Условное матожидание

Н. Киселев

Клуб теории вероятностей ФЭН ВШЭ

9 июня 2022 г.

Напоминание

1. Вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$. Сигма-алгебра (замкнутость относительно дополнения и счетного объединения). Вероятность (2 аксиомы).
2. Минимальная сигма-алгебра.

Определение

Пусть Ω - пространство элементарных исходов, $\mathcal{A}$ - какое-то семейство его подмножеств. Минимальной сигма-алгеброй, порожденной $\mathcal{A}$, называется семейство $\sigma(\mathcal{A})$, удовлетворяющее условиям:
1. $\mathcal{A} \subset \sigma(\mathcal{A})$
2. $\sigma(\mathcal{A})$ является σ-алгеброй
3. Если $\mathcal{G}$ - σ-алгебра и $\mathcal{A} \subset \mathcal{G}$, то $\sigma(\mathcal{A}) \subset \mathcal{G}$

3. Измеримость. Далее будем использовать обозначение $X \in \mathcal{F}$.

Напоминание

4. Интегрируемость.

Определение

Будем говорить, что f интегрируема, если $\int |f| d\mu < \infty$.

Замечание

В качестве f можно рассматривать случайную величину, так как случайная величина - функция $\Omega \to \mathbb{R}$.

5. Абсолютная непрерывность относительно меры.

Определение

ν абсолютно непрерывно относительно μ, если для любого измеримого A из $\mu(A) = 0$ следует $\nu(A) = 0$. Пишут $\nu \ll \mu$.

Напоминание

6. Теорема Радона-Никодима.

Теорема

Пусть μ и ν - σ-конечные (то есть все пространство может быть представлено в виде счетного объединения измеримых множеств конечной меры) меры на $(\Omega, \mathcal{F})$. Если $\nu \ll \mu$, то существует такая функция $f \in \mathcal{F}$, что для всех $A \in \mathcal{F}$

$$\int\limits_A f d\mu = \nu(A)$$

f обычно обозначают как $\frac{d\nu}{d\mu}$ - производная Радона-Никодима меры ν относительно меры μ.

Напоминание

7. Теорема Лебега о мажорируемой сходимости.

Теорема

Пусть $(X, \mathcal{F}, \mu)$ - пространство с мерой. Пусть $\{f_n\}_{n=1}^{\infty}$ и f - измеримые функции на X и $f_n(x) \to f$ почти всюду. Если существует определенная на том же пространстве интегрируемая функция g такая, что для всех n $|f_n(x)| < g(x)$ почти всюду, то функции f_n и f интегрируемы и

$$\lim_{n\to\infty} \int\limits_X f_n(x)\mu(dx) = \int\limits_X f(x)\mu(dx)$$

Формальное определение

Определение (Условное матожидание)

Пусть $(\Omega, \mathcal{F}_0, \mathbb{P})$ - вероятностное пространство, $\mathcal{F} \subset \mathcal{F}_0$ - σ-алгебра, и $X \in \mathcal{F}_0$ - случайная величина с $\mathbb{E}|X| < \infty$. **Условное** матожидание X **по сигма-алгебре** $\mathcal{F}$, $\mathbb{E}(X|\mathcal{F})$ - любая случайная величина Y, удовлетворяющая:

(i) $Y \in \mathcal{F}$

(ii) $\forall A \in \mathcal{F} \ \int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$

Формальное определение

Утверждение

Если Y удовлетворяет (i) и (ii), то Y интегрируема.

Доказательство.

Пусть $A = \{Y > 0\}$. Воспользовавшись (ii) дважды, получаем:

$$\int\limits_A Y d\mathbb{P} = \int\limits_A X d\mathbb{P} \leqslant \int\limits_A |X| d\mathbb{P}$$

$$\int\limits_{A^c} -Y d\mathbb{P} = \int\limits_{A^c} -X d\mathbb{P} \leqslant \int\limits_{A^c} |X| d\mathbb{P}$$

Сложив, получаем $\mathbb{E}|Y| \leqslant \mathbb{E}|X| \leqslant \infty$. $\square$

Утверждение

Если Y и Y' удовлетворяют (i) и (ii), то $Y = Y'$ почти наверное.

Доказательство.

