Теория вероятностей. Лекция пятнадцатая
Условная плотность и совместное распределение

Дмитрий Валерьевич Хлопин
glukanat@mail.ru

Институт математики и механики им. Н.Н.Красовского

18.12.2018

Переход от дискретного случая к общему

- борелевские множества и мера Лебега
- случайные величины и измеримые отображения
- функции распределения
- абсолютно случайные величины и не только
- математическое ожидание как интеграл
- условное математическое ожидание как интеграл
- совместные функции распределения и условное матожидание

Мы определили условное матожидание как интеграл, привели простые примеры. Однако до сих пор все условные матожидания считались лишь для случая, когда число гипотез не более чем счетно...
Что делать в континуальном случае?

Простейшие примеры, которые уже были

Пусть дана случайная величина ξ, $\mathbb{E}|\xi| < +\infty$. Тогда, для полной группы $H_1,\ldots,H_n,\ldots$ и $\mathcal{H} \triangleq \sigma(H_1,\ldots,H_n,\ldots)$ имеем $\mathcal{H}$-измеримую функцию

$$\mathbb{E}(\xi|\mathcal{H})|_{H_i} = \frac{\int_{H_i}\xi(\omega)\,\mathbb{P}(d\omega)}{\mathbb{P}(H_i)}.$$

В частности, для всякого множества $B \in \mathcal{F}$, при $\xi \triangleq 1_B$ имеем

$$\mathbb{E}(1_B|\mathcal{H})|_{H_i} = \mathbb{P}(B\,|\mathcal{H})|_{H_i} = \mathbb{P}(B\,|\,H_i).$$

Для не более чем счетного набора $B_i \in \mathcal{F}$, для $\xi \triangleq \sum_{k\in\mathbb{N}} a_k 1_{B_k}$ имеем

$$\mathbb{E}\Big(\sum_{k\in\mathbb{N}} a_k 1_{B_k}\Big|\mathcal{H}\Big) = \sum_{k\in\mathbb{N}} a_k \mathbb{P}(B_k\,|\,\mathcal{H}).$$

Собственно определение условного матожидания

Пусть даны случайная величина ξ и σ-алгебра $\mathcal{H}$.

Случайную величину $\zeta : \Omega \to \mathbb{R}$ будем называть (и обозначать $\mathbb{E}(\xi|\mathcal{H})$) условным математическим ожиданием действительной случайной величины ξ относительно $\mathcal{H}$, если она $\mathcal{H}$-измерима и удовлетворяет

$$\int_H \xi(\omega)\,\mathbb{P}(d\omega) = \int_H \zeta(\omega)\,\mathbb{P}(d\omega) \qquad \forall H \in \mathcal{H}.$$

Условное матожидание индикаторной функции 1_A (для $A \in \mathcal{F}$) будем также называть условной вероятностью и обозначать как $\mathbb{P}(A|\mathcal{H})$.

Из теоремы Радона–Никодима получены существование и единственность условного матожидания, если $\mathbb{E}|\xi|$ конечно.

Эквивалентное определение условного матожидания

Пусть даны случайная величина ξ и σ-алгебра $\mathcal{H}$.

Случайную величину $\zeta : \Omega \to \mathbb{R}$ будем называть (и обозначать $\mathbb{E}(\xi|\mathcal{H})$) условным математическим ожиданием действительной случайной величины ξ относительно $\mathcal{H}$, если она $\mathcal{H}$-измерима, и для всех $\mathcal{H}$-измеримых случайных величин η из $\mathbb{E}|\eta\xi| < +\infty$ следует $\mathbb{E}(\eta\xi) = \mathbb{E}(\eta\zeta)$, то есть

$$\int_\Omega \eta(\omega)\xi(\omega)\,\mathbb{P}(d\omega) = \int_\Omega \eta(\omega)\zeta(\omega)\,\mathbb{P}(d\omega).$$

Эквивалентность следует из прошлой лекции, фактически формула выше для η вида 1_H постулируется, значит выполнена для сумм, а значит и для их монотонных пределов.

Условное матожидание относительно случайной величины

Пусть ξ, η — случайные величины, и $\mathbb{E}|\xi| < +\infty$.

Случайную величину ζ будем называть (и обозначать $\mathbb{E}(\xi|\eta)$) условным математическим ожиданием ξ относительно η, если она $\sigma(\eta)$-измерима, и

$$\int_H \zeta(\omega)\,\mathbb{P}(d\omega) = \int_H \xi(\omega)\,\mathbb{P}(d\omega) \qquad \forall H \in \sigma(\eta).$$

Напомним, что случайная величина ζ измерима относительно σ-алгебры $\sigma(\eta)$ тогда и только тогда, когда для некоторой борелевской g почти всюду выполнено $\zeta(\omega) = g(\eta(\omega))$.

