«Дополнительные главы теории случайных процессов».
Лектор: Борисов Андрей Владимирович, проф. каф. МС
Borisych@me.com

Рекомендуемая литература по лекции:

1. Ширяев А.Н. Вероятность. М.: Наука, 1989.
2. Булинский А.В., Ширяев А.Н. Теория случайных процессов. М.: Физматлит, 2005.
3. Липцер Р.Ш., Ширяев А.Н. Статистика случайных процессов. М.: Физматлит, 1974.

Лекция 1. Аксиоматика ТВ. Условное математическое ожидание

1. *Аксиоматика теории вероятностей.*
2. *Случайные величины. Математическое ожидание.*
3. *Условное математическое ожидание (УМО).*
4. *Использование УМО в задачах оптимального байесовского оценивания*
5. *Задачи для самостоятельного решения.*

1. *Аксиоматика теории вероятностей*

Определение 1.1. Тройка (Ω, F, P), в которой Ω – пространство элементарных событий, F - σ-алгебра на Ω, а P – вероятностная мера на F, называют *вероятностным пространством*. Если σ-алгебра F включает в себя все свои подмножества множеств нулевой вероятностной меры, то соответствующее вероятностное пространство называется *полным*.

Ниже для напоминания представлены определения F и P.

Определение 1.2. Множество подмножеств Ω *образуют σ-алгебру* F, если F обладает следующими свойствами:
1) $\Omega \in F$,
2) $\forall A$: $A \in F \Rightarrow (\Omega \setminus A) \in F$,
3) $\forall \{A_n\}_{n \in N} \subseteq F \Rightarrow \sum_{n \in N} A_n \triangleq \bigcup_{n \in N} A_n \in F$, $\prod_{n \in N} A_n \triangleq \bigcap_{n \in N} A_n \in F$.

σ-алгебра $F = \{\emptyset, \Omega\}$ называется *тривиальной*.

Определение 1.3. Функция множества $P\{A\}$: $F \to R$ называется *вероятностной мерой*, если она удовлетворяет следующим свойствам:
1) $\forall A$: $A \in F$ $P\{A\} \geq 0$ (аксиома неотрицательности),
2) $P\{\Omega\} = 1$ (аксиома нормировки),
3) $\forall A, B$: $A, B \in F$, $AB = \emptyset \Rightarrow P\{A + B\} = P\{A\} + P\{B\}$ (аксиома аддитивности),
4) $\forall \{A_n\}_{n \in N} \subseteq F$: $A_1 \supseteq A_2 \supseteq A_3 \ldots$, $\prod_{n \in N} A_n = \emptyset \Rightarrow P\{A_n\} \to 0$ (аксиома непрерывности вероятности в нуле).

Определение 1.4. Пара (Ω, F) называется *измеримым пространством*. Примерами измеримых пространств служат $(R,\ B(R))$, $(R^n,\ B(R^n))$, где $B(.)$ – борелевская σ -алгебра на соответствующем пространстве.

Определение 1.5. Пусть (Ω, F) и (Ξ, G) – два измеримых пространства. Отображение $X(\omega)\colon\ \Omega\ \to \Xi$ называется $F|\,G$ – *измеримым*, если для $\forall\ G \in G\ \Rightarrow X^{-1}(G)\ \triangleq \{\omega \in \Omega\colon\ \ X(\omega) \in G\} \in F$.

2. Случайные величины. Математическое ожидание

Определение 1.6. Пусть (Ω, F, P) – вероятностное пространство, а (Ξ, G) – измеримое пространство. $F|\,G$ - измеримое отображение $X(\omega)\colon\ \Omega\ \to \Xi$ называется *случайным элементом*. В частном случае, если $(\Xi, G)\ =\ (R,\ B(R))$, то $X(\omega)$ – *случайная величина*, а если $(\Xi, G) = (R^n,\ B(R^n))$, то $X(\omega)$ – *n-мерный случайный вектор*.

Замечание 1.1. Благодаря тому, что отображение $X(\omega)$ - $F|\,G$ - измеримое, оно индуцирует на измеримом пространстве (Ξ, G) вероятностную меру.

Определение математического ожидания случайной величины (являющегося приложением интеграла Лебега) вводится за несколько этапов.

Определение 1.7. Подмножество $D \subseteq F$ называется *конечным разбиением пространства* Ω, если $D\ = \left\{D_1, ..., D_n\right\}$, $n\ <\ \infty$, и выполнены следующие свойства: $D_i D_j \equiv \emptyset$ при $i \neq j$, и $\sum\limits_{i=1}^{n} D_i =\ \Omega$. Также для простоты будем считать, что $P\{D_i\}\ >\ 0$.

