим через $L(X)$ множество случайных величин вида $g(X)$, где $g(\cdot)$ — произвольная измеримая функция. Нетрудно проверить, что $L(X)$ — (замкнутое) линейное подпространство в L^2. Тогда свойства 5 и 6 позволяют интерпретировать условное математическое ожидание $E(Y \mid X)$ как ортогональную проекцию вектора Y на подпространство $L(X)$. В этой интерпретации телескопическое свойство

Приложение: геометрическая интерпретация условного математического ожидания v

условного матожидания — это просто теорема о трех перпендикулярах, а оптимизационное свойство есть следствие того факта, что ортогональная проекция реализует минимальное расстояние от вектора до подпространства.

Для заметок