Мы знаем, что $\int\limits_A Y d\mathbb{P} = \int\limits_A Y' d\mathbb{P} \;\; \forall A \in \mathcal{F}$. Возьмем $A = \{Y - Y' \geqslant \varepsilon > 0\}$, получим

$$0 = \int\limits_A (X - X) d\mathbb{P} = \int\limits_A (Y - Y') d\mathbb{P} \geqslant \varepsilon \mathbb{P}(A)$$

Итого, $\mathbb{P}(A) = 0$. Это выполняется для всех ε, Y и Y' можно поменять местами и отсюда очевидно следует равенство почти наверное. □

Строго говоря, равенства вида $Y = \mathbb{E}(X|\mathcal{F})$ должны писаться как $Y = \mathbb{E}(X|\mathcal{F})$ *a.s.*, но мы опустим это.

Условное матожидание, существование

Утверждение

Условное матожидание $\mathbb{E}(X|\mathcal{F})$ существует.

Доказательство.

Для начала рассмотрим случай $X \geqslant 0$. Будем пользоваться теоремой Радона-Никодима. Положим $\mu = \mathbb{P}$ и $\nu(A) = \int\limits_A X d\mathbb{P}$ для $A \in \mathcal{F}$. Можно показать, что ν - это мера, и что $\nu \ll \mu$. Производная Радона-Никодима $\frac{d\nu}{d\mu} \in \mathcal{F}$ дает $\int\limits_A X d\mathbb{P} = \nu(A) = \int\limits_A \frac{d\nu}{d\mu} d\mathbb{P}$. Подставив $A = \Omega$, получим интегрируемость $\frac{d\nu}{d\mu}$, и только что мы показали выполнение свойства (ii). Итого, $\frac{d\nu}{d\mu} = E(X|\mathcal{F})$.

Общий случай выводится из простого представления $X = X^+ - X^-$. $\square$

Примеры

Интуиция об условном матожидании - о сигма-алгебре $\mathcal{F}$ можно думать как об описании информации, которая у нас имеется на руках. Про каждое событие $A \in \mathcal{F}$ мы знаем, произошло оно или нет. И тогда $\mathbb{E}(X|\mathcal{F})$ - это наше лучшее предположение о значении случайной величины X на основе имеющейся у нас информации.

1. $X \in \mathcal{F}$. Тогда $\mathbb{E}(X|\mathcal{F}) = X$.
   Из нашей интуиции это следует мгновенно. Более строго, (i) выполняется по предположению измеримости X, а (ii) выполняется очевидно - $\int\limits_A X d\mathbb{P} = \int\limits_A X d\mathbb{P}$ для всех $A \in \mathcal{F}$.

Примеры

2. Пусть X независимо от $\mathcal{F}$, то есть $\forall B \in \mathcal{R}$, $\forall A \in \mathcal{F}$ $\mathbb{P}\big(\{X \in B\} \cap A\big) = \mathbb{P}(X \in B)\mathbb{P}(A)$. Тогда $\mathbb{E}(X|\mathcal{F}) = \mathbb{E}X$.
   Это противоположный предыдущему примеру случай - если мы ничего не знаем о случайной величине X, лучшее предположение о ней - ее матожидание. Более строго:
   (i) выполняется, так как $\mathbb{E}X \in \mathcal{F}$ (это просто константа).
   Чтобы показать, что (ii) выполняется, рассмотрим случайную величину $\mathbb{I}_A$. X и $\mathbb{I}_A$ независимы, поэтому

$$\int_A X d\mathbb{P} = \int_A (X\mathbb{I}_A) d\mathbb{P} = \int_\Omega (X\mathbb{I}_A) d = \mathbb{E}(X\mathbb{I}_A) = \mathbb{E}X \cdot \mathbb{E}\mathbb{I}_A =$$

$$= \mathbb{E}X \cdot \mathbb{P}(A) = \int_A \mathbb{E}X d\mathbb{P}$$

Примеры

3. Рассмотрим конечное или счетное разбиение $\Omega = (\Omega_1, \Omega_2, \dots)$. Оно пораждает σ-алгебру $\mathcal{F} = \sigma(\Omega_1, \Omega_2, \dots)$. Тогда на множестве элементарных исходов Ω_i $\mathbb{E}(X|\mathcal{F}) = \dfrac{\mathbb{E}(X;\Omega_i)}{\mathbb{P}(\Omega_i)}$, где $\mathbb{E}(X;\Omega_1)$ - среднее значение X на Ω_i - $\int\limits_{\Omega_i} Xd\mathbb{P}$.
   Интуиция - информация, содержащаяся в Ω_i, задает область разбиения, в которой находится X, и тогда наше лучшее предположение - среднее значение X по этой области. Более строго:
   (i) выполняется, так как $\dfrac{\mathbb{E}(X;\Omega_i)}{\mathbb{P}(\Omega_i)}$ - константа на каждой области разбиения, то есть такая случайная величина измерима.