Замечание 1.2. Иногда конечное разбиение пространства Ω называют *группой гипотез*.

Определение 1.8. Пусть D – некоторое конечное разбиение пространства Ω. Случайная величина

$$X(\omega) =\ \sum_{i=1}^{n} x_i I_{D_i}(\omega) \qquad (1.1)$$

называется *простой*. Математическое ожидание простой случайной величины $X(\omega)$ определяется следующей формулой

$$E[X(\omega)]\ =\ \sum_{i=1}^{n} x_i P\{D_i\}. \qquad (1.2)$$

Пусть $X(\omega)$ – некоторая неотрицательная случайная величина на (Ω, F, P). Известно, что для нее можно построить последовательность простых неотрицательных случайных величин $X_k(\omega)$, такую, что $X_k(\omega) \uparrow X(\omega)$ при $k \to \infty$ для любого $\omega\ \in\ \Omega$.

Определение 1.9. *Математическим ожиданием* (интегралом Лебега) неотрицательной случайной величины $X(\omega)$ называется величина

$$E[X(\omega)] = E\left[X_k(\omega)\right]. \qquad (1.3)$$

Замечание 1.3. В силу данного определения математическое ожидание может быть бесконечным.

Определение 1.10. Пусть $X(\omega)$ – произвольная (знакопеременная) случайная величина. Случайная величина $X^+(\omega) = max(X(\omega), 0)$ называется *верхней срезкой* $X(\omega)$, а $X^-(\omega) =- min(X(\omega), 0)$ – *нижней срезкой*. При этом $X(\omega) \equiv X^+(\omega) - X^-(\omega)$. Математическое ожидание $E[X(\omega)]$ случайной величины определено, если $(E\left[X^+(\omega)\right], E\left[X^-(\omega)\right]) < \infty$. Если $(E\left[X^+(\omega)\right], E\left[X^-(\omega)\right]) < \infty$, то говорят, что $X(\omega)$ *имеет конечное математическое ожидание* (или является интегрируемой случайной величиной). Эквивалентное условие: $E[|X(\omega)|] < \infty$.

Замечание 1.4. Иногда математическое ожидание подчеркнуто записывается в форме интеграла Лебега:

$$E[X(\omega)] = \int_{\Omega} X(\omega)P(d\omega).$$

Свойства математических ожиданий

1. Для любого события $A \in F$ $P\{A\} = \int_A P(d\omega) = \int_{\Omega} I_A(\omega)P(d\omega) = E\left[I_A(\omega)\right]$, т.е. вероятность события A – математическое ожидание его индикаторной функции.
2. Если $E[X(\omega)]$ существует, и C – константа, то $E[CX(\omega)] = CE[X(\omega)]$.
3. Если $X(\omega)$ и $Y(\omega)$ – интегрируемые случайные величины, и $X(\omega) \leq Y(\omega)$, то и $E[X(\omega)] \leq E[Y(\omega)]$.
4. Если $X(\omega)$ – интегрируемая случайная величина, то $|E[X(\omega)]| \leq E[|X(\omega)|]$. Если g=g(x) – выпуклая функция и $g(X(\omega))$ также является интегрируемой случайной величиной, то

$$g(E[X(\omega)]) \leq E[g(X(\omega))] \qquad (1.4)$$

– *неравенство Йенсена*.

5. Если $X(\omega)$ – интегрируемая случайная величина, и $A \in F$, то $I_A(\omega)X(\omega)$ – также интегрируемая случайная величина.
6. Если $X(\omega)$ и $Y(\omega)$ – интегрируемые случайные величины, то

$$E[X(\omega) + Y(\omega)] = E[X(\omega)] + E[Y(\omega)].$$

7. Если $X(\omega) = 0$ P-п.н., то $E[X(\omega)] = 0$.
8. Если $X(\omega) = Y(\omega)$ P-п.н. и $X(\omega)$ и $Y(\omega)$ – интегрируемые случайные величины, то $E[X(\omega)] = E[Y(\omega)]$.
9. $X(\omega)$ и $Y(\omega)$ – интегрируемые случайные величины, и $E[I_A(\omega)X(\omega)] \leq E[I_A(\omega)Y(\omega)]$ для любого $A \in F$, то $X(\omega) \leq Y(\omega)$ P-п.н.