Условное матожидание относительно случайной величины. Эквивалентное определение

Пусть ξ, η — случайные величины, и $\mathbb{E}|\xi| < +\infty$.

Случайную величину $\zeta = h(\eta)$ будем называть (и обозначать $\mathbb{E}(\xi|\eta)$) условным математическим ожиданием ξ относительно η, если для некоторой борелевской функции h для всех борелевских g из $\mathbb{E}(g(\eta)\xi)$ следует $\mathbb{E}(g(\eta)\xi) = \mathbb{E}(g(\eta)h(\eta))$, то есть

$$\begin{aligned}\int_\Omega g(\eta(\omega))\xi(\omega)\,\mathbb{P}(d\omega) &= \int_\Omega g(\eta(\omega))h(\eta(\omega))\,\mathbb{P}(d\omega)\\ &= \int_\mathbb{R} g(y)h(y)(\eta\#\mathbb{P})(dy).\end{aligned}$$

Те же примеры в терминах функций

Пусть $\eta \overset{\triangle}{=} \sum_{i\in\mathbb{N}} c_i 1_{H_i}$, где c_i различны, а $H_i \cap H_j = \varnothing$. Для всякой случайной величины ξ ($\mathbb{E}|\xi| < +\infty$) имеем $\sigma(\eta)$-измеримую функцию

$$\mathbb{E}(\xi|\eta)|_{\eta=c_i} = \frac{\int_{H_i} g(\omega)\,\mathbb{P}(d\omega)}{\mathbb{P}(H_i)}.$$

В частности, для всякого множества $B \in \mathcal{F}$, для $\xi \overset{\triangle}{=} 1_B$ имеем

$$\mathbb{E}(1_\xi\,|\eta)|_{\eta=c_i} = \mathbb{P}(B\,|\,\eta = c_i) = \frac{\mathbb{P}(\omega \in B, \eta = c_i)}{\sum_{j\in\mathbb{N}} \mathbb{P}(\omega \in B, \eta = c_j)}.$$

Для не более чем счетного набора $B_i \in \mathcal{F}$, для $\xi \overset{\triangle}{=} \sum_{k\in\mathbb{N}} a_k 1_{B_k}$ имеем

$$\mathbb{E}\Big(\sum_{k\in\mathbb{N}} a_k 1_{B_k}\Big|\eta\Big)\Big|_{\eta=c_i} = \sum_{k\in\mathbb{N}} a_k \mathbb{P}(B_k\,|\,\eta = c_i) = \sum_{k\in\mathbb{N}} a_k \frac{\mathbb{P}(\xi = a_k, \eta = c_i)}{\sum_{j\in\mathbb{N}} \mathbb{P}(\xi = a_k, \eta = c_j)}.$$

Количество гипотез H_i всё ещё счётно!!

Совместная функция распределения

Совместной функцией распределения векторной случайной величины $\vec{\xi} = (\xi_1, \ldots, \xi_m)$ называют отображение $F_{\vec{\xi}} : \mathbb{R}^m \to [0,1]$, заданное по правилу:

$$F_{\vec{\xi}}(\vec{a}) \triangleq \mathbb{P}\{\omega \,|\, \forall k \in \{1, \ldots, m\}\ \xi_k(\omega) \le a_k\} \quad \forall \vec{a} = (a_1, \ldots, a_m) \in \mathbb{R}^m.$$

Предупреждение: как уже говорилось, иногда в определении функции распределения пишут строгое неравенство. При этом часть свойств, по большому счету аналогичных, меняется, причем для многомерных распределений отследить эти изменения существенно сложнее...

Пример. Найти вероятность $\mathbb{P}\{\omega \,|\, \xi_1(\omega) \in (a', a''], \xi_2(\omega) \in (b', b'']\}$.

$$F_{(\xi_1,\xi_2)}(a'', b'') - F_{(\xi_1,\xi_2)}(a'', b') - F_{(\xi_1,\xi_2)}(a', b'') + F_{(\xi_1,\xi_2)}(a', b').$$

Свойства совместной функции распределения [с-но]

Для любой векторной случайной величины $\vec{\xi} = (\xi_1, \ldots, \xi_m)$, в любой точке $\vec{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$, по каждой переменной ξ_k $(k \in \{1, 2, \ldots, m\})$:

1. $F_{\vec{\xi}}$ не убывает;
2. $\lim\limits_{x_k \to -\infty} F_{\vec{\xi}}(x_1, x_2, \ldots, x_m) = 0$;
3. $\lim\limits_{x_1, x_2, \ldots, x_m \to +\infty} F_{\vec{\xi}}(x_1, x_2, \ldots, x_m) = 1$;
4. существуют односторонние пределы $F_{\vec{\xi}}(x_1, \ldots, x_{k-1}, x_k + 0, x_{k+1}, \ldots, x_m)$ и $F_{\vec{\xi}}(x_1, \ldots, x_{k-1}, x_k - 0, x_{k+1}, \ldots, x_m)$;
5. функция $F_{\vec{\xi}}$ непрерывна справа, то есть $F_{\vec{\xi}}(x_1, \ldots, x_{k-1}, x_k + 0, x_{k+1}, \ldots, x_m) = F_{\vec{\xi}}(x_1, \ldots, x_m)$.

Доказательства такие же, как и для обычной функции распределения. Но полезно разобраться, почему в этих пределах сходимость x_i прописана по-разному. Для этого полезно еще одно свойство...

Функции маргинальных распределений [с-но]

Для всякой векторной случайной величины $\vec{\xi} = (\xi_1, \ldots, \xi_m)$ функции маргинальных распределений можно написать следующим образом: для всех $y_1, \ldots, y_m \in \mathbb{R}$

$$\begin{aligned}
F_{\xi_k}(y_k) &= \lim_{x_i \to +\infty (\forall i \neq k)} F_{\vec{\xi}}(x_1, x_2, \ldots, x_{k-1}, y_k, x_{k+1}, \ldots, x_m), \\
F_{(\xi_1, \xi_2)}(y_1, y_2) &= \lim_{x_i \to +\infty (\forall i > 2)} F_{\vec{\xi}}(y_1, y_2, x_3, \ldots, x_m), \\
\ldots
\end{aligned}$$

Подумать: докажите, что ξ_1, ξ_2 независимы тогда и только тогда, когда $F_{(\xi_1, \xi_2)}(y_1, y_2) = F_{\xi_1}(y_1) F_{\xi_2}(y_2)$ для всех $y_1, y_2 \in \mathbb{R}$.

Абсолютно непрерывные векторные случайные величины

Назовем векторную случайную величину $\vec{\xi} = (\xi_1, \ldots, \xi_m)$ абсолютно непрерывной, если для всех $(x_1, \ldots, x_m) \in \mathbb{R}^m$ выполнено

$$F_{\vec{\xi}}(x_1, \ldots, x_m) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \ldots \int_{-\infty}^{x_m} f_{\vec{\xi}}(y_1, \ldots, y_m) dy_m \, dy_{m-1} \ldots dy_1$$

для некоторой борелевской почти всюду неотрицательной функции $f_{\vec{\xi}} : \mathbb{R}^m \to \mathbb{R}$, называемой ее плотностью.

Подумать: если $f_{\vec{\xi}}$ достаточно гладкая ($m+1$ раз непрерывно дифференцируема), то $f_{\vec{\xi}} = \frac{\partial^m F_{\vec{\xi}}}{\partial x_1 \ldots \partial x_m}$.

Подумать: а если порядок интегрирования будет другой?

Подумать: почему каждый следующий внутренний интеграл законен (предыдущий измерим или интегрируем)?

Подумать: почему вся функция заведомо интегрируема?

В качестве ответов...

Порядок интегрирования можно менять

Пусть заданы измеримые пространства $(\Omega_1, \mathcal{F}_1), \ldots, (\Omega_n, \mathcal{F}_n)$ с мерами μ_i (вероятностными или лебеговскими мерами).
По теореме Каратеодори однозначно восстанавливается мера $\mu = \mu_1 \otimes \ldots \otimes \mu_n$ на $\mathcal{F} = \mathcal{F}_1 \otimes \ldots \otimes \mathcal{F}_n = \sigma(\{A_1 \times \ldots \times A_n \,|\, A_i \in \mathcal{F}_i \,\forall i\})$ (над $\Omega = \Omega_1 \times \ldots \times \Omega_n$) правилом $\mu(A_1 \times \ldots \times A_n) = \mu_1(A_1) \cdot \ldots \cdot \mu_n(A_n)$.