Примеры

3. Чтобы показать, что (ii) выполняется, заметим, что достаточно проверить это свойство для $A = \Omega_i$ (в $\mathcal{F}$ могут быть только disjoint объединения событий Ω_i, пересечений быть не может, так как это разбиение, и тогда можно просто применить свойство интеграла Лебега для disjoint событий):

$$\int\limits_{\Omega_i} \frac{\mathbb{E}(X;\Omega_i)}{\mathbb{P}(\Omega_i)} d\mathbb{P} = \mathbb{E}(X;\Omega_i) = \int\limits_{\Omega_i} X d\mathbb{P}$$

Примеры

4. Перейдем к более привычной записи условного матожидания, а именно матожидания относительно другой случайной величины.

Определение

$$\mathbb{E}(X|Y) = \mathbb{E}(X|\sigma(Y)),$$

где $\sigma(Y)$ - σ-алгебра, порожденная случайной величиной Y.

Примеры

5. Свяжем рассмотренное определение условного матожидания с определением, встречающимся в обычных курсах - $\mathbb{E}\big(g(X)|Y\big) = h(Y)$, где

$$h(y) = \frac{\int g(x)f(x,y)dx}{\int f(x,y)dx}$$

Покажем, что это и правда условное матожидание:
(i) выполняется, так как $h(Y) \in \sigma(Y)$ (h - адекватная функция от случайной величины Y, которая естественно измерима относительно σ-алгебры, порожденной самой собой).
(ii) выполняется, так как если $A \in \sigma(Y)$, то $A = \{w : Y(w) \in B\}$ для некоторого $B \in \mathcal{R}$, и тогда

$$\int\limits_A h(Y)d\mathbb{P} = \int\limits_B \int h(y)f(x,y)dxdy = \int\limits_B \int g(x)f(x,y)dxdy =$$

$$= \mathbb{E}\big(g(X)\mathbb{I}_B(Y)\big) = \mathbb{E}\big(g(X);A\big) = \int g(X)d\mathbb{P}$$

Примеры

6. Пусть X и Y - независимые случайные величины, φ - функция с $\mathbb{E}|\varphi(X,Y)| < \infty$, и $g(x) = \mathbb{E}\big(\varphi(x,Y)\big)$. Тогда
$$\mathbb{E}\big(\varphi(X,Y)|X\big) = g(X)$$
(i) выполнено очевидно - $g(X) \in \sigma(X)$.
Чтобы показать (ii), опять запишем $A \in \sigma(X)$ как $A = \{X \in C\}$, и из независимости и определения g получаем
$$\int\limits_A \varphi(X,Y)d\mathbb{P} = \mathbb{E}\big(\varphi(X,Y)\mathbb{I}_C(X)\big) = \int\int \varphi(x,y)\mathbb{I}_C(x)\nu(dy)\mu(dx) =$$
$$= \int g(x)\mathbb{I}_C(x)\mu(dx) = \mathbb{E}\big(g(X)\mathbb{I}_C(X)\big) = \int\limits_A g(X)d\mathbb{P}$$

Свойства

1. $\mathbb{E}(aX + Y|\mathcal{F}) = a\mathbb{E}(X|\mathcal{F}) + \mathbb{E}(Y|\mathcal{F})$
2. Если $X \leqslant Y$, то $\mathbb{E}(X|\mathcal{F}) \leqslant \mathbb{E}(Y|\mathcal{F})$
3. Если $X_n \geqslant 0$ и $X_n \uparrow X$ с $\mathbb{E}X < \infty$, то $\mathbb{E}(X_n|\mathcal{F}) \uparrow \mathbb{E}(X|\mathcal{F})$
   Доказательство: пусть $Y_n = X - X_n$, $Y_n \downarrow 0$. Достаточно показать $\mathbb{E}(Y_n|\mathcal{F}) \downarrow 0$. Из убывания и ограниченности Y_n $Z_n = \mathbb{E}(Y_n|\mathcal{F}) \downarrow Z_\infty$ - сходится к какому-то пределу. По определению для $A \in \mathcal{F}$

$$\int_A Z_n d\mathbb{P} = \int_A Y_n d\mathbb{P}$$

   Так как $Y_n \downarrow 0$, то по теореме Лебега о мажорируемой сходимости

$$\int_A Z_\infty d\mathbb{P} = \lim_{n\to\infty} \int_A Z_n d\mathbb{P} = 0$$

   Это верно для всех $A \in \mathcal{F}$, поэтому $Z_\infty = 0$.