10. (Теорема о монотонной сходимости). Пусть $\{X_n(\omega)\}$, $Y(\omega)$ – случайные величины. Если $X_n(\omega) \geq Y(\omega)$ для всех n, $E[Y(\omega)] > -\infty$ и $X_n(\omega) \uparrow X(\omega)$ при $n \to \infty$, то $E[X_n(\omega)] \uparrow E[X(\omega)]$.
11. (Лемма Фату). Пусть $\{X_n(\omega)\}$, $Y(\omega)$ – случайные величины. Если $X_n(\omega) \geq Y(\omega)$ для всех n, $E[Y(\omega)] > -\infty$, то $E\left[\underline{lim}\, X_n(\omega)\right] \leq \underline{lim}\, E[X_n(\omega)]$. Если $|X_n(\omega)| \leq Y(\omega)$, и $E[Y(\omega)] < \infty$, то $E\left[\underline{lim}\, X_n(\omega)\right] \leq \underline{lim}\, E\left[X_n(\omega)\right] \leq \overline{lim}\, E\left[X_n(\omega)\right] \leq E[\overline{lim}\, X_n(\omega)]$.
12. (Теорема Лебега о мажорируемой сходимости). Пусть $\{X_n(\omega)\}$, $Y(\omega)$ – случайные величины, $|X_n(\omega)| \leq Y(\omega)$, $E[Y(\omega)] < \infty$ и $X_n(\omega) \to X(\omega)$ P-п.н. Тогда $E[X(\omega)] < \infty$, $E[X_n(\omega)] \to E[X(\omega)]$ и $E[|X_n(\omega) - X(\omega)|] \to 0$ при $n \to \infty$.
13. Если $X(\omega)$ и $Y(\omega)$ – независимые интегрируемые случайные величины, то $E[X(\omega)Y(\omega)] = E[X(\omega)]E[Y(\omega)]$.

Замечание 1.5. Свойства (безусловных) математических ожиданий приведены здесь для того, чтобы впоследствии сравнить их со свойствами условных математических ожиданий. Вообще, свойств математических ожиданий гораздо больше (см. книгу А.Н. Ширяева «Вероятность»).

3. Условное математическое ожидание (УМО)

Определение 1.11. Пусть $A,\ B \in F$, $P\{B\} > 0$. Тогда *условная вероятность события A при условии B* определяется формулой

$$P\{B\} = \frac{P\{AB\}}{P\{B\}}. \qquad (1.5)$$

Определение 1.12. Пусть $X(\omega)$ – интегрируемая случайная величина, $B \in F$, $P\{B\} > 0$. Тогда *условное математическое ожидание X относительно события B* определяется формулой

$$E[B] = \frac{E[I_B(\omega)X(\omega)]}{E[I_B(\omega)]} = \frac{E[I_B(\omega)X(\omega)]}{P\{B\}}. \qquad (1.6)$$

Замечание 1.6. Условная вероятность одного события относительно другого, а также условное математическое ожидание относительно события – неслучайные величины. По сути, это – *апостериорные* характеристики, которые вычисляются после проведения опыта, в результате которого произошло событие B (и нам об этом известно). На практике дополнительная информация гораздо богаче: обычно наблюдаются случайные величины, векторы, процессы. Как учитывать такую информацию? Как учитывать возможную информацию априорно, до получения результатов опыта?

Определение 1.13. Пусть $X(\omega)$ – интегрируемая случайная величина, $D \subseteq F$ – конечное разбиение. Тогда *УМО* $X(\omega)$ *относительно разбиения* D называется величина

$$E[D] = \sum_{i=1}^{n} E\left[D_i\right] I_{D_i}(\omega) = \sum_{i=1}^{n} \frac{E\left[I_{D_i}(\omega)X(\omega)\right]}{E\left[I_{D_i}(\omega)\right]} I_{D_i}(\omega). \qquad (1.7)$$

Замечание 1.7. УМО относительно разбиения уже является функцией от случайного события, т.е. случайной величиной!

Замечание 1.8. УМО относительно разбиения позволяет построить условную вероятность относительно разбиения. При этом условная вероятность в этом случае уже будет случайной, но обладать свойствами вероятности. Действительно, пусть $G \subseteq F$ - еще одно конечное разбиение: $G = \{G_1, ..., G_m\}$. Построим по формуле (1.7) УМО $E[D] = P\{G_k|D\}$, $k=1,...,m$. Легко проверить, что $0 < P\{D\} < 1$, и $\sum_{k=1}^{m} P\{D\} = 1$ Р-п.н.