Теорема Фубини [без д-ва]. Пусть даны борелевская функция $f : \mathbb{R}^{k+l} \to \mathbb{R}$ и отображения $\xi_1 : \Omega_1 \to \mathbb{R}^k, \xi_2 : \Omega_2 \to \mathbb{R}^l$.
Тогда $f(\xi_1, \xi_2)$ μ-интегрируема в том и только в том случае, когда отображения $x \mapsto \int_{\Omega_2} f(x,y)\,\mu_2(dy)$, $y \mapsto \int_{\Omega_1} f(x,y)\,\mu_1(dx)$ μ_1- и μ_2-интегрируемы при μ_2-почти всех y и μ_1-почти всех x.
При этом,

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f(x,y)\mu(d(x,y)) &= \int_{\Omega_1} \int_{\Omega_2} f(x,y)\,\mu_2(dy)\,\mu_1(dx) \\ &= \int_{\Omega_2} \int_{\Omega_1} f(x,y)\,\mu_1(dx)\,\mu_2(dy).\end{aligned}$$

Эквивалентное в силу теоремы Фубини определение

1^o Порядок интегрирования можно менять: для всех $a_1,\ldots,a_m\in\mathbb{R}$

$$\int_{\prod_{i=1}^m(-\infty,a_i]} f_{\vec{\xi}}(\vec{y})\,\lambda(d\vec{y}) = \int_{-\infty}^{a_{s(1)}}\cdots\int_{-\infty}^{a_{s(m)}} f_{\vec{\xi}}(y_1,\ldots,y_m)dy_{s(m)}\ldots dy_{s(1)}$$

для любой перестановки $s:\{1,2,\ldots,m\}\to\{1,2,\ldots,m\}$.

Назовем векторную случайную величину $\vec{\xi}=(\xi_1,\ldots,\xi_m)$ абсолютно непрерывной, если для всех $(x_1,\ldots,x_m)\in\mathbb{R}^m$ выполнено

$$F_{\vec{\xi}}(x_1,\ldots,x_m) = \int_{(-\infty,x_1]\times\cdots\times(-\infty,x_m]} f_{\vec{\xi}}(\vec{y})\,\lambda(d\vec{y})$$

для некоторой борелевской почти всюду неотрицательной функции $f_{\vec{\xi}}:\mathbb{R}^m\to\mathbb{R}$, называемой ее плотностью.

Свойства плотности

Перейдем в 1^o от индикаторов и их сумм к измеримым функциям...
2^o Для любой g такой, что $\mathbb{E}g(\xi)$ конечно, и любой перестановки s

$$\begin{aligned}\mathbb{E}g(\xi) &\triangleq \int_{\mathbb{R}^m} g(\vec{y}) f_{\vec{\xi}}(\vec{y})\,\lambda(d\vec{y}) \\ &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} g(\vec{y}) f_{\vec{\xi}}(y_1, \ldots, y_m) dy_{s(m)} \ldots dy_{s(2)}\, dy_{s(1)};\end{aligned}$$

в частности, $\mathbb{P}\{\omega \,|\, \vec{\xi}(\omega) \in A\} = \int_A f_{\vec{\xi}}(\vec{x})\,\lambda(d\vec{x})$ при всех $A \in \mathcal{B}$ и

$$\mathbb{E}\vec{\xi} \triangleq \int_{\mathbb{R}^m} \vec{x} f_{\vec{\xi}}(\vec{x})\,\lambda(d\vec{x}).$$

3^o Из теоремы Радона-Никодима плотность $f_{\vec{\xi}}$ λ-почти всюду неотрицательна и единственна; а среди непрерывных функций, если она есть, то ровно одна.

Итак, как и в скалярном случае...

Следствие. Для $\vec{\xi}$ следующие условия эквивалентны:

1. $\vec{\xi}\#\mathbb{P}$ абсолютно непрерывна относительно меры λ;

1'. $\vec{\xi}$ абсолютно непрерывна с плотностью $f_{\vec{\xi}}:\mathbb{R}^m\to\mathbb{R}$;

2. для некоторой λ-суммируемой $f_{\vec{\xi}}:\mathbb{R}^m\to\mathbb{R}$ при всех $A\in\mathcal{B}$ $(\vec{\xi}\#\mathbb{P})(A)=\int_A f_{\vec{\xi}}(\vec{x})\,\lambda(d\vec{x})$;

2'. для некоторой λ-суммируемой $f_{\vec{\xi}}:\mathbb{R}^m\to\mathbb{R}$ при всех $(y_1,\ldots,y_m)\in\mathbb{R}^m$ выполнено
$$F_{\vec{\xi}}(y_1,\ldots,y_m)=\int_{-\infty}^{y_1}\cdots\int_{-\infty}^{y_m}f_{\xi}(x_1,\ldots,x_m)\,dx_m\ldots dx_1;$$

3. для некоторой λ-суммируемой $f_{\vec{\xi}}:\mathbb{R}^m\to\mathbb{R}$ при всех $h:\mathbb{R}^m\to\mathbb{R}$ таких, что $\mathbb{E}h(\vec{\xi})$ конечно, выполнено
$$\mathbb{E}h(\vec{\xi})=\int_{\mathbb{R}^m}h(\vec{x})f_{\vec{\xi}}(\vec{x})\,\lambda(d\vec{x}).$$

При этом, любые так полученные функции $f_{\vec{\xi}}$ почти всюду совпадают.