Свойства

4. Если φ выпуклая и $\mathbb{E}|X|, \mathbb{E}|\varphi(X)| < \infty$, то $\varphi\big(\mathbb{E}(X|\mathcal{F})\big) \leqslant \mathbb{E}\big(\varphi(X)|\mathcal{F}\big)$
5. $\mathbb{E}(\mathbb{E}(X|\mathcal{F})) = \mathbb{E}X$
   Доказательство: следует из свойства (ii) условного матожидания, взяв $A = \Omega$.

Свойства

6. Если $\mathcal{F}_1 \subset \mathcal{F}_2$, то:
   1. $\mathbb{E}\big(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_2\big) = \mathbb{E}(X|\mathcal{F}_1)$
   2. $\mathbb{E}\big(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1\big) = \mathbb{E}(X|\mathcal{F}_1)$

   Неформально - меньшая σ-алгебра всегда побеждает (вспомните интуицию с информацией).

   Доказательство:

   Первое равенство верно, так как $\mathbb{E}(X|\mathcal{F}_1) \in \mathcal{F}_1$, а $\mathcal{F}_1 \subset \mathcal{F}_2$, т.е. $\mathbb{E}(X|\mathcal{F}_1) \in \mathcal{F}_2$, и тогда согласно примеру 1 получаем, что равенство верно.

   Для доказательства второго заметим, что по определению $\mathbb{E}(X|\mathcal{F}_1) \in \mathcal{F}_1$, т.е. (i) выполнено. Также если $A \in \mathcal{F}_1 \subset \mathcal{F}_2$, то

$$\int_A \mathbb{E}(X|\mathcal{F}_1) d\mathbb{P} = \int_A X d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{F}_2) d\mathbb{P},$$

   откуда получаем выполнение (ii).

Свойства

7. Если $X \in \mathcal{F}$ и $\mathbb{E}|Y|$, $\mathbb{E}|XY| < \infty$, то $\mathbb{E}(XY|\mathcal{F}) = X\mathbb{E}(Y|\mathcal{F})$
Доказательство: (i) выполняется, так как $X \in \mathcal{F}$ по предположению, $\mathbb{E}(Y|\mathcal{F})$ по определению. Докажем (ii) для самого элементарного случая $X = \mathbb{I}_B$ для какого-то $B \in \mathcal{F}$. В этом случае для $A \in \mathcal{F}$

$$\int\limits_A \mathbb{I}_B \mathbb{E}(Y|\mathcal{F})d\mathbb{P} = \int\limits_{A\cap B} \mathbb{E}(Y|\mathcal{F})d\mathbb{P} = \int\limits_{A\cap B} Yd\mathbb{P} = \int\limits_A \mathbb{I}_B Yd\mathbb{P}$$

Отсюда по линейности получаем, что (ii) выполнено и для простой случайной величины X, а далее для X, $Y > 0$ построим приближение $X_n \uparrow X$ и воспользуемся теоремой Лебега о мажорируемой сходимости. Наконец, для общего случая разложим X и Y на положительную и отрицательную части.

8. Если $\mathbb{E}X^2 < \infty$, то $\mathbb{E}(\mathbb{X}|\mathcal{F})$ - такая случайная величина $Y \in \mathcal{F}$, которая минимизирует $\mathbb{E}(X - Y)^2$
Замечание: это утверждение показывает геометрическую интерпретацию условного матожидания.
$L^2(\mathcal{F}_0) = \{Y \in \mathcal{F}_0 : \mathbb{E}Y^2 < \infty\}$ - Гильбертово прстранство, $L^2(\mathcal{F})$ - замкунтое подпространство. Тогда $\mathbb{E}(X|\mathcal{F})$ - проекция X на это подпространство, то есть точка из $L^2(\mathcal{F})$, наиболее близкая к X.
Доказательство: взяв $Z \in L^2(\mathcal{F})$, по свойству 7 получим $Z\mathbb{E}(X|\mathcal{F}) = \mathbb{E}(ZX|\mathcal{F})$ (вопрос - почему можно применить свойство 7?). Взяв матожидания, получаем

$$\mathbb{E}\big(Z\mathbb{E}(X|\mathcal{F})\big) = \mathbb{E}\big(\mathbb{E}(ZX|\mathcal{F})\big) = \mathbb{E}(ZX) \Leftrightarrow \mathbb{E}\big(Z(X - \mathbb{E}(X|\mathcal{F}))\big) = 0$$

Положив $Y = \mathbb{E}(X|\mathcal{F}) - Z$, легко видеть, что минимум достигается при $Z = 0$.