Замечание 1.9. Легко заметить, что УМО относительно разбиения описывает нам УМО относительно дискретной случайной величины!

Определение 1.14. Пусть $X(\omega)$ – неотрицательная интегрируемая случайная величина, $G \subseteq F$ – σ-подалгебра F. Условное математическое ожидание случайной величины $X(\omega)$ относительно σ-подалгебры G – расширенная случайная величина $E[X|G](\omega)$, обладающая следующими свойствами:

1) $Y(\omega) = E[X|G](\omega) - G$ - измерима (т.е. $Y^{-1}(B(R)) \subseteq G$),
2) для любого $A \in G$ выполняется равенство

$$\int_A X(\omega)P(d\omega) = \int_A E[X|G](\omega)P(d\omega), \qquad (1.8)$$

или, что эквивалентно

$$E\Big[(E[G](\omega) - X(\omega))I_A(\omega)\Big] \equiv 0 \qquad (1.8')$$

для любого $A \in G$.

Для произвольной интегрируемой случайной величины $X(\omega)$ УМО $E[X|G](\omega)$ считается определенным, если $(E[G](\omega),\ E[G](\omega)) < \infty$, и определяется формулой

$$E[G](\omega) = E[G](\omega) - E[G](\omega). \qquad (1.9)$$

Замечание 1.10. Рассмотрим вероятностное пространство (Ω, F, P) и измеримое пространство (Ω, G), где $G \subseteq F$ – σ-подалгебра F. Рассмотрим функцию $Q\colon G \to R$:

$$Q(A) = \int_A X(\omega)P(d\omega).$$

Можно проверить, что $Q(A)$ является (знакопеременной) мерой на (Ω, G), причем $Q \ll P$. По теореме Радона-Никодима существует производная Радона-Никодима $\frac{dQ}{dP}(\omega)$, для которой верно равенство

$$Q(A) = \int_A X(\omega)P(d\omega) = \int_A \frac{dQ}{dP}(\omega)\, P(d\omega).$$

С другой стороны, из (1.8) следует, что $E[G](\omega) = \frac{dQ}{dP}(\omega)$ P-п.н. В последнем равенстве важно следующее: $X(\omega)$ – не есть производная Радона-Никодима Q по P, т.к. она в общем случае F -измерима, а не G-измерима. А $E[G]$ - G-измерима по определению!

Замечание 1.11. В соответствии с теоремой Радона-Никодима УМО $E[G](\omega)$ определено не единственным образом, а только с точностью до множеств вероятностной меры ноль.

Замечание 1.12. Пусть $D \subseteq F$ – конечное разбиение, $G = \sigma\{D\}$. Если $X(\omega)$ – интегрируемая случайная величина, то

$$E[D](\omega) = E[G](\omega) \qquad P - \text{п. н.}$$

Т.о. Определения 1.13 и 1.14 непротиворечивы.

Определение 1.15. Пусть $A \in F$ – некоторое случайное событие. УМО $E[G](\omega) = P\{A|G\}$ называется условной вероятностью события A относительно – σ-подалгебры G.

Замечание 1.13. На практике доступная информация никогда не бывает доступной в σ -подалгебр: обычно это случайные величины, векторы или процессы. Пусть $X(\omega)$ – интегрируемая случайная величина, $Y(\omega)$ – некоторая случайная величина. Легко проверить, что $F^Y = Y^{-1}(B(R)) \subseteq F$ - σ-подалгебра σ-алгебры F. Она называется σ -подалгеброй, порожденной случайной величиной $Y(\omega)$. Тогда по определению УМО случайной величины $X(\omega)$ относительно случайной величины $Y(\omega)$ определяется следующим равенством $E[Y](\omega) = E\left[F^Y\right](\omega)$. Аналогичным образом определяется σ -подалгебра F^Y в случае, если $Y(\omega)$ – случайный вектор. Если $Y(s, \omega)$, $s \in [0, t]$ – случайный процесс, то $F^Y = \sigma\{Y(s, \cdot),\ s \in [0, t]\}$ - минимальная σ-подалгебра, содержащая все подалгебры $F^{Y(s,\cdot)}$.

Замечание 1.14. Как вычислять $E[Y](\omega)$? Верно следующее утверждение. Пусть $X(\omega)$ –интегрируемая случайная величина, $Y(\omega)$ – случайная величина. Тогда существует такая борелевская функция $g = g(y)$, что $E[Y](\omega) = g(Y(\omega))$ P-п.н. Таким образом, УМО – это некоторая функция от наблюдений. Аналогичные утверждения верны и для случая наблюдений в форме случайного вектора и случайного процесса.