Плотности маргинальных распределений

4^o Все маргинальные распределения абсолютно непрерывной векторной случайной величины абсолютно непрерывны, при этом

$$F_{\xi_1}(x) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty} f_{\vec{\xi}}(y_1, y_2 \dots, y_m)dy_m \dots dy_2 dy_1,$$

$$f_{\xi_1}(x) = \int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty} f_{\vec{\xi}}(x, y_2 \dots, y_m)dy_m \dots dy_2.$$

$$F_{\xi_2,\xi_3,\dots,\xi_m}(x_2,\dots,x_m) = \int_{-\infty}^{x_2}\dots\int_{-\infty}^{x_m}\int_{-\infty}^{+\infty} f_{\vec{\xi}}(y, y_2 \dots, y_m)dy dy_m \dots dy_2,$$

$$f_{\xi_2,\xi_3,\dots,\xi_m}(x_2,\dots,x_m) = \int_{-\infty}^{+\infty} f_{\vec{\xi}}(y, x_2, \dots, x_m)dy.$$

Подумать: верно ли обратное, если все маргинальные распределения абсолютно непрерывны, будет ли таким их совместное распределение?

Подумать: абсолютно непрерывны ли относительно друг друга меры $\vec{\xi}\#\mathbb{P}$, $(\xi_2,\dots,\xi_m)\#\mathbb{P}$ и $\xi_1\#\mathbb{P}$? На каких σ-алгебрах они заданы?

Плотность независимых случайных величин

Теорема. Пусть $f_{(\xi,\eta)}$ существует. Тогда ξ, η независимы тогда и только тогда, когда $f_{(\xi,\eta)}(x,y)$ равно $f_\xi(x) f_\eta(y)$ для почти всех (x,y).

Доказательство. ξ, η независимы тогда и только тогда, когда

$$\mathbb{P}((\xi,\eta)(\omega) \in A \times B) = \int_{A\times B} f_{(\xi,\eta)}(x,y)\, \lambda(d(x,y)), \qquad (*)$$

$$\mathbb{P}(\xi(\omega) \in A)\mathbb{P}(\eta(\omega) \in B) = \int_{A\times B} f_\xi(x) f_\eta(y)\, \lambda(d(x,y)) \qquad (**)$$

равны для всех $A, B \in \mathcal{B}$.

Теперь если ξ_1, ξ_2 независимы, то из $(*) = (**)$ для всевозможных промежутков вида $(-\infty, z]$ в качестве A, B, плотностью для (ξ, η) (по определению) является и $f_\xi(x) f_\eta(y)$, то есть $f_{(\xi,\eta)}(x,y) = f_\xi(x) f_\eta(y)$ выполнено почти всюду.

Если же $f_{(\xi,\eta)}(x,y)$ и $f_\xi(x) f_\eta(y)$ почти всюду совпадают, то всегда $(*) = (**)$, а следовательно и ξ, η также независимы.

Формула свертки

Дана плотность $f_{(\xi_1,\xi_2)}$, нужно найти плотность для $\xi_1+\xi_2$. Для всех $x\in\mathbb{R}$ имеем

$$\begin{aligned} F_{\xi_1+\xi_2}(x) &= \mathbb{P}\{\omega|\xi_1(\omega)+\xi_2(\omega)\le x\} \\ &= \int_{\{(y_1,y_2)|y_1+y_2\le x\}} f_{(\xi_1,\xi_2)}(y_1,y_2)\lambda(d(y_1,y_2)) \\ &= \int_{\{(y_1,y_3)|y_3\le x\}} f_{(\xi_1,\xi_2)}(y_1,y_3-y_1)\lambda(d(y_1,y_3)) \\ &= \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{(\xi_1,\xi_2)}(y_1,y_3-y_1)dy_1dy_3. \end{aligned}$$

Теперь, по определению плотности,

$$f_{\xi_1+\xi_2}(x) = \int_{-\infty}^{+\infty} f_{(\xi_1,\xi_2)}(y,x-y)dy \qquad \forall x\in\mathbb{R}.$$

Пусть η известно, тогда можно оценить ξ через $\mathbb{E}(\xi|\eta)$

Пусть $m = 2, \vec{\xi} = (\xi, \eta) : \Omega \to \mathbb{R}^2$. Зададим функцию $h : \mathbb{R} \to \mathbb{R}$ правилом

$$h(y) \triangleq \frac{\int_{\mathbb{R}} x f_{(\xi,\eta)}(x, y)\, dx}{\int_{\mathbb{R}} f_{(\xi,\eta)}(x, y)\, dx} \qquad \forall y \in \mathbb{R}.$$

Проверим, что для любой борелевской функции g с $\mathbb{E}|g(\eta)\xi| < +\infty$,

$$\begin{aligned}
\mathbb{E}(g(\eta)\xi) &= \int_{\mathbb{R}^2} g(y) x f_{(\xi,\eta)}(x, y)\, \lambda(d(x, y)) \\
&= \int_{\mathbb{R}} g(y) \int_{\mathbb{R}} x f_{(\xi,\eta)}(x, y)\, dx dy \\
&= \int_{\mathbb{R}} g(y) \frac{\int_{\mathbb{R}} x f_{(\xi,\eta)}(x, y)\, dx}{\int_{\mathbb{R}} f_{(\xi,\eta)}(x, y)\, dx} f_\eta(y) dy \\
&= \int_{\mathbb{R}} g(y) h(y) (\eta \# \mathbb{P})(dy) = \mathbb{E}(g(\eta) h(\eta)).
\end{aligned}$$

Итак,

$$\mathbb{E}(\xi|\eta)(\omega) = h(\eta(\omega)) \qquad \forall \omega \in \Omega.$$

Регрессия. Условная плотность

Скалярные отображения

$$h(y) \triangleq \frac{\int_{\mathbb{R}} x f_{(\xi,\eta)}(x,y)\,dx}{\int_{\mathbb{R}} f_{(\xi,\eta)}(x,y)\,dx} \qquad \forall y \in \mathbb{R},$$

$$f_{(\xi|\eta)}(x|y) \triangleq \frac{f_{(\xi,\eta)}(x,y)}{\int_{\mathbb{R}} f_{(\xi,\eta)}(z,y)\,dz} = \frac{f_{(\xi,\eta)}(x,y)}{f_\eta(y)} \qquad \forall x,y \in \mathbb{R}$$

называют регрессией ξ на η и условной плотностью ξ относительно η. При этом, для $\eta(\omega) = y$ имеет место

$$\mathbb{E}(\xi|\eta)(\omega) = h(\eta(\omega)) = h(y) = \int_{\mathbb{R}} x f_{(\xi|\eta)}(x|y)dx.$$

Подумать: проверьте, что при почти всех фиксированных x $f_{(\xi|\eta)}(x|y)$ — плотность какой-то меры относительно другой (каких?).

Абсолютно непрерывный случай

Пусть (ξ,η) — абсолютно непрерывная векторная случайная величина. Тогда,

$$
\begin{aligned}
\mathbb{E}\xi &= \int_{\mathbb{R}} x f_\xi(x)\,dx, && \text{(математическое ожидание)}\\
f_{(\xi|\eta)}(x|y) &= \frac{f_{(\xi,\eta)}(x,y)}{\int_{\mathbb{R}} f_{(\xi,\eta)}(z,y)\,dz} = \frac{f_{(\xi,\eta)}(x,y)}{f_\eta(y)}, && \\
& && \text{(условная плотность)}\\
\mathbb{E}(\xi|\eta)(\omega)|_{y=\eta(\omega)} &= \int_{\mathbb{R}} x f_{(\xi|\eta)}(x|y)\,dx. && \text{(условное матожидание)}
\end{aligned}
$$

При этом,

$$\int_{\mathbb{R}} f_\xi(x)\,dx = 1, \qquad \int_{\mathbb{R}} f_{(\xi|\eta)}(x|y)\,dy = 1.$$

Подумать: вместо λ можно взять любую меру, относительно которой $(\xi,\eta)\#\mathbb{P}$ абсолютно непрерывна. Возьмем, например, считающую меру... Тогда мы получим...