Определение 1.16. Пусть $X(\omega)$ – интегрируемая случайная величина, $Y(\omega)$ – случайная величина, P_Y – вероятностное распределение $Y(\omega)$ на измеримом пространстве $(R, B(R))$. Борелевская функция $g = g(y)$ называется УМО $X(\omega)$ при условии *Y*=*y*, если для любого борелевского множества $A \in B(R)$ выполнено равенство

$$\int_{\{\omega: Y(\omega) \in A\}} X(\omega) P(d\omega) = \int_A g(y) P_Y(dy)$$

Замечание 1.15. Пусть $Y(\omega)$ – случайная величина. Тогда согласно предыдущему замечанию существует такая функция $P(A, y)$, что
1) Для любого фиксированного $A \in F$ функция $P(\cdot, \omega)$ – измерима,
2) $P(A, Y(\omega)) = P\{A|Y\}\ P - \text{п. н.}$

Тогда $P(A, y)$ называется *условным распределением относительно случайной величины* $Y(\omega)$. Ясно, что функция $P(A, y)$ определена неединственным образом. Если $P(A, y)$ обладает дополнительным свойством

3) для любого фиксированного y функция $P(A, \cdot)$ является вероятностной мерой,

то такая функция $P(A, y)$ называется *регулярной версией условного распределения*.

Свойства условных математических ожиданий

1. Если C – константа, то $E[C|G] = C$
2. Если $X(\omega)$ и $Y(\omega)$ – интегрируемые случайные величины, и $X(\omega) \leq Y(\omega)$, $P -$ п. н. то и $E[X(\omega)|G] \leq E[Y(\omega)|G]$ $P -$ п. н.
3. Если $X(\omega)$ – интегрируемая случайная величина, то $|E[X(\omega)|G]| \leq E[|X(\omega)||G]$ $P -$ п. н. Если g=g(x) – выпуклая функция и $g(X(\omega))$ также является интегрируемой случайной величиной, то

$$g(E[X(\omega)|G]) \leq E[g(X(\omega))|G] \qquad P - \text{п. н.} \qquad (1.10)$$

– *неравенство Йенсена*.

4. Если $X(\omega)$ и $Y(\omega)$ – интегрируемые случайные величины, a и b – константы, то

$$E[aX(\omega) + bY(\omega)|G] = aE[X(\omega)|G] + bE[Y(\omega)|G] \quad P - \text{п. н.}$$

5. Если $X(\omega)$– интегрируемая случайная величина, $G_0 = \{\Omega, \emptyset\}$ – тривиальная $\sigma -$ алгебра, то

$$E\left[X(\omega)|G_0\right] = E[X(\omega)] \qquad P - \text{п. н.}$$

6. Формула взятия повторного математического ожидания:

$$E\left[E\left[X(\omega)|G_0\right]\right] = E[X(\omega)] \qquad (1.11)$$

7. Пусть $G_1 \subseteq G_2 \subseteq F$ - вложенные $\sigma -$подалгебры, $X(\omega)$– интегрируемая случайная величина, тогда

$$E\left[E\left[X(\omega)|G_1\right]|G_2\right] = E\left[E\left[X(\omega)|G_2\right]|G_1\right] = E\left[X(\omega)|G_1\right] \qquad P - \text{п. н.} \qquad (1.12)$$

8. Если $X(\omega)$ - интегрируемая случайная величина, не зависящая от G (т.е. $E\left[I_A(\omega)I_B(\omega)\right] = E\left[I_A(\omega)\right]E[I_B(\omega)]$ для любых $A \in G$, $B \in F^X$), то

$$E[X(\omega)|G] = E[X(\omega)] \qquad P - \text{п. н.}$$

9. Если $X(\omega)$ - интегрируемая случайная величина, а $Y(\omega)$ - интегрируемая G – измеримая случайная величина, то

$$E[Y(\omega)X(\omega)|G] = Y(\omega)E[X(\omega)|G] \qquad P - \text{п. н.} \qquad (1.13)$$

10. Пусть $\{X_n(\omega)\}$, $Y(\omega)$ – случайные величины.
Если $|X_n(\omega)| \leq Y(\omega)$, $E[Y(\omega)] < \infty$ и $X_n(\omega) \to X(\omega)$ $P-$ п. н. при $n \to \infty$, то $E[X_n(\omega)|G] \to E[X(\omega)|G]$ и $[|X_n(\omega) - X(\omega)||G] \to 0$ $P-$ п. н.
Если $X_n(\omega) \geq Y(\omega)$, $E[Y(\omega)] > -\infty$ и $X_n(\omega) \uparrow X(\omega)$ $P-$ п. н. при $n \to \infty$, то $E[G] \uparrow E[G]$ $P-$ п. н.
Если $X_n(\omega) \geq Y(\omega)$, $E[Y(\omega)] > -\infty$ то $E\left[X_n(\omega)|G\right] \leq \underline{lim}\ E\left[X_n(\omega)G\right]$ $P-$ п. н.