Сравните: дискретный случай

Пусть (ξ, η) — векторная случайная величина, принимающая не более чем счетное число значений. Тогда,

$$\begin{aligned}
\mathbb{E}\xi &= \sum_{i\in\mathbb{N}} a_i \mathbb{P}(\xi = a_i), \qquad \text{(математическое ожидание)} \\
\mathbb{P}(\xi = a_i | \eta = b_j) &= \frac{\mathbb{P}(\xi = a_i, \eta = b_j)}{\sum_{k\in\mathbb{N}} \mathbb{P}(\xi = a_i, \eta = b_k)} = \frac{\mathbb{P}(\xi = a_i, \eta = b_j)}{\mathbb{P}(\eta = b_j)}, \\
&\qquad \text{(условная вероятность)} \\
\mathbb{E}(\xi|\eta)(\omega)|_{y=\eta(\omega)} &= \sum_{i\in\mathbb{N}} a_i \mathbb{P}(\xi = a_i | \eta = b_j). \qquad \text{(условное матожидание)}
\end{aligned}$$

При этом,

$$\sum_{i\in\mathbb{N}} \mathbb{P}(\xi = a_i) = 1, \qquad \sum_{k\in\mathbb{N}} \mathbb{P}(\xi = a_i | \eta = b_k) = 1.$$

Свойства условного математического ожидания

0^0 $\mathbb{E}(1_A|\mathcal{H}) = \mathbb{P}(A|\mathcal{H})$ для всех $A \in \mathcal{F}$.

1^0 $\mathbb{E}(\xi|\mathcal{H}) \geq 0$, если $\xi(\omega) \geq 0$ для почти всех $\omega \in \Omega$.

2^0 $\mathbb{E}(\xi_1|\mathcal{H}) \geq \mathbb{E}(\xi_2|\mathcal{H})$, если $\xi_1(\omega) \geq \xi_2(\omega)$ для всех $\omega \in \Omega$.

3^0 $\mathbb{E}(c|\mathcal{H}) = c$ для вырожденной случайной величины $c \in \mathbb{R}$;
более того, $\mathbb{E}(\xi|\mathcal{H}) = \xi$ для $\mathcal{H}$-измеримой ξ,
в частности, всегда $\mathbb{E}(f(\xi)|\xi) = f(\xi)$, $\mathbb{E}(\xi|\xi) = \xi$.

4^0 $\mathbb{E}(c\xi|\mathcal{H}) = c\mathbb{E}(\xi|\mathcal{H})$ для $c \in \mathbb{R}$,
более того, $\mathbb{E}(\eta\xi|\mathcal{H}) = \eta\mathbb{E}(\xi|\mathcal{H})$, если η $\mathcal{H}$-измерима.

5^0 $\mathbb{E}(\xi_1|\mathcal{H}) + \mathbb{E}(\xi_2|\mathcal{H}) = \mathbb{E}(\xi_1 + \xi_2|\mathcal{H})$.

6^0 $\eta_1\mathbb{E}(\xi_1|\mathcal{H}) + \eta_2\mathbb{E}(\xi_2|\mathcal{H}) = \mathbb{E}(\eta_1\xi_1 + \eta_2\xi_2|\mathcal{H})$ для $\mathcal{H}$-измеримых η_1, η_2.

7^0 $\mathbb{E}(\xi\eta|\mathcal{H}) = \mathbb{E}(\xi|\mathcal{H})\mathbb{E}(\eta|\mathcal{H})$, если $\xi 1_H, \eta 1_H$ независимы для всех $H \in \mathcal{H}$.

Условное матожидание как основание перпендикуляра

18^0 Для случайной величины $\xi \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ и σ-алгебры $\mathcal{H}$ среди всех $\eta \in \mathcal{L}^2(\Omega, \mathcal{H}, \mathbb{P})$ ошибка $\mathbb{E}(\xi - \eta)^2$ минимальна только при $\eta = \mathbb{E}(\xi|\mathcal{H})$.

Доказательство.

Рассмотрим $\zeta \in \mathcal{L}^2(\Omega, \mathcal{H}, \mathbb{P})$. По свойству 4^o $\mathbb{E}(\zeta\xi|\mathcal{H}) = \zeta\mathbb{E}(\xi|\mathcal{H})$ и

$$\begin{aligned}
\mathbb{E}((\xi - \mathbb{E}(\xi|\mathcal{H}) + \zeta)^2) &= \mathbb{E}((\xi - \mathbb{E}(\xi|\mathcal{H})^2) + \mathbb{E}(\zeta^2) \\
&\quad -2\mathbb{E}(\zeta\xi) + 2\mathbb{E}(\mathbb{E}(\zeta\xi|\mathcal{H})) \\
&= \mathbb{E}((\xi - \mathbb{E}(\xi|\mathcal{H}))^2) + \mathbb{E}(\zeta^2) \\
&\geq \mathbb{E}((\xi - \mathbb{E}(\xi|\mathcal{H}))^2).
\end{aligned}$$

Тогда ошибка минимальна (и равна $\mathbb{E}((\xi - \mathbb{E}(\xi|\mathcal{H}))^2)$) в точности при $\zeta = 0$, что и требовалось доказать.