Если $X_n(\omega) \geq 0$, то $E\left[\sum X_n(\omega)|G\right] = \sum E[X_n(\omega)|G]$ $P-$ п. н.

11. Пусть $X(\omega)$ - интегрируемая случайная величина, $\ldots \subseteq G_{-1} \subseteq G_0 \subseteq G_1 \subseteq G_2 \subseteq \ldots F$ - последовательность вложенных вложенные σ –подалгебр, $G_{-\infty} = \bigcap_n G_n$, $G_{+\infty} = \sigma\{\bigcup_n G_n\}$, тогда

$$E\left[G_n\right] \to E\left[G_{-\infty}\right] \text{ при } n \to -\infty\ P - \text{п. н.,}$$

$$E\left[G_n\right] \to E\left[G_{+\infty}\right] \text{ при } n \to +\infty\ P - \text{п. н.}$$

4. *Использование УМО в задачах оптимального байесовского оценивания*

Математический аппарат УМО широко применяется в системном анализе стохастических динамических систем при решении задач анализа, оценивания и управления. В данном курсе будет представлено использование УМО в задачах СК-оптимального байесовского оценивания.

Определение 1.17. Пусть $X(\omega)$ – ненаблюдаемая интегрируемая случайная величина, $Y(\omega)$ – наблюдаемая случайная величина. Измеримая борелевская функция $\overline{X} = \overline{X}(Y)$ называется *оценкой* X, построенной по наблюдениям Y. Величина $\Delta = \overline{X} - X$ называется ошибкой оценки $\overline{X}$. Пусть задана некоторая функция $\rho = \rho(\Delta)$, называемая функцией потерь и функционал $J(\overline{X}) = E[\rho(\overline{X} - X)]$, называемый *критерием оптимальности*. Задача

$$J(\overline{X}) \to \qquad (1.14)$$

называется задачей *(оптимального) байесовского оценивания величины* X *по наблюдениям* Y *в смысле критерия* J. При этом оценки $\overline{X}$ таковы, что значение критерия $J(\overline{X})$ определено. Оценка $\hat{X}$, доставляющая критерию $J(\overline{X})$ минимум, называется *(абсолютно) оптимальной в смысле критерия* $J(\overline{X})$.

Определение 1.18. Пусть X – некоторый класс допустимых функций $\overline{X}$ (допустимых оценок). Задача

$$J(\overline{X}) \to \qquad (1.14')$$

называется задачей *условно-оптимального байесовского оценивания величины* X *по наблюдениям* Y *в смысле критерия* J *в классе допустимых оценок* X. Оценка $\hat{X}$, доставляющая критерию $J(\overline{X})$ минимум на X, называется *условно-оптимальной в смысле критерия* $J(\overline{X})$ *в классе допустимых оценок* X.

Замечание 1.16. Обычно функции потерь ρ обладают очевидными свойствами.
1) $\rho(0) = 0$,
2) $\rho(x) \geq 0$,
3) $\rho(x) \geq \rho(y)$, если $\|x\| > \|y\|$.

Замечание 1.17. В задачах (1.14), (1,14') нужно найти оптимальную оценку $\hat{X}$ (либо доказать, что ее не существует). При этом значение $J(\hat{X})$ характеризует точность оптимальной оценки. Примечательно, что в большинстве случаев, даже если удается аналитически получить $\hat{X}$, не удается получить $J(\hat{X})$.

Определение 1.19. Частный случай задачи (1.14) для квадратичной функции потерь $\rho = \|\overline{X} - X\|_2^2$, называется задачей *оптимального в среднеквадратическом смысле оценивания (СК-оптимального оценивания)*.

Теорема 1.1. $\hat{X} = E[X|Y]$ – СК-оптимальная оценка.

Доказательство:

Пусть $\overline{X} = \overline{X}(Y)$ – произвольная оценка. Тогда

$$J(\overline{X}) = E\left[\|\overline{X} - X\|_2^2\right] = E\left[\|(\overline{X} - E[Y]) + (E[Y] - X)\|_2^2\right] = E\left[E\left[\|(\overline{X} - E[Y]) + (E[Y] - X)\|_2^2 | Y\right]\right] =$$

Замечание 1.18. В большом количестве задач байесовского оценивания (для разных функций потерь) оптимальная оценка выражается через условное распределение X относительно Y. Ниже представлены некоторые частные случаи вычисления УМО и условных распределений.

Пример 1.1. (*УМО дискретной случайной величины по наблюдению дискретной случайной величины*). Совместное распределение X и Y задано двумерным рядом распределения

	y_1	y_2	…	y_m
x_1	p_{11}	p_{12}	…	p_{1m}
x_2	p_{21}	p_{22}	…	p_{2m}
…	…	…	…	…
x_n	p_{n1}	p_{n2}	…	p_{nm}

$$p_{ij} = P\{X = x_i, Y = y_i\}.$$

Тогда

$$E[Y] = \sum_{i=1}^{n} x_i \sum_{j=1}^{m} \frac{p_{ij}}{\sum_{k=1}^{n} p_{kj}} I_{\{y_j\}}(Y), \qquad (1.15)$$

$$P\{Y\} = \sum_{j=1}^{m} \frac{p_{ij}}{\sum_{k=1}^{n} p_{kj}} I_{\{y_j\}}(Y), \qquad (1.16)$$

$$P\{Y = y_j\} = \frac{p_{ij}}{\sum_{k=1}^{n} p_{kj}}. \qquad (1.17)$$

Пример 1.2. (*УМО непрерывной случайной величины по наблюдению непрерывной случайной величины*). Совместное распределение X и Y задано двумерной плотностью распределения $f_{XY}(x,y)$. Без ограничения общности будем считать, что для всех (x,y) $f_{XY}(x,y)>0$.

$$E[Y] = \frac{\int_{-\infty}^{\infty} x f_{XY}(x,Y)dx}{\int_{-\infty}^{\infty} f_{XY}(u,Y)du}, \qquad (1.18)$$

$$f_{X|Y}(Y = y) = \frac{f_{XY}(x,y)}{\int_{-\infty}^{\infty} f_{XY}(u,y)du}. \qquad (1.19)$$

Пример 1.3. (*Теорема о нормальной корреляции)*. Пусть блочный случайный вектор

$$[X\ Y\ ] \sim N\left(\left[m_X\ m_Y\right],\ \left[k_{XX}\ k_{XY}\ k_{YX}\ k_{YY}\ \right]\right),$$

тогда условное распределение X относительно Y является гауссовским со средним

$$E[Y] = m_X + k_{XY} k_{YY}^{+} \left(Y - m_Y\right), \qquad (1.20)$$

и условной ковариационной матрицей

$$cov(Y) = E[Y] = k_{XX} - k_{XY} k_{YY}^{+} k_{YX}, \qquad (1.21)$$

где k_{YY}^{+} - матрица, псевдообратная к k_{YY} по Муру-Пенроузу.

Пример 1.4. (*Байесовская классификация)*. Пусть X – дискретный случайный вектор с рядом распределения

X	e_1	e_2	…	e_n
p	p_1	p_2	…	p_n

($e_1, \ldots,$ e_n – единичные векторы-столбцы пространства R^n). Пусть $V = V(\omega)$ – нормированная непрерывная случайная величина со строго положительной плотностью распределения $\varphi_V(v)$, $a = \left(a_1, ..., a_n\right)$, $b = \left(b_1, ..., b_n\right)$ – заданные вектор-столбцы, причем все компоненты b_k – строго положительные. Наблюдение Y описывается моделью

$$Y = aX + bXV. \qquad (1.22)$$

Тогда

$$E[Y] = \sum_{i=1}^{n} \frac{\frac{p_i}{b_i}\varphi_V\left(\frac{Y-a_i}{b_i}\right)}{\sum_{j=1}^{n}\frac{p_j}{b_j}\varphi_V\left(\frac{Y-a_j}{b_j}\right)} e_i, \qquad (1.23)$$

$$P\{Y = y\} = \frac{\frac{p_i}{b_i}\varphi_V\left(\frac{y-a_i}{b_i}\right)}{\sum_{j=1}^{n}\frac{p_j}{b_j}\varphi_V\left(\frac{y-a_j}{b_j}\right)}. \qquad (1.24)$$