17^0 [с-но]: при тех же условиях $\mathbb{D}\xi = \mathbb{E}((\xi - \mathbb{E}(\xi|\mathcal{H}))^2) + \mathbb{D}\mathbb{E}(\xi|\mathcal{H})$.

Существование матожидания через теорему Лебега: [с-но]

8^0 $\mathbb{E}(\mathbb{E}(\eta|\mathcal{G}))$ существует и конечно тогда и только тогда, когда $\eta \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, в частности, из существования $\mathbb{E}\eta$ следует существование $\mathbb{E}(\eta|\mathcal{G})$.

9^0 $\mathbb{E}(\xi|\mathcal{H})$ существует, если ξ ограничено, то есть $\xi \in \mathcal{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$.

10^0 $\mathbb{E}(|\xi\eta|\,|\mathcal{H})$ существует, если $\xi, \eta \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

11^0 $\mathbb{E}(|\xi|\,|\mathcal{H}), \mathbb{E}(\xi|\mathcal{H})$ существуют, если $\xi, \eta \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

12^0 $g(\mathbb{E}(\xi|\mathcal{H})) \leq \mathbb{E}(g(\xi)|\mathcal{H})$ для любой выпуклой (вниз) скалярной функции g; при этом, если существует матожидание справа, то и слева тоже существует.

13^0 $\mathbb{E}(\sum_{i=1}^\infty \xi_i|\mathcal{H})$ существует и равно $\sum_{i=1}^\infty \mathbb{E}(\xi_i|\mathcal{H})$, если конечен ряд $\sum_{i=1}^\infty \mathbb{E}|\xi_i|$.

“Условные” свойства условного матожидания [с-но]

Предполагая, что все нужные условные матожидания существуют...

14^0 $\mathbb{E}(\xi|\mathcal{G}) = \xi$ тогда и только тогда, когда ξ $\mathcal{G}$-измерима.

15^0 $\mathbb{E}(\xi|\mathcal{G}) = \mathbb{E}\xi$, если ξ независима относительно $\mathcal{G}$;

в частности, $\mathbb{E}(\xi|\mathcal{G}) = \mathbb{E}\xi$ для $\mathcal{G} = \{\Omega, \varnothing\}$.

16^0 $\mathbb{E}(\xi|\mathcal{H}) = \mathbb{E}(\mathbb{E}(\xi|\mathcal{G})|\mathcal{H})$ при $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$;

в частности, $\mathbb{E}(\xi|\eta) = \mathbb{E}(\mathbb{E}(\xi|\eta, \zeta)|\eta)$,

$\mathbb{E}\mathbb{E}(\xi|\mathcal{G}) = \mathbb{E}\xi$,

$\mathbb{P}(A) = \mathbb{E}\mathbb{P}(A|\mathcal{G})$ для всех $A \in \mathcal{F}$

(формула полной вероятности),

$\mathbb{E}(\xi|\mathcal{G}) = \mathbb{E}(\xi|\mathbb{E}(\xi|\mathcal{G}))$ (свойство Дуба).

Предел условных матожиданий

Теорема Лебега. Если последовательность случайных величин ξ_n сходится к ξ п.в., и для случайной величины $\varphi \in \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ имеют место $\mathbb{E}\varphi < +\infty$ и $|\xi_n| \le \varphi$ при всех натуральных n, то ξ также суммируема, при этом

$$\mathbb{E}(\xi|\mathcal{H}) = \lim_{n\to\infty} \mathbb{E}(\xi_n|\mathcal{H}) \in \mathcal{L}^p(\Omega, \mathcal{H}, \mathbb{P}).$$

Теорема Леви. Если последовательность случайных величин ξ_n монотонна, их матожидания $\mathbb{E}\xi_n$ ограничены, то случайная величина $\xi \triangleq \lim_{n\to\infty} \xi_n$ также суммируема, при этом

$$\mathbb{E}(\xi|\mathcal{H}) = \lim_{n\to\infty} \mathbb{E}(\xi_n|\mathcal{H}) \in \mathcal{L}^1(\Omega, \mathcal{H}, \mathbb{P}).$$

Лемма Фату. Для всякой последовательности неотрицательных случайных величин ξ_n

$$\mathbb{E}(\liminf_{k\to\infty} \xi_k|\mathcal{H}) \le \liminf_{k\to\infty} \mathbb{E}(\xi_k|\mathcal{H}).$$