Проверим, что $E[Y]$ определяется формулой (1.23). Для этого нужно проверить выполнение определения УМО. Очевидно, что $E[Y]$ в форме (1.23) – измеримая функция наблюдения Y. Осталось только проверить истинность (1.8) или (1.8'). Прежде всего, по формуле полной вероятности наблюдения Y имеют плотность распределения, определяемую формулой

$$\varphi_Y(y) = \sum_{j=1}^{n}\frac{p_j}{b_j}\varphi_V\left(\frac{y-a_j}{b_j}\right). \qquad (1.25)$$

Далее, пусть $B \in B(R)$ – произвольное борелевское множество, тогда

$$\int\limits_{\{\omega: Y(\omega) \in B\}} X(\omega)P(d\omega) = \sum_{i=1}^{n} \int\limits_{\{\omega: Y(\omega) \in B,\ X(\omega)=e_i\}} X(\omega)P(d\omega) = \sum_{i=1}^{n} e_i \int\limits_{\{\omega: Y(\omega) \in B,\ X(\omega)=e_i\}} P(d\omega) = \sum_{i=1}^{n} e_i \int_B \ldots$$

С другой стороны,

$$\int\limits_{\{\omega: Y(\omega) \in B\}} \sum_{i=1}^{n} \frac{\frac{p_i}{b_i}\varphi_V\left(\frac{Y(\omega)-a_i}{b_i}\right)}{\sum_{j=1}^{n}\frac{p_j}{b_j}\varphi_V\left(\frac{Y(\omega)-a_j}{b_j}\right)} e_i P(d\omega) = \int_B \sum_{i=1}^{n} \frac{\frac{p_i}{b_i}\varphi_V\left(\frac{y-a_i}{b_i}\right)}{\sum_{j=1}^{n}\frac{p_j}{b_j}\varphi_V\left(\frac{y-a_j}{b_j}\right)} e_i P_Y(dy) = \int_B \sum_{i=1}^{n} \frac{\frac{p_i}{b_i}\varphi_V\left(\frac{y-a_i}{b_i}\right)}{\sum_{j=1}^{n}\frac{p_j}{b_j}\varphi_V\left(\frac{y-a_j}{b_j}\right)} e_i \sum_{j=1}^{n}\frac{p_j}{b_j}\varphi_V\left(\ldots\right.$$

Выполнение равенства (1.8) проверено.

5. *Задачи для самостоятельного решения*

Задача 1.1. Пусть $X(\omega)$ - $F|G$ - измеримый случайный элемент. Доказать, что $X^{-1}(G) \subseteq F$ является σ-подалгеброй F.

Задача 1.2. Сформулировать утверждения, «симметричные» свойствам 10, 11 математического ожидания.

Задача 1.3. Пусть X – дискретный случайный вектор с рядом распределения

X	e_1	e_2	…	e_n

p	p_1	p_2	…	p_n

($e_1,\dots,$ e_n – единичные векторы-столбцы пространства R^n). Обозначим $p=(p_1,\dots, p_n)^T$ – вектор-столбец распределения. Найти $E[X]$, $cov(X,X)$. Пусть $Y=g(X)$ – произвольное преобразование вектора X. Доказать, что оно всегда может быть представлено в линейной форме: $Y=GX$.

Задача 1.4. Пользуясь Определением 1.10 ответить на вопрос: у случайной величины, имеющей распределение Коши среднее значение бесконечно или не существует?

Задача 1.5. Пусть $X(\omega)$ – случайная величина, $V(\omega)$ – независимая от $X(\omega)$ случайная величина, имеющая положительную плотность распределения $\varphi_V(v)$, $Y(\omega) = X(\omega) + V(\omega)$. Найти условное распределение Y относительно X. (Подсказка: у этого распределения существует плотность).

Задача 1.6. Исходя из определения УМО относительно σ-подалгебры доказать справедливость Замечания 1.13.

Задача 1.7. Пусть $X(\omega)$, $Z(\omega)$– независимые одинаково распределенные случайные величины, $Y(\omega) = X(\omega) + Z(\omega)$. Найти $E[Y]$.

Задача 1.8. Проверить теорему о нормальной корреляции для скалярного случая.

Задача 1.9. Доказать, что в случае функции потерь $\rho(\overline{X} - X) = |\overline{X} - X|$ оптимальная байесовская оценка совпадает с медианой условного распределения.

Задача 1.10. Доказать истинность формул (1.15) – (1.19).

Задача 1.11. Доказать истинность формулы (1.25).